

A CLASSIC REISSUE

Foundations for Microwave Engineering

SECOND EDITION

Robert E. Collin

Foundations for Microwave Engineering

SECOND EDITION

Donald G. Dudley, Series Editor

muiaUomjor Microwatt Engineering. Second Edition, covers the major topics of microwave engineering- Its presentation defines the accepted standard for both advanced undergraduate and graduate level courses on microwave engineering. An essential reference book for the practicing microwave engineer, it features.-

- Planar transmission lines, as well as an appendix that describes in detail conformal mapping methods for their analysis and attenuation characteristics.
- Small aperture coupling and its application in practical components such as directional couplers and cavity coupling
- Printed circuit components with an emphasis on techniques such as even and odd mode analysis and the use of symmetry properties
- Microwave linear amplifier and oscillator design using solid-state circuits such as varactor devices and transistors.

Foimdatoni for Microwave Engineering, Second Edition, has extensive coverage of transmission lines, waveguides, microwave circuit theory, impedance matching, and cavity resonators. It devotes an entire chapter to fundamental microwave tubes, as well as other chapters on periodic structures, microwave filters, small signal solid-state microwave amplifier and oscillator design, and negative resistance devices and circuits. Completely updated in 1992, it is being issued by the IEEE Press in response to requests from our many members, who found u an invaluable textbook and an enduring reference for practicing microwave engineers.

<oul ihe Author

Robert E. Collin is the author or coauthor of more than 150 technical papers and five books on electromagnetic nSeory and applications. His classic text, *Field Theory of Guided Waves*, is also a volume in the series. Professor Collin has had a long and distinguished academic career at Case Western Reserve University. In addition to his professorial duties, he has served as chairman of the Department of Electrical Engineering and as interim dean of engineering.

Professor Collin is a life fellow of the IEEE and a member of the Microwave Theory and Techniques Society and the Antennas and Propagation Society (APSI). He is a member of the U.S. Commission B of URSI and a member of the Geophysical Society. Other honors include the Diekman Award from Case Western Reserve University for distinguished graduate teaching, the IEEE APS Distinguished Career Award (1992), the IEEE Schelkunoff Prize Paper Award (1992), the IEEE Electromagnetics Award (1998), and an IEEE Third Millennium Medal in 2000. In 1990 Professor Collin was elected to the National Academy of Engineering.

The IEEE Press Series on Electromagnetic Wave Theory offers outstanding coverage of the field. It consists of new titles of contemporary interest, as well as reissues and revisions of recognized classics by established authors and researchers. The series emphasizes works of long-term archival significance in electromagnetic waves and applications. Designed specifically for graduate students, researchers and practicing engineers, the series provides affordable volumes that explore and explain electromagnetic waves beyond the undergraduate level.



Subscribe to our free Electrical Engineering eNewsletter at
www.wiley.com/enewsletters

Visit www.wiley.com/ieee

ISBN D-7f1D3-bD31-1



>WILEY-
'INTERSCIEN'

780780"36rmr)"

An IEEE Press Classic Reissue

**FOUNDATIONS FOR
MICROWAVE ENGINEERING**

SECOND EDITION

The IEEE Press Series on Electromagnetic Wave Theory consists of new titles as well as reprintings and revisions of recognized classics that maintain long-term archival significance in electromagnetic waves and applications.

Series Editor
Donald G. Dudley
University of Arizona

Advisory Board
Robert E. Collin
Case Western Reserve University

Akira Ishimaru
University of Washington

D. S. Jones
University of Dundee

Associate Editors

ELECTROMAGNETIC THEORY, SCATTERING, AND DIFFRACTION
Ehud Heyman
Tel-Aviv University

INTEGRAL EQUATION METHODS
Donald R. Wilson
University of Houston

DIFFERENTIAL EQUATION METHODS
Andreas C. Cangellans
University of Arizona

ANTENNAS, PROPAGATION, AND MICROWAVES
David R. Jackson
University of Houston

BOOKS IN THE IEEE PRESS SERIES ON ELECTROMAGNETIC WAVE THEORY

- Christopoulos. C, *The Transmission-Line Modeling Methods: TLM*
Clemmow, R C. *The Plane Wave Spectrum Representation of Electromagnetic Fields*
Collin. R. E.. *Field Theory of Guided Waves*, Second Edition
Collin, R. E.. *Foundations for Microwave Engineering*, Second Edition
Dudley. D. G., *Mathematical Foundations for Electromagnetic Theory*
Elliot. R. S., *Electromagnetics: History, Theory, and Applications*
Felsen. L. B.. and Marcuvitz. N.. *Radiation and Scattering of Waves*
Harrington. R. F, *Field Computation by Moment Methods*
Hansen et al, *Plane-Wave Theory of Time-Domain Fields: Near-Field Scanning Applications*
Ishimaru. A., *Wave Propagation and Scattering in Random Media*
Jones, D. S.. *Methods in Electromagnetic Wave Propagation*. Second Edition
Lindell. I. V., *Methods for Electromagnetic Field Analysis*
Peterson el al.. *Computational Methods for Electromagnetics*
Tai. C. T, *Generalized Vector and Dyadic Analysis: Applied Mathematics in Field Theory*
Tai. C. T. *Dyadic Green Functions in Electromagnetic Theory*. Second Edition
Van Bladel, J., *Singular Electromagnetic Fields and Sources*
Volakis et al., *Finite Element Method for Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications*
Wail. J.. *Electromagnetic Waves in Stratified Media*

An IEEE Press Classic Reissue

FOUNDATIONS FOR MICROWAVE ENGINEERING

SECOND EDITION



IEEE Press Series on
Electromagnetic Wave Theory

Robert E. Collin

*Professor of Electrical Engineering
Case Western Reserve University
Cleveland, OH*

IEEE Antennas & Propagation Society, *Sponsor*

IEEE Microwave Theory and Techniques Society, *Sponsor*

The Institute of Electrical and Electronics Engineers, Inc., New York

WILEY-
INTERSCIENCE

A JOHN WILEY & SONS, INC.. PUBLICATION

© 2001 THE INSTITUTE OF ELECTRICAL AND ELECTRONICS
ENGINEERS, INC. 3 Park Avenue, 17th Floor, New York, NY 10016-5997
Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400, fax 978-750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, e-mail: permcoordinator@wiley.com.

For general information on our other products and services please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Printed in the United States of America

1 0 9 8 7 6 5 4 3 2

ISBN 0-7803-6031-1

Library of Congress Cataloging-in-Publication Data

Collin, Robert E.

Foundations for microwave engineering / Robert E. Collin.- 2nd ed.

p. cm. - (IEEE Press series on electromagnetic wave theory)

Originally published : New York : McGraw Hill, c1992.

"An IEEE Press classic reissue."

Includes bibliographical references and index.

ISBN 0-7803-6031-1

I. Microwave devices. I. Title. II. Series.

TK7876 .C645 2000
621.381'3--dc21

00-053874

FOREWORD TO THE REISSUED EDITION

The purpose of the IEEE Press Series on Electromagnetic Wave Theory is to publish books of long-term archival significance in electromagnetics. Included are new titles as well as reprints and revisions of recognized classics. The book *Foundations for Microwave Engineering*, by Robert E. Collin, is by any measure such a classic. The original edition of the book appeared in 1966 and remained in print until the appearance of the second edition in 1992, a span of 26 years.

In the second edition, Professor Collin completely updated and modernized his book to include the many advances that had occurred in microwave engineering since the appearance of the original edition. That the second edition has gone out of print has caused concern among many of my colleagues in the IEEE Antennas and Propagation Society (APS) and the IEEE Microwave Theory and Techniques Society (MTT). We at the IEEE Press are delighted to be able to overcome this difficulty by introducing a reprint of the second edition into our Series on Electromagnetic Wave Theory. The book is a thorough and in-depth exposition on microwave engineering. Furthermore, it will make an excellent companion to Professor Collin's book, *Field Theory- of Guided Waves*, also included in the series.

Professor Collin has been a valued colleague for many years. He is the author or coauthor of five books and more than 150 technical papers. His contributions to electromagnetics span a wide range of subjects and have brought him international respect and many awards. Among these are election to the National Academy of Engineering, the IEEE Electromagnetics Field Award, the IEEE/APS Distinguished Career Award, an IEEE/APS Schelkunoff Prize Paper Award, and the IEEE Third Millennium Medal.

It is with pleasure that I welcome this book into the series.

Donald G. Dudley
University of Arizona
Series Editor
IEEE Press Series on Electromagnetic Wave Theory

IEEE Press
445 Hoes Lane, P.O. Box 1331
Piscataway, NJ 08855-1331

IEEE Press Editorial Board
Robert J. Herrick, *Editor in Chief*

M. Akay	M. Eden	M.S. Newman
J. B. Anderson	M. E. El-Hawary	M. Padgett
P. M. Anderson	R. F. Hoyt	W.D. Reeve
J. E. Brewer	S. V. Kartalopoulos	G. Zobrist
	D. Kirk	

Kenneth Moore, *Director of IEEE Press*
Catherine Faduska, *Senior Acquisitions Editor*
Linda Matarazzo, *Associate Acquisitions Editor*
Marilyn G. Catis, *Marketing Manager*
Mark Morrell, *Associate Production Editor*

IEEE Antennas & Propagation Society, *Sponsor*
AP-S Liaison to IEEE Press, Robert Mailloux

IEEE Microwave Theory and Techniques Society, *Sponsor*
MTT-S Liaison to IEEE Press, Karl Varian

Cover design: William T. Donnelly, *WTDesign*

CONTENTS

Preface

xv

1	Introduction	1
1.1	Microwave Frequencies	1
1.2	Microwave Applications	3
1.3	Microwave Circuit Elements and Analysis	6
	References	16
2	Electromagnetic Theory	17
2.1	Maxwell's Equations	17
2.2	Constitutive Relations	23
2.3	Static Fields	28
2.4	Wave Equation	31
2.5	Energy and Power	33
2.6	Boundary Conditions	39
2.7	Plane Waves	44
	Plane Waves in Free Space	44
2.8	Reflection from a Dielectric Interface	49
	1. Parallel Polarization	49
	2. Perpendicular Polarization	52
2.9	Reflection from a Conducting Plane	53
2.10	Potential Theory	56
•2.11	Derivation of Solution for Vector Potential	59
2.12	Lorentz Reciprocity Theorem	62
	Problems	65
	References	70
3	Transmission Lines and Waveguides	71
	Part 1 Waves on Transmission Lines	72
3.1	Waves on An Ideal Transmission Line	72
3.2	Terminated Transmission Line: Resistive Load	78

3.3	Capacitive Termination	82
3.4	Steady-State Sinusoidal Waves	85
3.5	Waves on a Lossy Transmission Line	86
	Loss-Free Transmission Line	88
	Low-Loss Transmission Line	89
3.6	Terminated Transmission Line: Sinusoidal Waves	89
	Terminated Lossy Line	94
	Part 2 Field Analysis of Transmission Lines	96
3.7	Classification of Wave Solutions	96
	TEM Waves	99
	TE Waves	100
	TM Waves	102
3.8	Transmission Lines (Field Analysis)	104
	Lossless Transmission Line	104
	Transmission Line with Small Losses	108
3.9	Transmission-Line Parameters	112
3.10	In homogeneously Filled Parallel-Plate Transmission Line	117
	Low-Frequency Solution	121
	High-Frequency Solution	123
3.11	Planar Transmission Lines	125
3.12	Microstrip Transmission Line	130
	Low-Frequency Solutions	136
	Microstrip Attenuation	153
	High-Frequency Properties of Microstrip Lines	158
	Attenuation	163
3.13	Coupled Microstrip Lines	164
3.14	Strip Transmission Lines	170
	Attenuation	171
3.15	Coupled Strip Lines	173
3.16	Coplanar Transmission Lines	175
	Attenuation	178
	High-Frequency Dispersion	180
	Part 3 Rectangular and Circular Waveguides	180
3.17	Rectangular Waveguide	181
	TE Waves	182
	Power	186
	Attenuation	187
	Dominant TE ₁₀ Mode	190
	TM Modes	193
3.18	Circular Waveguides	194
	TM Modes	194
	TE Modes	196
3.19	Wave Velocities	198
	Phase Velocity	199
	Group Velocity	200
	Energy-Flow Velocity	204
3.20	Ridge Waveguide	205
3.21	Fin Line	208
	Problems	210
	References	219

CHAPTER 1

INTRODUCTION

The purpose of this introductory chapter is to provide a short, and admittedly incomplete, survey of what the microwave engineering field encompasses. Section 1.2 presents a brief discussion of many of the varied and sometimes unique applications of microwaves. This is followed by a third section in which an attempt is made to show in what ways microwave engineering differs from the engineering of communication systems at lower frequencies. In addition, a number of microwave devices are introduced to provide examples of the types of devices and circuit elements that are examined in greater detail later on in the text.

1 MICROWAVE FREQUENCIES

The descriptive term *microwaves* is used to describe electromagnetic waves with wavelengths ranging from 1 cm to 1 m. The corresponding frequency range is 300 MHz up to 30 GHz for 1-cm-wavelength waves. Electromagnetic waves with wavelengths ranging from 1 to 10 mm are called millimeter waves. The infrared radiation spectrum comprises electromagnetic waves with wavelengths in the range 1 μ m (10^{-6} m) up to 1 mm. Beyond the infrared range is the visible optical spectrum, the ultraviolet spectrum, and finally x-rays. Several different classification schemes are in use to designate frequency bands in the electromagnetic spectrum. These classification schemes are summarized in Tables 1.1 and 1.2. The radar band classification came into use during World War II and is still in common use today even though the new military band classification is the recommended one.

In the UHF band up to around a frequency of 1 GHz, most communications circuits are constructed using lumped-parameter circuit compo-

4 Circuit Theory for Waveguiding Systems	220
4.1 Equivalent Voltages and Currents	221
4.2 Impedance Description of Waveguide Elements and Circuits	224
One-Port Circuits	224
Lossless One-Port Termination	228
*4.3 Foster's Reactance Theorem	230
*4.4 Even and Odd Properties of Z_m	232
4.5 N-Port Circuits	233
Proof of Symmetry for the Impedance Matrix	235
Proof of Imaginary Nature of [Z] for a Lossless Junction	236
Normalized Impedance and Admittance Matrices	237
4.6 Two-Port Junctions	238
Some Equivalent Two-Port Circuits	245
4.7 Scattering-Matrix Formulation	248
Symmetry' of Scattering Matrix	250
Scattering Matrix for a Lossless Junction	251
4.8 Scattering Matrix for a Two-Port Junction	254
4.9 Transmission-Matrix Representation	257
Voltage-Current Transmission Matrix	257
Wave-Amplitude Transmission Matrix	259
*4.10 Signal Flow Graphs	260
*4.11 Generalized Scattering Matrix for Power Waves	268
*4.12 Excitation of Waveguides	276
Probe Coupling in a Rectangular Waveguide	276
Radiation from Linear Current Elements	
Radiation from Current Loops	
*4.13 Waveguide Coupling by Apertures	
Aperture in a Transverse Wall	286
Aperture in Broad Wall of a Waveguide	290
Problems	
References	294
5 Impedance Transformation and Matching	
5.1 Smith Chart	304
5.2 Impedance Matching with Reactive Elements	308
Single-Stub Matching	309
5.3 Double-Stub Matching Network	312
5.4 Triple-Stub Tuner	317
5.5 Impedance Matching with Lumped Elements	319
Circuit Q and Bandwidth	325
5.6 Design of Complex Impedance Terminations	330
5.7 Invariant Property of Impedance Mismatch Factor	334
5.8 Waveguide Reactive Elements	339
Shunt Inductive Elements	340
Shunt Capacitive Elements	341
Waveguide Stub Tuners	342

5.9	Quarter-Wave Transformers	343
5.10	Theory of Small Reflections	347
5.11	Approximate Theory for Multisection Quarter-Wave Transformers	348
5.12	Binomial Transformer	350
5.13	Chebyshev Transformer	352
•5.14	Chebyshev Transformer (Exact Results)	356
5.15	Filter Design Based on Quarter-Wave-Transformer Prototype Circuit	360
	Junction Capacitance and Length Compensation	365
5.16	Tapered Transmission Lines	370
	Exponential Taper	372
	Taper with Triangular Distribution	372
*5.17	Synthesis of Transmission-Line Tapers	373
*5.18	Chebyshev Taper	380
*5.19	Exact Equation for the Reflection Coefficient	383
	Problems	387
	References	393
6	Passive Microwave Devices	394
6.1	Terminations	394
	Variable Short Circuit	395
6.2	Attenuators	397
	Electronically Controlled Attenuators	400
6.3	Phase Shifters	404
	Rotary Phase Shifter	404
	Electronically Controlled Phase Shifters	409
6.4	Directional Couplers	413
	Directional-Coupler Designs	416
	Coupled-Line Directional Couplers	427
	Branch-Line Directional Coupler	432
	Lange Directional Coupler	434
6.5	Hybrid Junctions	435
	Magic T	435
	Hybrid Ring	437
6.6	Power Dividers	442
6.7	Microwave Propagation in Ferrites	450
6.8	Faraday Rotation	460
6.9	Microwave Devices Employing Faraday Rotation	464
	Gyrator	464
	Isolator	466
	Resonance Isolator	467
6.10	Circulators	468
	Three-Port Circulator	471
	Field Analysis of Three-Port Circulator	473
6.11	Other Ferrite Devices	476
	Problems	476
	References	479

7	Electromagnetic Resonators	481
7.1	Resonant Circuits	481
7.2	Transmission-Line Resonant Circuits	485
	Series Resonance; Short-Circuited Line	485
	Open-Circuited Line	487
	Antiresonance	488
7.3	Microstrip Resonators	490
	Circular Disk Resonator	496
7.4	Microwave Cavities	500
	Rectangular Cavity	500
	Cylindrical Cavity	504
7.5	Dielectric Resonators	508
7.6	Equivalent Circuits for Cavities	517
	Aperture-Coupled Cavity	517
	Loop-Coupled Cavity	523
*7.7	Field Expansion in a General Cavity	525
	Cavity Field Expansions in Terms of Short-Circuit Modes	527
	Electric Field Expansion	528
	Orthogonality Properties	529
	Magnetic Field Expansion	531
	Orthogonality Properties	531
	Relationship between E _z and H _z Modes	532
*7.8	Oscillations in a Source-Free Cavity	533
	Cavity with Lossy Walls	534
	Degenerate Modes	536
*7.9	Excitation of Cavities	538
*7.10	Cavity Perturbation Theory	541
	Problems	545
	References	548
8	Periodic Structures and Filters	550
8.1	Capacitively Loaded Transmission-Line-Circuit Analysis	551
8.2	Wave Analysis of Periodic Structures	557
8.3	Periodic Structures Composed of Unsymmetrical Two-Port Networks	559
8.4	Terminated Periodic Structures	560
8.5	Matching of Periodic Structures	563
8.6	k_0 -fl Diagram	564
*8.7	Group Velocity and Energy Flow	566
8.8	Floquet's Theorem and Spatial Harmonics	569
8.9	Periodic Structures for Traveling-Wave Tubes	571
	Periodic Structures for Millimeter-Wave Traveling-Wave Tubes	W1
8.10	Sheath Helix	580
*8.11	Some General Properties of a Helix	583
8.12	Introduction to Microwave Filters	585
8.13	Image-Parameter Method of Filter Design	587

8.14	Filter Design by Insertion-Loss Method	591
8.15	Specification of Power Loss Ratio	592
	Maximally Flat Filter Characteristic	593
	Chebyshev Filter	593
8.16	Some Low-Pass-Filter Designs	595
8.17	Frequency Transformations	598
	Frequency Expansion	599
	Low-Pass to High-Pass Transformation	599
	Low-Pass to Bandpass Transformation	600
	Period Bandpass Mapping	602
8.18	Impedance and Admittance Inverters	603
8.19	A Microstrip Half-Wave Filter	617
8.20	Microstrip Parallel Coupled Filter	626
8.21	Quarter-Wave-Coupled Cavity Filters	635
8.22	Direct-Coupled Cavity Filters	639
8.23	Other Types of Filters	642
	Problems	642
	References	647
9	Microwave Tubes	648
9.1	Introduction	648
9.2	Electron Beams with dc Conditions	650
	Ion-Neutralized Beam	650
	Beam with Axially Confined Flow	651
	Brillouin Flow	652
9.3	Space-Charge Waves on Beams with Confined Flow	654
9.4	Space-Charge Waves on Unfocused Beams	661
9.5	Ac Power Relations	667
9.6	Velocity Modulation	670
9.7	Two-Cavity Klystron	678
	Excitation of a Cylindrical Cavity	679
	Cavity Excitation by a Velocity-Modulated Beam	683
9.8	Reflex Klystron	686
9.9	Magnetron	690
9.10	O-Type Traveling-Wave Tube	692
9.11	AZ-Type Traveling-Wave Tube	699
9.12	Gyrotrons	701
	Field-Particle Interaction in a Gyrototron	703
9.13	Other Types of Microwave Tubes	708
	Problems	709
	References	712
10	Solid-State Amplifiers	713
10.1	Bipolar Transistors	716
	Transistor Biasing	720
10.2	Field-Effect Transistors	721
	FET Biasing	724

10.3	Circle-Mapping Properties of Bilinear Transformations	725
10.4	Microwave Amplifier Design Using S,, Parameters	726
10.5	Amplifier Power Gain	728
	Derivation of Expressions for Gain	730
10.6	Amplifier Stability Criteria	735
	Conditionally Stable Devices	740
10.7	Constant Power-Gain Circles	744
	Properties of the Constant Gain Circles	746
	Stable Devices	746
	Unstable Devices	750
10.8	Basic Noise Theory	760
	Filtered Noise	762
	Noise in Active Devices	765
	Noisy Two-Port Networks	766
10.9	Low-Noise Amplifier Design	767
	Noise Figure	768
	Noise Figure for Cascaded Stages	770
	Constant Noise-Figure Circles	772
10.10	Constant Mismatch Circles	776
	Constant Input Mismatch Circle	778
	Output Impedance-Mismatch Circle	780
10.11	Microwave Amplifier Design	780
	Single-Stage Amplifier Design	781
	Design of Second Stage for a Two-Stage Amplifier	788
10.12	Other Aspects of Microwave Amplifier Design	793
	Problems	795
	References	798
11	Parametric Amplifiers	799
11.1	<i>p-n</i> Junction Diodes	800
11.2	Manley-Rowe Relations	BIM
11.3	Linearized Equations for Parametric Amplifiers	807
11.4	Parametric Up-Converter	809
11.5	Negative-Resistance Parametric Amplifier	814
11.6	Noise Properties of Parametric Amplifiers	821
	Problems	829
	References	830
12	Oscillators and Mixers	831
12.1	Gunn Oscillators	
	Gunn Oscillator Circuits	
12.2	IMP ATT Diodes	837
12.3	Transistor Oscillators	840
12.4	Three-Port Description of a Transistor	
12.5	Oscillator Circuits	
12.6	Oscillator Design	851

Y2.7	Mixers	856
	Linear Mixer Operation	861
	Nonlinear Mixer Operation	862
12.8	Mixer Noise Figure	864
12.9	Balanced Mixers	865
12.10	Other Types of Mixers	868
12.11	Mixer Analysis Using Harmonic Balancing	869
	Problems	873
	References	875
 Appendices		
I	Useful Relations from Vector Analysis	876
1.1	Vector Algebra	876
1.2	Vector Operations in Common Coordinate Systems	877
	Rectangular Coordinates	877
	Cylindrical Coordinates	877
	Spherical Coordinates	878
1.3	Vector Identities	879
1.4	Green's Identities	880
II	Bessel Functions	881
II. 1	Ordinary Bessel Functions	881
II.2	Modified Bessel Functions	883
	References	885
III	Conformal Mapping Techniques	886
III.1	Conformal Mapping	886
III.2	Elliptic Sine Function	889
III.3	Capacitance between Two Parallel Strips	892
III.4	Strip Transmission Line	896
III.5	Conductor Loss	898
III.6	Conductor Losses for a Microstrip Transmission Line	903
III.7	Attenuation for a Coplanar Line	905
IV	Physical Constants and Other Data	911
IV. 1	Physical Constants	911
IV.2	Conductivities of Materials	912
IV.3	Dielectric Constants of Materials	912
IV.4	Skin Depth in Copper	912
Index		913

PREFACE

The first edition of *Foundations for Microwave Engineering* was published in 1966. The text has remained continuously in use since that time, but it has become clear that it no longer gives an adequate account of modern microwave engineering practice. Since the publication of the first edition there has been a dramatic advance in the microwave field brought about by the development of solid state transistors that can provide amplification and signal generation well into the millimeter wavelength region. Along with the widespread use of solid state devices, compatible transmission line structures and passive components were developed that could be integrated with the solid state devices into compact miniaturized microwave systems. These developments made it mandatory that the text be thoroughly revised if it were to continue serving the needs of the student and the practicing microwave engineer.

In the revised addition I have adhered to the same general philosophy that governed the preparation of the first edition. Fundamental principles are stressed and complete derivations are provided for all significant formulas and relationships. All important fundamental concepts and principles are covered to the extent possible within a text of reasonable size. The applications of basic theory and principles are illustrated through detailed analysis of a large number of important components that find widespread use in practical microwave systems.

Chapter 1 is an updated introductory chapter. Chapter 2 is essentially the same as in the original edition and provides a comprehensive summary of basic electromagnetic theory that is needed as background for proper understanding of the rest of the text. Many students will already have knowledge of this material before they pursue a course in microwave engineering. For these students, Chapter 2 will serve as a concise reference or review of familiar material.

Chapter 3 is very different from that in the first edition. The first part of this chapter provides a more basic introduction to waves on transmission

lines using distributed circuit models. The propagation of pulse signals is also covered. The second part of this chapter is a long section covering the characteristics of planar transmission lines, such as microstrip lines, coupled microstrip lines, strip lines, and coplanar lines or waveguides. The treatment is considerably broader than what is available in any other current text. Most of the formulas for the quasi-TEM mode parameters are derived using conformal mapping methods in a new Appendix III and are not just quoted from the literature. Several new formulas for attenuation have been derived as well as suitable modifications of existing formulas to account for anisotropic substrates. The last part of the chapter covers the basic properties of rectangular and circular waveguides, as in the original edition.

Chapter 4 develops the basic microwave circuit theory and includes detailed discussions of the impedance, admittance and scattering matrix descriptions of microwave junctions. New material has been added on signal flow graphs and the generalized scattering matrix for power waves. The material on small aperture coupling has been updated to include radiation reaction that will account for power transmission through an aperture and thereby lead to physically meaningful equivalent circuits for small apertures.

Chapter 5 treats a number of topics related to impedance matching and transformations. The old topic of impedance matching with lumped reactive elements has been revived because this is now frequently used in microwave integrated circuits. The design of complex load terminations has also been included because this is required for microwave solid state amplifier design. The available power at any point in a lossless reciprocal network is an invariant quantity. This concept is explained in terms of the impedance mismatch factor. The invariance of the impedance mismatch factor places an important constraint on the design of interstage matching networks in a microwave amplifier and is used in Chapter 10 in the design of microwave amplifiers. The last part of Chapter 5 discusses multisection quarter-wave transformers and tapered transmission lines. A new example of a microstrip half-wave filter design based on the quarter-wave transformer as a prototype circuit has been included.

A variety of passive components are described along with detailed analysis in Chapter 6. In addition to those components described in the original edition, new material has been added on coupled-microstrip-line directional couplers, the branch-line coupler, hybrid junctions, and the Wilkinson power divider. New material on electronic controlled attenuators and phase shifters has also been added.

Chapter 7 on resonators has been expanded to include new material on microstrip resonators and dielectric resonators. The old material on Fabry-Perot resonators has been deleted in order to make room for a short section on cavity perturbation theory.

Chapter 8 on periodic structures and filters now includes a detailed treatment of gap-coupled and edge-coupled microstrip filters. The treatment

of admittance and impedance inverters was rewritten in order to more fully explain the use of inverters in filter design.

Apart from a brief discussion of gyrotron tubes, Chapter 9 on microwave tubes remains essentially the same as in the first edition.

The old Chapter 10 on masers has been replaced by a new chapter on microwave solid state amplifier design. This chapter gives a complete discussion of the scattering matrix approach to small signal narrow band amplifier design. The treatment is self-contained and all important relations for gain, stability, and low noise design are derived. A design strategy for low noise single stage and double stage amplifiers is developed along with considerations for the necessary tradeoffs that must be made between input and output VSWR, gain, low noise figure, and stability.

The original Chapter 11 on parametric amplifiers has been retained without any change.

A new Chapter 12 on oscillators and mixers has been added. This chapter is of limited scope because of the need to keep the overall length of the text within reasonable bounds. Solid state oscillators using Gunn devices and IMPATT diodes are described in a qualitative way only. An introduction to transistor oscillator design based on small signal scattering matrix parameters is provided. Included in this discussion is the relationship between the two-port and three-port scattering matrix description of a transistor because this is needed in order to efficiently analyze the effect of an impedance inserted in series with one of the transistor leads for feedback purposes,

Many textbooks provide introductory treatments of diode mixers without any consideration of the embedding network. Such treatments do not provide a good understanding of diode mixers because it is the impedance properties of the embedding network that determine the diode voltages at the various harmonic frequencies. The introductory treatment of diode mixers in Chapter 12 does include the embedding network and this should provide the student with a more complete understanding of mixer analysis and design. The last part of the chapter describes the harmonic balancing method for the analysis of mixers.

I have tried to provide a broad, comprehensive, and self-contained treatment of the fundamental theory and principles, and the methods of analysis and design that are the foundations for microwave engineering. There are, of course, limitations because all books must have a finite length. Many references have been included for the benefit of the reader who wishes to pursue a given topic in greater depth or refer to the original papers that a lot of the material has been based on. This text, in many respects, is a compilation of the work of a great many people. Unfortunately, it has not been possible to always give proper credit to those who were the originators of new concepts and the inventors of new devices.

It is my belief that the revised edition will prove to be useful for both senior elective as well as beginning graduate level courses in microwave engineering, and will also serve as a useful reference source on fundamental

TABLE 1.1
Frequency band designation

Frequency band	Designation	Typical service
3-30 kHz	Very low frequency (VLF)	Navigation, sonar
30-300 kHz	Low frequency (LF)	Radio beacons, navigational aids
300-3,000 kHz	Medium frequency (MF)	AM broadcasting, maritime radio. Coast Guard communication, direction finding
3-30 MHz	High frequency (HF)	Telephone, telegraph, and facsimile; shortwave international broadcasting; amateur radio; citizen's band; ship-to-coast and ship-to-aircraft communication
30-300 MHz	Very" high frequency (VHF)	Television, FM broadcast, air-traffic control, police, taxicab mobile radio, navigational aids
300-3,000 MHz	Ultrahigh frequency (UHF>	Television, satellite communication, radiosonde, surveillance radar, navigational aids
3-30 GHz	Superhigh frequency (SHF)	Airborne radar, microwave links, common-carrier land mobile communication, satellite communication
30-300 GHz	Extreme high frequency (EHF)	Radar, experimental

TABLE 1.2
Microwave frequency band designation

Frequency	Microwave band designation	
	Old	New
500-1,000 MHz	VHF	C
1-2 GHz	L	D
2-3 GHz	S	E
3-4 GHz	S	F
4-6 GHz	G	G
6-8 GHz	C	H
8-10 GHz	X	I
10-12.4 GHz	X	J
12.4-18 GHz	Ku	J
18-20 GHz	K	J
20-26.5 GHz	K	8
26.5-40 GHz	Ka	K

principles for the practicing microwave engineer. There is clearly much more material in the revised edition than can be covered in a one semester course. The last four chapters alone would provide sufficient material for a one semester course on active microwave circuits.

As an instructor I have always believed that it was very important to fully understand where formulas came from and how they are derived in order to present the material to students in a meaningful way. It is for this reason that I have attempted to make the text self-contained. In presenting many of the topics to undergraduate students I will only outline the basic approach used and will omit the details. It is my hope that other instructors will also view the detailed derivations that are provided in the text as a useful source of information in preparing a microwave engineering course and not as material that must always be presented in class. A number of topics that can be omitted in an undergraduate course are identified by a star. The problems based on these sections are also identified by a star.

In recent years the microwave engineering course that I have taught to seniors at Case Western Reserve University has drawn heavily on the material in Chapters 3 through 5, which is very basic core material. In addition, topics have been selected from Chapters 6 and 7 on components and resonators in order to illustrate the application of basic microwave circuit theory. The last quarter of the semester has been generally devoted to microwave solid state amplifier design along with a brief coverage of oscillators and mixers.

A better selection of problems and a solutions manual has been prepared for the revised edition. Over the past several years I have also prepared a number of short stand alone computer programs that provide useful tools to remove the drudgery of solving many of the homework problems. These programs are included on a floppy disk along with user instructions as part of the solutions manual. The programs cover the calculation of the characteristics of various planar transmission lines, including attenuation; the cutoff frequency, propagation constant, and attenuation of the dominant mode in rectangular and circular waveguides; impedance transformation along a transmission line; input and output impedances, admittances, and reflection coefficients for a linear two-port, which can be described in terms of impedance, admittance, or scattering matrix parameters; double-stub and lumped element impedance matching with frequency scans; two-port and three-port scattering matrix parameters for a transistor; and a rather long program that implements a design strategy for low noise one- and two-stage microwave amplifiers with various imposed constraints. Students have generally found these programs to be of significant help in problem solving. They have enjoyed working with the microwave amplifier design program. Without a computer program, the design of a microwave amplifier using potentially unstable devices and subject to various constraints on gain, noise figure, and input and output VSWR, is not feasible for students to carry out. The scope of each program

has been purposefully limited in order to ensure that the student will be fully aware of the solution strategy involved.

Many users of the first edition have provided me with helpful comments on the original material. In addition, I have received many helpful comments and suggestions from the following reviewers of the material for the revised edition. They are Chin-Lin Chen, Purdue University; M. Yousif El-Ibiary, University of Oklahoma; Irving Kaufman, Arizona State University; Stuart Long, University of Houston; Glenn S. Smith, Georgia Institute of Technology; and Robert J. Weber, Iowa State University. For the most part then suggestions and recommendations have been incorporated.

The new material for the revised edition was typed by Sue Sava. I would like to acknowledge the professional skill with which she prepared this material as well as her willingness to rearrange her schedule so as to meet various deadlines.

The last acknowledgment is to my wife Kathleen, who was willing to give up many other activities so that the revision could be carried out. Her encouragement and support of the project never faltered, and without it the revision could not have been undertaken.

Robert E. Collin

nenta. In the frequency range from 1 up to 100 GHz, lumped circuit elements are usually replaced by transmission-line and waveguide components. Thus by the term *microwaue engineering* we shall mean generally the engineering and design of information-handling systems in the frequency range from 1 to 100 GHz corresponding to wavelengths as long as 30 cm and as short as 3 mm. At shorter wavelengths we have what can be called optical engineering since many of the techniques used are derived from classical optical techniques. The characteristic feature of microwave engineering is the short wavelengths involved, these being of the same order of magnitude as the circuit elements and devices employed.

The short wavelengths involved in turn mean that the propagation time for electrical effects from one point in a circuit to another point is comparable with the period of the oscillating currents and charges in the system. As a result, conventional low-frequency circuit analysis based on Kirchhoff's laws and voltage-current concepts no longer suffices for an adequate description of the electrical phenomena taking place. It is necessary instead to carry out the analysis in terms of a description of the electric and magnetic fields associated with the device. In essence, it might be said, microwave engineering is applied electromagnetic fields engineering. For this reason the successful engineer in this area must have a good working knowledge of electromagnetic field theory.

There is no distinct frequency boundary at which lumped-parameter circuit elements must be replaced by distributed circuit elements. With modern technological processes it is possible to construct printed-circuit inductors that are so small that they retain their lumped-parameter characteristics at frequencies as high as 10 GHz or even higher. Likewise, optical components, such as parabolic reflectors and lenses, are used to focus microwaves with wavelengths as long as 1 m or more. Consequently, the microwave *engineer will frequently employ* low-frequency lumped-parameter circuit elements, such as miniaturized inductors and capacitors, as well as optical devices in the design of a microwave system.

MICROWAVE APPLICATIONS

The great interest in microwave frequencies arises for a variety of reasons. Basic among these is the ever-increasing need for more radio-frequency-spectrum space and the rather unique uses to which microwave frequencies can be applied. When it is noted that the frequency range 10^9 to 10^{12} Hz contains a thousand sections like the frequency spectrum from 0 to 10^9 Hz, the value of developing the microwave band as a means of increasing the available usable frequency spectrum may be readily appreciated.

At one time (during World War II and shortly afterward), microwave engineering was almost synonymous with radar (Radio Detection And Ranging) engineering because of the great stimulus given to the development of microwave systems by the need for high-resolution radar capable of

detecting and locating enemy planes and ships. Even today radar, in its many varied forms, such as missile-tracking radar, fire-control radar, weather-detecting radar, missile-guidance radar, airport traffic-control radar, etc., represents a major use of microwave frequencies. This use arises predominantly from the need to have antennas that will radiate essentially all the transmitter power into a narrow pencil-like beam similar to that produced by an optical searchlight. The ability of an antenna to concentrate radiation into a *narrow beam* is limited by *diffraction effects*, which in turn are governed by the relative size of the radiating aperture in terms of wavelengths. For example, a parabolic reflector-type antenna produces a pencil beam of radiated energy having an angular beam width of $140^\circ/(Z/A_0)$, where D is the diameter of the parabola and A_0 is the wavelength. A 90-cm (about 3 ft) parabola can thus produce a 4.7° beam at a frequency of 10'' Hz, i.e., at a wavelength of 3 cm. A beam of this type can give reasonably accurate position data for a target being observed by the radar. To achieve comparable performance at a frequency of 100 MHz would require a 300-ft parabola, a size much too large to be carried aboard an airplane.

In more recent years microwave frequencies have also come into widespread use in communication links, generally referred to as microwave links. Since the propagation of microwaves is effectively along line-of-sight paths, these links employ high towers with reflector or lens-type antennas as repeater stations spaced along the communication path. Such links are a familiar sight to the motorist traveling across the country because of their frequent use by highway authorities, utility companies, and television networks. A further interesting means of communication by microwaves is the use of satellites as microwave relay stations. The first of these, the Telstar, launched in July 1962, provided the first transmission of live television programs from the United States to Europe.

Since that time a large number of satellites have been deployed for communication purposes, as well as for surveillance and collecting data on atmospheric and weather conditions. For direct television broadcasting the most heavily used band is the C band. The up-link frequency used is in the 5.9- to 6.4-GHz band and the receive or down-link frequency band is between 3.7 and 4.2 GHz. For home reception an 8-ft-diameter parabolic reflector antenna is commonly used. A second frequency band has also been allocated for direct television broadcasting. For this second band the up-link frequency is in the 14- to 14.5-GHz range and the down-link frequencies are between 10.95 and 11.2 GHz and 11.45 and 11.7 GHz. In this band a receiving parabolic antenna with a 3-ft diameter is adequate. At the present time this frequency band is not being used to *any great extent in the United States*. It is more widely used in Europe and Japan.

Terrestrial microwave links have been used for many years. The TD-2 system was put into service in 1948 as part of the Bell Network. It operated in the 3.7- to 4.2-GHz band and had 480 voice circuits, each occupying a

3.1-kHz bandwidth. In 1974, the TN-1 system operating in the 10.7- to 11.7-GHz band was put into operation. This system had a capacity of 1,800 voice circuits or one video channel with a 4.5-MHz bandwidth. Since that time the use of terrestrial microwave links has grown rapidly.

At the present time most *communication systems* are shifting to the use of digital transmission, i.e., analog signals are digitized before transmission. Microwave digital communication system development is progressing rapidly. In the early systems simple modulation schemes were used and resulted in inefficient use of the available frequency spectrum. The development of 64-state quadrature amplitude modulation (64-QAM) has made it possible to transmit 2,016 voice channels within a single 30-MHz RF channel. This is competitive with FM analog modulation schemes for voice. The next step up is the 256-QAM system which is under development.

For the ready processing and handling of a modulated carrier, modulation sidebands can be only a few percent of the carrier frequency. It is thus seen that the carrier frequency must be in the microwave range for efficient transmission of many television programs over one link. Without the development of microwave systems, our communications facilities would have been severely overloaded and totally inadequate for present operations.

Even though such uses of microwaves are of great importance, the applications of microwaves and microwave technology extend much further, into a variety of areas of basic and applied research, and including a number of diverse practical devices, such as microwave ovens that can cook a small roast in just a few minutes. Some of these specific applications are briefly *discussed* below.

Waveguides periodically loaded with shunt susceptance elements support slow waves having velocities much less than the velocity of light, and are used in linear accelerators. These produce high-energy beams of charged particles for use in atomic and nuclear research. The slow-traveling electromagnetic waves interact very efficiently with charged-particle beams having the same velocity, and thereby give up energy to the beam. Another possibility is for the energy in an electron beam to be given up to the electromagnetic wave, with resultant amplification. This latter device is the traveling-wave tube, and is examined in detail in a later chapter.

Sensitive microwave receivers are *used in radio astronomy* to detect and study the electromagnetic radiation from the sun and a number of radio stars that emit radiation in this band. Such receivers are also used to detect the noise radiated from plasmas (an approximately neutral collection of electrons and ions, e.g., a gas discharge). The information obtained enables scientists to analyze and predict the various mechanisms responsible for plasma radiation. Microwave radiometers are also used to map atmospheric temperature profiles, moisture conditions in soils and crops, and for other remote-sensing applications as well.

Molecular, atomic, and nuclear systems exhibit various resonance phenomena under the action of periodic forces arising from an applied

electromagnetic field. Many of these resonances occur in the microwave range; hence microwaves have provided a very powerful experimental probe for the study of basic properties of materials. Out of this research on materials have come many useful devices, such as some of the nonreciprocal devices employing ferrites, several solid-state microwave amplifiers and oscillators, e.g., masers, and even the coherent-light generator and amplifier (laser).

The development of the laser, a generator of essentially monochromatic (single-frequency) coherent-light waves, has stimulated a great interest in the possibilities of developing communication systems at optical wavelengths. This frequency band is sometimes referred to as the *ultramicrowave* band. With some modification, a good deal of the present microwave technology can be exploited in the development of optical systems. For this reason, familiarity with conventional microwave theory and devices provides a good background for work in the new frontiers of the electromagnetic spectrum.

The domestic microwave oven operates at 2,450 MHz and uses a magnetron tube with a power output of 500 to 1000 W. For industrial heating applications, such as drying grain, manufacturing wood and paper products, and material curing, the frequencies of 915 and 2,450 MHz have been assigned. Microwave radiation has also found some application for medical hyperthermia or localized heating of tumors.

It is not possible here to give a complete account of all the applications of microwaves that are being made. The brief look at some of these, as given above, should convince the reader that this portion of the radio spectrum offers many unusual and unique features. Although the microwave engineering field may now be considered a mature and well-developed one, the opportunities for further development of devices, techniques, and applications to communications, industry, and basic research are still excellent.

1.3 MICROWAVE CIRCUIT ELEMENTS AND ANALYSIS

At frequencies where the wavelength is several orders of magnitude larger than the greatest dimensions of the circuit or system being examined, conventional circuit elements such as capacitors, inductors, resistors, electron tubes, and transistors are the basic building blocks for the information transmitting, receiving, and processing circuits used. The description or analysis of such circuits may be adequately carried out in terms of loop currents and node voltages without consideration of propagation effects. The time delay between cause and effect at different points in these circuits is so small compared with the period of the applied signal as to be negligible. It might be noted here that an electromagnetic wave propagates a distance of one wavelength in a time interval equal to one period of a sinusoidally

time-varying applied signal. As a consequence, when the distances involved are short compared with a wavelength λ_0 (λ_0 = velocity of light/frequency), the time delay is not significant. As the frequency is raised to a point where the wavelength is no longer large compared with the circuit dimensions, propagation effects can no longer be ignored. A further effect is the great relative increase in the impedance of connecting leads, terminals, etc., and the effect of distributed (stray) capacitance and inductance. In addition, currents circulating in unshielded circuits comparable in size with a wavelength are very effective in radiating electromagnetic waves. The net effect of all this is to make most conventional low-frequency circuit elements and circuits hopelessly inadequate at microwave frequencies.

If a rather general viewpoint is adopted, one may classify resistors, inductors, and capacitors as elements that dissipate electric energy, store magnetic energy, and store electric energy, respectively. The fact that such elements have the form encountered in practice, e.g., a coil of wire for an inductor, is incidental to the function they perform. The construction used in practical elements may be considered just a convenient way to build these devices so that they will exhibit the desired electrical properties. As is well known, many of these circuit elements do not behave in the desired manner at high frequencies. For example, a coil of wire may be an excellent inductor at 1 MHz, but at 50 MHz it may be an equally good capacitor because of the predominating effect of interturn capacitance. Even though practical low-frequency resistors, inductors, and capacitors do not function in the desired manner at microwave frequencies, this does not mean that such energy-dissipating and storage elements cannot be constructed at microwave frequencies. On the contrary, there are many equivalent inductive and capacitive devices for use at microwave frequencies. Their geometrical form is quite different, but they can be and are used for much the same purposes, such as impedance matching, resonant circuits, etc. Perhaps the most significant electrical difference is the generally much more involved frequency dependence of these equivalent inductors and capacitors at microwave frequencies.

Low-frequency electron tubes are also limited to a maximum useful frequency range bordering on the lower edge of the microwave band. The limitation arises mainly from the finite transit time of the electron beam from the cathode to the control grid. When this transit time becomes comparable with the period of the signal being amplified, the tube ceases to perform in the desired manner. Decreasing the electrode spacing permits these tubes to be used up to frequencies of a few thousand megahertz, but the power output is limited and the noise characteristics are poor. The development of new types of tubes for generation of microwave frequencies was essential to the exploitation of this frequency band. Fortunately, several new principles of operation, such as velocity modulation of the electron beam and beam interaction with slow electromagnetic waves, were discovered that enabled the necessary generation of microwaves to be carried out.

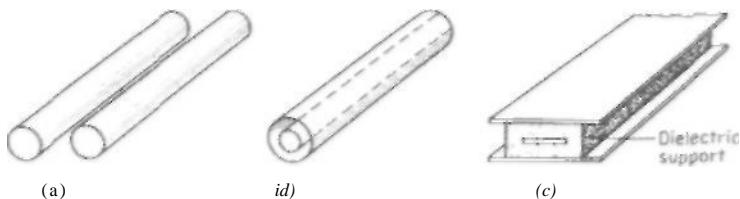


FIGURE 1.1

Some common transmission lines, (a) Two-conductor line; (b) coaxial line; (c) shielded strip line.

These fundamental principles with applications are discussed in a later chapter.

For low-power applications microwave tubes have been largely replaced by solid-state devices, such as transistors and negative resistance diodes. However, for high-power applications microwave tubes are still necessary.

One of the essential requirements in a microwave circuit is the ability to transfer signal power from one point to another without radiation loss. This requires the transport of electromagnetic energy in the form of a propagating wave. A variety of such structures have been developed that can guide electromagnetic waves from one point to another without radiation loss. The simplest guiding structure, from an analysis point of view, is the transmission line. Several of these, such as the open two-conductor line, coaxial line, and shielded strip line, illustrated in Fig. 1.1, are in common use at the lower microwave frequencies.

At the higher microwave frequencies, notably at wavelengths below 10 cm, hollow-pipe waveguides, as illustrated in Fig. 1.2, are often preferred to transmission lines because of better electrical and mechanical properties. The waveguide with rectangular cross section is by far the most common type. The circular guide is not nearly as widely used.

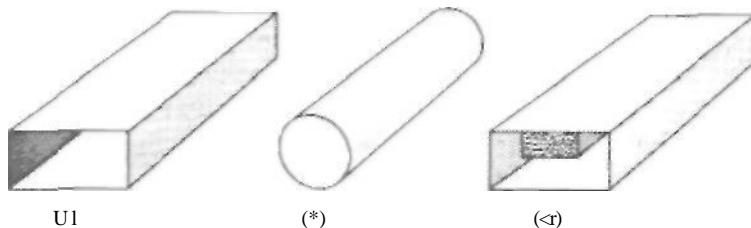


FIGURE 1.2

Some common hollow-pipe waveguides, (a) Rectangular guide; (b) circular guide; (c) ridge guide.

The ridge-loaded rectangular guide illustrated in Fig. 1.2c is sometimes used in place of the standard rectangular guide because of better impedance properties and a greater bandwidth *of operation*. In *addition to* these standard-type guides, a variety of other cross sections, e.g., elliptical, may also be used.

Another class of waveguides, of more recent origin, is surface waveguides. An *example of this type is a conducting wire coated with a thin layer* of dielectric. The wire diameter is small compared with the wavelength. Along a structure of this type it is possible to guide an electromagnetic wave. The wave is bound to the surface of the guide, exhibiting an amplitude decay that is exponential *in the radial direction away from the surface*, and hence is called a surface wave. Applications are mainly in the millimeter-wavelength range since the field does extend a distance of a wavelength or so beyond the wire, and this makes the effective guide diameter somewhat large in the centimeter-wavelength *range*. A *disadvantage of surface waveguides and open-conductor transmission lines is that radiation loss occurs whenever other obstacles are brought into the vicinity of the guide*.

The development of solid-state active devices, such as bipolar transistors *and, more notably, field-effect transistors (FET)*, has had a *dramatic impact* on the microwave engineering field. With the availability of microwave transistors, the focus on waveguides and waveguide components changed to a focus on planar transmission-line structures, such as *microstrip lines and coplanar transmission lines*. These *structures*, shown in Fig. 1.3, can be manufactured using printed-circuit techniques. They are compatible with solid-state devices in that it is easy to connect a transistor to a microstrip circuit but difficult to incorporate it as part of a waveguide circuit. By using gallium-arsenide material *it has been possible to design field-effect transistors that provide low noise and useful amplification at millimeter wavelengths*. At the lower microwave frequencies hybrid integrated microwave circuits are used. In hybrid circuit construction the transmission lines and *transmission-line components, such as matching elements*, are manufactured first and then the solid-state devices, such as diodes and transistors, are soldered into place. The current trend is toward the use of monolithic microwave integrated circuits (MMIC) in which both the transmission-line circuits and active devices *are fabricated on a single chip*. A variety of broadband MMIC amplifiers have been designed. The development of MMIC circuits for operation at frequencies up to 100 GHz is well under way.

A unique property of the transmission line is that a satisfactory analysis of its properties may be carried out by treating it as a network with distributed parameters and solving for the voltage and current waves that may propagate along the line. Other waveguides, although they have several properties similar to transmission lines, must be treated as electromagnetic boundary-value problems, and a solution for the electromagnetic fields must be determined. Fortunately, this is readily accomplished for the common

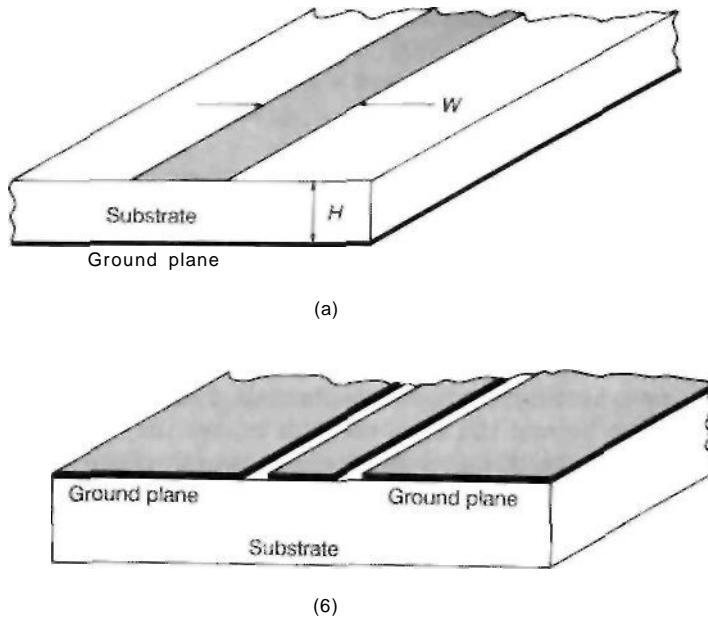


FIGURE 1.3
(a) microstrip transmission line; (b) coplanar transmission line.

waveguides used in practice. For waveguides it is not possible to define unique voltage and current that have the same significance as for a transmission line. This is one of the reasons why the field point of view is emphasized at microwave frequencies.

Associated with waveguides are a number of interesting problems related to methods of exciting fields in guides and methods of coupling energy out. Three basic coupling methods are used: (1) probe coupling, (2) loop coupling, and (3) aperture coupling between adjacent guides. They are illustrated in Fig. 1.4, and some of them are analyzed later. These coupling

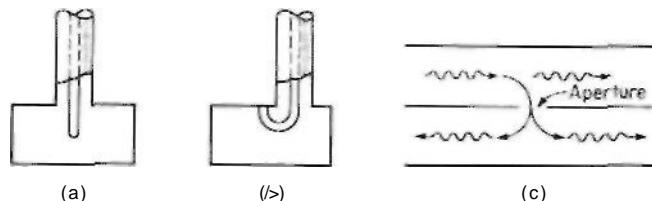


FIGURE 1.4
Basic methods of coupling energy into and out of waveguides, (a) Probe coupling; (b) loop coupling; (c) aperture coupling.

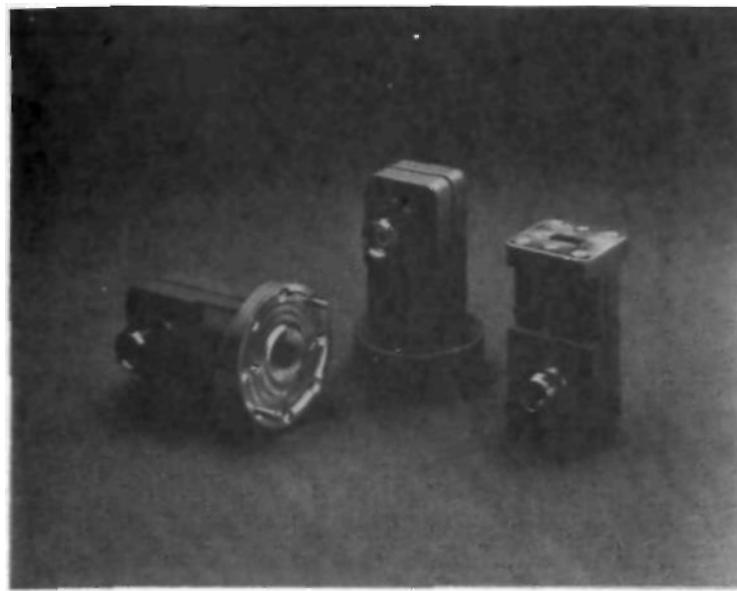


FIGURE 1.5

Waveguide-to-coaxial-line transitions that use probe coupling as shown in Fig. 1.4a. (Photograph courtesy of Ray Moskaluk, Hewlett-Packard Company.)

devices are actually small antennas that radiate into the waveguide. A photograph of a waveguide-to-coaxial-line transition is shown in Fig. 1.5.

Inductive and capacitive elements take a variety of forms at microwave frequencies. Perhaps the simplest are short-circuited sections of transmission line and waveguide. These exhibit a range of susceptance values from minus to plus infinity, depending on the length of the line, and hence may act as either inductive or capacitive elements. They may be connected as either series or shunt elements, as illustrated in Fig. 1.6. They are commonly referred to as stubs and are widely used as impedance-matching elements. In a rectangular guide thin conducting windows, or diaphragms, as illustrated in Fig. 1.7, also act as shunt susceptive elements. Their

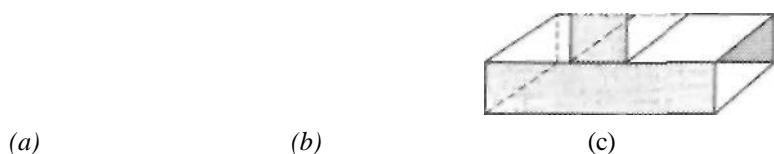
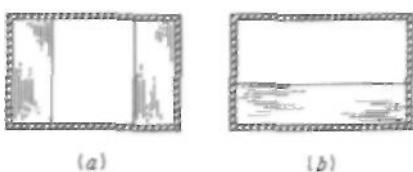
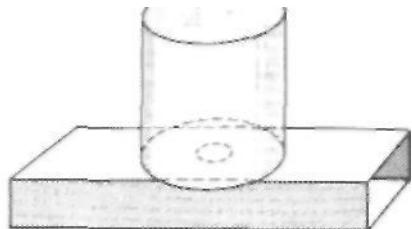


FIGURE 1.6

Stub-type reactive elements, (a) Series element; (b) shunt element; (c) waveguide stub.

**FIGURE 1.7**

Shunt susceptive elements in a waveguide.
(a) Inductive window; (b) capacitive window.

**FIGURE 1.8**

Cylindrical cavity aperture coupled to rectangular waveguide.

inductive or capacitive nature depends on whether there is more magnetic energy or electric energy stored in local fringing fields.

Resonant circuits are used both at low frequencies and at microwave frequencies to control the frequency of an oscillator and for frequency filtering. At low frequencies this function is performed by an inductor and capacitor in a series or parallel combination. Resonance occurs when there are equal average amounts of electric and magnetic energy stored. This energy oscillates back and forth between the magnetic field around the inductor and the electric field between the capacitor plates. At microwave frequencies the *LC* circuit may be replaced by a closed conducting enclosure, or cavity. The electric and magnetic energy is stored in the field within the cavity. At an infinite number of specific frequencies, the resonant frequencies, there are equal average amounts of electric and magnetic energy stored in the cavity volume. In the vicinity of any one resonant frequency, the input impedance to the cavity has the same properties as for a conventional *LC* resonant circuit. One significant feature worth noting is the very much larger *Q* values that may be obtained, these being often in excess of 10^4 , as compared with those obtainable from low-frequency *LC* circuits. Figure 1.8 illustrates a cylindrical cavity that is aperture coupled to a rectangular waveguide. Figure 1.9 is a photograph of a family of waveguide low-pass filters. The theory and design of microwave filters is given in Chap. 8. A photograph of a family of waveguide directional couplers is shown in Fig. 1.10. The design of directional couplers is covered in Chap. 6. The photograph in Fig. 1.11 shows a family of coaxial-line GaAs diode detectors.

When a number of microwave devices are connected by means of sections of transmission lines or waveguides, we obtain a microwave circuit. The analysis of the behavior of such circuits is carried out either in terms of equivalent transmission-line voltage and current waves or in terms of the amplitudes of the propagating waves. The first approach leads to an equivalent-impedance description, and the second emphasizes the wave nature of the fields and results in a *scattering-matrix* formulation. Both approaches are used in this book. Since transmission-line circuit analysis forms the basis, either directly or by analogy, for the analysis of all microwave circuits,

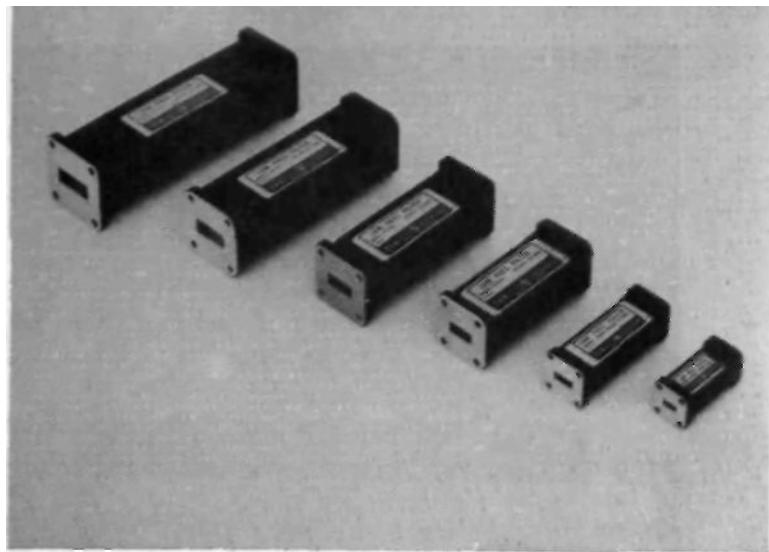


FIGURE 1.9

A family of waveguide low-pass filters for various microwave frequency bands. (Photographs courtesy of Ray Moskaluk, Hewlett Packard Company.)

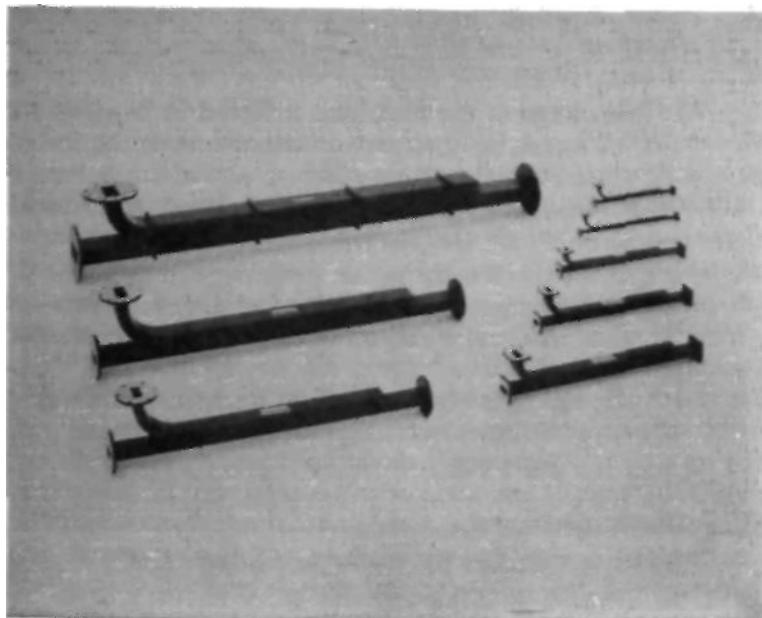


FIGURE 1.10

A family of waveguide directional couplers for various microwave frequency bands. (Photographs courtesy of Ray Moskaluk, Hewlett Packard Company.)



FIGURE 1.11

Coaxial-line GaAs diode detectors for various microwave frequency bands. (Photographs courtesy of Ray Moskaluk, Hewlett Packard Company.)

a considerable amount of attention is devoted to a fairly complete treatment of this subject early in the text. This material, together with the field analysis of the waves that may propagate along waveguides and that may exist in cavities, represents a major portion of the theory with which the microwave engineer must be familiar.

The microwave systems engineer must also have some understanding of the principles of operation of various microwave tubes, such as klystrons, magnetrons, and traveling-wave tubes, and of the newer solid-state devices, such as masers, parametric amplifiers, and microwave transistors. This is required in order to make intelligent selection and proper use of these devices. In the text sufficient work is done to provide for this minimum level of knowledge of the principles involved. A treatment that is fully adequate for the device designer is very much outside the scope of this book.

Solid-state oscillators for use as local oscillators in receiver front ends have largely replaced the klystron. Solid-state oscillators for low-power transmitters are also finding widespread use. Thus the future for microwave engineering is clearly in the direction of integrated solid-state circuits and the development of the necessary passive components needed in these circuits, which are also compatible with the fabrication methods that are used.

In the light of the foregoing discussion, it should now be apparent that the study of microwave engineering should include, among other things, at least the following:

1. Electromagnetic theory
2. Wave solutions for transmission lines and waveguides
3. Transmission-line and waveguide circuit analysis
4. Resonators and slow-wave structures
5. Microwave oscillators and amplifiers
6. Antennas
7. Microwave propagation
8. Systems considerations

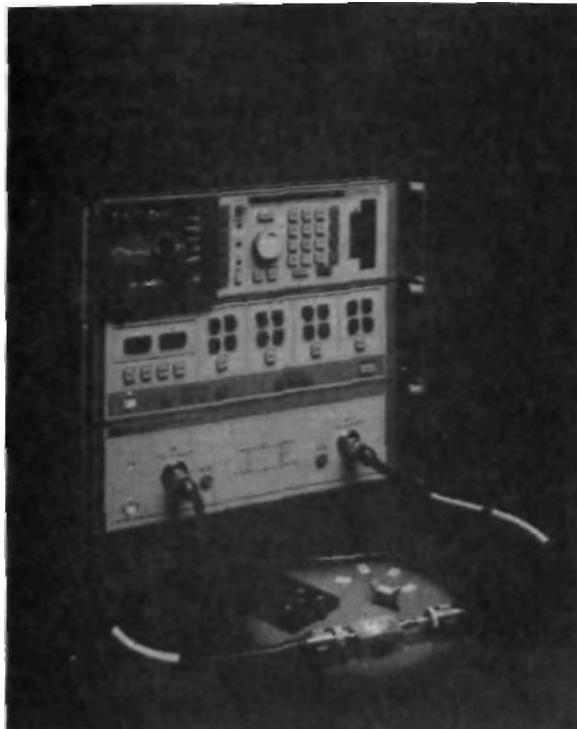


FIGURE 1.12
A microwave network analyzer used to measure scattering matrix parameters. (Photographs courtesy of Ray Moskaluk, Hewlett Packard Company.)

Apart from the last three, these are the major topics covered in the text. It is not possible to discuss in any great detail more than a few of the many microwave devices available and in current use. Therefore only a selected number of them are analyzed, to provide illustrative examples for the basic theory being developed. The available technical literature may be, and should be, consulted for information on devices not included here. Appropriate references are given throughout the text.

The number of topics treated in this text represents a good deal more than can be covered in a one-semester course. However, rather than limit the depth of treatment, it was decided to separate some of the more specialized analytical treatments of particular topics from the less analytical discussion. These specialized sections are marked with a star, and can be eliminated in a first reading without significantly interrupting the continuity of the text.¹ The student or engineer interested in the design of microwave devices, or in a fuller understanding of various aspects of microwave theory, is advised to read these special sections.

As in any engineering field, measurements are of great importance in providing the link between theory and practice at microwave frequencies.

¹ Problems based on material in these sections are also marked by a star.

Space does not permit inclusion of the subject of microwave measurements in this text. A number of excellent texts devoted entirely to microwave measurements are available, and the reader is referred to them.

There are a variety of commercially available instruments that enable microwave measurements to be carried out automatically with computer control. The photograph in Fig. 1.12 shows a network analyzer equipped to measure the scattering-matrix parameters of a microwave device. The scattering-matrix parameters, as a function of frequency, can be displayed on a Smith chart. The scattering-matrix parameters are commonly used in place of the usual impedance and admittance parameters to characterize a microwave device and are described in Chap. 4.

REFERENCES

1. Historical Perspectives of Microwave Technology, *IEEE Trans.*, vol. MTT-32, September, 1984, Special Centennial Issue.
2. Kraus, J. D.: "Antennas," 2nd ed., McGraw-Hill Book Company, New York, 1988.
3. Collin, R. E.: "Antennas and Radiowave Propagation," McGraw-Hill Book Company, New York, 1985.
4. Stutzman, W. L., and G. A. Thiele: "Antenna Theory and Design," John Wiley & Sons, Inc., New York, 1981.
5. Elliott, R. S.: "Antenna Theory and Design," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1981.
6. Balanis, C. A.: "Antenna Theory, Analysis, and Design," Harper & Row Publishers, Inc., New York, 1982.
7. Pratt, T., and C. W. Bostian: "Satellite Communications," John Wiley & Sons, New York, 1986.
8. Ivanek, F. (ed.): "Terrestrial Digital Microwave Communications," Artech House Books, Norwood, Mass., 1989.
9. Skolnik, M. I.: "Introduction to Radar Systems," McGraw-Hill Book Company, New York, 1962.
10. Montgomery, C. G.: "Technique of Microwave Measurements," McGraw-Hill Book Company, New York, 1947.
11. Ginzton, E. L.: "Microwave Measurements," McGraw-Hill Book Company, New York, 1957.
12. Bailey, A. E. (ed.): "Microwave Measurement," Peter Peregrinus, London, 1985.
13. Okress, E. C.: "Microwave Power Engineering," Academic Press, New York, 1968.
14. Ulaby, F. T., R. K. Moore, and A. K. Fung: "Microwave Remote Sensing: Active and Passive. Microwave Remote Sensing, Fundamentals and Radiometry," vol. 1, Addison-Wesley, Reading, Mass., 1981.

CHAPTER 2

ELECTROMAGNETIC THEORY

MAXWELL'S EQUATIONS

Electric and magnetic fields that vary with time are governed by physical laws described by a set of equations known collectively as Maxwell's equations. For the most part these equations were arrived at from experiments carried out by several investigators. It is not our purpose here to justify the basis for these equations, but rather to gain some understanding of their physical significance and to learn how to obtain solutions of these equations in practical situations of interest in the microwave engineering field. The electric field \mathbf{f} and magnetic field \mathbf{B} are vector fields and in general have amplitudes and directions that vary with the three spatial coordinates x , y , z and the time coordinate t . In mks units, which are used throughout, the electric field is measured in volts per meter and the magnetic field in webers per square meter. Since these fields are vector fields, the equations governing their behavior are most conveniently written in vector form.

The electric field \mathbf{E} and magnetic field \mathbf{B} are regarded as fundamental in that they give the force on a charge q moving with velocity \mathbf{v} ; that is,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (2.1)$$

It is assumed that the reader is familiar with vector analysis. However, for convenient reference, a number of vector formulas and relations are summarized in App. I.

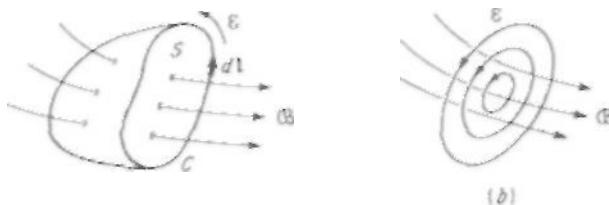


FIGURE 2.1
Illustration of Faraday's law.

where F is the force in newtons, g is the charge measured in coulombs, and v is the velocity in meters per second. This force law is called the Lorentz force equation. In addition to the \mathbf{E} and \mathbf{B} fields, it is convenient to introduce two auxiliary field vectors, namely, the electric displacement \mathbf{D} and the magnetic intensity \mathbf{H} . These are related to \mathbf{E} and \mathbf{B} through the electric and magnetic polarization of material media, a topic covered in the next section. In this section we consider fields in vacuum, or *free space*, only. In this case the following simple relationships hold:

$$\frac{1}{\mu_0} = \frac{\mathbf{D}}{\mathbf{B}} \quad (2.2a)$$

$$\mathbf{B} = \mu_0 \mathbf{H} \quad (2.2b)$$

where $\mu_0 = 4\pi \times 10^{-7} \text{ H/m}$ and is called the permeability of vacuum, and $\epsilon_0 = 10^{-9}/36\pi = 8.854 \times 10^{-12} \text{ F/m}$ and is known as the permittivity of vacuum.

One of the basic laws of electromagnetic phenomena is Faraday's law, which states that a time-varying magnetic field generates an electric field. With reference to Fig. 2.1, let C denote an arbitrary closed curve that forms the boundary of a nonmoving surface S . The time rate of change of total magnetic flux through the surface S is $d(\phi_B \cdot dS)/dt$. According to Faraday's law, this time rate of change of total magnetic flux is equal to the negative value of the total voltage measured around C . The latter quantity is given by $-\oint_C \mathbf{E} \cdot d\mathbf{l}$. Hence the mathematical statement of Faraday's law is

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = - \frac{1}{\mu_0 \epsilon_0} \iint_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \quad (2.3)$$

The line integral of \mathbf{E} around C is a measure of the circulation, or "curling up," of the electric field in space. The time-varying magnetic field may be properly regarded as a vortex source that produces an electric field having nonzero curl, or circulation. Although (2.3) is in a form that is readily interpreted physically, it is not in a form suitable for the analysis of a physical problem. What is required is a differential equation that is equivalent to (2.3). This equation may be obtained by using Stokes' theorem from vector analysis, which states that the line integral of a vector around a closed contour C is equal to the integral of the normal component of the

curl of this vector over any surface having C as its boundary. The curl of a vector is written $\nabla \times S''$ (App. I), and hence (2.1) becomes

$$\oint_{J^*} \nabla \times dS = \int_{J^*} \nabla \times g \cdot dS = - \frac{d}{dt} [a \cdot dS]$$

Since S is completely arbitrary, the latter two integrals are equal only if

$$\nabla \times g = - \frac{d}{dt} \quad (2.4)$$

which is the desired differential equation describing Faraday's law. The curl is a measure of the circulation of a vector field at a point.

Helmholtz's theorem from vector analysis states that a vector field is completely defined only when the curl, or circulation, of the field, and also its divergence, are given at every point in space. Now the divergence (or convergence) of field lines arises only if a proper source (or sink) is available. The electric field, in addition to having a curl produced by the vortex source $-BSS/dt$, has a divergence produced by electric charge. Gauss' law states that the total flux of $9i = e_0 f$ from a volume V is equal to the net charge contained within V . If p represents the charge density in coulombs per cubic meter, Gauss' law may be written as

$$\oint_S e_0 \nabla \cdot dS = \int_V p dV \quad (2.5)$$

This equation may be converted to a differential equation by using the divergence theorem to give

$$\oint_S e_0 g \cdot dS = \int_V \nabla \cdot s_0 \nabla dV = \int_V p dV$$

Since V is arbitrary, it follows that

$$\nabla \cdot g = -as = p \quad (2.6)$$

where $\nabla \cdot g$ is the divergence of g , that is, a measure of the total outward flux of g from a volume element, divided by the volume of the element, as this volume shrinks to zero. Since both the curl and divergence of the electric field are now specified, this field is completely determined in terms of the two sources, HSt/dt and p .

To complete the formulation of electromagnetic phenomena, we must now relate the curl and divergence of the magnetic field to their sources. The vortex source that creates the circulation, or curl, of the magnetic field $\nabla \times H$ is the current. By current is meant the total current density, the conduction current density f measured in amperes per square meter, the displacement current density dV/Ht , and the convection current pv consisting of charge in motion if present. Convection current is not included in this chapter. However, in the chapter dealing with microwave tubes, convection current plays a central role and is discussed in detail there. The displacement current density fLS/dt was first introduced by Maxwell, and leads to

the possibility of wave motion, as will be seen. Mathematically, the circulation of \mathbf{J} around a closed contour C bounding a surface S as in Fig. 2.1 is given by

$$\oint_C \mathbf{J} \cdot d\mathbf{l} = \iint_S \frac{\partial \mathbf{J}}{\partial n} \cdot d\mathbf{S} \quad (2.7)$$

Application of Stokes' law to the left-hand side yields

$$\iint_S V \times \mathbf{J} \cdot d\mathbf{S} = \iint_S \frac{\partial V}{\partial t} \cdot d\mathbf{S} + \iint_V \mathbf{J} \cdot d\mathbf{S}$$

from which it may be concluded that

$$\nabla \times \mathbf{J} = - \frac{\partial V}{\partial t} \quad (2.8)$$

Since magnetic charge, as the dual of electric charge, does not exist in nature, it may be concluded that the divergence of \mathbf{J} is always zero; i.e., the flux lines of \mathbf{J} are always closed since there are no charges for them to terminate on. Thus the net flux of \mathbf{J} through any closed surface S is always zero; i.e., just as much flux enters through the surface as leaves it. Corresponding to (2.5) and (2.6), we thus have

$$\mathbf{i}^* \cdot d\mathbf{S} = 0 \quad (2.9)$$

$$\nabla \cdot \mathbf{i}^* = 0 \quad (2.10)$$

Conduction current, of density f , is the net flow of electric charge. Since charge is conserved, the total rate of flow of charge out of a volume V is equal to the time rate of decrease of total charge within V , as expressed by the equation

$$\nabla \cdot \mathbf{i}^* = - \frac{dQ}{dt} \quad (2.11)$$

This is the continuity equation, and it may be converted to a differential equation by using the divergence theorem in the same manner as was done to derive (2.6) from (2.5). It is readily found that

$$\nabla \cdot \mathbf{i}^* = - \frac{\partial \rho}{\partial t} = 0 \quad (2.12)$$

This equation may also be derived from (2.8) and (2.6). Since the divergence of the curl of any vector is identically zero, the divergence of (2.8) yields

$$\nabla \cdot (\nabla \times \mathbf{J}) + \mathbf{J} \cdot \nabla \times \mathbf{V} = 0$$

Using (2.6) converts this immediately into the continuity equation (2.12). If the displacement current density $d\mathbf{D}/dt$ had not been included as part of the total current density on the right-hand side of (2.8), that equation would

have led to the conclusion that $\nabla \cdot \mathbf{f} = 0$, a result inconsistent with the continuity equation unless the charge density was independent of time.

In summary, the four equations, known as Maxwell's equations, that describe electromagnetic phenomena in vacuum are

$$\nabla \times \mathbf{f} = - \frac{d\mathbf{S}}{at} \quad (2.13a)$$

$$\nabla \times \mathbf{A} = - + / \quad (2.136)$$

$$\nabla \cdot \mathbf{p} = 0 \quad (2.13c)$$

$$\nabla \cdot \mathbf{A} = 0 \quad (2.13d)$$

where in (2.136) the convection current $p\mathbf{v}$ has not been included. The continuity equation may be derived from (2.136) and (2.13c), and hence contains no additional information. Although $-d\mathbf{J}\delta/dt$ may be regarded as a source for $i\mathbf{A}$, and $H\delta i/dt$ as a source of \mathbf{B} , the ultimate sources of an electromagnetic field are the current f and charge p . For time-varying fields, that charge density p which varies with time is not independent of f since it is related to the latter by the continuity equation. As a consequence, it is possible to derive the time-varying electromagnetic field from a knowledge of the current density f alone.

It is not difficult to show in a qualitative way that (2.13a) and (2.136) lead to wave propagation, i.e., to the propagation of an electromagnetic disturbance through space. Consider a loop of wire in which a current varying with time flows as in Fig. 2.2. The conduction current causes a circulation, or curling, of the magnetic field around the current loop as in Fig. 2.2a (for clarity only a few flux lines are shown). The changing magnetic field in turn creates a circulating, or curling, electric field, with field lines that encircle the magnetic field lines as in Fig. 2.2b. This changing electric field creates further curling magnetic field lines as in Fig. 2.2c, and so forth. The net result is the continual growth and spreading of the electromagnetic field into all space surrounding the current loop. The

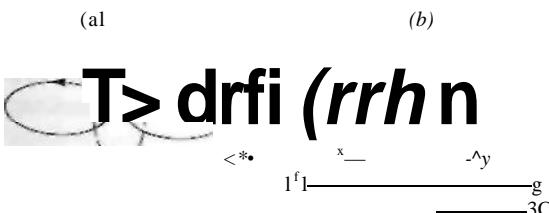


FIGURE 2.2
The growth or generation of an electromagnetic wave from a current loop.

disturbance moves outward with the velocity of light. A little thought will show that the same characteristic mutual effect between two quantities must always exist for wave motion. That is, quantity A must be generated by quantity B , and vice versa. For example, in an acoustical wave the excess pressure creates a motion of the adjacent air mass. The motion of the air mass by virtue of its inertia in turn creates a condensation, or excess pressure, farther along. The repetition of this process generates the acoustical wave.

For the most part, as at lower frequencies, it is sufficient to consider only the steady-state solution for the electromagnetic field as produced by currents having sinusoidal time dependence. The time derivative may then be eliminated by denoting the time dependence of all quantities as e^{Jal} and representing all field vectors as complex-phasor space vectors independent of time. Boldface roman type is used to represent these complex-phasor space vectors. For example, the mathematical representation for the electric field $\hat{E}(x,y,z,t)$ will be $E(x,y,z)e^{J\omega t}$. Each component of E is in general complex, with a real and imaginary part; thus

$$\mathbf{E} = \mathbf{a}_r(\mathbf{f}_{xr} + jE_{xi}) + \mathbf{a}_v(\mathbf{f}_{vr} + jE_{vi}) + \mathbf{a}_s(\mathbf{f}_{zs} + jE_{zi}) \quad (2.14)$$

where the subscript r refers to the real part and the subscript v refers to the imaginary part. Each component is allowed to be complex in order to provide for an arbitrary time phase for each component. This may be seen by recalling the usual method of obtaining g'' from its phasor representation. That is, by definition,

$$Z(x,y,z,t) = \operatorname{Re}[E(x,y,z,t)e^{\lambda t}] \quad (2.15)$$

Thus

$$E_x = \operatorname{Re}[(E_{xr} + jE_{xi})e^{-j\omega t}]$$

$$= V_x r + W_i \cos(i\omega t + \phi)$$

where $\phi = \tan^{-1} H E_{xi}/E_{xr}$. Unless E_x had both an imaginary part jE_{xi} a real part E_{xr} , the arbitrary phase angle ϕ would not be present. As general rule, the time factor $e^{J\omega t}$ will not be written down when the phase representation is used. However, it is important to remember both the that such a time dependence is implied and also the rule (2.15) for obtain the physical field vector from its phasor representation. The real imaginary parts of the space components of a vector should not be confused with the space components; for example, E_{xr} and E_{xi} are not two components of E_x since the component $a_x E_x$ is always directed along the axis in space, with the real and imaginary parts simply accounting for arbitrary time phase or origin.

A further point of interest in connection with the phasor representation is the method used for obtaining the time-average value of a fie quantity.

For example, if

$$\mathbf{r} = a_x E_1 \cos(\omega t + \phi_1) + a_y E_2 \cos(\omega t + \phi_2) + a_z E_3 \cos(\omega t + \phi_3)$$

the time-average value of $\langle g \rangle$ is

$$\begin{aligned} & \frac{1}{T} \int_{-T/2}^{T/2} \mathbf{r} \cdot d\mathbf{r} = \frac{1}{T} \int_{-T/2}^{T/2} [E_1 \cos(\omega t + \phi_1) + E_2 \cos(\omega t + \phi_2) + E_3 \cos(\omega t + \phi_3)] dt \\ &= \pm(E_x^2 + E_y^2 + E_z^2) \quad (2.16) \end{aligned}$$

where T is the period, equal to $2\pi/\omega$. The same result is obtained by simply taking one-half of the scalar, or dot, product of \mathbf{E} with the complex conjugate \mathbf{E}^* ; thus

$$\mathbf{E}^* \cdot \mathbf{E} = |\mathbf{E} \cdot \mathbf{E}^*| = |[(J\mathbf{f} + \mathbf{E}^*) + (4\pi\epsilon_0 + 2\mu_0) + (\mathbf{C} + \mathbf{S}^*)]| \quad (2.17)$$

since $\mathbf{E} \cdot \mathbf{E}^* = (E_x + iE_{-x}, E_y + iE_{-y}, E_z + iE_{-z}) \cdot (E_x^* - iE_{-x}^*, E_y^* - iE_{-y}^*, E_z^* - iE_{-z}^*) = E_x^* E_x + E_y^* E_y + E_z^* E_z + i(E_x^* E_{-y} - E_x E_{-y}^* + E_x^* E_{-z} - E_x E_{-z}^* + E_y^* E_{-x} - E_y E_{-x}^* + E_y^* E_{-z} - E_y E_{-z}^* + E_z^* E_{-x} - E_z E_{-x}^* + E_z^* E_{-y} - E_z E_{-y}^*)$, etc. This is equal to (2.16), since $\mathbf{f}^* = \mathbf{f}$, etc.

By using the phasor representation, the time derivative d/dt may be replaced by the factor $j\omega$ since $dE^*/dt = j\omega E^*$. Hence Maxwell's equations, with steady-state sinusoidal time dependence, become

$$\nabla \times \mathbf{E} = -j\omega \mathbf{B} \quad (2.18a)$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \mathbf{D} \quad (2.18b)$$

$$\nabla \cdot \mathbf{D} = \rho \quad (2.18c)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.18d)$$

CONSTITUTIVE RELATIONS

In material media the auxiliary field vectors \mathbf{D} and \mathbf{B} are defined in terms of the polarization of the material and the fundamental field quantities \mathbf{E} and \mathbf{H} . The relationships of \mathbf{D} to \mathbf{E} and of \mathbf{B} to \mathbf{H} are known as constitutive relations, and must be known before solutions for Maxwell's equations can be found.

Consider first the electric case. If an electric field \mathbf{E} is applied to a material body, this force results in a distortion of the atoms or molecules in such a manner as to create effective electric dipoles with a dipole moment \mathbf{P} per unit volume. The total displacement current is the sum of the vacuum displacement current $\partial\mathbf{D}/\partial t$ and the polarization current $\partial\mathbf{P}/\partial t$. To avoid accounting for the polarization current $\partial\mathbf{P}/\partial t$ explicitly, the

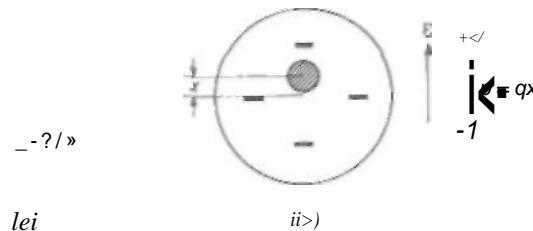


FIGURE 2.3
Model for determining the polarization of an atom.

displacement vector $\mathbf{3}$ is denoted as

$$\mathbf{9I} = e_0 \mathbf{g} + \mathbf{\&} \quad (2.19)$$

whence the total displacement current density can be written as $33i/dt$.

For a great many materials the polarization $\mathbf{\&}$ is in the direction of the *electric field* \mathbf{g} , although rarely will \mathbf{fP} have the same time phase as $\mathbf{\&}$. A simple classical model will serve to illustrate this point. Figure 2.3a shows a model of an atom consisting of a nucleus with charge q surrounded by a spherically symmetrical electron cloud of total charge $-q$. The application of a field \mathbf{I} displaces the electron cloud an effective distance A : as in Fig. 2.36. This displacement is resisted by a restoring force kx proportional to the displacement (Prob. 2.1). In addition, dissipation, or damping, effects are present and result in an additional force, which we shall assume to be proportional to the velocity. If m is the effective mass of the electron cloud, the dynamical equation of motion is obtained by equating the sum of the inertial force md^2x/dt^2 , viscous force $m\mathbf{v} \cdot \mathbf{dx}/dt$, and restoring force kx to the applied force $-q\mathbf{E}$ thus

$$m \frac{d^2x}{dt^2} + m\mathbf{v} \cdot \frac{dx}{dt} + kx = -q\mathbf{E} \quad (2.20)$$

When $\mathbf{I} = E_x \cos \omega t$, the solution for x is of the form $x = -A \cos(\omega t) + \Phi$

If $E_x \cos \omega t$ is represented by the phasor E_x , and x by the phasor X , the solution for X is readily found to be

$$X = -\omega^2 m + j(\omega m + k)$$

and hence

$$x = \operatorname{Re}(Xe^{j\omega t}) = A \cos(\omega t + \Phi)$$

where

$$\frac{(q/m)E_x}{[(y^2 + \langle \mathbf{v}^2 \rangle)^{1/2}]}$$

$$d = \tan \frac{\Phi}{\omega^2 - \omega_0^2}$$

and we have replaced k/m by ω_0^2 .

The dipole moment is p_x , where

$$p_x = -qx = -\frac{q^* E_x}{\omega_0^2 (1 + \omega_0^2 \tau^2)} \cos(\omega t + \phi) \quad (2.21)$$

For N such atoms per unit volume the polarization per unit volume is $\langle P_x \rangle = Np_x$ and the displacement δ_x is given by

$$\delta_x = e_0 E_x \cos \text{tot H} \frac{Nq^2 E_x}{m U a f^2} \cos((\omega t + 45^\circ))$$

This equation may also be put into the following form:

$$\delta_x = E_x \left[e_0 (v^2 \omega_0^2 - \omega^2) + \frac{Nq^2/m^2}{(1 + \omega^2/\omega_0^2)^{1/2}} + \frac{(a \omega v e_0)}{(1 + (\omega^2/\omega_0^2)^{1/2})^2} \right] \cos(\text{tot} - B) \quad (2.22)$$

where $\theta = \tan^{-1} \frac{\omega_0^2}{\omega^2} - \tan^{-1} \frac{Nq^2/m^2}{\omega^2} = \frac{\theta_0}{1 + (\omega^2/\omega_0^2)^{1/2}}$

Two points are of interest in connection with (2.22). One is the linear relationship between δ and θ , and hence between S and ϵ . The second is the phase lag in δ relative to E whenever damping forces are present.

The phase difference between ϵ , W , and δ makes it awkward to handle the relations between these quantities unless phasor representation is used. In phasor representation (2.21) and (2.22) become

$$\frac{\theta}{\omega} = \frac{\omega_0^2 - \omega^2 + j\omega \nu}{\omega^2 + j\omega \nu} \quad (2.23)$$

$$D_x = \frac{e_0 (\omega_0^2 - \omega^2 + j\omega \nu) + Nq^2/m}{\omega^2} \frac{H}{\omega} \angle E_x \quad (2.24)$$

$$\theta = \theta_0 + ja\nu$$

In general, for linear media, we may write

$$P = \epsilon_0^* \omega^* E \quad (2.25)$$

where ϵ_0^* is a complex constant of proportionality called the electric susceptibility. The equation for D becomes

$$D = \epsilon_0 E + P = \epsilon_0 (1 + \epsilon_0^*) E \\ = \epsilon_r E = \epsilon_r \epsilon_0 E = (\epsilon' - j\epsilon'') E \quad (2.26)$$

where $\epsilon = \epsilon_0 (1 + \epsilon'')$ is called the permittivity, and $\epsilon_r = \epsilon/\epsilon_0$, the dielectric constant of the medium. Note that ϵ is complex whenever damping effects are present and that the imaginary part is always negative. A positive imaginary part would imply energy creation instead of energy loss. [The reader may verify from (2.22) that θ is always positive.]

Loss in a dielectric material may also occur because of a finite conductivity a . The two mechanisms are indistinguishable as far as external effects

related to power dissipation are concerned. The curl equation for \mathbf{H} may be written as

$$\nabla \times \mathbf{H} = jw(e' - je'')\mathbf{E} + \langle T\mathbf{E} \rangle$$

where $\mathbf{J} = cr\mathbf{E}$ is the conduction current density in the material. We may also write

$$\nabla \times \mathbf{H} = jio(e' - j[e'' + \langle \sigma \rangle])\mathbf{E} + (we'' + cr)\mathbf{E} \quad (2.27)$$

where by $e'' + cr/io$ may be considered as the effective imaginary part of the permittivity, or $we'' + a$ as the total effective conductivity.

The loss tangent of a dielectric medium is defined by

$$\tan \delta = \frac{we}{\sigma} \quad (2.28)$$

Any measurement of $\tan \delta$, always includes the effects of finite conductivity cr . At microwave frequencies, however, we'' is usually much larger than a because of the large value of w .

Materials for which \mathbf{P} is linearly related to \mathbf{E} and in the same direction as \mathbf{E} are called linear isotropic materials. Nonlinear effects generally occur only for very large applied fields, and as a consequence are rarely encountered in microwave work. However, nonisotropic material is of some importance. If the crystal structure lacks spherical symmetry such as that in a cubic crystal, it may be anticipated that the polarization per unit volume will depend on the direction of the applied field. In Fig. 2.4 a two-dimensional sketch of a crystal lacking cubic symmetry is given. The polarization produced when the field is applied along the x axis may be greater than that when the field is applied along the y or z axis because of the greater ease of polarization along the x axis. In this case we must write

$$\mathbf{D}_T = e_{xx}\mathbf{E}_x \quad D_y = e_{yy}\mathbf{E}_y \quad D_z = e_{zz}\mathbf{E}_z \quad (2.29)$$

where e_{xx} , e_{yy} and e_{zz} are, in general, all different. The dielectric constants $e_{rx} = e_x^* \times A_o > e_{rv} = e_y^* v_A > e_{rz} = e_z^* r^* = e \ll A_o a^T M$ known as the principal dielectric constants, and the material is said to be anisotropic. If the coordinate system used had a different orientation with respect to the crystal structure,

CBT.. — & "

6-4> - & '

FIGURE 2.4

A noncubic crystal exhibiting anisotropic effects.

the relation between D and E would become

$$\begin{aligned} D_x &= e_{xx}E_x + e_{xy}E_y + e_{xz}E_z \\ D_y &= e_{yx}E_x + e_{yy}E_y + e_{yz}E_z \\ D_z &= e_{zx}E_x + e_{zy}E_y + e_{zz}E_z \end{aligned}$$

or in matrix form,

$$\begin{vmatrix} D \\ A \\ D \\ z \end{vmatrix} = \begin{vmatrix} e_{xx} & e_{xy} & e_{xz} \\ e_{yx} & e_{yy} & e_{yz} \\ e_{zx} & e_{zy} & e_{zz} \end{vmatrix} \begin{vmatrix} E_x \\ E_y \\ E_z \end{vmatrix}; \quad (2.30)$$

Only for a particular orientation of the coordinate system does (2.30) reduce to (2.29). This particular orientation defines the principal axis of the medium. For anisotropic media the permittivity is referred to as a tensor permittivity (a tensor of rank 2 may be represented by a matrix). For the most part the materials dealt with in this text are isotropic. Nevertheless, an awareness of the existence of anisotropic media and of the nature of the constitutive relations for such media is important.

For the magnetic case, H is defined by the constitutive relation

$$M_0 H = B - /x_0 M \quad (2.31)$$

where M is the magnetic dipole polarization per unit volume. For most materials (ferromagnetic materials excluded), M is linearly related to B and hence to H . By convention this is expressed by the equation

$$M = *_m H \quad (2.32)$$

where $*_m$ is called the magnetic susceptibility. Substituting (2.32) into (2.31) gives

$$B = M_0(M + H) = M_0(1 + A_m)H = A_t H \quad (2.33)$$

where $\mu_r = f_0(l + x_m)$ is called the permeability.

As in the electric case, damping forces cause μ_r to be a complex parameter with a negative imaginary part; that is, $M = M_r - j\mu_r^2 -$. Also, there are magnetic materials that are anisotropic; in particular, ferrites are anisotropic magnetic materials of great usefulness at microwave frequencies. These exhibit a tensor permeability of the following form:

$$[M] = \begin{matrix} M_r & j\mu_r^2 & 0 & 1 \\ -VM_2 & M_r & 0 & 0 \\ 0 & 0 & M_a & 0 \end{matrix} \quad (2.34)$$

when a static magnetic field is applied along the axis for which the permeability is μ_r . A discussion of ferrites and their uses is presented later; so further comments on their anisotropic properties is deferred until then.

In Sec. 2.1 care was taken to write Maxwell's equations in a form valid not only in vacuum but also in material media. Thus (2.13) and (2.18) are

valid in general, but with the constitutive relations of this section replacing the free-space relations (2.2). Note, however, that it is not possible to write, in general, constitutive relations of the form $\mathbf{D} = \epsilon\mathbf{E}$, $\mathbf{B} = \mu\mathbf{H}$, when \mathbf{D} and \mathbf{B} , and likewise \mathbf{H} and \mathbf{B} , are not in time phase. For arbitrary time dependence we must write instead $\mathbf{D} = \epsilon_0(\mathbf{E} + \mathbf{E}_0)$, $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{H}_0)$ and relate \mathbf{D} and \mathbf{B} to \mathbf{E} and \mathbf{H} through the dynamical equation of motion governing the polarization mechanism. This difficulty may be circumvented by using the phasor representation for which relations such as $\mathbf{D} = \epsilon\mathbf{E}$ are perfectly valid because the complex nature of ϵ accounts for the difference in time phase.¹ It should be pointed out, however, that for many materials used at frequencies up to and including microwaves, the losses are so small that \mathbf{D} and \mathbf{B} , and also \mathbf{H} and \mathbf{B} , are very nearly in time phase. In such cases constitutive relations such as $\mathbf{D} = \epsilon_0\mathbf{E}$, $\mathbf{B} = \mu_0\mathbf{H}$ apply with negligible error. Significant departure in time phase between \mathbf{D} and \mathbf{B} or \mathbf{H} and \mathbf{B} occurs only in the vicinity of a natural resonance frequency of the equation of motion for the polarization.

2.3 STATIC FIELDS

For electric and magnetic fields that are independent of time, the electric and magnetic fields are not coupled, and likewise the current and charge are not coupled. Putting all time derivatives equal to zero in (2.13) yields²

$$\nabla \times \mathbf{E} = 0 \quad (2.35a)$$

$$\nabla \cdot \mathbf{eE} = p \quad (2.35b)$$

$$\nabla \times \mathbf{H} = \mathbf{J} \quad (2.36a)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.36b)$$

$$\nabla \cdot \mathbf{J} = 0 \quad (2.36c)$$

The last equation is the continuity equation for the special case $d\rho/dt = 0$.

The static electric field has zero curl, or circulation, and this means that the line integral of \mathbf{E} around any arbitrary closed contour is zero. This property is just the condition that permits \mathbf{E} to be derived from the gradient of a scalar potential function Q ; that is, since $\nabla \times \nabla Q = 0$ is identically zero, we may put

$$\mathbf{E} = -\nabla Q \quad (2.37)$$

¹The situation here is like that encountered in ac circuit analysis, where in phasor notation the voltage V equals the current I multiplied by the impedance Z ; that is, $V = IZ$. An Ohm's law of this sort cannot be written for the physical voltage and current, for if $V = \text{Re}(V e^{j\omega t}) = V \cos \omega t$, then $J^* = rM/e^{-j\omega t} = [V/(R^2 + X^2)^{1/2}] \cos(\omega t - \phi)$, where $i = \tan^{-1} |X/R|$. Clearly, V cannot be equated to J multiplied by a constant because of the difference in phase. ²For static fields we are using boldface roman type to represent the physically real vector fields.

$$\bullet \gg / \\ x = a$$

$$\Phi = 0 \quad / = 0$$

FIGURE 2.5

A simple potential problem.

Substituting (2.37) into (2.356) and assuming that e is a constant independent of the coordinates give

$$-\nabla \cdot \mathbf{E} = V^2 \Phi = - - \quad (2.38)$$

This equation is known as Poisson's equation. When $p = 0$, Laplace's equation

$$\nabla^2 \Phi = 0 \quad (2.39)$$

is obtained. The basic field problem in electrostatics is to solve Poisson's or Laplace's equation for a potential function Φ that satisfies specified boundary conditions.

As a simple example consider two infinite conducting planes at $x = 0, a$, as in Fig. 2.5. Let charge be distributed with a density $p = p_0 x$ between the two plates.^t It is required to find a Φ which is a solution of Poisson's equation and which equals zero on the plane $x = 0$ and V on the plane $x = a$. The potential will depend on x only; so (2.38) becomes

$$\frac{d^2 \Phi}{dx^2} = -p_0 x \quad |_{x=0} = 0$$

Integrating this equation twice gives $\Phi = -p_0 x^3 / 6e_0 + C_1 x + C_2$. Imposing the boundary conditions at $x = 0, a$ yields $0 = C_2$,

$$Po^a$$

^tThe example is somewhat artificial since the assumed charge distribution is not a stable one; i.e., the electric field it produces would cause the charge distribution to change.

and hence $C_2 = 0$, $C_1 = V/a + p_0 a^2/6\epsilon_0$. The solution for Φ is thus

$$\Phi = -\frac{p_0 x^3}{6\epsilon_0} + \frac{p_0 a^2 x}{6\epsilon_0} + \frac{V}{a}$$

The electric field between the two plates is

$$E = -V\hat{x} = -a, \quad \hat{x} = a \frac{i P_0 *^2}{2\epsilon_0} - \frac{P_0 \hat{x}}{6\epsilon_0} - \frac{V}{a} \hat{J}$$

The solution for the electrostatic field is greatly facilitated by introduction of the scalar potential Φ . For the same reason it is advantageous to introduce a potential function for the solution of magnetostatic problems. Since B always has zero divergence, it may be derived from the curl of a vector potential A ; that is,

$$B = \nabla \times A \quad (2.40)$$

This makes the divergence of B vanish identically because the divergence of the curl of a vector is identically zero. Using (2.40) in (2.36a) and assuming that p_0 is constant yields the equation

A vector identity of use here is $\nabla \times (\nabla \times A) = \nabla \nabla \cdot A - \nabla^2 A$. The divergence of A may be placed equal to zero without affecting the value of B derived from the curl of A , and hence the equation for A is

$$\nabla^2 A = -M \hat{J} \quad (2.41)$$

This equation is a vector Poisson's equation. In rectangular coordinates, (2.41) represents three scalar Poisson's equations, the first being

$$*K = -f_i J_K \quad (2.42)$$

In a curvilinear coordinate system, such as a cylindrical coordinate system, (2.41) cannot be written in such a simple component form. The reason is that, for example, $\nabla^2 a_r A_r$ does not equal $a_r \nabla^2 A_r$ because, even though the unit vector a_r is of constant length, its orientation varies from point to point since it is always directed along the radius vector from the origin to the point under consideration. The evaluation of $\nabla^2 A$ in curvilinear coordinates is made by using the vector identity quoted above to give $\nabla^2 A = \nabla \cdot A - \nabla \times (\nabla \times A)$. These latter operations are readily carried out.

The interest in static field solutions at microwave frequencies arises because the field distribution over a cross-sectional plane of a transmission line is a static field distribution and because static field solutions are good approximate solutions to the actual fields in the vicinity of obstacles that are small compared with the wavelength. The potential theory introduced above may be extended to the time-varying case also, and this is done in a following section.

WAVE EQUATION

For convenience, the two curl equations are repeated here:

$$\nabla \times \mathbf{r} = \frac{\partial}{\partial t} \quad (2.43a)$$

$$\nabla \times \mathbf{f} = \frac{\partial^2 \mathbf{f}}{\partial t^2} \quad (2.43b)$$

where it is assumed for the present that the current density J is zero in the region of interest. These equations, together with the assumed constitutive relations $S = \epsilon_0 \mu_0 = 1/\rho c^2$, may be combined to obtain a separate equation for each field. The curl of (2.43a) is

$$\nabla \times \nabla \times \mathbf{f} = \frac{\partial^2 \mathbf{f}}{\partial t^2} = -\rho \cdot \frac{\partial^2 \mathbf{f}}{\partial t^2}$$

Using (2.43b) and expanding $\nabla \times \nabla \times \mathbf{f}$ now yields

$$\nabla \nabla \cdot \mathbf{f} - \nabla^2 \mathbf{f} = -\rho \frac{\partial^2 \mathbf{f}}{\partial t^2},$$

Since ρ is assumed zero and ϵ_0 is taken as a constant, $\nabla \cdot \mathbf{f} = 0$, and we obtain

$$\nabla^2 \mathbf{f} - \rho c^2 \mathbf{f} = 0 \quad (2.44)$$

which is a three-dimensional wave equation. The velocity of propagation v is equal to $(\rho \epsilon_0)^{-1/2}$. In free space v is equal to the velocity of light c . To illustrate the nature of the solutions of (2.44), consider a case where \mathbf{f} has only an x component and depends only on the z coordinate. In this instance

$$\frac{\partial^2 f_x}{\partial z^2} - \frac{\partial^2 f_x}{\partial t^2} = 0$$

Any function of the form $f(z - vt)$ is a solution of this equation since

$$\frac{\partial^2 f_x}{\partial z^2} - \frac{\partial^2 f_x}{\partial t^2} = 0$$

and hence

$$\frac{\partial^2 f_x}{\partial z^2} - \frac{\partial^2 f_x}{\partial t^2} = 0$$

This solution is illustrated in Fig. 2.6 and clearly represents a disturbance propagating in the positive z direction with velocity v . An equally valid solution is $f(z + vt)$ and represents a disturbance propagating in the negative z direction.

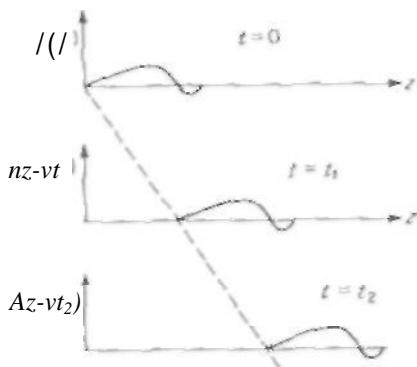


FIGURE 2.6
Propagation of a disturbance $f(z - vt)$.

By eliminating the electric field, it is readily found that the magnetic field "also satisfies the wave equation (2.44). In practice, however, we solve the wave equation for either \mathbf{E} or \mathbf{H} and then derive the other field by using the appropriate curl equation. When constitutive relations such as $\mathbf{G} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ cannot be written, the polarization vectors \mathbf{D} and \mathbf{J} must be exhibited explicitly in Maxwell's equations. Wave equations for \mathbf{E} and \mathbf{H} may still be derived, but \mathbf{G} and \mathbf{B} will now enter as equivalent sources for the field (which they actually are). The derivation is left as a problem at the end of this chapter.

For harmonic time dependence, the equation obtained in place of (2.44) is

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0 \quad (2.45)$$

where $k^2 = \omega^2/c^2$. This equation is referred to as the Helmholtz equation, or reduced wave equation. The constant k is called the wave number and may be expressed in the form

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda} = \frac{2\pi}{\lambda} \quad (2.46)$$

where the wavelength λ is equal to c/ω . In free space the wave number will be written as k_0 , and is equal to $2\pi/\lambda_0 = 2\pi/c_0$. The magnetic field \mathbf{H} , as may be surmised, satisfies the same reduced wave equation.

In a medium with finite conductivity σ , a conduction current $\mathbf{J}_c = \sigma\mathbf{E}$ will exist, and this results in energy loss because of Joule heating. The wave equation in media of this type has a damping term proportional to σ and the first time derivative of the field. In metals, excluding ferromagnetic materials, the permittivity and permeability are essentially equal to their free-space values, at least for frequencies up to and including the microwave range. Thus Maxwell's curl equations become

$$\nabla \times \mathbf{H} = -\mu_0 \frac{\partial \mathbf{E}}{\partial t} \quad \nabla \times \mathbf{E} = \epsilon_0 \sigma \mathbf{H} + \mathbf{J}_c$$

Elimination of X'' in the same manner as before leads to the following wave equation for &:

$$\nabla^2 i? - n_0 c r - n_0 e_0 \wedge - = 0 \quad (2.47)$$

The magnetic field $\#^*$ also satisfies this equation. For the time-harmonic case damping effects enter in through the complex nature of e and $/x$, and hence the wave number k . It should be recalled here that, as shown by (2.27), a finite conductivity $c r$ is equivalent to an imaginary term in the permittivity e . In the present case the equivalent permittivity is $e = e_0 - j c r / u >$ and the Helmholtz equation is

$$\nabla^2 E + w^2 j i_0 e j 1 - j - I E = 0 \quad (2.48)$$

$$I - w f_0 ;$$

In metals the conduction current $o \cdot E$ is generally very much larger than the displacement current $w e_0 E$, so that the latter may be neglected. For example, a is equal to $5.8 \times 10^7 \text{ S/m}$ for copper, and at a frequency of 10^{10} Hz , $a/e_0 = 0.55$, which is much smaller than tr . Only for frequencies in the optical range will the two become comparable. Thus (2.47) may be simplified to

$$\nabla^2 g - - M_0 \wedge r - = 0 \quad (2.49)$$

and (2.48) reduces to

$$\nabla^2 E - > M_0 \circ - E = 0 \quad (2.50)$$

Equation (2.49) is a diffusion equation similar to that which governs the flow of heat in a thermal conductor.

ENERGY AND POWER

When currents exist in conductors as a result of the application of a suitable potential source, energy is expended by the source in maintaining the currents. The energy supplied by the source is stored in the electric and magnetic fields set up by the currents or propagated (radiated) away in the form of an electromagnetic wave. Under steady-state sinusoidal time-varying conditions, the time-average energy stored in the electric field is

$$W = R e - / E - D * d V = - / c' E - E * r f V \quad (2.51a)$$

$$\frac{1}{4} \bullet v \qquad \frac{1}{4} - V$$

If e is a constant and real, (2.51a) becomes

$$W_e = \frac{1}{4} \bullet v (E - E * d V) \quad (2.516)$$

The time-average energy stored in the magnetic field is given by

$$\frac{1}{4} \int_V \text{Re} \left[\mathbf{H}^* \cdot \mathbf{B} \right] dV = \frac{1}{4} \int_V \text{ii}[\mathbf{H}] \cdot \mathbf{H}^* dV \quad (2.52a)$$

which, for ii real and constant, becomes

$$W_m = \int_V \mathbf{H} \cdot \mathbf{H}^* dV \quad (2.52b)$$

These expressions for W_e and W_m are valid only for nondispersive media, i.e., media for which ϵ and τ_i can be considered independent of ω in the vicinity of the angular frequency ω with which the fields vary. In general, when the losses are small, so that $\epsilon'' \ll \epsilon'$ and $\tau'' \ll \tau_i$, we have

$$W_e = \frac{1}{4} \int_V \mathbf{E} \cdot \mathbf{E}^* dV \quad (2.53a)$$

$$W_m = \int_V \mathbf{H} \cdot \mathbf{H}^* dV \quad (2.53b)$$

for the time-average stored electric and magnetic energy.

The above equations for the time-average energy in a dispersive medium may be established by considering a classical model of the polarization mechanism similar to that discussed in Sec. 2.2. In a unit volume let the effective oscillating charge of the dipole distribution be $-a$ with an effective mass m . Let the damping force be equal to mu times the velocity of the charge. This damping force takes account of collision effects and loss of energy by radiation from the oscillating charge. The equation of motion for the polarization charge displacement u is

$$\frac{d^2u}{dt^2} = -\frac{du}{dt}$$

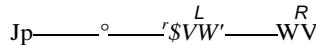
where u is parallel to the direction of the field \mathbf{H} . In this equation k is elastic constant giving rise to the restoring force. This constant arises from the Coulomb forces acting on the displaced charge, and hence is of electrical origin. The dipole polarization $3''$ is $-qu$, and the polarization current $f = d.9/dt$. Introducing the polarization current into the equation of motion gives

$$m \frac{df_B}{dt} = mu = -\frac{k_{rt}}{q} u$$

This equation is formally the same as that which describes the current in a series LCR circuit with an applied voltage V equal to W and with

$$\frac{m}{q^2}, \quad \frac{mv}{Q}, \quad \frac{q^2}{k}$$

An equivalent circuit describing the polarization is illustrated in Fig. 2.7.



C FIGURE 2.7
Equivalent circuit for polarization current.

a time dependence $e^{-\omega t}$ is assumed and phasor notation is used,

$$\begin{aligned} & R - JX \\ & \quad J P - EY - EWTX^{\wedge} \end{aligned}$$

where Y is the input admittance and $X = c/L - 1/\omega C$. Since $P = e_0 x_e E$ and $J_p = j(\omega P)$, we see that

$$ft>e_0 A_{-e} = <\!> e_0 \{x' e^{-JX'e}\} = \sim JY = \frac{-X-jR}{R^2 + X^2}$$

and hence

$$\sim We = {}_D \overset{-X}{R^2} \overset{2j}{T_X^{\wedge}} \quad (2.54a)$$

The time-average power loss associated with the polarization is the same as the power loss in R in the equivalent circuit. This is given by

$$Pi = \frac{1}{2} \overset{\wedge}{E_p} \hat{R} = \frac{R}{R^2 + X^2} \sim EE^* o e_0 x_e'' \quad (2.55)$$

per unit volume. This equation shows that $(Df_0^{\wedge} = we)$ is an equivalent conductance. The time-average energy stored in the system is of two forms. First there is the kinetic energy of motion, that is, $\langle m(du/dt)^2 \rangle$ averaged over a cycle, and this is equal to the magnetic energy stored in the inductor in the equivalent circuit. This time-average kinetic energy per unit volume is given by

$$U_m = \frac{1}{4} \overset{\wedge}{U_p} j; = \frac{I}{4} \overset{\wedge}{R^2} \overset{2}{X'} \quad (2.56a)$$

The second form of stored energy is the potential energy associated with the charge displacement. The time-average value of this energy is equal to the time-average electric energy stored in the capacitor C in the equivalent circuit, and is given by

The total time-average energy stored per unit volume is $U = U_m + U_e$. Note

that U is not given by $\frac{1}{2}E\epsilon_0x_e^2$. The latter expression gives

$$= \frac{1}{4}EK \frac{1/w^2C - L}{fl^2 + X^2} = U - u$$

or the difference between the potential and kinetic energy stored.

To obtain an expression for the total stored energy, note that

$$\frac{d}{do} \frac{L + 1/w^2C}{R^2 + X^2} = \frac{2X}{R^2 + X^2}$$

For a low-loss system, $R^2 \ll X^2$, and we then have $1 - 2X^2/(R^2 + X^2) \approx 1$; so

$$\frac{d}{dx} \frac{I - X}{WTx^2} = \frac{L + 1/w^2C}{R^2 + X^2}$$

Multiplying this expression by $E\epsilon_0$ now gives the total time-average energy stored, as comparison with (2.56a) and (2.56b) shows. Thus the final expression for the time-average electric energy stored in a volume V is given by the volume integral of $U = U_e + U_u$, plus the free-space energy density $\epsilon_0(E^2 - E^{*2})/4$ and is

$$W = \frac{1}{4} \int_V (E \cdot E^*) dV$$

$$= \frac{1}{4} \int_V \frac{d(oe')}{E \cdot E^*} dV$$

since $e' = \epsilon_0(l + x'^2)$. This equation is the result given earlier by (2.53a).

A similar type of model may be used to establish (2.53c) for average stored magnetic energy. It should be pointed out that under varying conditions the average stored energy associated with either electric or magnetic polarization includes a kinetic-energy term. This term is negligible at low frequencies and also when e' and j' are essentially independent off the range of a of interest. When this energy is not negligible, the modified expressions for stored energy must be used.

Although (2.53) is more general than (2.51) and (2.52), we shall, in the majority of instances, use the latter equations for the stored energy, thereby tacitly assume that we are dealing with material that is nondissipative or very nearly so.

The time-average power transmitted across a closed surface S is given by the integral of the real part of one-half of the normal component of the

complex Poynting vector $\mathbf{E} \times \mathbf{H}^*$; that is,

$$P = \frac{1}{2} \epsilon_0 \mathbf{E} \cdot \mathbf{H}^* \cdot dS \quad (2.57)$$

The above results are obtained from the interpretation of the complex Poynting vector theorem, which may be derived from Maxwell's equations as follows: If the divergence of $\mathbf{E} \times \mathbf{H}^*$, that is, $\nabla \cdot (\mathbf{E} \times \mathbf{H}^*)$, is expanded, we obtain

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}^*) = (\nabla \times \mathbf{E}) \cdot \mathbf{H}^* - (\nabla \times \mathbf{H}^*) \cdot \mathbf{E}$$

From Maxwell's equations $\nabla \times \mathbf{E} = -jwB$ and $\nabla \times \mathbf{H}^* = -jwD^* + \mathbf{J}^*$, and hence

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}^*) = -jwB \cdot \mathbf{H}^* + jwD^* \cdot \mathbf{E} - \mathbf{E} \cdot \mathbf{J}^*$$

The integration of this equation throughout a volume V bounded by a closed surface S gives the complex Poynting vector theorem; i.e.,

$$\begin{aligned} \text{If } & \int_V (\nabla \cdot (\mathbf{E} \times \mathbf{H}^*)) dV = \int_V (\mathbf{E} \times \mathbf{H}^*) \cdot d\mathbf{S} \\ - \int_{JM} & = \frac{1}{2} \int_V (-jwB \cdot \mathbf{H}^* + jwD^* \cdot \mathbf{E} - \mathbf{E} \cdot \mathbf{J}^*) dV \\ & = -jw \int_V (B \cdot \mathbf{H}^* - D^* \cdot \mathbf{E} - \mathbf{E} \cdot \mathbf{J}^*) dV - \frac{1}{2} \int_S \mathbf{E} \cdot \mathbf{J}^* dS \end{aligned} \quad (2.58a)$$

where the divergence theorem has been used on the left-hand side integral. The above result may be rewritten as

$$1, \quad , jB \cdot H^* - E \cdot D^* \mid$$

$$+ \frac{1}{2} \int_V (-jwB \cdot \mathbf{H}^* + jwD^* \cdot \mathbf{E} - \mathbf{E} \cdot \mathbf{J}^*) dV \quad (2.586)$$

where $-dS$ is a vector element of area directed into the volume V . If the medium in V is characterized by parameters $\epsilon = \epsilon' - j\epsilon''$, $\mu = \mu' - j\mu''$, conductivity σ , the real and imaginary parts of (2.58) may be equated to give

$$\begin{aligned} \text{R} \epsilon \frac{1}{2} \int_V \mathbf{E} \times \mathbf{H}^* \cdot (-dS) &= \frac{1}{2} \int_V f(\mu'' \mathbf{H} \cdot \mathbf{H}^* + \epsilon'' \mathbf{E} \cdot \mathbf{E}^*) dV \\ &+ \frac{1}{2} \int_M a \mathbf{E} \cdot \mathbf{E}^* dV \end{aligned} \quad (2.59a)$$

$$\text{Im} - f h \mathbf{E} \times \mathbf{H}^* \cdot (-dS) = \frac{1}{2} \int_V \frac{\mathbf{H} \cdot \mathbf{H}^*}{\mu'} - \frac{\mathbf{E} \cdot \mathbf{E}^*}{\epsilon'} dV \quad (2.596)$$

Equation (2.59a) is interpreted to state that the real electromagnetic power transmitted through the closed surface S into V is equal to the power loss produced by conduction current $\sigma \mathbf{E}$, resulting in Joule heating plus the

power loss resulting from polarization damping forces. Note that we'' could be interpreted as an equivalent conductance, as pointed out in Sec. 2.2. The equation also shows that $/i''$ and e'' must be positive in order to represent energy loss, and hence the imaginary parts of e and $/J$ must be negative. Equation (2.596) states that the imaginary part of the complex rate of energy flow into V is equal to $2w$ times the net reactive energy W_m — stored in the magnetic and electric fields in V . The complex Poynting vector theorem is essentially an energy-balance equation.

A result analogous to the above may be derived for a conventional network, and serves to demonstrate the validity of the interpretation of (2.58). Consider a simple series RLC circuit as in Fig. 2.7. If the current in the circuit is I and the applied voltage is V , the complex input power given by

$$\frac{1}{2}VI^* = \frac{1}{2}ZII^* = \frac{1}{2}\text{ff}^* R + jaL -$$

The time-average power loss in R , magnetic energy stored in the field around L , and electric energy stored in the field associated with C are given, respectively, by

$$P_r = -RII^* \quad W_m = -jUl^* \quad W_e = \frac{1}{4} \frac{II^*}{w^2 C}$$

since the voltage across C is I/coC . Hence

$$\hat{V}I^* = \pm ZII^* - P_r + 2jco(W_m - W_e)$$

which has the same interpretation as (2.58). This equation may also be solved for the impedance Z to give

$$Z = \frac{V}{I^*} \quad (2.60)$$

and provides a general definition of the impedance of a network in terms of the associated power loss and stored reactive energy. The factor $|I^*$ in the denominator serves as a normalization factor, and is required in order to make Z independent of the magnitude of the current at the input to the network.

In the case of a general time-varying field, an expansion of $\nabla \cdot W X \#$ and substitution from Maxwell's equations (2.13) lead to the following Poynting vector theorem for general time-varying fields:

$$, / \quad ax \quad M$$

$$+ e_0 \hat{\epsilon}' \frac{d}{dt} + \& \frac{d}{dt} + \& S dV$$

Since $\hat{J} \cdot d\mathbf{C}/M = \nabla \cdot f(\mathbf{J}) \cdot \hat{\mathbf{A}}/dt$, etc., and the electric and magnetic polarization currents are $f_p = d\mathbf{E}/dt$, $f_m = i\chi_0(d\mathbf{H}/dt)$, we have

$$+ (\mathbf{y} \cdot \nabla f + \mathbf{J} \cdot \mathbf{r}_p) + \sigma - \epsilon f_m \cdot \mathbf{dV} \quad (2.61)$$

where $-\mathbf{rfS}$ is an element of surface area directed into V . This equation states that the rate of energy flow into V is equal to the time rate of change of the free-space field energy stored in V plus the rate of energy dissipation in Joule heating arising from the conduction current f and, in addition, the instantaneous rate of energy supplied in maintaining the polarization currents. If J and \mathbf{B} , and also \mathbf{E} and \mathbf{H} are in phase, there is no energy loss associated with the polarization currents. If these quantities are not in phase, some energy dissipation takes place, leading to increased heating of the material.

If the susceptibilities χ_e and χ_m can be considered as constants, so that $d\mathbf{E}/dt = \epsilon_0 \chi_e(d\mathbf{H}/dt)$ and $d\mathbf{H}/dt = \mu_0 \chi_m(\mathbf{B} - \mathbf{B}_0)$, then (2.61) becomes

$$\hat{\mathbf{r}} \cdot \hat{\mathbf{x}} \cdot (-\mathbf{rfS}) = -/\mathbf{v} (\hat{\mathbf{A}} - \hat{\mathbf{A}}^2) \hat{\mathbf{A}} + /_{\mathbf{v}} \mathbf{f} \cdot \hat{\mathbf{A}} \mathbf{v}, \quad (2.62)$$

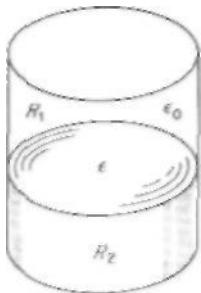
which is the usual form of the Poynting vector theorem. The first term on the right is now interpreted as the instantaneous rate of change of the total electric and magnetic energy stored in the volume V .

The susceptibilities can usually be considered as true constants whenever the inertia and damping forces are small compared with the elastic restoring force in the dynamical equation describing the polarization. For example, with reference to (2.54a), this is the case when k is much greater than $t\omega_0^2$ or $\omega^2 m$, that is, when $1/\omega C$ is large compared with wL and R , so that

$$\langle \mathbf{r} \cdot \mathbf{t} \rangle = \mathbf{c} = \mathbf{q}_-$$

BOUNDARY CONDITIONS

In order to find the proper and unique solutions to Maxwell's equations for situations of practical interest (these always involve material bodies with boundaries), a knowledge of the behavior of the electromagnetic field at the boundary separating material bodies with different electrical properties is required. From a mathematical point of view, the solution of a partial differential equation, such as a wave equation, in a region V is not unique unless boundary conditions are specified, i.e., the behavior of the field on the boundary of V . Boundary conditions play the same role in the solution of

**FIGURE 2.8**

A cylindrical cavity partially filled with a dielectric medium.

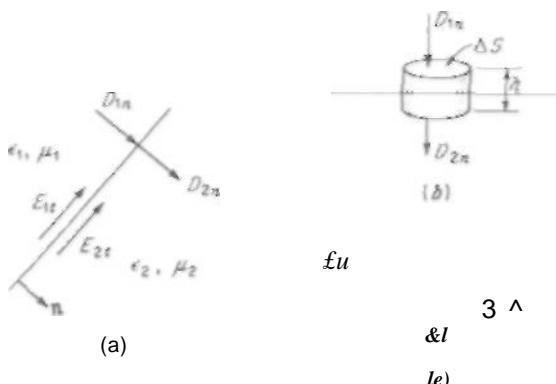
partial differential equations that initial conditions play in the solution of the differential equations that govern the behavior of electric circuits.

As an example, consider the problem of finding a solution to Maxwell's equations inside a cylindrical cavity partially filled with a dielectric medium of permittivity ϵ , as in Fig. 2.8. In practice, the solution is obtained by finding general solutions valid in the two regions labeled R_1 and R_2 . These general solutions must satisfy prescribed conditions on the metallic boundaries and in addition contain arbitrary amplitude constants that can be determined only from a knowledge of the boundary conditions to be applied at the air-dielectric boundary separating regions R_1 and R_2 .

The integral form of Maxwell's equations provides the most convenient formulation in order to deduce the required boundary conditions. Consider two media with parameters ϵ_1, μ_1 and ϵ_2, μ_2 in Fig. 2.9a. If there is no surface charge on the boundary, which is the usual case for nonconducting media, the integral of the displacement flux over the surface of the small "coin-shaped" volume centered on the boundary as in Fig. 2.9b gives, in the limit as h tends to zero,

$$\lim_{fc \rightarrow 0} \oint D \cdot dS = D_{2n} A S - D_{1n} A S = 0$$

or $D_{2n} = D_{1n} = n \quad D_2 = n \quad D_1$ (2.63)

**FIGURE 2.9**
Boundary between two different media.

where n denotes the normal component. The limit $h \rightarrow 0$ is taken so that the flux through the sides of the coin-shaped region vanishes. Equation (2.63) simply states that the displacement flux lines are continuous in the direction normal to the boundary. A similar result clearly must hold also for the magnetic flux lines since $\nabla \cdot \mathbf{B} = 0$, and hence, by analogy,

$$\mathbf{n} \cdot \mathbf{B}_2 = \mathbf{n} \cdot \mathbf{B}_1 \quad (2.64)$$

To obtain boundary conditions on the tangential components of the electric field \mathbf{E} and magnetic field \mathbf{H} , the circulation integrals for \mathbf{E} and \mathbf{H} are used. If for the contour C in Fig. 2.9c, the width h is made to approach zero, the magnetic flux flowing through this contour vanishes and

$$\begin{aligned} \lim_{k \rightarrow 0 \sim C} \oint_C (\mathbf{E} \cdot d\mathbf{l}) &= \lim_{h \rightarrow 0} -ja \int_S \mathbf{B} \cdot r f S = 0 \\ &= E_{2l} M - E_{lt} M \end{aligned}$$

or

$$E_u = E_{2l} \quad (2.65)$$

For the same contour C the total displacement current directed through the contour vanishes as $h \rightarrow 0$, so that

$$\begin{aligned} \lim_{h \rightarrow 0 \sim C} \oint_C (\mathbf{H} \cdot d\mathbf{l}) &= \lim_{h \rightarrow 0} [jut \int_S D \cdot r f S] = 0 \\ &= (H_{2l} - H_u) M \end{aligned}$$

or

$$H_{2l} = H_u \quad (2.66)$$

where t denotes the components tangential to the boundary surface. These latter relations state that the components of \mathbf{E} and \mathbf{H} tangent to the boundary are continuous across the boundary; i.e., the tangential components on adjacent sides of the boundary are equal at the boundary surface.

For the boundary conditions at the surface separating a good conductor (any metal) and free space or air, some simplification is possible. As shown in a later section, the electromagnetic field can penetrate into a conductor only a minute distance at microwave frequencies. The field amplitude decays exponentially from its surface value according to $e^{-\mu u/l}$, where u is the normal distance into the conductor measured from the surface, and δ is called the skin depth. The skin depth is given by

$$\mathbf{S} = \frac{1}{W/lf} \quad (2.67)$$

For copper ($\sigma = 5.8 \times 10^7 \text{ S/m}$) at a frequency of 10^{10} Hz , the skin depth is $6.6 \times 10^{-5} \text{ cm}$, truly a very small distance. Likewise, the current $J = trE$ is concentrated near the surface. As the conductivity is made to approach infinity, δ approaches zero and the current is squeezed into a narrower and narrower region and in the limit $a \rightarrow 0$ becomes a true surface current. Since the skin depth is so small at microwave frequencies for metals, the approximation of infinite conductivity may be made with negligible error (an

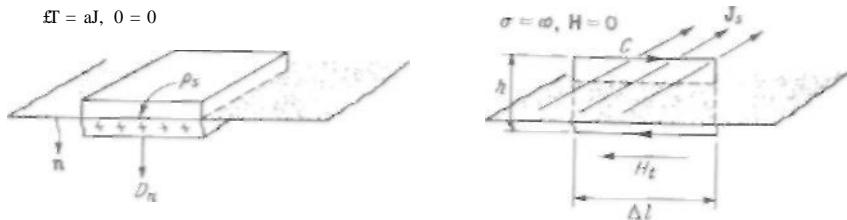


FIGURE 2.10
Boundary- of a perfect conductor.

exception is when attenuation is to be calculated, since then infinite conductivity implies no loss). For infinite conductivity the field in the conductor must be zero. Since the flux lines of B are continuous and likewise since the tangential component of E is continuous across the boundary, it is necessary that

$$\mathbf{n} \cdot \mathbf{B} = 0 \quad (2.68a)$$

$$\mathbf{E}_t = \mathbf{n} \times \mathbf{E} = 0 \quad (2.68b)$$

at the surface of a perfect conductor. This same argument cannot be applied to the normal component of D and the tangential component of H because, as noted above, a surface current J_s will exist on the surface in the limit $\sigma \rightarrow \infty$. Applying Maxwell's equation

$$\oint \mathbf{H} \cdot d\mathbf{l} = \mu_0 j_V \mathbf{V} \cdot dS + \oint \mathbf{j} \cdot dS$$

to the contour C illustrated in Fig. 2.10 gives

$$\begin{aligned} \lim_{h \rightarrow 0} \oint \mathbf{H} \cdot d\mathbf{l} &= H \cdot M = \lim_{h \rightarrow 0} (\mu_0 j_V \mathbf{V} \cdot dS + \lim_{h \rightarrow 0} \oint \mathbf{j} \cdot dS) \\ &= \lim_{A \rightarrow 0} hJM - J_C M \end{aligned}$$

or in vector form,

$$\mathbf{n} \times \mathbf{H} = \mathbf{J}, \quad (2.68c)$$

Note that the field in the conductor goes to zero, that the total displacement current through C vanishes as $h \rightarrow 0$, but that hj tends to the limiting value J_s as the conductivity is made infinite and h is made to approach zero. Associated with the surface current is a charge of density p_s on which the normal displacement flux lines terminate. Hence, at the surface of a perfect conductor,

$$\mathbf{n} \cdot \mathbf{D} = D_n = p_s \quad (2.68d)$$

When it is desired to take into account the large but finite conductivity (as would be the case in attenuation calculations), an impedance boundary

condition may be used with little error. The metallic surface exhibits a surface impedance Z_m , with equal resistive and inductive parts, given by

$$Z_m = \frac{1 + j}{\sigma d_s} \quad (2.69)$$

At the surface a surface current exists, and the relation between this and the electric field tangent to the surface is

$$\mathbf{E}_s = Z_m \mathbf{J}_s \quad (2.70)$$

Note that the tangential electric field cannot be zero for finite conductivity, although it may be very small. Now $\mathbf{n} \times \mathbf{H} = \mathbf{J}_s$, so that

$$\mathbf{E}_s = Z_m \mathbf{J}_s = Z_m \mathbf{n} \times \mathbf{H} \quad (2.71)$$

From (2.69) it is seen that the resistive part of the surface impedance is equal to the dc resistance per square of a unit square of metal of thickness d_s . In a later section the above results are verified; so further comments are reserved until then.

In practice, it suffices to make the tangential components of the fields satisfy the proper boundary conditions since, when they do, the normal components of the fields automatically satisfy their appropriate boundary conditions. The reason is that when the fields are a solution of Maxwell's equations, not all the components of the field are independent. For example, when the tangential part of the electric field is continuous across a boundary, the derivatives of the tangential component of electric field with respect to coordinates on the boundary surface are also continuous. Thus the curl of the electric field normal to the surface is continuous, and this implies continuity of the normal component of \mathbf{B} . More specifically, if the xy plane is the boundary surface and E_x, E_y are continuous, then $i>E_x/dx, 9E_x/dy, 9E_y/dx$, and dE_y/dy are also continuous. Hence $-j\omega B_z = SE_z/H_x - dE_x/dy$ is continuous. For the same reasons continuity of the tangential components of \mathbf{H} ensures the continuity of the normal component of \mathbf{D} across a boundary.

In addition to the boundary conditions given above, a boundary condition must be imposed on the field solutions at the edge of a conducting body such as a wedge. The edge condition requires that the energy stored in the field in the vicinity of an edge of a conducting body be finite. This limits the maximum rate at which the field intensities can increase as the edge is approached.^t A detailed analysis shows that at the edge of a two-dimensional perfectly conducting wedge with an internal angle ϕ , the field components normal to the edge must not increase any faster than r^α , where r is

^tJ. Meixner, The Behavior of Electromagnetic Fields at Edges, N.Y. Univ. Inst. Math. Set. Res. Rept., vol EM-72. December, 1954. The theory is also discussed in R. E. Collin, "Field Theory of Guided Waves," chap. 1, IEEE Press. Piscataway, N.J.. 1991. revised edition.

the perpendicular radial distance away from the edge and

$$\frac{n\pi}{2\pi} - \langle f \rangle$$

where the integer n must be chosen so that a is greater than or equal to — A at least.

When solving for fields in an infinite region of space, the behavior of the field at infinity must also be specified. This boundary condition is called a radiation condition, and requires that the field at infinity be a wave propagating a finite amount of energy outward, or else that the field vanish so fast that the energy stored in the field and the energy flow at infinity are zero.

2.7 PLANE WAVES

In this section and the two following ones we shall introduce wave solutions by considering plane waves propagating in free space and reflection of a plane wave from a boundary separating free space and a dielectric, or conducting, medium. The latter problem will serve to derive the boundary conditions given by (2.68) to (2.71) in the preceding section.

Plane Waves in Free Space

The electric field is a solution of the Helmholtz equation

$$\nabla^2 E = -k^2 E$$

This vector equation holds for each component of E , so that

$$\nabla^2 E_i + k^2 E_i = 0 \quad i = x, y, z \quad (2)$$

The standard procedure for solving a partial differential equation is the method of separation of variables. However, this method does not work for all types of partial differential equations in all various coordinate systems, and when it does not work, a solution is very difficult, if not impossible, to obtain. For the Helmholtz equation the method of separation of variables does work in such common coordinate systems as rectangular, cylindrical, and spherical. Hence this method suffices for the class of problems discussed in this text. The basic procedure is to assume for the solution a product of functions each of which is a function of one coordinate variable only. Substitution of this solution into the partial differential equation then separates the partial differential equation into three ordinary differential¹ equations which may be solved by standard means.

In the present case let $E_x = f(x)g(y)h(z)$. Substituting this expression into (2.72) gives

$$ghf'' + fhg'' + fgh'' + klfgh = Q$$

where the double prime denotes the second derivative. Dividing this equation by fgh gives

$$\frac{f''}{f} + \frac{g''}{g} + \frac{h''}{h} = 0 \quad (2.73)$$

Each of the first three terms in (2.73), such as f''/f , is a function of a single independent variable only, and hence the sum of these terms can equal a constant $-kl$ only if each term is constant. Thus (2.73) separates into three equations:

$$\begin{aligned} \frac{f''}{f} &= k_x^2, \\ \frac{g''}{g} &= k_y^2, \\ \frac{h''}{h} &= k_z^2, \end{aligned} \quad (2.74)$$

where k_x, k_y, k_z are called separation constants. The only restriction so far on k_x, k_y, k_z is that their sum must equal ω , that is,

$$k_x^2 + k_y^2 + k_z^2 = \omega^2 \quad (2.75)$$

so that (2.73) will be satisfied.

Equations (2.74) are simple-harmonic differential equations with exponential solutions of the form $e^{\pm jk_x x}$, $e^{\pm jk_y y}$, $e^{\pm jk_z z}$. As one suitable solution for E_x we may therefore choose

$$E_x = A e^{j(k_x x - k_y y - k_z z)} \quad (2.76)$$

where A is an amplitude factor. This solution is interpreted as the x component of a wave propagating in the direction specified by the propagation vector

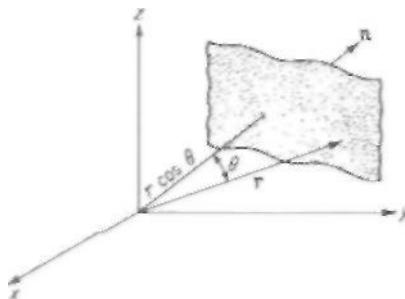
$$\mathbf{k} = \mathbf{aA} + \mathbf{aA}^\perp + \mathbf{aA}^\parallel \quad (\text{Eqn 2.76})$$

because the scalar product of \mathbf{k} with the position vector

$$\mathbf{r} = a_x x \mathbf{i} + a_y y \mathbf{j} + a_z z \mathbf{k}$$

equals $k_x x + k_y y + k_z z$ and is k_0 times the perpendicular distance from the origin to a plane normal to the vector \mathbf{k} , as illustrated in Fig. 2.11. The \mathbf{k} vector may also be written as $\mathbf{k} = n k_0$, where n is a unit vector in the direction of \mathbf{k} and k_0 is the magnitude of \mathbf{k} by virtue of (2.75).

Although (2.76) gives a possible solution for E_x , this is not the complete solution for the electric field. Similar solutions for E_y and E_z may be found. The three components of \mathbf{E} are not independent since the divergence relation $\nabla \cdot \mathbf{E} = 0$ must hold in free space. This constraint means

**FIGURE 2.11**Illustration of plane normal to vector \mathbf{k} .

that only two components of \mathbf{E} can have arbitrary amplitudes. However, for $\nabla \cdot \mathbf{E} = 0$ to vanish everywhere, all components of \mathbf{E} must have the same spatial dependence, and hence appropriate solutions for E_y and E_z are

$$E_y = Be^{-ikr} \quad E_z = Ce^{-ikr}$$

with B and C amplitude coefficients. Let \mathbf{E}_0 be the vector $a_x A + a_y B + a_z C$; then the total solution for \mathbf{E} may be written in vector form as

$$\mathbf{E} = \mathbf{E}_0 e^{-ikr} \quad (2.78)$$

The divergence condition gives

$$\nabla \cdot \mathbf{E} = \nabla \cdot \mathbf{E}_0 e^{-ikr} = \mathbf{E}_0 \cdot \nabla e^{-ikr} = -k \cdot \mathbf{E}_0 e^{-ikr} = 0$$

or $k \cdot \mathbf{E}_0 = 0 \quad (2.79)$

since $\nabla e^{-ikr} = -jke^{-ikr}$, as may be verified by expansion in rectangular coordinates. The divergence condition is seen to constrain the amplitudes A, B, C so that the vector \mathbf{E}_0 is perpendicular to the direction of propagation as specified by \mathbf{k} . The solution (2.78) is called a uniform plane wave since the constant-phase surfaces given by $\mathbf{k} \cdot \mathbf{r} = \text{const}$ are planes and the field \mathbf{E} does not vary on a constant-phase plane.

The solution for \mathbf{H} is obtained from Maxwell's equation

$$\nabla \times \mathbf{E} = -V_0 \omega \mathbf{A}_0 \mathbf{H}$$

which gives

$$\begin{aligned} \mathbf{H} &= \frac{1}{V_0 \omega \mu_0} \nabla \times \mathbf{E}_0 e^{-ikr} = \frac{1}{V_0 \omega \mu_0} \mathbf{E}_0 \times \mathbf{n} e^{-ikr} \\ &\leftarrow \frac{1}{V_0 \omega \mu_0} \mathbf{E}_0 \times \mathbf{n} e^{-ikr} = \frac{\mathbf{n} \times \mathbf{E}}{V_0 \omega \mu_0} \\ &\sqrt{\frac{\epsilon_0}{\mu_0}} \mathbf{n} \times \mathbf{E} = \mathbf{y}_0 \mathbf{n} \times \mathbf{E} \end{aligned} \quad (2.80)$$

where $Y_0 = \epsilon_0 / (\mu_0 V_0)$ has the dimensions of an admittance and is called the intrinsic admittance of free space. The reciprocal $Z_0 = 1/Y_0$ is called the

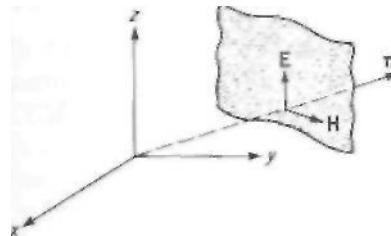


FIGURE 2.12

Space relationship between E, H, and n in a TEM

intrinsic impedance of free space. Note that H is perpendicular to E and to n, and hence both E and H lie in the constant-phase planes. For this reason this type of wave is called a transverse electromagnetic wave (TEM wave). The spatial relationship between E, H, and n is illustrated in Fig. 2.12.

The physical electric field corresponding to the phasor representation (2.78) is

$$E = Re(E_0 e^{jk \cdot r + J_0 \theta}) = E_0 \cos(k \cdot r - \omega t) \quad (2.81)$$

where, for simplicity, E₀ has been assumed to be real. The wavelength is the distance the wave must propagate to undergo a phase change of 2π. If we let A₀ denote the wavelength in free space, it follows that

$$|k|A_0 = k_0 A_0 = 2\pi$$

so that

$$k_0 = \frac{2\pi}{c} = \frac{2\pi}{\lambda} \quad (2.82)$$

This result is the familiar relationship between wavelength A₀, frequency f = c/λ, and velocity c in free space. A wavelength in a direction other than that along the direction of propagation n may also be defined. For example, along the direction of the x axis the wavelength is

$$\lambda_x = \frac{2\pi}{k_x} \quad (2.83)$$

and since k_x is less than k₀, λ_x is greater than A₀. The phase velocity is the velocity with which an observer would have to move in order to see a constant phase. From (2.81) it is seen that the phase of E is constant as long as k · r - wt is constant. If the angle between k and r is θ, then k · r - ωt = k₀r cos θ - ωt. Differentiating the relation

$$k_0 r \cos \theta - \omega t = \text{const}$$

gives

$$\frac{dr}{dt} = \frac{\omega}{k_0 r \cos \theta} \quad (2.84)$$

for the phase velocity u_p in the direction r. Along the direction of propagation, cos θ = 1 and v_p = c/k₀ = c. In other directions, the phase velocity is

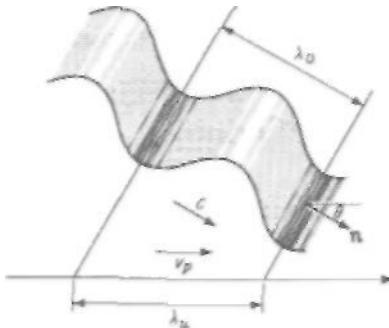


FIGURE 2.13

A wave propagating obliquely to the u axis.

greater than c . These results may be understood by reference to Fig. 2.13. When the wave has moved a distance A_0 along the direction n , the constant-phase-plane intersection with the u axis has moved a distance $\lambda_u = A_0 \sec \theta$ along the direction u . For this reason the wavelength and phase velocity along u are greater by a factor $\sec \theta$ than the corresponding quantities measured along the direction of propagation n .

The time-average rate of energy flow per unit area in the direction n is given by

$$P = |\operatorname{Re} E \times H^* \cdot n| = |\operatorname{Re} y_0 E \times (n \times E^*) \cdot n| = |Y_0 E_0 \cdot E_g| \quad (2.85)$$

The time-average energy densities in the electric and magnetic fields of a TEM wave are, respectively,

$$U_e = \sim E \cdot E^* = \wedge E_0 \cdot E^*$$

$$U_m = \frac{1}{4} H \cdot H^* = \frac{1}{4} y_0 (n \times E) \cdot (n \times E^*) = \frac{1}{4} E_0 \cdot E^* = U_e$$

and are seen to be equal. Since power is a flow of energy, the velocity v_g of energy propagation is such that

$$(U_e + U_m)v_g = P$$

$$\frac{1}{2} \frac{\omega_0 \lambda_0}{c} \frac{|E_0|^2}{\epsilon_0} = \frac{Y_0}{4} \frac{v_g}{c} \quad (2.86)$$

Thus, for a TEM wave in free space, the energy in the field is transported with a velocity $c = 3 \times 10^8$ m/s, which is also the phase velocity. Since phase velocity is independent of frequency, a modulated carrier or signal will have all its frequency components propagated with the same velocity c . Hence the signal velocity is also the velocity of light c . Later on, in the study of waveguides, situations arise where the phase velocity is dependent on frequency and consequently is not equal to the velocity of energy propagation or the signal velocity.

REFLECTION FROM A DIELECTRIC INTERFACE

In Fig. 2.14 the half-space $z > 0$ is filled with a dielectric medium with permittivity ϵ (dielectric constant $\epsilon_r = \epsilon/\epsilon_0$; index of refraction $77 = \sqrt{\epsilon}$). A TEM wave is assumed incident from the region $2 < 0$. Without loss in generality, the xy axis may be oriented so that the unit vector n_i specifying the *direction of incidence* lies in the xz plane. It is *convenient to* solve this problem as two special cases, namely (1) parallel polarization, where the electric field of the incident wave is coplanar with n_i , and the interface normal, i.e., lies in the xz plane, and (2) perpendicular polarization, where the electric field of the incident wave is perpendicular to the plane of incidence as defined by n_i , and the interface normal, i.e., along the y axis. An incident TEM wave with arbitrary polarization can always be decomposed into a linear sum of perpendicular and parallel polarized waves. The reason for treating the two polarizations separately is that the reflection and transmission coefficients, to be defined, are different for the two cases.

Parallel Polarization

Let the incident TEM wave be

$$E_i = E_1 e^{-jkn_i z} \quad H_i = y_0 n_i \times E_i \quad (2.87)$$

where E_i lies in the xz plane. Part of the incident power will be reflected, and the remainder will be transmitted into the dielectric medium. Let the reflected TEM wave be

$$E_r = E_2 e^{-jk(n_2 + n_3)z} \quad H_r = y_0 n_2 \times E_r \quad (2.88)$$

where n_2 and E_2 are to be determined. In the dielectric medium the solution for a TEM wave is the same as that in free space, but with ϵ_0 replaced by ϵ . Thus, in place of $A_0 = c\omega^2 f_0 \epsilon_0$ and $Y_0 = \omega^2 \epsilon_0 / n_0$, the parameters $k = \omega j n_0 \epsilon = -qk_0$ and $Y = \omega j \epsilon / n_0 = I^2 \rho$ are used, where $77 = \sqrt{\epsilon} \epsilon_r$ is the index of refraction. The transmitted wave in the dielectric may be

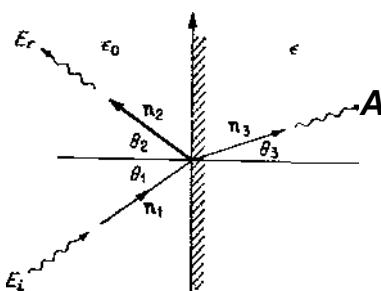


FIGURE 2.14
Plane wave incident on a dielectric interface.

expressed by

$$E_r = E_3 e^{j(kz - \omega t)} \quad H_r = Y n_3 \times E_r \quad (2.89)$$

with E_3 and n_3 as yet unknown.

The boundary conditions to be applied are the continuity of the tangential components of the electric and magnetic fields at the interface plane $z = 0$. These components must be *continuous for all values of x and y* on the $z = 0$ plane, and this is possible only if the fields on adjacent sides of the boundary have the same variation with x and y . Hence we must have

$$k_0 n_{lx} = k_0 n_{2l} = k n_{ax} = T k_0 n_{3x} \quad (2.90)$$

i.e., the propagation phase constant along x must be the same for all waves. Since n_{ly} was chosen as zero, it follows that $n_{2y} = n_3 = 0$ also. The unit vectors n_1, n_2, n_3 may be expressed as

$$n_1 = a_x \sin \theta_1 + a_z \cos \theta_1,$$

$$n_2 = a_x \sin \theta_2 + a_z \cos \theta_2$$

$$n_3 = a_x \sin \theta_3 + a_z \cos \theta_3$$

Equation (2.90) gives

or

$$e_i = B_2 \quad (2.91)$$

which is the well-known Snell's law of reflection; in addition, (2.90) gives

$$\sin \theta_1 = \frac{1}{n_3} \sin \theta_2, \quad (2.92)$$

which is also a well-known result specifying the angle of refraction θ_3 in terms of the angle of incidence θ_2 and the index of refraction n_3 .

The incident electric field E_s has components $E_{lx} = E_s \cos \theta_1$,

$$E_{lz} = -E_s \sin \theta_1$$

since $n_1 \cdot E_r$ must equal zero. Note that E_s is used to denote the magnitude of the vector E_s . Since the incident electric field has no y component, the reflected and transmitted electric fields also have zero y components.^f Expressing all fields in component form, i.e.,

$$E_{2x} = E_2 \cos \theta_2$$

$E_{2z} = E_2 \sin \theta_2$, $E_{3l} = E_3 \cos \theta_3$, $E_{3z} = -E_s \sin \theta_3$, and imposing the boundary condition of continuity of the x component at $z = 0$ yields the relation

$$E_s \cos \theta_1 + E_2 \cos \theta_2 = E_3 \cos \theta_3$$

^f If the reflected and transmitted electric fields were assumed to have a y component, boundary conditions which must apply would show that these are, indeed, zero.

$$\text{or } (E_l + E_2) \cos e_x = E_r f l - \frac{\sin^2 \theta_3}{V} \frac{1/e_r - \sin^2 \theta_1}{(2.93)}$$

by using (2.91) and (2.92). Apart from the propagation factor, the magnetic field is given by

$$H_x = * > ! X E, = Y_0 a_y (-n_{l1} E_u + n_{l2} E_u) = y_0 a_y E,$$

$$H_2 = -Y_0 a_y E_2$$

$$H_3 = 7 a_y \epsilon_3$$

and has only a y component. Continuity of this magnetic field at the boundary requires that

$$Y_0(E_l E_2) = Y E, = v Y_0 E_3 \quad (2.94)$$

If a reflection coefficient τ and a transmission coefficient T , are introduced according to the following relations:

amplitude of reflected electric field E_2

amplitude of incident electric field E_x

amplitude of transmitted electric field E_z ,

$$\tau = \frac{E_2}{E_x} = \frac{(c_r - \sin^2 \theta_1)^{1/a}}{\pi \cos \theta_1} \quad (2.956)$$

then the boundary conditions (2.93) and (2.94) may be expressed as

$$1 + T, = \tau, \frac{(c_r - \sin^2 \theta_1)^{1/a}}{\pi \cos \theta_1} y \quad (2.96a)$$

$$1 - V_x = 777, \quad (2.966)$$

These equations may be solved to give the Fresnel reflection and transmission coefficients for the case of parallel polarization, namely,

$$T, = \frac{f e_r - \sin^2 \theta_1}{(e_r - \sin^2 \theta_1)} \frac{e_r \cos \theta_1}{e_r + \cos \theta_1}, \quad (2.97a)$$

$$T, = \frac{277 \cos \theta_1 j}{(e_r - \sin^2 \theta_1)^{1/2} + e_r \cos \theta_1}, \quad (2.976)$$

An interesting feature of r is that it vanishes for an angle of incidence $\theta_t = \theta_{ft}$, called the Brewster angle, where, from (2.97a),

$$e_r - \sin^2 \theta_{fc} = e^2 \cos^2 \theta_b$$

$$\text{or } \sin \theta_6 = \frac{I * r}{(I * r)^{1/2}} \quad (2.98)$$

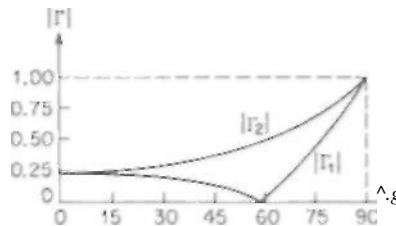


FIGURE 2.15

Modulus of reflection coefficient at a dielectric interface for $\epsilon_r = 2.56$, $11''$, I parallel polarization, $|r|$ perpendicular polarization.

At this particular angle of incidence all the incident power is transmitted into the dielectric medium. In Fig. 2.15 the reflection coefficient $|r|$, is plotted as a function of θ_i , for polystyrene, for which $\epsilon_r = 2.56$.

2 Perpendicular Polarization

For perpendicular polarization the roles of electric and magnetic fields are interchanged so that the electric field has only a y component. The fields may, however, still be expressed in the form given by (2.87) to (2.89), but with E_1 , E_2 , and E_3 having y components only. As in the previous case, the boundary conditions must hold for all values of x and y on the $z = 0$ plane. Therefore Snell's laws of reflection and refraction again result; i.e., (2.91) and (2.92) must be satisfied. In place of the boundary conditions (2.93) and (2.94), we have

$$\mathbf{E}_1 + \mathbf{E}_2 = \mathbf{E}_3 \quad (2.99a)$$

$$Y_0(\mathbf{E}_1 - E^{\wedge} \cos \theta_i) = Y \mathbf{E}_3 \cos \theta_i \quad (2.99b)$$

Introducing the following reflection and transmission coefficients:

$$\mathbf{r}_2 =$$

into (2.99) yields

$$1 + T_2 = T_2 \quad (2.100a)$$

$$\mathbf{i} - \mathbf{r}_2 = T_2 \frac{(\epsilon_r - \sin^2 \theta_i)^{1/2}}{\cos \theta_i}, \quad (2.100b)$$

The Fresnel reflection and transmission coefficients for the case of perpendicular polarization thus are

$$\mathbf{r}_2 = \frac{\cos \theta_i - (\epsilon_r - \sin^2 \theta_i)^{1/2}}{(\epsilon_r - \sin^2 \theta_i)^{1/2} + \cos \theta_i} \quad (2.101a)$$

$$T_2 = \frac{2 \cos \theta_i}{U_r - \sin^2 \theta_i^{1/2} + \cos \theta_i}, \quad (2.101b)$$

A notable difference for this case is the nonexistence of a Brewster angle for which T_2 vanishes. For comparison with the case of parallel polarization, F_2 is plotted in Fig. 2.15 for $\epsilon_r = 2.56$.

REFLECTION FROM A CONDUCTING PLANE

The essential features of the behavior of the electromagnetic field at the surface of a good conductor may be deduced from an analysis of the simple problem of a TEM wave incident normally onto a conducting plane. The problem is illustrated in Fig. 2.16, which shows a medium with parameters ϵ, μ, a filling the half-space $z > 0$. Let the electric field be polarized along the x axis so that the incident and reflected fields may be expressed as

$$\begin{aligned} E_i &= E_i a_x e^{-jk_0 z} \\ H_i &= Y_0 E_i a_y e^{-jk_0 z} \\ E_r &= r_i E_i a_r e^{+jk_0 z} \\ H_r &= -Y_0 r_i E_i a_v e^{+jk_0 z} \end{aligned} \quad (2.1026)$$

where T is the reflection coefficient.

In the conducting medium the conduction current $\text{tr}E$ is much greater than the displacement current $j\epsilon_0\mu E$, so that Helmholtz's equation reduces to (2.50); i.e.,

$$\nabla^2 E - j\omega n a E = 0$$

The transmitted field is a solution of

$$\frac{\delta^x}{\mathbf{J}_z^{1/2}} - j\omega \mu H - T E = 0$$

since no x or y variation is assumed. The solution for a wave with an x component only and propagating in the z direction is

$$E_t = f_z a_{r,t} e^{+jk_0 z} \quad (2.103a)$$

$$f_z = \frac{1}{2} \left(a_r + \frac{a_v}{Y_0} \right)$$

FIGURE 2.16
A TEM wave incident normally on a conducting plane.

with a corresponding magnetic field

$$\mathbf{H} \cdot \mathbf{J} = \frac{1}{V} \mathbf{E} \times \mathbf{B}, = \frac{y}{B_v E_3 e^{-y}} \quad (2.1036)$$

$$\text{where } y = (ja^2 i^2 r)^{1/2} = \frac{*^+ J}{\omega} \quad (2.104)$$

and the skin depth $S_s = (\omega / \pi \sigma / 2)^{1/2}$. The propagation constant $y = a + jk$ has equal phase and attenuation constants. In the conductor the fields decay by an amount e^{-y} in a distance of one skin depth S_s , which is a very small distance for metals at microwave frequencies (about 10^{-5} cm). The intrinsic impedance of the metal is Z_m where

$$(jwn^*)^{1/2} \quad T S_s \quad (2.105)$$

and is very small compared with the intrinsic impedance $Z_0 = (f_0 / \epsilon_0)^{1/2}$ of free space. For example, for copper at 10¹⁰ MHz, $Z_m = 0.026(1 + j)$ ft as compared with 377 ft for Z_0 . Note that (2.1036) may be written as

$$H_z = -a_y E_3 e^{-y} = Y_m 8i_v E_3 e^{-y}$$

which shows that the ratio of the magnitudes of the electric field to the magnetic field for a TEM wave in a conductor is the intrinsic impedance Z_m .

Returning to the boundary-value problem and imposing the boundary conditions of continuity of tangential fields at the boundary plane $z = 0$ give

$$(l + r)E_l = E_a = TE_l \quad (2.106a)$$

$$(1 - F)F_0 \mathbf{J} \cdot \mathbf{J}_i = H_z = Y_m E_3 = Y_m TE_x \quad (2.106b)$$

where $E_3/E_l = T$, the transmission coefficient. Solving (2.106) for the reflection coefficient V and T yields

$$F = \frac{Z_m}{Z_m + Z_0} \sim \frac{Z_m}{Z_0} \quad (2.107c)$$

$$T = 1 + r = \frac{Z_m}{Z_m + Z_0} \quad (2.107d)$$

$$Z_m + Z_0$$

Since $|Z_m|$ is small compared with Z_0 , the reflection coefficient R is almost equal to -1 and the transmission coefficient T is very small. Almost all the incident power is reflected from the metallic boundary. As the conductivity σ is made to approach infinity, the impedance Z_m approaches zero and in the limit $T = -1$ and $R = 0$. Hence, for a perfect conductor, the tangential electric field at the surface is zero and the tangential magnetic field has a value equal to twice that of the incident wave.

The current density in the conductor is $\mathbf{J} = \sigma \mathbf{E}_t = aT^*E^*e^{-yz}$. The total current per unit width of conductor along y is

$$J_s = \int_0^\infty J dz = \sigma T^*a, \quad \int_0^\infty e^{-yz} dz = \frac{aT^*a}{y} A/m$$

This result may also be expressed in the following form:

$$< \frac{2aZ_m E_t}{(Z_m + Z_0)j} on^* \quad (2.108)$$

by substituting for T from (2.1076) and replacing y by $jiof/Z_m$ from (2.105). As $\langle T \rangle^*$, the limiting value of J_s becomes

$$J_s = -s \cdot a, = 2y, \epsilon_{ia}, \quad (2.109)$$

since $Z_m \rightarrow 0$ and $\langle rZ_t^* \rangle \rightarrow yw/x$. This current exists only on the surface of the conductor since, as $a \rightarrow \infty$, the skin depth $\delta \rightarrow 0$; that is, the field decays infinitely fast with distance into the conductor. When a is infinite, $T = -1$ and the total tangential magnetic field at the surface is $2Y_u E_a$ and equal in magnitude to J_s . In vector form the boundary conditions at the surface of a perfect conductor are thus seen to be

$$\mathbf{n} \times \mathbf{E} = 0 \quad (2.110a)$$

$$\mathbf{n} \times \mathbf{H} = \mathbf{J}, \quad (2.110b)$$

where n is a unit outward normal at the conductor surface.

For finite conductivity the current density at the surface is $\sigma T \mathbf{E}_t$, and the magnetic field at the surface is $Y_m T E_t$. In terms of these quantities the total current per unit width may be expressed as

$$\begin{matrix} \langle TTE_t \rangle & aZ_m \\ y & y \end{matrix}$$

In other words, the total current per unit width is equal to the tangential magnetic field at the surface.

The time-average power transmitted into the conductor per unit area is given by the real part of one-half of the complex Poynting vector at the surface, and is

$$P_s = |\operatorname{Re}(\mathbf{E} \times \mathbf{H}^*) \cdot \mathbf{a}_2| = \pm TT^*E_t E_f \operatorname{Re} Y_m = |TT^*E_t E_f| a_s \quad (2.111)$$

The reader may readily verify that this is equal to the result obtained from a volume integral of $\mathbf{J} \cdot \mathbf{J}^*$; that is,

$$\oint_S \mathbf{J} \cdot d\mathbf{S}$$

Equation (2.111) may be simplified with little error by making the following

approximation:

$$\begin{aligned} aTT^* & \quad 4<\!rZ_{,,}Z^* \\ (Z_m + Z_0)(Z_{ra} + Z_0)^* & \\ 4aZ_{,,}Z^* & \quad 8 \\ Z_0^2 & \quad a8^*Z^* \end{aligned}$$

whence (2.111) becomes

$$P_{<\sim 2} = \frac{1}{(2Y_0E_l)(2Y_uEt)} >^{2112}$$

Note that $2Y_0E_l$ is the value of the magnetic field, tangent to the surface that would exist if a were infinite. Hence an excellent approximate technique for evaluating power loss in a conductor is to find the tangential magnetic field, say H_n that would exist for a perfect conductor, and then compute the power loss according to the relation

$$P_r = |Re(H_n H^* Z_m)| = \pm Re(J_s J^* Z_m) \quad (2.113)$$

This procedure is equivalent to assuming that the metal exhibits a surface impedance $Z_{,,}$ and the current is essentially the same as that which would exist for infinite conductivity.

The procedure outlined above for power-loss calculations is widely used in microwave work. Although the derivation was based on a consideration of a very special boundary-value problem, the same conclusions result for more complex structures such as conducting spheres and cylinders. In general, the technique of characterizing the metal by a surface impedance $Z_{,,}$ and assuming that the current J_s is the same as that for infinite conductivity is valid as long as the conductor surface has a radius of curvature at least a few skin depths in magnitude.

2.10 POTENTIAL THEORY

The wave solutions presented in the previous sections have all been source-free solutions; i.e., the nature of the sources giving rise to the field was not considered. When it is necessary to consider the specific field generated by a given source, as in antenna problems, waveguide and cavity coupling, etc., this is greatly facilitated by introducing an auxiliary vector potential function A . As will be seen, the vector potential A is determined by the current source, and the total electromagnetic field may be derived from A .

Since $\mathbf{V} \cdot \mathbf{B} = 0$ always, this condition will hold identically if \mathbf{B} is expressed as the curl of a vector potential A since the divergence of the curl of a vector is identically zero. Thus let

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2.114)$$

The assumed time dependence $e^{j\omega t}$ is not written out explicitly in (2.114)

since this is a phasor representation. The curl equation for \mathbf{E} gives

$$\nabla \times \mathbf{E} = -\nu' w \mathbf{B} = -jcoV \times \mathbf{A}$$

or

$$\nabla \times (\mathbf{E} + j\omega \mathbf{A}) = 0$$

Now the curl of the gradient of a scalar function Φ is identically zero; so the general integral of the above equation is

$$\mathbf{E} + j\omega \mathbf{A} = -\nabla \Phi$$

or

$$\mathbf{E} = -j\omega \mathbf{A} - \nabla \Phi \quad (2.115)$$

Substituting this expression into the $\nabla \times \mathbf{H}$ equation gives

$$\nabla \times \mathbf{H} = -\nabla \times \nabla \times \mathbf{A} = j\omega \mathbf{E} + \mathbf{J} = w^2 c \mathbf{A} - j\omega e V \Phi + \mathbf{J} \quad (2.116)$$

Up to this point the divergences of \mathbf{A} and $\nabla \Phi$ have not been specified [note that (2.114) specifies the curl of \mathbf{A} only]. Therefore a relation between $\nabla \cdot \mathbf{A}$ and Φ may be chosen so as to simplify (2.116). Expanding $\nabla \times \nabla \times \mathbf{A}$ to give $\nabla \nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A}$ enables us to write (2.116) as

$$\nabla \nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A} = k^2 \mathbf{A} - j\omega e / i \nabla^* + j\omega \mathbf{J}$$

where $k^2 = w^2 p.e.$ If now the following condition is specified:

$$\nabla \nabla \cdot \mathbf{A} = -j\omega e / i \nabla^* \Phi$$

or

$$\nabla \cdot \mathbf{A} = -j\omega e / i \nabla^* \Phi \quad (2.117)$$

this equation simplifies to

$$\nabla^2 \mathbf{A} + k^2 \mathbf{A} = -IM \mathbf{J} \quad (2.118)$$

Thus \mathbf{A} is a solution of the inhomogeneous Helmholtz equation, the current \mathbf{J} being the source term. The condition imposed on $\nabla \cdot \mathbf{A}$ and Φ in (2.117) is called the Lorentz condition in honor of the man first to propose its use.

In the preceding derivation three of Maxwell's equations have been used and are therefore satisfied. The fourth equation, $\nabla \cdot \mathbf{D} = p$, must also hold, and this will be shown to be the case provided the Lorentz condition is obeyed. Hence the three equations (2.114), (2.115), and (2.118), together with the Lorentz condition (2.117), are fully equivalent to Maxwell's equations. To verify the equation $\nabla \cdot \mathbf{D} = p$, take the divergence of (2.115) to obtain

$$\nabla \cdot \mathbf{E} = -\nu' w \nabla \cdot \mathbf{A} - e \nabla^2 \Phi \quad (2.119)$$

where e is a constant. Using the Lorentz condition yields

$$\nabla \cdot \mathbf{D} = -jcoe \nabla \cdot \mathbf{A} - \nabla^2 \frac{1}{-\nu' w} = \frac{1}{\nu' w} \nabla \cdot (\nabla^2 \mathbf{A} + k^2 \mathbf{A}) = -\frac{1}{\nu' w} \nabla \cdot \mathbf{J}$$

by using (2.118) and noting that $\nabla^2 \nabla \cdot \mathbf{A} = \nabla \cdot \nabla^2 \mathbf{A}$; that is, these differential operators commute. Now $\nabla \cdot \mathbf{J} = -jcop$ from the continuity equation;

so we obtain

$$\nabla D = - (-j\omega \rho) = p$$

If, instead of eliminating $\nabla \cdot A$ in (2.119), $\nabla \cdot A$ is eliminated by use of the Lorentz condition, we get

$$\nabla \cdot D = p = -j\omega \epsilon \nabla \times H - \mu \nabla^2 D$$

or

$$\nabla^2 D + \mu \nabla^2 T = -p \quad (2.120)$$

Hence the scalar potential ϕ is a solution of the inhomogeneous scalar Helmholtz equation, with the charge density p as a source term.

For the time-varying field, J and p are not independent, and hence the field can be determined in terms of A and J alone. The scalar potential can always be found from the Lorentz relation, and p from the continuity equation, but explicit knowledge of these is not required in order to solve radiation problems. For convenience, the pertinent equations are summarized here:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2.121a)$$

$$\mathbf{E} = -j\omega \mathbf{A} - \nabla \phi = -j\omega \mathbf{A} + \frac{\nabla V}{V} \cdot \mathbf{A} - \frac{k^2 A}{V} + \frac{\nabla V}{V} \cdot \mathbf{A} \quad (2.121b)$$

$$\nabla^2 \mathbf{A} + k^2 \mathbf{A} = -\nabla \times \mathbf{J} \quad (2.121c)$$

where the Lorentz condition was used to eliminate $\nabla \phi$ in (2.121b). Note that, in rectangular coordinates, (2.121c) is three scalar equations of the form

$$\nabla^2 A_x + k^2 A_x = -\partial J_x / \partial z$$

but that, in other coordinate systems, $\nabla^2 \mathbf{A}$ must be expanded according to the relation $\nabla^2 \mathbf{A} = \nabla \nabla \cdot \mathbf{A} - \nabla \times (\nabla \times \mathbf{A})$.

The simplest solution to (2.121c) is that for an infinitesimal current element $J(x', y', z') = J(r')$ located at the point (x', y', z') , as specified by the position vector $r' = a_x x' + a_y y' + a_z z'$, as in Fig. 2.17. This solution is

$$\mathbf{A}(x, y, z) = \mathbf{A}(r) = \frac{1}{4\pi r} \int \frac{J(r')}{r'} dV \quad (2.121d)$$

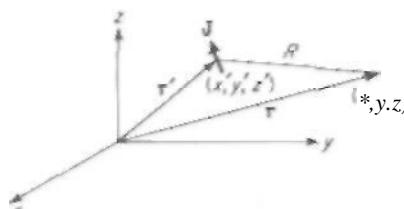


FIGURE 2.17
Coordinates used to describe vector potential a current sheet.

where $R = |\mathbf{r} - \mathbf{r}'|$ is the magnitude of the distance from the source point to the field point at which \mathbf{A} is evaluated; i.e.,

$$R = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}^{1/2}$$

and $J(\mathbf{r}')\mathbf{rfV}'$ is the total source strength. In terms of this fundamental solution, the vector potential from a general current distribution may be obtained by superposition. Thus, adding up all the contributions from each infinitesimal current element gives

$$\mathbf{A}(\mathbf{r}) = \frac{-jkR}{4\pi r' V} \int dV \frac{J(x'y'z')}{R} = \frac{-jkR}{4\pi r' V} \int dV \frac{J(\mathbf{r}')}{R} \frac{-\mathbf{J}^*|\mathbf{r}-\mathbf{r}'|}{4\pi J_y |\mathbf{r} - \mathbf{r}'|} \quad (2.123)$$

where the integration is over the total volume occupied by the current. Note that the solution for \mathbf{A} as given by (2.122) is a spherical wave propagating radially outward from \mathbf{J} and with an amplitude falling off as $1/R$. The solution (2.123) is a superposition of such elementary spherical waves.

*2.11 DERIVATION OF SOLUTION FOR VECTOR POTENTIAL

In this section a detailed derivation of the solution to the inhomogeneous Helmholtz equation for a unit current source is given. A unit source is a source of unit strength, localized at a point in space (a familiar example is a point charge). Such a unit source in a three-dimensional space is a generalization of a unit current impulse localized at a time t' along the time coordinate. A current pulse is represented by the Dirac delta function $\delta(t - t')$ in circuit theory, where $\delta(t - t')$ has the property

$$\delta(t - t') = 0 \quad t \neq t' \quad (2.124a)$$

and at $t = t'$ it becomes infinite but is integrable to give

$$\int \delta(t - t') dt = 1 \quad (2.124b)$$

A further property is that, for any function $f(t)$ which is continuous at t'

$$\int_{t'-\epsilon}^{t'+\epsilon} f(t) \delta(t - t') dt = f(t') \quad (2.124c)$$

This result follows since \mathbf{r} can be chosen so small that, in the interval $t' - T < t < t' + T$, the function $f(t)$ differs by a vanishing amount from $f(t')$ since $f(t)$ is continuous at t' . Hence (2.124c) may be written as

$$\int_{t'-\epsilon}^{t'+\epsilon} f(t) \delta(t - t') dt = f(t')$$

by virtue of (2.124b).

As the preceding discussion has shown, the delta function is a convenient mathematical way to represent a source of unit strength localized at a point along a coordinate axis, in the above example along the time axis. In an n -dimensional space a product of N delta functions, one for each coordinate, may be used to represent a unit source. Thus, in three dimensions, a unit source is represented by

$$\delta(x-x')\delta(y-y')\delta(z-z') = S(r - r') \quad (2.125)$$

where $S(r - r')$ is an abbreviated notation for the product of the three one-dimensional delta functions. The source function $S(r - r')$ has the following properties:

$$\delta(r - r') = 0 \quad r \neq r' \quad (2.126a)$$

\int

$$\frac{\delta(\mathbf{F}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')\mathbf{i}\nabla-\mathbf{P}(-))}{-\nabla} * \cdot \cdot \cdot \nabla \quad (2.126c)$$

$\backslash \mathbf{O} \quad r' \text{ not in } V$

where F is an arbitrary vector (or scalar) function that is continuous at r' , that is, at x', y', z' . These properties follow from the properties of the one-dimensional delta functions that make up $S(r - r')$.

A unit current source directed along the unit vector a at r' may be expressed as $\mathbf{J} = a\delta\mathbf{r} - r'$. The vector potential is a solution of

$$\nabla^2 A + k^2 A = -xa\delta(r - r') \quad (2.127)$$

Since the current is in the direction a , the vector potential must also be in this direction, and hence $A = Aa$. Equation (2.127) may therefore be written as a scalar equation:

$$\nabla^2 A + k^2 A = -ta\delta(r - r') \quad (2.128)$$

At all points $r \neq r'$, A is a solution of

$$\nabla^2 A + k^2 A = 0 \quad (2.129)$$

If the source point r' is considered as the origin in a spherical coordinate system, then, since no angle variables occur in the source term in (2.128), the solution for A must have spherical symmetry about the source point r' . Thus, in terms of the spherical radial coordinate $R = |r - r'|$, which is the radial distance from the origin at r' , (2.129) is a function of R only and may be written as

$$\text{or } \frac{\frac{R^2}{dR} \frac{dR}{dR}}{\frac{dR^2}{dR^2} + \frac{2dA}{R dR} + k^2 A} = 0 \quad (\nabla^2 A)$$

after expressing the independent part of ∇^2 in spherical coordinates.

anticipation of a spherical-wave solution, let $A = f(R)e^{-jkR}$. Substitution in (2.130) leads to the following equation for $f(R)$:

$$\frac{d^2f}{dr^2} + \frac{2f}{r} = -\frac{2jk}{r}$$

which is readily verified to have the solution $f \approx C/R$, where C is an arbitrary constant. Consequently, the solution to (2.129) is $A = Ce^{-jkR}/R$. This solution is singular at $R = 0$, and the singularity must correspond to that of the source term at this point.

To determine the constant C , integrate (2.128) throughout a small sphere of radius r_0 centered on r' and use the delta-function property (2.1266) to obtain

$$\begin{aligned} & \int_{r_0}^{2r_0} \int_V V^2 A + k^2 A R^2 \sin \theta d\theta d\phi dR \\ & - \int_V (\nabla^2 A + k^2 A) dV = \int_V -n/8(r - r') dV = -u \end{aligned}$$

Now the integral of the term $k^2 R^2 A$, which is proportional to R^2 , will vanish as r_0 tends to zero. Hence, for sufficiently small r_0 ,

$$\int_V V^2 A dV = -a$$

Since $\nabla^2 A = \nabla \cdot \nabla A$, the divergence theorem may be used to give

$$\int_V V^2 A dV = \int_S \nabla A \cdot dS = j \int_S \nabla A \cdot a_r r^2 d\Omega$$

since $dS = a_r r^2 d\Omega$, where $d\Omega$ is an element of solid angle. Since A is a function of R only, $\nabla A = a_r (dA/dR)$, and hence

$$\int_S \nabla A \cdot dS = \int_S \frac{t_r}{r^2} \frac{\partial A}{\partial r} d\Omega = \frac{t_r}{r_0^2} \frac{9A}{6} = \frac{3A}{4r_0^2} = -n$$

Evaluating dA/dR for $R = r_0$ shows that

$$\frac{-?A}{4ir/rr^2} \frac{5fl}{= -4irCr_*} \frac{/f}{r_0^2} \frac{e^{-jkro}}{\backslash}$$

in the limit as r_0 tends to zero. Hence $4irC = jx$, or $C = u/4\pi$, in order for the singularity in the solution for A to correspond to that for a unit source.

The above solution for the vector potential from a unit source, namely,

$$A(r) = \frac{M}{47r} \frac{e^{-jk(r-r')}}{|r - r'|} \quad (2.131)$$

is clearly a function of both the source point and field point. Since (2.131) is the solution for a unit source, it is often called a Green's function and

denoted by the symbol G as

$$G(r|r') = G(r|r')a = G(x,y,z|x',y',z')a = \frac{H}{AT} \frac{e^{-y^*|r-r'|}}{|T|} \quad (2.132)$$

because, by definition, a Green's function is the solution of a different equation for a unit source.

The vector potential from a general current distribution may now expressed in the form

$$A(r) = \frac{-/J(r')}{ATT} \frac{,/-\star f r - r' i}{r - r'} dV = \frac{\int J(r')G(r|r') dV}{-V} \quad (2.133)$$

since any current distribution J may be considered as a sum of weight unit sources.

2.12 LORENTZ RECIPROCITY THEOREM

The Lorentz reciprocity theorem is one of the most useful theorems in solution of electromagnetic problems, since it may be used to deduce number of fundamental properties of practical devices. It provides the basis for demonstrating the reciprocal properties of microwave circuits and showing that the receiving and transmitting characteristics of antennas are the same. It also may be used to establish the orthogonality properties of the modes that may exist in waveguides and cavities. Another important use is in deriving suitable field expansions (analogous to a Fourier series expansion) for the fields radiated or coupled into waveguides and cavities probes, loops, or coupling apertures.

To derive the theorem, consider a volume V bounded by a closed surface S as in Fig. 2.18. Let a current source J₁ in V produce a field E₁, H₁, while a second source J₂ produces a field E₂, H₂. These fields satisfy Maxwell's equations; so

$$\nabla \times E_1 = -j'w/iH_1, \quad \nabla \times H_1 = jcoeE_1 + J_1,$$

$$\nabla \times E_2 = -joffiH_2, \quad \nabla \times H_2 = jWE_2 + J_2$$

Expanding the relation $\nabla \cdot (E_1 \times H_2 - E_2 \times H_1)$ and using Maxwell's equations show that

$$\begin{aligned} & \nabla \cdot (E_1 \times H_2 - E_2 \times H_1) \\ &= (\nabla \times E_1) \cdot H_2 - (V \times H_2) \cdot E_1 - (V \times E_2) \cdot H_1 + (V \times H_1) \cdot E_2 \\ &= -J_2 \cdot E_1 + \langle J_1 \cdot E_2 \rangle \end{aligned} \quad (2.134)$$

In any waveguide or cavity an infinite number of field solutions are possible. Any one solution is called a mode for the same reason that the various solutions for vibrating strings and membranes are called modes. Orthogonality of modes is discussed in Sec. 3.14.

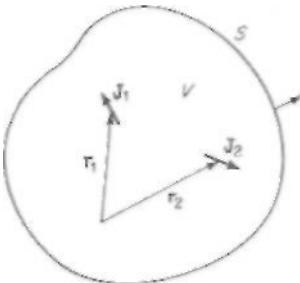


FIGURE 2.18
Illustration of the Lorentz reciprocity theorem.

Integrating both sides over the volume V and using the divergence theorem give

$$\begin{aligned} \int_V \nabla \cdot (E_1 \times H_2 - E_2 \times H_1) dV &= \int_S (E_1 \times H_2 - E_2 \times H_1) \cdot n dS \\ &= \int_V (E_2 \times J_1 - E_1 \times J_2) dV \quad (2.135) \end{aligned}$$

where n is the unit outward normal to S .

Equation (2.135) is the basic form of the Lorentz reciprocity theorem.! For a number of typical situations that occur, the surface integral vanishes. If S is a perfectly conducting surface, then $n \times E_1 = n \times E_2 = 0$ on S . Since $E_1 \times H_2 \cdot n = (n \times E_1) \cdot H_2$, etc., the surface integral vanishes in this case. If the surface S is characterized by a surface impedance Z_{ss} , then, according to (2.71),

$$E_1 = -Z_{ss} n \times H_1 \quad \text{or} \quad n \times E_1 = -Z_{ss} n \times (n \times H_1)$$

[note that in (2.71) n points into the region occupied by the field, and hence the minus sign is used here, since n is directed out of V]. Consequently,

$$\begin{aligned} (n \times E_1) \cdot H_2 - (n \times E_2) \cdot H_1 &= -Z_{ss} [n \times (n \times B_1) \cdot H_2 + Z_{ss} [n \times (n \times H_2)] \cdot H_1] \\ &= Z_{ss} (n \times H_1) \cdot (n \times H_2) - Z_{ss} (n \times H_2) \cdot (n \times H_1) = 0 \end{aligned}$$

and the surface integral vanishes again.

tin anisotropic media with nonsymmetric permittivity or permeability tensors, a modified form must be used. See, for example, R. F. Harrington and A. T. Villeneuve, Reciprocity Relations for Gyrotropic Media, *IRE Trans.*, vol. MTT-6, pp. 308-310. July, 1958.

Another example where the surface integral vanishes is when $\mathbf{J} \cdot \mathbf{n}$ is chosen as a spherical surface at infinity for which $\mathbf{n} = \mathbf{a}_r$. The radiated field at infinity is a spherical TEM wave for which

$$\mathbf{H} = \mathbf{Y} \mathbf{a}_r \mathbf{X} \mathbf{E} = \mathbf{I} \cdot \mathbf{I} \quad \mathbf{a}_r \times \mathbf{E}$$

Therefore

$$(n \times E_r) \cdot H_2 - (n \times E_2) \cdot H_1 = Y(a_r \times E_r) \cdot (a_r \times E_2) - F(a_r \times E_2) \cdot (a_r \times E_1) = 0$$

and the surface integral vanishes.

Actually, for any surface S which encloses all the sources for the field the surface integral will vanish. This result may be seen by applying (2.135) to the volume V bounded by S and the surface of a sphere of infinite radius. There are no sources in this volume, and since the surface integral over the surface of the sphere with infinite radius is zero, we must have, from (2.135),

$$\oint_S (E_j \times H_2 - E_2 \times H_j) \cdot (-n) dS = 0 \\ = (f)_{\sim S} (E_r \times H_2 - E_2 \times B_y \cdot n) dS$$

Hence the surface integral taken over any closed surface S surrounding all the sources vanishes.

When the surface integral vanishes, (2.135) reduces to

$$(E_r J_2 dV) = (E_2 J_1 dV) \quad (2.136)$$

If J_1 and J_2 are infinitesimal current elements, then

$$E_1(r_2) - J_2(r_2) = E_2(r_1) - J_1(r_1) \quad (2.137)$$

which states that the field E_x produced by J_x has a component along J_2 equal to the component along J_1 of the field radiated by J_2 when J_1 and J_2 have unit magnitude. The form (2.137) is essentially the reciprocity principle used in circuit analysis except that E and J are replaced by voltage V and current I . The applications of the reciprocity theorem are illustrated at various points throughout the text and hence are not discussed further at this time.

- 2.1. An atom of atomic number Z has a nuclear charge Ze and Z electrons revolving around it. As a model of this atom, consider the nucleus as a point charge and treat the electron cloud as a total charge $-Ze$ distributed uniformly throughout a sphere of radius r_0 . When an external field E is applied, the nucleus is displaced an amount x . Show that a restoring force $x(Ze)^2/4\pi r_0^2 e_0$ is produced and must be equal to ZeE . Thus show that the induced dipole moment is $p = 4\pi r_0^2 e_0 E$ and is linearly related to E .

- 2.2. In a certain material the equation of motion for the polarization is

$$\mathbf{IF} \frac{d^2 g}{dt^2} + \omega^2 p = 2\epsilon_0 \nabla \cdot \mathbf{E}$$

where W is the total field in the dielectric. Find the relation between ω and W when $r = Re(\epsilon - i\omega)$ and E is real. If $\omega_0 = 10^9$ and $v = 10^9$ over what frequency range can a relationship such as $2 = eg = e, \omega + 9^9$ be written if it is assumed that the criterion to be used is that the phase difference between \mathbf{E} and \mathbf{H} should not exceed 5° ? Plot the magnitude and phase angle of the dielectric constant $\epsilon_r = \epsilon/\epsilon_0 = (e' - ie'')/\epsilon_0$ as a function of ω .

- 2.3. A dielectric material is characterized by a matrix (tensor) permittivity

$$\begin{array}{c|ccc} & \text{xx} & \text{yy} & \text{zz} \\ \hline & 7 & 3 & -2v^5 \\ & 0 & 3 & -2v'0 \\ & -4 & -2/6T5 & -2v'0!5 \\ f_{\sim} & & & 10 \end{array}$$

when referred to the xyz coordinate frame. If the coordinate axis is rotated into the principal axis u, v, w , the permittivity is exhibited in diagonal form:

$$\mathbf{H} = \begin{pmatrix} s & 0 & 0 \\ 0 & e_m & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Find the principal axis and the values of the principal dielectric constants e_{uu}/ϵ_0 , etc.

Hint: By definition, along a principal axis a scalar equation such as $A_u = e_{uu}E_u$ holds. In general, if D is directed along a principal axis, then

$$\begin{array}{c|ccc} |D_x| & 7 & 3 & -2V/0JT \\ |D_y| & 0 & 3 & -2/61* \\ |D_z| & 4 & -2vfr5 & -2v^5 \\ \hline & 10 & & \end{array} \begin{array}{c|c} |E_x| & E \\ |E_y| & E \\ |E_z| & B \end{array}$$

or in words, when D is directed along a principal axis, it is related to E by a scalar constant A . The above constitutes a set of three homogeneous equations, of which the first is

Verify that, for a solution, the following determinant must vanish:

$$\begin{vmatrix} 7 - 4A/\epsilon_0 & 3 & -2\sqrt{0!5} \\ 3 & 7 - 4A/\epsilon_0 & -2v/\omega f \\ -2Z6\sim If & -2v^5 & 10 - 4A/\epsilon_0 \end{vmatrix} = 0$$

This cubic equation gives three roots for A, which may be identified as ϵ_w for one root, say ϵ_{uw} the components of a vector direction along the corresponding principal axis are proportional to the cofactors above determinant. This type of problem is called a matrix eigenvalue problem. The A's are the eigenvalues.

Answer: $3\epsilon_n = 2\epsilon_n$

Unit vectors along the principal axis are

$$\hat{a}_x = 0.5\hat{a}_x - \frac{1}{\sqrt{3}}\hat{a}_y - \frac{1}{\sqrt{3}}\hat{a}_z$$

$$\hat{a}_y = 0.5\hat{a}_x + 0.5\hat{a}_y + \frac{1}{\sqrt{3}}\hat{a}_z$$

$$\hat{a}_z = \frac{1}{\sqrt{3}}\hat{a}_x - \frac{1}{\sqrt{3}}\hat{a}_y + 0.5\hat{a}_z$$

- 2.4. In the interior of a medium with conductivity σ and permittivity ϵ , free charge density $p_0(x,y,z)$ at time $t = 0$. Show that the charge decays according to

$$p_0 e^{-\sigma t/\epsilon}$$

Evaluate the relaxation time T for copper for which $\sigma = 5.8 \times 10^7 \text{ S/m}$, $\epsilon = 80\epsilon_0$. Find T for sea water also for which $\sigma = 4 \text{ S/m}$ and $\epsilon = 80\epsilon_0$. If the relaxation time is short compared with the period of an applied time-harmonic field, there will be negligible accumulation of free charge and $\nabla \cdot D$ may be assumed to zero. What is the upper frequency limit for which this is true in the case of copper and sea water, i.e., the frequency for which T is equal to the period?

Hint: Use the continuity equation, Ohm's law, and the divergence theorem for D.

- 2.5. Show that, when the relaxation time for a material is small compared with the period of the time-harmonic field, the displacement current may be neglected in comparison with the conduction current.
- 2.6. Consider two concentric spheres of radii a and b . The outer sphere is kept at potential V, and the inner sphere at zero potential. Solve Laplace's equation in spherical coordinates to find the potential and electric field between the spheres. Take $b > a$.
- 2.7. Solve Laplace's equation to find the potential and electric field between coaxial cylinders of radii a and b if the center cylinder is kept at a potential V and the outer cylinder at zero potential. Take $b > a$.
- 2.8. Derive (2.45) from (2.18).
- 2.9. Derive (2.47).
- 2.10. Express the scalar Helmholtz equation $\nabla^2 V + k^2 V = 0$ in cylindrical coordinates. If $V = f(t)g(r)h(z)$, find the differential equations satisfied by f , g , and h .

- 2.11.** When material polarization ξP and Jt are explicitly taken into account, show that the wave equations satisfied by ψ and 2 are

$$\nabla^2 \psi - M_0 e_0 - i = \frac{Si^*}{-VV^* + n_a e_0 - t - Vx \cdot \nabla \times S}$$

$$\frac{\partial^2 \psi}{dt^2} - \frac{\partial^2 \psi}{at^2} = \frac{df}{at} - \frac{d}{at} = \frac{1}{e_0} \nabla \cdot \nabla \times \psi$$

Note that $\nabla \cdot \nabla S = 0$; so $\nabla \cdot JT = -V \cdot JT$ and $\nabla \cdot \mathbf{J} = p$; so $\nabla \cdot \mathbf{E} = p - \nabla \cdot \mathbf{J}$. Examination of the source terms in the above equations shows that $d\xi P/dt$ is a polarization current analogous to conduction current f .

- 2.12.** Derive (2.62).

- 2.13.** Between two perfectly conducting coaxial cylinders of radii a and b , $b > a$, the electromagnetic field is given by

$$E = a_r E_0 r \sim V^* e^{jkr} \quad H = \frac{1}{2\pi} Y_0 E_0 r^{-1} e^{-jkr}$$

where $k_0 = \omega(ji_0 \epsilon_0)^{1/2}$, $Y_0 = (e_0/\mu_0)^{1/2}$. Find the potential difference between the cylinders and the total current on the inner and outer cylinders. Express the power in terms of the voltage and current, and show that it is equal to that computed from an integration of the complex Poynting vector over the coaxial-line cross section. Show that the characteristic impedance of the line is $V/I = (Z_n/2i7) \ln(6/a) = 601 \ln(6/a)$, where V is the voltage and I is the total current on one cylinder.

- 2.14.** A round wire of radius r_0 much greater than the skin depth δ_s has a uniform electric field E applied in the axial direction at its surface. Use the surface-impedance concept to find the total current on the wire. Show that the ratio of the ac impedance of the wire to the dc resistance is

Evaluate this ratio for copper at $f = 10^6$ Hz for $a = 5.8 \times 10^7$ S/m, $r_0 = 0.1$ cm, $ix = n_0$.

- 2.15.** The half-space $z > 0$ is filled with a material with permittivity ϵ_r and permeability ix^* / μ_0 . A parallel polarized plane TEM wave is incident at an angle θ_i , as in Fig. 2.14. Find the reflection and transmission coefficients for the electric field. Does a Brewster angle exist for which the reflection coefficient vanishes?
- 2.16.** Repeat Prob. 2.15 for the case of a perpendicular polarized incident wave. Does a Brewster angle exist? If so, obtain an expression for it.
- 2.17.** The half-space $z > 0$ is filled with a material with permeability M and permittivity ϵ . When a plane wave is incident normally on this material, show that the reflection and transmission coefficients are

$$\mathbf{r} = \frac{Z - Z_0}{Z_0} \hat{\mathbf{z}} \quad T = \mathbf{i} + \mathbf{r} = \frac{2Z}{Z + Z_0}$$

where $Z = (ix/\epsilon)^{1/2}$, $Z_0 = (n_0/e_0)^{1/2}$. Choose an electric field with an x component only.

- 2.18. The half-space $z > 0$ is filled with a material of permittivity ϵ_2 and $f_i = ix_0$. A second sheet with permittivity ϵ_1 is placed in front. A plane wave incident normally on the structure from the left, as illustrated. Verify that the reflection coefficient at the first interface vanishes if $\epsilon_1 = (\epsilon_2 \epsilon_0)^{1/2}$ and the thickness $d = jA_0(\epsilon_0/\epsilon_1)^{1/2}$. The electric field may be assumed to have an component only. The matching layer is known as a quarter-wave transform (actually an impedance transformer). This matching technique is used reduce reflections from optical lenses and is called *lens blooming*, or coated lenses.

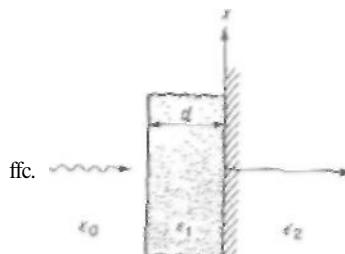


FIGURE P2.18

- 2.19. In terms of the vector potential A from a short current element $A?J_0a$, located at the origin, show that the radiated electric and magnetic fields are

$$\mathbf{H} = -\mathbf{a}^{\wedge} \sin ffe^{-jk_nr}$$

$$\mathbf{E} = \frac{r_0 \& 2}{2\pi} \frac{jZ_0}{k_u} \mathbf{a}_r \cos ffe^{-jk_nr}$$

$$\frac{1}{4\pi} \frac{QAZJZQ}{k_0} \frac{0}{r} \frac{J^* \mathbf{a}_r}{r^2} = \mathbf{a}_r \sin ffe^{-jk_nr}$$

Hint: Use (2.122) and (2.121), and express A as components in a spherical coordinate system r, θ, ϕ . Note that $a_z = a_r \cos \theta - a_\phi \sin \theta$.

- 2.20. A dielectric may be characterized by its dipole polarization P per unit volume. If $p = J = 0$ and P is taken into account explicitly, show that, if a vector potential A is introduced according to $B = \nabla \times A$, then A is a solution of

$$\nabla^2 A + k^2 A = -j\omega n_p P$$

and that the fields are given by

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \mathbf{E} = \frac{\nabla V \cdot \mathbf{A} + k^2 A}{-7\pi M_0 \epsilon_0}$$

Note that a Lorentz condition is used. Thus an electric dipole P is equivalent to a current element jwP , or alternatively, a current element $\langle J \rangle$ considered as an electric dipole $P = J/\langle \rangle$.



FIGURE P2.21

2.21. A small current loop constitutes a magnetic dipole $M = /Sa$, where $/$ is the current, S the area of the loop, and a a vector normal to the plane of the loop and pointing in the direction that a right-hand screw, rotating in the direction of the current, would advance. The field radiated by such a current loop, with linear dimensions much smaller than a wavelength, may be obtained by a potential theory analogous to that given in Prob. 2.20 by treating the loop as a magnetic dipole M . Thus replace B by $f_i H + \#_M$ in Maxwell's equation and treat M as a source term. Since p is zero, $V \cdot D = 0$, and this permits D to be expressed as $D = -V \times A_m$, where A_m is a magnetic-type vector potential. By paralleling the development in the text for the potential A , show that the following relations are obtained:

$$\nabla^2 A_m + k f A_m = -j u i e_0 M$$

$$D = -V \times A_m,$$

$$H = k A_m + V V \cdot A_m$$

Hence, for a z-directed magnetic dipole at the origin,

$$A_{m,i} = \frac{j t o f i e_0 M}{4\pi r^3} \quad < j > i < r >$$

from which the fields are readily found. Note that in this problem M represents the magnetic dipole source density in Maxwell's equations, but in the solution for the vector potential it represents the total magnetic dipole strength. It would have been more consistent to use $M S(r - r')$ to represent the source density, where $\delta(r - r')$ is the three-dimensional Dirac delta function which has the property

$$\int_V S(r - r') dV = 1 \quad r \text{ in } V$$

2.22. Consider an arbitrary current element J , in front of a perfectly conducting plane. This current radiates a field E , having zero tangential components on the conducting plane. Use the Lorentz reciprocity theorem to show that a current J_3 parallel to the conducting plane and an infinitesimal distance in front of it does not radiate.

REFERENCES

1. Kraus, J. D.: "Electromagnetics," 3rd ed., McGraw-Hill Book Company, New York, 1981
2. Hayt, W. H., Jr.: "Engineering Electromagnetics," 5th ed., McGraw-Hill Book Company, New York, 1989.
3. Schwarz, S. E.: "Electromagnetics for Engineers," Holt, Rinehart, and Winston, Inc., Philadelphia, 1990.
4. Johnk, C. T. A.: "Engineering Electromagnetic Fields and Waves." 2nd ed., John Wiley & Sons, Inc., New York, 1988.
5. Wait, J. R.: "Electromagnetic Wave Theory." Harper & Row Publishers, Inc., New York, 1985.
6. Stratton, J. A. "Electromagnetic Theory," McGraw-Hill Book Company, New York, 1941
7. Shen, L. C., and J. A. Kong: "Applied Electromagnetism," Brooks-Cole, Calif., 1983.

CHAPTER 3

TRANSMISSION LINES AND WAVEGUIDES

This chapter is a long one and for this reason has been divided into three parts, namely:

- Part 1—Waves on transmission lines
- Part 2—Field analysis of transmission lines
- Part 3—Rectangular and circular waveguides

The three parts are closely related but independent with the exception of Sec. 3.7, which is needed as an introduction to both Parts 2 and 3. With the exception of this section, the three parts can be studied independently and in any order.

In Part 1 we give an introduction to waves on transmission lines using a distributed-circuit model of the transmission line. By using the distributed-circuit model, we are able to study the excitation and propagation of current and voltage waves on a transmission line without the need to invoke Maxwell's equations.

The electrical characteristics of a transmission line such as the propagation constant, attenuation constant, characteristic impedance, and the distributed-circuit parameters can only be determined from a knowledge of the fields surrounding the transmission line. Thus in Part 2 we carry out a detailed field analysis of transmission lines. This part also includes an extensive discussion of planar transmission-line structures such as the microstrip line.

Part 3 presents the theory for waves in hollow rectangular and circular waveguides (pipes). In the beginning section of Part 2, we show that Maxwell's equations can be separated into equations that describe three types of waves. These are transverse electromagnetic waves (TEM) transverse electric (TE), and transverse magnetic (TM) waves. The TEM wave is the principal wave that can exist on a transmission line. The TE and TM waves are characterized by having no axial component of electric and magnetic field respectively. The TE and TM waves are the fundamental wave types that can exist in hollow-pipe waveguides. Hollow-pipe waveguides do not support TEM waves. The ability to reduce Maxwell's equations into three set of equations, one set for each wave type, facilitates the analysis of transmission lines and waveguides. Thus this decomposition of Maxwell's equations is carried out in the first section of Part 2.

PART I WAVES ON TRANSMISSION LINES

In this section we introduce the topic of voltage and current waves on a two-conductor transmission line by using a distributed-circuit model of the transmission line. This allows us to explore a number of fundamental properties of one-dimensional waves without having to consider the electromagnetic fields in detail. The distributed-circuit-model approach has limitations and in general must be replaced by a detailed solution for the electromagnetic field associated with the guiding structure if we want to determine the distributed-circuit parameters. The field analysis of transmission lines is presented in Part 2.

3.1 WAVES ON AN IDEAL TRANSMISSION LINE

In Fig. 3.1a we show a two-conductor transmission line consisting of parallel round conductors (wires). The conductors will be assumed to be perfect, i.e., have infinite conductivity. The conductors extend from $z = -\infty$, thus forming a semiinfinite transmission line. At $z = 0$ a voltage source with internal resistance R_g is connected to the transmission line. The generator produces a voltage $V(t)$ that is impressed across the transmission line. If the generator is switched on at time $t = 0$, a current $I(t)$ will begin to flow into the upper conductor. A return current $S(t)$ will then flow on the lower conductor since current flow through the generator must be continuous. The return current is produced by the action of the electric field established between the two conductors. Since the transmission line is semiinfinite in length, there is no direct conducting path between the upper and lower conductors. However, there is a distributed capacitive

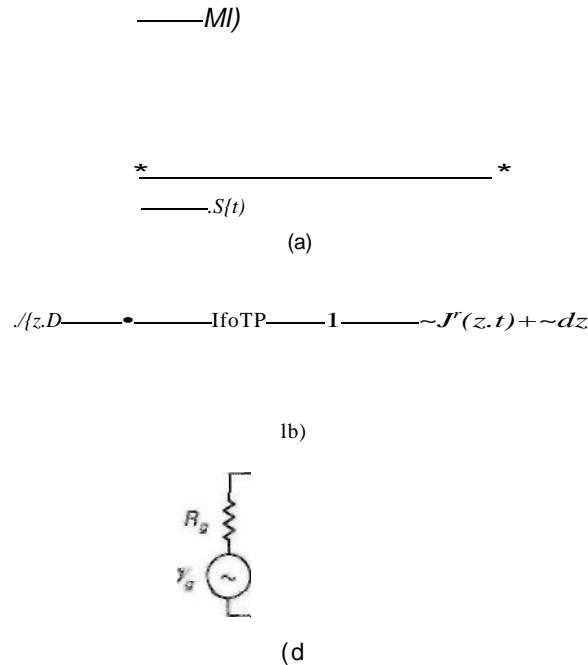


FIGURE 3.1
(a) An ideal two-conductor transmission line connected to a voltage generator; (b) equivalent circuit of a differential section of the transmission line with no loss; (c) equivalent circuit seen by the generator.

per meter between the two conductors; so we have a capacitive or displacement current flowing from the upper conductor to the lower conductor.

The electric current results in a magnetic field around the conductors and consequently the transmission line will also have a distributed series inductance L per meter. We can model a differential section dz of this transmission line as a series inductance Ldz and a shunt capacitance Cdz as shown in Fig. 3.16. If the conductors had finite conductivity, we would also need to include a series resistance in the equivalent circuit of a differential section. However, we are assuming that the conductors are perfect; so the series distributed resistance R per meter is zero.

Since electrical effects propagate with a finite velocity v (the speed of light in vacuum), it should be clear that the voltage $V(z,t)$ and current $J(z,t)$ at some arbitrary point z on the transmission line will be zero until a time z/v has elapsed after switching the generator on. We will show that the generator launches voltage and current waves on the transmission line that propagate with a finite velocity. The equations that describe these waves are established by applying Kirchhoff's circuit laws to the equivalent circuit of a differential section of the transmission line, along with a specification of the terminal relationships (boundary conditions) that must hold at the generator end.

At some arbitrary point z on the transmission line, let the voltage and current be given by $V(z,t)$, $J(z,t)$. At a differential distance dz further

along, the voltage and current have changed by small amounts (dV/dz) d_z and $(d.f/dz)d_z$; so the output voltage and current at $z + dz$ will be

$$\begin{aligned} J(z + dz, t) &= T(z, t) \frac{d}{dz} \\ \mathcal{V}(z + dz, t) &= f(z, t) + \frac{d\mathcal{V}(z, t)}{dz}; \end{aligned}$$

The sum of all potential drops around the circuit must be zero; so we have

$$\begin{aligned} -\frac{d\mathcal{V}}{dz} + L \frac{dJ}{dt} + T + \frac{BV}{dz} dz &= 0 \\ \text{or } \frac{d\mathcal{V}(z, t)}{dz} - \frac{dS(z, t)}{dt} &= -Li \quad (3.1a) \end{aligned}$$

The sum of currents flowing into the output node must also be zero; so we can write

$$\begin{aligned} J - C \frac{d\mathcal{V}}{dt} - S - \frac{dJ}{dz} dz &= 0 \\ \text{or } \frac{dS(z, t)}{dt} - \frac{dJ(z, t)}{dz} &= dT \quad (3.1b) \end{aligned}$$

These two partial differential equations describe the relationship between the voltage and current waves on the transmission line.

We can obtain an equation for the voltage $\mathcal{V}(z, t)$ by differentiating (3.1a) with respect to z and using (3.1b) to eliminate the current; thus

$$\begin{aligned} \frac{d^2\mathcal{V}(z, t)}{dz^2} &= -L \frac{d^2S}{dzdt} = -L \frac{d^2T}{dzdt} = -L \frac{dH'}{dt} \\ \text{or } \frac{d^2\mathcal{V}(z, t)}{dz^2} - LC \frac{d^2T}{dt^2} &= 0 \quad (3.2a) \end{aligned}$$

In a similar way we obtain

$$LC \frac{d^2J}{dt^2} = 0 \quad (3.2b)$$

The product LC has the dimensions of one over velocity squared. These equations are one-dimensional wave equations and describe waves propagating with a velocity⁸

$$\boxed{P-J \perp r / LC}$$

⁸For an ideal transmission line in air, $v = c = 3 \times 10^8$ m/s, the velocity of light.

Consider the equation

$$\frac{dV}{dz} = 1 - \frac{d^2V}{dz^2}$$

We can readily show that any two arbitrary functions of the form $f^+(t - z/v)$ and $f^-(t + z/v)$ are solutions of this equation. If we let $w = t - z/v$ then we have

$$\begin{aligned} \frac{df^+(t-z/v)}{dz} &= \frac{9f(w)}{dz} = \frac{df^*(w)}{dw} \frac{dw}{dz} = \frac{lT(w)}{v dw} \\ \text{and } \frac{a^*r(u)}{dz^2} &= \frac{i^*T(")}{v^2 dw^2} \end{aligned}$$

For d^2r/dt^2 we get d^2f/dw^2 . Consequently,

$$\frac{a^2f}{dz^2} - \frac{1}{v^2} \frac{d^2f^+}{dt^2} - \frac{d^2r}{dw^2} \frac{I}{v^2} \frac{1}{v^2 - v^2} = 0$$

so $f^*U \sim z/v$ is clearly a solution of the one-dimensional wave equation, as is $f^-(t + z/v)$.

The function $f^*(t - z/v)$ is the same as the function f^*it but delayed in time by an amount z/v which equals the distance z divided by the velocity of propagation v . We interpret this solution as a wave propagating in the positive z direction and identify this solution with a superscript + sign. The other solution represents a wave propagating in the $-z$ direction and is identified by the - sign as a superscript.

The general solution for the voltage waves on the transmission line is

$$T(z, t) = V^+ r f(t - z/v) + V^- r f(t + z/v) \quad (3.4)$$

where V^+ and V^- are amplitude constants. By using (3.16) we see that

$$\frac{ay}{dz} = -\frac{JirJf}{clv} \frac{1}{1 - \frac{at}{v}} \quad St$$

If we assume that S is of the form

$$S(z) = rr(t - z/v) - rr(t + z/v)$$

then

$$\frac{''}{dz} \frac{1}{v} \frac{r}{r} \frac{f}{f} \frac{1}{1 - \frac{at}{v}} \frac{1}{1 - \frac{at}{v}}$$

upon using

$$\begin{aligned} af^* &= df^* \frac{d(t + z/v)}{dt} = 1 af \\ az &= a(t + z/v) \end{aligned}$$

An examination of these equations shows that the assumed solution for

$J'i.z, t)$ is compatible with that for the voltage $p'Cztt)$ if we choose

$$\mathbf{r} = \mathbf{VCV}^{\Delta} \quad \mathbf{l} = \mathbf{UCV}^{-}$$

The parameter νC has the dimensions of an admittance and is also equal to C/L . The characteristic admittance Y_c of the transmission line is defined by this parameter. The reciprocal parameter is called the characteristic or surge impedance of the transmission line. It is given by

$$-\Delta = yr \quad (3.)$$

By using this parameter the solution for the current waves on the transmission line can be expressed in the form

$$V^- \quad i \quad z \mid \quad V^- \quad I \quad z$$

The negative sign preceding the wave with amplitude V^- indicates a reversal in the direction of current flow for the wave propagating in the $-z$ direction.

For the transmission-line circuit shown in Fig. 3.1, the generator launch voltage and current waves propagating in the $+z$ direction. Since the transmission line extends to infinity, no waves propagating in the $-z$ direction will be present. Later on we will consider a transmission line that is terminated at $z = \infty$ with either a resistance, capacitance, or a combination of these elements. Waves propagating in both the $+z$ and $-z$ directions will then exist. For the present case the voltage and current waves on the transmission line are assumed to be

$$nzj) = v + r[t - z -]$$

$$s(z,t) = rr[t - -]$$

with $V^+ = Z_c$. At the generator end $z = 0$ the terminal conditions require that

$$S(0,t) = S_e$$

where S_B is the current supplied by the generator. These terminal conditions can be expressed in the form

$$\ll Z'' r(0 + v + r < ' >$$

$$V^*$$

from which we find that

$$V_{\Delta \diamond} = Y_{\pm} R^{\Delta} \quad (3.7)$$

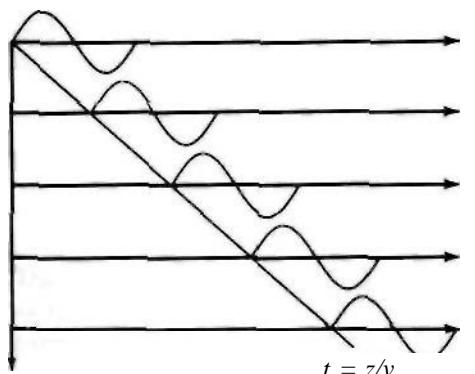
The voltage wave launched on the transmission line is thus given by

$$\Delta \Delta - Z^{\Delta} A' - \Delta) \quad (3.8a)$$

with a corresponding current wave

$$\bullet \Delta \bullet < > - V n r^{\Delta} - \mathbf{f} \quad (3.86)$$

At any point on the transmission line, the voltage waveform is the same as that produced by the generator but delayed in time and reduced in amplitude by the factor $Z_s/(R_g + Z_c)$. The voltage reduction is the usual voltage



(a)

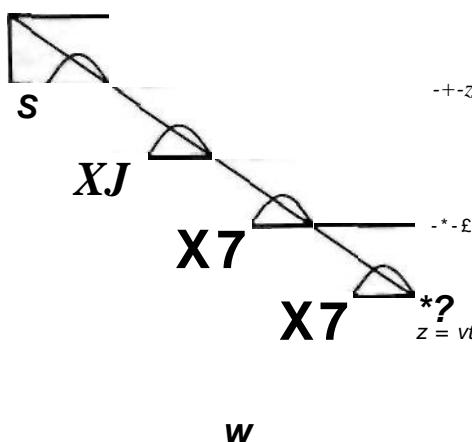


FIGURE 3.2

Time-distance and distance-time plots of voltage waveform $V(z, t)$ on a transmission line for a single-cycle sinusoidal generator voltage pulse.

division factor associated with the equivalent circuit shown in Fig. 3.2c p the infinite line the generator sees only an equivalent impedance Z_c equal to the transmission-line characteristic impedance.

In Fig. 3.2a we show a time-distance plot of the voltage waveform on transmission line for the case when the generator produces a single cycle of a sinusoidal waveform, i.e.,

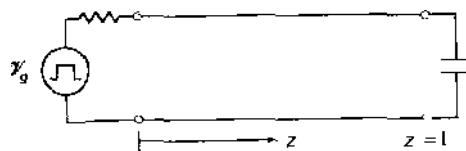
$$\gamma_g(t) = V_0 \sin t \quad 0 < t < 2\pi$$

Figure 3.2a shows $\gamma(z,t)$ as a function of t at various distances z , while Fig. 3.2b shows $\gamma''(z,t)$ as a function of z for various values of t . In the latter plot note that the leading edge of the waveform is the initial voltage produced by the generator at time $t = 0$ and hence the waveform appears reversed when plotted as a function of z .

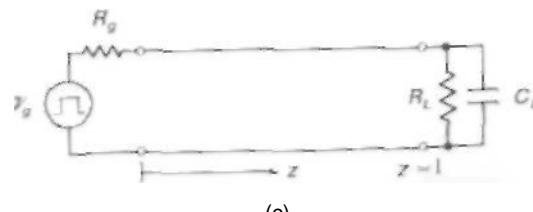
3.2 TERMINATED TRANSMISSION LINE: RESISTIVE LOAD

In Fig. 3.3 we show a transmission line terminated at a distance l from generator in a load resistance R_L . At the load end the terminal conditions are

$$\begin{aligned} \gamma(M) &= n \gamma \# i & (3.9) \\ s(i,t) &= .y_L & (3.8) \end{aligned}$$



(b)



(c)

FIGURE 3.3
The terminated transmission

If we choose R_L equal to the characteristic impedance Z_L , then $2A = J'_L R_i = J' i Z_c$. For a wave propagating in the $+z$ direction, $A(z,t) = Z_c J(z,t)$ so that at $z = l$, $T(l,t) = Z^A(l,t)$, which satisfies the load terminal condition. Thus by choosing $R_L = Z_c$ the forward propagating wave will be completely absorbed by the load resistor and no reflected wave will be generated at the load end. Thus, in order to avoid a reflected wave, such as a reflected pulse in a digital circuit application, the transmission line should be terminated in its characteristic impedance.

When $R_L \neq Z_c$ the terminal conditions at the load end cannot be satisfied without introducing a reflected wave. The incident wave at $z = M_s$ given by

where V^+ is the amplitude of ψ_j relative to ψ_g . In order for a reflected wave to combine with the incident wave so as to satisfy the terminal conditions (3.9), the reflected wave must have the same time dependence as that of the incident wave. Hence the form of the reflected wave will be

/ z 21

The argument must contain the factor $t + z/v$ plus additional delay factors, so that at $z = I$ the reflected wave has the form $T_g(t - l/v)$. The reflected current wave is given by

1

At the load end the total current on the transmission line must equal the current J'_L flowing through R_L ; thus

$$\frac{1}{-vr-TyrJt--} = S,$$

The total *voltage* on the transmission line must equal the load voltage; so

$$(V_+ + V_-) V_g l t - \backslash = V_L = f_L R_b$$

When we divide this equation by the first one, we obtain

$$V^* + V^- = R_L$$

which gives

$$V$$

$$RL + Z.$$

(3.1)

The parameter T_L is called the load voltage reflection coefficient. amplitude $V\sim$ is that of the reflected voltage wave and $V\sim$ is the amplif* of the incident voltage wave. The ratio is determined by the *conditions the load end only*.

Once a reflected wave has been launched from the load terminatio the total voltage on the transmission line will consist of the incident volt wave plus the reflected voltage wave until the time at which the refle wave reaches the generator end. If the generator internal impedance R equals the characteristic impedance Z_c , then the reflected wave is absor" at the generator end. If $R_g \neq Z_c$ then the reflected wave is reflected at generator end to produce another forward propagating wave. For the reflected wave at the generator end, the terminal conditions are obtained short-circuiting the voltage generator. Thus the reflected wave sees a termi nation R and will be reflected with a reflection coefficient T given by

$$! \quad \frac{R_e - Z_c}{R_a + Z_c} \quad (3.11)$$

As long as the generator continues to produce a voltage $2^{\wedge}(0)$, it continu to launch a first forward propagating wave with voltage $\wedge(0, t)$ $Z_c T_g(t)/(R_g + Z_c)$ and with current $J^{\wedge}(0, t) = S_g$. Thus the superposi of a reflected wave at the generator end requires the launching of a seco forward propagating wave with a voltage amplitude that cancels that of reflected wave at the generator terminals, i.e., the generator is treated being short-circuited. The second forward propagating wave will also dergo reflection at the load termination, so that as time proceeds we will e up with a multitude of forward propagating and reflected waves on transmission line. This collection of waves can be described as follows:

$$\begin{aligned} v|s, t) &= v + v_t(t - \wedge + rjr^{\wedge}lt + \wedge -) \\ &+ r_g r_L V \cdot \cdot \cdot ; \quad t - \frac{2I}{v} U|t - \frac{2I}{v} \\ &+ r_g ilv^+ T_{-e} t + \frac{2 - 4I}{v} U|t - \frac{3I}{v} \\ &+ r^* riv + y_t t - \frac{z + 4I}{U} U t - \frac{4I}{v} + \end{aligned} \quad (3.1)$$

where $V^+ = Z_c/(R_g + Z_c)$ and $U(t - a)$ is the unit step function wb» equals zero for $t < a$ and equals unity for $t \geq a$. The unit step function &

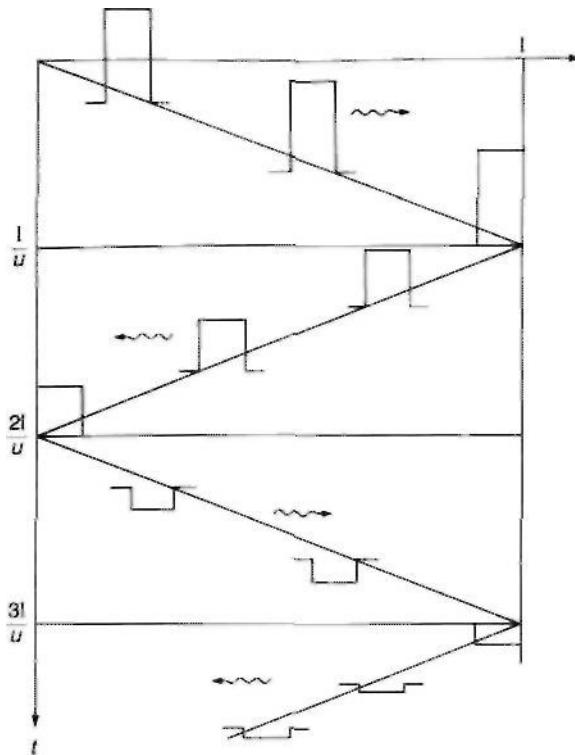


FIGURE 3.4
Distance-time plot of a pulse undergoing multiple reflection on a transmission line when $r_L = -T_e - 0.5$. Reflection at the generator end causes a reversal in the polarity of the pulse.

convenient function to use to specify when a waveform begins. In the case of multiple reflected waves on a transmission line, each reflected wave begins after time delays of l/u , $2l/v$, $3l/v$, etc., corresponding to the time delay to propagate a certain number of times back and forth between the generator end and load end.^t The current wave can be obtained by multiplying the forward propagating waves by Y_c and the backward propagating waves by $-Y_g$. When the generator voltage 2^0 exists for only a finite time interval, the total voltage wave on the line will decay toward zero since each successive reflected wave is multiplied by a reflection coefficient, either T_e or r_L , which is less than one in magnitude, and hence successive waves are of diminishing amplitude.

The sequence of multiple reflected waves can be illustrated in a distance-time plot. In Fig. 3.4 we show this type of plot for a generator producing a rectangular pulse. We have chosen $R_L = 3Z_c$ and $R_g = Z_c/3$ so that $Y_L = 0.5$, $Y_g = -0.5$. When the reflection coefficient is negative, the

^tThe unit step functions were introduced for clarity in describing the physical process but are actually not required in (3.12) since $2^{*(t-T)} = 0$ for $t < T$.

sign of the reflected voltage wave is reversed and this is illustrated in p-3.4. In high-speed digital circuits using interconnecting transmission U_n multiple reflected pulses are undesirable and can be avoided by terminal' each transmission line in a resistance equal to its characteristic impedan

3.3 CAPACITVE TERMINATION

When the transmission line is terminated in a reactive element such as capacitor as shown in Fig. 3.36, the reflected wave will have a waveform different from that of the incident wave. The solution for the reflected wave is readily found from the condition that the sum of the incident plus reflected voltage wave at $z = l$ must satisfy the terminal conditions. The incident wave is again chosen to be $\mathcal{V}_i(t, z) = V^+ \mathcal{Z}_S U - l/v$. The reflected wave is initiated at time $t_1 = l/v$ and will propagate from the point $z = l$ toward the generator. Therefore it is of the form

$$\mathcal{V}_r\left(t - \frac{l}{v} + \frac{z-l}{v}\right) = \mathcal{V}_r\left(t + \frac{z-2l}{v}\right)$$

For a capacitor we require

$$SL = c \cdot \sim dT = Cl \frac{dni, t}{dt} = dW_r(t-l/v) \frac{dt}{dt}$$

where $\mathcal{V}_r(t - l/v)$ is the reflected voltage wave at $z = l$. In addition, we the condition

$$SL = Y_C \quad t - \sim$$

From these two equations we obtain

$$d^2t-l/v) \quad 1 \quad / \quad /$$

$$dt \quad \mathbf{c}_L \mathbf{Z}_C \quad \bullet \\ = -\mathbf{V}^+ \frac{dVJt-l/v}{dt} \quad + \frac{V^*}{\mathbf{c}_L \mathbf{Z}_C} \frac{''K't}{---} \quad (3.13)$$

For a specific example we will consider the case when the generator produces a rectangular pulse given by

$$y_g(t) = 1 \quad 0 < t < T$$

The right-hand side of (3.13) will now become the source function

$$\frac{V^+}{CZ_C} \frac{1}{t - \frac{-\Delta t}{CZ_C}} T = \frac{t - \frac{-\Delta t}{CZ_C}}{t + \frac{J}{CZ_C} + \frac{slt}{CZ_C}} T$$

where $\delta(t - a)$ is the Dirac delta function or impulse function that arises from the derivative of the rectangular pulse. We can integrate (3.13) by introducing the integrating factor $e^{i/r}$ where $T = CZ_C$. We note that

$$\frac{d}{dt^K - r} = \frac{I dV_r}{dt} = \frac{1}{T}$$

so consequently

$$\begin{aligned} \int'_{l/v}^{-7/v} \frac{-\gamma/v e^{i/T} dt}{dt} &= \left[\frac{TU}{v} - \frac{1}{v} \right] e^{i/r} - 7^\wedge(0) e^{i/L} \\ &= -f(t - \frac{-\Delta t}{CZ_C}) - \frac{\Delta t}{CZ_C} T - \frac{1}{v} dt \\ &= -V f_{l/v} \frac{S|t - \frac{-\Delta t}{CZ_C}|}{T} - \frac{1}{v} dt \\ &\quad \begin{array}{c} I \\ - \end{array} \frac{I}{v} < t < \frac{l}{v} \\ &\quad V^+ (e^{i/T} - 2e^{i/UT}) \\ &\quad V^+ (-2e^{i/UT} + 2e^{(t+vT)/vT}) \quad t > - + T \end{aligned}$$

Since we have included the impulse functions as derivatives of the applied rectangular pulse, the lower limit of integration is regarded to be just before $t = l/v$ and thus $2^\wedge(0)$ is equal to zero since it corresponds to $2^\wedge(0 -)$. Hence we obtain

$$\begin{aligned} H^* r &= \frac{V^+ (1 - 2e^{-i(l-vT)/vT})}{V^+ (-2e^{i/UT} + 2e^{(t+vT)/vT}) e^{-i/r}} \quad \begin{array}{c} I \\ - \end{array} \frac{l}{v} < t < \frac{l}{v} + \frac{m}{T} \\ &\quad \begin{array}{c} I \\ - \end{array} t > - + T \quad v \end{aligned} \quad (3.14)$$

At $t = l/v$ the reflected wave has an amplitude equal to $-V^+$ which cancels that of the incident pulse. This is consistent with the requirement that initially the capacitor C_L is uncharged and must have a zero voltage across it. The capacitor charges to a final voltage level

$$\begin{aligned} K(T) &= 2V^+ (1 - e^{-i/T}) \\ &= 2V^+ (1 - e^{-i/T}) \end{aligned}$$

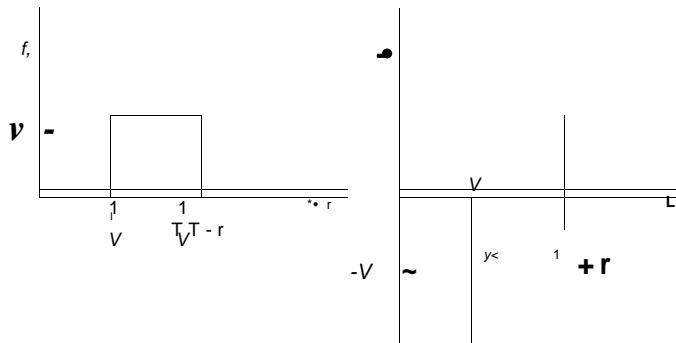


FIGURE 3.5
Incident and reflected voltage waveforms at $z = L$.

at $t = l/v + T$ and then discharges toward zero. The apparent discontinuous change that occurs in the reflected voltage wave at $t = l/v + T$ is caused by the sudden drop to zero volts for the incident pulse, and in order to match the voltage across the capacitor, the reflected voltage wave must have a positive jump of value V^* . The incident and reflected voltage waveforms at $z = L$ are shown in Fig. 3.5. The reflected voltage waveform will propagate toward the generator and will begin to initiate a new forward propagating wave at time $t = 2l/v$. Clearly the capacitor has made a significant change in the waveform of the reflected wave.

The analysis for the case of a capacitor-resistor termination as shown in Fig. 3.3c is similar. The terminal conditions are

$$\hat{y} = \hat{G}_L \frac{d\hat{y}_r}{dt} \quad t \sim$$

$$t \sim$$

so in place of (3.13) we have

$$\frac{dK}{dt} = \frac{i}{I} - \frac{1}{Z_c} = -\frac{dy}{dt} \quad C_L Z_c \quad r, \quad (3.15)$$

The solution is similar to that for (3.13) except that the charging time constant is now $T_X = C_L R_L Z_c / (R_L + Z_c)$. Initially, V_r has a value equal to $-V^*$ as before. In this case the capacitor charges toward a final voltage equal to

$$V^+ \mid l = \frac{R_L Z_c}{R_L + Z_c} \cdot \frac{2R_L}{RL} \cdot \mathbf{F}^*$$

determined by the steady-state voltage across R_L if C_L was absent. A'

$t = l/u + T$ the capacitor voltage will be

The reflected-wave voltage at this time will be $V_c \sim V^*$. When t becomes greater than $l/v + T$ the incident voltage wave pulse drops to zero volts so the reflected-wave voltage jumps to a value equal to V_c and will then decay toward zero with a time constant T_c .

STEADY-STATE SINUSOIDAL WAVES

When the generator produces a sinusoidal voltage $\mathcal{V}_g(t) = V \cos \omega t$, the steady-state voltage waves on the transmission line will be of the form $\cos \alpha(t - z/v)$ and $\cos \phi(t + z/v)$. The steady state is achieved, for all practical purposes, after a few multiple reflections have occurred, since the amplitude of the successive reflected wave decreases quite rapidly because it is multiplied by V_g or Y_L upon each reflection. The solution for steady-state sinusoidal waves is most conveniently obtained using phasor analysis. The generator voltage is represented by $V_g e^{j\omega t}$. The voltage and current waves on the transmission line will then also have an $e^{j\omega t}$ time dependence. The differential equations (3.1a) and (3.16) now become (the common time factor $e^{j\omega t}$ is dropped)

$$\frac{dV(z)}{dz} = -jwLI(z) \quad (3.16a)$$

$$\frac{dl(z)}{dz} = -jioCV(z) \quad (3.16b)$$

where $V(z)$ and $I(z)$ are complex phasor amplitudes. By eliminating the current we find that $V(z)$ satisfies the equation

$$\frac{d^2V(z)}{dz^2} + \frac{u^2}{c^2} = 0 \quad (3.16c)$$

The solution for $V(z)$ is of the form

$$V(z) = V^+ e^{-j\beta z} + V^- e^{j\beta z} \quad (3.17a)$$

with a corresponding solution

$$I(z) = I^+ e^{-j\beta z} - I^- e^{j\beta z} \quad (3.17b)$$

for the current waves. The constant $\beta = io/v$ is the propagation phase constant. As before the current amplitudes are related to the voltage amplitudes through the characteristic impedance of the line, i.e.,

$$I^+ = Y_c V^+ \quad I^- = Y_c V^-$$

When the time factor is restored, it is readily seen that e'

corresponds to a wave propagating in the $+z$ direction, while $e^{jiz+Ju''} j_s$ wave propagating in the $-z$ direction. In the next section we will show that for a transmission line with finite conducting wires and possibly also surrounded with lossy dielectric materials, the waves attenuate in amplitude as they propagate. For this case the wave solutions are of the form

$$V = V_+ e^{-Jf^* - \alpha z} + V_- e^{Jf^* + \alpha z} \quad (3.180)$$

$$I = \mathbf{r}^* \cdot \mathbf{i}^* - \mathbf{f}^* \cdot \mathbf{h}^* \quad (3.186)$$

where α is the attenuation constant.

3.5 WAVES ON A LOSSY TRANSMISSION LINE

Conductors used in a transmission line will always have a finite conductivity and will therefore exhibit some series resistance. Furthermore, because of the skin effect the current flows in a thin layer at the surface of the conductor, the effective thickness of the layer being equal to the skin depth δ_S given by $(2/\omega i\mu)^{1/2}$ [see (2.104)]. Consequently, the series resistance increases with an increase in the frequency of operation. In order to account for this resistance, a distributed series resistance R per meter must be included in the distributed circuit used to model the transmission line.

The two conductors in a transmission line are usually maintained parallel to each other by supporting them in a dielectric structure. For example, a coaxial transmission line is filled with a dielectric medium in order to keep the center conductor coaxial with the outer shield. Dielectric materials usually have a negligible conductance but do have a small amount of dielectric loss due to polarization loss in the dielectric. Consequently, a shunt conductance G per meter is added to the distributed circuit to account for this loss. Thus, for a lossy transmission line, the equivalent circuit of a differential length dz is chosen to be that shown in Fig. 3.6.

If the voltage and current at the input are $T(z, t)$, $H(z, t)$ and if the voltage and current at the output are

$$V_+ - \frac{dV}{dz} dz \quad I_+ + \frac{dI}{dz} dz$$

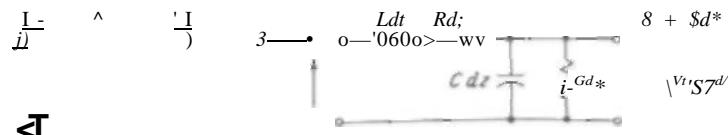


FIGURE 3.6
Equivalent circuit of a differential length of transmission line.

then Kirchhoff's laws give

$$\text{yr} \cdot \left\langle \frac{\partial}{\partial z} r + \frac{\partial^2 \gamma'}{\partial z^2} \right\rangle J = JR dz + L dz \frac{dS}{dt}$$

or $\frac{\partial \gamma'}{\partial z} = -J^r R - L \frac{dS}{dt}$ (3.19a)

Similarly,

$$S - \left\langle S + \frac{dS}{dz} dz \right\rangle = \gamma' G dz + C dz \frac{\partial \gamma'}{\partial t}$$

or $\frac{\partial \gamma'}{\partial z} = -rG - C - \frac{\partial \gamma'}{\partial t}$ (3.196)

The first equation states that the potential difference between the input and output is equal to the potential drop across R and L . The second equation states that the output current is less than the input current by an amount equal to the shunt current flowing through C and G . Differentiating (3.19a) with respect to z and (3.196) with respect to time t gives

$$\frac{d^2 \gamma'}{dz^2} = -R \frac{dS}{dz} - L \frac{d^2 \gamma'}{dt^2} \quad (3.20a)$$

$$\frac{d^2 \gamma'}{dt^2} = -G \frac{d^2 \gamma'}{dz^2} - C \frac{d^2 T}{dt^2} \quad (3.206)$$

Using (3.196) and (3.206) in (3.20a) now gives the following equation for the line voltage γ' :

$$\frac{d^2 \gamma''}{dz^2} = \frac{I}{R \cdot G \cdot r + C \cdot \pi^2} \frac{d^2 \gamma'}{dt^2} + L \frac{I}{dt} \frac{d \gamma'}{dt} + C \frac{d^2 T}{dt^2}$$

or $\frac{d^2 T}{dt^2} = -(RC + LG) \frac{d^2 \gamma'}{dz^2} - LC \frac{d \gamma'}{dt} - RG \gamma' = 0$ (3.21)

The current γ satisfies this one-dimensional wave equation also. If a solution in the form of a propagating wave

$$\gamma = \text{Re}(V e^{-\gamma z + i \omega t})$$

is assumed, substitution into (3.21) shows that the propagation constant γ must be a solution of

$$y^2 - j(u(RC + LG) + w^2 LC - RG) = 0 \quad (3.22)$$

If only the steady-state sinusoidally time-varying solution is desired, phasor notation may be used. If we let V and I represent the voltage and current without the time dependence $e^{j\omega t}$, the basic equations (3.19) may be

written as

$$\frac{dV}{dz} = -(R+j\omega L)I \quad (3.23a)$$

$$\frac{dI}{dl} = -(G+j\omega C)V \quad (3.23b)$$

The wave equation (3.21) becomes

$$-\nabla^2(RG-a^2LC)V-j\omega(RC+LG)V=0 \quad (3.24)$$

The general solution to (3.24) is

$$V = V_+ e^{-y_+ z} + V_- e^{-y_- z} \quad (3.25)$$

where $y = a + j\omega t$ is given by

$$y = \sqrt{-RG + j\omega L} \quad (3.26)$$

from (3.22). The constants V_+ and V_- are arbitrary amplitude constants for waves propagating in the $+z$ and $-z$ directions, respectively. The solution for the current I may be found from (3.23a), that is

$$I = f^* \langle r^{y_+} - / - e^{y_-} \rangle^2 = \dots \quad (3.27)$$

The parameter

$$Z = \frac{R+j\omega L}{j\omega C} = \frac{\sqrt{(R+j\omega L)^2}}{\omega C} \quad (3.28)$$

is the characteristic impedance of the line since it is equal to the ratio V^+/I^+ and V_-/I_- . Note that $y = \sqrt{(R+j\omega L)(G+j\omega C)}$.

Loss-Free Transmission Line

For a line without loss, i.e., for which $R = G = 0$, the propagation constant is

$$y = \beta = \sqrt{\omega L/C} \quad (3.29)$$

and the characteristic impedance is pure real and given by

$$Z_c = \sqrt{\omega L/C} \quad (3.30)$$

According to the field analysis, β is also equal to $\sqrt{\omega/C}$, and hence

$$LC = M_e \quad (3.31)$$

for a transmission line. This result may also be verified from the solution for L and C , as shown later in the section on transmission-line parameters

Using (3.31) in (3.30) shows that the characteristic impedance is also given by

$$C'' V C^2 = CV c \quad (3.32)$$

where Z is the intrinsic impedance of the medium. The characteristic impedance differs from the intrinsic impedance Z by a factor e/C , which is a function of the line configuration only.

Low-Loss Transmission Line

For most microwave transmission lines the losses are very small; that is, $R \ll coL$ and $G \ll toC$. When this is the case, the term RG in the expression (3.26) for y may be neglected. A binomial expansion then gives

$$y \approx j*JW + \frac{1}{W} - \frac{R}{W} + \frac{G}{W} = a + jp \quad (3.33)$$

To first order the characteristic impedance is still given by (3.30) or (3.32). Thus the phase constant for a low-loss line is

$$\beta = wvX_C \quad (3.34a)$$

and the attenuation constant a is

$$a = \sqrt{(R^2 + G^2)/W} \quad (3.34b)$$

where $Y_c = Z_c^{-1} = |C/L|$ is the characteristic admittance of the transmission line.

3.6 TERMINATED TRANSMISSION LINE: SINUSOIDAL WAVES

In this section the properties of a transmission line terminated in an arbitrary load impedance Z_L are examined. This will serve to illustrate how the forward and backward propagating waves can be combined to satisfy the boundary conditions at a termination. Figure 3.7 illustrates schematically a transmission line terminated in a load impedance Z_L . The line is assumed lossless and with a characteristic impedance Z_c and a propagation constant

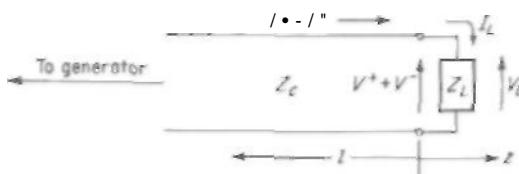


FIGURE 3.7
Terminated transmission line.

$y = 7/3$. It should be noted that at microwave frequencies conventional low-frequency resistors, inductors, or capacitors, when connected across the two conductors of a transmission line, may behave as impedance elements with quite different characteristics from the low-frequency behavior.

If a voltage wave $V^+ e^{j\beta z}$ with an associated current $I^+ e^{j\beta z}$ is incident on the termination, a reflected voltage wave $V^- e^{j\beta z}$ with a current $-I^- e^{j\beta z}$ will, in general, be created. The ratio of the reflected and incident wave amplitudes is determined by the load impedance only. At the load the total line voltage must equal the impressed voltage across the load and the line current must be continuous through the load. Hence, if Z_L is located at $z = 0$,

$$V = V_+ + V_- = V_L \quad (3.35a)$$

$$i = r - r = h \quad (3.35b)$$

But $r = F_C V^T$, $r = Y_C V$, and $V_L/I_L = Z_L$ by definition of load impedance. Therefore

$$V^+ + V^- = V_L \quad (3.36a)$$

$$V^+ - V^- = -V_L \quad (3.36b)$$

The ratio of V to V^* is usually described by a voltage reflection coefficient T defined as

$$T = \frac{V^+ - V^-}{V^+ + V^-} \quad (3.37)$$

In place of (3.36) we may write

Dividing one equation by the other yields

$$\frac{1 + r_L}{r_L - r_C} = T \quad (3.38)$$

The quantity Z_L/Z_C is called the normalized load impedance (load impedance measured in units of Z_C), and $(1 + T_L)/(1 - V_L)$ is then the normalized input impedance seen looking toward the load at $z = 0$. The normalized load impedance will be expressed as \bar{Z}_L , with the bar on top signifying normalized impedance in general. Solving for the voltage reflection coefficient T gives

$$\frac{\bar{Z}_L - Z_C}{Z_L + Z_C} = \frac{Z_L/Z_C - 1}{Z_L/Z_C + 1} = \frac{\bar{Z}_L - 1}{\bar{Z}_L + 1} \quad (3.39)$$

Analogous to a voltage reflection coefficient, a current reflection coefficient r , could also be introduced. In the present case

$$r = \frac{-r}{Y_L} = \frac{Y_r V^-}{Y V^+} = -r_L$$

In this text, however, only the voltage reflection coefficient will be used; so the adjective "voltage" can be dropped without confusion.

The incident voltage wave can be considered as transmitting a voltage V_- across the load, and a voltage transmission coefficient T can be defined as giving V_L in terms of V^+ ; thus

$$V_L = TV^+ = (1 + r_L)V^+$$

$$\text{So } P = 1 + T_x \quad (3.40)$$

A corresponding current transmission coefficient is not used in this book.

Returning to (3.39), it is seen that if $Z_L = Z_c$, the reflection coefficient is zero. In this case all the power in the incident wave is transmitted to the load and none of it is reflected back toward the generator. The power delivered to the load in this case is

$$P = i \operatorname{Re}(V^+ V^*) = i |V^+|^2 F_c = i |y_j|^2 y_L \quad (3.41)$$

The load is said to be matched to the transmission line when $y_L = 0$.

If Z_L does not equal Z_c , the load is mismatched to the line and a reflected wave is produced. The power delivered to the load is now given by

$$\begin{aligned} P &= |\operatorname{Re}(V^+ V^-)| = |\operatorname{Re}[(V^+ + V')(V^+ - V')^*]| \\ &= i \operatorname{Re}[y_c(V^+ + V^-)(V^+ - V^-)^*] \\ &= |\operatorname{Re}[Y_c |V^+|^2 (1 + r)(1 - r)^*]| \\ &= |y_c i V^+|^2 (i - ir_j^2) \end{aligned} \quad (3.42)$$

The final result states the physically obvious result that the power delivered to the load is the incident power minus that reflected from the load.

In the absence of reflection, the magnitude of the voltage along the line is a constant equal to $|V^+|$. When a reflected wave also exists, the incident and reflected waves interfere to produce a standing-wave pattern along the line. The voltage at any point on the line ($z < 0$) is given by

$$v = v_+ e^{-rf}, \quad +Y_L V^+ e^{jlkz}$$

and has a magnitude given by

$$|v| = |V^+| \sqrt{1 + r^2} e^{-2rf} = |V^+| \sqrt{1 + r^2} e^{-2rf}$$

where $f = -z$ is the positive distance measured from the load toward the



generator, as in Fig. 3.7. Let Y_L be equal to $pe^{j\theta}$, where $p = |r_j|$; then

$$\begin{aligned}
 IVI &= |V^+| |1 + pe^{j(\phi - kx)}|^2 \\
 &= |V^*| [(1 + p \cos(\phi - 2pl))^2 + p^2 \sin^2(\phi - 2pl)]^{1/2} \\
 &= |V^+| \{(1 + p \cos(\phi - 2pl) - p \sin(\phi - 2pl))\}^{1/2} \\
 &= |V^+| (1 + p)^2 - 4p \sin^2(\phi - 2pl) \quad (3.43)
 \end{aligned}$$

This result shows that IVI oscillates back and forth between maximum values of $|V^+|(1 + p)$ when $\phi - 6\pi/2 = mr$ and minimum values $|V^+|(1 - p)$ when $\phi - 6\pi/2 = mr + TT/2$, where n is an integer. These results also agree with physical intuition since they state that voltage maxima occur when the incident and reflected waves add in phase and that voltage minima occur when they add 180° out of phase. Successive maxima and minima are spaced a distance $d = ir/f_i = XTT/2TT = A/2$ apart, where A is the wavelength for TEM waves in the medium surrounding the conductors. The distance between a maximum and the nearest minimum is $A/4$.

Since the current reflection coefficient is equal to $-r_L$ the current waves subtract whenever the voltage waves add up in phase. Hence current maxima and minima are displaced $A/4$ from the corresponding voltage maxima and minima. Figure 3.8 illustrates the voltage and current standing-wave patterns that result when Z_L is a pure resistance equal to $3Z_C$.

The ratio of the maximum line voltage to the minimum line voltage is called the voltage standing-wave ratio S ; thus

$$\frac{|V^+|(1+p)}{|V^+|(1-p)} = \frac{1+p}{1-p} \quad (3.44)$$

This is a parameter of considerable importance in practice for the following reasons: At microwave frequencies instruments for the direct absolute measurement of voltage or current are difficult to construct and use. On the other hand, devices to measure relative voltage or current (or electric or magnetic field) amplitudes are easy to construct. A typical device is a small probe inserted into the region of the electric field around a line. The output of the probe is connected to a crystal rectifier, and produces an output¹ current which is a measure of the relative electric field or voltage at the probe position. By moving the probe along the line, the standing-wave ratio can be measured directly in terms of the maximum and minimum probe currents.

¹The symbol p denotes both charge density and the modulus of the reflection coefficient. The context makes it clear which quantity is under discussion; so confusion should not occur.

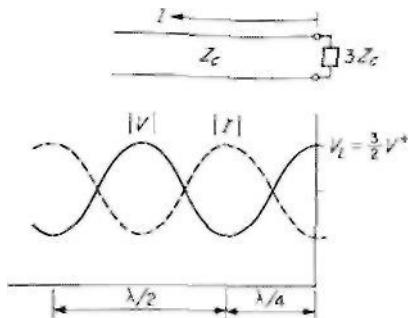


FIGURE 3.8

Voltage and current standing-wave patterns on a line terminated in a load impedance equal to $3Z_c$.

currents. The location of a voltage minimum can also be measured, and this permits the phase angle δ of V_L to be calculated. Since p is known from the measured value of S , V_L is specified, and the normalized load impedance may be calculated from (3.38).

Although the reflection coefficient was introduced as a measure of the ratio of reflected- to incident-wave amplitudes at the load, the definition may be extended to give the corresponding voltage ratio at any point on the line. Thus, at $z = -l$, the reflection coefficient is

$$!'/() = \frac{V e^{-jlt}}{V^+ e} = \frac{y}{p-2jei} = r p-W \quad (3.45)$$

where $r_L = V/V''$ denotes the reflection coefficient of the load. The normalized impedance, seen looking toward the load, at $z \sim -l$, is

$$\begin{aligned} Z_{ss} &= \frac{V}{V^* e^{ipl} + V^- e^{-Jli}} \\ &= \frac{i + i'(0)}{i - r(0)} \frac{1 + V_L e^{-2jp}}{i - r, e^{-a - \ll}} \end{aligned} \quad (3.46)$$

By replacing Y_L by $(Z_L - Z_e)/(Z_L + Z_e)$ and $e^{\pm Jpl}$ by $\cos pi \pm j \sin pi$, this result may be expressed as

$$Z_{ss} = \frac{Z_s^*}{Z_r} \frac{Z_L + jZ_e \tan pi}{Z_e + jZ_L \tan pi} \quad (3.47)$$

A similar result holds for the normalized input admittance; so

$$Y_{ss} = \frac{Y_L + jY_c \tan pi}{Y_c + jY_L \tan pi} \frac{Y_L + j \tan pi}{1 + jY_L \tan pi} \quad (3.48)$$

Of particular interest are two special cases, namely, $pi = TT$ or $I = A/2$ and

$pi = -rr/2$ or $I = A/4$, for which

$$/ = 2 >$$

(3.49 «)

$$zji = 4 \quad = Z, \quad (3.496,$$

The first is equivalent to an ideal one-to-one impedance transform[^] whereas in the second case the impedance has been inverted with resDerf to Z_c .

The terminal conditions at the generator end are readily established by using (3.47) to evaluate the input impedance Z_{in} seen looking toward the load at the generator end. If the generator with open-circuit voltage V has an internal impedance Z_g , then by the usual voltage division formula the total transmission-line voltage V at $z = -I$ will be given by

ft.

$$Z_{in} + Z_c$$

But V is the sum of the forward-propagating-wave and reflected-wave voltages at $z = -I$, that is,

$$\begin{aligned} V &= V V'' + V e^{-jel} \\ &= V^+ e^{'''}(l + T_L e^{-2jf''}) \end{aligned}$$

By using this expression we can solve for V^+ which is found to be

$$(Z_{in} + Z_g)(e^+ + Y_L e^- M)$$

$$\begin{aligned} &\underline{Z_{in}(Z_L + Z_c)V_g} \\ &2(Z_{in} + Z_g)(Z_L \cos pi + jZ_c \sin pi) \end{aligned} \quad (3.50)$$

Terminated Lossy Line

In the case of a lossy line with propagation constant $y \sim JP^{+a}$ previous equations hold except that jp must be replaced by $jp + a < wne$. a is usually so small that, for the short lengths of line used in experimental setups, the neglect of a is justified. Nevertheless, it is of so interest to examine the behavior of a lossy transmission line terminated in load Z_L . One simplifying assumption will be made, and this is that characteristic impedance Z_c can still be considered real. This assumption certainly valid for low-loss lines of the type used at microwave frequent

A detailed calculation justifying this assumption for a typical case is called for in Prob. 3.18.

Clearly, the presence of an attenuation constant a does not affect the definition of the voltage reflection coefficient T_L for the load. However, at any other point a distance l toward the generator, the reflection coefficient is now given by

$$f(l) = Y_L e^{-*W \cdot l} \quad (3.51)$$

As l is increased, T decreases exponentially until, for large l , it essentially vanishes. Thus, whenever a load Z_L is viewed through a long section of lossy line, it appears to be matched to the line since Y is negligible at the point considered. This effect may also be seen from the expression for the input impedance, namely,

$$\begin{aligned} 1 + r_L e^{-*W \cdot l} &= Z_L + Z_c \tanh(jpl + al) \\ &+ Z_L t \operatorname{tanh}(jpl + al) \end{aligned} \quad (3.52)$$

which approaches Z_c for l large since $\tanh x$ approaches 1 for x large and not a pure imaginary quantity.

The losses also have the effect of reducing the standing-wave ratio S toward unity as the point of observation is moved away from the load toward the generator. As the generator is approached, the incident-wave amplitude increases exponentially whereas the reflected-wave amplitude decreases exponentially. The result is a standing-wave pattern of the type illustrated in Fig. 3.9. For illustrative purposes a relatively large value of a has been assumed here.

The power delivered to the load is given by

$$P_L = i \operatorname{Re}(VJl) = |V_L|^2 G_L = -|V^+|^2 (1 - |r_L|^2) \quad (3.53)$$

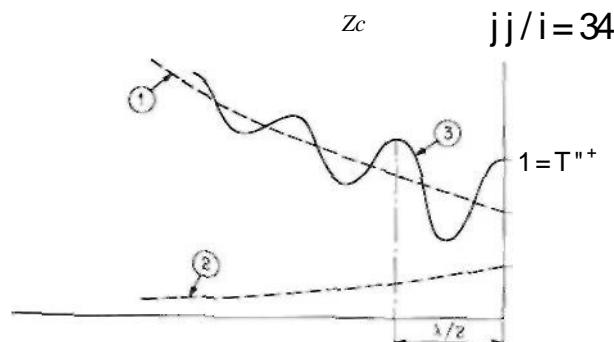


FIGURE 3.9
Voltage standing-wave pattern on a lossy transmission line.
(1) Envelope of incident-wave amplitude;
(2) envelope of reflected-wave amplitude;
(3) standing-wave pattern.

as before. At some point $z = -l$, the power directed toward the load i_8

$$\begin{aligned}
 P(l) &= \pm \operatorname{Re}(V7^*) = \wedge |V|^2 Y_C = \wedge |v^+ e''|^2 [1 - |r(/)|^2] \\
 &= \sim W - \wedge (e^{2\pi l} \sim |r_L|^2)
 \end{aligned} \tag{3.54}$$

where irk'' has been replaced by **ITJ**. Of the power given by (3.54) only, that portion corresponding to P , as given by (3.53) is delivered to the load. The remainder is dissipated in the lossy line, this remainder being given by

$$\wedge O - \wedge - y 1 \wedge 1 V' - i) \tag{3.55}$$

PART 2 FIELD ANALYSIS OF TRANSMISSION LINES

The first section of this part will show that Maxwell's equations can be reformulated so as to describe three classes of waves, TEM, TE, and TM waves. The TEM wave is the principal wave on transmission lines. From the solution for the electric and magnetic fields for the TEM wave, we will be able to establish that there are unique voltage and current waves associated with the TEM wave. We will also be able to evaluate the distributed-circuit parameters R , L , C , and G for a transmission line. The field analysis thus provides a theoretical basis for treating the transmission line as a distributed circuit.

After the basic equations for TEM, TE, and TM waves have been derived, we present the field analysis for transmission lines that support TEM waves. This is followed by several sections dealing with planar transmission lines. Many of the planar lines that we examine support only quasi-TEM waves but can be analyzed as transmission lines once their equivalent distributed-circuit parameters have been determined.

3.7 CLASSIFICATION OF WAVE SOLUTIONS

The transmission lines and waveguides analyzed in this chapter are characterized by having axial uniformity. Their cross-sectional shape &* electrical properties do not vary along the axis, which is chosen as ^{the} axis. Since sources are not considered, the electric and magnetic fields

solutions of the homogeneous vector Helmholtz equation, i.e.,

$$\nabla \cdot \mathbf{E} + k\% \mathbf{E} = 0 \quad \nabla^2 \mathbf{H} + k^* \mathbf{H} = 0$$

The type of solution sought is that corresponding to a wave that propagates along the z axis. Since the Helmholtz equation is separable, it is possible to find solutions of the form $f(z)g(x,y)$, where f is a function of z only and g is a function of x and y or other suitable transverse coordinates only. The second derivative with respect to z enters into the wave equation in a manner similar to the second derivative with respect to time. By analogy with e^{-jpt} as the time dependence, the z dependence can be assumed to be e^{-jpkz} . This assumption will lead to wave solutions of the form $\cos(w \pm kz)$ and $\sin(w \pm kz)$, which are appropriate for describing wave propagation along the z axis. A wave propagating in the positive z direction is represented by e^{-jpkz} , and e^{jpkz} corresponds to a wave propagating in the negative z direction. With an assumed z dependence e^{-jpkz} , the del operator becomes $\nabla = \nabla_t + V_z = \nabla_t - jk a_z$, since $V_z = a_z d/dz$. Note that ∇_t is the transverse part and equals $a_x d/dx + a_y d/dy$ in rectangular coordinates. The propagation phase constant fi will turn out to depend on the waveguide configuration.

Considerable simplification of Maxwell's equations is obtained by decomposing all fields into transverse and axial components and separating out the z dependence. Thus let (the time dependence e^{jpt} is suppressed)

$$\begin{aligned} \mathbf{E}(*, y, z) &= E_x(x, y, z) + E_z(x, y, z) \\ &= e(x, y)e^{-jli*} + e_z(x, y)e^{-jpkz} \end{aligned} \quad (3.56)$$

$$\begin{aligned} \mathbf{H}(x, y, z) &= H_x(x, y, z) + U_z(x, y, z) \\ &= h(x, y)e^{-jli*} + h_z(x, y)e^{-jpkz} \end{aligned} \quad (3.57)$$

where E_x, E_z , H_x , H_z are the transverse (x and y) components, and E_z, H_z are the axial components. Note also that $e(x, y), h(x, y)$ are transverse vector functions of the transverse coordinates only, and $e_z(x, y), h_z(jc, y)$ are axial vector functions of the transverse coordinates.

Consider the $\nabla \times \mathbf{E}$ equation, which may be expanded to give

$$\nabla \times \mathbf{E} = (\nabla_t - jk a_z) \times (e + e_z) e^{-jpkz} = -jton_0/h + h j e^{-jpkz}$$

$$\text{or } \nabla_t \times e - jpa_z \times e + \nabla_t \times e_z - j/3a_z \times e_z = -jwM_0h - jwM_0h$$

The term $a_z \times e = 0$, and $\nabla_t \times e_2 = \nabla_t \times a_z e = -a_z \times \nabla_t e$. Note also that $\nabla_t \times e$ is directed along the z axis only, since it involves factors such as $a_x \times a_x, a_y \times a_y, a_z \times a_z$, whereas $a_z \times e$ and $\nabla_t \times e$, have transverse components only. Consequently, when the transverse and axial

components of the above equation are equated, there results

$$\nabla \times \mathbf{e} = -j\mu_0 h_z \quad (3.58_Q)$$

$$\nabla \times \mathbf{e}_z - jpa_t \times \mathbf{e} = -\mathbf{a}, \quad \nabla \cdot \mathbf{V}_z - j/a = -ju > n_0 h \quad (3.586)$$

In a similar manner the $\nabla \times \mathbf{H}$ equation yields

$$\nabla \times \mathbf{h} = j\mu_0 e_0 \mathbf{e}_z \quad (3.58c)$$

$$\mathbf{a} \times \nabla \cdot \mathbf{h} + jpa_z \times \mathbf{h} = -jwe_0 \mathbf{e} \quad (3.58d)$$

The divergence equation $\nabla \cdot \mathbf{B} = 0$ becomes

$$\begin{aligned} \nabla \cdot \mathbf{B} &= \nabla \cdot \mu_0 \mathbf{H} = (\nabla, -jpa_2) \cdot (\mathbf{h} + \mathbf{h}_0)_{M_o C} \\ &= (\nabla, -\mathbf{h} - \mathbf{a}, -\mathbf{h} J_{M_o e})^T = 0 \end{aligned}$$

or

$$\nabla \cdot \mathbf{h} = \mathbf{M} \mathbf{K} \quad (3.58e)$$

Similarly, $\nabla \cdot \mathbf{D} = 0$ gives

$$\nabla \cdot \mathbf{e} = j0 \mathbf{e}_z \quad (3.58/-)$$

This reduced form of Maxwell's equations will prove to be very useful in formulating the solutions for waveguiding systems.

For a large variety of waveguides of practical interest it turns out that all the boundary conditions can be satisfied by fields that do not have all components present. Specifically, for transmission lines, the solution of interest is a transverse electromagnetic wave with transverse components only, that is, $E_z = H_z = 0$, whereas for waveguides, solutions with $E_z = 0$ or $H_z = 0$ are possible. Because of the widespread occurrence of such field solutions, the following classification of solutions is of particular interest.

1. Transverse electromagnetic (TEM) waves. For TEM waves, $E_z = H_z = 0$. The electric field may be found from the transverse gradient of a scalar function $\phi(x, y)$, which is a function of the transverse coordinates only and is a solution of the two-dimensional Laplace equation.
2. Transverse electric (TE), or H , modes. These solutions have $E_z = 0$, but $H_z \neq 0$. All the field components may be derived from the axial component H_z of magnetic field.
3. Transverse magnetic (TM), or E , modes. These solutions have $H_z = 0$ but $E_z \neq 0$. The field components may be derived from E_z .

In some cases it will be found that a TE or TM mode by itself will not satisfy all the boundary conditions. However, in such cases linear combinations of TE and TM modes may be used, since such linear combinations always provide a complete and general solution. Although other possible types of wave solutions may be constructed, the above three types are the most useful in practice and by far the most commonly used ones.

The appropriate equations to be solved to obtain TEM, TE, or TM modes will be derived below by placing E_z and H_z , E_x and H_x respectively, equal to zero in Maxwell's equations.

For TEM waves $E_z = H_z = 0$; SO (3.3) reduces to

$$\nabla \cdot \mathbf{E} = 0 \quad (3.59a)$$

$$(ia_z \times \mathbf{E}) = wM_0\mathbf{h} \quad (3.59b)$$

$$\nabla \times \mathbf{H} = 0 \quad (3.59c)$$

$$/3 a_z \times \mathbf{H} = -we_0\mathbf{E} \quad (3.59d)$$

$$\nabla \cdot \mathbf{H} = 0 \quad (3.59e)$$

$$\nabla \cdot \mathbf{E} = 0 \quad (3.59f)$$

The vanishing of the transverse curl of \mathbf{E} means that the line integral of \mathbf{E} around any closed path in the xy plane is zero. This must clearly be so since there is no axial magnetic flux passing through such a contour. Although $\nabla \cdot \mathbf{H} = 0$ when there are no volume currents present, the line integral of \mathbf{H} will not vanish for a transmission line with conductors on which axial currents may exist. This point will be considered again later when transmission lines are analyzed. Equation (3.59a) is just the condition that permits \mathbf{E} to be expressed as the gradient of a scalar potential. Hence let

$$e(x,y) = -V, \langle S \rangle(x,y) \quad (3.60)$$

Using (3.59f) shows that Φ is a solution of the two-dimensional Laplace equation

$$\nabla^2 \Phi(x,y) = 0 \quad (3.61)$$

The electric field is thus given by

$$\mathbf{E}(x,y,z) = -\nabla, \langle S \rangle(x,y)e^{iz}$$

But this field must also satisfy the Helmholtz equation

$$\nabla^2 \mathbf{E}_z + k^2 \mathbf{E}_z = 0$$

Since $\nabla^2 = \nabla^2 - j/3a_z$, $\nabla^2 = \nabla_z^2 + (k^2 - j^2)$, that is, the second derivative with respect to z gives a factor $-j/J_z$, this reduces to

$$\nabla^2 \mathbf{E}_z + (k^2 - j^2) \mathbf{E}_z = 0$$

$$\text{or } \nabla^2 \Phi + (k^2 - j^2) \Phi = 0$$

This shows that $j/3 = \pm k_0$ for TEM waves, a result to be anticipated from the wave solutions discussed in Chap. 2. The magnetic field may be found

from the $\nabla \times \mathbf{E}$ equation, i.e., from (3.596); thus

$$\overset{\text{WMO}}{*_{\mathbf{o}}} \mathbf{h} = a_2 \mathbf{X} \mathbf{e} = Z_0 \mathbf{h} \quad (3.62)$$

In summary, for TEM waves, first find a scalar potential Φ which is a solution of

$$\nabla^2 \Phi(*,;y) = 0 \quad (3.63a)$$

and satisfies the proper boundary conditions. The fields are then given by

$$\mathbf{E} = \mathbf{E}_z = \mathbf{e}^{* * * *} = -\mathbf{V}/De^{* * * 2} \quad (3.63b)$$

$$\mathbf{H} = \mathbf{H}_z = \pm \mathbf{h} e^{* * * \times * *} = \pm Y_0 a_z \mathbf{X} \cdot \mathbf{e}^{* * * - * * *} \quad (3.63c)$$

where $k_0 = w/(z_0 e_0)^{1/2}$, $Y_0 = (e_0/j_0)^{1/2}$, and $e^{-jk_0 z}$ represents a wave propagating in the $+z$ direction and $e^{jk_0 z}$ corresponds to wave propagation in the $-z$ direction. For TEM waves, Z_0 is the wave impedance, and from (3.63c) it is seen that, for wave propagation in the $+z$ direction,

$$\frac{V}{7T} = -\frac{F''}{W} = z_o \quad (3.64a)$$

whereas for propagation in the $-z$ direction,

For transverse electric (TE) waves, h_z plays the role of a potential function from which the rest of the field components may be obtained. The magnetic field \mathbf{H} is a solution of

$$\nabla^2 H + k^2 H = 0$$

Separating the above equation into transverse and axial parts and replacing ∇^2 by $\nabla^2 - B^2$ yield

$$\nabla^2 k_l(x,y) + k^2 h_l(x,y) = 0 \quad (3.65a)$$

$$\nabla^2 h + k^2 h = 0 \quad (3.65b)$$

where $k^2 = k_z^2 - B^2$ and a z dependence $e^{-ik_z z}$ is assumed. Unlike the case of TEM waves, B^2 will not equal k^2 for TE waves. Instead, B is determined by the parameter k^2 in (3.65a). When this equation is solved, subject to appropriate boundary conditions, the eigenvalue k^2 will be found to be a function of the waveguide configuration.

The Maxwell equations (3.58) with $e_s = 0$ become

$$\nabla \cdot \mathbf{X} \mathbf{e} = -7'w/i_0 h_2 \quad (3.66a)$$

$$\nabla \times \mathbf{X} \mathbf{e} = co^{\wedge} h \quad (3.66b)$$

$$\nabla \cdot \mathbf{X} \mathbf{h} = 0 \quad (3.66c)$$

$$\mathbf{a}_z \cdot \nabla \times \mathbf{V} / i_z + jpa_z \times \mathbf{h} = -jwe_0 e \quad (3.66d)$$

$$\nabla \times \mathbf{V} = -jlih_2 \quad (3.66e)$$

$$\nabla \cdot \mathbf{e} = 0 \quad (3.66f)$$

The transverse curl of (3.66c) gives

$$\nabla \cdot (\nabla \times \mathbf{h}) = \nabla \cdot \nabla \times \mathbf{h} - \nabla^2 \mathbf{h} = 0$$

Replacing $\nabla \cdot \mathbf{h}$ by jph_2 from (3.58e) and $\nabla^2 \mathbf{h}$ by $-k'^2 h$ from (3.656) leads to the solution for \mathbf{h} in terms of h_z ; namely,

$$h = -\nabla \times V; \quad (3.67)$$

To find \mathbf{e} in terms of \mathbf{h} , take the vector product of (3.666) with \mathbf{a}_z , to obtain

$$\nabla \times (\mathbf{a}_z \times \mathbf{X} \mathbf{e}) = p[(a_z \cdot \mathbf{e}) \mathbf{a}_z - (\mathbf{a}_z \cdot \mathbf{e}) \mathbf{a}_z] = -jSe = \nabla \times \mathbf{X} \mathbf{h}$$

$$\text{or } \mathbf{e} = -\nabla \times \mathbf{a}_z \times \mathbf{h} = -Z_a \mathbf{a}_z \times \mathbf{h} \quad (3.68)$$

The factor $k_0 Z_0 / p$ has the dimensions of an impedance, and is called the wave impedance of TE, or H , modes. It will be designated by the symbol Z_h , so that

$$Z_h = \nabla \times \mathbf{a}_z \times \mathbf{h} \quad (3.69)$$

Thus, in component form, (3.68) gives

$$T \cdot T \cdot - \cdot ? \cdot - \cdot \nabla \times \mathbf{a}_z \times \mathbf{h} \quad (3.70)$$

for a wave with z dependence e^{-jklz} .

The remaining equations in the set (3.66) do not yield any new results; so the solution for TE waves may be summarized as follows: First find a solution for h_z , where

$$\nabla^2 h_z + k^2 h_z = Q \quad (3.71a)$$

$$\text{Then } \mathbf{h} = -S \nabla \times \mathbf{a}_z \times \mathbf{h} \quad (3.71b)$$

$$\text{and } \mathbf{e} = -Z_h \mathbf{a}_z \times \mathbf{h} \quad (3.71c)$$

where $P = (k^2 - k^2)^{1/2}$ and $Z_h = \pm \infty$

Complete expressions for the fields are

$$\mathbf{H} = \pm h e^{j\phi} \mathbf{e}_z + h j t^{1/2} \mathbf{e}_x \quad (3.71d)$$

$$\mathbf{E} = \mathbf{E}_z = e e^{j\phi} \mathbf{e}_y \quad (3.71e)$$

Note that in (3.71c?) the sign in front of h is reversed for a wave propagating in the $-z$ direction since h will be defined by (3.716), with β positive regardless of whether propagation is in the $+z$ or $-z$ direction. The sign in front of e does not change since it involves the factor $/3$ twice, once in the expression for h and again in Z_h . Only the sign of one of e or h can change if a reversal in the direction of energy flow is to occur. That is, the solution for a wave propagating in the $-z$ direction can be chosen as $\mathbf{E} = -e e^{j\phi z}$, $\mathbf{H} = (h - h_z) e^{j\phi z}$ or as $\mathbf{E} = e e^{j\phi z}$, $\mathbf{H} = (-h + h_z) e^{j\phi z}$. One solution is the negative of the other. The latter solution is arbitrarily chosen as the standard in this text.

"M Waves

The TM, or E , waves have $h_z = 0$, but the axial electric field e_z is not zero. These modes may be considered the dual of the TE modes in that the roles of electric and magnetic fields are interchanged. The derivation of the equations to be solved parallels that for TE waves, and hence only the final results will be given.

First obtain a solution for e_z , where

$$V^* e_t + k^2 e_z = 0 \quad (3.72a)$$

subject to the boundary conditions imposed. This will serve to determine the eigenvalue k^2 . The transverse fields are then given by

$$\mathbf{E}_z = e e^{j\phi z} = - \frac{iB}{\tan V} e^{j\phi z} \quad (3.72b)$$

$$\mathbf{H}_z = \pm h e^{j\phi z} = \pm Y_e a_z \times e e^{j\phi z} \quad (3.72c)$$

where $f3 = (k^2 - k^2)^{1/2}$ and the wave admittance Y_e for TM waves is given by

$$Y_e = Z_h = \frac{k}{J} Y_0 \quad (3.72a-)$$

The dual nature of TE and TM waves is exhibited by the relation

$$Z_e Z_h = Z_0 \quad (3.73)$$

which holds when both types of waves have the same value of p and is derivable from (3.69) and (3.72d). The complete expression for the electric field is

$$\begin{aligned} \mathbf{E} = \mathbf{E}_+ + \mathbf{E}_- &= \mathbf{e} \mathbf{e}^T \wedge \pm \mathbf{e}_z \mathbf{e}^f * \\ &= (-\wedge \mathbf{V}, \mathbf{e}_2 \pm \mathbf{e}_1 \mathbf{e}' \wedge) \end{aligned} \quad (3.74)$$

It is convenient to keep the sign of e the same for propagation in either the $+z$ or $-z$ direction. Since $\mathbf{V} \cdot \mathbf{E} = 0$, that is, $\mathbf{V}_+ \cdot \mathbf{E}_+ + dE_z dz = 0$, this requires that the z component of electric field be $-e_z e^{jilz}$ for a wave propagating in the $-z$ direction, because $\mathbf{V}_- \cdot \mathbf{E}_-$ does not change sign, whereas dE_z/dz does, in view of the change in sign in front of ji . The transverse magnetic field must also change sign upon reversal of the direction of propagation in order to obtain a change in the direction of energy flow. For reference, this sign convention is summarized below. The transverse variations of the fields are represented by the functions e , h , e_z and h_z independent of the direction of propagation. Waves propagating in the $+z$ direction are then given by

$$\mathbf{E} = \mathbf{E}^+ = (\mathbf{e} + \ll,)^* - \bullet^* \quad (3.75a)$$

$$\mathbf{H} = \mathbf{r} = (\mathbf{h} + \mathbf{h}_z \mathbf{j} \mathbf{e}' \wedge) \quad (3.75b)$$

For propagation in the $-z$ direction the fields are

$$\mathbf{E} - \mathbf{B} \mathbf{T} = (\mathbf{e} - \bullet,)^* \bullet \quad (3.76a)$$

$$\mathbf{H} = \mathbf{H}^- = (-\mathbf{h} + \mathbf{h}_z \mathbf{j} \mathbf{e}' \wedge)^* \quad (3.76b)$$

Additional superscripts (+) or (-) will be used when it is necessary to indicate the direction of propagation. The previously derived equations for TEM, TE, and TM modes are valid in a medium with electrical constants e , ii , provided these are used to replace e_0 , n_0 . A finite conductivity can also be taken into account by making e complex, i.e., replacing e by $e - ja/co$.

The wave impedance introduced in the solutions is an extremely useful concept in practice. The wave impedance is always chosen to relate the transverse components of the electric and magnetic fields. The sign is always such that if i,j,k is a cyclic labeling of the coordinates and propagation is along the positive direction of coordinate k , the ratio $\wedge/ril = (Z_w)_k$ is positive. Here $(Z_w)_k$ is the wave impedance referred to the k axis as the direction of propagation. If i,j,k form an odd permutation of the coordinates, then EJH_j is negative. The usefulness of the wave-impedance concept stems from the fact that the power is given in terms of the

transverse fields alone. For example, for TE waves,

$$\begin{aligned}
 P &= -RejExK^* \int_a z dx dy \\
 &= \frac{1}{2} \operatorname{Re} \left[e \cdot X h^* \cdot a_z \right] dx dy \\
 &\quad + \frac{1}{2} \operatorname{Re} \left[f Z J a \cdot X h \right] \cdot \left(h^* \cdot a \right) dx dy \\
 &= \frac{1}{2} \int_s / h \cdot h^* dx dy = \frac{1}{2} \int_s / f e \cdot e^* dx dy
 \end{aligned}$$

upon expanding the integrand. Thus the wave impedance enables the power transmitted to be expressed in terms of one of the transverse fields alone. A further property of the wave impedance, which will be dealt with later, is that it provides a basis for an analogy between conventional multiconductor transmission lines and waveguides.

3.8 TRANSMISSION LINES (FIELD ANALYSIS)

Lossless Transmission Line

A transmission line consists of two or more parallel conductors. Typical examples are the two-conductor line, shielded two-conductor line, and coaxial line with cross sections, as illustrated in Fig. 3.10. Initially, it will be assumed that the conductors are perfectly conducting and that the medium surrounding the conductors is air, with $\epsilon \sim \epsilon_0$, $\mu \sim \mu_0$. The effect of small losses will be considered later.

When the conductors are completely surrounded by a uniform dielectric medium, the principal wave that can exist on the transmission line is a TEM wave. The electric field for this wave can be found from the scalar potential which is a solution of Laplace's equation in the transverse plane. Microstrip lines and other planar transmission lines do not have the dielectric medium completely surrounding the conductors and therefore do not support a pure TEM wave. In this case it is found that only in the low-frequency limit does the dominant mode of propagation approach that of a TEM wave. We refer to the principal wave on these lines as a

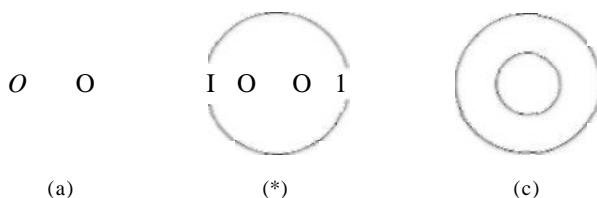


FIGURE 3.10

Cross sections of typical transmission lines, (a) Two-conductor line; (b) shielded two-conductor line; (c) coaxial line-

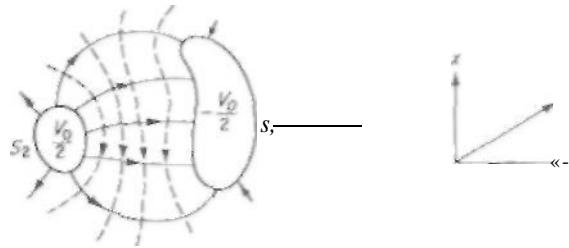


FIGURE 3.11

Cross section of a general two-conductor line showing transverse field patterns.

quasi-TEM wave. The solution for the electric and magnetic fields of the quasi-TEM wave requires a separate solution for both the electric and magnetic fields in order to determine the distributed-circuit line parameters R , L , C , and G . This is because the electric and magnetic fields are no longer related in the simple way that they are for the TEM wave. The solution for the magnetic field can be found by solving for the vector potential function as will be shown later. In this section and the following one, we consider only transmission lines that support a TEM wave.

With reference to Fig. 3.11, let the one conductor be at a potential $V_0/2$ and the other conductor at $-V_0/2$. To determine the field of a TEM wave, a suitable potential $\Phi(A; y)$ must be found first. It is necessary that $\langle t \rangle$ be a solution of

$$\nabla^2 A = 0$$

and satisfy the boundary conditions

$$\begin{aligned} \Phi &= V_0 && \text{on } S_2 \\ \Phi &= 0 && \text{on } S_1 \end{aligned}$$

Since $\langle P \rangle$ is unique only to within an additive constant, we could equally well choose $\Phi = V_0$ on S_2 and $\Phi = 0$ on S_1 . If a solution for Φ is possible, a TEM mode or field solution is also possible. When two or more conductors are present, this is always the case. The solution for Φ is an electrostatic problem that can be solved when the line configuration is simple enough, as exemplified in Fig. 3.10.

The fields are given by (3.63), and for propagation in the $+z$ direction are

$$E = E_x = e e^{-jko_z} = -\nabla \Phi e^{-jko_z} \quad (3.77a)$$

$$H = H_y = y_1 a_1 X e^{kx^2} \quad (3.77b)$$

The line integral of \mathbf{e} between the two conductors is

$$\begin{aligned} \oint_{S_1} \mathbf{e} \cdot d\mathbf{l} &= \int_{S_1}^{S_2} \mathbf{e} \cdot d\mathbf{l} \\ &= \oint_{S_1} \mathbf{e} \cdot d\mathbf{l} = " [\$ (S2) - *(S1) \mathbf{J}] = -V_0 \quad I \end{aligned}$$

Associated with the electric field is a unique voltage wave

$$V = V_0 e^{-j\omega t} \quad (3.78)$$

since the line integral of \mathbf{e} between S_1 and S_2 is independent of the path chosen because \mathbf{e} is the gradient of a scalar potential.

The line integral of \mathbf{h} around one conductor, say S_2 , gives

$$\oint_{S_2} \mathbf{h} \cdot d\mathbf{l} = \oint_{S_2} \mathbf{J}_s \cdot d\mathbf{l} = I_0$$

by application of Ampere's law, $\nabla \times \mathbf{H} = j\omega \mathbf{I} + \mathbf{J}$, and noting that there is no axial displacement flux D_z for a TEM mode. On the conductors the boundary conditions require $\mathbf{n} \times \mathbf{e} = 0$ and $\mathbf{n} \times \mathbf{h} = \mathbf{J}_s$, where \mathbf{n} is a unit outward normal and \mathbf{J}_s is the surface current density. Since \mathbf{n} and \mathbf{h} lie in a transverse plane, the current \mathbf{J}_s is in the axial direction. In the region remote from the conductors, $\nabla \times \mathbf{h} = 0$, but the line integral around a conductor is not zero because of the current that exists. The current on the two conductors is oppositely directed, as may be verified from the expression $\mathbf{n} \times \mathbf{h} = \mathbf{J}_s$. Associated with the magnetic field there is a unique current wave

$$I = I_0 e^{-j\omega t} \quad (3.79)$$

Since the potential ϕ is independent of frequency, it follows that the transverse fields \mathbf{e} and \mathbf{h} are also independent of frequency and are, in actual fact, the static field distributions which exist between the conductors if the potential difference is V_0 and currents I_1, I_2 exist on S_1, S_2 respectively. The magnetic lines of flux coincide with the equipotential lines, since \mathbf{e} and \mathbf{h} are orthogonal, as seen from (3.776).

Example 3-1 Coaxial line. Figure 3.12 illustrates a coaxial transmission line for which the solution for a TEM mode will be constructed. In cylindrical coordinates r, θ, z , the two-dimensional Laplace equation is

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2} = 0$$

or for a potential function independent of the angular coordinate θ ,

$$r dr \backslash dr$$

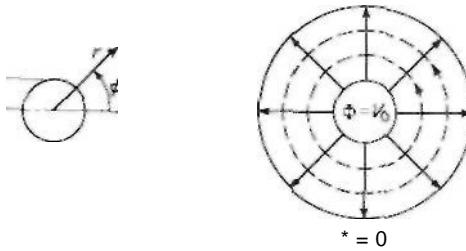


FIGURE 3.12
Coaxial transmission line.

Integrating this equation twice gives

$$\$ = C, \ln r - C,$$

Imposing the boundary conditions $4\phi = V_0$, at $r = a$, $\phi = 0$ at $r = b$, gives

$$V_0 = C, \ln a + C_2 \quad 0 = C, \ln b + C,$$

and hence $C_2 = -C, \ln b, C, = V_0 / [\ln(b/a)]$,

$$* = \frac{V_0}{\ln(b/a)} \frac{\ln(r/b)}{\ln(a/b)} \quad (3.80)$$

The electric and magnetic fields of a TEM mode propagating in the $+z$ direction are given by (3.77) and are

$$E = -a_r \frac{\partial \phi}{\partial r} * = -\frac{r}{b} \frac{\partial \phi}{\partial r} T \frac{\partial T}{\partial r}$$

$$In \left(\frac{ft}{a} \right) \frac{1}{r} \frac{1}{\ln(b/a)} \frac{1}{r} \frac{1}{\ln(b/a)} \quad (3.81a)$$

$$H = y_0 a_r \times E = \frac{y_0 V_0}{\ln(b/a)} \frac{1}{r} \frac{1}{\ln(b/a)} \quad (3.81b)$$

The potential difference between the two conductors is obviously V_0 ; so the voltage wave associated with the electric field is

$$V = V_0 e^{-j(kz - \omega t)} \quad (3.82)$$

The current density on the inner conductor is

$$n \times H = a_r \times H = \frac{W_0}{\ln(b/a)} \frac{a}{r} e^{-j(kz - \omega t)}$$

The total current, apart from the factor $e^{-j(kz - \omega t)}$, is

$$I_n = \frac{1}{a \ln(b/a)} I_a d(b) = \frac{Y_g V_0 \ln(b/a)}{\ln(b/a)} \quad (3.83)$$

The current on the inner surface of the outer conductor is readily shown to be equal to I_0 also, but directed in the $-z$ direction. The current wave associated with the magnetic field is therefore

$$I = V_0 e^{-j(kz - \omega t)} \quad (3.84)$$

The power, or rate of energy flow, along the line is given by

$$P = \frac{1}{2} \operatorname{Re} \int_{q_0}^j \rho \frac{dhrTr}{dr} E \times H^* \cdot asdrdt = \frac{1}{2} \frac{Y_a Vrf}{\ln(b/a)} j \int_{q_0}^{rb} \rho dr$$

$$Tr Y_0 Vi \ln(o/a) \quad (3.85)$$

The power transmitted is seen to be also given, as anticipated, by the expression

$$\frac{1}{2} \operatorname{Re}(V7^*) = \frac{1}{2} Y_0 I^2 = \frac{1}{2} \rho^2 n(b/a)$$

The characteristic impedance of the line is defined by the ratio

$$\frac{\rho}{o} = Z_c \quad (3.86)$$

in terms of which the power may be expressed as $P = jZ_c I^2 = \%Y_c VQ$, where Y_c is the characteristic admittance of the line and equal to Z_c^{-1} . The characteristic impedance is a function of the cross-sectional shape of the transmission line.

Transmission Line with Small Losses

Practical transmission lines always have some loss caused by the finite conductivity of the conductors and also loss that may be present in the dielectric material surrounding the conductors. Consider first the case when the conductors are surrounded by a dielectric with permittivity $\epsilon = \epsilon' - j\epsilon''$ but the conductors are still considered to be perfect. The presence of a lossy dielectric does not affect the solution for the scalar potential Φ . Consequently, the field solution is formally the same as for the ideal line, except that k_u and Y_0 are replaced by $k = k_0(\epsilon_r - j\epsilon''_r)^{1/2}$ and $Y = Y_0(\epsilon_r - j\epsilon''_r)^{1/2}$, where the dielectric constant $c_r = \epsilon'_r - j\epsilon''_r = \epsilon/\epsilon_0$. For small losses such that $\epsilon'' \ll \epsilon'_r$ the propagation constant is

$$jk = a + j/3 = y K^{1/2} * \frac{1}{\sqrt{1 - y^2}} f^{1/2} * \frac{\epsilon''^{1/2}}{(4)^{1/2} * o} + \frac{e'' k}{\sqrt{1 - y^2}}$$

Thus

$$a = \frac{6r}{2(\epsilon_r)^{1/2}} \quad (3.87a)$$

$$B = (\epsilon'_r)^{1/2} k_0 \quad (3.87b)$$

where a is the attenuation constant and B is the phase constant. The wave consequently attenuates according to e^{-ar} as it propagates in the $+z$ direction.

It will be instructive to derive the above expression for a by means of a perturbation method that is widely used in the evaluation of the attenua-

tion, or damping, factor for a low-loss physical system. This method is based on the assumption that the introduction of a small loss does not substantially perturb the field from its loss-free value. The known field distribution for the *loss-free* case is *then* used to evaluate the loss in the *system*, and from this the attenuation constant can be calculated. In the present case, if $\epsilon'' = 0$, the loss-free solution is

$$E = -V/ber^m \quad H = Ya_2XE$$

where $k = (e'_r)^{1/2}k_0$ and $Y = (e'_r)^{1/2}Y_0$. When ϵ'' is small but not zero, the imaginary part of e , that is, ϵ'' , is equivalent to a conductivity

$$a = (oe'') = we_0\epsilon''$$

A conductivity a results in a shunt current $J \sim ixE$ between the two conductors. The power loss per unit length of line is

$$P_s = \frac{1}{2\omega J_s} \int_S J \cdot J^* dS = \frac{wE''}{2J_s} \int_S E \cdot E^* dS \quad (3.88)$$

where the integration is over the cross section of the line, and the loss-free solution for E is used to carry out the evaluation of P_s . Since loss is present, the power propagated along the line must decrease according to a factor e^{-2az} . The rate of decrease of power propagated along the line equals the power loss. If the power at $z = 0$ is P_0 , then at z it is $P = P_0 e^{-2az}$. Consequently,

$$\frac{dP}{dz} = -P_i = 2aP_0 e^{-2az} = 2aP \quad (3.89)$$

which states that the power loss at any plane z is directly proportional to the total power P present at this plane. The power propagated along the line is given by

$$\begin{aligned} P &= \frac{1}{2} \operatorname{Re} \int_S E \times H^* \cdot a, dS \\ &= \frac{Y}{2} \operatorname{Re} \int_S E \times (a_2 \times E^*) \cdot a, dS = \frac{Y}{2} \int_S (E \cdot E^*) dS \end{aligned}$$

Hence the attenuation a is given by

$$\frac{P}{IP} = \frac{a}{2Y} = \frac{we''}{2Y_0(e'_r)^{1/2}} = A_s \cdot 2(0')^{1/2}$$

which is the same as the expression (3.87a). For this example the perturbation method does not offer any advantage. However, often the field solution for the lossy case is very difficult to find, in which case the perturbation method is extremely useful and simple to carry out by comparison with other methods. The case of transmission lines with conductors having finite conductivity is an important example of this, and is discussed below.

If the conductors of a transmission line have a finite conductivity, σ_{Qev} exhibit a surface impedance

$$Z_{ss} = 111 \quad (3.90)$$

where $\delta_s = (2/\omega/c\sigma)^{1/2}$ is the skin depth (Sec. 2.9). At the surface the electric field must have a tangential component equal to $Z_m J_s$, where J_s is the surface current density. Therefore it is apparent that an axial component of electric field must be present, and consequently the field is no longer that of a TEM wave. The axial component of electric field gives rise to a component of the Poynting vector directed into the conductor, and this accounts for the power loss in the conductor. Generally, it is very difficult to find the exact solution for the fields when the conductors have finite conductivity. However, since $|Z_{ss}|$ is very small compared with Z_0 , the axial component of electric field is also very small relative to the transverse components. Thus the field is very nearly that of the TEM mode in the loss-free case. The perturbation method outlined earlier may be used to evaluate the attenuation caused by finite conductivity.

The current density J_s is taken equal to $\mathbf{n} \times \mathbf{H}$, where \mathbf{n} is the unit outward normal to the conductor surface and \mathbf{H} is the *loss-free* magnetic field. The power loss in the surface impedance per unit length of line is

$$\begin{aligned} p_s &= -\operatorname{Re} Z_{ss} \mathbf{J}_s \cdot \mathbf{J}_s^* \\ &= \int_{-S_2}^{S_2} (\mathbf{n} \times \mathbf{H})^* \cdot (\mathbf{n} \times \mathbf{H}) dl \\ &= \int_{-S_2}^{S_2} \mathbf{H}^* \cdot \mathbf{H} dl \end{aligned} \quad (3.91)$$

where $R_{ss} = l/trS_s$ is the high-frequency surface resistance, and

$$\begin{aligned} (\mathbf{n} \times \mathbf{H})^* \cdot (\mathbf{n} \times \mathbf{H}) &= \mathbf{n} \cdot \mathbf{H} \times (\mathbf{n} \times \mathbf{H})^* \\ &= \mathbf{n} \cdot [(\mathbf{H} \cdot \mathbf{H}^*)\mathbf{n} - (\mathbf{H} \cdot \mathbf{n})\mathbf{H}^*] = \mathbf{H} \cdot \mathbf{H}^* \end{aligned}$$

since $\mathbf{n} \cdot \mathbf{H} = 0$ for the infinite-conductivity case. The integration is taken around the periphery $S_1 + S_2$ of the two conductors. The attenuation constant arising from conductor loss is thus

$$a = \frac{\operatorname{Re} Z_{ss}}{IP} = \frac{\int_{-S_2}^{S_2} \mathbf{H} \cdot \mathbf{H}^* dl}{2Z_0 \int_{-S_2}^{S_2} \mathbf{H} \cdot \mathbf{H}^* dS} \quad (3.92)$$

where the power propagated along the line is given by

$$\operatorname{Re} -\mathbf{J} \mathbf{E} \times \mathbf{H}^* \cdot \mathbf{a}_2 dS = -Z \mathbf{J} \mathbf{H} \cdot \mathbf{H}^* dS$$

and Z is the intrinsic impedance of the medium; that is, $Z = (iJ_e/e)^{1/2}$.

When both dielectric and conductor losses are present, the attenuation constant is the sum of the attenuation constants arising from each cause, provided both attenuation constants are small.

Example 3.2 Lossy coaxial line. Let the coaxial line in Fig. 3.12 be filled with a lossy dielectric ($\epsilon = \epsilon' - j\epsilon''$), and let the conductors have finite conductivity a . For the loss-free case $U'' = 0$, $a = \infty$) the fields are given by (3.81), with k_0 and Y_0 replaced by $k = \{\epsilon'/\epsilon_0\}^{1/2}k_0$.

Thus

$$\mathbf{E} = \frac{YV_U}{\ln(6/a)} \frac{e}{r} e^{jkz} \quad (3.93a)$$

$$\mathbf{H} = \frac{YV_U}{\ln(b/a)} \frac{a}{r^2} \frac{e}{r} e^{jkz} \quad (3.93b)$$

The power propagated along the line is

$$P_n = \frac{1}{2} \text{Re} \int d\mathbf{A} \cdot \mathbf{H}^* \mathbf{E} = \frac{1}{2} \text{Re} \int d\mathbf{A} \cdot \mathbf{H}^* \mathbf{E}$$

The power loss P_n from the lossy dielectric is, from (3.88),

$$P_n = \frac{Y^2 V_U^2}{2 \ln(6/a)} \frac{e^{2jkz}}{r^2} \quad (3.95a)$$

The power loss from finite conductivity is given by (3.91), and is

$$P_n = \frac{R_m r Y^2 V_U^2}{2 [\ln(6/a)]^2} \frac{e^{2jkz}}{r^2} \quad (3.95b)$$

Hence the attenuation constant a for the coaxial line is given by

$$a = \frac{P_n + P_n}{2P} = \frac{\frac{we''}{2Y}}{\frac{R_m Y}{2 \ln(6/a)}} \frac{b + a}{ab} = \frac{R_m}{2(e')^{1/2}} \frac{b + a}{2Z \ln(6/a)} \quad (3.96)$$

For the lossy case the propagation constant is consequently taken as

$$a + j\beta = a + jk$$

with a given by (3.96).

TRANSMISSION-LINE PARAMETERS

In this section the field analysis to determine the circuit parameters ϵ_0 , C , and G for a transmission line is examined in greater detail. This will serve further to correlate the field analysis and circuit analysis of transmission lines.

Consider first the case of a loss-free line such as that illustrated in Fig. 3.11. When the scalar potential ϕ has been determined, the charge density on the conductors may be found from the normal component of electric field at the surface; that is, $p_s = \epsilon n \cdot e = -\epsilon n \cdot \nabla \phi = -\epsilon \nabla \phi / dn$, where ϵ is the permittivity of the medium surrounding the conductors. The total charge Q per unit length on conductor S_2 is

$$Q = \epsilon n \cdot e dl$$

The total charge on the conductor S_2 is $-Q$ per meter. The potential of S_2 is V_0 , and hence the capacitance C per unit length is

$$C = \frac{Q}{V_0} = \frac{\epsilon n \cdot e dl}{V_0} \quad (3.97)$$

The total current on S_2 is

$$I_0 = \int_{S_2} h dl = \int_{S_2} Y n \cdot e dl = Y Q$$

since $h \cdot n = Y|e| = Yn \cdot e$ at the surface of S_2 because the normal component of h and the tangential component of e are zero at the perfectly conducting surface S_2 . The characteristic impedance of the line is given by

$$V_0 = \sqrt{\epsilon Z}$$

(3.98)

A knowledge of the capacitance per unit length suffices to determine the characteristic impedance.

To determine the inductance L per unit length, refer to Fig. 3.13, which illustrates the magnetic flux lines around the conductors. Since h &

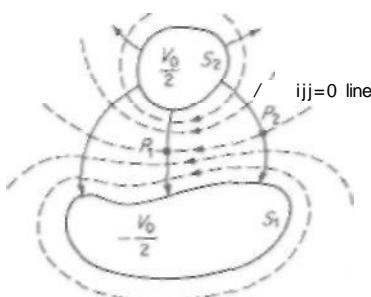


FIGURE 3.13
Magnetic flux lines in a transmission line.

orthogonal to \mathbf{e} , these coincide with the equipotential lines. All the flux *lines* from the $\$ = 0$ to the $4> = V_0/2$ line link the current on S_2 . The flux linkage is the total flux cutting any path joining the $\Phi = 0$ line to the surface S_2 . If a path such as P_1S_2 or P_2S_2 is chosen, which is orthogonal to the flux lines, this path coincides with a line of electric force. The flux cutting such a path is

$$\oint_{P_1} \mathbf{f} \cdot d\mathbf{l} = \int_{P_1} Y^{Sa} - e d\lambda = t i Y_{\frac{1}{2}}$$

since $Ih| = Y/eI$ for a TEM wave. The inductance of one conductor of the line is

$$T * \mathbf{v}^{\mathbf{V} \circ}$$

The inductance of both conductors per unit length is twice this value; so

$$L = ,MY^{\wedge} = \mu Y Z_e \quad (3.99)$$

From this relation and (3.98) it is seen that $Z = \lambda Z_c/L = CZ_c/e$, and hence

$$f = \frac{M}{e} = \frac{y Z_c}{L} = \frac{CZ_c}{e}$$

which gives

$$Z_c = y \quad (3.100)$$

Equations (3.98) and (3.99) also show that

$$fie = LC \quad (3.101)$$

for a transmission line. The above expressions for C and L can also be obtained from the definitions based on stored energy. The derivation is left as a problem.

If the dielectric has a complex permittivity $e = e' - j\epsilon''$, where ϵ'' includes the conductivity of the dielectric if it is not zero, the total shunt current consists of a displacement current I_D and a conduction current I_s . The current leaving conductor S_2 per unit length is

$$I = I_D + I_s = jioe(p \mathbf{e} \cdot \mathbf{n} dl) = joje'(p \int_{S_z} \mathbf{e} \cdot \mathbf{n} dl) + a>e''(p \int_{S_z} \mathbf{e} \cdot \mathbf{n} dl)$$

where the first integral on the right gives the displacement current and the second integral gives the conduction current. The total shunt admittance is given by $Y = jcoC + G = (I_s + I_D)/V_0$, and hence it is seen that

$$\mathbf{G} = \mathbf{v}^{\mathbf{V}} = \frac{cpe''}{\int_{\mathcal{H}} \mathbf{V}_{\mathcal{H}}} = \frac{1}{C} \quad (3.102)$$

since $jwC = I_D/V_0$ and $jcoC/jcoe' = C/e'$. This relation shows that Q differs from C by the factor e/e' .

The transmission-line loss from finite conductivity may be accounted for by a series resistance R per unit length provided R is chosen so that

The right-hand side gives the total power loss per unit length arising from the high-frequency resistance of the conductors. In terms of this quantity the resistance R is thus defined as

$$R = R_{\gg} \approx \frac{1}{4\pi^2} \int_{S_t+S_2} J^2 dl \quad (3.104)$$

where $R_m = l/(r\delta_s)$ and δ_s is the skin depth.

A further effect of the finite conductivity is to increase the series inductance of the line by a small amount because of the penetration of the magnetic field into the conductor. This skin-effect inductance L_s is readily evaluated on an energy basis. The surface impedance Z_m has an inductive part $jX_m = j/\omega\delta_s$ equal in magnitude to R_m . The magnetic energy stored in X_m is (note that X_m is equivalent to a surface inductance $X_m/\omega = L_m$)

$$W_m = \frac{1}{4\pi^2} \int_{S_t+S_2} J^2 dl$$

$$= \frac{1}{4\pi^2} \int_{S_t+S_2} J^2 dl + \frac{1}{4\pi^2} \int_{S_t+S_2} R_m J^2 dl = \frac{1}{4\pi^2} \int_{S_t+S_2} (R_m + \omega L_s) J^2 dl$$

by using (3.103) to replace the integral. Defining L_s by the relation

$$|LJ^2| = W_m$$

gives

$$\omega L_s = R \quad (3.105)$$

The series inductive reactance of the line is increased by an amount equal to the series resistance. However, for low-loss lines, $R \ll \omega L_s$, so that $L_s \approx R$ and the correction is not significant for most practical lines. The inductance L_s is called the internal inductance since it arises from flux linkage internal to the conductor surfaces.

It should not come as a surprise to find that $\omega L_s = R$ since both the inductive reactance and resistance arise from the penetration of the current and fields into the conductor. The effect of this penetration into the conductor by an effective distance equal to the skin depth δ_s is correctly accounted for in a simplified manner by introduction of the surface impedance $Z_m = (1 + j)/\omega\delta_s$.

Example 3.3 Coaxial-line parameters. For the coaxial line of Fig. 3.12 the potential ϕ is given by

$$\phi = V_0 \ln(r/a)$$

The charge on the inner conductor is

$$Q = e \int_0^r a_r \cdot ead \ln(b) = e \int_{\ln(a/b)}^{\ln(6/a)} ad \ln(b) dr$$

$$m = \star r^\wedge = J^\wedge L$$

Hence the capacitance per unit length is

$$C = \frac{e}{V_0} = \frac{Q}{\ln(b/a)} \quad (3.106)$$

since the capacitance arises only from that part of the charge associated with e' whereas e'' gives rise to the shunt conductance.

The magnetic field is given by (3.936) as

$$\frac{YV}{\ln(6/a)} \frac{a}{r}$$

The current I_0 is

$$I_0 = \frac{r^z u}{J_0 h} \cdot a.ad4 = \frac{2irYV}{\ln(6/a)} Q$$

Thus the characteristic impedance is

$$Z_c = \frac{V_0}{I_0} = \frac{YV}{\ln(6/a)} \quad (3.107)$$

The flux linking the center conductor is

$$m(o/a) \cdot n \cdot r$$

Consequently, the inductance per unit length is

$$L = \frac{ib}{r} = \frac{ixYV_0}{2TTYV_0} \frac{6}{a} \ln \frac{y}{2T} \frac{b}{a} \quad (3.108)$$

from which it is seen that $LC = ye'$ and $Z_c = (LC)^{1/2}$.

The shunt conductance G is given by $we''/C'e'$, and is

$$G = \frac{te''}{6} \frac{2-rre'}{\ln(6/a)} \frac{2-nwe''}{\ln(6/a)} \quad (3.109)$$

To find the series resistance, the power loss in the inner and outer conductors must be evaluated. This was done in Example 3.2, with the result

[Eq. (3.956)]

$$\frac{1}{2} \mathbf{R} \mathbf{I} \mathbf{S} = \mathbf{P}_{12} \frac{R_m T r Y^2 V_0^2 b + a}{[ln(b/a)f]}$$

Solving for R gives

$$R = \frac{R_{12}}{ITT} \frac{b + a}{ab} \quad (3.112)$$

The internal inductance L_s is equal to R/lo ; so the total series line inductance per unit length is

$$L_s = h^{in} a + 2,0, ab^2 r \quad (3.113)$$

The distributed-circuit parameters R , L , C , and G for a transmission line can also be determined from an evaluation of the stored electric and magnetic field energy and the power loss per unit length. Energy storage in the magnetic field is accounted for by the series inductance L per unit length, whereas energy storage in the electric field is accounted for by the distributed shunt capacitance C per unit length. Power loss in the conductors is taken into account by a series resistance R per unit length. Finally, the power loss in the dielectric may be included by introducing a shunt conductance G per unit length. Suitable definitions for the parameters L , C , R , and G based on the above concepts are

$$L = \frac{\sim r W}{*o'os} \quad (3.112a)$$

$$C = \frac{-^n}{W} \int_S E \cdot E^* dS \quad (3.112b)$$

$$R = -^n(f) \quad HH'dl \quad (3.112c)$$

$$G = -^n \int E \cdot E^* dS \quad (3.112d)$$

where $/,,$ is the total current on the line, and V_0 the potential difference. These expressions are obtained, for example, by equating the magnetic energy $\mid l_0 lo L = W_m$ stored in the equivalent series inductance L to the expression for W_m in terms of the field. The above definitions are readily shown to be equivalent to the other commonly used definitions such as

magnetic flux linkage 113f1j

total current

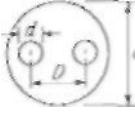
total charge per unit length 113ft

voltage difference between conductors

$$G = \frac{\text{total shunt current}}{\text{voltage difference between conductors}} \quad (3.113e)$$

Parameters of some common transmission lines are given in Table 3.*-j

TABLE 3.1
Parameters of common transmission lines*

	R
	$\cosh^{-1} \frac{D}{2R_r}$ $\frac{2R_r}{nd}$ $[(D/d)^2 - 1]^{1/2}$
	$\frac{1}{2\pi} \frac{Ho}{ve}; \quad \ln \frac{o}{a}$ $2n \cdot \frac{a}{a-b} J$ $\ln \frac{2p}{1+g^*} \quad 2ff, \quad I + 2p^2$ $1 + 4p^2 \quad + \text{---} \mathbf{V} \quad (1 + g^*)^{-1} \cdot 4p -$

For all TEM transmission lines

$$\frac{C}{Z_0} = \frac{(M_0 \mu_0)^{*}}{GZ} L = (M_0 \epsilon_0)^{*} Z_t \quad G = \frac{\epsilon_0 C}{R K} \quad /?, \quad \frac{-1}{5} \sqrt{5}^{\text{M}} \backslash^{1/U} \quad \frac{2\pi I}{U}$$

310 INHOMOGENEOUSLY FILLED PARALLEL-PLATE TRANSMISSION LINE

In Fig. 3.14a we show a parallel-plate transmission line (waveguide) partially filled with dielectric material having a permittivity $\epsilon = \epsilon_r \epsilon_0$, where ϵ_r is the dielectric constant. The plates are infinitely wide and spaced a distance b apart. The dielectric sheet has a thickness a and rests on the bottom plate.

The purpose for studying this particular waveguide is that it exhibits a number of characteristics that are similar to those of the microstrip transmission lines examined in the following section. We will show that the dominant mode of propagation in the waveguide under consideration is an E mode and that as the frequency approaches zero this mode becomes a TEM mode. Furthermore, in the low-frequency limit, the propagation con-

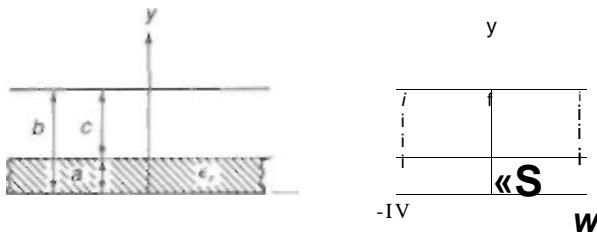


FIGURE 3.14

(a) Partially loaded parallel-plate waveguide; (b) parallel-plate waveguide with magnetic wall at $x = +W$.

stant can be found in terms of the distributed capacitance; and inductance per meter by the usual transmission-line formula $\ell/\omega = covLC$. As the frequency increases i increases faster than to , in which case we say that the transmission line exhibits dispersion.

Another feature that can be easily described for this waveguide is the existence of a surface-wave mode of propagation that consists of a field concentrated near the air-dielectric interface.

Since the analytic solution for the partially filled parallel-plate waveguide is readily constructed, this waveguide serves as a useful example to provide some physical insight into the properties of microstrip transmission lines.

An electric wall is a surface on which the tangential electric field must be zero. A good conductor such as copper provides a surface with a very small skin-effect surface impedance (see Sec. 2.9). When we let the conductivity σ become infinite, we obtain an electric wall on which the boundary condition $\mathbf{n} \times \mathbf{E} = 0$ holds. The dual of an electric wall is a magnetic wall on which the tangential magnetic field is zero, i.e., the boundary condition $\mathbf{n} \times \mathbf{H} = 0$ holds. The magnetic wall does not have a physical realization but is, nevertheless, a useful theoretical concept. In practice, a magnetic wall can be inserted into a field region, without disturbing the field, along any surface on which the tangential magnetic field is zero. Such surfaces usually correspond to certain symmetry planes in a given problem. In addition to the above boundary conditions, Maxwell's equations show that on an electric wall the normal component of \mathbf{H} is zero, that is, $\mathbf{n} \cdot \mathbf{H} = 0$. The dual boundary condition $\mathbf{n} \cdot \mathbf{E} = 0$ holds on a magnetic wall.

For the E mode that we will consider in the partially filled parallel-plate waveguide, we will assume that the fields do not depend on the x coordinate but are functions of y and z only. A consequence of this assumption is that only the field components E_y , E_z and H_x are present. Thus we can place a magnetic wall along any $x =$ constant surface without disturbing the field. We will now assume that magnetic walls are inserted at $x = \pm W$ as shown

in Fig. 3.146. By means of this artifice, we are able to talk about a closed waveguide structure, closed by *electric* walls at $y = 0, b$ and by *magnetic* walls at $x = \pm W$.

In order to find the solutions for E modes having the z dependence $e \sim J^p z$, we must find solutions for the axial electric field component $e_z(y)$ first. In an ideal transmission line the propagation constant equals that for plane TEM waves in the surrounding medium. In the structure under investigation we have a nonuniform medium, namely dielectric in the region $0 < y < a$ and air in the region $a < y < b$. Consequently, we can anticipate that the propagation phase constant $/3$ for the dominant mode will take on an intermediate value, i.e.,

$$k_0 = \text{toj}(i_0 e_0) < f_i < y f e^\wedge k_0 = k$$

The equation satisfied by e_{Sy} is (3.72a) which is repeated below

$$\nabla^2 e_z + k^2 e_z = 0$$

Since we assume no variation with x , the transverse laplacian operator becomes simply d^2/dy^2 . In this equation $k^2 = k_r^2 - P^2$ in the air region and equals $k^2 - l^2$ in the dielectric region. The propagation constant $(3$ must be the same in both regions because the tangential electric and magnetic fields must match at the air-dielectric interface for all values of z . For convenience, we will let $k_c = I$ in the dielectric region and let it equal p in the air region. We thus require that $p^2 - k_r^2 = I^2 - k^2$ or

$$l^2 - p^2 \sim (e_r - l)k^2 \quad (3.114)$$

In the two regions the axial electric field is thus a solution of

$$\frac{d^2 e_z}{dy^2} = 0 \quad 0 \leq y < a \quad (3.115a)$$

$$a \leq y < b \quad (3.115b)$$

along with the boundary conditions

$$e_A y) = 0 \quad y = 0, 6 \quad (3.116a)$$

$$e_z(y) \text{ continuous at } y = a \quad (3.116b)$$

$$\left. \frac{e_r}{I^2} \frac{de_2}{dy} \right|_a = \left. \frac{1}{7^2} \frac{de_2}{dy} \right|_a \quad (3.116c)$$

The third boundary condition comes from the requirement that H_x be continuous across the air-dielectric interface. The transverse fields are given

by (3.726) and (3.72c). The generic form of the equations is

$$P \quad Jfiae** -J/5*$$

$$''* \quad *'k*dy^e$$

upon using a. $X_a = -a$. In the equation for ff, the wave admittance y is given by kY/fi in the dielectric and by $(k_0 Y_0)/p$ in the air region where $kY = o>yJ/x_0\epsilon i/e/fx_0 = fr^a o^a o^-$. Thus we have

$$e_v(y) = \begin{cases} -\frac{iR}{p} \frac{\partial e}{\partial y} & \text{dielectric region} \\ -\frac{j\beta}{p^2} \frac{\partial e_z}{By} & \text{air region} \end{cases} \quad (3.117a)$$

$$K(y) = \begin{cases} \frac{je, k_0 Y_0}{p^2} \frac{de_2}{3y} & \text{dielectric region} \\ \frac{jk_0 Y_0}{p^2} \frac{de_z}{dy} & \text{air region} \end{cases} \quad (3.1176)$$

An examination of the expression for $h_x(y)$ shows that continuity of h_x at $y = a$ gives the boundary condition specified by (3.116c). We also find that in an inhomogeneously filled waveguide, the wave impedance, defined by the ratio $-E_y/H_x$, is not constant since it has a different value in the air region from that in the dielectric region.

The reader can readily verify that the solutions of (3.115a) and (3.1156) that satisfy the boundary conditions at $y = 0$ and b are

$$C_j(y) = C_1 \sin ly \quad 0 < y < a$$

$$^e z(y) = C_2 \sin p(b - y) \quad a < y < b$$

where C_1 and C_2 are unknown amplitude constants. The boundary condition (3.1166) requires that

$$C_1 \sin la = C_2 \sin pc$$

where $c = b - a$. The last boundary condition (3.116c) requires

$$\frac{*r}{I} C_1 \cos la - \frac{*}{p} C_2 \cos pc$$

When we divide the first equation by the second one, we obtain

$$/ \tan la = -e_r p \tan pc \quad (3.118)$$

This transcendental equation must be solved simultaneously with (3.114) to determine the allowed values of β and p . There will be an infinite number of solutions; consequently an infinite number of E modes are possible. Since $\beta/3$ is given by

$$(i = y/k^2 - p^2 = y/k^2 - I^2) \quad (3.119)$$

most of the modes will be nonpropagating since increasing values of p and β give $p > k_0$ which makes $\beta/3$ imaginary. When $\beta/3$ is imaginary the z dependence is of the form $e^{-il\beta z}$ and the field decays exponentially from the point at which it is excited. These nonpropagating modes are called evanescent modes.

We note from (3.119) that a value of $\beta/3$ between k_0 and k can occur only if p is imaginary. Thus we must consider the possibility that an imaginary p , say $p = jp_0$, is a possible solution to (3.118). If we let I_0 be the corresponding value of I , then our relevant equations become

$$Z_0 \tan l_0 a = e_r p_0 \tanh p_0 c \quad (3.120a)$$

$$l^2 + p^2 = (e_r - l)k_l \quad (3.120b)$$

We consider solutions of these equations and the corresponding fields in the low- and high-frequency limits in the next two subsections.

Solution

When the frequency is very low, $k'l$ is a very small number (at 1 MHz, k_0 equals 0.02094 rad/m); hence β_0 and p_0 are then also small. We will assume that a is at most a few centimeters, then $l_0 a$ and $p_0 c$ are also small and we can replace the tangent function and the hyperbolic tangent function by their arguments. Thus (3.120a) becomes

Upon using (3.120b) we readily find that

$$(e_r - l)k^2 - p^2 = \frac{a}{a + e_r c}$$

$$a + e_r c$$

The solution for $\beta/3$ in the low-frequency limit is thus

$$P = 1/*_0 + Pi = \sqrt{\frac{a}{a + e_r c}}*_0 = \sqrt{e_r} *_0 \quad (3.12D)$$

where e_r , given by this equation, is called the effective dielectric constant.

We will now show that this equals $\text{co} \angle /LC$, where L and C are the stat" distributed inductance and capacitance per meter for the given structure

If we have a uniform current density J_z on the inner surface of n. upper plate and $-J_z$ on the inner surface of the lower plate, the magneti field between the plates will be given by $H_x = J_z$. The time-average stored magnetic energy per unit length is given by

$$4 Jn' \frac{W}{2}$$

We equate this to $\text{^L} If$ where the total current $I = 2WJ_z$ and then find that

$$L = \frac{W}{2} \quad (3.122)$$

The distributed capacitance C per meter is found by considering the capacitance of the dielectric and air regions as represented by two equivalent parallel-plate capacitances C_{rf} and C_a in series where

$$C_{rf} = \frac{e_r e_0 2W}{C_a + C_d} = \frac{e_0 2W}{e_r c + a}$$

The capacitance C_d is that of a parallel-plate capacitor of width $2W$, unit length, plate spacing a , and filled with dielectric. C_a is the capacitance of the air-filled section which has a spacing c .

The series capacitance is given by

$$\frac{C_n C_d}{C_a + C_d} = \frac{2We_r e_0}{e_r c + a} \quad (3.123)$$

The product $LC = e_r e_0 i_0 b / (e_r c + a)$ which gives the solution for $\beta = i_0 \sqrt{LC}$ equal to that in (3.121).

The expressions for the fields can be written down in simplified form using the small argument approximations and the relationship $C_2 * C_1 \sin l_0 a / j \sinh p_0 c = -jC_1 \hat{C} \hat{p} \hat{c}$ obtained from the boundary condition requiring continuity of e_z at $y = a$. We readily find that in the region $0 < y < a$,

$$e_z = C \hat{y} \quad (3.124a)$$

$$e_y = -\mathbf{v} \wedge -\mathbf{i} \wedge \mathbf{s}$$

$$\begin{aligned} J^r r k_0 Y_0 & U_r (e_r c + a) \\ *0 & V (e_r - 1)c \end{aligned}$$

and in the air region

$$e_s = \frac{c}{c} \hat{c} \wedge (b-y) \quad (3.124\text{rf})$$

$$\begin{array}{c} JPUa \quad jfi \\ \hline / \quad * \\ / \quad (\epsilon_r - 1)c \end{array} \quad (3.124\text{e})$$

$$\begin{array}{c} jk_0 Y_0 \quad le_r(e_r c + a) \\ \hline K = IQ \quad c, y, c, y / (6p-1)c \end{array} \quad (3.124/)$$

We note that in the low-frequency limit e_z vanishes as k_0 , and hence $/_0$, approach zero, while e_v and h_x remain constant. If we define the voltage V between the upper and lower plates by the line integral of e_y , then

The total 2-directed current on the upper plate is $t_z = 2WJ_z = 2WH_X$, and hence the characteristic impedance is given by

$$Z_c = \frac{V}{I_z} = \frac{Z_0}{2W} \sqrt{\frac{(\epsilon_r c + a)b}{\epsilon_r}} = \sqrt{\frac{L}{C}} \quad (3.125)$$

Thus we find that in the low-frequency limit the dominant mode of propagation in the partially filled parallel-plate waveguide becomes a TEM mode and the waveguide may be analyzed as a transmission line. The propagation constant and characteristic impedance are determined by the static distributed inductance and capacitance. In general, at low frequencies the mode of propagation would be called a quasi-TEM mode since the axial electric field e_z , even though it is small, is not zero. At high frequencies the mode of propagation is an E mode and departs significantly from a TEM mode in its field distribution.

Solution

At high frequencies k_0 , and hence l_0 and p_0 , are large. In this case p_0c is large so we can replace $\tanh p_0c$ by unity and (3.120a) gives

$$Z_0 \tan J_0 a = \epsilon_r p_0 = \epsilon_r V (\epsilon_r - 1)^* o - ' o \quad (3.126)$$

upon using (3.1206) to eliminate p_0 . This equation is independent of the

plate separation δ . The solution for e_z can be approximated as follows:

$$e_t(y) = C \sin l_0 y \quad 0 < y < a \quad (3.127_a)$$

$$e_s(y) = C_j \sinh p_0 (b - y)$$

$$= C \frac{\sinh p_0(b - y)}{\sinh p_0(b - a)}$$

$$e_{Pl} < b-y \rangle$$

$$e^*$$

$$= C_1 \sin(\ell_0 \omega) e^{-\ell_0 y} \quad a < y < b \quad (3.127_b)$$

This is a field that decays exponentially away from the air-dielectric surface and does not depend on δ as long as $p_0 c = p_0(b - a)$ is large. This field is guided by the dielectric sheet on the ground plane (lower conductor) even if the upper plate is removed to infinity. This type of mode is called a surface-wave mode because its field is confined close to the guiding surface. The axial electric field for this surface-wave mode is illustrated in Fig. 3.15.

The first root for ℓ_0 , in the eigenvalue equation (3.126) occurs for $l_0 a < 7r/2$ or $\ell_0 < 7r/2a$. Thus, as k_0 approaches infinity, ℓ_0 remains bounded but p_0 will become large because $p_0 = (e_r - Dk^2)^{1/2}$. Consequently, for large enough k_0 we will have $1\% \ll k^2$ and then $\ell_0^2 = k^2$. As we go from zero frequency to very high frequencies, the propagation constant varies from a low value of j/e^{k_0} given by (3.121) to an asymptotic value of j/e^{k_0} . We see that ℓ_0 is not a linear function of ω or k_0 and for this reason is said to exhibit dispersion. The term *dispersion* arises from a consideration of signal propagation. In our discussion of waveguides later on in this chapter, we will show that a signal consisting of a band of frequencies will have its frequency components dispersed whenever ℓ_0 is not a linear function of ω . This is caused by the phase velocity v_p , given by the relation $\ell_0 = co/v_p$ and thus $v_p = \omega/c = 1/j\ell_0 n_{li} e_r$, being a function of ω . The ratio ℓ_0^2/ω gives the effective dielectric constant at any frequency. In Fig. 3.16 we show a plot of e_z versus frequency for the case when $e_r = 1$, $a = 0.4$ cm, and $\delta = 1$ cm. This curve is derived by solving the pair of equations (3.120). Microstrip transmission lines exhibit similar dispersion characteristics.

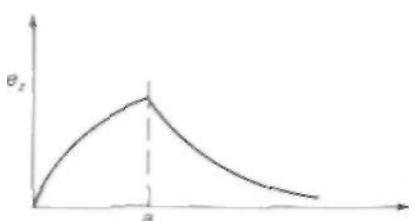


FIGURE 3.15
Axial electric field for surface-wave mode.

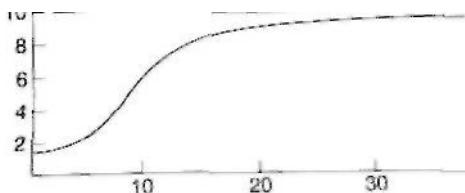


FIGURE 3.16
Effective dielectric constant as a function of frequency for $c_r = 10$, $j = 40 \text{ GHz}$, $Q = 0.4 \text{ cm}$, and $b = 1 \text{ cm}$.

A second surface-wave mode solution can be found from (3.120a) with l_{0a} in the range $T\pi < l_{0a} < 3\pi/2$ provided $y_{er} - 1/\epsilon_0$ is larger than v so that (3.120b) can also be satisfied. For large $k_{C1}a$ many surface-wave modes can propagate. In addition to the surface-wave modes, there are also an infinite number of solutions to (3.118) for real values of p . The higher-order solutions have values of p on the order of ntr/b in value, where n is an integer. Provided rr/b is greater than k_0 , these values of p will give imaginary values of fi and hence nonpropagating modes. The cutoff occurs when $p = k_0$ giving $\beta = 0$. Thus at cutoff

$$\beta \tan la = -e,k_0 \tan k_{0c} = \sqrt{e,k_0} \tan ^2 e,k_0 a$$

since $\beta = k$ for $(i = 0)$. This equation reduces to

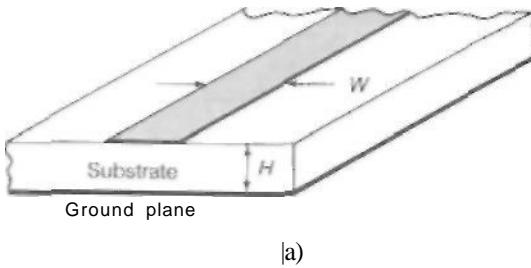
$$\tan yj7_{\sim r,k_0 a} = -y/tT_r \tan k_{0c}$$

which can be solved for the values of k_0 at which the various modes cease to propagate.

We will not consider the partially filled parallel-plate waveguide any further even though a good deal more could be said about its mode spectrum. The purpose of our discussion is to highlight those features that will be displayed by microstrip transmission lines, which is the next topic taken up.

PLANAR TRANSMISSION LINES

A planar transmission line is a transmission line with conducting metal strips that lie entirely in parallel planes. The most common structure is one or more parallel metal strips placed on a dielectric substrate material adjacent to a conducting ground plane. A planar transmission line that is widely used is the microstrip line shown in Fig. 3.17. It consists of a single conducting strip of width W placed on a dielectric substrate of thickness H and located on a ground plane. By image theory this transmission line is equivalent to a line consisting of two parallel conducting strips placed opposite each other on a dielectric sheet of thickness $2H$ as also shown in



(a)

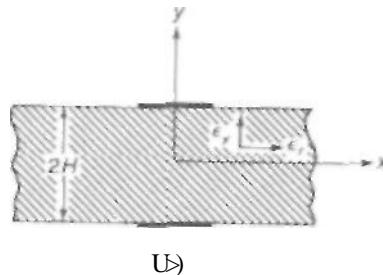


FIGURE 3.17

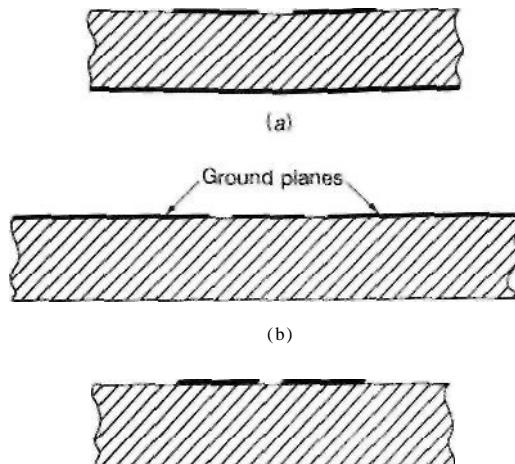
(a) The microstrip transmission line;
(b) equivalent parallel strip line obtained by using image theory.

Fig. 3.17. Typical dimensions for a microstrip line are substrate thickness of 0.25 to 1 mm and strip widths of 0.1 to 5 mm.

The microstrip transmission line can be fabricated using conventional printed-circuit-board techniques which result in good mechanical tolerances and a low cost.

In addition to the microstrip line, there are many other planar-transmission-line structures that are used for various purposes. A number of these other transmission-line configurations are shown in Figs. 3.18 to 3.20. The coupled microstrip line shown in Fig. 3.18a is used in directional couplers. The coupled microstrip line supports two modes of propagation. The even mode of propagation has the same voltage and current on the two strips, while the odd mode of propagation has opposite voltages and currents on the two strips.

The coplanar transmission line shown in Fig. 3.18c consists of a single strip mounted between two ground planes on the same side of the dielectric substrate. The coplanar line has an advantage over the microstrip line in that shunt connection of components to the ground plane can be made on the same side of the substrate. In addition, it allows the series connection of components to be made with equal facility to that for microstrip lines. The coplanar strip line shown in Fig. 3.18c is similar to the coplanar line in that all conducting strips are in the same plane (coplanar). It is less desirable than the coplanar line because it is not balanced relative to a ground plane¹¹⁶ and thus wave propagation on this line is more strongly influenced by nearby conductors such as a shielding enclosure. In practice, a shielding

**FIGURE 3.18**

(a) Coupled microstrip lines; (b) coplanar transmission line; (c) coplanar strip transmission line.

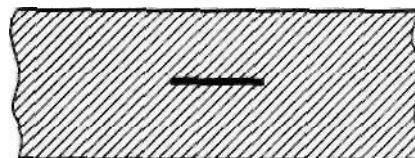
wmy//A mmm

FIGURE 3.19

)d suspended mi-



lb)



enclosure for a microwave circuit is needed to reduce spurious radiation from the circuit, eliminate electromagnetic coupling with nearby circuit, and for environmental protection.

In Fig. 3.19 are illustrated the suspended and inverted suspended microstrip lines which are quite similar to the conventional microstrip line but involve less dielectric substrate material. Figure 3.20a shows a slot line. The open slot line is not as widely used as the microstrip and coplanar lines are. The shielded slot line, which is typically a slot line placed inside a rectangular waveguide as shown in Fig. 3.20b, is called a fin line and has found to be useful for many circuit applications. Only shunt-connected components can be used with a slot line.

The final transmission-line structure illustrated here is the strip line shown in Fig. 3.20c. This line consists of a strip placed between two parallel plates that function as ground planes. The strip may be rigid enough to be suspended in air or it may be sandwiched between two dielectric sheets as shown in the figure. The strip line was often used in microwave filters and couplers before the other forms of planar transmission lines became popular. The strip line is somewhat more difficult to fabricate but has certain advantages for special applications to filters and couplers. Many directional coupler, power divider, and filter designs using strip lines were developed in the period from 1955 to 1975. An excellent reference source for design data for strip-line circuits is the book by Howe listed in the references at the end of this chapter.

The methods used to fabricate planar-transmission-line structures and related circuit elements are compatible with integrated circuit fabrication and have allowed the development of microwave integrated circuits (MIC circuits). In integrated microwave circuits the active devices and all interconnecting transmission lines, impedance-matching elements, needed capacitors and resistors, etc., are fabricated on the same chip. In these applications the microstrip and coplanar transmission lines are the ones most easily adapted for on-chip fabrication. In MIC circuits the substrate thickness and line widths are generally much smaller than in hybrid circuits. The term *hybrid* is used to describe integrated microwave circuits where the discrete components such as transistors, capacitors, and resistors are soldered in place.

The dielectric substrate material used in a planar transmission line must have low loss, i.e., a small loss tangent. A large dielectric constant results in a shorter propagation wavelength and hence a more compact circuit. The substrate material should have good mechanical strength, be easy to machine, and have good thermal conduction. When active devices are mounted into a planar-transmission-line circuit, the heat generated by the active device is in part conducted away to the ground plane through the substrate material. It is difficult to use metal heat sinks in a microwave circuit because these large metal structures would interact with the electric field.

magnetic field in an undesirable, and often unpredictable, manner. Consequently, in power amplifier circuits substrate materials with good thermal conductivity are required. Dielectric materials used in low-frequency circuit boards are generally too lossy to be used for microwave transmission lines. The dielectric constant and thickness must be maintained to a high level of uniformity in the manufacturing of substrates because otherwise the fabricated transmission lines will not perform according to the specified design since the propagation phase constant and characteristic impedance both depend on these parameters. Uniform thickness and dielectric constant is particularly important in the design of filters and impedance-matching elements whose dimensions are critical. Once a printed microwave structure such as a filter has been constructed, it is not *very easy* to add external tuning elements to bring the constructed filter performance into specifications.

A commonly used substrate material is polytetrafluoroethylene (PTFE)^t which has a dielectric constant of 2.1 and a loss tangent of 0.0002 at 1 MHz and around 0.0005 at microwave frequencies. This material has excellent resistance to chemicals used in the photoetching process. In order to increase the mechanical strength, it can be loaded with woven fiberglass mat or glass microparticles. This increases the dielectric constant to the range 2.2 to 3. The use of glass fiber results in some anisotropy in the dielectric constant. In the manufacturing process the glass fibers are generally aligned parallel with the substrate so the dielectric constant along the substrate is typically 5 to 10 percent larger than that normal to the substrate. By using ceramic powders as fillers, notably titanium oxide, much larger dielectric constants can be obtained. Typical values are in the range 5 to 15.

Ceramic materials such as aluminum oxide (alumina) and boron nitride, as well as the glasslike material sapphire, are also used (*or substrates*). These materials are very difficult to machine. Alumina is perhaps the most commonly used material. It has excellent thermal conductivity. For integrated microwave circuits the usual semiconductor materials germanium, silicon, and gallium arsenide are used. These substrate materials have a high dielectric constant and may exhibit some conductivity depending on the doping level.

In Table 3.2 we have summarized the important properties of a number of substrate materials. In this table ϵ_r is the dielectric constant along the substrate and ϵ_y is the dielectric constant normal to the substrate.^{\$}

^tThis material is commonly known as Teflon, which is a registered trade name of Du Pont. In order to keep the notation as simple as possible, we use ϵ_y instead of ϵ_{ry} for the relative permittivity (dielectric constant) in the y direction.

TABLE 3.2
Properties of substrate materials

Material	Loss tangent	Thermal conductivity	Ma. hioaba^
PTFE/woven glass	2.84	2.45	0.001-0.002
PTFE/microfiberglass	2.26	2.2	0.0005-0.001
•CuFion	2.1	2.1	0.0004
*RT/Duroid 5880	2.26	2.2	0.001
*RT/Duroid 6006	6.36	6	Medium
•Epsilam 10	13	10.3	Medium
Boron nitride	5.12	3.4	Good
Silicon	11.7-12.9	11.7-12.9	0.001-0.003
Germanium	16	16	Medium
Gallium arsenide	12.9	12.9	0.0005-0.001
Alumina	9.6-10.1	9.6-10.1	0.0005-0.002
Sapphire	9.4	11.6	0.0002
Beryllium oxide	6.7	6.7	0.001-0.002

¹CuFion is a registered trademark of Polyflon Company. It is a Teflon material electroplated with copper. RT/Duroid is a registered trademark of Rogers Corporation. Rogers Corporation also manufactures substrates with dielectric constants around 10. Epsilam 10 is a registered trademark of the 3M Company. It is a ceramic-filled Teflon material.

The data in Table 3.2 have been compiled from a variety of sources.^t Since the dielectric constant and loss tangent are frequency dependent and also influenced by the material processing, the listed data should be viewed as representative values at microwave frequencies.

Substrate materials are usually plated with copper in 0.5-, 1-, or 2-oz weights (amount of copper per square foot). The use of 1-oz copper weight gives a plating thickness of 0.0014 in. Gold plating is sometimes used on top of the copper to prevent oxidation of the metal. In integrated microwave circuit construction a metalization thickness of a few microns is typical. One-half oz copper-clad board has a metalization thickness of 18 μ m.

3.12 MICROSTRIP TRANSMISSION LINE

In a microstrip transmission line the dielectric material does not completely surround the conducting strip and consequently the fundamental mode of propagation is not a pure TEM mode. At low frequencies, typically below a

^tH. Howe, "Stripline Circuit Design," Artech House Books, Dedham, Mass., 1974.

T. Laverghetta, Microwave Materials: The Choice is Critical, *Microwave J.*, vol. 28, p. 163, 1985.

M. N. Afsar and K. J. Button, Precise Millimeter-Wave Measurements of Complex Refractive Index, Complex Dielectric Permittivity and Loss Tangent of GaAs, Si, SiO_2 , Al_2O_3 , BeO, Macor, and Glass, *IEEE Trans.*, vol. MTT-31, pp. 217-223, 1983.

Some data were also obtained from manufacturers' literature.

few gigahertz for practical microstrip lines, the mode is a quasi-TEM mode. In the frequency range up to a gigahertz or somewhat higher, the microstrip transmission line can be characterized in terms of its distributed capacitance and inductance per meter in a manner similar to what was found for the partially loaded parallel-plate transmission line in the previous section. Unfortunately, there are *no* simple *closed-form* analytic expressions that can be derived for describing the field distribution or the characteristics of planar transmission lines. Formal solutions can be derived and evaluated on a computer and have been used to compile data on the characteristics of these transmission-line structures. Static field analysis has also been extensively used to obtain the low-frequency characteristics. However, even the static field analysis is quite complex.

The analysis of planar transmission lines can be based directly on a solution for the electric and magnetic fields in the structure. An alternative approach is to first solve for the scalar and vector potential functions and from these find the corresponding electromagnetic field. In actual fact the propagation constant and characteristic impedance can be found from the potentials without a detailed consideration of the fields. The advantage of using the scalar and vector potentials in the analysis is that this approach provides a direct link to the quasistatic solutions in terms of more familiar low-frequency concepts.

In this section we will develop the essential equations to be satisfied by the scalar and vector potentials for a microstrip transmission line. From these equations we then obtain simplified ones that will describe the quasi-TEM mode of propagation at low frequencies. The term low frequency is a relative one. It is the ratio of fine dimensions to wavelength that determines whether a microstrip line can be adequately described in terms of the quasi-TEM mode of propagation. In MIC circuits with line widths as small as 100 μm , the low-frequency region *can* extend as high as 20 to 30 GHz. Even though space does not permit a full development of analytic methods suitable for solving planar-transmission-line problems, some insight into the properties of these structures is obtained from the basic formulation of the relevant equations.

After we have presented the theoretical foundations, typical dispersion curves and graphical results for characteristic impedances are given for a number of important substrate materials and a range of microstrip conductor widths.

The vector and scalar potential functions are solutions of Helmholtz equations as described in Sec. 2.10 when the sources are located in vacuum (air). For the microstrip line shown in Fig. 3.17, two added complications enter due to the presence of the dielectric in only part of the region of interest and the anisotropic nature of some substrate materials. For this reason we need to derive new equations to be satisfied by the potential functions. The substrate material will be characterized by a dielectric constant ϵ_u in the y direction which is normal to the interface and by a

dielectric constant ϵ_r in the x and z directions. The unknown charge and current densities on the conducting microstrip will be denoted by p and J . These source densities are concentrated along y since they exist only on the microstrip which is assumed to have negligible thickness. The source concentration can be described by introducing the delta function $\delta(y - H)$, localize the sources at $y = H$. Thus we can write

$$p(x,y,z) = p_s(x,z)\delta(y - H) \quad (3.128_a)$$

$$J(x,y,z) = J_s(x,z)\delta(y - H) \quad (3.128_b)$$

where J_s and p_s now describe surface densities rather than volume densities.

We will assume that the dielectric constants $\epsilon_y(y)$ and $\epsilon_r(y)$ are functions of y that are constant in the substrate and undergo a rapid change in value to unity as the interface is crossed into the air region. The reason for doing this is that the equations we then obtain for the potentials will automatically give us the boundary conditions needed to properly join the solutions for the potentials in the substrate region to those in the air region.

We begin the derivation by letting

$$\mathbf{B} = \nabla \times \mathbf{A}$$

From Maxwell's equation

$$\nabla \times \mathbf{E} = -jw\mathbf{B} = -j\omega V \times \mathbf{A}$$

we get $\nabla \times (\mathbf{E} + j\omega \mathbf{A}) = 0$ which has the general solution

$$\mathbf{E} = -j\omega \mathbf{A} - \mathbf{V} <t>$$

where $<P$ is a scalar potential function. Up to this point we have followed the same steps as in Sec. 2.10. Maxwell's curl equation for the magnetic field is

$$\nabla \times \mathbf{H} = j\omega D + \mathbf{J}$$

For an anisotropic dielectric we have

$$\mathbf{D} = \epsilon_0 \epsilon_r (E_x a_x + E_z a_z) + \epsilon_0 \epsilon_y E_y a_y$$

We can replace \mathbf{H} by $P-Q$ $\nabla \times \mathbf{A}$ to obtain

$$\nabla \times \nabla \times \mathbf{A} = \nabla V - \mathbf{A} - \nabla^2 \mathbf{A} = j<\omega> D + M_0 \mathbf{J}$$

and express \mathbf{D} in terms of the potentials as follows:

$$\mathbf{D} = -\epsilon_0 \epsilon_r j u (A_x a_x + A_z a_z) + a_x \mathbf{a}_x + a_z \mathbf{a}_z$$

$$-e_0 \epsilon_y j A_y a_y + \mathbf{a}_y \sim dy$$

By adding and subtracting a term to the y component that includes the factor e_r , we can reexpress D in the form

$$D = -e_r e_0 (\nabla A + V D) - e_0 (e_v - e_r) j a > A_y a_y + a -$$

We wish to eliminate the $V \cdot A$ term in the equation for A by setting it equal to the gradient of another function. For this purpose we now express $e_r V$ in the form $V(e_r P) - 4>v V_r$, where V_r has only a y component since e_r is a function of y only. We can set $V \cdot A$ equal to $-j i o e_0 n_0 V(e_r P)$ which gives the Lorentz condition

$$V \cdot A = -j w e_0 e_r n_0 < t > \quad (3.129)$$

The equation for the vector potential now becomes

$$\begin{aligned} -V' A &= j(o n_Q - y W_0 e_r A + e_0 * V e_r \\ &\quad - e_0 (e_y - e_r) j m a_y A_y + a_v H_0 J \end{aligned}$$

The current J does not have a y component so the x and z components of this equation are

$$V * A_x + e_r(y) k l A_x = -\gg_o J_x \quad (3.130a)$$

$$V' A_z + e_r(y) k l A_2 = -\wedge J_z \quad (3.130b)$$

while the y component becomes

$$\begin{aligned} V \% + \epsilon_y(y) A_g A_v &= -j o n_0 e_0 * \frac{\partial}{\partial y} (W - e_r) * y \\ &= Ju/J - o e_0 (6_v - e_r) \frac{d<P>}{tly} + <t>(H)(e_r - l)S(y - H) \end{aligned} \quad (3.131)$$

where $-(e_r - l)S(y - H)$ expresses the derivative of the step change that occurs in e_r as y crosses the interface at H , that is, $e_r(y)$ changes from e_r to unity.

The equations for A_x and A_2 are of the same form as derived in Sec. 2.10, but the equation for A_v is new. The equation for A_v is coupled to the scalar potential ϕ at the boundary even if we have an isotropic substrate. Thus boundary conditions require the presence of an A_v component even though there is no y component of current.

A separate equation for the scalar potential is obtained by using Gauss' law $\nabla \cdot D = \rho$ and the Lorentz condition (3.129). Thus we find the following equations:

$$\begin{aligned} \nabla \cdot D &= \nabla \cdot -e_r e_0 (j\omega A_y + V) - e_0 (e_v - e_r) j \frac{\partial a_v A_y}{\partial y} + a_v \frac{\partial}{\partial y} \\ &= -\epsilon_0 j \omega V - (e_r A_y) + V \cdot (-e_r V^*) + j \omega (e_y - e_r) A_y \\ &= -\epsilon_0 M^* J \cdot A + A \cdot V_{cr} + V \cdot (-e_r V^*) + j \omega (e_y - e_r) A_y \end{aligned}$$

By replacing $\nabla \cdot A$ with $-j\omega \epsilon_0 / j\omega e_r$ from the Lorentz condition, we obtain

$$= -\rightarrow \wedge \text{if} - \rightarrow W \wedge - C_r) \wedge - P_-$$

The last step is to simplify this equation using

$$\begin{aligned} \frac{\partial}{\partial y} (e_v - e_r) A_y &= (e_v - e_r) \frac{dA_y}{dy} + A_v \frac{\partial e_r}{\partial y} - A_r \frac{\partial e_v}{\partial y} \\ \text{and } V \cdot e_r V &= e_r V^2 P + \frac{f \epsilon_r d \omega}{f l y} \end{aligned}$$

By using these expressions a number of terms cancel and we obtain the final form

$$\begin{aligned} \frac{f^2 \omega^2 S}{dx'} + \frac{f l \omega^2}{f l z'} - \frac{C_y}{f l y} + 6 \pi f c_0^2 \epsilon_0^* &= - \frac{f l \omega^2}{f l z'} + (e_y - e_r) \frac{d^2 A_y}{dy^2} \\ &= - \frac{f l \omega^2}{f l z'} + (e_y - e_r) A_y (H) 5(y - H) - j \omega (e_y - e_r) \frac{d^2 A_y}{dy^2} \end{aligned} \quad (3.132)$$

where we have also used

$$A > T = -(e_y - e_r) A_y (H) 8(y - H)$$

This equation also displays a coupling between the scalar potential and $A_y(H)$ at the interface as well as coupling within the substrate whenever ϵ_y does not equal ϵ_r , that is, for anisotropic substrates. In the air region $y > H$, both ϵ_r and ϵ_v are replaced by unity in (3.130) to (3.132).

After the above lengthy derivation we can now obtain simpler equations to be solved in each separate region along with boundary conditions to use in joining the solutions at the interface $y = H$. The source terms p and J when expressed in the form (3.128) contain the $\delta(y - H)$ factor. In order for the left-hand sides of (3.130) to (3.132) to equal the corresponding right-hand sides, we must obtain a delta function $\delta(y - H)$ from the derivative of the potentials with respect to y . In (3.130a) and (3.130e) this requires that dA_x/dy and dA_z/dy have a step change at the interface so that the second derivative with respect to y will produce a delta function. The required step change can be found by integrating both sides of the equation over a vanishingly small interval centered on $y = H$. The integral of terms not involving a derivative with respect to y will vanish since these terms must be continuous at $y = H$ and the interval length vanishes. For example, if A_s were not continuous at $y = H$, the second derivative with respect to y would generate a singular term corresponding to the derivative of the delta function and no such term exists on the right-hand side of the equation. Thus from (3.130a) we obtain

$$\lim_{\substack{y \\ \rightarrow 0}} / \frac{\mathbf{r}^{\text{ff}} + \mathbf{r}^{\text{rtt}}, \mathbf{A}}{H^+ - V^-} dy = \frac{d\mathbf{A}}{dy} \Big|_{H^-}^{H^+}$$

$$\lim_{\substack{y \\ \rightarrow 0}} rM^+ - n_0 J_{sx}(x, z) \delta(y - H) dy = -IM_0 J_S$$

or

$$\left. \frac{W}{H^-} \right|_{H^+} = -\mathbf{p} \& \mathbf{J}^* \quad (3.133a)$$

In a similar way we obtain

$$\left. \frac{dy}{n} \right|_n = -Mo^s \quad (3.1336)$$

The notation H^+ and H^- means evaluating the derivative on adjacent sides of the interface at $y = H$. These two equations state that the tangential components of the magnetic field must be discontinuous across the current sheet J_s since from the equation $\mathbf{B} = \mathbf{V} \times \mathbf{A}$:

$$17 = *Hx = -Mo^s$$

In a similar way we obtain the following boundary conditions by integrating

(3.131) and (3.132) about a small interval centered on $y = H$:

$$\frac{dA'''}{dy} = y \epsilon_0 \epsilon_r (\epsilon_r - 1) \langle t \rangle (H) \quad (3.133c)$$

$$\frac{d^2A_y}{dy^2} = -f_o + M \epsilon_r \langle t \rangle A_y (H) \quad (3.133d)$$

A term such as $(\epsilon_v - \epsilon_r) d \langle P \rangle / dy$ that occurs in (3.131), and a similar term occurring in (3.132), does not contribute because

$$\lim_{y \rightarrow H^-} (\epsilon_v - \epsilon_r) - dy = \lim_{y \rightarrow H^+} (\epsilon_v - \epsilon_r) - dy = 0$$

since $d \langle J \rangle / dy$ is continuous in the interval $H^- < y < H$ and for $y > H$ we have $\epsilon_v = \epsilon_r = 1$. The boundary conditions on $\langle t \rangle$ reflect the fact that the total y -directed electric field has a contribution from A_y so that the discontinuity in D_y across the charge layer is given by

$$-\epsilon_v \frac{dA_y}{dy} + \int_{y^-}^{y^+} i \omega A_y dy + \epsilon_v \epsilon_0 \frac{d^2A_y}{dy^2} \Big|_{y^-}^{y^+} = P$$

which is (3.133d).

By using the above boundary conditions, we can solve (3.130) to (3.132) in each respective region away from the interface. Thus we need only to solve the following homogeneous equations, subject to the specified boundary conditions, in the substrate region:

$$(v^2 + k^2) A_x = 0 \quad (3.134a)$$

$$(3.1346)$$

$$V^2 A_y + e_v k^2 A_y = j(o \mu_0 \epsilon_0 (\epsilon_v - \epsilon_r)) \frac{dA_y}{dy} \quad (3.134c)$$

$$\frac{\partial^2 A_y}{\partial z^2} + \frac{d^2 A_y}{dz^2} = e_v \epsilon_0 \epsilon_r V^2 A_y = -X(v - \epsilon_r) \frac{dA_y}{dy} \quad (3.134a'')$$

In the air region the equations to be solved are obtained by setting $\epsilon_r = 1$. There is no volume coupling between A_x and A_y in the air region or in an isotropic substrate region. Since we are interested in wave solutions representing waves propagating in the z direction, we can assume that the dependence is e^{jktz} . The second derivative with respect to z can then be replaced by $-/3^2$. The common factor e^{-ipz} can be deleted from the equations just as e^{-imz} was dropped for convenience.

Low-Frequency Solutions

We can obtain the equations to be solved in the zero-frequency limit by assuming that the potentials and the source terms can be expanded as

power series in to . Thus we let

$$A = A^{\circ} + \alpha A^1 + \alpha^2 A^2 + \dots \quad (3.135a)$$

$$\Phi = \Phi^{\circ} + \Phi^1 + to^2 \Phi^2 + \dots \quad (3.135b)$$

$$J = J^{\circ} + w J^1 + a r J^2 + \dots \quad (3.135c)$$

$$P = p^{\circ} + t o P^1 + w' Y + \dots \quad (3.135d)$$

The parameter $k_0 = io^2 p_o e_0$ is of second order in to . The propagation constant (5) can be expressed in the form $fi = \hat{k}_0$, where e_e is a frequency-dependent effective dielectric constant. Consequently, p^2 is also of second order in to .

We now substitute these power-series expansions into (3.130) to (3.132) and equate all zero-order terms to obtain the following lowest-order equations:

$$(a? + \hat{A}_x^{\circ} - / *_0 J,^{\circ}) \quad (3.136a)$$

$$\hat{\frac{Z}{dx}}^{\circ} + \hat{\frac{A}{dy}}^l = -n_0 J? \quad (3.136b)$$

$$P = a^2 ,$$

$$\frac{d^2}{r^2} \left(\frac{a}{r} \right) \left(\frac{a}{r} \right) \left(\frac{p^{\circ}}{e_0} \right) = -\frac{1}{r^2} \frac{d^2}{r^2} \left(\frac{e_r}{r} \right) \left(\frac{e_v}{r} \right) \left(\frac{J^{\circ}}{e_0} \right) \quad (3.136d)$$

Further information is obtained from the Lorentz condition (3.129) which gives

$$\frac{3A^{\circ}}{dx} + \frac{M^{\circ}}{dy} = 0 \quad (3.137a)$$

$$-MAI = ->^o * r^* \circ \quad (3.137b)$$

In the air region e_r and e_v are set equal to unity. From the continuity equation relating current and charge, namely,

$$V \cdot J = -j c o p$$

we obtain

$$= 0 \quad (3.138a)$$

$$-jpj? = -jwp^{\circ} \quad (3.138b)$$



In the above equations the e^{-jitz} factor is not included but any derivative with respect to z was replaced by $-j/5$.

Since J^o must be zero at the edges $x = \pm W/2$ of the microstrip we conclude that J^o is zero because the integral of dJ^o/dx is at most constant. Hence, to lowest order, there is no x-directed current on the microstrip and A'' is zero. The Lorentz condition then requires that $dA^o/az = 0$ and hence $A^o = 0$ also since a constant A^o is a trivial solution and would not produce any magnetic field contribution. Thus, to lowest order we only have to solve for an A^o and a scalar potential ϕ^o . If we assume the microstrip to be at a potential V , then the boundary condition on ϕ^o is that it equals V on the microstrip and equals zero on the ground plane.

We can integrate the continuity equation (3.1386) across the microstrip line to get

$$0// = "Q^o \quad (3.139)$$

where γ_z^o is the total 2-directed current on the microstrip and Q^o is the total charge. On the microstrip the axial electric field must be zero. To lowest order this boundary condition is

$$El = ->A^o - \frac{\partial \phi^o}{az} = -ju>A^o + jp< b^o = 0$$

$$\text{or} \quad wA^o = pV \quad (3.140)$$

Hence A^o is also constant on the microstrip.

We will show shortly that the inductance L per unit length of the microstrip line is given by the equation

$$I?L = A_z^o \quad (3.141)$$

The capacitance per unit length is given by

$$C = — \quad (3.142)$$

By using these expressions to eliminate Q^o in (3.139) and A'' in (3.140), we obtain the pair of equations

$$pi? = coCV \quad (3.143c)$$

$$\gg U,,^o = PV \quad (3.143d)$$

from which we find that

$$p^2 = u>^2LC \quad (3.1440)$$

$$h^\bullet \quad \wedge \sim \quad (3.1446)$$

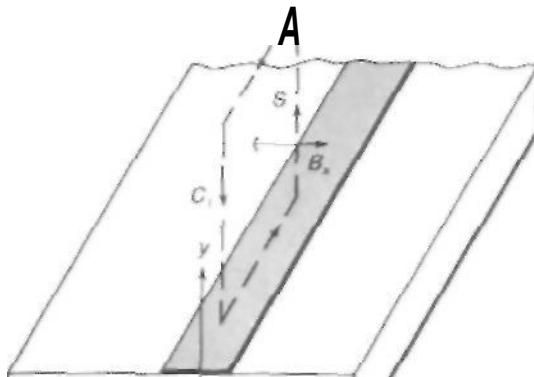


FIGURE 3.21
Surface used to find the magnetic flux linkage in a microstrip line.

We have thus been able to show in a rigorous way that in the low-frequency limit the microstrip line can be analyzed as a static field problem and that its propagation constant and characteristic impedance are determined by the low-frequency distributed capacitance and inductance. The analysis leading up to (3.144) is quite general and applies to other planar transmission lines as well.

At this point we return to the promised derivation of (3.141) giving the line inductance. With reference to Fig. 3.21 we note that the magnetic flux ip linking the microstrip per unit length is given by the integral of B_x over the area extending from the microstrip to infinity. Thus

$$ip = \int I \cdot B \cdot a_x dy dz$$

By using $B = v^* \times A$ and Stokes' law, we can write

$$\begin{aligned} ip &= \int_0^f \int_H V \times A \cdot a, dy dz \\ &= 6|d| \end{aligned}$$

where C_1 is the boundary of the area. Since A'' is zero and A° is zero at infinity and is constant on the microstrip, we obtain $\oint = A^\circ_z$ for the flux linkage. The inductance is given by $i/\omega f$ and this gives (3.141).

We can also derive equations for the next level of approximation. However, the solution of these equations is not much easier than the solution of the original equations; so it is not worthwhile to develop the power-series solutions beyond the lowest order. Thus the equations to be

solved are

$$\overset{;>^2}{\mathbf{a}^*} \overset{d^2}{\mathbf{dp}} \mathbf{A}^* = 0 \quad y < H, y > H \quad (3.145aj)$$

$$\frac{dx^2}{dx^2} \quad e_r \frac{dy^2}{dy^2} \quad y < H \quad (3.1456)$$

$$\overset{*^2}{\mathbf{g}x^*} \quad \overset{*^2}{\mathbf{a}_y^2} \mid " \quad y > H \quad (3.145c)$$

with the boundary conditions

$$\overset{ff}{\mathbf{M}^{\circ}} = -\infty \quad (3.145d)$$

$$\overset{P<}{\mathbf{a}^{*0}} \quad \overset{-\epsilon_r}{\mathbf{dy}} \quad \overset{y}{\gg} \quad (3.145e)$$

$$\Delta^{\circ} = V \quad \text{on microstrip}$$

$$A^{\circ} = \text{constant} \quad \text{on microstrip}$$

Along the interface and away from the microstrip, the right-hand sides of (3.145f) and (3.145e) are zero. In addition, Δ° and A_z° must be zero on the ground plane in order to make the tangential electric field vanish on this surface.

The equations for A'' do not depend on the dielectric constants of the substrate material. Hence the line inductance is the same as for an air-filled line. But for an air-filled transmission line with distributed capacitance C_a , we have

$$\sqrt{LC_a} = y > oe^{\lambda} \quad \text{and hence} \quad L =$$

so we can find L by finding the distributed capacitance of an air-filled microstrip line. By introducing C_a in place of L , the solutions for $/3$ and $\&c$ can be expressed in the form

$$/ \wedge \mathbf{w} \mathbf{i} \mathbf{c} / - k_0 = j r_e k_0 \quad (3.146a)$$

$$= 1 / \mathbf{c} : \quad \sqrt{\frac{L}{C_a}} \sqrt{\frac{C^*}{C}} \quad \cdot \cdot \cdot \quad (3.1466)$$

where Z_{c0} is the characteristic impedance of the air-filled fine and the ratio C/C_a gives the low-frequency equivalent (effective) dielectric constant e_{eq}

The effect of having an anisotropic dielectric substrate does not add any additional complication. If we introduce a new variable $u = (e_r/e_y)^{1/2}y$, then upon using

$$S \frac{d}{du} \frac{du}{dy} = \sqrt{\frac{I e_s}{\epsilon_y}} \frac{d}{du}$$

we find that (3.1456) reduces to

$$\frac{d^2}{dx^* du^l} \quad (3.147)$$

When $y = H$ the corresponding value of u is $U_r/e_y^{1/2}H$; so the solution of (3.147) is that for a microstrip with an equivalent substrate thickness H_v given by $(\epsilon_r/\epsilon_y)^{1/2}H = H_v$. The boundary condition (3.145e) becomes

$$h_{H^+} \leftarrow \epsilon_y \sqrt{\frac{\epsilon_r}{\epsilon_y}} \frac{\partial \Phi^0}{\partial u} \Big|_{H_v} \quad d< p^0 \quad 9y_a - y^A S \frac{dS^0}{du} \Big|_{H_v} = - - \quad (3.148)$$

which shows that the equivalent dielectric constant of the substrate should be taken as the geometric mean $e_g = \sqrt{e_r e_y}$. Thus, by modifying the substrate thickness and introducing the equivalent dielectric constant, the solution for the distributed capacitance C for the case of an anisotropic substrate can be reduced to that for an isotropic substrate. The distributed capacitance C_a is that for the unsealed microstrip line.

The unit of length does not enter directly into the differential equations for the potentials. Thus x and y can be in units of meters, centimeters, inches, or any other convenient unit. What this means is that the distributed capacitance and inductance per unit length is dependent only on the ratio of strip width to substrate thickness, i.e., on W/H . If we have found C and L for a given set of values for W and H on a per-meter basis, then if we change W to sW and H to sH , where s is a scaling factor, both C and L on a per-meter basis do not change. Hence the characteristic impedance, effective dielectric constant, and propagation constant β for any planar transmission line is invariant to a scaling of the cross-sectional dimensions. However, the attenuation caused by conductor loss does not scale since the series resistance is inversely proportional to the conductor widths. The attenuation due to conductor losses will double if the conductor size is reduced by a factor of 2. The scaling law is clearly illustrated for an ideal parallel-plate capacitor with a plate dimension of W^2 and separation H and having a capacitance $\epsilon_0 W^2 H / (2d)$. Clearly keeping the ratio W/H fixed keeps the capacitance unchanged.

A variety of methods exist for solving the two-dimensional Laplace equation (3.147). For planar transmission lines the conformal mapping method is widely used, generally along with some approximations that are necessary because of not having a dielectric medium filling all of the space

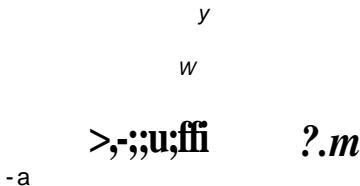


FIGURE 3.22

A microstrip line with perfectly conductive, side walls inserted at $x = \pm a$ with $ct \gg y$

around the conductors. A number of useful solutions obtained by conformal mapping methods are described in App. III. We will refer to some of these solutions as needed.

In order to illustrate the general method of solution, we will develop a Fourier series solution to (3.145) which will turn out to provide an efficient method to obtain the parameters of a microstrip transmission line. In order to use the Fourier series method, we place perfectly conducting (electric) walls at $x = \pm a$ as shown in Fig. 3.22. Provided a is chosen equal to 10VV or $10H$, whichever is larger, the sidewalls have a negligible effect on the field which is concentrated near the microstrip.

We can expand the unknown charge density p_s into a Fourier series of the form

$$P_s(x) = \sum_{n=1,3,\dots} P_n \cos \frac{n\pi x}{2a} \quad (3.149a)$$

The charge coefficients P_n are given by

$$P_n = \frac{1}{a} \int_{-W/2}^{W/2} p_s(x') \cos \frac{n\pi x'}{2a} dx' \quad (3.149b)$$

The charge density is an even function of x because of the symmetry involved; so only a cosine series is needed. The functions are chosen so that they vanish at $x = \pm a$, a required boundary condition for the potential; so only odd integers n are used.

The potential $\phi(x,y)$ can also be expanded into Fourier series; so we let

$$\begin{aligned} \phi(x,y) &= \sum_{n=1,3,\dots} f_n(y) \cos \frac{n\pi x}{2a} & y > H_e \\ &= \sum_{n=1,3,\dots} g_n(y) \cos \frac{n\pi x}{2a} & 0 < y < H_e \end{aligned} \quad (3.150)$$

where $f_n(y)$ and $g_n(y)$ are to be found. We are using an effective substrate thickness H_e , so that an anisotropic substrate can be accommodated. We have dropped the superscript 0 since it is understood from the context that we are solving for a static potential field.

Each Fourier term in the expansion of Φ must be a solution of Laplace's equation. Hence we require that

$$\lambda^2 + \lambda^2 w f_B(y) \cos \lambda = 0$$

which gives

$$\frac{df_n(y)}{y} - f_n(y) = 0$$

$$- \lambda - - \lambda g_n(y) = 0$$

where $\lambda^2 = (nv/2a)^2$. In the region $y < H_e$, a suitable solution for $g_n(y)$ that vanishes on the ground plane is

$$g_n(y) = C_n \sin n\pi y/H_e$$

where C_n is an unknown constant. In the region $y > H_e$, we need a solution that will vanish as y approaches infinity; so we choose

$$f_n(y) = D_n e^{-ny/H_e}$$

where D_n is another unknown constant. At $y = H_e$, the two potential functions must match; so we have

$$C_n s m h_n H_e = D_n e^{-nH_e/H_e}$$

The Fourier series expansion of the charge density p_s represents this charge density as sheets of charge $p_s \cos n\pi y/H_e$ that extend from $x = -a$ to $x = a$. By superimposing an infinite number of such charge sheets, we obtain a charge density p_s that is nonzero only on the microstrip $-W/2 < x < W/2$. The boundary condition (3.148) is applied to each Fourier term to obtain

$$\begin{aligned} df_n &= dg_n \\ H_e & \quad \quad \quad \end{aligned}$$

$$\text{or } -w_n D_n e^{-w_n y/H_e} - e_g w_n C_n \cosh w_n y/H_e \sim 0$$

We now have two equations which we can solve to find C_n and D_n . The solutions are

$$\begin{aligned} C_n &= \frac{e_0 u_{n0} (\sinh w_n H_e + e_g \cosh w_n H_e)}{w_n^2 - e_g^2} \\ D_n &= \frac{e_0 u_{n0} (\sinh w_n H_e + e_g \cosh w_n H_e)}{w_n^2 - e_g^2} \end{aligned}$$

where $e_g = J_s^* e_s$.

We now substitute our solutions for f_n and g_n into (3.150) and (3.1496) for p_{ss} . Thus we obtain

$$\begin{aligned} *(*,y) - \int_{-W/2}^{W/2} & \left[\cos w_n x \cos u_n x' \right. \\ & + e_g \cosh w_n H_e) \\ & \times \sinh /^n y \\ & \left. \times \text{siimi}_{n=1}^{\infty} i 2 e^{-\lambda f_{ir3}} P_s(x') dx' \right] \end{aligned} \quad (3.151)$$

where the upper term is for $y < H_e$ and the lower one is for $y > H_e$. The factor multiplying $P_s(x')$ under the integral sign represents the Green's function for this problem. It is designated by the symbol $G(x,y;x',y')$ so in abbreviated form we express (3.151) as

$$*(*,?) - \int_{-W/2}^{W/2} G(x,y;x',H_e) p_s(x') dx' \quad (3.152)$$

The last boundary condition to be imposed is the requirement that $\Phi = V$ on the microstrip; thus

$$V = \int_{-W/2}^{W/2} G(x,H_e;x;H_e) p_s(x') dx' \quad - \leq x < \frac{W}{2} \quad (3.153)$$

This is an integral equation whose solution would determine the unknown charge density $p_s(x')$. Once we know the charge density, we can calculate the total charge on the microstrip using

$$\int_{-W/2}^{W/2}$$

and find $C = Q/V$.

Integral equations are most often not solvable by analytic means. However, various numerical schemes exist for obtaining good approximate solutions. The most popular method is the *Method of Moments*. In "ris" method the first step is to choose a finite number of basis functions and to expand $p_s(x')$ in terms of these in the form

$$N \quad W \quad W$$

where Q_n are unknown coefficients. The basis functions could be the unit height pulse functions shown in Fig. 3.23, the cosine functions $\cos 2\pi nrx'$ or any other reasonable set that would give a good approximation to p_M

tfi

*N

W
2W
2

FIGURE 3.23

Unit height pulse functions for expanding the charge density.

When this substitution is made in (3.153), we obtain

$$V - \frac{N}{L} Q_n G_n(x) = \frac{W}{2} < * < \frac{W}{2} \quad (3.154)$$

where

$$G_n(x) = \int_{-W/2}^x G(x, H_i; x', H_i) *_n(*) dx'$$

The next step is to convert (3.154) into a matrix equation for the unknowns either by making both sides of the equation equal at N points in x along the microstrip, or by using weighting functions to make N weighted averages of both sides equal. We can choose the $\langle l \rangle_m(x)$ as weighting functions, in which case the method is called Galerkin's method. Other choices for the weighting functions can also be made. If we use Galerkin's method, we obtain

$$\mathbf{f} \cdot G_{nm} Q_a = V_m \quad m = 1, 2, \dots, N \quad (3.155)$$

where the matrix elements are given by

$$\begin{aligned} G_{nm} &= \int_{-W/2}^{W/2} G_n(x) \mathbf{4}_{m,n}(x) dx \\ &= \int_{-W/2}^{W/2} G(x, x) U_x \langle l \rangle_m(*) dx \end{aligned}$$

and the components V_m of the source vector are given by

$$V_m = \int_{-W/2}^{W/2} (\mathbf{V}_m \cdot \mathbf{4}_m(x)) dx \quad m = 1, 2, \dots, N$$

The system of linear equations given in (3.155) can be solved for the unknown charge amplitude coefficients Q_{ri} . If N is chosen sufficiently large, we will obtain a good approximation for the charge density.

If we know a priori how $p_r(x')$ is distributed on the strip, or a close approximation to it, we could get an excellent approximate solution using very few basis functions. From conformal mapping solutions we know that on an isolated infinitely long conducting strip of width W and with total

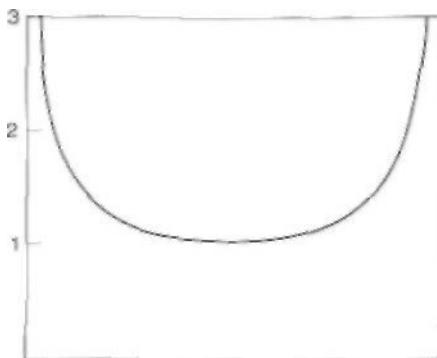


FIGURE 3.24
Charge density on an idea) isolated strip of width W .

charge Q per meter that the charge density is given by

$$P_s(x) = \frac{2Q}{\pi W y^7!} \frac{x^2/(W/2)}{(x^2/(W/2))'} \quad (3.156)$$

This charge distribution is illustrated in Fig. 3.24. At sharp corners and edges the charge density always exhibits an infinite behavior. However, the density is never so singular that it cannot be integrated. The reader can readily verify that if the substitution $x = (W/2)\sin \theta$, $dx = (W/2)\cos \theta d\theta$ is made, then

$$\int_{-w/2}^{w/2} P_s(x) dx' \wedge r'^2 d\theta = Q \quad I$$

In a microstrip line the charge density is influenced by the dielectric substrate and the ground plane but surprisingly (3.156) remains a good approximation. By using a two- or three-term polynomial along with (3.156), i.e., choosing

$$P_s(x) = \frac{Q_0 + Q_1 x^2 + Q_2 x^4}{1 - x^2/(W/2)^2} \quad (3.157)$$

an excellent approximate solution for $p_s(x)$ will be obtained.

We do not plan to solve the integral equation (3.153) numerically using the Method of Moments. It does require considerable numerical computation to evaluate the matrix elements because we only have the Green function expressed as an infinite Fourier series. What we are going to do is to express the integral equation in a form that we can interpret as representing the capacitances between a conducting strip *in air* above a ground plane with spacings H_e , $2H_e$, $3H_e$, etc. We can then make use of a conformal mapping solution for a pair of strips in air to evaluate the distributed capacitance C for the microstrip line.

Let us choose $e = 1$ but still retain the strip spacing above the ground plane as H_e . By using (3.151) in the integral equation (3.153) for this case, we obtain (note that $\sinh w_n H_r + \cosh w_n H_e = e^{w_n H_r}$)

$$Q' \frac{rW/2}{L} \frac{* \text{cosec } W_{nX} \cos W_{nX'}}{7} \frac{(l - e^{-*nH_e})}{(l - e^{-*nH_r})}$$

$$\int Q' dx' = V \quad (3.158)$$

where we have set $y = H_e$ and multiplied and divided by the total charge Q . The part multiplied by $-c^{-2} \text{cosec } W_{nX}$ represents the effect of the ground plane which is equivalent to that of the image strip at $y = -H_e$ and having a charge density $-p_s$. A solution of (3.158) provides the solution for the problem of a strip above a ground plane, as well as the solution to the problem of two parallel strips in air and separated by a distance $2H_p$. The capacitance per meter for a strip above a ground plane is given by

$$Q$$

$$C_a(H_e) = \sim$$

while that between two strips separated by a distance $2H_e$ is

We see that the integral, involving the normalized charge density, represents $1/CJHJ$.

Consider now the factor

$$\frac{\sinh w_n H_v}{\sinh w_n H_e + e_g \cosh w_n H_c}$$

that occurs in the Green's function in (3.151) at $y = H_e$. We can express this term in the form

$$\begin{aligned} & \frac{B'' \text{ft}}{(1 + e_g) e^w} - \frac{1}{(1 - e_j c)^w} \\ & (1 + e_g) e^w - (1 - e_j c)^w \end{aligned} \quad I - C - \frac{I}{e^w} + C$$

$$\frac{1 + e_g}{e_g} \left(1 - e^{-2 \pi n} \right) \left(1 - e^{-2 \pi n} \right)^{-1}$$

where $77 = (1 - e_{gg})/d + e_{gg}$. The last factor is now expanded into a power

series to obtain

$$\frac{1}{1 + e^{-2u''H_e}} \sum_{m=0}^{\infty} (-1)^m e^{-imw_l H_e} = Y_0 - \sum_{m=0}^{\infty} (-1)^{m+1} e^{-imw_l H_e}$$

We now add $-ij'''$ to both series which has no net effect because of the minus sign in front of the second series. This gives

$$\frac{1}{1 + e^{-2u''H_e}} \sum_{m=0}^{\infty} (-1)^m e^{-2u''mH_e} + \sum_{m=0}^{\infty} (-1)^{m+1} e^{-2u''(m+1)H_e}$$

The $m = 0$ term is zero in the first series so we can change m to $m + 1$ and still sum this for $m = 0, 1, 2, 3, \dots$ without changing its value; thus we get

$$\frac{i}{1 + e^{-2u''H_e}} \sum_{m=0}^{\infty} (-1)^{m+1} (1 - e^{-2(m+1)H_e}) + \sum_{m=0}^{\infty} (-1)^{m+1} (l - e^{-2(m+1)H_e}) \\ \cdot (n - 1) L * 7^m (1 - e^{-2mH_e}) \quad (3.159)$$

Upon using this expansion in the Green's function, the integral equation (3.153) can be expressed in the form

$$4e, rW/2 \cos w_n x \cos \phi, A \\ \cdot = 0 (1 + 6, y - 72, i, 3, \dots, 4e_0 w, a) \\ (1 - , - + - , *) \mathbf{M} \mathbf{i} \mathbf{l} (\&, \mathbf{V}) \quad (3.160)$$

We note that the m th term considered by itself is an integral equation of the same form as that in (3.158) apart from the multiplying factor $4e^{77m}/(1 + e_g)^2$. This integral equation would provide a solution for the capacitance of a strip in air spaced a distance $(m + l)H_e$ above the ground plane. If we assumed that the normalized charge density $p_s(x')/Q$ was the same for any strip, independent of the spacing above the ground plane, each integral would produce a constant voltage V_m but different from V . With increased strip spacing and constant total charge on each strip, the integral has to give a larger voltage since $Q = C_{al}(m + l)H_e V_m$ and $C_{al}(m + 1) > C_{al}(m)$ the capacitance between the strip and the ground plane, decreases with increasing m . The approximation that the charge density is the same independent of strip-ground-plane spacing is a necessary one to make sense there is only one charge density expression in the integral equation. Then

approximation is a good one and by using it we can express (3.160) in the form

since by our assumption

$$\frac{W/2}{\epsilon - W/2, \approx 1.3} \cdot \frac{\epsilon \cos \theta \cos \phi}{4 \pi H C} = \frac{\epsilon \cos \theta \cos \phi}{4 \pi H C}$$

$$Q = \int_{-W/2}^{W/2} V(x) dx = C_a [(m + 1) f_{ce}]$$

From (3.161) we now obtain the following solution for the distributed capacitance $C = Q/V$:

$$C = \frac{1}{1 + CJLH} \sum_{m=1}^M \frac{J_m}{x_m} j_m = \frac{C_D [(m + 1) f_{ce}]}{1 + CJLH} \quad (3.162)$$

where M is the number of terms that are kept. Since $-q$ is negative the series is an alternating one. Typically, from 10 to 40 terms are needed for good accuracy. The evaluation can be done on a computer very quickly and requires only a simple program to implement. However, we do require an expression for the capacitance between a strip in air as a function of the spacing above the ground plane, which is given below.

The exact solution for the capacitance between a strip of width $2W$ and a distance H above a ground plane is given in App. III, along with tabulated values as a function of $2W/H$. For practical applications it is desirable to have simple formulas that will enable the capacitance to be evaluated with an accuracy of 1 percent or better. A number of investigators have proposed such formulas which are based on approximate analytic solutions, along with empirical adjustment of various numerical constants so as to achieve the desired accuracy. The following formulas give excellent results for the

H. A. Wheeler, Transmission Line Properties of a Strip on a Dielectric Sheet on a Plane, *IEEE Trans.*, vol. MTT-25, pp. 631-647, August, 1977.

E. O. Hammerstad, Accurate Models for Microstrip Computer-Aided Design, *IEEE MTT-S Int. Microwave Symp. Dig.*, pp. 407-409, 1980.

E. O. Hammerstad, Equations for Microstrip Circuit Design, *Proc. European Micro. Conf. Hamburg, W. Germany*, pp. 268-272, September, 1975. Equations (3.163) and (3.166) come from this publication.

S. Y. Poh, W. C. Chew, and J. A. Kong, Approximate Formulas for Line Capacitance and Characteristic Impedance of Microstrip Line, *IEEE Trans.*, vol. MTT-29, pp. 135-142, February, 1981.

capacitance per meter of a strip of width W at a height H above a plane and with air dielectric:

$$\mathbf{C} = \frac{2776,}{TBH} \ln \frac{W}{w} \quad H * 1 \quad (3.163a)$$

$$\mathbf{C} = \frac{W}{H} + 1.393 + 0.667 \ln \frac{(W)}{\sqrt{H}} + 1.444 \quad H > 1 \quad (3.1636)$$

These formulas give results that agree with those tabulated in App. m *. within 1/4 percent.

The effect of finite thickness T for the microstrip on the distributed capacitance is normally negligible. If necessary the effect of finite thickness can be included by using an effective width W_e , where W_e is given by the following expressions due to Gunston and Weale:t

$$W_e = W + 0.398T \left| 1 + \ln \frac{4\pi T W}{T} \right| \quad \begin{cases} W & < 1 \\ H & \end{cases}$$

$$= W + 0.398: r | 1 + \ln \frac{2H}{W} \quad \begin{cases} W & > 1 \\ H & \end{cases}$$

The above expressions can be used in (3.162) to evaluate the capacitance C for a microstrip line with an isotropic or anisotropic dielectric substrate. The effective dielectric constant e_e for a microstrip line is given by

$$\frac{C_a}{e_l} \quad (3.164)$$

where C_a is the capacitance of the unsealed air-filled line. The characteristic impedance is given by

$$\boxed{1/5- \frac{i^{\lambda} V_{Mo} \times}{e, C,}} \quad (3.165)$$

Even though computations based on (3.162) are straightforward, there is an easier way to find e_e . Schneider has presented a remarkably simp' formula for the effective dielectric constant of a microstrip line with an isotropic substrate.¹ This formula was modified by Hammerstad to improve

¹M. A. R. Gunston and J. R. Weale, Variation of Microstrip Impedance with Strip Thickness. *Electron. Lett.*, vol. 5, p. 697, 1969.

²M. V. Schneider, Microstrip Lines for Microwave Integrated Circuits, *Bell System Tech. J.* vol. 48, pp. 1422-1444, 1969. Schneider's formula has a numerical coefficient of 10 instead of 12 multiplying the H/W term.

the accuracy. The modified formula, which we will refer to as the S-H formula, is

$$+ 1 - e - \sqrt{(\dots) 12H^{-1/2} \frac{T}{\sqrt{WH}}} \quad (3.166)$$

where $F(e_r, H) = 0.02(e_r - 1)(1 - W/H)^2$ for $W/H < 1$ and equals zero for $W/H > 1$. The last term accounts for the reduction in e_e caused by the finite thickness of the microstrip. We have checked the accuracy of this formula against the results obtained by solving the integral equation and found the agreement to be better than 1 percent for $0.25 < W/H < 6$ and $1 < e_r < 16$. We can also adapt the S-H formula to treat the case of an anisotropic substrate as follows: For an anisotropic substrate we replace the spacing parameter H by the effective spacing H_e given by

$$H_e = \sqrt{\frac{\epsilon_r}{\epsilon_y}} H$$

and use the geometric mean $e_g = \sqrt{\epsilon_r \epsilon_y}$ for the dielectric constant e_g . The S-H formula gives the capacitance of a microstrip line with spacing H_e and dielectric constant e_g relative to an air-filled line with the same spacing H_e ; thus

$$C(e_g, H_e) = e + 1 - e_{\infty} - 1 / H_e^{1/2}, \quad tT v$$

The effective dielectric constant is, however, given by the ratio of $C(e_g, H_e)$ to the capacitance of the air-filled unsealed line according to (3.164). Hence, for the case of an anisotropic substrate, we have

$$\epsilon_{\infty} = \frac{e + 1 - e_g - 1}{1 + 12 - \frac{H A^{-1}}{+ F(c, ff.)}} \quad m**>$$

The capacitances for the air-filled lines are readily computed using (3.163).

For comparison purposes, typical results obtained for e_e using (3.162), formulas (3.166) or (3.167), and those obtained by solving the integral equation are given in Table 3.3. Also listed are the number of terms needed in the formula (3.162) to give a numerical convergence of 0.3 percent. Overall, all three methods give values for e_e that are in close agreement. Equation (3.162) gives values that are on the high side for wide strips and substrates with large dielectric constants. This is caused by the variation in charge density with strip spacing above the ground plane, which is more

TABLE 3.3
Comparison of values for effective dielectric constant using different formulas

W/H	Integral equation		Integral equation			$e_r = 5.12, e_v = 3.4t$
	Eq. (3.162)	Eq. (3.166)	Eq. (3.162)	Eq. (3.166)		
$\epsilon_r = 2$						
0.25	1.588	1.589 an	1.583	2.671	2.675(15)	2.69
0.5	1.61	1.612(7)	1.605	2.694	2.698(15)	2.721
1	1.645	1.649 (7)	1.639	2.731	2.734(15)	2.731
2	1.696	1.699(7)	1.689	2.797	2.802 (16)	2.794
.	1.762	1.761 (8)	1.75	2.906	2.929 (16)	2.890
6	1.801	1.799 (8)	1.789	2.979	2.944 (16)	2.963
$\epsilon_r = 6$						
$\epsilon_r = 10$						
0.25	3.896	3.896 (22)	3.913	6.195	6.192(36)	6.244
0.5	4.003	4.004 (22)	4.025	6.387	6.40 (37)	6.445
v	4.173	4.169(22)	4.193	6.69	6.70 (37)	6.748
2	4.428	4.447 (23)	4.444	7.15	7.16(38)	7.201
4	4.763	4.837 (23)	4.75	7.757	7.946 (39)	7.750
6	4.966	5.012(23)	4.943	8.127	8.303 (40)	8.098

tBoron nitride.

jThe numbers in parentheses are the number of terms used in the numerical solution.

pronounced for wide strips. From the tabulated results it can be inferred that the modified Schneider's formula will be acceptable for most applications since even the dielectric constant of the substrate is often not known to an accuracy much better than 1 percent. The following two examples will illustrate the application of the above formulas to microstrip lines.

Example 3.4. A microstrip line uses a substrate with dielectric constant $\epsilon_r = 9.7$ (alumina) and thickness 0.5 mm. The strip width is also 0.5 mm. We wish to find the effective dielectric constant, the characteristic impedance, and the microstrip wavelength at a frequency of 2 GHz.

Since the substrate is isotropic, we use (3.166) to find ϵ_e . Thus since $W/H = 1$,

$$10.7 \sqrt{8.7} (1 + 12)^{-1/2} = 6.556$$

In order to evaluate Z_c using (3.165), we first find C_a using (3.163). Thus

$$\epsilon_0 m(8 + |) = 2.978$$

From (3.165) we get

$$\frac{120\pi}{2.978\sqrt{8.7}} = \frac{49.44}{2.978\sqrt{6.556}}$$

At 2 GHz the propagation constant is $\beta = \sqrt{2\pi f \mu_0} = 1.0725 \text{ rad/cm}$. Hence $A = 2\pi/\beta = 5.858 \text{ cm}$. The wavelength can also be found using $\lambda = A_0/\beta = 5.858/\sqrt{2\pi f \mu_0} = 0.5 \text{ mm}$.

Example 3.5. A microstrip line uses a sapphire substrate 1 mm thick and having $\epsilon_r = 9.4$, $\epsilon_y = 11.6$. We want to find the effective dielectric constant and characteristic impedance for the case when the strip is 0.5 mm wide.

Since this substrate is anisotropic, we first find $e_g = J_e \epsilon_y = 10.44$ and $H_e = y_j e_g / \sqrt{\epsilon_y} = 9.44 \text{ mT}$. We now use (3.167) to find e_s . By using (3.163a) we get the ratio

$$\frac{C_a(H_e)}{C_a(B)} = \frac{2.347}{2.26} = 1.0385$$

From (3.167)

$$e_s = 1.0385 \cdot \frac{11.44}{2} \cdot \frac{9.44}{1 + 12 \cdot \frac{0.9}{0.5}} + 0.02 \times 9.44 \times \frac{0.9}{0.5} = 7.02$$

From (3.163a) we get $CJ.H = 2.26 \epsilon_0$ and using (3.165) gives

$$Z = \frac{120\pi}{2.26/0.2} = 62.96 \text{ n}$$

Microstrip Attenuation

Dielectric losses and conductor losses will introduce attenuation. The attenuation caused by the finite conductivity of the conductors is accounted for by the series resistance R , while attenuation caused by dielectric loss is modeled by the shunt conductance G in the distributed circuit model of the microstrip line. The separate attenuation constants are given by

$$a_d = \frac{G}{2Z_C} \quad (3.168a)$$

$$a_c = \frac{R}{2Z_C} \quad (3.168b)$$

and the total attenuation is given by

$$a = a_c + a_d \quad (3.168c)$$

The attenuation in decibels per unit length is obtained by multiplying a by 8.686.

We will first examine the attenuation caused by dielectric loss. The dielectric loss arises when the permittivity ϵ is complex, that is, $\epsilon = \epsilon' - j\epsilon''$. The loss tangent

$$\tan \delta = \frac{\epsilon''}{\epsilon'} = 81$$

is the usual given parameter for a dielectric material. Maxwell's equation

$$\mathbf{V} \times \mathbf{H} = j\omega eE + \sigma \mathbf{E} = \mathbf{y} \cdot \mathbf{W} \mathbf{E} + (\epsilon_r \omega^2 + \sigma) \mathbf{E}$$

shows that $\epsilon_r \omega^2$ can be viewed as the effective conductivity of a dielectric when the ohmic conductivity σ is zero. Normally we can assume that $\sigma = 0$ except for a semiconductor substrate, in which case σ will depend on the doping level.

The electric energy stored in the substrate region of the microstrip line is given by

$$W_{el} = \frac{1}{4} \int_V \mathbf{E} \cdot \mathbf{E}^* dV$$

where V_l is the volume of the substrate region per unit length of line. The power loss due to dielectric loss is given by

$$P_d = \frac{\epsilon_r \omega^2}{2} \int_V \mathbf{E} \cdot \mathbf{E}^* dV$$

Thus we see that

If the dielectric filled all of the space around the microstrip, we could equate W_{el} to $CV^2/4$ and P_d to $GV^2/2$ and thereby obtain

$$G = \frac{we}{\epsilon} C$$

However, for a partially filled line some of the electric energy is located in the air region that occupies a volume we will call V_2 . Consequently, we have

$$W_{el} + W_{e2} = \int_V \mathbf{E} \cdot \mathbf{E}^* dV + \int_{V_2} \mathbf{E} \cdot \mathbf{E}^* dV = \frac{CV^2}{4}$$

If we had an air-filled line, we could write

$$Wei + We_2 = \frac{e}{4} \int_V \mathbf{E} \cdot \mathbf{E}^* dV + \frac{e}{4} \int_{V_2} \mathbf{E} \cdot \mathbf{E}^* dV = \frac{CV^2}{4}$$

with the understanding that the electric field in the two cases will not be the same. There is no simple exact way to determine how the electric energy is split between the two regions. There is, however, an approximate method to find the division of the total energy between the two regions and based on the assumption that the volume integrals of $\mathbf{E} \cdot \mathbf{E}^*$ in the two cases are approximately the same. By making this assumption we can write

$$eh + e_0 I_2 = CV^2$$

$$e_0 I_1 + e_0 I_2 = C_a V^2$$

where I_x and I_2 represent the values of the integrals over V_1 and V_2 , respectively. The above two equations can be solved for $/_1$ and $/_2$ to give

$$e - e_0 \quad e \sim e_0$$

where $e'_r = e'/e_0$. The fraction of the total energy in the dielectric region is

$$<./, \quad e'_r(C - C_{,,}) \quad e'_r e, -1 \quad e'_r$$

where we have used $C = e_c C_{,,}$. The parameter α : is called the filling factor. If q was found independently, then we could solve for e_e to get $e_e = e'_r q + (1 - q)$. The parameter q is the ratio of the integral of $E \cdot E^*$ over the volume V_y to the integral over the total volume $V_1 + V_2$, that is, $q = /_1/(/_1 + I_2)$ and clearly represents a filling factor.

With the above assumption regarding no change in the volume integrals for the two cases, we see that G , as given earlier, should be reduced by the same fraction by which the total electric energy was split since there is essentially no loss in the air region. Hence an estimate for G is

$$G = \sim^J Z - r - C \quad (3.170)$$

By using $Z_e = yeJy^t_0e_0/C$ we obtain

$$a_d = \frac{GZ_C}{\sim^J} = \frac{e'_r e, - 1}{\tan S}, \quad (3.171)$$

for the attenuation constant due to dielectric loss. In the derivation we used $a > ijnoe_0 = k_0 = 27r/A_0$. As an example if $e'_r = 9.7$, $e_e \sim 6.55$, and $\tan S = 2 \times 10^{-4}$ we get $a_d = 1.52 \times 10^{-3}$ Np/wavelength. In decibel units this equals 0.013 dB/wavelength, which is a relatively small value. The attenuation caused by conductor losses will be significantly larger. Equation (3.171) is valid for isotropic substrates only.

We now turn our attention to evaluating the attenuation caused by finite conductivity of the microstrip and the ground plane. The continuity equation (3.138b) shows that the current density J along the conductors varies the same way as the charge density. Thus on a conducting strip of width W the current density will be similar to the charge density given by (3.156). At the edge of an infinitely thin strip, the current density will increase inversely proportional to the square root of the distance from the edge and becomes infinite at the edge. However, the density can be integrated to give a finite value for the total current. But since power-loss calculations require integrating the square of the current density, we would find that for the current density on an infinitely thin strip we would obtain infinite power loss. In practice, the conductors have a finite thickness and the current density is less singular at the edge and the power loss is finite.

Consequently, it is necessary to take into account the finite thickness of t_L conductors. In addition, it is necessary to determine how the total c_{ur} divides between the two faces of the microstrip since the presence /" ground plane results in an unequal division of the current for strip width greater than one-half of the spacing above the ground plane.

The current distribution, current division, and power loss can be evaluated using conformal mapping techniques. In order to obtain useful formulas, some approximations are necessary. The analysis for a microstrip line is carried out in App. III and the results obtained are repeated here (note that in App. III the strip width is $2w$ and the thickness is $2r$, where here we use W for the width and T for the thickness). The normalized series distributed resistance for the microstrip is R_l where

$$R_l = \frac{I}{LR} \left(1 + \frac{1}{m} \right) \frac{4\pi W}{T} \quad (3.172)$$

The loss ratio LR is given by

$$\text{LR} = \frac{W}{H} \frac{1}{\sqrt{H}} \frac{1}{W} \quad \text{IV} \quad \frac{W}{H} < 0.5 \quad (3.173)$$

$$\text{LR} = 0.94 + 0.132 \frac{W}{H} - 0.0062 \frac{1}{\sqrt{H}} \quad 0.5 < \frac{W}{H} < 10$$

The loss ratio gives the increase in resistance that results from an unequal division of the current. The normalized series resistance R_2 of the ground plane is given by

$$\frac{R_2}{R_m} = \frac{\frac{W}{H}}{\frac{W}{H} + 5.8 + 0.03H/W} \quad 0.1 < \frac{W}{H} < 10 \quad (3.174)$$

This formula states that the effective width of the ground plane is $W + 5.8H$ and having uniform current density. The skin-effect resistance R_m is given by $R_m = (w/i/a)^{1/2}$. For copper with a conductivity of 5.8×10^7 S/m. w^6 have $R_m = 8.22 \times 10^{-3}/f$ ft where f is in gigahertz. The total series resistance is $R_s + R^*$ and thus upon using (3.346) we get

$$a_{ss} = \frac{x?}{2Z_C} + R_2 \quad (3.1751)$$

for the attenuation caused by conductor losses. For the quasi-TEM mode the magnetic field, and hence the conductor losses, do not depend on substrate material.

The equations presented above predict somewhat higher theoretical attenuation than that obtained from a formula developed by Pucel, M», and Hartwig using Wheeler's incremental inductance rule.t Our form¹¹*

tR. A. Pucel, D. J. Masse, and C. P. Hartwig, Losses in Microstrip, *IEEE Trans.*, vol. *4T*, pp. 342-350, June, 1966.

appear to be in better agreement with experimental results. For practical microstrip lines surface roughness can increase the attenuation by as much as 50 percent or more depending on the scale of surface roughness relative to the skin depth. The etching process does not produce a perfectly flat end face at the sides of the strip. Some undercutting of the edge occurs along with some roughness, which will also result in an increase of the attenuation above the theoretical values.

Example 3.6. We wish to find the attenuation for a microstrip line using a copper strip of width 0.5 mm, a spacing of 0.5 mm above the ground plane, an alumina substrate with $e_r = 9.7$, and a loss tangent of 2×10^{-4} . The strip thickness T is 0.02 mm. The frequency of operation is 4 GHz.

The effective dielectric constant and characteristic impedance were found in Example 3.4 and are $e_e = 6.556$, $Z_c = 49.44 \text{ ohms}$. The wavelength of operation is 7.5 cm. The attenuation caused by dielectric loss was calculated after (3.171) was presented and is $1.52 \times 10^{-3} \text{ Np/wavelength}$ or $2.02 \times 10^{-4} \text{ Np/cm}$.

For this microstrip line $W/H = 1$ and $W/T = 25$. From (3.173) we find that the loss ratio is 1.0658 and (3.172) gives

$$R_1 = \frac{1.0658 \times 8.22 \times 10^{-3} \text{ vf / 1}}{0.05} + \frac{\ln 100TT}{ir^2} \\ = 3.157 \times 10^{-5} \text{ n/cm}$$

By using (3.174) we obtain

$$R_* \sim \frac{8.22 \times 10^{-3/4}}{0.05} = 4.81 \times 10^{1/2} \text{ fl/cm} \\ 5.8 + 0.03$$

For this case the loss in the microstrip is a factor of 6.5 greater than that in the ground plane. The reason for this is the high current density near the edges in the microstrip as compared with a more uniform current distribution over a wider area on the ground plane. By using (3.175) we find that the attenuation due to conductor loss is

$$" \bullet : \frac{3.157 \times 10^{-1} + 4.81 \times 10^{-2}}{2 \times 49.44} = 3.68 \times 10^{-3} \text{ Np/cm}$$

This attenuation is 18.4 times greater than that caused by dielectric loss. The total attenuation is

$$a = a_c + a_d = 3.88 \times 10^{-3} \text{ Np/cm}$$

which equals 0.0337 dB/cm. Surface roughness could result in a real attenuation 15 to 25 percent higher than this at 4 GHz. In this example we neglected the correction to e_c and Z_c due to the finite thickness T since it is small.

High-Frequency Properties of Microstrip Lines

The equations describing the quasi-TEM mode can be used with accent v accuracy for frequencies up to 2 to 4 GHz for a substrate thickness of 1 mm . For a substrate 0.5 mm thick, the upper frequency limit would be 4 GHz . When these limits are exceeded, it is necessary to take into account the frequency dispersion of the effective dielectric constant and the characteristic impedance with frequency. At the higher frequencies the electric field becomes more confined to the region between the microstrip* and ground plane. The greater concentration of the field in this region results in an increase in the effective dielectric constant as well as increase

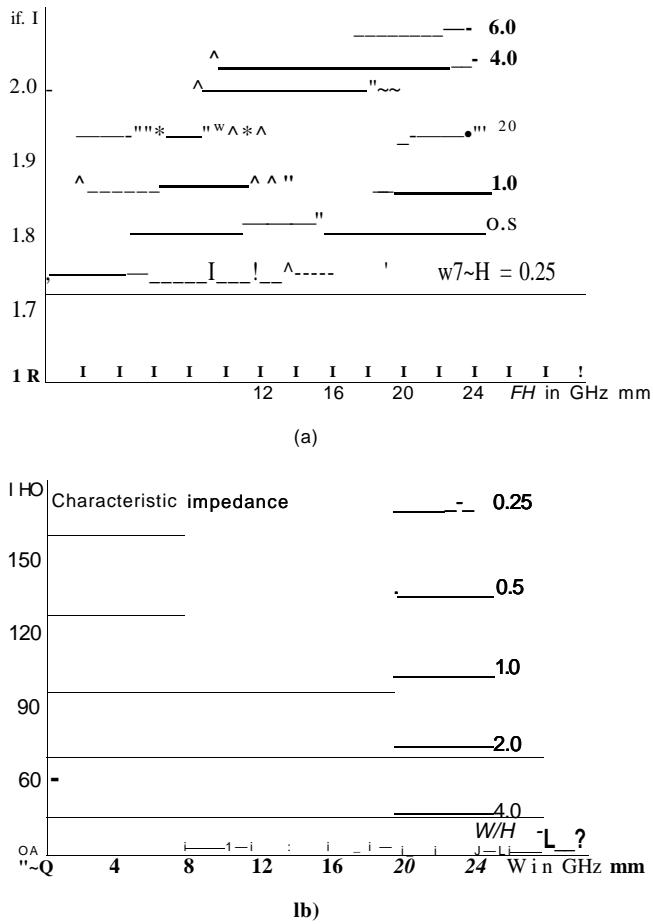
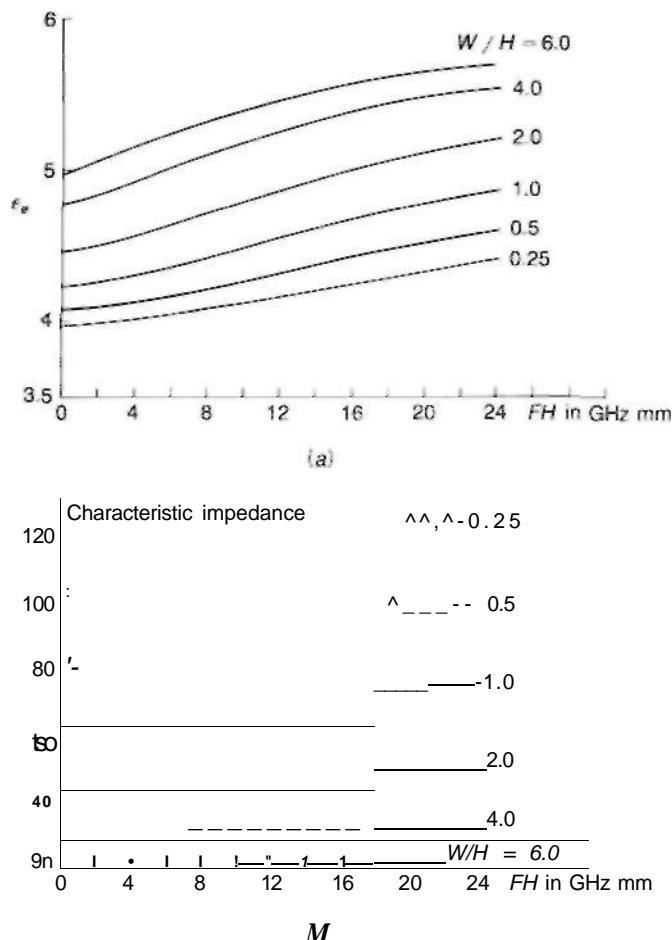


FIGURE 3.25

(a) Effective dielectric constant for a PTFE/microfiber glass substrate with $\epsilon_r = 2.26$, $t_y = 0.001 \text{ mm}$.
 (b) Characteristic impedance.

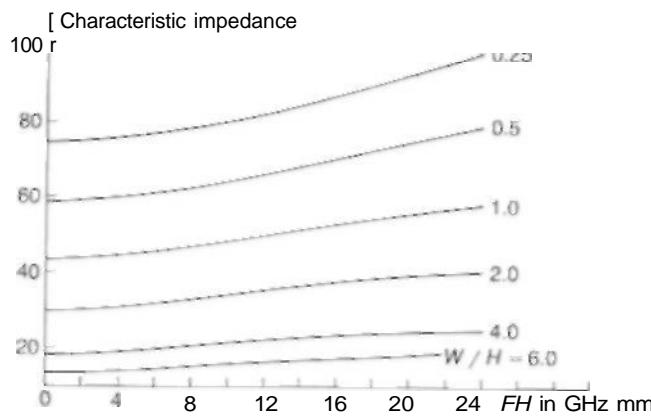
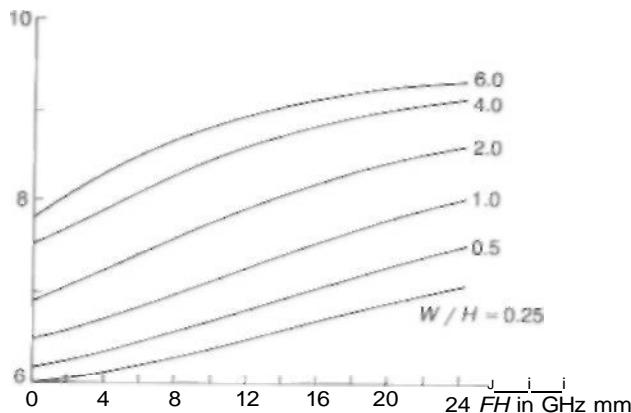
%%

**FIGURE 3.26**

(a) Effective dielectric constant for an RT/Duroid 6006 substrate having $\epsilon_r = 6.36$, $\epsilon_y = 6$; (b) characteristic impedance.

attenuation because of a greater concentration of the electric field in the substrate which has some loss. The conductor loss also increases because the skin-effect resistance R_m increases and more of the current flows on the inner face of the microstrip.

In order to determine the effective dielectric constant at high frequencies, it is necessary to carry out a full wave analysis, i.e., the complete set of equations given earlier for the potentials must be solved. In Figs. 3.25 to 3.28 we show the dispersive properties for four common substrate materials, a PTFE/microfiber glass substrate with $\epsilon_r = 2.26$, $\epsilon_y = 2.2$, RT/Duroid 6006 with $\epsilon_r = 6.36$, $\epsilon_y = 6$, alumina with $\epsilon_r = 9.7$, and gallium arsenide



(b)

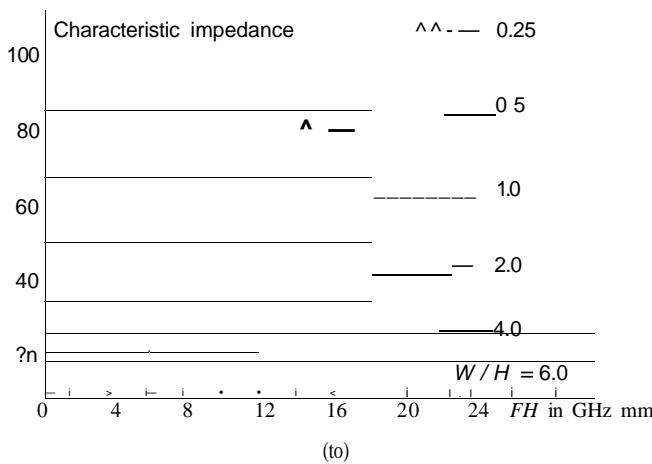
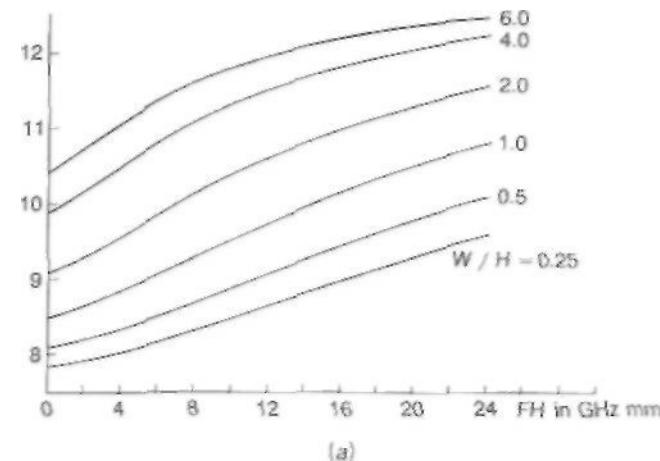
FIGURE 3.27

(a) Effective dielectric constant for an alumina substrate with $c_r = 9.7$; (b) characterise impedance.

with $\epsilon_r = 12.9$. An examination of these figures shows that dispersion effects are more pronounced for wider strips and large dielectric constants⁸.

Along with each figure giving the effective dielectric constant as a function of normalized frequency is a figure showing the characterise impedance as a function of normalized frequency. The normalized frequency is the actual frequency in gigahertz multiplied by the substrate thickness⁸ millimeters. Thus, for substrates 0.5 mm thick, the frequency range covered is 0 to 48 GHz.

When the propagating mode is not a TEM or quasi-TEM mode, there is no unique value for the characteristic impedance because it is not $P^{\circ S^9} \backslash J$ to define a unique value for the voltage as given by the line integral of electric field between the ground plane and the microstrip. The character

**FIGURE 3.28**

(a) Effective dielectric constant for gallium-arsenide substrate with $\epsilon_r = 12.9$; (b) characteristic impedance.

tic impedances given in Figs. 3.25 to 3.28 are based on using the following definition for the equivalent voltage:

$$\mathbf{V} = -/\mathbf{V}^* = /> \mathbf{A} > \int_{\mathbf{r}_0}^{\mathbf{r}_P} dy$$

where the path of integration is along a straight line from the ground plane to the center of the microstrip. The current I is chosen as the total 2-directed current on the microstrip and Z_c was calculated from the ratio V/I . In general, the power flow along the microstrip transmission line will not equal $|V|I$. The characteristic impedance can be defined in terms of the power flow P by choosing either the current I or voltage V according to the

above definitions and using

$$\frac{Z_1^2}{2Z_C}$$

to find Z_c . These two equations will give different values for the characteristic impedance with neither one being equal to V/I .

The lack of a unique value for the characteristic impedance is not a great disadvantage since microstrip junctions and discontinuities can be described by equivalent circuits using any convenient definition for the characteristic impedance. In Chap. 4 we will find that an equivalent transmission-line circuit theory can be formulated for any waveguiding system and does not require that there be a unique characteristic impedance associated with the propagating modes.

For computer-aided design (CAD) of microstrip circuits, it is important to have simple formulas that can be used to find the effective dielectric constant. Many different formulas have been proposed. The most accurate one that covers the full range of parameter values likely¹ to be encountered was developed empirically by Kobayashi and is given below:

$$< " > - ' - i t o w u " \quad (3m6)$$

where

$$K = 0.75 + (0.75 \sim 0.332c^{1/73})W/H$$

$$h = \frac{47.746}{H\epsilon_r - \epsilon_e(0)} \tan^{-1} \frac{\sqrt{\epsilon_e(0) - 1}}{\epsilon_r - \epsilon_e(0)}$$

$$m = m_Q m_c < 2.32$$

$$m_n = 1 + \frac{47.746}{H\epsilon_r - \epsilon_e(0)} + 0.32(1 + JW/H)^{-3}$$

$$m = \frac{I + YTWH^{0.15} \sim 0.235e^{-a45f/ro}}{H - 0.17} \quad \frac{W}{H} > 0.7$$

In these formulas H is in millimeters, the frequency f is in gigahertz, and whenever the product m_0m_c is greater than 2.32 the parameter m is chosen equal to 2.32. The effective dielectric constant at the frequency f is $e_e(f)$ and $e_c(0)$ is the quasistatic value which can be found using (3.166). It requires only a simple computer program to evaluate $e_e(f)$ using Kobayshi's formula. The accuracy is estimated to be within 0.6 percent for $0.1 < W/H < 10$, $1 < e_r < 128$ and for any value of H/A_0t .

In microstrip circuit design where junctions of microstrip lines with different widths are involved, it is necessary to characterize the junction in terms of an equivalent circuit. The parameters of this equivalent circuit will depend on frequency. The equivalent characteristic impedances that are assigned to the microstrip lines are arbitrary and often are simply chosen to have normalized values of unity. Any impedance level change that occurs at the junction is incorporated as part of the equivalent circuit of the junction. For these reasons we will not quote any of the formulas that have been proposed for evaluating characteristic impedance as a function of frequency because they are of limited use in practice.

Simple formulas giving the attenuation of microstrip lines at high frequencies do not exist. Equation (3.171) can be expected to give a good estimate for the attenuation due to dielectric loss provided e_e is replaced by the effective dielectric constant that applies at the frequency of interest.

A realistic evaluation of the attenuation caused by the finite conductivity of the conductors requires evaluation of the current density on the microstrip and the ground plane at the frequency of interest. In general, the current tends to be more uniform across the microstrip at high frequencies, particularly for wide strips. In Fig. 3.29 we show several computed current distributions at frequencies of 10 and 20 GHz for an alumina substrate 1 mm thick. The quasistatic current distribution is also shown. The current density has been normalized so that the total current on the microstrip equals W . In view of the tendency for the current density to become nearly uniform at high frequencies for wide strips, the attenuation constant can be estimated with fair accuracy by assuming uniform current density over a width W on both the microstrip and the ground plane. In this limit the attenuation caused by the conductor loss for wide strips that are not too thin is given by

$$\alpha_c = \frac{\gamma \ell}{WZ_C} \quad (3.177)$$

^tThe author has verified the accuracy of Kobayshi's formula by comparison with calculated numerical results for $0.25 < W/H < 6$ and $2 < e_r < 12$.

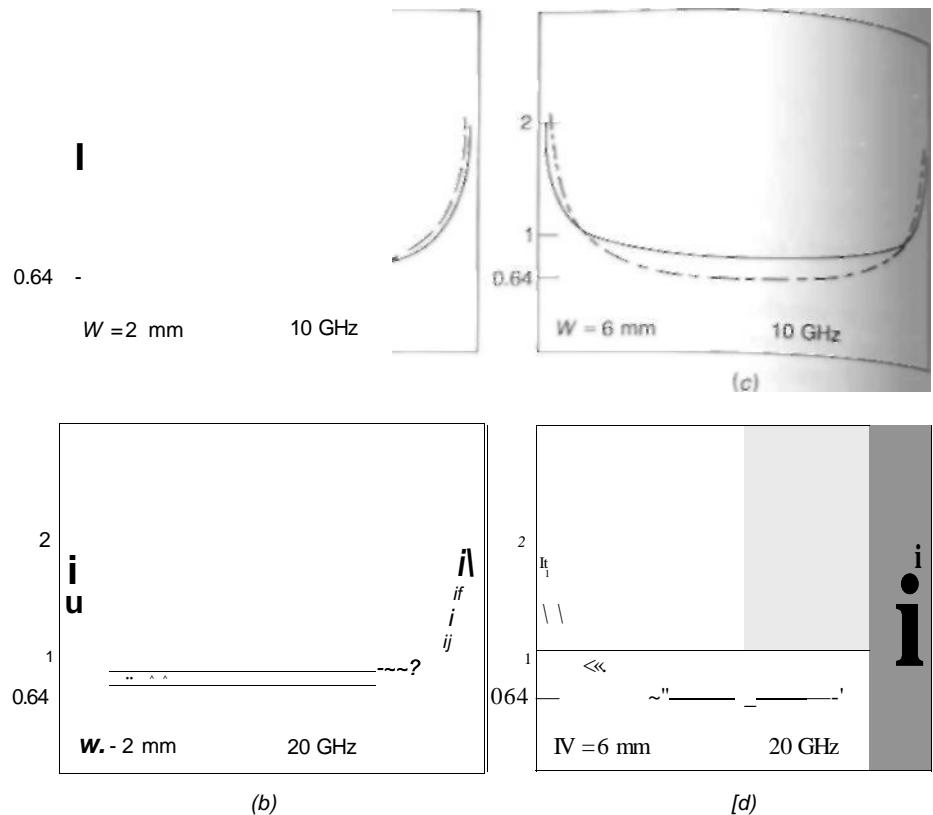
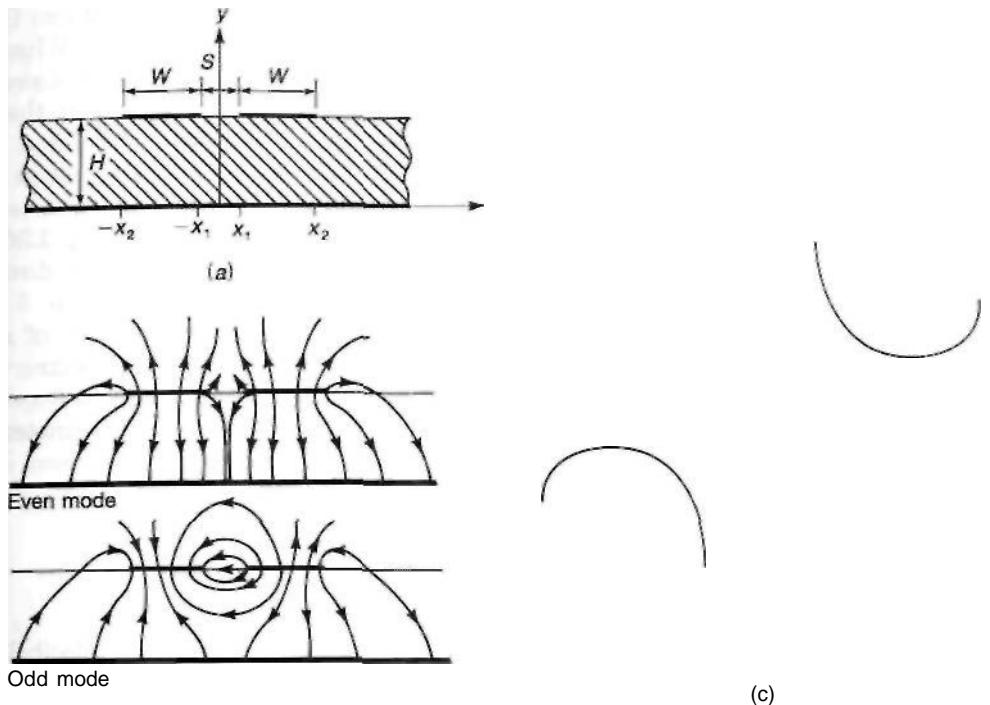


FIGURE 3.29
Current distribution on the microstrip for an alumina substrate with $H = 1 \text{ mm}$ and two different widths. The broken curves give the quasistatic distribution, (a) $W = 2 \text{ mm}$, $f = 10 \text{ GHz}$; (b) $W = 2 \text{ mm}$, $f = 20 \text{ GHz}$; (c) $W = 6 \text{ mm}$, $f = 10 \text{ GHz}$; (d) $W = 6 \text{ mm}$, $f = 20 \text{ GHz}$.

For narrow strips and high frequencies, no simple formulas for attenuation appear to be available. For narrow strips, say $W/H < 1$, the quasistatic formula (3.175) is probably a reasonably good estimate since M^* current density does not depart significantly from the quasistatic distribution for narrow strips.

3.13 COUPLED MICROSTRIP LINES

When two conducting strips of width W are placed side by side dielectric substrate above a ground plane as shown in Fig. 3.30a, we obtain a coupled microstrip line. Since this is a three-conductor transmission there are two fundamental quasi-TEM modes of propagation. The y mode is the mode corresponding to both strips being at the same potential^{*}

**FIGURE 3.30**

(a) Coupled microstrip line; (b) the electric field distribution for the even and odd modes; (c) the current distribution for the even and odd modes.

and on which the same currents exist. The odd mode corresponds to the strips being at opposite potentials, $-V$ and V , relative to the ground plane. For the odd mode the currents on the two strips are also equal in amplitude but of opposite sign. A sketch of the electric field lines for the two modes is shown in Fig. 3.306. For isolated strips in air, i.e., with no ground plane and substrate present, the theoretical current distributions for the two modes are:

$$J'(x) = V(*^2 - *?) (*\$ - *^2) \quad (3.178a)$$

$$\langle I^*(*) \rangle = \wedge (*^2 - *?) (*! - *^2) \quad (3.178b)$$

For the even mode the factor x in the numerator reduces the arnpUrtur the singular behavior near the inner edges at $x = \pm x_v$. When x_t equals * °^a the current singularity at $\pm r$, vanishes and $J''(x)$ becomes the exDe^{^0} current density on a single strip $2x_2$ units wide. For the odd mode current singularity at the inner edges $\pm x_1$ is more like a $1/x$ simil when *! is very small. This is caused by the strong electric field acrn very narrow slit with the adjacent conductors at opposite potential TO current distribution for the two modes is shown in Fig. 3.30c.

The coupled microstrip line is used in various directional coupler designs and these applications will be discussed in Chap. 6. The imports parameters describing the quasi-TEM mode properties of the coupled microstrip line are the even- and odd-mode effective dielectric constants ϵ' ϵ'' that determine the two propagation constants, and the even- and odd-mode characteristic impedances Z'_c , Z''_c . An important parameter in directional coupler design is the coupling coefficient, which is given in terms of the characteristic impedances of the two modes by

$$C = \frac{ye}{j} \quad (3.179,$$

The coupling coefficient is commonly expressed in decibel units, that is, $20 \log C$. In a coupled microstrip line it is not practical to achieve much more than a 2.5: 1 impedance ratio; so strong coupling cannot be realized in a simple coupled microstrip directional coupler. However, other designs are available, so this is not a problem for the microwave circuit engineer.

There are no simple formulas giving the quasi-TEM properties of coupled microstrip lines that have an accuracy comparable to that for microstrip lines. Bahl and Bhartia list formulas characterizing coupled microstrip lines that give results which are acceptable for noncritical applications.¹ The computer program CMST implements these formulas. We have checked the accuracy against numerical results obtained from a full wave solution for an alumina substrate. The effective dielectric constants were found to be accurate to within 3 percent. The characteristic impedance for the even mode was also found to be accurate to within 3 percent. For the odd-mode characteristic impedance, the error was as large as 8 percent when $W/H = 1$, $S/H = 0.25$, but considerably less for larger values of W/H and S/H .

The dispersion properties for coupled microstrip lines on a variety of different substrates have been computed by Morich.² In Figs. 3.31 to 3.34

9 Son⁵

tj., Bahl and P. Bhartia. "Microwave Solid State Circuit Design," p. 28, John Wiley & Sons Inc., New York, 1988.

JM. Morich, Broadband Dispersion Analysis of Coupled Microstrip on Anisotropic Substrates, by Perturbation-Iteration Theory, M.S. Thesis, Case Western Reserve University, dev⁸, Ohio, May, 1987.

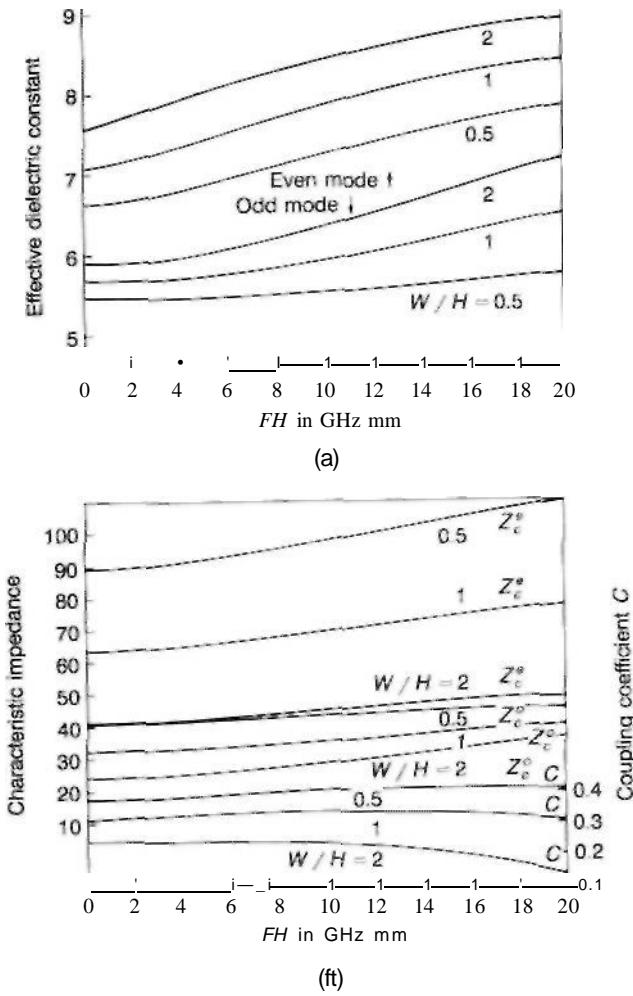
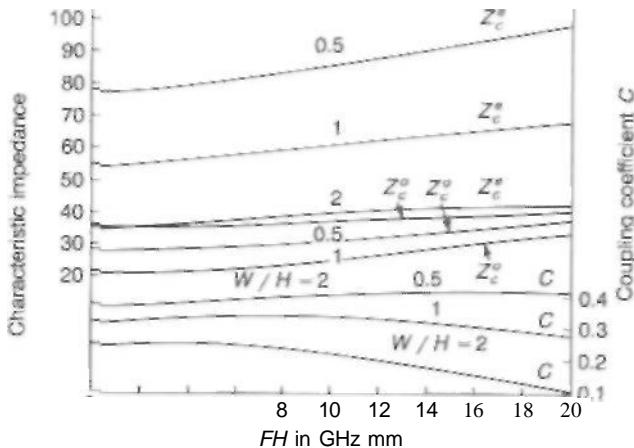
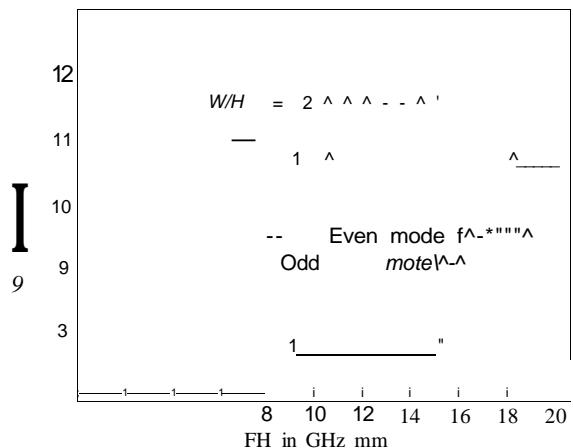


FIGURE 3.31

Dispersion characteristics of a coupled microstrip line on an alumina substrate. $S/H = 0.25$, $\epsilon_r = 9.7$. (a) Even- and odd-mode effective dielectric constant; (b) even- and odd-mode characteristic impedance and coupling coefficient C .



(6)

FIGURE 3.32

Dispersion characteristics of coupled microstrip line on a gallium-arsenide substrate. $S''/0.25$, $\epsilon_r = 12.9$, (a) Even- and odd-mode effective dielectric constant; (6) even- and odd-mode characteristic impedance and coupling coefficient C .

rises to a value of 0.524 at $FH = 20$ GHz mm. A coupling coefficient of 0.1 is needed for a 6-dB directional coupler. The effective dielectric constant for the odd mode is smaller than that for the even mode because a larger percentage of the electric field energy is located in the air region—capacitance between closely spaced parallel strips at opposite potential is large so the characteristic impedance of the odd mode is smaller than that for the even mode for the normal range of parameters involved. For strips with large spacing, there is very little coupling and the two impedances will be almost the same.

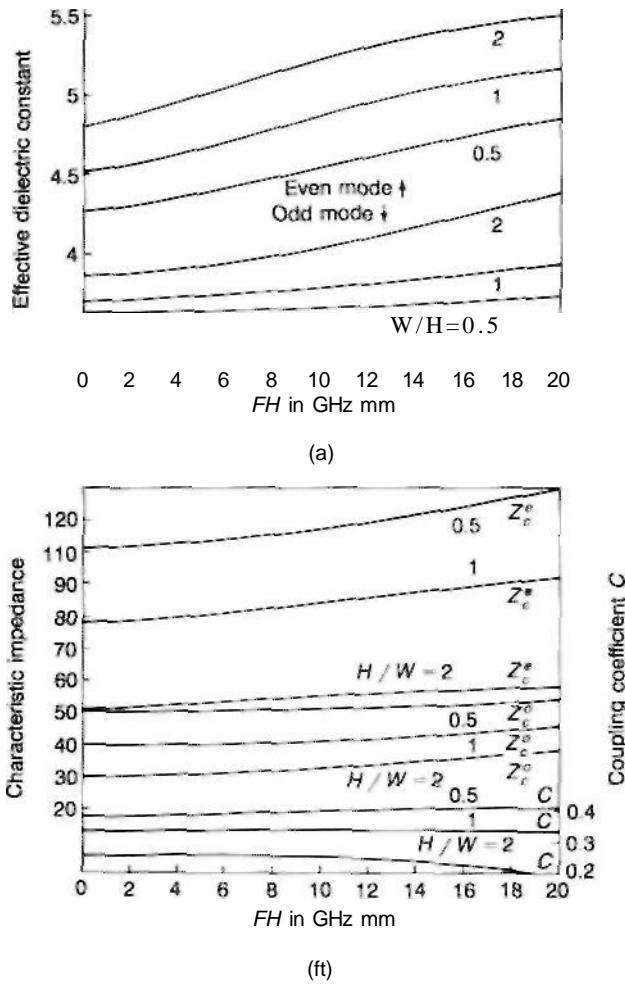


FIGURE 3.33
Dispersion characteristics of coupled microstrip line on an RT/Duroid 6006 substrate. $S/H = 0.25$, $\epsilon_r = 6.36$. (a) Even- and odd-mode effective dielectric constant; (b) even- and odd-mode characteristic impedance and coupling coefficient C .

STRIP TRANSMISSION LINES

The basic strip transmission line consists of a conducting strip embedded in a dielectric medium between two ground planes as shown in Fig. 3.34a. Figures 3.34b and c illustrate coupled strip lines. For broadside CQ_U, strips as shown in Fig. 3.34c, the coupling coefficient is significantly great* than for coplanar strips. Thus the strip line with broadside coupled string suitable for directional couplers where a coupling of 3 dB is required. Since the dielectric completely surrounds the strips, the strip line and the COUDIBH strip line support pure TEM modes of propagation. There is no frequency dispersion or change in effective dielectric constant with frequency. Consequently, for the coupled strip line the even and odd modes of propagation have the same phase velocity, which is also a desirable feature for directional coupler design.

For the symmetrical strip line and coupled strip line, the TEM-mode characteristic impedance is readily found from the distributed capacitance with the latter determined by conformal mapping techniques. The distributed capacitance for the symmetrical strip line is found in App. in.

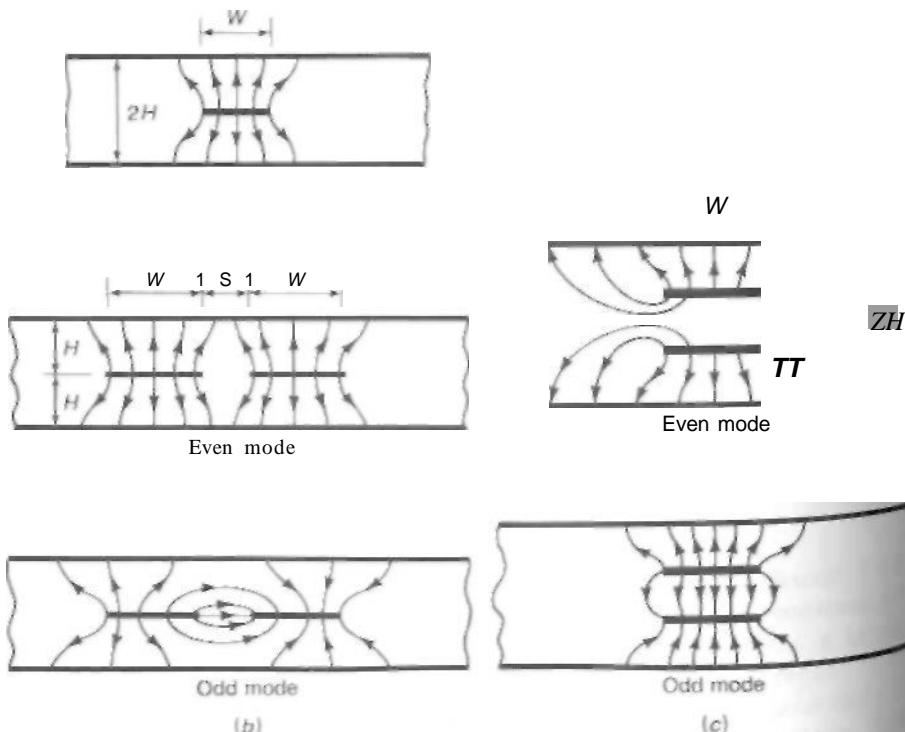


FIGURE 3.34

(a) The basic strip-line configuration; (b) coupled strip line using coplanar strips; (c) coupled strip line using broadside coupled strips. The electric field lines for the TEM modes are shown.

From the given expression (III.27) and using (III. 13c), we have

$$Z_r = \frac{Z_0 K(k)}{* & K(k')} \quad \frac{Z_0 K}{A j \gamma_r K'} \quad (3.180a)$$

where

$$k = \frac{\cosh -rt}{\cosh -rt} \quad \frac{W}{4H} \quad k' = \tanh iv \quad W$$

$$W = \frac{K}{\ln 2} \quad \frac{1}{TWT} \quad 0.7 < k < 1$$

$$\frac{K}{K'} = \frac{1}{\ln 2} \quad \frac{1}{TWT} \quad 0 < k < 0.7$$

and ϵ_r is the dielectric constant of the dielectric material that completely surrounds the center strip. For $W > 2H$ the formula for Z_c reduces to the simple form

$$Z_c = \frac{1}{\sqrt{8\pi Z_r (\ln 2 + TTW/4H)}} \quad W > 2H \quad (3.180b)$$

For a very narrow strip

$$Z_c = \frac{16H}{2nyfc_* \ln TWT} \quad W < 0AH \quad (3.180c)$$

Since the dielectric material completely fills the strip line, the attenuation due to dielectric loss is given by (3.87a) and is

$$\alpha J = \frac{ire_*}{Aov^{\lambda}} \quad iTi/\epsilon'_r \tan \phi, \quad (3.181)$$

Formulas for the attenuation in a strip line due to conductor losses and assuming an inner conductor of elliptical cross section have been derived using conformal mapping techniques.^t In App. III it is shown that the series resistance of an isolated conductor of elliptical cross section is greater than that for a conductor with a rectangular cross section. By using an equivalent center thickness T_e for the conductor with elliptical cross section, the two series resistances will be equal provided T_e is chosen to be

$$T_e = e^{-\lambda^2} \quad \frac{4WT}{\lambda} \quad (3.182)$$

^tR. E. Collin, *loc. cit.*, chap. 4, eqs. (73), (74), and (76).

We will assume that this equivalence is also a good approximation for a conductor placed between two ground planes. On this basis the attenuation by the center conductor is

$$\alpha_{c,} = \frac{16Z_cHK^2k'}{16Z_cHK^2k' - \ln \frac{16Hk'}{k_r T_e}} \quad (3-183Q)$$

and that due to the ground planes is

$$\alpha_{g,}^2 = \frac{\frac{R_n W}{2}}{64Z_c H^2 K^2 k'} \quad (3-1836)$$

For wide and narrow strips these formulas reduce to the following simplified forms:

$$\alpha_{c,}^l = \frac{R_j 7_r}{\sim W ir} \frac{\frac{8H}{\ln \frac{f_e}{2} + JH}}{\frac{W}{\sim}} \quad W > 2H \quad (3184a)$$

o » * + !S

$$\alpha_{c,}^w = \frac{rr R^2 W}{8Z_0 H^2 \frac{\ln 2}{\sqrt{2}} + \frac{4}{4/7}} \quad W > 2H \quad (3.1846)$$

$$\alpha_{r,} = \frac{2Rj e_r n}{2Rj e_r n} \quad W < OAH \quad (3.184c)$$

$$\alpha_{g,} = \frac{16H}{2Z_0 H n} \quad W < OAH$$

Suitable formulas for evaluating K and K' are given by (III.13) in App III. The total attenuation for a strip line is given by

$$\alpha = \alpha_d + \alpha_{c,} + \alpha_{g,} \quad (3-184)$$

Example 3.7. A strip line has a ground-plane spacing $2H = 1$ cm and uses centered copper conducting strip of width $W = 1$ cm and thickness $T = 0.01$ cm. The dielectric filling material has a dielectric constant $\epsilon_r = 2.2$ and a tangent equal to 10^{-3} . We want to find the characteristic impedance and attenuation at a frequency of 10 GHz. For this line $W/H = 2$, so (3.1806)^c be used. Thus

$$Z_c = \frac{120TT^2}{8\pi^2(\ln 2 + 7r/2)} = 44.09 \Omega$$

The wavelength of operation is 3 cm; so by using (3.181) we find the attenuation due to dielectric loss to be

$$\begin{aligned} & \text{irr}^2 \times 10^{-s} \\ & \ll * - \frac{3}{3} - \\ & = 1.55 \times 10^{-1} \text{ Np/cm} \quad \text{or} \quad 1.35 \times 10^{-2} \text{ dB/cm} \end{aligned}$$

From (3.182) we obtain $T_e = 0.00742$ cm for the equivalent thickness. We can use (3.184a) to find

$$\begin{aligned} & \frac{8.22 \times 10^{-3} \text{ vTO} J22v}{240.7 \times 0.5} \frac{m^-}{\ln 2 + -/2} \frac{0.00724}{2} \\ & = 3.034 \times 10^{-4} \text{ Np/cm} \quad \text{or} \quad 2.635 \times 10^{-3} \text{ dB/cm} \end{aligned}$$

From (3.184b) we obtain

$$\begin{aligned} & 8.22 \times 10^{-3} \text{ vTO} J22v \\ & a_{n2} \sim 8 \times 12077 \times 0.5^2 (\ln 2 + -/2) \\ & = 7.096 \times 10^{-5} \text{ Np/cm} \quad \text{or} \quad 6.163 \times 10^{-4} \text{ dB/cm} \end{aligned}$$

The total attenuation is 0.0167 dB/cm. For this transmission line the dielectric loss is greater than the conductor loss.

According to a formula for strip-line conductor attenuation developed by Cohn, we would get $a_c = 3.236 \times 10^{-3}$ dB/cm, which agrees very closely with 3.25×10^{-3} dB/cm obtained with our formulas.t

COUPLED STRIP LINES

For coplanar strips of width W and spacing S , the even- and odd-mode characteristic impedances are given by\$

$$\begin{aligned} & c \quad 4/\epsilon_r \quad K(k_t) \\ & r \quad J^K K_i K_l \\ & \quad i_l T_r K(k_n) \end{aligned} \quad (3.186)$$

tSee H. Howe. "Strip Line Circuit Design," eq. (1.5), Artech House Books, Dedham, Mass. 1974.

iS. B. Cohn, Shielded Coupled Strip Transmission Line. *IRE Trans.*, vol. MTT-I, pp. 29-37, October, 1955.

S. B. Cohn, Characteristic Impedances of Broadside-Coupled Strip Transmission Lines, *IRE Trans.*, vol. MTT-8, pp. 633-637, November, 1960.

S. B. Cohn, Thickness Corrections for Capacitive Obstacles and Strip Conductors, *IRE Trans.*, vol. MTT-8, pp. 638-644, November, 1960.

where

$$(w \quad W) \quad [TT \quad W+ \quad S]$$

and

$$\mathbf{K} = \mathbf{V}\mathbf{I} - *? \quad \mathbf{Ab} = \mathbf{I}/\mathbf{I} - *,^2.$$

Note that $\mathbf{KXA}^T = \mathbf{J}^T(\mathbf{A})$. For the evaluation of K and $\#$ see (111.13).

For coplanar strips with a thickness T ,

$$\begin{aligned} Z' &= \frac{Z_0(2H-T)}{4j7\sim_r(W+(H/TT)CA_s)} \quad i = e,o \quad W > Q.m \quad (3.187) \\ A=l &+ \frac{\ln[1 + \tanh(7rS/4H)]}{\ln 2} \\ A_- &= 1 + \frac{\ln[1 + \coth(irS/4/f)]}{\ln 2} \\ &\quad (4H - T) \quad T(4H - 7) \\ &\quad 2tf \ln \frac{(2H - T)^2}{(2H - T)^2} \end{aligned}$$

For broadside coupled strips as shown in Fig. 3.34c, the even- and odd-mode characteristic impedances are given by [valid for $W > 0.35S$, $W > 0.7/7(1 - S/2H)$]

$$2\Lambda^{-f} \quad ST \gg^W + \sqrt{0.4413 + \frac{1}{\ln \frac{2H}{2H-S-2T} - \frac{S+2T}{2H-S-2T} - \frac{2H}{2T}}} \quad (3.188a)$$

$$2\Lambda^{-f} \quad \frac{W}{2f} - S - 2T + \frac{W}{S+C} \circ V^+ - \frac{2}{-} \quad I + \mathbf{S} \mathbf{I} \quad \frac{T}{-S} - \frac{T}{S} \quad (3.1886)$$

where

$$P'' = \frac{2// - 2T}{rS} \quad \ln \frac{2tf - 2T}{2H-S-2T} + \frac{8}{2H-S-2T} \ln \frac{2H-2T}{2T}$$

All of the characteristic impedances given above refer to the impedance of one strip to the ground planes. froff¹

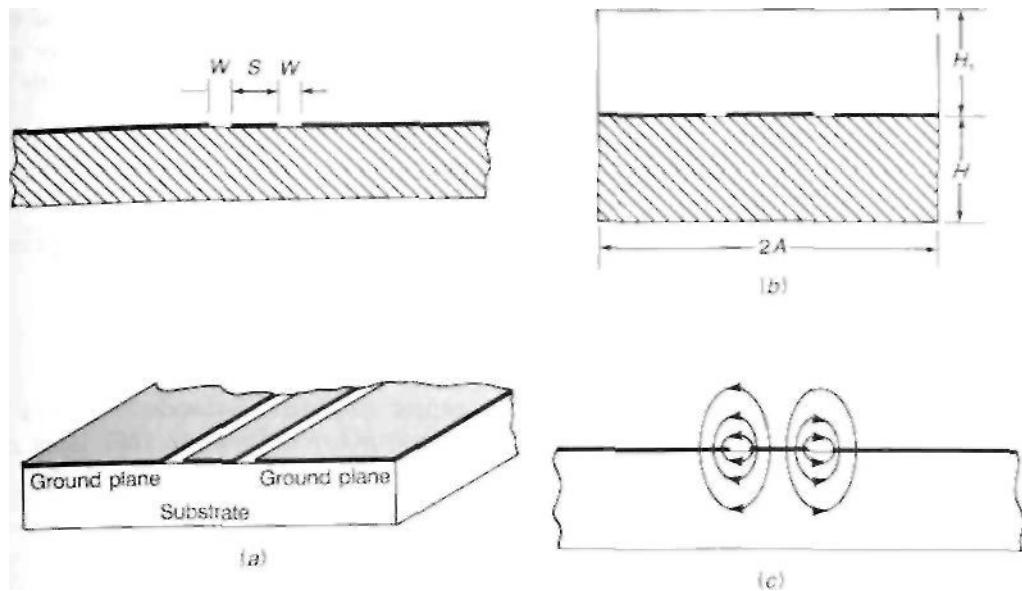


FIGURE 3.35

(a) Basic coplanar transmission line; (b) a shielded coplanar transmission line; (c) electric field distribution.

COPLANAR TRANSMISSION LINES

An illustration of a coplanar transmission line is shown in Fig. 3.35a. A shielded coplanar line is shown in Fig. 3.35b. The strip width is S and the strip to ground-plane spacing is W . The coplanar line is often called a coplanar waveguide (CPW). The most significant advantage that a coplanar line has over a microstrip line is the ability to connect active and passive circuit components in shunt from the conducting strip to the ground plane on the same side of the substrate. In a microstrip line a connection to the ground plane requires drilling a hole through the substrate which is somewhat difficult for ceramic materials such as alumina. Figure 3.35c shows the electric field distribution in a coplanar line.

The characteristics of a coplanar line at low frequencies can be determined by conformal mapping techniques. A solution for the coplanar line is given in App. III. For the ideal case when the ground planes are very wide relative to the slot width W and the dielectric substrate is very thick, the electric field has the same distribution as for a coplanar line in air. The reason for this is that the mapping of the air-dielectric boundary, shown as the intervals BC and EF in Fig. III.4, is the two parallel sides in the ideal parallel-plate capacitor shown in Fig. III.3. Thus the solution for the potential is not affected by the presence of the dielectric. The distributed

capacitance of the coplanar line is thus the capacitance of one-half of air-filled line plus one-half of that for a line completely surrounded by dielectric. This capacitance is given by III. 14 and for the line with dielectric on one side only we have

$$C \sim \frac{K}{2e_0} + \frac{K'}{2e_r e_0} - \frac{(e_r + 1)}{4} \frac{K}{e_r} \quad (3.189)$$

where $K(k)$ and $K' = K(k')$ are again the complete elliptical integrals of the first kind. The modulus k is given by the ratio

$$k = \frac{1}{Y/k} = \frac{S/2}{W + S/2} = \frac{S}{S + 2W} \quad <^{3190}$$

since in Fig. III.4 the center conductor extends from -1 to 1 and the ground planes begin at $u = \pm(l/k)$. From (3.189) it is clear that the effective dielectric constant for the coplanar line is given by

$$\frac{e_r + 1}{2} \quad (3.191)$$

Consequently, the characteristic impedance is given by

$$Z_C = \frac{Z_Q K'}{T T^A} \quad <^{3192}$$

When the substrate material is anisotropic with a dielectric constant e_y in the direction perpendicular to the air-dielectric interface and e_r in the direction parallel to the conductors, then e_r should be replaced by $e_g = yjz_r e_y$ in (3.189) and (3.191).

In most applications it is necessary to provide shielding of a microwave circuit. If the shield dimensions are large, the shield will not produce a significant effect on the line characteristics. In monolithic microwave integrated circuits, the substrate is very thin and fragile; so it is desirable to use another ground plane below the substrate to mechanically strengthen the overall circuit and to also provide a better heat sink to help dissipate the power generated by active devices. Ghione and Naldi have considered number of different shield arrangements as well as coupled coplanar line*. We will present results only for the shielded coplanar line shown in Fig. 3.356. For this structure the effect of the sidewalls, spaced by amount $2A$, is negligible provided $2A > 10(S + 2W)$. We will assume that this is the case. Ghione and Naldi make the assumption that the air-dielectric interface in the slot regions can be replaced by magnetic walls in order to simplify

*G. Ghione and C. U. Naldi, Coplanar Waveguides for MMIC Applications: Effect of P and S Shielding, Conductor Backing, Finite Extent Ground Planes, and Line-to-Line Coupling. *Trans., vol. MTT-35*, pp. 260-267, March, 1987.

conformal mapping solution. The assumption is a correct one only for the case when the upper and lower shields are spaced the same distance from the coplanar line, that is, $H = H_Y$ in Fig. 3.356. For the line dimensions encountered in most practical applications of coplanar lines, the assumption does not introduce significant error. The advantage of this assumption is that it decouples the air and dielectric regions; so it is only necessary to find the capacitance of the air- and dielectric-filled sections separately.

It was found that the effective dielectric constant is given by

$$\epsilon_e = 1 + q/e_r - 1 \quad (3.193a)$$

where the filling factor q is given by

$$q = \frac{K_j k J / K_j k' J}{K(k_2) / K(k_2) + K(k) / K(k)} \quad (3.193*);$$

K is the complete elliptic integral of the first kind, k and k'_2 are the complementary moduli given by $y_1 = k$ and $y_2 = k'_2$, and

$$*^1 = \frac{\tanh(i\pi S/4H)}{\tanh[Lf_1(S+2W)/H]} \quad (3.193c)$$

$$*^3 = \frac{\tanh(TrS/4f_1)}{\tanh[7r(S+2W)/4ff_1]} \quad (3.193c/)$$

The characteristic impedance is given by

$$Z_n = \frac{Z_0}{2\gamma T_e [K(k_1) / K(k)]} + \frac{K(k_2) / K(k_2)}{K(k_2) / K(k_2)} \quad (3.194)$$

The ratios K/K' are easily evaluated using III.13c. For an anisotropic substrate e_r should be replaced by $e_g = Je_r e_Y$ and H should be replaced by the effective thickness $H_e = Je_r e_Y H$. When H and ϵ_r are very large, &, and k_2 become equal to k given by (3.190) and the filling factor q equals 0.5.

If the lower shield plate is replaced by a magnetic wall, we obtain an approximate model of a coplanar line with a substrate of finite thickness H and with an upper shield at a spacing H , above the coplanar conductors. For this case Ghione and Naldi give

$$q = \frac{K(k_2) / K(k'_a)}{K(k_2) / K(k'_a) + K(k) / K(k')} \quad (3.195a)$$

$$Z_n = \frac{2j r_e [K(k_2) / K(k'_a) + K(k) / K(k)]}{\epsilon_r} \quad (3.195b)$$

where

$$\begin{aligned} & \sinh(7rS/4H) \\ & \quad \beta = \sinh[\text{ir}(S + 2W)/4H] \end{aligned}$$

and k_2 is given by (3.193d), k is given by (3.190), and e_c is given by (3.193a). When H is made infinite k_2 becomes equal to k . The result structure is an unshielded coplanar line with a substrate having a finite thickness H .

Example 3.8. A coplanar line with upper and lower shielding and used in *-MMIC circuit has the following dimensions: $S = 50 \mu\text{m} = 0.05 \text{ mm}$, $\text{IV} = \text{jr}' \text{ nm}$, $H = 250 \text{ } \mu\text{m} = 0.25 \text{ mm}$, $\text{II}_c = 800 \text{ Mm}$. The substrate material is gallium arsenide with $e_r \approx 12.9$. We want to find the effective dielectric constant and characteristic impedance. The first step is to find k_x and k_2 using (3.193c) and (3.194c). A straightforward evaluation gives $k_x = 0.3547$, $k_2 = 0.33547$. We now use (III. 13c) to find $K(k_x)/K(k) = 0.6573$, $K(k_2)/K(k_2) = 0.6414$. By using (3.193a) and (3.1936), we get $q = 0.506$ and $e_c = 7.0228$. The last calculation is for Z_c using (3.194) and gives $Z_c = 54.77 \text{ } \Omega$. This example shows that for the shield spacings used very little effect on the coplanar-line characteristics occurred as is apparent from the fact that $k_x = k_2$ and $q = 0.5$. relationships that hold exactly when no shields are used.

Attenuation

The attenuation in a coplanar line caused by dielectric loss is given by the same formula as for a microstrip line, i.e., by (3.171). By using the filling factor q this can be expressed as

$$q \tan \theta, \quad (3-196)$$

K

This formula can be used for shielded and unshielded lines as long as the appropriate filling factor q is used.

For thin conductors with thickness T less than $0.05S$ and with a slot width $W > 0.3S$, the formula for the unshielded coplanar-line attenuation derived in App. III can be used. The center conductor has a series resistance per unit length given by

$$\alpha_t = 4S(1 - k^2)K^2/k \quad (3.197")$$

The distributed series resistance of the ground planes is given by

$$\alpha_g = 4S(1 - k^2)K^2/k \quad 77 + 1,1 \frac{4tr(S + 2W)}{f} \frac{1}{\ln TTT} \quad (3.197*)$$

where $k = S/(S + 2W)$. The attenuation due to conductor loss is given by

$$* - \frac{R_1 + R_2}{S} \quad \text{B-19c}$$

In *normal* situations the additional attenuation introduced by upper and lower shields is small if H and ℓ , are greater than AW .

Example 3.9. A coplanar line has a copper strip of width 0.6 mm, slot width $W = 0.6$ mm, metal thickness $T = 0.005S = 3 \mu m$. The dielectric used is alumina with a dielectric constant of 9.7 and a loss tangent equal to 2×10^{-4} . We want to find the attenuation caused by dielectric loss and conductor loss at a frequency of 4 GHz. For this line $k = S/(S + 2W) = 0.333$. The effective dielectric constant, from (3.191), is 5.35. From (3.192) and (III.13c) we obtain

$$Z_r = \frac{120\pi T}{4j\sqrt{M}} \ln \frac{1 + \sqrt{1 + q^2}}{1 - q^2} = 63.7 \text{ } il$$

where we used $k' = \sqrt{1 + k^2} = 0.9428$. From (3.196) we obtain

$$a_{rf} = \frac{9.7\pi T}{7.5\sqrt{5.35}} \cdot 0.5 \times 2 \times 10^{-4} \times 8.686 = 1.526 \times 10^{-3} \text{ dB/cm}$$

where $q = (e_r - 1)/(e_r + 1) = 0.5$ was used. To evaluate $K(k) = K(0.333)$, we use (111.13c) to obtain

$$\begin{aligned} K(0.333) &= \frac{1}{1 + k'} \ln \frac{1 - k'}{Xl + k'} = \frac{1}{1.9428} \ln 0.294 = 1.9428 \\ &= 1.617 \end{aligned}$$

From (3.197) we get $i_2 = 19.25i_m/\text{cm}$ and $R_2 = 5.97i_{mm}/\text{cm}$. We now use $R_m = 8.22 \times 10^{-3} \text{ V}$ and (3.197c) to get

$$\% = (2.158 \times 10^{-2} + 6.69 \times 10^{-3}) = 2.827 \times 10^{-2} \text{ dB/cm}$$

The ground planes contribute 23.7 percent of the conductor loss. The total attenuation is $2.98 \times 10^{-2} \text{ dB/cm}$.

Jackson has made some loss calculations for coplanar and microstrip transmission lines at a frequency of 60 GHz and found that for typical line dimensions and characteristic impedances *greater* than 50 il , the coplanar line has a smaller attenuation.^t The substrate considered had a dielectric constant of 12.8 and a loss tangent of 6×10^{-3} . The maximum strip width considered by Jackson was 0.3 mm. For these line dimensions the quasistatic formulas given above can be used. We have verified that indeed the quasistatic formulas give essentially the same attenuation. The lower loss in coplanar lines appears to be due to being able to use a wider center conductor for a given impedance as compared with that for a microstrip line. However, this is not always

^tR. W. Jackson, Considerations in the Use of Coplanar Waveguide for Millimeter-Wave Integrated Circuits *IEEE Trans.*, vol. MTT-34, pp. 1450-1456, December, 1986.

necessarily true since the relative attenuation of the two types of transmission lines will depend on other factors such as substrate thickness and dielectric constant also.

High-Frequency Dispersion

Coplanar transmission lines exhibit dispersion effects similar to that of microstrip lines. There are less available computed results for the effect of dielectric constant and characteristic impedance for coplanar lines than what is available for the microstrip line. The coplanar line can be viewed as two coupled slot lines and from this point of view it is clear that there are two modes of propagation that are quasi-TEM in character. The reader is referred to the paper by Nakatani and Alexopoulos for typical dispersive properties of a coplanar line.^t In many integrated microwave circuit applications, the line dimensions are so small relative to the wavelength that the quasi-TEM formulas can be used even though the frequency may be as high as 50 GHz.

PART 3 RECTANGULAR AND CIRCULAR WAVEGUIDES

Hollow-pipe waveguides do not support a TEM wave. In hollow-pipe waveguides the waves are of the TE and TM variety. The waveguide with a rectangular cross section is the most widely used one. It is available in sizes for use at frequencies from 320 MHz up to 333 GHz. The WR-2300 waveguide for use at 320 MHz has internal dimensions of 58.42 in by 29.1 in and is a very large duct. By contrast, the WR-3 waveguide for use at 333 GHz has internal dimensions of 0.034 in by 0.017 in and is a very miniature structure. The standard WR-90 X-band waveguide has internal dimensions of 0.9 in by 0.4 in and is used in the frequency range of 8.2 to 12.5 GHz. The rectangular waveguide is widely used to couple transmitters and receivers to the antenna. For high-power applications the waveguide is filled with an inert gas such as nitrogen and pressurized in order to increase the voltage breakdown rating.

Circular waveguides are not as widely used as rectangular waveguides but are available in diameters of 25.18 in down to 0.239 in to cover the frequency range 800 MHz up to 116 GHz.

^tA. Nakatani and N. G. Alexopoulos, Toward a Generalized Algorithm for the Modeling of Dispersive Properties of Integrated Circuit Structures on Anisotropic Substrates, *IE&E*, vol. MTT-33, pp. 1436-1441, December, 1985.

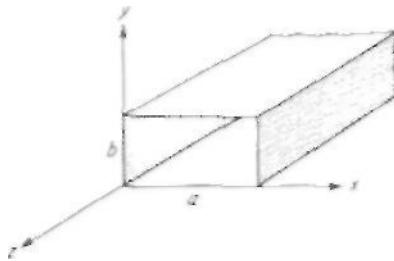


FIGURE 3.36
Rectangular waveguide.

RECTANGULAR WAVEGUIDE

The rectangular waveguide with a cross section as illustrated in Fig. 3.36 is an example of a waveguiding device that will not support a TEM wave. Consequently, it turns out that unique voltage and current waves do not exist, and the analysis of the waveguide properties has to be carried out as a field problem rather than as a distributed-parameter-circuit problem.

In a hollow cylindrical waveguide a transverse electric field can exist *only* if a time-varying axial *magnetic field* is present. Similarly, a transverse magnetic field can exist only if either an axial displacement current or an axial conduction current is present, as Maxwell's equations show. Since a TEM wave does not have any axial field components and there is no center conductor on which a conduction current can exist, a TEM wave cannot be propagated in a cylindrical waveguide.

The types of waves that can be supported (propagated) in a hollow empty waveguide are the TE and TM modes discussed in Sec. 3.7. The essential properties of all hollow cylindrical waveguides are the same, so that an understanding of the rectangular guide provides insight into the behavior of other types as well. As for the case of the transmission line, the effect of losses is initially neglected. The attenuation is computed later by using the perturbation method given earlier, together with the loss-free solution for the currents on the walls.

The essential properties of empty loss-free waveguides, which the detailed analysis to follow will establish, are that there is a double infinity of possible solutions for both TE and TM waves. These waves, or modes, may be labeled by two identifying integer subscripts n and m , for example, TE₁₁,..,. The integers n and m pertain to the number of standing-wave interference maxima occurring in the field solutions that describe the variation of the fields along the two transverse coordinates. It will be found that each mode has associated with it a characteristic cutoff frequency f_c below which the mode does not propagate and above which the mode does propagate. The cutoff frequency is a geometrical parameter dependent on the waveguide cross-sectional configuration. When f_c has been determined, it is found that the propagation factor i is given by

$$\mathbf{P} \sim \{\mathbf{H} - \mathbf{k}^* \mathbf{y}^*\} \quad (3.198)$$

where $k_n = w/t_0 c_0$ and $k_{\#} = 2Trf_c/p_0 e_0$. The guide wavelength is readj seen to be given by

$$A_s = \frac{2TT}{T} \frac{A_0}{''(1-A|/A'')^{1/8}} = \frac{\wedge}{yJl-C/f^*} \quad (3-199)$$

where A_0 is the free-space wavelength of plane waves at the frequen /*= W/2TT. Since k_c differs for different modes, there is always a lower band of frequencies for which only one mode propagates (except when k_c may ho the same for two or more modes). In practice, waveguides are almost universally restricted to operation over this lower-frequency band for which only the dominant mode propagates, because of the difficulties associated with coupling signal energy into and out of a waveguide when more than one mode propagates. This latter difficulty arises because of the different values of the propagation phase constant /3 for different modes, since this means that the signal carried by the two or more modes does not remain in phase as the modes propagate along the guide. This necessitates the use of separate coupling probes for each mode at both the input and output and thus leads to increased system complexity and cost.

Another feature common to all empty uniform waveguides is that the phase velocity v_p is greater than the velocity of light c by the factor A_g/A_0 . On the other hand, the velocity at which energy and a signal are propagated is the group velocity v_g and is smaller than c by the factor $A_0/A_{\#}$. Also, since $/3$, and hence A_s , u , and v_g , are functions of frequency, any signal consisting of several frequencies is dispersed, or spread out, in both time and space as it propagates along the guide. This dispersion results from the different velocities at which the different frequency components propagate. If the guide is very long, considerable signal distortion may take place. Group and signal velocities are discussed in detail in Sec. 3.19.

With some of the general properties of waveguides considered, it is now necessary to consider the detailed analysis that will establish the above properties and that, in addition, will provide the relation between k_c and the guide configuration, the expressions for power and attenuation, etc. The case of TE modes in a loss-free empty rectangular guide is considered first.

TE Waves

For TE, or H , modes, $e_z = 0$ and all the remaining field components can J> determined from the axial magnetic field h_z by means of (3.71). The a» field h , is a solution of

$$*\mathbf{K} + *?*, = \mathbf{0}$$

$$d^2 h, \quad d^2 k_z \quad -\mathbf{0}$$

or

(3.200)

If a product solution $h_z = f(x)g(y)$ is assumed, (3.200) becomes

$$\frac{1}{fdx^2} \frac{d^2f}{dx^2} + \frac{1}{g} \frac{d^2g}{dy^2} + *? \ll 0$$

after substituting fg for h_z and dividing the equation by fg . The term $\frac{1}{fdx^2} \frac{d^2f}{dx^2}$ is a function of x only, $\frac{1}{g} \frac{d^2g}{dy^2}$ is a function of y only, and $k'l$ is a constant, and hence this equation can hold for all values of x and y only if each term is constant. Thus we may write

$$\begin{aligned} III - H & \quad d^2f \\ \frac{1}{fdx^2} & \quad * \quad \text{or} \quad dz^2 \\ \frac{1}{g} \frac{d^2g}{dy^2} & \quad d^2g \\ & \quad v \quad dy^2 \end{aligned}$$

where $k_x + k^2 = k^2m$ order that the sum of the three terms may vanish. The use of the *separation-of-variables* technique has reduced the partial differential equation (3.200) to two ordinary simple-harmonic second-order equations. The solutions for f and g are easily found to be

$$f = A_1 \cos k_x x + A_2 \sin k_x x$$

$$g = S_2 \cos k_y y + B_2 \sin k_y y$$

where A_1, A_2, B_1, B_2 are arbitrary constants. These constants, as well as the separation constants k_x, k_y , can be further specified by considering the boundary conditions that h_z must satisfy. Since the normal component of the transverse magnetic field h must vanish at the perfectly conducting waveguide wall, (3.716) shows that $n \cdot V_h = 0$ at the walls, where n is a unit normal vector at the walls. When this condition holds, tangential e will also vanish on the guide walls, as (3.71c) shows. The requirements on h_z are thus

$$\frac{d}{dx} h_z = 0 \quad \text{at } x = 0, a$$

$$\frac{\partial}{\partial y} h_z = 0 \quad \text{at } y = 0, b$$

where the guide cross section is taken to be that in Fig. 3.36. In the solution for f , the boundary conditions give

$$-k_x A_1 \sin k_x x + A_1 A_2 \cos k_x x = 0 \quad \text{at } x = 0, a$$

Hence, from the condition at $x = 0$, it is found that $A_2 = 0$. At $x = a$, it is

necessary for $\sin k_x a = 0$, and this specifies k_x to have the values

$$mr \quad 1 = 0, 1, 2, \dots$$

In a similar manner it is found that $B_2 = 0$ and

$$* \bullet = rrnr \quad m = 0, 1, 2, \dots$$

Both n and m equal to zero yields a constant for the solution for h and n other field components; so this trivial solution is of no interest.

If we use the above relations and put $A, B, = A_{nm}$, the solutions for h , are seen to be

$$h = A \cdot \cos_{\frac{a}{a}} \frac{nirx}{a} \cos_{\frac{miry}{a}} \quad (3.201)$$

for $n = 0, 1, 2, \dots$; $m = 0, 1, 2, \dots$; $n = m * 0$. The constant A_{nm} is an arbitrary amplitude constant associated with the nmth mode. For the nmth mode the cutoff wave number is designated k_c , given by

$$\begin{aligned} & \text{c. } \ll m \\ & \text{niry} \quad I \text{ mir}^{-1/2} \end{aligned} \quad (3.202)$$

and is clearly a function of the guide dimensions only. The propagation constant for the nmth mode is given by

$$\begin{aligned} & y_{nm}^{1/2} = JPnm = J\{kl \cdot K.nm\} \\ & f_{\frac{a}{a}}^{27r} \quad mr \quad m^{T-1/2} \end{aligned} \quad (3.203)$$

When $k_0 > k_{c,nm}$, f_{nm} is pure real and the mode propagates; when $k_0 < k_{c,nm}$, then y_{nm} is real but (l_{nm}) is imaginary and the propagation factor is e^{-yn}, U_t which shows that the mode decays rapidly with distance $|z|$ from the point at which it is excited. This decay is not associated with energy loss, but is a characteristic feature of the solution. Such decaying, or evanescent, modes may be used to represent the local diffraction, or fringing, fields that exist in the vicinity of coupling probes and obstacles in waveguides. frequency separating the propagation and no-propagation bands is designated the cutoff frequency $f_{c,nm}$. This is given by the solution of that is,

$$\begin{aligned} & Ic.nm \quad C \quad c \quad riv- \quad rrnr^{-1/2} \\ & 'c.nm \quad 2^{\wedge} \quad et \quad + \end{aligned} \quad (3.204)$$

where c is the velocity of light. The cutoff wavelength is given by

$$A_c = \frac{2ab}{(n^2 + m^2)^{1/2}} \quad (3.205)$$

A typical guide may have $a \sim 26$, in which case

$$A_{c,nm} = \frac{2a}{(n^2 + 4m^2)^{1/2}}$$

and $A_{c,10} = 2a$, $A_{c,01} = a$, $A_{c,..} = 2a/\sqrt{5}$, etc. In this example there is a band of wavelengths from a to $2a$, that is, a frequency band

$$\begin{matrix} c & & c \\ & & \\ 2c & ' & a \end{matrix}$$

for which only the H_{10} mode propagates. This is the dominant mode in a rectangular guide and the one most commonly used in practice. Above the frequency c/a , other modes may propagate; so the useful frequency band in the present case is a one-octave band from $c/2a$ to c/a .

The remainder of the field components for the TE_{im}, or H_{nm} , mode are readily found from (3.201) by using (3.71). The results for the complete $nmth$. solution are

$$H_z = A_{,,,} \cos \frac{mrx}{a} \cos \frac{mny}{b} g^{+*} \quad (3.206a)$$

$$H_x = \pm j - \frac{2}{k} \frac{A_{nm}}{cnm} - \sin \frac{a}{a} \cos \frac{miry}{b} \quad (3.206b)$$

$$H_y = i j p \frac{P_{nm}}{k_{cnm}} \frac{rrmr}{b} \cos \frac{mrx}{a} \sin \frac{mny}{v} \quad (3.206c)$$

$$E_x = Z_{h,nm} A_{,,} m j - \frac{B_{nm}}{a} \frac{rrnr}{b} \cos \frac{nirx}{a} \sin \frac{miry}{b} \quad (3.206a-)$$

$$E_y = -Z_{hiHm} A_{amj-f} \frac{B_{nm}}{kcnm} \frac{nv}{a} \sin \frac{mrx}{a} \cos \frac{miry}{b} e^{-\frac{miry}{v}} W \quad (3.206c)$$

where the wave impedance for the $nmth$ H mode is given by

$$Z_{h,nm} = \lambda Z_0 \quad (3.207)$$

When the mode does not propagate, $Z_{h,nm}$ is imaginary, indicating that there is no net energy flow associated with the evanescent mode. A general field with $E_2 = 0$ can be described in a complete manner by a linear superposition of all the H_{nm} modes.

For a propagating H_{nm} mode the power, or rate of energy flow, in the positive z direction is given by

$$\begin{aligned} & \frac{1}{2} \operatorname{Re} \left(\int_0^a E_x H_z - E_z H_x \right) dx dy \\ & = \frac{1}{2} \operatorname{Re} \left(\int_0^a Z_{nm} E_z^2 + H_x^2 \right) dx dy \quad (3.208) \end{aligned}$$

If we substitute from (3.206d) and (3.206c) and note that

$$\begin{aligned} & \frac{\partial A_{nm}}{\partial x} = \frac{n v x}{a} - \frac{r m n r y}{6} \sin \left(\frac{m \pi y}{a} \right) \quad \text{for } n \neq 0 \\ & = \begin{cases} \frac{a^2}{4} & n = 0, m \neq 0 \\ \frac{a^2}{4} & n \text{ or } m = 0 \end{cases} \end{aligned}$$

we find that

$$\begin{aligned} P_{nm} &= \frac{1}{2} \operatorname{Re} \left(\int_0^a E_z^2 + H_x^2 \right) dy \\ &= \frac{1}{2} \operatorname{Re} \left(\int_0^a \epsilon_{0m} E_z^2 + \frac{R_{f,nm}}{r_m} H_x^2 \right) dy \\ &= \frac{1}{2} \operatorname{Re} \left(\int_0^a \epsilon_{0m} \left| A_n J_{nm} \right|^2 \sin^2 \left(\frac{n \pi x}{a} \right) + \frac{R_{f,nm}}{r_m} \left| H_{nm} \right|^2 \right) dy \quad (3.209) \end{aligned}$$

where ϵ_{0m} is the Neumann factor and equal to 1 for $m = 0$ and equal to 2 for $m > 0$.

If two modes, say the H_{nm} and H_{rs} modes, were present simultaneously, it would be found that the power is the sum of that contributed by each individual mode, that is, $P_{nm} + P_{rs}$. This is a general property of loss-free waveguides. This power orthogonality arises because of the orthogonality of the functions (eigenfunctions) that describe the transverse variation of the fields when integrated over the guide cross section; e.g.,

$$\int_0^a \sin \frac{n \pi x}{a} \sin \frac{r \pi x}{a} dx = 0$$

Even when small losses are present the energy flow may be taken to be to* contributed by each individual mode, with negligible error in all cases except when two or more degenerate modes are present. Degenerate modes are modes which have the same propagation constant y , and for these § presence of even small losses may result in strong coupling between * modes.

If the waveguide walls have finite conductivity, there will be a continuous loss of power to the walls as the modes propagate along the guide. Consequently, the phase constant $j\tilde{f}$ is perturbed and becomes $y = a + j/3$, where a is an attenuation constant that gives the rate at which the mode amplitude must decay as the mode progresses along the guide. For practical waveguides the losses caused by finite conductivity are so small that the attenuation constant may be calculated using the perturbation method outlined in Sec. 3.8 in connection with lossy transmission lines. The method will be illustrated for the dominant H_{10} mode only. For the H_{nm} and also the E_{nm} modes, the calculation differs only in that somewhat greater algebraic manipulation is required.

For the H_{10} mode, the fields are given by (apart from the factor $e^{-Jf_0 v^*}$)

$$\begin{aligned} i_z &= A_{10} \cos \frac{\pi x}{a} \\ h &= j \frac{P_{TM}}{k_c k} A_{10} \frac{\pi}{a} \sin \frac{\pi x}{a} \\ &\approx -z h_{10} \cdot A_{10} \frac{.7010}{a} \sin \frac{\pi x}{a} \end{aligned}$$

as reference to (3.206) shows. From (3.209) the rate of energy flow along the guide is

$$P_{\text{flow}} = I \cdot ab \cdot \frac{1}{4} \int_{-a}^{+a} h_{10}^2 dx$$

The currents on the lossy walls are assumed to be the same as the loss-free currents, and hence are given by

$$J_s = n \times H$$

where n is a unit inward directed normal at the guide wall. Thus, on the walls at $x = 0, a$, the surface currents are

$$\begin{aligned} J_s &= a_x \times H = -a_v A_{10} & x = 0 \\ &= -a_r \times H = -a_v A_{10} & x = a \end{aligned}$$

whereas on the upper and lower walls the currents are

$$\begin{aligned} a \times H &= -a_s \frac{J_{Pi0}}{s} A_{10} \frac{\pi}{a} \sin \frac{\pi x}{a} + a_u A_{10} \cos \frac{\pi x}{a} & \mathbf{y} = 0 \\ -a_v \times H &= a_u \frac{J_{PlQ}}{s} A_{10} \frac{\pi}{a} \sin \frac{\pi x}{a} - a_u A_{10} \cos \frac{\pi x}{a} & y = \end{aligned}$$

With a finite conductivity σ , the waveguide walls may be characterized exhibiting a surface impedance given by

$$Z_{\text{surf}} = \frac{1 + j}{tr8_a} = (\sigma + jk)^{\#},$$

where S_t is the skin depth. The power loss in the resistive part $R = \rho e^2 / 8\pi^2$ per unit length of guide is

$$\begin{aligned} & \int_{\text{guide walls}} J^* dl \\ & RJA_{101} = 2 \int_0^a dy + 2 \int_0^a r \sin \theta dr \int_0^a \frac{J^*}{a} dx + 2 \int_0^a \cos \theta \frac{J^*}{a} dx \end{aligned}$$

Since $A_{101} = \pi a^2$, the above gives

$$P_L = RJA_{101} = \frac{a}{b} \frac{P_w}{c_{10}} + \frac{a}{2}$$

If P_0 is the power at $z = 0$, then $P_0 e^{-2az}$ is the power in the guide at any z . The rate of decrease of power propagated is

$$\frac{dP_w}{dz} = 2aP_0 = P_L$$

and equals the power loss, as indicated in the above equation. The attenuation constant a for the H_{10} mode is thus seen to be

$$\begin{aligned} a &= 2P_0 \left(\frac{a}{b} + \frac{a}{2} \right) \\ &= ab \left(\frac{I_{10}}{I_{10}} \right) \frac{v_{h,w}}{v_{i,o}} \\ &= ab \frac{k_0 Z_0}{(2.68^2 M + a k_0^2) N_p / m} \end{aligned} \quad (3.210)$$

The attenuation for other TE_{nm} modes is given by the formula in Table 3.4, which summarizes the solutions for TE_{111} , and also TM_{nn} modes. In Fig. 3.37 the attenuation for the TE_{10} mode in a copper rectangular guide is given as a function of frequency. To convert attenuation given in nepers to decibels, multiply by 8.686.

The theoretical formulas for attenuation give results in good agreement with experimental values for frequencies below about 5,000 MHz. At higher frequencies, measured values of α may be considerably higher depending on the smoothness of the waveguide surface. If surface imperfections of the order of magnitude of the skin depth 8_s are present, it is read-

TABLE 3.4
Properties of modes in a rectangular guidet

TE modes			TM modes		
tf.-	$\cos \frac{mrx}{a} \cos \theta$	$6^{1/2} e^{-jO_M z}$	0		
E_x	0		$\sin \frac{mrx}{a} \sin \frac{mry}{b} e^{-jn_{TM} z}$		
E_y	$Z_{nm} H_x$		$\frac{-\# \ll \infty, \infty^*}{\cos} \frac{\gg \cdot \cdot \cdot}{\sin} \frac{"f.v}{e^{-j\pi/m^2}}$		
H_z	$J P_{nm} n \sim$	$\frac{mrx}{a} \frac{mry}{b} e^{-j\omega z}$	\ll, Z_{nm}		
H_x	$J P_{nm} mTT$	$\frac{mrx}{a} \sin \frac{mry}{b} e^{-j\omega z}$	$\frac{\xi}{Z_{nm}} e^{j\omega z}$		
$Z_{h,nm} = 3 \text{ nm}^{z_0}$					
$b_{c,nm}$					
P_{nm}					
$*c_{c,nm}$					
$a = 2$					

$tfl_{in} = (\epsilon_0 \mu_0 / 2IT)^{1/2}$, $\epsilon_0 = 1$ for $in = 0$ and 2 for $m > 0$. The expression for ϵ_J is not valid for degenerate modes.

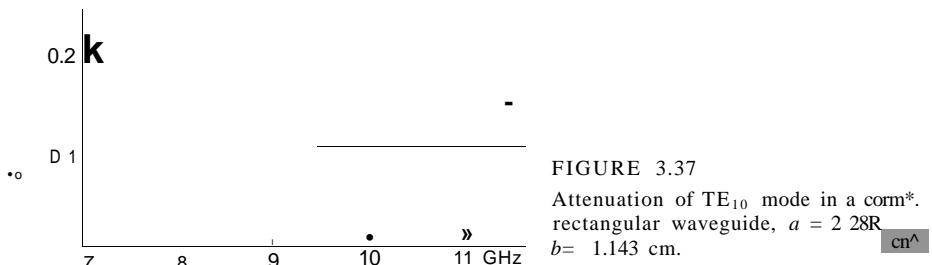


FIGURE 3.37

Attenuation of TE_{10} mode in a corrugated rectangular waveguide, $a = 2.28\text{ cm}$, $b = 1.143\text{ cm}$.

appreciated that the effective surface area is much greater, resulting in greater loss. By suitably polishing the surface, the experimental values of attenuation are found to be in substantial agreement with the theoretical values.^t

Dominant TE_{10} Mode

Since the TE_{10} mode is the dominant mode in a rectangular guide, and also the most commonly used mode, it seems appropriate to examine this mode in more detail. From the results given earlier, the field components for this mode are described by the following (propagation in the $+z$ direction assumed):

$$H = A \cos \frac{\pi x}{a} e^{-\frac{\pi k_z z}{a}} \quad (3.211a)$$

$$H = i A s m \frac{jB}{k_c} e^{\frac{jli}{a}} \quad (3.211b)$$

$$E_y = -j A Z_h \sin \frac{\pi x}{a} e^{-\frac{\pi k_z z}{a}} \quad (3.211c)$$

where the subscript 10 has been dropped for convenience since this discussion pertains only to the TE_{10} mode. The parameters s , k_c and Z_h are given by

$$k_c = \frac{\pi}{a} \quad (3.212a)$$

$$s = \left| k_0 - \left(\frac{Z_h}{a} \right)^{1/2} \right| \quad (3.212*)$$

$$Z_h = -\frac{1}{s} \mathbf{f} = -\frac{1}{s} \mathbf{V} \mathbf{Z} \quad (3.212c)$$

^tSee J. Allison and F. A. Benson, Surface Roughness and Attenuation of Precision Drilled Chemically Polished, Electropolished, Electroplated and Electroformed Waveguides, *Proc. (London)*, vol. 102, pt. B, pp. 251-259, 1955.

The guide wavelength A_g , is

$$A_g = \frac{2\pi r}{\sqrt{1 - (A_0/2r)^2}} \quad (3.212d)$$

since the cutoff wavelength $A_c = 2a$. The phase and group velocities are

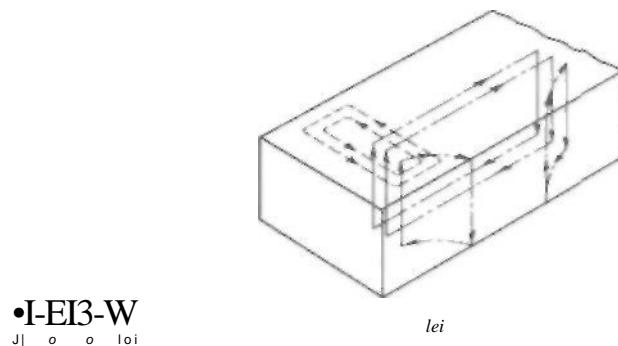
$$v_p = \frac{c}{\sqrt{1 - (A_0/2r)^2}} \quad (3.212c)$$

$$\text{and } A_g = T; \quad (3.212/*)$$

and are discussed in detail in Sec. 3.19.

In Fig. 3.38 the magnetic and electric field lines associated with the TE_{10} mode are illustrated. Note that the magnetic flux lines encircle the electric field lines; so these can be considered to be the source (displacement current) for the magnetic field. On the other hand, the electric field lines terminate in an electric charge distribution on the inner surface of the upper and lower waveguide walls. This charge oscillates back and forth in the axial and transverse directions and thus constitutes an axial and transverse conduction current that forms the continuation of the displacement current. The total current, displacement plus conduction, forms a

$i-r-n-r$



lb)

FIGURE 3.38

Magnetic and electric field lines for the TE_{10} mode, (a) Transverse plane; (b) top view; (c) mutual total current and magnetic field linkages.

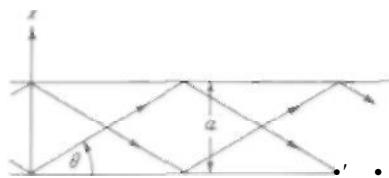


FIGURE 3.39
Decomposition of TE_{10} mode into two plane w

closed linkage of the magnetic field lines, and as such may be regarded being generated by the changing magnetic flux these enclose. This completes the required mutual-support action between the electric and magnet* fields which is required for wave propagation.

The fields for a TE_{10} mode may be decomposed into the sum of two plane TEM waves propagating along zigzag paths between the two waveguide walls at $x = 0$ and $x = a$, as in Fig. 3.39. For the electric field we have

$$E = \frac{\mu_0 t}{2 k_r} (e^{jT^* \langle x - JP^* \rangle} - e^{-jnx \langle x - JPz \rangle})$$

If $17/a$ and (i) are expressed as

$$\frac{\pi}{a} = k_0 \sin \theta \quad B = k_0 \cos \theta$$

the relation $(TT/O)^2 + B^2 = k^2$ still holds. The electric field is now given by

$$Z_h \quad 8 \\ ** \quad 2 \quad A /$$

which is clearly two plane waves propagating at angles $\pm \theta$ with respect to the z axis, as illustrated. Alternatively, the field may be pictured as a plane wave reflecting back and forth between the two guide walls. As shown in Sec. 2.7, the constant phase planes associated with these obliquely propagating plane waves move in the z direction at the phase velocity $c/\cos \theta$ | Bc/k_0 , and this is the reason why the phase velocity of the TE_{10} mode exceeds the velocity of light. Since the energy in a TEM wave propagates with the velocity c in the direction in which the plane wave propagates, this energy will propagate down the guide at a velocity equal to the component of c along the z axis. This component is $v_g = c \cos \theta = (k_0/B)c$ and is * group velocity for the TE_{10} mode. When $\theta = TT/2$, the plane waves reflect back and forth, but do not progress down the guide; so the mode is cut off.

The above decomposition of the TE_{10} mode into two plane waves may be extended to the $TE_{m,n}$ modes also. When n and m are both different from zero, four plane waves result. Although such superpositions of P¹ waves may be used to construct the field solutions for rectangular gu¹, this is a rather cumbersome approach. However, it does lend insight into why the phase velocity exceeds that of light, as well as other properties the modes.

For TM modes, h_z equals zero and e_z plays the role of a potential function from which the remaining field components may be derived. This axial electric field satisfies the reduced Helmholtz equation

$$\nabla^2 e_z + k^2 e_z = 0 \quad (3.213)$$

of the same type encountered earlier for h_z , that is, (3.200). The solution may be found by using the separation-of-variables method. In the present case the boundary conditions require that e_z vanish at $x = 0, a$ and $y = 0, b$. This condition requires that the solution for e_z be

$$e_z = A_{nm} \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{b} \quad (3.214)$$

instead of a product of cosine functions which was suitable for describing h_z . Again, there are a doubly infinite number of solutions corresponding to various integers n and m . However, unlike the situation for TE modes, $n = 0$ and $m = 0$ are not solutions. The cutoff wave number is given by the same expression as for TE modes; i.e.,

$$\frac{n\pi}{a} = \sqrt{\left(\frac{TO}{7T}\right)^2 - 1} \quad (3.215)$$

and the propagation factor γ_{nm} by

$$\gamma_{nm} = (\lambda^2 - k^2)^{1/2} \quad (3.216)$$

The lowest-order propagating mode is the $n = m = 1$ mode, and this has a cutoff wavelength equal to $2ab/\sqrt{a^2 + b^2}$. Note that the TE_{10} mode can propagate at a lower frequency (longer wavelength), thus verifying that this is the dominant mode. It should also be noted that for the same values of n and TO , the TE_{nm} and TM_{nm} modes are degenerate since they have the same propagation factor. Another degeneracy occurs when $a = b$, for in this case the four modes $TE_{11}, TE_{Wf11}, TM_{11},$ and TM_{Wf11} all have the same propagation constant. Still further degeneracies exist if a is an integer multiple of 6, or vice versa.

The rest of the solution for TM modes is readily constructed using the general equations (3.72). A summary of this solution is given in Table 3.4. The TM modes are the dual of the TE modes and apart from minor differences have essentially the same properties. For this reason it does not seem necessary to repeat the preceding discussion.

In any hollow waveguide the dominant mode is a TE mode because the boundary conditions $e_z = 0$ for TM modes always require e_z to have a greater spatial variation in the transverse plane than that for h_z for the lowest-order TE mode, and hence the smallest value of k_c occurs for TE modes. Hence a TE mode has the lowest cutoff frequency, i.e., is the dominant mode.

3.18 CIRCULAR WAVEGUIDES

Figure 3.40 illustrates a cylindrical waveguide with a circular cross section of radius a . In view of the cylindrical geometry involved, cylindrical coordinates are most appropriate for the analysis to be carried out. Since the general properties of the modes that may exist are similar to those for a rectangular guide, this section is not as detailed.

TM Modes

For the TM modes a solution of

$$V, \nabla^2 e_2 + fe_2 = 0$$

is required such that e_2 will vanish at $r = a$. When we express the transverse laplacian ∇^2 in cylindrical coordinates (App. I), this equation becomes

$$\frac{d^2 e_2}{dr^2} + \frac{1}{r} \frac{de_2}{dr} - \frac{1}{r^2} \frac{d^2 e_2}{d\phi^2} = 0 \quad (3.217)$$

The separation-of-variables method may be used to reduce the above to two ordinary differential equations. Consequently, it is assumed that a product solution $f(r)g(\phi)$ exists for e_2 . Substituting for e_2 into (3.217) and dividing the equation by fg yield

$$\frac{1}{r} \frac{d^2 f}{dr^2} + \frac{1}{r} \frac{df}{dr} - \frac{1}{r^2} \frac{d^2 g}{d\phi^2} = k^2 \quad (3.217)$$

Multiplying this result by r^2 gives

$$r^2 \frac{d^2 f}{dr^2} + r \frac{df}{dr} - \frac{1}{r^2} \frac{d^2 g}{d\phi^2} = k^2 r^2 \quad (3.218)$$

The left-hand side is a function of r only, whereas the right-hand side depends on ϕ only. Therefore this equation can hold for all values of the variables only if both sides are equal to some constant v^2 . As a result, (3.217) is seen to separate into the following two equations:

$$\frac{d^2 f}{dr^2} + \frac{1}{r} \frac{df}{dr} - k^2 r^2 = 0 \quad (3.218a)$$

$$\frac{d^2 g}{d\phi^2} = 0 \quad (3.218*)$$

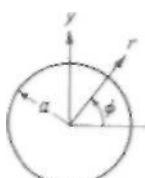


FIGURE 3.40
The circular cylindrical waveguide.

TABLE 3.5
Values of p_{nm} for TM modes

n	P_{nl}	$P \gg 2$	$P \ll 3$
0	2.405	5.520	8.654
1	3.832	7.016	10.174
2	5.135	8.417	11.620

In this case the field inside the waveguide must be periodic in $r f >$ with period $2\pi/\lambda$, that is, single-valued. It is therefore necessary to choose v equal to an integer n , in which case the general solution to (3.218b) is

$$g(f) = A_1 \cos n f + A_2 \sin n f$$

where A_1 and A_2 are arbitrary constants.

Equation (3.218a) is Bessel's differential equation and has two solutions (a second-order differential equation always has two independent solutions) $J_n(k_c r)$ and $Y_n(k_c r)$, called Bessel functions of the first and second kind, respectively, and of order n . For the problem under investigation here, only $J_n(k_c r)$ is a physically acceptable solution since $Y_n(k_c r)$ becomes infinite at $r = 0$. The final solution for e_z may thus be expressed as

$$e_z(r, f) = (A_1 \cos n f + A_2 \sin n f) J_n(k_c r) \quad (3.219)$$

Reference to App. II shows that $J_n(x)$ behaves like a damped sinusoidal function and passes through zero in a quasiperiodic fashion. Since e_z must vanish when $r = a$, it is necessary to choose $k_c a$ in such a manner that $J_n(k_c a) = 0$. If the m th root of the equation $J_n(x) = 0$ is designated p_{nm} , the allowed values (eigenvalues) of k_c are

$$k_{c,nm} = \sim \quad (3.220)$$

The values of p_{nm} for the first three modes for $n = 0, 1, 2$ are given in Table 3.5. As in the case of the rectangular guide, there are a doubly infinite number of solutions.

Each choice of n and m specifies a particular TM_{nm} mode (eigenfunction). The integer n is related to the number of circumferential variations in the field, whereas m relates to the number of radial variations. The propagation constant for the nm th mode is given by

$$\theta_{nm} = *S \cdot \Lambda f \quad (3.221)$$

t Y_n is also called a Neumann function.

the cutoff wavelength by

$$\frac{2\pi c}{P_{c, m}} \quad (3.222)$$

and the wave impedance by

$$\frac{\rho_0 c}{\mu_0} \quad (3.223)$$

A cutoff phenomenon similar to that for the rectangular guide exists. P_0 the dominant TM mode, $A_r = 2\pi r/a = 2.61a$, a value 30 percent greater than the waveguide diameter.

Expressions for the remaining field components may be derived by using the general equations (3.72). Energy flow and attenuation may be determined by methods similar to those used for the rectangular guide. A summary of the results is given in Table 3.6.

TE Modes

The solution for TE modes parallels that for the TM modes with the exception that the boundary conditions require that dh/dr vanish at $r = a$. An appropriate *solution* for A , is

$$h_z(r, \theta) = (B_1 \sin n\theta + B_2 \cos n\theta) J_n(k_c r) \quad (3.224)$$

with the requirement that

$$\frac{dJ_n(k_c r)}{dr} = 0 \quad \text{at } r = a \quad (3.225)$$

The roots of (3.225) will be designated by p'_{nm} ; so the eigenvalues k_{cnm} are given by

$$k_c = \sqrt{\frac{p'_{nm}}{r^2}} \quad (3.226)$$

Table 3.7 lists the values of the roots for the first few modes. Note that $P_{om} = P_{im}$ since $dJ_Q(x)/dx = -J^Q(x)$, and hence the TE_{0m} and TM_{1n} modes are degenerate.

The first TE mode to propagate is the TE_n mode, having a cutoff wavelength $A_{cn} = 3.41a$. This mode is seen to be the dominant mode of the circular waveguide, and is normally the one used. A sketch of the n lines in the transverse plane for this mode is given in Fig. 3.41-ff. The attenuation constant for the dominant TE_U mode is given by

$$R_c = \frac{1.841^2 r^{1/2}}{k a^2} \quad (3.227)$$

TABLE 3.6
Properties of modes in circular waveguides

	TE modes	TM modes
H_z	$\frac{J_{nl} P_{nm}}{a^2}$	$\frac{E_z}{\sin n \theta}$
E_r	$Z_{h,m} H_t$	$\frac{J_{nm} P_{nm}}{\sin n \theta}$
$E_{z,..}$	$Z_{h,nm,Hr}$	$\frac{J_{nm} P_{nm}}{\sin n \theta}$
P_{nr}	$kl - l''$	$A? - \frac{1}{2} \pi^{1/2}$
Power	$\frac{Z_0^2 \rho P_{nm}''}{J_{nm}^2} = \frac{P_{nn}}{Z_0}$	$\frac{P_{nn}}{Z_0}$
A_g	$(p_m)^2 - n^2$	$a^2 - t^2$

TABLE 3.7
Values of p'_{nm} for TE modes

n	P_{n1}	P_{n2}	P_{n3}
0	3.832	7.016	10.174
1	1.841	5.331	8.536
2	3.054	6.706	9.970

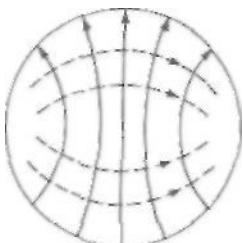


FIGURE 3-41
Field lines for the TE_n mode in a circular waveguide.

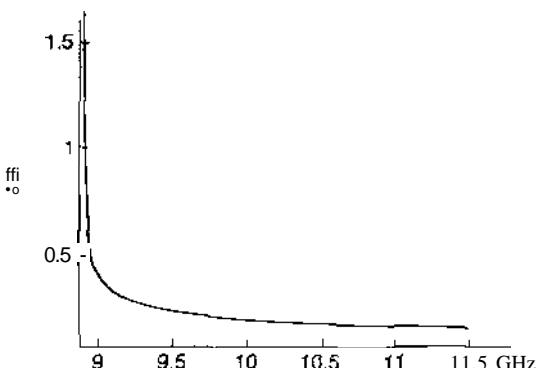


FIGURE 3-42
Attenuation of dominant TE_0 mode in a circular waveguide. Diameter = * cm.

Figure 3.42 shows the attenuation in decibels per meter for a copper waveguide with an internal diameter of 2 cm. For this guide the cut-off frequency is 8.79 GHz. In the normal operating range from 9 to 11 GHz, the attenuation drops from 0.36 dB/m at 9 GHz to 0.11 dB/m at 11 GHz.

19 WAVE VELOCITIES

In any system capable of supporting propagating waves, a number of velocities occur that pertain to signal propagation, energy propagation, wavefront propagation, etc. These various velocities are examined below in the context of propagation in waveguides.

Velocity

The phase velocity of a wave in a waveguide was introduced earlier and shown to be equal to

$$v_p = f \cdot C \quad (3.228)$$

for an air-filled guide. Here k_g is the guide wavelength, A_0 the free-space wavelength, and c the velocity of light. The phase velocity is the velocity an observer must move with in order to see a constant phase for the wave propagating along the guide. It is noted that the phase velocity is greater than the velocity of light, and since the principle of relativity states that no signal or energy can be propagated at a velocity exceeding that of light, the physical significance of the phase velocity might very well be questioned.

The clue to the significance of the phase velocity comes from the recognition that this velocity entered into wave solutions that had a steady-state time dependence of $e^{j\omega t}$. A pure monochromatic (single-frequency) wave of this type exists only if the source was turned on at $t = -\infty$ and is kept *on* for all future time as well. This is clearly not a physically realizable situation. In actual fact the source must be turned on at some finite time, which can be chosen as $t = 0$. The generated signal is then of the form illustrated in Fig. 3.43. Associated with the sudden steplike beginning of the signal is a broad frequency band, as a Fourier analysis shows. If this signal is injected into the guide at $z = 0$, an observer a distance I farther along the guide will, in actual fact, see no signal until a time I/c has elapsed. In other words, the wavefront will propagate along the guide with a velocity c . At the time I/c , the observer will begin to see the arrival of the transient associated with the switching on of the signal. After a suitable period of time has elapsed, the transient will have died out, and the observer will then see the steady-state sinusoidally varying wave. Once steady-state conditions prevail, the phase velocity can be introduced to describe the velocity at which a constant phase point appears to move along the guide. Note, however, that there is no information being transmitted along the guide once steady-state conditions have been established. Thus the phase velocity is not associated with any physical entity such as a signal, wavefront, or energy flow velocity. The term signal is used here to denote a time function that can convey

$Sit)$



FIGURE 3.43
Sinusoidal signal applied at time $t = 0$.

information to the observer. Thus the step change at $t = 0$ is a signal K once steady-state conditions are achieved, the observer does not receive' more information. A better understanding of the above features will vZ obtained after the group velocity, discussed below, has been examined

Group Velocity

The physical definition of group velocity is the velocity with which a sign 1 consisting of a very narrow band of frequency components propagates. Th appropriate tool for the analysis of this situation is the Fourier transform rr a time function is denoted by $f(t)$, this function of time has associated with it a frequency spectrum $F(a)$ given by the Fourier transform of $f(t)$,

$$F(a) = \int f(t)e^{-jat} dt \quad (3.229o)$$

Conversely, if the spectrum $F(a)$ is known, the time function may be found from the inverse Fourier transform relation

$$f(t) = \frac{1}{2\pi} \int F(a)e^{jat} da \quad (3.2296)$$

From Eq. (3.2296) it is seen that the Fourier transform represents $f(t)$ as a superposition of steady-state sinusoidal functions of infinite duration. These relations are a generalization of the Fourier series relations. If the time function is passed through a device having a response $Z(w)$ that is a function of frequency, e.g., filter, the output time function $f_0(t)$ will have a frequency spectrum $F(a)Z(a)$, and hence, by (3.2296), is given by

$$f_0(t) = \int F(a)Z(a)e^{jat} da$$

In general, $Z(a) = |Z(a)|e^{-j\phi(a)}$; so

$$f_0(t) = \int F(a)|Z(a)|e^{j(a-t)} da \quad (3.230)$$

If the output f_0 is to be an exact reproduction of the input, then in (3.230) it is necessary for $|Z(a)|$ to equal a constant A , and $\phi(a)$ must be a linear function of a , say $aw + b$. In this case

$$f_0(t) = \frac{A}{2\pi} \int F(a)e^{j(a-wt-b)} da \quad (3.231)$$

If we put $t \sim a = t'$, the above becomes

$$f_0(t) = \frac{A}{2\pi} \int F(a)e^{j(a-wt'-b)} da = A e^{-jw(t'-b)} \quad (3.231)$$

as comparison with (3.2296) shows. Thus the output time function is

$$U^* + \ll -W) = Ae^{-\alpha t} an = Ae^{-\alpha t} f(t - a) \quad (3.232)$$

i.e., an exact duplicate of the input, apart from a constant multiplier and a time delay a . Thus the conditions given on $|Z|$ and i/α are those sufficient for a distortion-free system.

Now, in a waveguide, the transverse variations of the field are independent of frequency. The only essential frequency-dependent part of the field solution is the propagation factor e^{jPx} since

$$0 = (fc - *f)^{1/a} = (\wedge - ft?)$$

is a function of frequency. Thus a waveguide of length l , in which the field has a time dependence $e^{jP_0 t}$, $w > 0$, can be considered as a frequency filter with a response $e^{-\alpha l}$. Since αl is not a linear function of to , it may be anticipated that some signal distortion will occur for propagation in a waveguide. For an ideal TEM-wave transmission line, $\alpha l = k_0 = io/c$ and distortion-free transmission is possible. However, practical lines have an attenuation which depends on frequency (R_m and αl), and this will produce distortion. Fortunately, for narrowband signals neither waveguides nor transmission lines produce significant distortion unless very long lines are used.

Consider now a time function $f(t)$ having a band of frequencies between $-f_m$ and $+f_m$. This signal is used to modulate a carrier of frequency fc resulting in a resultant signal $S(t) = f(t)cos co_c t = Re[f(t)e^{j\omega_c t}]$

$$S(t) = f(t)cos co_c t = Re[f(t)e^{j\omega_c t}] \quad (3.233)$$

If $F(fco)$ is the spectrum of $f(t)$, the spectrum of $S(t)$ is

$$F_s(w) = f J \int_{-\infty}^{\infty} f(t) e^{-j\omega_c t} dt, \\ = -[F(ai - aj_c) + F(a > + w_c)] \quad (3.234)$$

These spectra are illustrated in Fig. 3.44.

For positive co the waveguide response is $e^{-\alpha l}$. For negative w the response must be chosen as $e^{\alpha l}$ since, if the time variation is e^{-jQt} , the sign in front of (i) must be positive for propagation in the positive z direction. In other words, all physical systems will have a response such that $|Z(w)|$ is an even function of co and $iZ(w)$ is an odd function of co . Since P is an even function, the sign must be changed for $co < 0$. These even and odd symmetry properties are required simply so that the output time function is real, a physical requirement. The output spectrum for the

$$F(u) \quad F_s i(u)$$



FIGURE 3.44
Spectrum of $\frac{d}{dt}$ ($/$ land Sit).

waveguide is thus

$$F_{ss}(u) = |[F(w - to)e^{jw^3/\lambda'} + F(a > -r) e^{jPM}]|$$

and the output signal is

$$S_{ss}(t) \sim -f F_0(u) e^{-jdu} \quad (3.235)$$

The analysis that follows is simplified if the signal is represented in complex form as $f(t)e^{j\omega_c t}$ with a spectrum $F(u - co_c)$. In this case only the positive half of the spectrum needs to be considered, and the output signal is given by

$$S_{ss}(0) = \text{Re} - C' * F(v - \omega_c) e^{-j\omega_c w} da \quad (3.236)$$

since $F(a > -w_c)$ is zero outside the band $w_c - w_m < w < w_e + w_m$. If the band is very narrow, $w_{ss} \ll w_c$, then $\frac{d}{d\omega} F(\omega)$ may be approximated by the first few terms in a Taylor series expansion about a_e . Thus

$$P(a) = \frac{d}{da} F(a) \approx \frac{1}{2} \frac{d^2 F}{da^2} (w - w_c)^2 + \dots \quad (3.237)$$

Retaining the first two terms only and letting $\frac{d}{d\omega} F(\omega_c) = 0_O$ and $d(5/du)_c = P_O$ at w_c , (3.236) gives

$$S_0(t) = \text{Re} - r^{+w_m} F(w) - a J e^{j(t - M e - M + J f V * da)}$$

If this is compared with (3.231) and (3.232), it is seen that

$$\begin{aligned} S_0(t) &= \text{Re} e^{j(t - M e - M + J f V * da)} - P'_0 l e^{jw_c(t - P'_0 l)} \\ &\quad - \mathbf{fit} - \mathbf{M}^* \gg \mathbf{t}^* \mathbf{j} - \mathbf{M} \end{aligned} \quad (3.238)$$

To the order of approximation used here, the input modulating signal $/$ is reproduced without distortion but with a time delay $P'_0 L$. This is anticipated since $\frac{d}{d\omega} F(\omega)$ was approximated by a linear function of ω in the

$\omega_c - (\omega_m \text{ to } \omega_c + \omega_m)$ (distortion-free condition). The signal delay defines the group velocity v_g which is equal to the distance / divided by the delay time; thus

$$v_g = \frac{1}{W^*} \frac{\mathbf{M} \mathbf{y}^T}{da} \quad (3.239)$$

This is also the signal velocity. Note, however, that this velocity has significance only if the band, or "group," of frequencies making up the signal is so narrow that f_i may be approximated by a linear function throughout the frequency band of interest. If this is not the case, more terms in the expansion (3.237) must be retained and signal distortion will occur. In this case the group velocity as given by (3.239) is no longer the signal velocity. In fact, because of signal distortion, no unique signal velocity exists any longer. Different portions of the signal will travel with different velocities, and the resultant signal becomes dispersed in both time and space.

In the case of a waveguide,

$$v_g = \frac{d(c\omega^2/c^2 - k_0^2)^{1/2}}{d\omega} \quad (3.240)$$

c $\ll 0$ ω_0 ω_g

It is seen that $v_g < c$ and that $v_g v_p = c^2$ for a waveguide.

A typical plot of k_0 versus β for a waveguide is given in Fig. 3.45. From this plot it can be seen that for a narrow band of frequencies a linear approximation for β is good. Also note that for high frequencies (large k_0) β becomes equal to k_0 . Thus frequencies well above the cutoff frequency f_e suffer very little dispersion and propagate essentially with the velocity of light c . No frequency components below the cutoff frequency f_e can propagate along the guide.

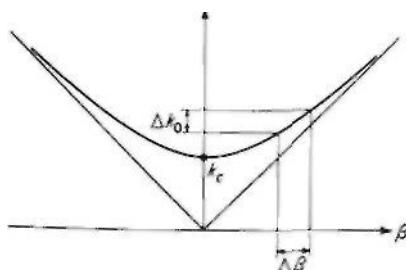


FIGURE 3.45
Plot of k_0 versus β for a waveguide.

The equality of the wavefront velocity and the velocity of light can be readily explained by means of Fig. 3.45. The switching on of a signal r_{es} in an initial transient that has a spectrum of frequencies extending from 0 to infinity. Any small group of frequencies at the high end of the spectrum have a group velocity equal to c since dk_0/dp equals unity for $k \rightarrow \infty$. Thus the high-frequency part of the transient will propagate along the guide with a velocity c. The lower-frequency components will propagate with smaller group velocities and arrive later.

Energy-Flow Velocity

Power is a flow of energy, and consequently there is a velocity of energy g_{0w} such that the average energy density in the guide multiplied by this velocity is equal to the power. In a waveguide it turns out that this velocity of energy flow is equal to the group velocity. A proof for the case of E modes will be given below, that for H modes being very similar.

For E modes the field is given by [see (3.72)]

$$\mathbf{H} = \frac{P}{\omega} \mathbf{E}, \quad \text{where } P = \frac{k_0 Y_0}{2} E^2 \omega$$

The average rate of energy flow, or power, is given by

$$P = \int_{\text{cross section}} \mathbf{E} \cdot \mathbf{H} dS = \frac{1}{2} k_0 Y_0 \int_{\text{cross section}} E^2 \omega dS \quad (3.241)$$

where the integration is over the guide cross section.

The energy density in the magnetic field per unit length of guide is

$$\frac{1}{4} \mu_0 H^2 = \frac{1}{4} \mu_0 \left(\frac{P}{k_0 Y_0} \right)^2 = \frac{P^2}{4 k_0 Y_0} \quad (3.242)$$

The energy density in the electric field per unit length of guide is equal to that in the magnetic field. This is readily shown to be the case by using the complex Poynting vector theorem, which states that (Sec. 2.5)

$$\int_S \mathbf{E} \cdot \mathbf{H}^* dS = P + 2j\omega(W_m - W_e)$$

where the integration is over the guide cross section, and the term on the right gives the power transmitted past the plane S plus $2j\omega$ times the reactive energy stored in the guide beyond the plane S. Since the integral of the complex Poynting vector over a cross section for a propagating wave in a loss-free guide is real, it follows that $W_m = W_e$. In addition, since the location of the transverse plane S is arbitrary, it also follows that energy densities U_m and U_e per unit length of guide are equal.

The velocity of energy flow may now be found from the relation

$$v = \frac{P}{U_{\text{eff}} + U_{\text{ext}}} = \frac{P}{2t/\epsilon_0} = \frac{k_0 Y_0}{P} \frac{H^2}{M \sigma W} \quad (3.243)$$

and comes out equal to the group velocity as stated earlier.

RIDGE WAVEGUIDE

For a rectangular waveguide with a width a equal to twice the height ft, the maximum bandwidth of operation over which only the dominant TE_{10} mode propagates is a 2:1 band. For some system applications it is necessary to have a waveguide that operates with only a single mode of propagation over much larger bandwidths. A transmission line supporting only a TEM mode can fulfill this requirement but must then have cross-sectional dimensions that are small relative to the shortest wavelength of interest. A coaxial transmission line will support higher-order TE and TM modes in addition to the TEM mode. Thus, to avoid excitation of a higher-order mode of propagation, the outer radius must be kept small relative to the wavelength. The small cross section implies a relatively large attenuation; so some other form of waveguide is needed. The ridge waveguide illustrated in Fig. 3.46 was developed to fulfill this need for a single-mode waveguide capable of operating over a very broad band. Physically, it is easy to understand why the ridge waveguide has a very large frequency band of operation. The center section of width W and spacing S functions very much like a parallel-plate transmission line and consequently the ridge waveguide has a much lower cutoff frequency for the same width and height as does the

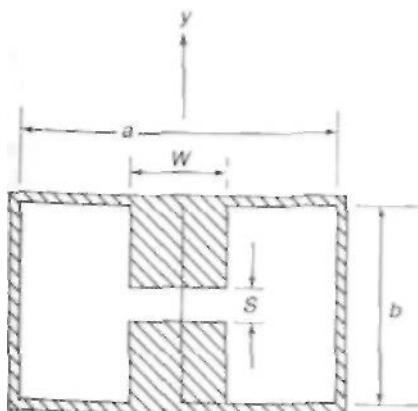


FIGURE 3.46
Ridge waveguide.

conventional rectangular waveguide. Operation over bandwidths of 5% more is possible.

The ridge waveguide, when uniformly filled with dielectric, which may be air, has the same general properties as the rectangular and circular waveguides discussed earlier. If we can determine the cutoff wavelength for the dominant mode, then at any frequency above the cutoff frequency ω_c the propagation constant jk is given by

$$P = yjkl - (2tr/X_c f)$$

At the cutoff frequency ($i = 0$) and the electromagnetic field has no variation with the axial coordinate z . The cutoff wave number $k_c = 2ir/\lambda_c$ can be found using the transverse-resonance method as described below.

The transverse-resonance method is based on finding the resonant frequency for the transmission-line circuit that provides a model for the waveguide cross section. At cutoff we can view the electromagnetic field as a uniform plane wave with components E_y and H_z that propagates in the x direction and is incident onto a second parallel-plate transmission line of reduced height. The equivalent transmission-line circuit is that of two parallel-plate transmission lines of height b , length $(a-W)/2$, and short-circuited at $x = 0$ and a . These two transmission lines are connected to another parallel-plate line of height S and length W and placed between the first two as shown in Fig. 3.47. At the junction where the height changes, a local fringing electric field occurs and stores electric field energy in the vicinity of the step. The effect of this local fringing electric field is taken into account by a shunt capacitive susceptance jB at each junction as shown in Fig. 3.47.

The standing-wave field pattern along the x direction can exist only at the resonant frequency for the transmission-line circuit shown in Fig. 3.47. For the dominant mode the voltage is a maximum and the current is zero at the midsection. Thus, at $x = a/2$, the impedance looking in the x direction toward the short circuit must be infinite. The corresponding input admittance will be zero. At the step the admittance looking toward the short circuit is

$$Y = -jY_{ci} \cot k_c \frac{a-W}{2} + jB$$

By using the formula for admittance transformation along a transnu

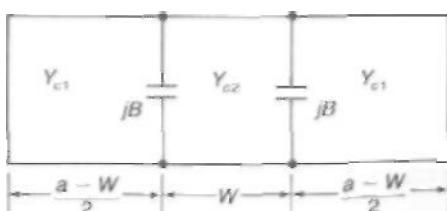
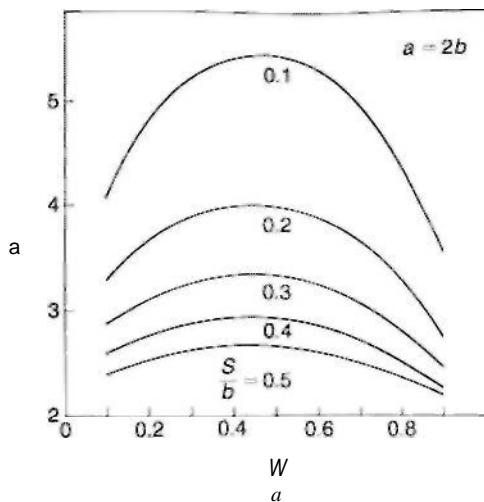


FIGURE 3.47
Equivalent transmission-line circuit of a section of ridge waveguide,

**FIGURE 3.48**

Normalized cutoff wavelength K_c/a for a ridge waveguide.

line, we obtain

$$Y_{-..} = Y_{c2} \frac{Y + jY_{c3}}{Y_{c2} + jY} \tan k_e W/2$$

In order for Y to vanish, we must have

$$jB - jY_{cl} \cot k_t \frac{a-W}{a} + jY_{c2} \tan k_e \frac{W}{2} = 0 \quad (3.244)$$

which is the transverse-resonance condition. The two characteristic admittances are inversely proportional to the parallel-plate spacing; thus $Y_{c2} = (b/S)Y_{ci}$. An approximate expression for the normalized shunt capacitive susceptance can be found using quasistatic conformal mapping and is

$$B = \frac{26}{1 - \ln 4u + \frac{1}{-u^2} + \left(1 - \frac{1}{u^2}\right)^{\frac{1}{\lambda}}} \quad u = \frac{a}{a-W} < 0.5 \quad (3.245)$$

The eigenvalue equation (3.244) is a transcendental equation. The computer program RIDGEWG solves (3.244) for the normalized cutoff wavelength A_c/a .

Figure 3.48 shows typical results that are obtained. The numerical results obtained from (3.244) agree within 1 percent of the values given by Hopfer and Hoefer and Burton over the commonly used range of parameters.¹

¹N. Marcuvitz, "Waveguide Handbook," MIT Radiation Lab Series, vol. 10, reprinted by Boston Technical Publishers, Inc., 1964.

²S. Hopfer, The Design of Ridged Waveguides *IRE Trans.*, vol. MTT-3, pp. 20-29, October, 1955.

W. J. R. Hoefer and M. N. Burton, Closed Form Expression for the Parameters of Finned and Ridged Waveguides, *IEEE Trans.*, vol. MTT-30, pp. 2190-2194, December 1982.

3.21 FIN LINE

If the width W of the ridges in the ridge waveguide is very small, we obtain a fin line as shown in Fig. 3.49a. Usually the fins are metal foils on a thin dielectric substrate mounted in the E plane of a rectangular waveguide for the dominant mode the current flows in the axial direction; so good electrical contact between the fins and the waveguide is not essential. The fin line is a shielded slot line. The fin line can be matched to the rectangular waveguide by means of a tapered section or by using one or more quarter wave impedance transformers as shown in Figs. 3.49b and c.

The fin line is suitable for use in microwave circuits that incorporate two-terminal devices such as diodes. Transistors cannot be connected to a fin line since they are three-terminal devices.

The cutoff wavelength for a fin line may also be found by using the transverse-resonance method. The fins produce a shunt normalized capacitive susceptance across the center of the waveguide given by

$$\frac{B}{26} \left| \begin{array}{c} \text{XT} \\ \ll 2 \quad \alpha i^T \bullet \quad n \end{array} \right| = \frac{1}{\pi} \left(\frac{\epsilon - *}{\epsilon + *} \bullet \right) \quad (3.246)$$

where

$$A = \ll i$$

$$P_2 = 2a| + a^* - 1$$

$$P_3 = 4a? + \%a_{xal} - 3_{U1}$$

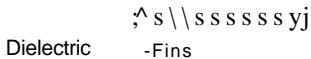
$$P_4 = 8a? + 3a^+ + 24ajfa| - 8a? - 4a| + 1$$

$$"1 = \cos^{2^{rs}} Jb$$

$$a_2 = sm —$$

$$r_{,,} = \left| WHxf \right|^{1/2}$$

The equivalent circuit of the fin-line cross section consists of two short circuited transmission lines of length $a/2$ with jB connected at the center



bsgc^ssss^sss^{5 s 5 v 5 • 5 5 !}

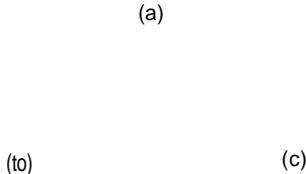


FIGURE 3.49

(a) Fin line; (b) tapered matching section; (c) quarter-wave matching section.

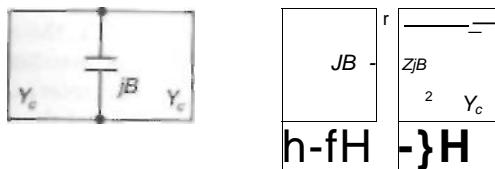


FIGURE 3.50

Equivalent circuit of fin-line cross section.

shown in Fig. 3.50. The resonance condition is

$$\frac{Y_r}{K} = \frac{jB}{2K} \cdot \frac{\cot \frac{ira}{J}}{A} = 0 \quad (3.247)$$

The computer program FINLINE solves this transcendental equation for the normalized cutoff wavelength A_c/a for $S/b = 0.1$ to 0.9 . Typical numerical results are shown in Fig. 3.51 for the case where $a = 26$. These numerical results agree within 1 percent or better with those given by Hoefer and Burton.^t

The propagation constant f_i is given by the same formula as for a conventional waveguide, i.e.,

$$e = (H - k^*_c)^2$$

When the fins are mounted on a dielectric substrate, a correction is needed for the propagation constant if the dielectric has an appreciable thickness

^tW, J. R. Hoefer and M. N. Burton. *loc. cit.*

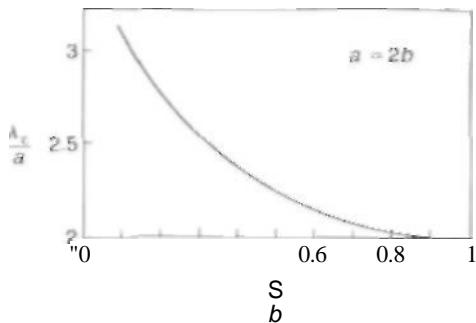


FIGURE 3.51
Normalized cutoff wavelength λ_c/a for 50% dielectric-filled waveguide.

and a large dielectric constant. Empirical formulas for this case are available.

PROBLEMS

- 3.1. For the ideal transmission line shown in Fig. P3.1, the switch is closed at $t = 0$ and opened 1 μ s later. The characteristic impedance of the line is 50 Ω . The load resistance is also 50 Ω . The battery has an internal resistance of 10 Ω .
- Sketch the voltage across R_L as a function of time for a line 300 m long. The wave velocity $c = 3 \times 10^8$ m/s.
 - Sketch the voltage waveform across R_L when $R_L = 25 \Omega$ and the line is 900 m long.
 - Sketch the voltage waveform across R_L when $R_L = 100 \Omega$ and the line is 900 m long.
 - Repeat (6) and (c) for a line 75 m long.

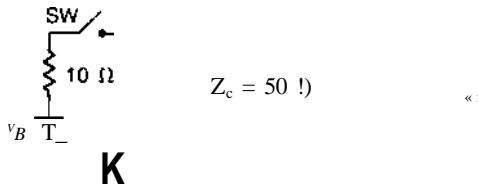


FIGURE P3.1

- 3.2. Let a generator with internal resistance R_g be connected to a transmission line of length l and having a characteristic impedance Z_c . The line is terminated in a load resistance R_L . Let $r = l/v$ be the one-way propagation time delay. The generator produces a pulsed waveform $p(t)$, $0 < t < T$. Show that the voltage across R_L is given by

$$V_{RL} = \frac{1}{Z_c} [R_g + Z_c] p(t - r) + \frac{1}{Z_c} V_g p(t - 2r) + \frac{1}{Z_c} R_g p(t - 3r)$$

Hint: See (3.12) and consider the total line voltage at $z = l$.

- 3.3. A pulse generator produces a sawtooth pulse $P(t) = LOt/T$, $0 < t < T$, where $T = 1(T^8)$ s. The generator has an internal resistance $R_g = 200 \Omega$ and is connected to a transmission line with $Z_c = 50 \Omega$. The line is l meters long and terminated in a load resistance R_L . The wave velocity equals 3×10^8 m/s.
- (a) Find and sketch the load voltage as a function of time when $l = 3$ m and $R_L = 200 \Omega$.
 - (b) Repeat (a) when $R_L = 12.5 \Omega$.
 - (c) Find an analytic expression for the voltage across R_L when $R_L = 200 \Omega$ and $l = 12$ m.
 - (d) Make a distance-time plot of the line voltage for (c).
- 3.4. A battery with voltage of 10 V is connected in series with a 50-JJ resistor to the input of a 50-fl transmission line at time $t = 0$. The transmission line, of length 12 m, is terminated in a 1-jF capacitor. Find and sketch the voltage across the capacitor as a function of time.

Hint: Apply Thevenin's theorem.

Answer:

$$V_c = 10[1 - e^{-t/(l/c)}]^{1/2} \quad / \text{ in microseconds}$$

- 3.5. In the circuit illustrated in Fig. P3.5, the battery is connected at $t = 0$. Find and sketch the voltage across R_L as a function of time. Assume that $R_L = R_g = Z_c = 50 \Omega$, $C = 1 \text{ pF}$, $I = 300 \text{ mH}$, and $v = 3 \times 10^8 \text{ m/s}$.

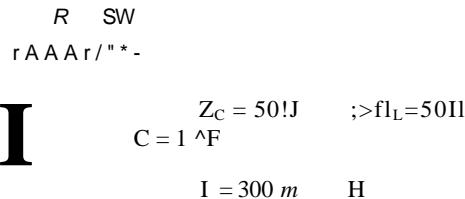


FIGURE P3.5

- 3.6. The resistor R_g is replaced by a capacitor $C_L = 1 \text{ pF}$ in Fig. P3.5. Find the voltage across C_L during the time interval $1 \text{ us} < t < 3 \text{ us}$.

Answer:

$$V_L = 5[l + e^{-(t/l)^2}] - 10e^{-(t/l)^2}$$

where t is in microseconds.

- 3.7. Consider an ideal loss-free transmission line of length l , as shown in Fig. P3.7. The far end is short-circuited. At the input end a battery of voltage V_0 is switched across the line at time $t = 0$. Sketch the voltage wave on the line at the middle over the time interval $0 < t < ll/c$.

-2

FIGURE P3.7

r
"or"
1—VW—o
*0

$$Z_c = R_a$$

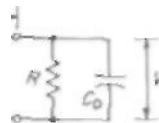


FIGURE P3.8

- 3.8. Consider the transmission-line circuit illustrated in Fig. P3.8. At time $t = 0$, a battery of voltage V_0 is switched across the line at the input. Determine the output load voltage V as a function of time.

Hint: This transient problem may be solved in a manner similar to that used in low-frequency circuit theory. The governing equations for the transmission line are

$$\frac{\partial^2 V}{\partial z^2} - \frac{B \cdot J \cdot F}{U} \frac{3}{Hz} = -Q \frac{dV}{dt},$$

The time derivative may be eliminated by taking the Laplace transform. The transformed solutions for V and S are

$$Y \stackrel{J}{=} e^{-W} + V \sim e^{-sW} \quad V e^{-sW} + I \sim e^{sW}$$

By transforming the circuit equations for the load termination and the input voltage, the resultant equations may be solved for the Laplace transform of the load voltage. The inverse transform then gives the load voltage as a function of time.

The foregoing procedure may be simplified by first replacing the battery by a source $V_0 e^{-sW}$ and obtaining the transfer function $V/V_0 = Z_s(ju)$ as a function of u for this steady-state problem. Replacing ju by s then gives the transfer function in the s domain. The Laplace transform of the output voltage is then

$$V(s) = -Z_s \frac{(s)}{s}$$

since the Laplace transform of the input step voltage is V_0/s . The output voltage is obtained from the inverse Laplace transform of $V(s)$.

- 3.9. Obtain expressions for the voltage and current standing-wave patterns on a lossless open-circuited transmission line. Sketch these patterns. Assume an r time dependence.

- 3.10. A transmission line with $Z_c = 50$ ft is terminated in an impedance $25 + j11$. Find the reflection coefficient, standing-wave ratio, and fraction of the incident power delivered to the load.

Answer: $VSWR = 2.618$, 80 percent.

- 3.11. Verify (3.47) and compute Z_{in} at a distance $A_0/4$ from the termination given in Prob. 3.10.

- 3.12. On a transmission line with $Z_c = 50$ ft, the voltage at a distance $0.4A_0$, the load is $4 + j2$. The corresponding current is 0.1 A. Determine the normalized load impedance.

Answer: $0.145 + j0.397$.

- 3.13. A 50-ft transmission line is terminated by a 75-ft load resistor. Find the distance z from the load at which $Y_{in} = 0.02 + j0.01$. By connecting a $\lambda/4$

susceptance $-jB$ across the line at this point, the load will be matched to the transmission line. Explain why this is so. The distance λ can be expressed in terms of the wavelength λ_0 .

- 3.14.** Figure P3.14 illustrates a three-conductor transmission line. Since potential is arbitrary to within an additive constant, the shield S_0 , can be chosen to be at zero potential. Show that there are two linearly independent solutions for TEM waves in this transmission line. If S_0 encloses N conductors, how many TEM-wave solutions are possible?

Hint: Note that the potential can be arbitrarily specified on S_1 and S_2 .

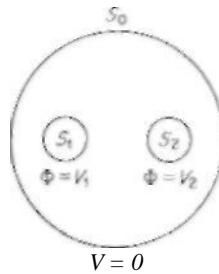


FIGURE P3.14

- 3.15.** Show that power transmitted along a transmission line is given by

$$\mathbf{P} = \mathbf{y}/|\mathbf{V}_{\langle t \rangle}|^2 \mathbf{d}^* \mathbf{dy}$$

For Prob. 3.14 show that this equals $jfV_1/\lambda + V_2I_2$ by using Green's first identity (App. I) to convert the surface integral to a contour integral around the conductor boundaries. I_x and λ , are the total currents on S_1 and S_2 .

- 3.16.** Consider a three-conductor transmission line as shown in Fig. P3.14 but assume that the cross sections of S_1 and S_2 are not the same. Let Φ_a and Φ_b be two different solutions for the potential field. For Φ_a , let V_{ab} , λ_{ab} , and $\lambda_{2Q} \Delta \lambda_{2Q}$ be the voltage and currents on S_1 and S_2 . Similarly, for Φ_b , let V_{ba} , λ_{ba} , and $\lambda_{2Q} \Delta \lambda_{2Q}$ be the voltages and currents on S_1 and S_2 . For the TEM modes derived from $\Phi_a + \Phi_b$, show that the power flow is given by

$$HKi + v_{ln}(i_{a1} + i_{b1}) \cdot |(\mathbf{V}_{Q2} + \mathbf{v}_{z2})(\lambda_{a2} + z_{b2})|$$

It is convenient to choose the potentials so that the two TEM modes obtained from Φ_a and Φ_b have independent power flow. Show that this will be the case if the interaction term

$$(V_{al}I_{bl} + V_{hl}I_{al}) + (V_{o2}I_{b2} + V_{b2}I_{a2})$$

equals zero. Furthermore, show that the interaction term will vanish if the potentials are chosen so that

$$\begin{aligned} & \mathbf{C}_u \\ & K.2 " \quad v_{b2} \quad \mathbf{C}_0, + \mathbf{C}_1, \end{aligned}$$

where C_u is the capacitance between S_1 and S_0 , C_{22} is the capacitance between S_2 and S_0 , and C_{12} is the capacitance between S_1 and S_2 . For a

*4

symmetrical line, $C_u = C_{22}$ and the two modes correspond to the even odd modes.

Hint: Use the relations $\phi_{11} = C_M V_{el} + C_{12}(V_{B1} - V_{o2})$, $Q^A = \epsilon C_{12}(V_{o2} - V_{aj})$ and similar ones for the total charge Q_{hl} and Q_{11} , or $\phi_{11} = j^{22} Q_{11}$ in terms of V_6 , and V_{k2} . Also note that $\phi_{11} = (y_0/e_0)Q_{11}$, etc.

- 3.17.** Show that, for an air-filled coaxial line, minimum attenuation occurs when $x = 1 + x$, $x = b/a$. What is the corresponding characteristic impedance?
Hint: Hold the outer radius b constant and find da/da .
- 3.18.** Evaluate Z_c for a lossy coaxial line using (3.28) and computed values of R , L , and C . Assume $\epsilon_r = 3a = 1$ cm, $f = 10^9$ Hz, $a = 5.8 \times 10^{-7}$ S/rad, and $\epsilon = (2.56 - j0.0006)e_0$. Verify that

$$\text{Im } Z_c \ll \text{Re } Z_c \quad \text{and} \quad Z_c = \frac{M}{C}^{1/2}$$

See Table 3.1 for coaxial line parameters.

- 3.19.** Use the energy definitions of L and C [Eqs. (3.112)1 to derive the results given by (3.106) and (3.108) for a coaxial transmission line.
- 3.20.** A microstrip line has a substrate 1 mm thick and with a dielectric constant $\epsilon_r = 8$. The strip width $W = 2.5$ mm. Find the low-frequency effective dielectric constant and characteristic impedance.
Answer: $\epsilon_e = 5.953$, $Z_c = 32.13$ ohms.
- 3.21.** A microstrip line uses an anisotropic dielectric substrate with $\epsilon_r = 10$ and $\epsilon_z = 8$. The substrate is 0.5 mm thick and the strip width $W = 0.75$ mm. Find the low-frequency effective dielectric constant and characteristic impedance.
Answer: $\epsilon_e = 5.895$, $Z_c = 42.97$ ohms.
- 3.22.** A microstrip line has a 1-mm-thick dielectric substrate with a dielectric constant of 6. Use the computer program MSTP to generate data giving the effective dielectric constant and characteristic impedance as a function of strip width W . Use these data to design a microstrip system with an input line having $Z_r = 50$ ohms, an output line having $Z_c = 75$ ohms, and an intermediate quarter-wave transformer section with $Z_r = 50 \times 75$ ohms. Specify the three strip widths W_1 , W_2 , W_3 and the length of the quarter-wave transformer at 8 GHz $f = 2$ GHz (see Fig. P3.22).
Answer: Widths are 1.505 mm, 1.0125 mm, 0.649 mm, length = 1.1 cm.

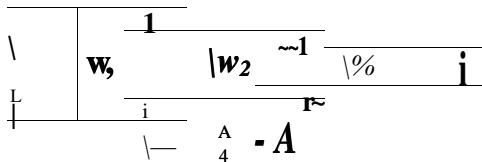


FIGURE P3.22

- 3.23.** Find the effective dielectric constant, characteristic impedance, and attenuation at 2 GHz for a microstrip line with the following parameters: $\epsilon_r = 10$, $\epsilon_z = 8$, $t = 2 \times 10^{-3}$ mm, $L = 1$ mm, strip width $W = 0.75$ mm, and strip thickness $T = 0.01$ mm. Use the computer program MSTP.

- 3.24.** Use the computer program MSTPD to generate dispersion data for the effective dielectric constant for a microstrip line having the following parameters: dielectric constant = 6, substrate thickness $H = 0.5$ mm, and strip width $W = 2$ mm. At what frequency has the effective dielectric constant increased by 5 percent more than the quasistatic value?
- 3.25.** In a monolithic microwave integrated circuit, gallium arsenide with $e_r = 12.9$ is used as a substrate material. Find the effective dielectric constant, characteristic impedance, and attenuation at 10 GHz for a microstrip line with the following parameters: substrate thickness = 0.1 mm, strip width = 0.05 mm, strip thickness = 0.002 mm, and loss tangent = 6×10^{-8} . For these dimensions the quasistatic parameters are accurate. The computer program MSTP can be used for the evaluation. What is the attenuation in decibels per wavelength (microstrip) for this microstrip line?
- 3.26.** A microstrip line has the following parameters: strip width $W = 1$ mm, substrate thickness = 1 mm, and anisotropic dielectric with $e_r = 6.5$, $e_v = 6$. Find the following: distributed capacitance C and inductance L per centimeter, characteristic impedance, effective dielectric constant, and the quasi-TEM-mode wavelength at 2 GHz.
- 3.27.** Use the computer program CMSTP to find the even- and odd-mode characteristic impedances and the voltage coupling coefficient C for a coupled microstrip line having the following parameters: strip width $W = 1$ mm, strip spacing $S = 0.1$ mm, substrate thickness = 1 mm, and substrate dielectric-constant = 9.7.
- 3.28.** A strip line has a ground-plane spacing of 2 mm, a strip width of 1 mm, and is filled with a dielectric medium with dielectric constant 2.3. Find the characteristic impedance.
- 3.29.** Use the computer program STPL to evaluate the characteristic impedance and attenuation of a strip line with the following parameters: ground-plane spacing = 2 mm, strip width $W = 0.5$ mm, strip thickness $T = 0.01$ mm, dielectric constant = 6, loss tangent = 0.006, and frequency of operation = 5 GHz. What is the ratio of the attenuation due to dielectric loss relative to that of conductor loss?
- 3.30.** A broadside coupled strip line is required for a 3-dB directional coupler. The even-mode characteristic impedance is to be 50 Ω . The voltage coupling coefficient is 0.707. From this information determine the required odd-mode characteristic impedance. Find the required strip width W and spacing S for this coupled strip line. The ground-plane spacing is 4 mm and the dielectric constant of the dielectric filling is 5. The strip thickness $T = 0.05$ mm. Use the computer program CSTPL. You will need to follow an iterative procedure to arrive at the required parameters. An accuracy of ± 0.5 percent is adequate.
- Hint:** Begin with $W = 3.5$ mm, $S = 0.5$ mm.
- 3.31.** In a monolithic microwave integrated circuit, a coplanar transmission line with the following parameters is used: strip width $S = 0.1$ mm, slot width $W = 0.1$ mm, strip thickness = 0.002 mm, substrate thickness = 0.5 mm, dielectric constant = 12.9, loss tangent = 0.0008, and frequency = 10 GHz. Use the computer program CPW to determine the characteristic impedance

and attenuation. If the strip thickness is increased to 0.005 mm, will it significantly reduce the attenuation?

- 3.32. Figure P3.32 shows a coplanar-transmission-line circuit for use in a v. amplifier circuit. The required input- and output-fine chara f r ^ impedances are 50 ohms and 72 ohms. The impedance of the quarter-wave section $\sqrt{50 \times 72} = 60$ ohms. The ground-plane spacing $2W + S$ is kept constant at 4 nun. The strip thickness is 0.002 mm. The substrate thickness is 0.41 * 0.3

the dielectric constant is 12.9 with a loss tangent of 0.001. Use the com" ^ program CPW to determine the required strip widths S_1 , S_2 , and S_r . Determine the length l of the quarter-wave matching section at a frequency f GHz. How much attenuation occurs in the quarter-wave section? An accuracy of ± 0.25 percent is sufficient.



FIGURE P3.32

- 3.33. In a planar transmission line, the attenuation is 0.25 dB/cm. By what fraction is the wave amplitude reduced in propagating a distance of 1 cm on this transmission line?
- 3.34. Derive the equations (3.72) for TM waves.
- 3.35. Find the cutoff frequency for the TE_{10} mode in a rectangular waveguide with dimensions 4 cm by 2 cm. Find the guide wavelength A_g and phase velocity at a frequency 25 percent higher than the cutoff frequency.
- 3.36. Derive the solution for a TE_{10} mode in a rectangular guide of wide dimension a and height b when the guide is filled with dielectric of permittivity ϵ_r . Show that the cutoff frequency is given by $f_c = c/2ae_r^{1/2}$, where c is the free-space velocity of light and ϵ_r is the dielectric constant. Show that the guide wavelength is smaller for a dielectric-filled guide than for an air-filled guide.
- 3.37. Obtain an expression for the attenuation of a TE_{10} mode in a dielectric-filled guide when $\epsilon = \epsilon_1 - j\epsilon_2$ but the walls are perfectly conducting. Obtain the exact expression and compare it with the results deduced by an approximate perturbation method.
- 3.38. Obtain a solution for an H wave in the parallel-plate transmission line centered dielectric slab as illustrated in Fig. P3.38. Assume that the plate is perfectly conducting and infinitely wide. Can a TEM wave propagate in this structure? Why?
- Hint:* Assume $h_x = \cos k_x x$ for $|x| < a/2$ and $h_x = Ae^{-k_x x}$ for $x > a/2$. Verify that $k_x^2 + p_x^2 = (\epsilon_r - D^2/\omega^2)^{1/2}$. Match the tangential fields at $x = a/2$ to obtain an equation for A and one relating the parameters p_x and D .

*0 «b

FIGURE P3.38

- 3.39. Obtain solutions for TE_{0,0} modes in the partially filled waveguide illustrated in Fig. P3.39.

Hint: Assume that

$$\mathbf{ft} = \begin{cases} \cos(Li) & 0 < x < t \\ A \cos(p(a - x)) & t < x < a \end{cases}$$

and match the tangential fields at x = t. Thus show that

$$\mathbf{0}^2 = kl - p^2 = \mathbf{e}_r^* \mathbf{g} - kl$$

and that

$$p \tan k_d t = -k_d \tan p d$$

Note that there are an infinite number of solutions for p and k_d corresponding to various TE_{0,0} modes. Obtain numerical values for /3, p, and k_d when k₀ = 2, t = 1 cm, d = 1.5 cm, and e_r = 4. Note that there is a lowest-order solution for p pure imaginary.

60

FIGURE P3.39

- 3.40. Obtain expressions for the surface currents of a TE₁₀ mode in a rectangular guide. A narrow slot may be cut in a waveguide along a current flow line without appreciably disturbing the field. Show that, for the TE₁₀ mode, narrow centered axial slots may be cut in the broad face of a rectangular guide. This principle is used in standing-wave detectors to provide suitable points of entry for a probe used to sample the interior waveguide field.
- 3.41. Use the computer program RECTWG to evaluate the parameters of a rectangular waveguide with width a = 1 cm, height b = 0.4 cm at a frequency of 20 GHz. How much attenuation will occur in a waveguide 5 m long?
- 3.42. For TE modes in a waveguide, write H_x = -I(z)V_h, E_y = V(z)a_z X V_h. Use Maxwell's equations to show that V_{iz}(z) and I(z) satisfy the transmission-

line equations

$$\frac{cV}{az} = -i\omega M_0^2 \quad \frac{dl}{az} \quad \text{at } z = 0, \dots + 777^1 - |V| \quad \text{at } z = L$$

Construct an equivalent distributed-parameter circuit for these modes p
TM modes put $E_z = -V(2)V_e, H_z = -JU/a_2 X V/e_{il}$ and show that

$$\frac{rfV}{dz} / \quad A_z^2 \backslash \quad d/z$$

Construct an equivalent distributed-parameter circuit for these modes.t

- 3.43. Consider an infinitely long rectangular guide. The guide is filled with dielectric for $z > 0$, having a dielectric constant ϵ_r . An H_w mode is incident from $z < 0$. At $z = 0$, a reflected r_{10} mode and a transmitted H_{10} mode are produced because of the discontinuity. Show that the reflection coefficient is given by $(Z_2 - Z_1)/(Z_2 + Z_1)$, where Z_1 is the wave impedance in the empty guide and Z_2 is the wave impedance in the dielectric-filled guide. Show that the ratio of the wave impedances equals the ratio of the guide wavelengths.
- 3.44. Find the surface currents for the H_{nl} mode in a circular guide.
- 3.45. Obtain an expression for power in a TE_n mode in a circular guide. (See App. II for Bessel-function integrals.)
- 3.46. Derive an expression for attenuation for $TE_{l,m}$ modes in a circular waveguide.
Answer: $a = R_m f_c^2 [aZ_0 f^l - l^2]$.
- 3.47. Find the attenuation in decibels per mile for an H_{lm} mode in a circular copper guide of 1 in diameter when operated at a frequency of 10 times the cutoff frequency.
- 3.48. Show that, in a coaxial line with inner radius a and outer radius b , there are solutions for $TE_{l,m}$ and $TM_{l,m}$ modes. A suitable solution for e_z and k_z is

$$[AJ_{l,m}(k_z r) + J_{l,m}(k_z r)] \cos n \phi$$

Obtain equations (transcendental in nature) for determining the cutoff wave number k_z by imposing proper boundary conditions at $r = a, b$.

- 3.49. Use the computer program RIDGEWG to find the cutoff wavelength and frequency for a ridge waveguide with dimensions $a = 1$ cm, $b = 0.4$ cm, width = 0.5 cm, and ridge spacing $S = 0.1$ cm. Compare this with the cutoff frequency of a standard waveguide with $a = 1$ cm and $b = 0.4$ cm.
- 3.50. Use the transverse-resonance technique to derive the eigenvalue equation for $TE_{l,m}$ modes in the partially filled rectangular guide of Prob. 3.39. Verify that the wave impedances in the x direction in the two regions are $k_0 Z_0/P$
 $k_z/k_d = k_0 Z_0/k_d$.
- 3.51. A rectangular waveguide with internal dimensions $a = 0.9$ in and $b = 0.4$ in (standard X-band waveguide) has a centered fin with a slot spacing $* \sim$ mm. Find the cutoff wavelength and compare it with that for the waveguide without fin loading. Use the computer program FINLINE.

- 3.52. The permittivity ϵ is generally a function of $e(u)$ of ω . Obtain an expression for the group velocity of a coaxial line filled with dielectric. Neglect the frequency dependence of the attenuation due to conductor loss.
- 3.53. A rectangular guide of dimensions $a = 26 = 2.5$ cm is operated at a frequency of 10^{10} Hz. A pulse-modulated carrier of the above frequency is transmitted through the guide. How much pulse delay time is introduced by a guide 100 m long?

REFERENCES

1. Ramo, S., J. R. Whinnery, and T. Van Duzer: "Fields and Waves in Communication Electronics," 2nd ed., John Wiley & Sons, Inc., New York, 1984.
2. Liboff, R. L., and G. C. Dalman: "Transmission Lines, Waveguides, and Smith Charts." MacMillan Publishing Company, New York. 1985.
3. Pozar, D. M.: "Microwave Engineering," Addison-Wesley Publishing Company, Reading, Mass. 1990.
4. Rizzi, P. A.: "Microwave Engineering—Passive Circuits," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1988.
5. Edwards, T. C.: "Foundations for Microstrip Circuit Design," John Wiley & Sons, Inc., New York, 1987.
6. Bahl, I., and P. Bhartia: "Microwave Solid State Circuit Design." John Wiley & Sons, Inc., New York, 1988.
7. Chang, K. (ed.): "Handbook of Microwave and Optical Components," vol. 1. John Wiley & Sons. Inc., New York, 1989.
8. Ishii, T. K.: "Microwave Engineering," 2nd ed., Harcourt Brace Jovanovich, New York, 1989.
9. Wolff, E. A., and R. Kaul: "Microwave Engineering and Systems," John Wiley & Sons, Inc., New York, 1988.
10. Baden Fuller, A. J.: "Microwaves," 3rd ed., Pergamon Press, New York, 1990.

CHAPTER 4

CIRCUIT THEORY FOR WAVEGUIDING SYSTEMS^t

At low frequencies the interconnection of resistors, capacitors, and inductors results in a circuit. Such circuits are normally linear, so that the superposition principle may be used to find the response when more than one exciting source is present. Kirchhoff's laws form the basis for the analysis, whether in terms of loop currents or node voltages. In these low-frequency circuits the various elements are connected by conducting wires, and generally the length of these connecting wires is not critical or important.

At microwave frequencies equivalent reactive and resistive elements may also be connected to form a microwave circuit. In place of connecting wires, transmission lines and waveguides are used. The length of the connecting link is often several wavelengths, and hence propagation effects become very important. The analysis of microwave circuits is therefore necessarily somewhat more involved than that for the low-frequency case.¹ The circuit theory of transmission-line circuits has been well developed many decades, and, as will be shown the circuit theory for waveguide systems is formally the same.

^tThe basic theory of microwave circuits is developed in C. G. Montgomery, B. H. Purcell, "Principles of Microwave Circuits," McGraw-Hill Book Company, 1948. Much of the material presented here in Secs. 4.1 to 4.9 must of necessity be sketched over its basic nature.

Many of the circuit-analysis techniques and circuit properties that are valid at low frequencies are also valid for microwave circuits. In actual fact, low-frequency circuit analysis is a special case of microwave circuit analysis. As a consequence, a study of microwave circuits provides a deeper physical insight into conventional circuit theory. In this chapter the physical basis for a circuit theory for waveguiding systems is developed. In later chapters we shall utilize this foundation in the study of impedance matching, waveguide devices, resonators, filters, etc.

EQUIVALENT VOLTAGES AND CURRENTS

At microwave frequencies voltmeters and ammeters for the direct measurement of voltages and currents do not exist. For this reason voltage and current, as a measure of the level of electrical excitation of a circuit, do not play a primary role at microwave frequencies. On the other hand, it is useful to be able to describe the operation of a microwave circuit in terms of voltages, currents, and impedances in order to make optimum use of low-frequency circuit concepts. For the most part this can be done. There is, however, a notable difference, namely, the nonuniqueness of the voltages and currents in most instances. It was noted in the preceding chapter that for the TEM wave on a transmission line there existed a voltage and a current wave uniquely related to the transverse electric and magnetic fields, respectively. In the case of TE and TM modes in a waveguide, no unique voltage or current waves exist that have the same physical significance as those associated with the TEM wave on a transmission line. This result might have been anticipated since the guide boundary is a closed conducting boundary, and one is at a loss as to the two points on the boundary between which the voltage should be measured. Furthermore, if voltage is defined as the line integral of the transverse electric field between two chosen points on the boundary it is found that for TM waves the line integral is zero (Probs. 4.1 and 4.2), whereas for TE waves the value of the line integral depends on the path of integration that is chosen. For these reasons the introduction of voltage and current waves, to be associated with waveguide modes, is done on an equivalent basis and has formal significance only. The basis for the introduction of equivalent voltages and currents is discussed below.

In the previous chapter it was shown that propagating waveguide modes have the following properties;

1. Power transmitted is given by an integral involving the transverse electric and transverse magnetic fields only.
2. In a loss-free guide supporting several modes of propagation, the power transmitted is the sum of that contributed by each mode individually.

3. The transverse fields vary with distance along the guide according to propagation factor $e^{\pm jkz}$, only.
4. The transverse magnetic field is related to the transverse electric field L by a simple constant, the wave impedance of the mode; i.e.,

$$Z_{\text{eff}} h = a \cdot X e$$

for a mode propagating in the $+z$ direction.

These properties suggest letting equivalent voltage and current waves be introduced proportional to the transverse electric and magnetic fields respectively, since the transverse fields have properties similar to those of the voltage and current waves on a transmission line. That is, in actual fact what is done.

A propagating waveguide mode may be expressed in general as

$$E = C^* e e^{jPz} + C^+ e_z e^{jQz} \quad (4.1)$$

$$H = \{T h - e^{-*^2} + C^\wedge e - e^{-\wedge^2} \quad (4.16)$$

for propagation in the $+z$ direction, and

$$E = C \sim e e^{jP'} - C \sim e_z e^{jP*} \quad (4.2a)$$

$$H = \sim \langle M w^* \rangle + C \sim i_z e^{jQz} \quad (4.26)$$

for propagation in the $-z$ direction. In (4.1) and (4.2), C^+ and $C \sim$ are arbitrary amplitude constants. Note also that if the mode is a TE or a TM mode, then e_z and h_z is zero accordingly. Let the following equivalent voltage and current waves be introduced:

$$V = f^* e^{-Jf i^*} + V \sim e^{jP'} \quad (4.3a)$$

$$I \sim I^+ e^{-jPz} - I \sim e^{jPz} \quad (4.36)$$

where $V^* = K_f i^*$, $V \sim = K_x C \sim$, and $I' = K_2 C^\wedge$, $I = K_1 C \sim$. K_f and K_2 are constants of proportionality that will establish the relationship between voltages and the transverse electric field and currents and the transverse magnetic field. In order to conserve power, it is necessary that

$$\lambda_z^+ (I^+)^* = \frac{C}{l} \int_S e x h^* - e_z^* dS$$

or

$$K^\wedge K Z = J e x h^* - a_z dS \quad (4.4)$$

By proper normalization of the functions e and h , the product $K_1 K_2 |e|^2$ is made equal to unity. Although (4.4) provides one relationship between L_h and K_2 , a second relationship is required before they are determined. A second relationship can be chosen in a variety of ways. For example-

voltage and current waves given by (4.3) may be thought of as existing on a fictitious transmission line that is equivalent to the waveguide. As such, it may be desirable to choose the characteristic impedance of this transmission line equal to unity, in which case

$$V^+ \quad V\sim \quad K,$$

As an alternative, it might be desirable to choose the characteristic impedance equal to the wave impedance, in which case

$$z_+ = \frac{V^+}{I^+} = \frac{V\sim}{I\sim} = z, \quad (4-6)$$

Other possibilities are obvious and equally valid. In this text either the definition (4.5) or (4.6) is used. The one that is used will be stated, or else it will be clear from the discussion which definition is being utilized. When the equivalent voltages and currents are chosen so that the equivalent transmission-line characteristic impedance is unity, we shall refer to them as *normalized* voltages and currents. Note that even though equivalent transmission lines may be used to represent a waveguide, the propagation constant of this line must be taken as that for the waveguide.

A waveguide supporting N propagating modes may now be formally represented as N fictitious transmission lines supporting equivalent voltage and current waves (from property 2 listed above for waveguide modes). Thus we have

$$V = \sum_{n=1}^N (V_n e^{-\lambda n} + V_n^* e^{\lambda n}) \quad (4.1a)$$

$$= \sum_{n=1}^N t_n (K e^{-\lambda n} - I_n e^{\lambda n})$$

$$= \sum_{n=1}^N E_n (Y_n e^{-\lambda n} - Y_n^* e^{\lambda n}) \quad (4.1b)$$

where the Y_n are arbitrarily chosen characteristic admittances for the equivalent transmission lines. When an obstacle is inserted into a waveguide supporting N modes of propagation, these modes are in general coupled together by the obstacle. This coupling can be described in terms of an equivalent circuit made up of impedance elements. This impedance description of obstacles in waveguides is developed in the next section. Once the equivalent voltage and current amplitudes have been determined, the wave-

guide fields are known from the relations

$$n = 1$$

$$N$$

$$n = 1$$

and the specified proportionality constants K_{1n} , K_{2n} for each mode. The axial field components may be found from (4.8) by the use of Maxwell's equations. Note that the equivalent current wave amplitude for propagation in the $\sim z$ direction is expressed by $-I_\sim$, and hence the corresponding transverse magnetic field is proportional to $-K_{1r} I_\sim$. When a waveguide supports several modes of propagation simultaneously at the same frequency, the number of electrical ports will exceed the number of physical ports. That is, power can be fed to a given load by means of any of the propagating modes, and all these modes may be common to a single physical waveguide input port.

4.2 IMPEDANCE DESCRIPTION OF WAVEGUIDE ELEMENTS AND CIRCUITS

One-Port Circuits

A one-port circuit (equivalent to a two-terminal network) is a circuit for which power can enter or leave through a single waveguide or transmission line. A short-circuited transmission line and a short-circuited waveguide containing a metallic post as illustrated in Fig. 4.1 are examples of one-port circuits.

For one-port devices of the above type, a knowledge of any two of the four quantities V^* , V_\sim , $V = V^* + V_\sim$, $I = I^+ - I^-$ will serve to describe the effect of the one-port device on an incident wave (it is assumed that the waveguide supports only one propagating mode). These quantities must, of course, be referred to a terminal plane such as t in Fig. 4.1 in order to be unambiguously specified. A terminal plane, or reference plane, is the equivalent of a terminal pair in a low-frequency network. In the present instance an impedance description is desired. If the total voltage and current at the

^tThe microwave equivalent-circuit theory presented in this chapter may be extended to include nonpropagating modes. However, when nonpropagating modes are included, the impedance and scattering matrices do not have the same properties as when only propagating modes are present at the terminal planes. See H. Haskal, Matrix Description of Waveguide Discontinuity in the Presence of Evanescent Modes, *IEEE Trans.*, vol. MTT-12, pp. 184-188, March, 1964.

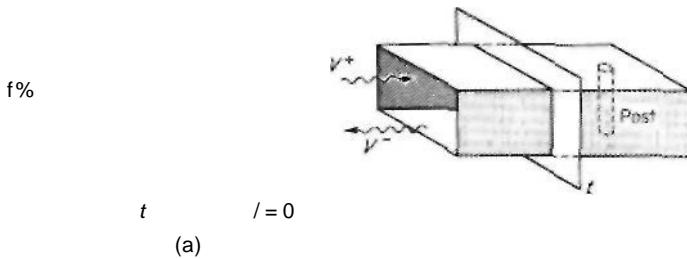


FIGURE 4.1

(a) Short-circuited coaxial line; (b) short-circuited waveguide with post.

terminal plane are

$$\begin{aligned} \mathbf{y} &= v^+ + V^- \\ i \sim r - r &= Y_c(v^+ - v^-) \end{aligned}$$

where Y_c is the equivalent characteristic admittance (actual characteristic admittance for the transmission line), the input impedance is given by

$$Z_m = \frac{V}{I} = \frac{V^+ - V^-}{V^* - V^-} Z_c \quad (4.9)$$

The complex Poynting vector may be used to establish the physical nature of one-port impedance functions. From (2.59) and (2.60) we have

$$-\int_0 E \times H^* \cdot n dS = P_r + 2jw(W_m - W_r) \quad (4.10)$$

where n is a unit inward normal to the closed surface S , P_r is the power dissipated in the volume bounded by S , and $W_m - W_r$ is the net reactive energy stored within S . For the surface S choose the terminal plane, the guide walls, and the short-circuiting plane. For perfectly conducting walls and short circuit, $n \times E = 0$; so the integral reduces to that over the terminal plane only. Thus

$$-\int_0 E \times H^* \cdot a_n dS = P_r + 2jw(W_m - W_r) \quad (4.11)$$

Now at the terminal plane the transverse fields are [see (4.8)]

$$E_x - KJ^l(V_r + V_t) = K^2 V_r \quad (4.12a)$$

$$H_z - J^l / - / - J^l h^* = J^l q / J^l h \quad (4.12b)$$

Hence (4.11) becomes

$$-(K^2 K t^2 V l^2 j e_x \cdot h^* \cdot a_n dS) = i v \cdot Z^* = P_r + 2jw(W_m - W_r) \quad (4.13)$$

If now V is replaced by IZ^A , we find that

$$Z_{in} = \frac{P_{in} + 2jw(W_m - W_e)}{(IP)} = R + jX \quad (4.14)$$

This relates the input impedance to the power loss and net reactive energy stored in the volume beyond the terminal plane. Inasmuch as the current may be an equivalent current, the corresponding impedance Z is an equivalent one also. Since P_{in} , W_m , and W_e are all proportional to $|V|^2$ — hence also proportional to $|J|^2$ in view of the linearity of the field equations — the equivalent resistance R and reactance X in (4.14) are independent of the amplitude of the incident wave.

By replacing J^* by Y^*V^* in (4.13), we obtain, after taking the complex conjugate,

(4.15)

for the input admittance of the one-port device. The susceptance B is positive (capacitive in nature) only if $W_e > W_m$.

The evaluation of an input impedance by means of the general definition (4.14) will be carried out for the simplest case, that of a short-circuited coaxial line. In the short-circuited coaxial line of Fig. 4.1, the fields in the one-port device are given by

$$E = \frac{V^*}{\ln(6/a)} \frac{a}{r} (e^{-j\alpha z} - e^{jk_0 l})$$

$$H = \frac{n/b/a}{\ln(b/a)} \frac{v}{r})$$

since the electric field must vanish at the short-circuited position $z = l$. If the terminal plane is located at $z = -l$, then

$$4 \propto 0 \quad J_a \quad J_l$$

$$\frac{27re_0|V^+|^2}{\ln(6/a)} \frac{-}{l} / \sin^2 k_0 z dz$$

$$\frac{\text{TT}6_0|V^-|^2}{l} \frac{\sin 2k_0 l}{2k_n}$$

Similarly, it is found that

$$W_m = \frac{1}{\ln(6/a)} \frac{\sin 2A_0}{I + \frac{2A_{in}}{2A_{in}}}$$

$$\ln(6/a)$$

The total current at the terminal plane at $z = -l$ is

$$I = Y_c V^+ (e^{jk_0 z} + *-**>) = 2Y_c V \cos k_0 l$$

Using (4.14) now gives

$$\begin{aligned} Z_m &= \frac{4ja > ire_0 |V^*|^2}{\ln(6/a)} \quad sm2k_0 l \\ &= \frac{j(OTre_0)}{A_0(4Y_r^2 |V^+|^2) \cos^2 V} \quad (416) \end{aligned}$$

on using the relations $Y_c = 2vY_0/[Jn(b/a)]$, $sm2k_0 l = 2 \sin kj \cos k_0 l$, and $k_0 = w/(i_0 e_0)^{1/2}$. This result for the input impedance of a short-circuited coaxial line, as obtained from the general definition (4.14), agrees with the simple computation based directly on expressions for the total voltage and current at the terminal plane. However, the purpose of introducing (4.14) was not as a computational tool, but rather for the physical insight it provides into the nature of the impedance function for a one-port circuit.

The second example of a one-port circuit as illustrated in Fig. 4.1 cannot be evaluated in as straightforward a manner as for the coaxial line because it does not consist of a uniform unperturbed waveguide. The presence of a conducting post within the termination results in induced currents on the post that will excite a multitude of waveguide modes. However, since it is assumed that only one mode propagates (the TE₁₀ mode), all the other modes decay exponentially in both directions away from the post. By choosing the terminal plane sufficiently far away from the post, the fields at this plane are essentially just those of the incident and reflected dominant modes. The evanescent modes excited by the post will store reactive energy, and this will contribute to the input reactance as viewed from the terminal plane, as reference to (4.14) shows. The presence of the post within the termination modifies the input reactance by changing the amplitude of the reflected dominant wave in just the right amount to account for the additional reactive energy stored.

As seen from the preceding discussion, it is important when dealing with waveguide structures to choose terminal planes sufficiently far away from obstacles that excite evanescent modes, so that only dominant-mode fields have significant amplitudes at these reference planes. This will ensure that all the reactive energy associated with the nonpropagating modes that make up the fringing field around the obstacle is taken into account in the expression for the input reactance. This precaution is particularly important in any experimental setup used to measure the impedance function for a particular obstacle. Once the impedance has been properly determined at a given terminal plane, it may be referred to any other terminal plane by

using the impedance-transformation formula

$$Z(l) = Z_0 \frac{Z_0 + iZ_{\text{eff}} f(l_2 - l_1)}{Z_0 + iZ(l) \tan[B(l_0) \tan(l_2 - l_1)]} \quad (4)$$

where l_x is the location of terminal plane 1 and l_2 specifies the location of the new terminal plane. In particular, shift in the terminal-plane position*¹! a multiple of $A_g/2$ leaves the impedance invariant. Thus an impedance can be referred to terminal planes located in the near vicinity of an obstacle where now it is understood that this impedance describes the effect of the obstacle on the dominant mode only, and does not imply that the total field at this particular terminal plane is that of the dominant mode only. In other words, the impedance description of a waveguide element or obstacle gives information on the effect this element has on the dominant propagating mode but does not give any information on the detailed field structure near the obstacle. Fortunately, the latter information is rarely required.

Lossless One-Port Termination

If there are no losses present in a one-port circuit, the input impedance is a pure reactance given by

$$jX^{\wedge} J^{\wedge} f31 \quad (18)$$

The assumption of a lossless structure is often a very good approximation for microwave circuits. If $W_m = W_e$, the input reactance vanishes and a condition of resonance exists. There are actually two possibilities, namely, $W_m = W_e$ but $I \neq 0$, and $W_m = W_e$ but $V \neq 0$. The first corresponds to a zero in the input reactance (series resonance), whereas the second corresponds to a zero in the input susceptance (parallel resonance) as given by

$$-\frac{*MW, -W_m}{VW^*} \quad (19)$$

When the input reactance is zero, the input susceptance must be infinite, which implies that $V = 0$ at the terminal plane. This latter condition is possible since, for a pure reactive termination, all the incident power is reflected, so that the total voltage along the waveguide is a standing wave, in the form $\sin(l/3)$. In the case of a zero for the susceptance function B , the reactance X must be infinite (have a pole) and J must vanish at the terminal plane. It may be anticipated, then, that the reactance and susceptance functions will have a number of zeros and poles, i.e., frequencies at which they vanish or become infinite. This behavior is clearly evident in the expression for the reactance of a short-circuited coaxial line, which is

$$jX = jZ_c \tan k_0 l = jZ_c \tan \frac{\pi}{c} \quad (*\bullet)$$

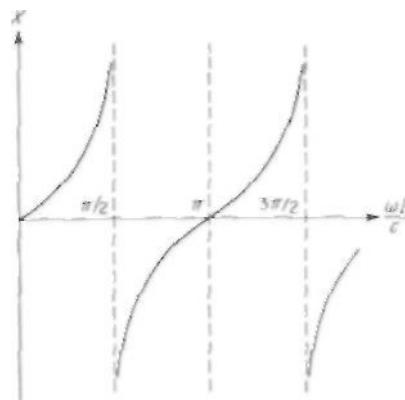


FIGURE 4.2
Input reactance of a short-circuited coaxial line.

A plot of X against frequency is given in Fig. 4.2. In particular, note that the slope of the reactance is *always* positive; that is, $dX/d\omega l > 0$. This positive-slope condition means that the poles and zeros of X must alternate as ωl is increased from zero to infinity. We shall show below that this is a general property of any reactive one-port circuit, a result known as Foster's reactance theorem. First, however, it will be instructive to rewrite (4.20) by using the infinite-product representation and also the partial-fraction expansion of the tangent function^a

$$\begin{aligned}
 X = 2c_0 \tan k_0 l &= \prod_{n=0}^{\infty} \frac{(ay)}{(n + i)^2 + b^2} \\
 &= Z_0 \cdot \prod_{n=1}^{\infty} \frac{1}{(n - 1)^2 + (k_0 l)^2} \quad (4.21)
 \end{aligned}$$

The first form contains the product of an infinite number of factors in both the numerator and denominator and clearly exhibits both the zeros and poles and their alternating occurrence. The second form exhibits the poles very clearly, but information on the zero locations is lost. In the vicinity of a pole, say that at $w = w_p = nvc/2l$, all terms in the partial-fraction expan-

^aE. A. Guillemin, "The Mathematics of Circuit Analysis," chap. 6, John Wiley & Sons, Inc., New York, 1949.

E. T. Copson, "Theory of Functions of a Complex Variable," Oxford University Press, Fair Lawn, N.J., 1935.

J. Pierpont, "Functions of a Complex Variable," Dover Publications, Inc., New York, 1955.

sion are small except the n th term, so that

$$X \ll \frac{2wc}{I(c_0, - u>)(io_n + a>)} \frac{Z_0}{l(w, - w)} \quad (4.22)$$

since $w \sim u>$. This behavior near a pole is similar to that for a simple i_C parallel network for which

$$X = \frac{-coL}{LC} \quad (4.23)$$

where $w_0 = (LC)^{1/2}$. However, the microwave network is a good deal more complicated, for it has an infinite number of poles and zeros, and not just double zero and a single pole as a simple parallel LC circuit has [the zero occur at $w = 0$, where wL vanishes, and at infinity, where $(wC)^{-1}$ vanishes]. These similarities and differences are important to note since they are characteristic of microwave networks in general, even though we have demonstrated some of these properties for a short-circuited coaxial line only.

*4.3 FOSTER'S REACTANCE THEOREM

The theorem that will now be proved is that the rate of change of the reactance X and susceptance B with $o>$ is positive. Once this result is established, it follows that the poles and zeros of a reactance function must alternate in position along the $o>$ axis. Figure 4.3 illustrates a general one-port reactive termination. The field within the termination satisfies Maxwell's equations

$$\nabla \times E = -ju\omega iH \quad \nabla \times H = ja\omega eE$$

The derivative with respect to w of the complex conjugate of these equations gives

$$V X \frac{\partial E}{\partial \omega} = ja\omega iH - \frac{3H^*}{diss} + ;H \quad \nabla X \frac{\partial H}{\partial \omega} = -jwe - \frac{dE^*}{y} - \frac{9a\epsilon}{y} E^* -$$

Consider next the quantity

$$\begin{aligned} \mathbf{E} \times \frac{(9H^*)}{(do)} \frac{dE^*}{(Ho)} X H &= V X E \frac{dE^*}{(do)} - E \cdot V X \frac{\partial H^*}{(00)} \\ &+ V X \frac{r?E^*}{(do)} H \frac{dE^*}{(do)} - V X H \end{aligned}$$

$V'', -r$



FIGURE 4.3
A one-port reactive termination,

Substituting from above gives

$$\mathbf{V} - \mathbf{E} \cdot \mathbf{X} = \frac{\langle \mathbf{H}^* \rangle}{dt} + \frac{\langle \mathbf{E}^* \rangle}{dt} - \frac{\mathbf{I}}{I} = j \frac{\mathbf{H} \cdot \mathbf{H}^*}{dt} + \frac{\mathbf{E} \cdot \mathbf{E}^*}{dt} - \frac{\mathbf{I} \cdot \mathbf{SB}^*}{dt} + j \mathbf{eo} / i \mathbf{H} \cdot \frac{\mathbf{H}^*}{dt} + \mathbf{eE} \cdot \frac{\mathbf{E}^*}{dt} - \mathbf{eE} \cdot \frac{\mathbf{H}^*}{dt}$$

The second term on the right-hand side vanishes; so we have

$$\mathbf{E} \cdot \mathbf{X} = \frac{d\mathbf{H}^*}{dt} + \frac{d\mathbf{E}^*}{dt} - \frac{\mathbf{I}}{I} = j \frac{\mathbf{H} \cdot \mathbf{H}^*}{dt} + \frac{\mathbf{E} \cdot \mathbf{E}^*}{dt}$$

If we integrate throughout the volume of the termination and use the divergence theorem on the left-hand side, we obtain

$$(6) \quad \frac{d\mathbf{H}^*}{dt} + \frac{d\mathbf{E}^*}{dt} - \frac{\mathbf{I}}{I} = -j \int_V \mathbf{H} \cdot \mathbf{H}^* dt + \mathbf{E} \cdot \mathbf{E}^* dt - 4j(W_m + W_e) \quad (4.24a)$$

where $W_m + W_e$ is the total time-average energy stored in the lossless termination, as reference to (2.53) in Sec. 2.5 shows, and dS is chosen directed into the volume.

Since both $\mathbf{n} \cdot \mathbf{E}$ and $\mathbf{n} \cdot \mathbf{X} dE/dt$, where \mathbf{n} is a unit inward normal, vanish on the perfectly conducting waveguide walls, the surface integral reduces to an integral over the terminal plane t only. On the terminal plane we have

$$\int_V \mathbf{H} \cdot \mathbf{E} dt + \int_V \mathbf{E} \cdot \mathbf{X} dt - \int_I \mathbf{I} dt = V \frac{dU^*}{dt} + \frac{dV^*}{dt} \quad (4.24b)$$

where V and I are the equivalent terminal voltage and current. Now $V = jIX$ for a lossless reactive termination; so

$$\begin{aligned} \int_V \mathbf{H} \cdot \mathbf{E} dt + \int_V \mathbf{E} \cdot \mathbf{X} dt - \int_I \mathbf{I} dt &= jX \int_V \frac{dU^*}{dt} dt + jI \int_V \frac{dV^*}{dt} dt \\ \text{Thus } V \frac{dU^*}{dt} + I \frac{dV^*}{dt} &= jXI \frac{dU^*}{dt} - jXI \frac{dV^*}{dt} - jB^* r_{out} \end{aligned}$$

and hence we find that (4.24b) yields

$$\begin{aligned} \frac{dX}{dt} &= jX \frac{dU^*}{dt} - jI \frac{dV^*}{dt} \\ \text{to } \frac{dX}{dt} &= jF \end{aligned} \quad (4.25)$$

The right-hand side is proportional to the total energy stored in the termination and can never be negative. Consequently, the slope of the reactance function must always be positive. If F is replaced by jBV in

(4.246), it is readily found that

$$\frac{dB}{dco} \propto \frac{4(W_m + W_e)}{W^*} \quad (4.26)$$

and hence the susceptance is also an increasing function of frequency m . The above relations also show that the frequency sensitivity of the reactance susceptance is proportional to the total average energy stored. These relations are readily verified in the case of simple LC reactive networks and a problem calling for this verification is given at the end of this chapter.

*4.4 EVEN AND ODD PROPERTIES OF Z ,

Before terminating the discussion dealing with one-port impedance functions, one further general property should be pointed out. This property is that the real part of $Z_m = R + jX$ is an even function of w , whereas the imaginary part is an odd function. The physical necessity of this was pointed out in the previous chapter in the section dealing with group velocity. The property stems from the requirement that the response of a circuit to a real-time-dependent driving function must also be real. That is, if $W(t)$ is the applied voltage at the terminal plane, the frequency spectrum is given by the Fourier transform:

$$V(a>) = \int e^{-at} r(0) dt \quad (4.27)$$

The frequency spectrum of the current that flows is

The current as a function of time is

$$v = 2irJ \cdot *R((o)) + jX(u>)$$

and must be a real function. This will be the case if

$$\begin{aligned} V(-a>) &= \underline{V^*(-o)} \\ R(-w) + jX(-<>) &\sim [\ll(*\gg) + jX(a>)]^* \end{aligned}$$

for then (4.29) becomes

$$\begin{aligned} J_{-f0} &= \frac{V(io)e^{j\pi}}{+jX(w)} \quad \stackrel{\text{J}_0^i}{=} \frac{V(o>)e^{ja}}{+jX(u>)} \\ \sim 2^{\wedge} L_x R(OJ) & \end{aligned}$$

which we can show is a real function. In the first integral on the rig¹

replace w by $\sim u >$ to obtain

$$\frac{V(-oj)e^{-j\omega t}}{\sigma - jQ} + \frac{R(-a)}{R(-a) + jX(-at)} = \frac{V(\infty)e^{j\omega t}}{R(u) + jX(w)} \quad \text{dot} \quad (4.30)$$

The two terms in the integrand are complex conjugates of each other, and hence the sum is real; i.e.,

$$S(t) = -\operatorname{Re} \left[\frac{V(a)e^{j\omega t}}{aim} + M \ll J^{dl} \right] \quad (4.31)$$

The condition specified on $V/R + jX$ is satisfied by V alone and by $R + jX$ alone. Clearly, from (4.27), $V(\sim a) = V^*U$. If

$$R(-to) + jX(-co) = [R(<\omega) + jX(<\omega)]^*$$

then

$$S(-\infty) - E(\infty) X(-w) \Rightarrow -X(\omega)$$

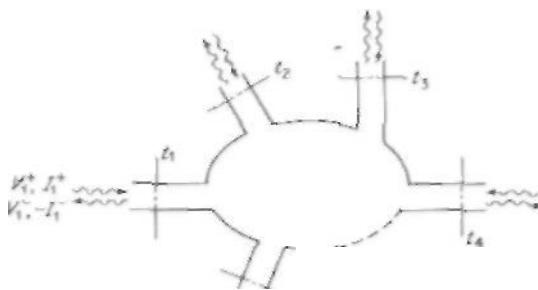
and R is an even function of w , whereas X is an odd function of to , as was to be proved. These even and odd properties are useful to know when approximate expressions for impedance functions are constructed from experimental data. For example, a series such as

$$a_{to} a_{-Aco'} + a_{5u} + \dots$$

could be used to represent X , but not to represent R , since the series is an odd function of ω .

AT-PORT CIRCUITS

Figure 4.4 illustrates the junction of JV waveguides or transmission lines (or a combination of the two) that terminate in a common region or junction. The region between the N chosen terminal planes may contain any arbitrary collection of passive elements. If each guide can support only one mode of propagation, this circuit constitutes an iV-port microwave circuit. If one or more of the guides can support several independent modes of propaga-



$\bullet v \cdot v O V / v$

FIGURE 4.4
An A-port microwave circuit.

tion, the number of electrical ports exceeds the number of mechanical ports. Each mode, since it carries power independently of all other modes, corresponds to an electrical port through which power may enter or leave the junction. To simplify the discussion, it will be assumed that each port supports only a single propagating mode. The extension of the theory to the case when some or all of the guides may support several propagating modes is more or less obvious.

Let the terminal planes be chosen sufficiently far from the junction so that the fields on the terminal planes are essentially just those of the incident and reflected dominant propagating modes. These fields may be uniquely defined at the terminal planes in terms of suitably defined equivalent voltages and currents. Clearly, the amplitudes of all the incident waves may be arbitrarily specified, i.e., chosen independently. The amplitudes of all the reflected waves are then determined by the physical properties of the junction; i.e., all the V_n^* are linear functions of the V_n . When the V_n^* and V_n are known, the corresponding currents I_n^* , I_n are known from the relations

$$I_n = \frac{V_n}{Z_{22}} \quad I_n^* = \frac{V_n^*}{Z_{22}}$$

Since Maxwell's equations are linear and the junction is assumed to be linear in its behavior, any N linearly independent combinations of the $4N$ quantities V_n^* , V_n , J_n , and I_n may be chosen as the independent variables to describe the electrical behavior of the junction. For an impedance description the total currents $J_n = J_n^* - J_n$ at the terminal planes are chosen as independent variables. The N total terminal-plane voltages $V_n = V_n^* + V_n$ are then the dependent variables, and are linearly related to the currents as follows:

V_n	Z_{22}	\mathbf{Z}_m	\mathbf{h}	(4.32)
V_N	Z_{NI}	Z_m	\mathbf{h}	

The matrix of elements Z_{ij} is the impedance matrix and provides a complete description of the electrical properties of the N -port circuit. Some of the properties of this impedance matrix are discussed below.

If the junction contains a nonreciprocal medium such as a plasma (ionized gas) or a ferrite with an applied dc magnetic biasing field (as are discussed in Chap. 6), then, in general, $Z_{ij} \neq Z_{ji}$; that is, the impedance matrix $[Z]$ is not symmetrical. The junction then requires $2N^2$ parameters to describe it completely since each Z_{ij} is complex and has two independent terms. If the junction does not contain any nonreciprocal media, $Z_{ij} = Z_{ji}$ and the impedance matrix is symmetrical. In this instance a total of $0.5(N^2 - N) = N(N + 1)$ independent parameters are required to describe the junction since $N^2 - N$ of the parameters are equal. Finally, if the junction is lossless—and many microwave junctions may be approximated as such with negligible error—then all the Z_{ij} must be pure imaginary.

since there can be no power loss within the junction. In this case there are only $N(N + 1)$ independent parameters required for a complete description. Any network containing the required number of resistive and reactive elements may be used as a representation for the junction at a given frequency. However, it must be kept in mind that when the frequency is changed, the values of the network elements (resistance, capacitance, and inductance) must also be changed. Rarely would any one particular network representation provide a complete description of the junction over a band of frequencies unless the network parameters are changed in value when the frequency is changed.

The foregoing discussion applies also to the admittance matrix $[Y]$, which relates the currents to the total voltages at the terminal planes; i.e.,

$$\begin{matrix} \mathbf{h} & M_1 & M_2 & M_N & \mathbf{V}, \\ & Y_{11} & & \cdot & V, \\ & Y_{12} & & \cdot & V, \\ & \vdots & & \ddots & \vdots \\ & Y_{N1} & & Y_{NN} & V_N \end{matrix} \quad (4.33)$$

The impedance and admittance matrices are reciprocals of each other; so

$$[Y] = [Z]^{-1} \quad (4.34)$$

of Symmetry for the Impedance Matrix

The symmetry of the impedance matrix is readily proved when the junction contains media characterized by scalar parameters u and e . Let incident-wave amplitudes V^* be so chosen that the total voltage V_n equals zero at all terminal planes except the i th plane. Let the corresponding field solution be E_j, H_j . Similarly, let a second solution E_j, H_j correspond to the case when incident-wave amplitudes are chosen so that all V_n equal zero except V_y . The Lorentz reciprocity theorem [Eq. (2.135)] gives

$$\oint (E_j \times H_j - E_j \times H_j) \cdot n \, dS = 0$$

when there are no sources within the closed surface S . Let S consist of the conducting walls bounding the junction and the N terminal planes. The integral over the walls vanishes if they are perfectly conducting or if they exhibit a surface impedance Z_m (Sec. 2.12). Therefore we obtain an integral over the terminal planes only, i.e.,

$$\sum_{n=1}^N \int (E_j \times H_j - E_j \times H_j) \cdot n \, dS = 0 \quad (4.35)$$

However, for the particular solutions considered here, $n \times E_j$ and $n \times E_j$, that is, E_{ti}, E_{tj} are zero on all terminal planes except t_f and t_i , respectively, since all V_n except V^* and V_y , have been chosen equal to zero. Thus (4.35)

becomes

$$\int_{J_i}^{\infty} \mathbf{E} \cdot \mathbf{H} \cdot \mathbf{n} dS = \int_{J_i}^{\infty} \mathbf{E} \cdot \mathbf{H} \cdot \mathbf{n} dS$$

where $(I,-)$, is the current at the terminal plane i arising from an annl' voltage at plane J_i , and similarly for $(/,,)$. From the admittance descrinf (4.33) of the junction, we have

$$\mathbf{f} \ll (/,,-\mathbf{W}) \text{ for } \mathbf{K} = \mathbf{e}, \gg \# \mathbf{i}$$

Hence (4.36) gives

$$\mathbf{V} \mathbf{V} \mathbf{Y} = \mathbf{V} \mathbf{V} \mathbf{Y}$$

or

$$\mathbf{Y} = \mathbf{Y} \quad (4.37)$$

which proves the symmetry of the admittance matrix. Since the reciprocal of a symmetrical matrix is a symmetrical matrix also, it follows that the impedance matrix is also symmetrical. The symmetry of the impedance and admittance matrices is a consequence of reciprocity. For nonreciprocal media, $/x$ or e (or both) are nonsymmetrical matrices, and (2.135) no longer applies. In this case the impedance matrix is no longer symmetrical. Nonreciprocal microwave devices are discussed in Chap. 6; so no further comments on these are made in this section.

*Proof of Imaginary Nature of [Z] for a Lossless Junction

For a lossless junction all the elements in the impedance and admittance matrices are pure imaginary. Let $[V]$ and $[I]$ be column matrices representing the terminal voltages and currents, respectively. The transposed matrices $[V]^T$, $[I]^T$, are row matrices of the form

$$INJ$$

The total complex power into the junction is

$$\sum_{n=1}^1 \sum_{m=1}^1 \mathbf{f}_n \mathbf{f}_m^T \mathbf{z}_{nm} \mathbf{i}_m = \mathbf{P}_j + 2 \mathbf{f}_n \mathbf{f}_m^T (\mathbf{w}_m - \mathbf{w}_e)$$

For a lossless junction $P_j = 0$ and the double sum must be pure imaginary. Since the f_n , can be chosen as independent variables, they may all be chose

as zero except for the n th one. In this case

$$\operatorname{Re}(\mathcal{Z}_n) = 0$$

or

$$\operatorname{Re}\mathcal{Z}_{nn} = 0$$

If all but \mathcal{Z}_{nn} and \mathcal{Z}_{mm} are chosen equal to zero, we obtain

$$\bullet \& z[\{i^*i_m + 1^*4\}2_{,,} + i_n i^*z_{nn} + i_m i^*z_{mm}] = 0$$

But $I_{,,}$, I^* , $\mathcal{Z}_{,,}/$, and $I^*I_m + IJ^*$ are all real quantities and Z_{nn} , Z_{mm} are imaginary; so this equation can hold only if

$$\operatorname{Re}\mathcal{Z}_{nn} = 0$$

Therefore all Z_{nm} are pure imaginary for a lossless junction.

Normalized Impedance and Admittance Matrices

Let us assume that we have chosen equivalent voltages and currents V_n , $I_{,,}$ so that the n th equivalent transmission line has a characteristic impedance Z_n given by

$$\begin{matrix} V^+ & V \\ n \sim & \sim F \sim T \end{matrix}$$

For an N -port circuit the impedance description is given by (4.32). The impedance-matrix elements are Z_{nm} . We now wish to redefine the equivalent voltages and currents so that each transmission line has unity characteristic impedance and to find the new impedance-matrix elements which will be designated by Z_{nm} and called normalized elements.

Let the new voltage and current amplitudes be V_f , V_{\sim} , $/^*$, and $I_{,,}$. In order to have the same power flow, we require

$$\bullet \& n^* I_n = 2^* V_n I_n$$

We can express the power flow in the following two equivalent ways also:

since for the normalized voltage and current amplitudes $V_n^*/I_n^* = 1$. From the last two relations for power flow, we see that the normalized voltage and current amplitudes are given by

$$i := y f z; i$$

From these relations it readily follows that

$$K = f c v_n /_{,,} = \{z \sim j_{,,}$$

and

$$V_{,,} = J Z_{\sim n} V_n \quad I_n = J T J_n$$

If we substitute from these expressions into (4.32), we obtain

$$\begin{array}{c|ccccc} \begin{array}{c} \text{V}_1 \\ \% \\ \% \end{array} & \left| \begin{array}{ccccc} \{Yy & 0 & 0 & 0 \\ & & & 0 \\ 0 & 0 &)/Y_N & V_f & "J \end{array} \right. \\ \hline \begin{array}{c} IT \\ \mathbf{0} \\ \bullet \end{array} & \begin{array}{ccccc} Q & ^{11} & 2,2 \\ 0 & Z_{\theta I} & "J \end{array} \\ \end{array}$$

W
"2.V

$$\begin{array}{c|ccccc} & & & \begin{array}{c} Y_1 \\ AW \end{array} \\ \begin{array}{c} fY_{\sim i} \\ \mathbf{0} \\ x \end{array} & \begin{array}{ccccc} 0 & 0 \\ fc & & & & 4 \end{array} \\ \hline \begin{array}{c} I \\ ffi \end{array} & \begin{array}{ccccc} & & & & J>J \end{array} \\ \begin{array}{c} YxZn \\ 21 \end{array} & \begin{array}{ccccc} <Y, Y_2Z_{X2} & & & & \end{array} \\ & \begin{array}{ccccc} & & & & \rangle/Y_t Y_N Z_{IN} \end{array} \\ & \begin{array}{ccccc} Y\%Z_{22} & & & & \end{array} \\ & \begin{array}{ccccc} & & & & \rangle/Y^N Z_{2N} \end{array} \\ \begin{array}{c} \rangle/Y^N Y_{\sim I} Z_I \\ .VI \end{array} & \begin{array}{ccccc} & & & & Yf, -Z_{NN} \end{array} \\ & \begin{array}{ccccc} & & & & "V \end{array} \end{array}$$

From this expression, we see that the elements of the normalized impedance matrix are given by

$$m \sim y/Z \sim Z_B \quad (4.38a)$$

A similar analysis shows that the elements of the normalized admittance matrix are given by

$$Y \quad (4.386)$$

4.6 TWO-PORT JUNCTIONS

At this point it seems advisable to examine the special case of the *two-P⁰**¹ junction rather than continue with the general theory of *V*-port circuit*. The derivation of some further properties of the *TV*-port junction is called for in some of the problems at the end of this chapter. Three examples of

two-port junctions are shown in Fig. 4.5. The first is the junction of two rectangular guides of unequal height (called an .E-plane step since the E vector of the TE_{10} mode lies in the plane containing the step geometry). The second is a symmetrical junction consisting of two similar rectangular guides joined by an intermediate guide of greater width. The third two-port junction consists of a typical coaxial-line-waveguide junction, where the center conductor of the coaxial line extends into the rectangular guide to provide an antenna radiating energy into or coupling energy out of the rectangular guide. A discussion of these particular junctions will serve to develop the general impedance description of two-port junctions or circuits.

Since evanescent modes are excited at each discontinuity, the terminal planes are chosen far enough away so that these decaying waves have negligible amplitudes at the terminal planes. Equivalent voltages and currents are introduced proportional to the total transverse electric and magnetic fields, respectively, at each terminal plane. For example, for the junction in Fig. 4.5a, let the incident and reflected transverse fields of the TE_{10} mode at each terminal plane be (coordinates x, y, z refer to the left-hand-side guide, and the primed coordinates x', y', z' refer to the right-hand-side guide, as in Fig. 4.6)

$$E_x = (CjV^{''''} + Cfe^{-\lambda' l} a_s \sin \theta) \quad \text{at } t_1$$

$$H_y = -Y_w(C/e^{jl^3}) - Cfe^{-\lambda' l} a_s \sin \theta \quad \text{at } t_1$$

$$E_y = (C, V^{3,2} + Cje^{-jl^1}) a_y \sin \theta \quad \text{at } t_2$$

$$H_z = YjCieP''* - C_2 e^{-\lambda l} a_x \sin \theta \quad \text{at } t_2$$

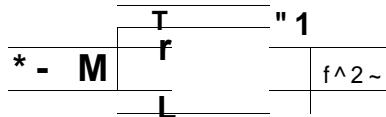
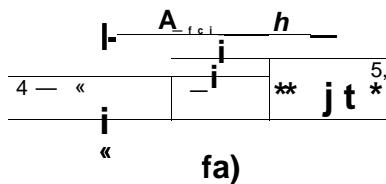
These expressions are obtained by choosing $h_s = (JTrY_w/lia)\cos(Trx/a)$ in order to simplify the expressions for the transverse fields. Let the equivalent voltages and currents be chosen as

$$V_s^+ = Kfif e^{Jsl} \quad V_f^- = Kfie \sim^m$$

$$V_2' = K_2 C_s e^{Jl^1*} \quad V_2 \sim = K_2 C_2 e \sim^{**}$$

$$I_2 = Y_w K_2 C_s \sim^l * \quad I_f = Y_w K_2 C_2 e \sim^{**}$$

In order to simplify the notation, we have expressed K_n as K_x and K_{2l} as K_2 and eliminated the second constants K_{12} and K_{22} by introducing the wave admittance using (4.6). Thus the characteristic impedance of the equivalent transmission line is equal to the wave impedance $Z_w = Y_w$ of the TE_{10} mode. To conserve power, it is necessary to choose K_x and K_2



(4)

ki**it***

[el]

FIGURE 4.5
Examples of two-port junctions.



FIGURE 4.6
Coordinates used to describe the junction in Fig. 4.5a.

SO that

$$Vtm)^* = Y_w K? \setminus Ctf = rjCjf \int_o^P \int_o^j \frac{d^2y}{dx} dx dy$$

or $K_1 = y[abj2]$ and

$$* ?(\#)'' = r^4 l^4 C^4 I^2 = y j C_2 \int_o^j \int_o^l \frac{d^2y}{dx} dx dy$$

or $K_2 = jab_2/2$.

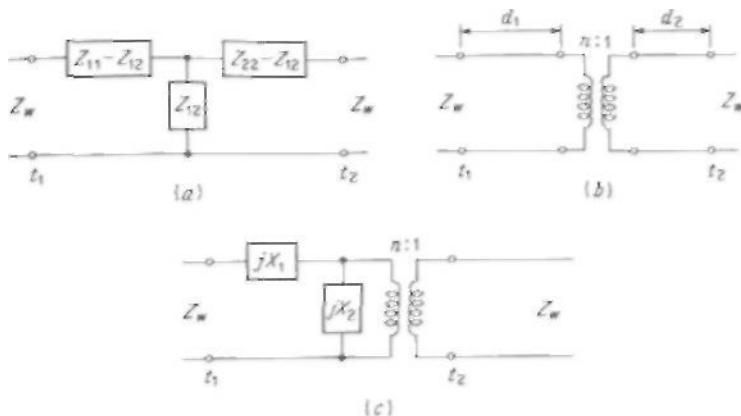


FIGURE 4.7
Equivalent circuits for a lossless two-port circuit.

If we use the above equivalent voltages and currents, we have

$$\begin{vmatrix} \mathbf{f} & \mathbf{v} \\ \mathbf{l} & \mathbf{h} \end{vmatrix} = \begin{vmatrix} 2j i & Z_{12} \\ Z_{1-i} & \Delta_{22} \end{vmatrix} \begin{vmatrix} \mathbf{V} \\ \mathbf{h} \end{vmatrix}$$

as a suitable description of the E-plane step. An equivalent circuit consisting of a T network joining two transmission lines as in Fig. 4.7a provides a convenient equivalent circuit for the junction. Other equivalent circuits are also possible. In particular, if the junction is lossless, any circuit consisting of three independent parameters may be used. Figure 4.76 illustrates a circuit consisting of two lengths of transmission lines of length d_1 and d_2 connected by an ideal transformer of turns ratio ml . Figure 4.7c is a variation of this circuit, where the transmission lines are replaced by reactive elements jX_1 and jX_2 . The parameters of any one of these circuits may be expressed in terms of the parameters of any of the others. The required derivations may be carried out in the usual manner.

The foregoing discussion applies equally well to the other two-port junctions of Fig. 4.5 provided suitably defined voltages and currents are introduced. Likewise, the equivalent circuits of Fig. 4.7 may be used to describe the behavior of these other junctions. Although general forms for the equivalent networks can be readily specified, the values of the parameters are not so easily found. In some cases the network parameters can be evaluated analytically, whereas in many other cases they must be determined by experimental measurements.!

+For typical analytical solutions and the methods employed, see R. E. Collin, "Field Theory of Guided Waves," 2nd ed., IEEE Press, Piscataway, N.J., 1991.

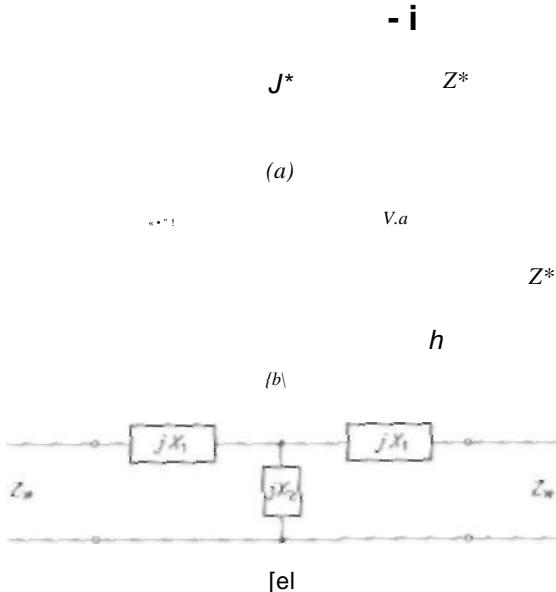


FIGURE 4.8
Equivalent circuits for a symmetrical two-port junction.

For the junction of Fig. 4.56, perfect symmetry exists about the midplane, and hence the equivalent impedance matrix has $Z_u = Z_{22}$. For a lossless junction of this type, the equivalent circuit can be any circuit containing two independent parameters in a symmetrical connection. Some typical circuits that may be used are illustrated in Fig. 4.8.

The foregoing discussion could be rephrased so as to apply to the admittance-matrix representation as well; i.e.,

$$\begin{matrix} h & & \\ & Y_{12} & Y_{22} & Y_{21} \\ & & & V-2 \end{matrix} \quad (4.40)$$

The basic equivalent circuit described by (4.40) is the II network illustrated in Fig. 4.9.

Example 4.1. To illustrate the use of equivalent circuits (assuming that the parameters are known at each frequency of interest), consider the coaxial line-waveguide junction of Fig. 4.5c. Let a generator of internal impedance γ_e be connected to the coaxial line a distance l_1 from the terminal plane t_1 . Let the output guide be connected to a load that is matched to the guide, i.e., γ_w^* presents an impedance I_w at the terminal plane t_2 . The overall circuit is illustrated in Fig. 4.10a. We wish to evaluate the power transmitted to

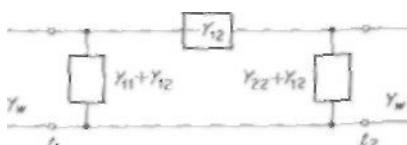


FIGURE 4.9
Equivalent circuit for admittance matrix of a two-port junction.

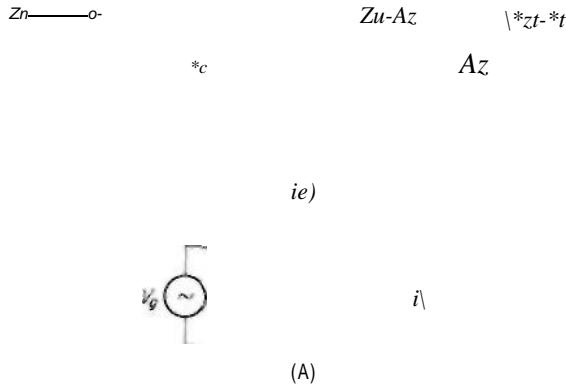


FIGURE 4.10
Equivalent circuit for generator connected to a coaxial-line-fed waveguide

load and the standing-wave ratio on the input line. The equivalent transformed load impedance at the plane t_x is found by conventional circuit analysis to be

$$\frac{Zf_2}{Z \cdot l - Z \parallel r_{22} + Z_u} \quad (4.41)$$

This impedance is transformed by the length l of coaxial line into an impedance

$$\frac{Z_L + jZ_c \tan \phi_l}{Z_s + jZ_L \tan \phi_l} \quad (4.42)$$

at the generator terminals as in Fig. 4.106. This reduced circuit is easily solved. The current supplied by the generator is

$$I_g = \frac{V_g}{Z_s + jZ_L} \quad (4.43)$$

and the power delivered to Z'_L is

$$P = \frac{1}{2} |I_g|^2 \operatorname{Re} Z_L \quad (4.44)$$

If the coaxial line and junction have negligible loss, this is also the power that is delivered to the load.

To compute the standing-wave ratio, note that the effective impedance terminating the coaxial line at t_x is Z_L . Thus a reflection coefficient R' given by

$$R' = \frac{Z_L - Z_c}{Z_L + Z_c} \quad (4.45a)$$

is produced. This reflection results in a standing-wave ratio S given by

$$S = \frac{1 + R'}{1 - R'} \quad (4.45b)$$

The reflection coefficient and standing-wave ratio depend only on the effective terminating impedance Z_L .

Maximum power will be delivered to the load if Z_g is made equal to the complex conjugate of Z_l , just as in the case of a low-frequency circuit. In low-frequency network theorems may also be applied. For example, Thevenin's theorem may be applied at the terminal plane t_2 to reduce the circuit to an equivalent generator with a new voltage V_g and a new internal impedance Z_{gt} . The new impedance is readily found by transforming Z_g to an equivalent impedance

$$Z_{gt} = Z_c^* + jZ_g \tan \pi l$$
(4.46)

at the plane t_2 . When viewed through the junction, this impedance appears as an impedance

$$Z_l \sim Z_c^* + jZ_g \tan \pi l$$
(4.47)

at the terminal plane t_2 .

The evaluation of the open-circuit voltage at t_2 is not quite as straightforward. However, by using Thevenin's theorem twice in succession, the desired result can be deduced with a minimum of labor. We first construct a Thevenin equivalent circuit at plane t_2 . The equivalent internal generator impedance to use here is Z_{ge} given above. Let the voltage waves produced by the generator when the coaxial line is open-circuited at I , be

$$V^+ = V'_+ e^{-j\beta z} + V'_- e^{+j\beta z}$$

where z is measured from the generator. At $z = I$ the coaxial line is open-circuited; so the total current

must vanish at $z = I$. Thus $V^+ = V'_+ e^{-2j\beta I}$. Hence, at the generator end where $z = 0$,

$$\begin{aligned} V(0) &= V'_+ (1 + e^{-2j\beta I}) \\ I(0) &= Y_c V'_+ (1 - e^{-2j\beta I}) \end{aligned}$$

But $I(0) = I_g$ and $V_g = I_g Z_g + V(0)$; so

$$\begin{aligned} V(0) &= V'_+ (1 + e^{-2j\beta I}) - V_g - I_g Z_g = V'_+ - I_g Z_g \\ &= V'_+ - Y_c V'_+ (l - e^{-2j\beta I}) \end{aligned}$$

When we solve for V' we obtain

$$V' = (I + Y_c Z_g) + (I - Y_c Z_g) e^{-2j\beta I} \quad (4.48)$$

The open-circuit voltage at t_1 is now readily found to be

$$\begin{aligned} V^+ &= V'_+ e^{-j\beta I} + V'_- e^{+j\beta I} - 2V'_+ e^{-2j\beta I} \\ &= (I + Y_c Z_g) + (I - Y_c Z_g) e^{-2j\beta I} \end{aligned} \quad (4.49)$$

The Thevenin equivalent circuit at t_1 is that illustrated in Fig. 4-H.

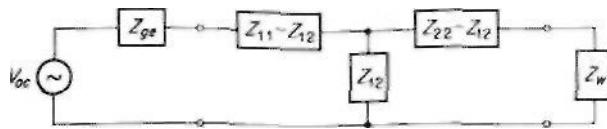


FIGURE 4.11
Thevenin equivalent circuit at terminal plane t_1 .

FIGURE 4.12
Thevenin equivalent circuit at terminal plane t_2 .

Application of Thevenin's theorem once more in the usual manner leads readily to the circuit of Fig. 4.12, where Z'_g is given by (4.47) and V^\wedge is given by

$$V^\wedge = \frac{Z_g V_{oc}}{Z_{ii}} \quad (4.50)$$

The reader may readily verify that the power delivered to Z_w as computed from the circuit of Fig. 4.12 is the same as that given by (4.44) (assuming no circuit losses), i.e..

$$P = \frac{1}{2} \frac{V^2}{Z'_e + Z_w} \quad (4.51)$$

and is equal to (4.44). When there are circuit losses, not all the power delivered to the equivalent load impedance Z'_L of (4.44) is absorbed in the load Z_w . However, (4.51) does give the correct power delivered to the load even if other circuit losses are present. An alternative way of solving the problem is to replace the coaxial line by an equivalent T network also (Prob. 4.9), in which case the circuit is reduced to a conventional lumped-parameter network.

e Equivalent Two-Port Circuits!

Figure 4.13 illustrates a number of useful equivalent circuits and some of their duals for representing lossless two-port junctions. The impedance parameters Z_n , Z_{22} , and Z_{Vi} are given below in terms of the network parameters, and vice versa. The same equations apply for the admittance parameters for the dual network (replace Z_n by V^\wedge , Z_{22} by Y_0 , etc.). Note that in the dual networks, the turns ratio of the ideal transformers are reversed. In addition, note that Z_0 in these circuits is an independent

^tThe material in this section has been reproduced in modified form from C. G. Montgomery, R. H. Dickc, and E. M. Purcell, "Principles of Microwave Circuits," pp. 105-108. McGraw-Hill Book Company, New York, 1948.

parameter or characteristic impedance and does not equal $(M_o/e)^{1/2}$

$$(a) \quad Z_n = -jZ_0 \cot \theta I \quad Z = Z_{22} - Z_u$$

$$Z_{22} = Z - jZ_0 \quad \cot \theta I \quad \cos \theta I = -\frac{i}{\sqrt{8}}$$

$$Z_{12} = \pm jZ_0 \csc \theta I \quad Z_0 = \sqrt{2} \cdot \frac{1}{Z_u^{1/2}}$$

$$(b) \quad Z_u = Z_l - j \cot \theta I \quad Z_x = Z_u + y f T z_{12}^*$$

$$Z_{22} = Z_2 - j C O t p l \quad Z_2 = Z_{22} + / 1 + Z f j,$$

$$Z_{12} = \pm 7 \csc \theta I \quad \sin P I = \pm \frac{1}{Z_{12}}$$

If $\theta I = \pi/2$, choose

$$Z_u = Z, \quad Z_l = Z,,$$

$$Z - Z_{22} = Z_2 \quad Z_2 = Z_{22}$$

$$Z_{12} = jZ_0 \quad Z_0 = -jZ_{12} * 1$$

$$(c) \quad Z_{12} = \frac{Z_{12}}{Z_0 \cot \theta I} \quad \cos \theta I = \frac{|Z_v| Z_{22}}{Z_{12}}$$

$$Z_{22} = -i \cot \theta I \quad Z,, = -jZ_u \frac{\sqrt{2}}{Z_u Z_{22}}$$

$$Z_{12} = j - \frac{1}{\cot \theta I} \quad n = \frac{u}{Z_{22}}$$

$$(rf) \quad Z_u = -j \cot \theta I \quad \cot \theta I = jZ_u$$

$$Z_{22} = Z - \frac{1}{\cot \theta I} \quad n = \frac{u}{Z_{22}}$$

$$Z_{x2} = - \frac{1}{\cot \theta I} \quad Z = Z_{22} \frac{Z_u Z_{12}}{1 - Z_f c}$$

$$(e) \quad Z_n = Z_{l+} - n \% \quad Z_i - Z_u - \frac{Z_{12}}{Z_{22}}$$

$$Z_{22} = Z_2 \quad Z_o = Z_{22}$$

$$Z_{12} = \pm n Z_2 \quad r_c = \frac{1}{Z_{22}}$$

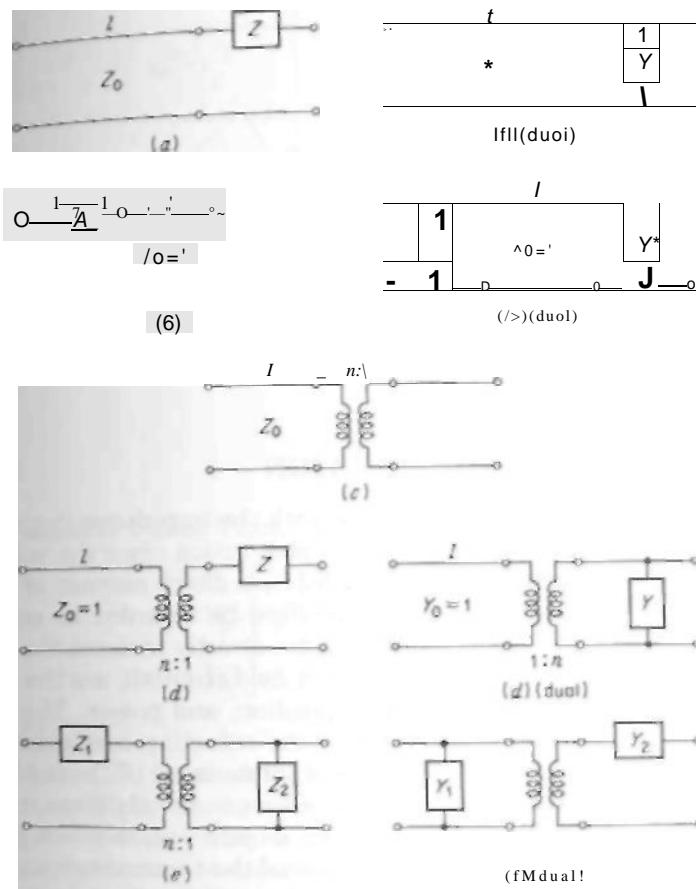


FIGURE 4.13
Some equivalent circuits for lossless two-port junctions.

An equivalent circuit incorporating a length of transmission line is particularly convenient to use since a shift in one (or both) of the terminal planes will reduce the circuit to a very simple form. For example, let the equivalent circuit of Fig. 4.13c be used to represent the junction in Fig. 4.14 between the terminal planes t_x and t_2 . If the terminal plane t_x is shifted a distance $'$ to the left so that $pU + I' = w$, that is, $I + I' = A_g/2$, then the equivalent circuit has a section of transmission line one-half guide wavelength long. But since impedance is invariant to a transformation through a half-wavelength-long section of transmission line, this being equivalent to a 1:1 turns-ratio transformer, the section may be removed and the equivalent circuit reduces to a single ideal transformer. This new circuit represents the junction between the new terminal planes t'_x and t_2 .



FIGURE 4.14

A junction and its equivalent circuit.

SCATTERING-MATRIX FORMULATION

The preceding section dealing with the impedance description of microwave circuits is in many respects an abstraction since the voltages, currents, and impedances cannot be measured in a direct manner at microwave frequencies. The quantities must therefore be regarded as secondary, or derived, quantities. The quantities that are directly measurable, by means of a small probe used to sample the relative field strength, are the standing-wave ratio, location of a field minimum position, and power. The first two quantities lead directly to a knowledge of the reflection coefficient. The measurement of power is needed only if the absolute value of the field in the device needs to be known. Another parameter that is directly measurable is the transmission coefficient through a circuit or junction, this being a relative measurement of the amplitude and phase of the transmitted wave as compared with those of the incident wave. In other words, the directly measurable quantities are the amplitudes and phase angles of the waves reflected, or scattered, from a junction relative to the incident-wave amplitudes and phase angles. Again, in view of the linearity of the field equations and most microwave devices, the scattered-wave amplitudes are linearly related to the incident-wave amplitudes. The matrix describing this linear relationship is called the scattering matrix.

Consider the N -port junction of Fig. 4.15. If a wave with an associate* equivalent voltage V_+^+ is incident on the junction at terminal plane t_1 , a reflected wave $S_n V_+^+ = V_{\sim}$ will be produced in line 1, where S_n is the reflection coefficient, or scattering coefficient, for line 1, with a wave incident on line 1. Waves will also be transmitted, or scattered, out of the other junctions and will have amplitudes proportional to V_+^+ . These amplitudes can be expressed as $V_{\sim} = S_{nl} V_f$, $n = 2, 3, \dots, N$, where S_{nl} is a transmission coefficient on line n from line 1. When waves are incident in all the other lines, the scattered wave in each line has contributions arising from all the other lines.

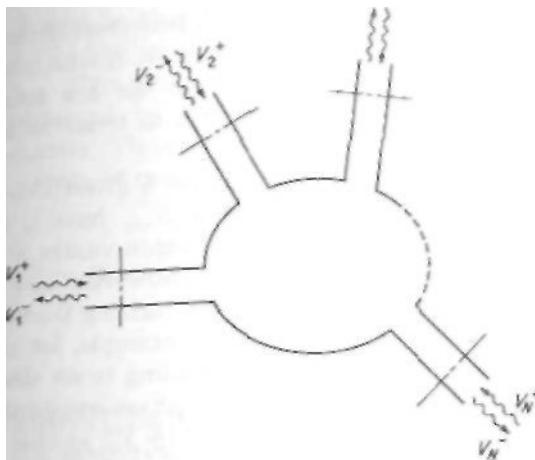


FIGURE 4.15
An N-port junction illustrating scattered waves.

incident waves. Thus, in general, we can write

$$\left| \begin{array}{c} \text{VR} \\ = \begin{array}{cccc} S_{11} & S_{12} & S_{13} & \\ S_{21} & S_{22} & S_{23} & \\ \hline S_{N1} & S_{N2} & S_{N3} & S_{NN} \end{array} \\ \text{V}; \end{array} \right| [v-] = [S][V^+] \quad (4.52a)$$

or

$$[v-] = [S][V^+] \quad (4.526)$$

where $[S]$ is called the scattering matrix.

When dealing with the scattering-matrix description of a junction, it is convenient to choose all the equivalent voltages (and currents, which, however, do not enter the picture explicitly) so that the power transmitted is given by $5|V_n|^2$ for all values of n . This corresponds to choosing the equivalent characteristic impedances equal to unity.^t The main reason for doing this is to obtain a symmetrical scattering matrix for reciprocal structures. If this normalization is not used, then because of different impedance levels in different lines, the scattering matrix cannot be symmetrical. Note that, with the assumed normalization, $V = V^* + V_\sim$ and $I = I^* - I_\sim$, and hence $V^+ = \sqrt{V + I}$ and $V_\sim = \sqrt{(V - I)}$. Thus the new variables V^+ and V_\sim are linear combinations of the variables V and I used in the impedance description. For this reason the currents do not enter into the scattering-matrix formulation. If desired, they may be calculated from the relation $I = V - V_\sim$.

^t A value different from unity would also be suitable, the only requirement being that all "ayes" have the same characteristic impedance, so that power will always be equal to a constant

When it is necessary to distinguish between normalized and unnormalized voltage and current amplitudes, then we use an overbar on the normalized variables. Throughout this section we are using only normalized variables so the overbar is not included, in order to keep the notation simple as possible.

At any particular frequency and for a given location of the terminal planes, the scattering-matrix elements S_{mn} have definite values. If the frequency is changed, these elements change values also, in a manner readily deduced analytically in general. However, at a fixed frequency a change in the scattering-matrix elements arising from a shift in the terminal-plane location is readily found. For example, let terminal plane t be shifted outward an amount Δz , corresponding to an electrical phase shift $\Delta \phi = P \Delta z / \lambda$ where P is the propagation phase constant for the n th line. If the incident-wave voltage is still denoted by V_i at this new terminal plane all the transmission coefficients S_{mn} for transmission into line m from line n must be multiplied by $e^{-j\Delta \phi}$ to account for the additional path length over which the waves must travel. The reflected wave in line n has traveled a distance $2\Delta z$ more relative to the incident wave at the new terminal plane. Thus the new value of S_{mn} is $e^{-j2\Delta \phi} S_{mn}$. Likewise, waves traveling from line m to line n must travel a distance Δz farther, and thus S_{nm} is changed to $e^{jB\Delta z} S_{nm}$. These results are readily expressed in the general case by the following transformation of the $[S]$ matrix into the new $[S']$ matrix:

$e^{-j\hat{J}_i}$	S_n	o_{12}	$\&2N$
$[S'] =$	$\therefore;$ «!	M'	'22
	$\therefore J \in$	-S, vi	S_{JV2}
			S^{NN-}
$, \neg J \gg i$			
	$, \neg J \gg 2$		(4.53)

where $\theta_{n+1} = P_n l_n$ is the outward electrical phase shift of the $n+1$ th plane.

Symmetry of Scattering Matrix

For a reciprocal junction the scattering matrix is symmetrical, $t^1 = S_{nm} = S_{mn}$, provided the equivalent voltages have been chosen so power is given by $j\sqrt{V_n}/l^2$ for all modes. The latter condition is eQ^{u1} . By choosing the characteristic impedance of all equivalent transmission lines used to represent the waveguides equal to unity. If the voltages a^{1*}

chosen in this fashion, $[S]$ will, in general, not be symmetrical. Problem 4-19 gives an example of a nonsymmetrical scattering matrix.

The proof of the symmetry property of the scattering matrix is readily obtained by utilizing the known symmetry property of the impedance matrix. Thus the symmetry of the scattering matrix is basically a consequence of reciprocity. For the normalization used in the present section,

$$\mathbf{V} = \mathbf{V}_+ + \mathbf{V}_- \quad \mathbf{i} = \mathbf{i}^+ - \mathbf{r} = \mathbf{v}^+ - \mathbf{v}$$

Thus, since $[\mathbf{V}] = [\mathbf{V}^+] + [\mathbf{V}] = [\mathbf{Z}][\mathbf{I}] = [\mathbf{Z}\mathbf{J}\mathbf{V}^+] - [\mathbf{Z}][\mathbf{V}^-]$ we have

$$\begin{aligned} ([\mathbf{Z}] &+ [\mathbf{U}])[\mathbf{V}^-] = ([\mathbf{Z}]-[\mathbf{U}])[\mathbf{V}^-] \\ [\mathbf{V}^-] &= i[\mathbf{Z}] + [\mathbf{U}]y^l[\mathbf{Z}] - [\mathbf{U})[\mathbf{V}^+] \end{aligned} \quad (4.54)$$

where $[\mathbf{U}]$ is the unit matrix. Comparing this result with (4.526) shows that the scattering matrix is related to the impedance matrix in the following manner:

$$[\mathbf{S}] = ([\mathbf{Z}] + [\mathbf{i}/]) -'([\mathbf{Z}] - [\mathbf{v}/]) \quad (4.55a)$$

Alternatively, we have

$$[\mathbf{v}^+] = k[\mathbf{v}] + in = k[\mathbf{z}] + WW$$

$$\text{and } [\mathbf{V}^-] = |([\mathbf{V}] - [\mathbf{I}]) = |([\mathbf{Z}] - [\mathbf{U}])[\mathbf{I}]$$

and this gives

$$[\mathbf{V}^-] = ([\mathbf{Z}]-[\mathbf{U}])[\mathbf{Z}] + [\mathbf{U}]^{-l}[\mathbf{V}^+]$$

$$\text{or } [\mathbf{S}] = ([\mathbf{Z}]-[\mathbf{U}])[\mathbf{Z}] + [\mathbf{U}]^{-l} \quad (4.55b)$$

The transpose of (4.55a) is

$$[\mathbf{s}]_r = ([\mathbf{z}] - [\mathbf{v}/]), ([\mathbf{z}] + [\mathbf{t}/]);'$$

But since the matrices in parentheses are symmetrical, they are equal to their transpose; e.g.,

$$az] - [u]), = [z] - [u]$$

$$\text{Hence } [\mathbf{S}]_r = ([\mathbf{Z}] - [\mathbf{V}])\{[\mathbf{Z}] + [\mathbf{U}]\}^{-l}$$

and using (4.55b) now gives

$$[\mathbf{S}]_r = [\mathbf{S}] \quad (4.56)$$

a result that can hold only if $[\mathbf{S}]$ is a symmetrical matrix.

Scatte

ringMat \wedge

for a Lossless Junction

or a lossless junction the total power leaving the N ports must equal the total incident power. The mathematical statement of this power-conservation

tion condition is

$$\sum_{n=1}^N |v_n|^2 = \sum_{n=1}^N |w_n|^2 \quad \text{•5?)}$$

This condition will impose a number of restrictions on the scattering parameters such as to reduce the total number of independent parameters to $N(N + 1)$, the same number of independent parameters as the impedance matrix for a lossless junction. Replacing V_n by V_n'

$$K - \sum_{i=1}^N s_{ni} v_i;$$

the power-conservation condition may be expressed as

$$\sum_{i=1}^N s_{ni} v_i = \sum_{n=1}^N |w_n|^2 \quad (4.58i)$$

The v_i are all independent incident voltages; so if we choose all $V_n = Q$ except V_s^+ , we obtain

$$\sum_{n=1}^N |s_n v_n|^2 = IV^2 \quad (4.59i)$$

$$\text{or } \sum_{n=1}^N |s_n v_n|^2 = E s_n v_n^* = I^2 \quad (4.60)$$

The index i is arbitrary; so (4.60) must hold for all values of i . Equation (4.60) states that for a lossless junction the product of any column of the scattering matrix with the conjugate of this same column equals unity.

In addition to the above constraint on the s_{nm} , a number of additional constraints may be derived. If we choose all $V_n^* = 0$ except V_s^+ and V_r^+ , (4.60) gives

$$\sum_{n=1}^N |s_{ns} v_n + s_{nr} v_r|^2 = \sum_{n=1}^N (s_{ns} v_n + s_{nr} v_r)(s_{ns} v_n + s_{nr} v_r)^* = |w_n|^2 + IV^2$$

Expanding the left-hand side gives

$$\begin{aligned} & \sum_{n=1}^N |B_{ns} v_n|^2 + \sum_{n=1}^N |s_{nr} v_r|^2 + \sum_{n=1}^N 2 \operatorname{Re}(s_{ns} v_n s_{nr} v_r^*) \\ & + \sum_{n=1}^N s_{nr} s_{ns}^* v_r^* (V_s^+)^* = |V_3^+|^2 + |V_r^H|^2 \end{aligned}$$

TTsine (4-59) results in a number of terms canceling, and we are left with

$$\sum_{r=1}^N [s_{ns}s_r v; (v; f + s_n^* s_{nr} v; (v_s^*)^*)] = 0$$

In view of the independent nature of $V/$ and V^* , choose, first of all, $V^+ = V^*$. We then obtain

$$\sum_{n=1}^N (S_{ns} S_n^* r + S_n^* s_{nr}) = 0 \quad (4.61a)$$

If, instead, we choose $V^* = jV^*$, with $V;$ real, we obtain

$$\sum_{n=1}^N (S_{ns} S_n^* r - S_n^* s_{nr}) = 0 \quad (4.61b)$$

Since neither V^* nor V^* is zero, both (4.61a) and (4.61b) can hold only if

$$\sum_{n=1}^N S_n X = 0 \quad **r \quad (4.62)$$

This equation states that the product of any column of the scattering matrix with the complex conjugate of any other different column is zero.

The conditions (4.60) and (4.62) are sufficient to restrict the number of independent parameters in the scattering matrix to $|MAT| + 1$. A matrix with elements that satisfy these conditions is called a unitary matrix. To illuminate this unitary property further, it will be instructive to rederive the above results by means of matrix algebra. The power-conservation condition (4.57) can be expressed as

$$\begin{aligned} [V^-][V^-]^* &= [V^+][V^+ r] \\ &= ([S][V^+]), ([S][V^+])^* \\ &= [V^-][S]_f [s_n V^+]^* \end{aligned}$$

Upon factoring this equation, we obtain

$$nH, U \cancel{f7} - [sL[sr][V^+ r] = 0$$

This equation can hold only if

$$[S]_f [Sf] = [U] \quad (4.63a)$$

$$IBf - IW \quad (4.63b)$$

mat⁶ & v is not zero. The result (4.63b) is the definition of a unitary matrix. The conditions (4.60) and (4.62) are obtained by carrying out the tnx multiplication called for in (4.63a).

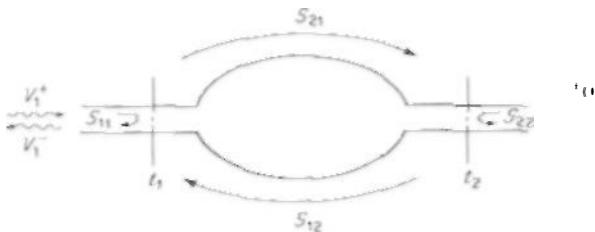


FIGURE 4.16
A two-port junction.

4.8 SCATTERING MATRIX FOR A TWO-PORT JUNCTION

Since many common microwave circuits are two-port junctions, the scattering-matrix description of these is examined in greater detail. With reference to Fig. 4.16, let the scattering-matrix parameters of the junction be S_{21} , S_{12} , and S_{22} . The incident- and scattered-wave amplitudes are related by

$$[V-] = [S][V+] \quad (4.64a)$$

or $Vf = S_U V? + S_{12} V?$ (4.64b)

$$V; = S_{21} V? + S_{22} Vi \quad (4.64c)$$

If the output guide is terminated in a matched load, $V_2 = 0$. From (4.64b) it is seen that S_{12} is the reflection coefficient in the input guide 1, with guide 2 terminated in a matched load. Also, S_{21} is the transmission coefficient into guide 2 from guide 1. Similar remarks, of course, apply to the parameters S_{22} and S_{12} .

If guide 2 is terminated in a normalized impedance Z_2 at the terminal plane t_2 , then $V_{2\sim}$ may be regarded as the incident wave on Z_2 , and $V_2 >$ s the wave reflected from Z_2 . The ratio must be equal to the reflection coefficient of the load; hence

$$\frac{Y_L}{V-}, \frac{z_2-i}{Z_2} = r, \quad (4.65)$$

Substituting into (4.64), we obtain

$$\begin{aligned} K - S_u V\sim &= S_{12} Vi = S_{12} r_L V_2 \\ &- S_{21} V? S_{22} r_L V_2 - V_2 \end{aligned}$$

Solving for Vf/V^* gives

$$V\backslash - c = S_{12} S_{21} r_L \quad (4.66)$$

which shows how the input reflection coefficient in guide 1 is modified if the output guide is not terminated in a matched load.

POT a reciprocal junction, $S_{12} = S_{21}$, and the scattering matrix contains at most, six independent parameters, which are the magnitudes and ^{nas}e angles of S_u , S_{12} , and S_{22} . If the junction is lossless, the scattering matrix contains only three parameters, since the $S_{,,}$, are related by conditions (4.60) and (4.62), which in the present case become

$$S_i \cdot S_f + S_{12}S_{f2} = 1 \quad (4.67a)$$

$$S_{22}S^*_2 + S_{12}S_{t2} = 1 \quad (4.67b)$$

$$S_nS^*_2 + S_{12}S_{\$2} = 0 \quad (4.67c)$$

The first two equations show that

$$|S_{,,}| = |S_{22}| \quad (4.68)$$

and hence the reflection coefficients in the input and output guides are equal in magnitude for a lossless junction. In addition, (4.67a) shows that

$$IS_{12}I = V_1 - IS_n|^2 \quad (4.69)$$

If we let $S_u = IS_u e^{-jB}$, $S_{22} = |S_u|e^{jB}$ and $S_{12} = (1 - |S_n|^2)^{1/2}e^{\alpha}$, then (4.67c) gives

$$IS_n l(1 - |S_{11}|^2)^{1/2} V^t \sim y^* + e^{\alpha} \rightarrow 0 = 0$$

or, equivalently,

$$e^{jH''} l^{+ \gg 2 \gg} = -eV^*$$

Thus

$$e_1 + e_2 = 20 - n \pm 2n -$$

and

$$\frac{e_1 + e_2}{2} + j\frac{n}{2} + nn \quad (4.70)$$

The two results (4.69) and (4.70) completely specify the transmission coefficient S_{12} in terms of the reflection coefficients S_u and S_{22} . Since $S_{,,}$ and S_{22} are readily measured and a knowledge of these suffices for the complete description of a lossless junction, the scattering matrix is a particularly convenient way of describing a lossless microwave two-port circuit.

The direct evaluation of the scattering-matrix parameters is illustrated by considering two simple examples. In Fig. 4.17a a shunt susceptance jB is connected across a transmission line with characteristic impedance Z_c . To ^{hnd} S_u , we assume the output line to be matched, so that $V_2 = 0$. The reflection coefficient on the input side is

$$S_i = \frac{y_c \sim \alpha_n}{Y_c + Y_{in}} = \frac{Y_c - Y_c - jB}{2Y_c + jB} = \frac{-jB}{2Y_c + jB}$$

rom symmetry considerations it is clear that $S_{22} = S_n$. The third parameter ^{Can b-} evaluated using (4.69) and (4.70) or by finding the transmittage α^2 with the output line matched. For a pure shunt element, we

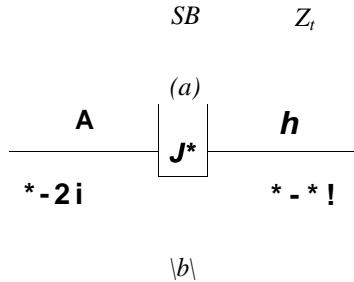


FIGURE 4.17

Shunt and series elements on a transmission line

must have $V_+ + Vf = V_{2\sim} = V + (l + S_u)$. Since $V^+ = S_{21}V_+$ also, we obtain

$$S_{21} = 1 + S_u - S_v - \frac{2K}{2Y_e + jB}$$

For the second example we consider a series reactance jX connecting lines with characteristic impedances Z_x and Z_2 , as in Fig. 4.176. In this example the characteristic impedances of the two lines are different; so we must first choose normalized voltages. Let V_+, Vf, V_J, V_2 be the actual transmission-line voltages for the waves that can exist on the input and output sides. Power flow for a single propagating wave is given by $j^2y_1^2V_+^2$ and $|y_2|V_2^2|^2$. If we choose normalized voltages $Vf = y_i^{1/2}V_1^{-1/2}$ and $V_J = Y_2^{1/2}V_2$, then power flow is given by $IVJI^2$ and $\pm|Vf|^2$ and is directly proportional to the voltage wave amplitude squared.

If the output line is matched, we have

$$\begin{array}{cccccc} Vf & Vf & & Z_m - Z_x & Z_2 - Z_x + jX \\ \hline V_+ & V^* & "u & Z_{to} + Z_t & Z_2 + Z_s + jX \end{array}$$

With the input line matched, we find

$$\frac{\%}{V_2^+} \quad \frac{?Z}{f^*} \quad , \quad \frac{Z_1 - Z_2 + jX}{Z_2 + Z_s + jX}$$

To find S_{21} , again consider the output line matched. On the input un¹ have $V_+ = V_? + Vf \Rightarrow V_?(1 + S_u)$ and

$$I > = Y_I(V_? - Vn) = Y_I V_?(l - S_{ll})$$

The current is continuous through a series element, and hence

$$-I_2 = 4 = \mathbf{A} = \mathbf{W}(\mathbf{i} - \mathbf{s}_u)$$

But $/_2 = W$; so $Y_2V_2 = Y_2a - S_u$. We now obtain

$$S_{21} - S_{12} - \frac{Y_2}{Z_1 + Z_2 + jX} \cdot \frac{V_2}{V_2} \cdot \frac{Z_1}{Z_2} \cdot \frac{1}{(1 - S_u)} \quad \text{5.}$$

$$h^{1/2} \frac{2Z}{Z_1 + Z_2 + jX} \quad \frac{2Z}{Z_x + Z_2 + jX} \quad 2^{\wedge} Z \cdot Z$$

The equality of S_{12} and S_{21} occurs because of the symmetrical manner in which Z_1 and Z_2 enter into this expression. If unnormalized voltages were used, the same expressions would be obtained for S_u and S_{22} , but for S_{21} and S_{12} we would obtain instead

$$S_{21} = \frac{2Z}{Z_1 + Z_2 + jX} \quad S_{12} = \frac{2Z}{Z_x + Z_2 + jX}$$

TRANSMISSION-MATRIX REPRESENTATION

When a number of microwave circuits are connected together in cascade, it is more convenient to represent each junction or circuit by a transmission matrix that gives the output quantities in terms of the input quantities. The reason for this is that, with such a representation, the matrix which describes the complete cascade connection may be obtained simply by multiplying the matrices describing each junction together. The independent variables may be chosen as the input voltages V_i , and currents I_i , the incident- and reflected-wave amplitudes V_i^+ , V_i^- on the input side, or any other convenient linearly independent quantities. When voltages and currents are chosen, we shall call the corresponding matrix the voltage-current transmission matrix. If incident- and reflected-wave amplitudes are chosen, we shall refer to the matrix as a wave-amplitude transmission matrix. To simplify the discussion, we shall consider the cascade connection of two-port circuits only. However, the general formulation is readily extended to cover the cascade connection of JV-port circuits.

The transmission-matrix formulation is of great value in analyzing nitely 1^on^g periodic structures such as those used in slow-wave circuits r traveling-wave tubes and linear accelerators. Since examples of these are alized in Chap. 8, we shall consider only the basic formulation in this section.

...^e-CW

ent

Transmission Matrix

sure 4.18a illustrates a two-port junction with input total voltage and as th e independent variables and the junction is linear, the dependent

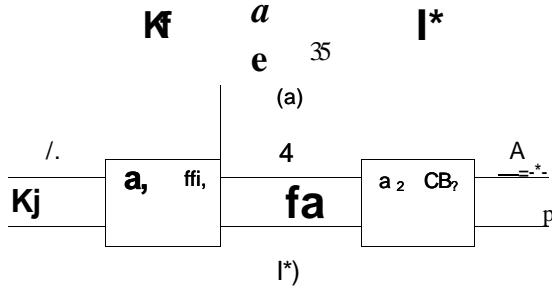


FIGURE 4.18
(a) A two-port junction; (fc)
cascade connection of two non-Memory Junctions.

variables V_j , I_i , are linearly related to V_2, I_2 . Consequently, we may write

$$V_x = s \cdot V_2 + M_2 \quad (4.71a)$$

$$J_j = W V_2 + \& I_2 \quad (4.71b)$$

where s , M , $\&$, and W are suitable constants that characterize the junction. Note that we have chosen the positive direction of current to be to the right at all terminals. This is done so that the output current J_2 becomes the input current to the next junction, etc., in a cascade connection as illustrated in Fig. 4.18b.

In matrix form (4.71) becomes

$$\begin{matrix} V_x \\ <\& \end{matrix} \quad (4.72)$$

The relationship of the voltage-current transmission matrix to the impedance matrix is readily found by rewriting the following equations in the form (4.71):

$$V_i = I_i Z_U - I_2 Z_{i2}$$

These equations may be solved to give

$$\begin{matrix} Z_{12}/Z_{i2} & [Z_U Z_{22} & Z_{i2}]/Z_{i2} & V_x \\ h & 1/Z_{i2} & Z^A/Z_{VI} & \end{matrix} \quad (4.1)$$

The Z_{nm} parameters of the junction are readily identified in terms of the Z_{nm} from this relation. The determinant of the voltage-current transmission matrix is

$$= 1$$

for a reciprocal junction, as is readily verified from (4.73).

For a cascade connection as illustrated in Fig. 4.186, we may write

$$\begin{vmatrix} v_1 \\ h \\ W \\ A \end{vmatrix} = \begin{vmatrix} k_i & \bullet A_j & fv_2 \\ *_i & *_{j+1} & V \\ si_s & 4U & V_3 \\ s_2 & ^{R_i} & V \end{vmatrix}$$

and therefore

$$\begin{aligned} \begin{vmatrix} hi \end{vmatrix} &= \begin{vmatrix} [j/2 & ^2 1 \\ r_2 & ^2] & v_1 \\ s f_x s/2 + \& \% & v_3 \end{vmatrix} \\ &= \begin{vmatrix} s/i \& t + M_x B_a & fv_3 \end{vmatrix} \end{aligned} \quad (4.75)$$

Thus the input quantities are readily found in terms of the output variables simply by multiplying the transmission matrices together. The ratio of the output voltage to current is determined by the load impedance.

Wave-Amplitude Transmission Matrix

The wave-amplitude transmission matrix relates the incident- and reflected-wave amplitudes on the input side of the junction to those on the output side. It bears the same relationship to the scattering matrix as the voltage-current transmission matrix does to the impedance matrix. Just as in the case of the voltage-current transmission-matrix representation, it is convenient to choose the variables in such a fashion that the output variables from one junction become the input variables for the next junction. With reference to Fig. 4.19a, we thus choose

(4.76a)

$$e T - FT \quad (4.76b)$$

$$c := v - \quad (4.76c)$$

$$c_2 = v; \quad (4.70a'')$$

$$cs = v_3 \quad (4.76e)$$

(4.76/-)

The superscript + refers to the amplitude of the wave propagating to the right, and the superscript - refers to the amplitude of the wave propagating to the left. The input and output quantities are linearly related; so we may write

$$\begin{matrix} "12 \\ "i1 & A_{22} & c, \end{matrix} \quad (4.77)$$

the A_{nm} are suitable constants that describe the junction. For the

$\hat{A}_{21} \quad \hat{A}_{22}$

[4!]

 $A_u \quad A_{,2}$ $lb)$

FIGURE 4.19

Wave-amplitude transmission-matrix representation of a junction.

cascade connection of Fig. 4.196, we have

$$\begin{vmatrix} A_n & A_{12} & A_u & A_{12} & c \\ A_{21} & A_{22} & A_{22} & A_{22} & \end{vmatrix} \quad (4.78)$$

In terms of the scattering matrix for the single junction, we have

$$\begin{vmatrix} Vr \\ v_* \\ e \\ t \\ \end{vmatrix} = \begin{vmatrix} cf \\ s_{12} \\ s_{22} \\ \end{vmatrix} = \begin{vmatrix} s_u & s_{12} \\ s_{12} & s_{22} \\ \end{vmatrix} \begin{vmatrix} V^+ \\ V^- \\ \end{vmatrix}$$

$$\begin{bmatrix} S_n & S_{12} \\ S_{12} & S_{22} \end{bmatrix} \begin{bmatrix} c^+ \\ c^- \\ \end{bmatrix} \quad (4.79)$$

These equations may be solved for c^+ and c^- to give

$$\begin{aligned} 1/S_{12} & \quad -S_{22}/S_{12} & c^+, \\ S_n/S_{12} & \quad (S_{12} - S_u S_v j/S_{12}) & c^-, \end{aligned} \quad (4.80)$$

from which the A_{nm} are readily identified in terms of the S_n . Note that the determinant of the $[A]$ matrix equals unity, i.e.,

$$A_n A_{22} - A_{21} A_{12} = 1 \quad (4.8D)$$

as is readily verified from (4.80). However, if the wave amplitudes have been normalized, so that power was given by $\hat{A}_{11}, \hat{A}_{22}, \dots$, etc., the determinant of the $[A]$ matrix would be different from unity in general (this follows from the nonsymmetry of the scattering matrix in this case). Some problems given at the end of this chapter illustrate these points.

*4.10 SIGNAL FLOW GRAPHS

A signal flow graph is a graphical representation of the relationships between a set of independent input variables that are linearly related to a set of dependent output variables. For example, the impedance-

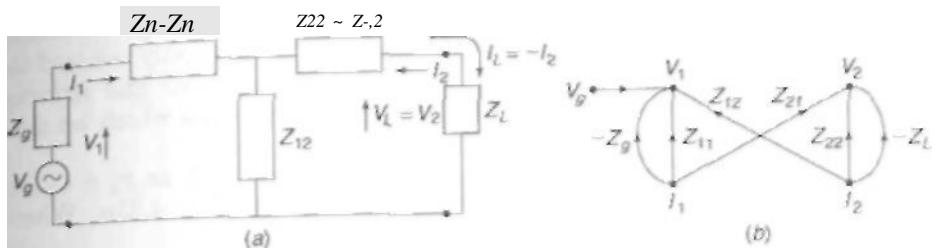


FIGURE 4.20

(a) A generator connected to a load through an impedance network; (b) signal-flow-graph representation of the linear system in (a).

description of a two-port network

$$V_1 = Z_u 7_1 + Z_{12} 7_2$$

$$V_2 = Z_{21} 7_1 + Z_{22} 7_2$$

may be represented by a signal flow graph by drawing four nodes to represent the variables V_1 , V_2 , 7_1 , and 7_2 and connecting lines having transmission factors that show how the inputs 7_1 and 7_2 feed signals to the output nodes labeled V_1 and V_2 . The graph is shown in Fig. 4.20. The node labeled V_1 has an input $Z_n 7_1$ from 7_1 and an input $Z_{12} 7_2$ from 7_2 . These inputs are represented by the directed line segments having transmission factors Z_n and Z_{12} as shown in Fig. 4.20. Similar connections go from node V_2 to nodes 7_1 and 7_2 . If the output is connected to a load Z_L , we have a further condition, namely, $V_2 \equiv -I_2 Z_L$, which can be represented as a directed line segment from the 7_2 node to the V_2 node and having a transmission factor $-Z_L$. If we connect a generator with voltage V_g and impedance Z_g to the input, then one additional equation $V_1 = V_g - I_1 Z_g$ is imposed on the system. To represent this equation on the graph, a node labeled V_g is added along with a directed line segment from this node to V_1 , and with unity transmission factor. Also an additional line segment from node 7_1 to node V_g , with transmission factor $-Z_g$ must be added. From this example it should be clear how a signal flow graph is constructed to represent a linear system.

Once the signal flow graph has been constructed, the solution giving any one variable such as V_1 , V_2 , 7_1 , or 7_2 in terms of the source variable V_g can be determined from the topology of the graph and the application of a set of formal rules known as Mason's rules.^f For complicated graphs the use of Mason's rules can be quite intricate and the chances of making a mistake

^f Mason, Feedback Theory—Some Properties of Signal Flow Graphs, *Proc IRE*, vol. 41, Or., '44-H56, 1953; also J. Mason, Feedback Theory—Further Properties of Signal-Flow Phs, *Proc. IRE*, vol. 44, pp. 920-926, 1956.

**
^

are quite high. For this reason we will not give Mason's rules but will instead show how the graph can be systematically reduced to a simple form which gives the desired solution by inspection. The reduction is carried applying five basic straightforward rules which we describe below

RULE 1. A pair of linear relations such as $x_2 = C_{21}x_1$, $x_3 = C_{32}x_2$ graphical representation shown in Fig. 4.21a. When there are no inputs to node 2, we have

$$x_3 = \wedge 2 V \sim [Z2] x_1$$

Thus rule 1 states that two series paths are equivalent to a single path from node 1 to node 3 with a transmission factor equal to the product of the transmission factors from node 1 to node 2 with that from node 2 to node 3 as shown in Fig. 4.21a.

RULE 2. If there are two or more parallel paths connecting node 2 to node 1, we have

$$\begin{aligned} x_2 &= A_{21}x_1 + B_{21}x_1 + C_{21}x_1 + \dots \\ &= (A_{21} + B_{21} + C_{21} + \dots) * i \end{aligned}$$

Rule 2 states the *obvious result* that several parallel connecting paths are equivalent to a single path with a transmission factor equal to the sum of those of the individual paths. This rule is illustrated in Fig. 4.21b.

RULE 3. Consider the linear relationships

$$x_2 = \wedge 2 1 x_1 + Q > 3 - r_3$$

$$x_3 = \wedge 3 2 * 2$$

which have the graph shown in Fig. 4.21c. With no other inputs to nodes 1 and 3, we can eliminate x_3 to obtain

$$\begin{aligned} x_2 &= \wedge 2 1 * 1 + \wedge' 2 3 \wedge 3 2 * 1 2 \\ * 2(1 - C_{23}C_{32}) &= C_{21}x_1 \\ x_o &= \begin{matrix} \wedge 2 1 * 1 \\ 1 - \wedge 2 3 \wedge 3 2 \\ - u o o x o \\ " C_{23}C_{32} \end{matrix} \end{aligned}$$

Thus rule 3 states that a feedback loop may be eliminated by dividing the input transmission factor by 1 minus the transmission factor around the loop which is the product $C_{23}C_{32}$ by rule 1. If there are several outputs from node 2, each input transmission factor is divided by 1 but all output transmission factors remain unchanged as shown in Fig. 4.21c. If node 3 is isolated, then the feedback loop becomes a self-loop gain $C_{23}C_{32}$

n

a

I

u²

O

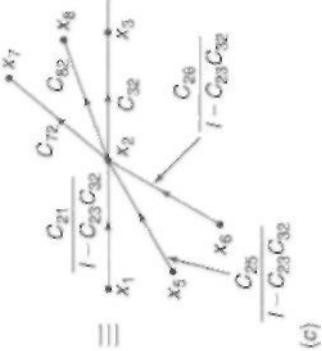
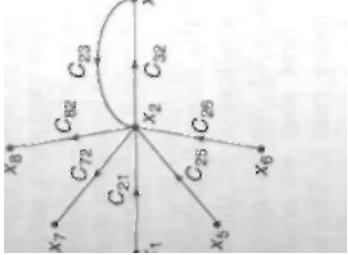


$\mathbf{0}^5$

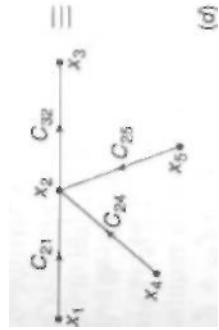
\mathbf{S}

\mathbf{o}

v



(c)



\mathbf{o} ci \mathbf{J}^1

« f OT3
-f C ~
« .2 < *
gg Bjl
E3 ^

!!

2-c

a:
o

*

RULE 4. Let node 2 have a single output and two or more inputs. Then we have the relationships

$$\begin{aligned} & - C_{21}x_1 + C_{23}x_3 \quad '24*4 \\ *,, * C_{n1}x_1 & = (C_{n2}C_{21})x_1 + (C_{n2}C_{23})x_3 + \dots \end{aligned}$$

This shows that x_2 can be split into a number of nodes x'_2, x_2, x''_2 etc with separate inputs $C_{21}x_1$ for node x'_2 , $C_{23}x_3$ for node x''_2 , etc., and all feed node n with a transmission factor C_{n2} . This rule is illustrated Fig. 4.21rf. If node 2 has a self-loop, it should be eliminated using *mlet* before the node is split.

RULE 5. Rule 5 is similar to rule 4 except there is only one input but several outputs from node 2. In this case each output can be considered as coming from a single node such as x'_2, x_2, x''_2 , etc., with each of these split nodes having the same input $C_{21}x_1$. This rule is shown in Fig. 4.21e. A self-loop at node 2 should be eliminated using rule 3 before the node is split.

The above rules are essentially those described by Kuhn.^t

We will now illustrate the application of the above rules to solve the circuit problem illustrated in Fig. 4.20. We will choose V_2 as the variable to be found in terms of V_g . As a first step we combine the parallel paths between nodes V_r, I_x and V_2, I_2 using rule 2. The new graph is redrawn in Fig. 4.22a. None of the rules we have given are of any help in reducing this graph any further and it seems as though we have come to an impasse. We can get around the dilemma by writing our load terminal condition in the form $I_2 = -V_2/Z_L$, which then provides an input to V_2 , the variable we are interested in. Hence we undo the use of rule 2 between nodes V_2 and I_2 and redraw the graph as in Fig. 4.22b. We now split node 2 using rule 4 to obtain the graph shown in Fig. 4.22c. This graph contains a self-loop which we eliminate by using rule 3, which requires dividing the input transmission factors by $1/(1 + Z_{22}/Z_L) = Z_L/(Z_{22} + Z_L)$. The result is the graph shown in Fig. 4.22e. Note that we also combined the transmission factors from node I_t to V_2 and from V_{jj} to I_2 using rule 1. We again run into a problem in reducing the graph in Fig. 4.22e because there is no input to node r. This is because of the way we chose to state the terminal conditions at a generator end. Instead of using $V_r = V_g - Z_g I_v$ we can use $I_j = "gl" K^A V_r / Z_g$ which shifts the generator input to an input to node I_x $**" \wedge_e$ provides an input to I_x from V_v . The new graph is shown in Fig. 4.5* I^A can reduce this graph by splitting node V_v into two nodes and applying¹ » 4 again. This leads to the graph shown in Fig. 4.22g which has a se

The self-loop is eliminated by applying rule 3 and results in the graph shown in Fig. 4.22A. This graph has a feedback loop so we apply rule 3 again to obtain

$$\begin{array}{ccccccc}
 & & & & \frac{Zq}{Z_r + Z_{22}} & & \\
 h & Z_g + z_u & 1 - & \frac{Z_{12}Z_{21}}{(Z_{11} + Z_g)(Z_{22} + Z_L)} & & & \\
 & & & & -2_{21}V_g & & \\
 & & (Z_n + Z_g)(Z_{22} + Z_J) & & Z_{l2}Z_{21} & &
 \end{array} \quad (4.82)$$

This example is simple enough so that we can easily solve for I_2 analytically and verify that (4.82) is the *correct* answer.

We initially chose to express the terminal conditions at the load and generator ends so as to cause difficulty in reducing the signal flow graph in order to illustrate the importance of choosing the correct *way to express* the terminal conditions. The terminal conditions should be expressed in a form that will result in all nodes in the signal flow graph having both input and output signals.

As a second example we will solve the same problem using a scattering-matrix representation of the impedance network; thus we use

$$\mathbf{V}_2 = \mathbf{S}_{21}\mathbf{V}\mathbf{R} + \mathbf{S}_{22}\mathbf{V}_2^+$$

At the load end we have $V_2 = F_L V_2^+$ where the load reflection coefficient is given by $F_L = (Z_L - Z_c)/(Z_L + Z_c)$ and Z_c is the characteristic impedance of the transmission line assumed to be connected between the network and the load impedance. This line is considered to have a negligible length. At the generator end we have the terminal conditions

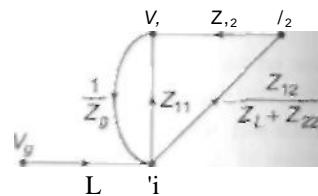
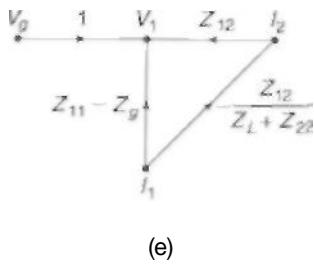
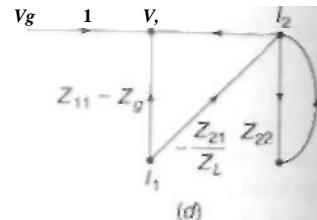
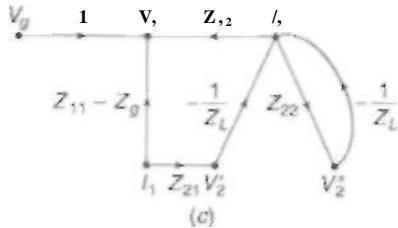
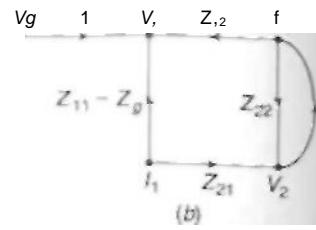
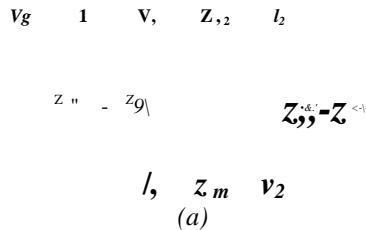
$$(V^+ - Vf)$$

which can be expressed as

where the *generator reflection coefficient* is given by

$$\begin{array}{c}
 \mathbf{r}_s \\
 \mathbf{s} \\
 \frac{z^* - z_s}{z_s + z_s}
 \end{array}$$

char⁸⁵*In assume that a transmission line of negligible length and with net a(Tristic imPendance Z_c is connected between the generator and the work. The signal flow graph for the above system of linear equations is



$$\begin{array}{c}
 \frac{Z_{12}}{Z_{11} - Z_g} k \\
 1 \\
 \frac{1}{Z_L + Z_{22}} \\
 \frac{1}{Z_g + Z_{11}} \\
 \frac{1}{Z_n} \quad I_1
 \end{array}$$

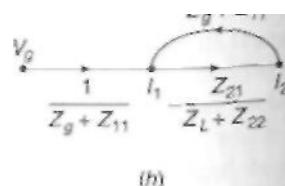
**FIGURE 4.22**

Illustration of steps used to reduce the signal flow graph in Fig. 4.206 to a simple form-

shown in Fig. 4.23a. Note that we have chosen to express the terminal conditions in such a form as to provide an input to the node V_2 ? v_f output from node V_f so as to avoid the problems encountered earlier in reducing the graph to a simple form. The procedure that we will follow parallels that used in the first example. We will choose V_2 as the variable W to be solved since this is the amplitude of the wave that is incident on the

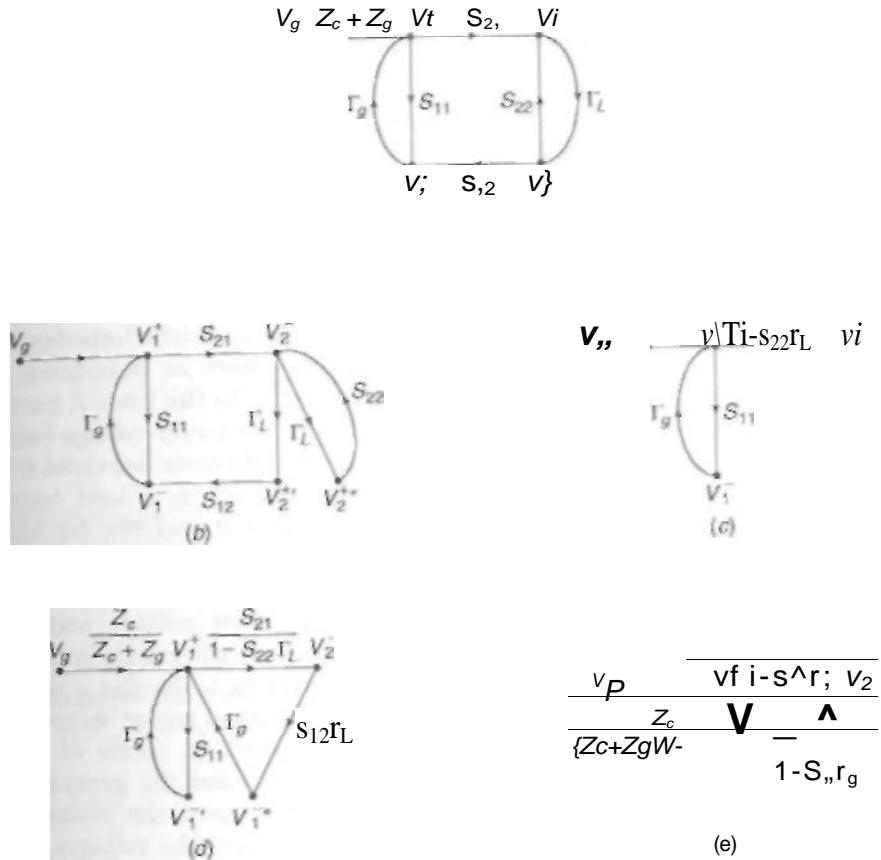


FIGURE 4.23

(o) Signal flow graph for the circuit in Fig. 4.20a but using a scattering-matrix formulation; (6)-(e) steps followed in reducing the signal flow graph to a simple form.

The first step is to split node V_f using rule 5 to obtain the graph shown in Fig. 4.23b. Rule 3 is now applied to eliminate the self-loop and produces the graph in Fig. 4.23c. A similar treatment of node V_f but using rule 4 followed by application of rules 2 and 3 results in the graphs shown in Figs. 4.23d and e. The final graph gives the desired solution upon applying rule 3 once more. Thus we find that

$$\frac{V_K Z_c S_{2l}}{(Z_c + Z_g)[(i - s_n V_g)(i - s_{22} ||) - s_{12} s_{2l} V_L r_t]} \quad (4.83)$$

The load voltage is given by $V_L = V_2(l + T_L)$.

For simple two-port networks, the use of signal flow graphs does not offer a great advantage over the algebraic method of obtaining a solution.

However, for three- and four-port networks, signal flow graph techniques provide a useful tool with significant savings in the effort required to find a desired solution.

*4.11 GENERALIZED SCATTERING MATRIX FOR POWER WAVES

If we have a load impedance equal to the characteristic impedance of transmission line to which it is connected, then all of the power in incident wave is delivered to the load. The reflected power will be zero. This would seem to represent an optimum situation. However, if the generator impedance does not equal the characteristic impedance of the interconnecting transmission line, we do not have an impedance match for maximum power transfer from the generator to the load. A partial standing wave on the transmission line can result in a larger voltage being applied to the load impedance and a greater amount of power delivered to the load. In general the generator should be terminated in a load impedance equal to the complex conjugate of the generator impedance for maximum power transfer. This conjugate impedance-matching criterion generally means that there will be a partial standing wave on the transmission line.

It is possible to introduce new voltage and current variables and corresponding wave amplitudes that will result in the conjugate impedance-matching condition being equivalent to having a zero reflection coefficient for the load. The new waves are called power waves. The scattering matrix that describes a microwave network in terms of incident, reflected, and transmitted power waves will be called the generalized scattering matrix. We will use script letters to represent the elements of the generalized scattering matrix and also to represent the voltages and currents associated with the power waves. The theory of the generalized scattering matrix and power waves is developed below. This theory provides a useful extension of the conventional theory for analyzing systems in which both the source and load impedances are complex. The theory can be developed by analogy to that of the conventional theory. We present the key relations from the conventional theory first and then use these to obtain similar relations for the generalized case.

For the circuit shown in Fig. 4.24a, let the voltage and current amplitudes on the transmission line be V^* , V^- and I^* , I^- at the load. Current amplitudes are related to the voltage amplitudes as follows:

V^*

V^-

$$j=j^*-r$$

$$\frac{z_g - z_t}{W} = \frac{v^* + v^-}{z_L} - v_g Q^* t$$

(a)

IOI

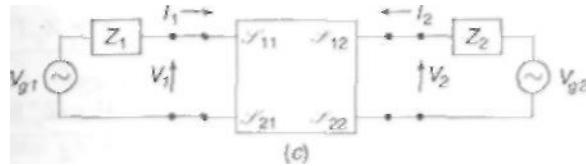


FIGURE 4.24

(a) A generator connected to a load impedance by means of a transmission line; (b) a circuit with complex source and load impedances; (c) a two-port network.

When $Z_t = Z_c$ there is no reflected wave since $T_L = 0$. Thus $V \approx 0$ and the voltage on the line at the generator end will be $V^* e^{j\theta}$ and equals $V_e/2$ when $Z = Z_c$. The incident power is given by

$$P_{\text{inc}} = P_{\text{ava}} = \frac{IV^* I^2}{2Z_c} = \frac{IV^*}{8Z_c} \quad (4.84)$$

and is equal to the available power P_{ava} from the source. If $Z \neq Z_c$ then the power delivered to the load is given by

$$P_{\text{load}} = (i - ir)^2 P_{\text{inc}} = (i - ir, i^2) P_{\text{ava}} \quad [4.85]$$

The total voltage and current at the load are

$$V = V^+ + V^- \quad i = r - r =$$

We can express the wave amplitudes V^+ and V^- in terms of V and i by the relations

$$\frac{V + ZJ}{2} \quad (4.86a)$$

$$V - ZJ \quad (4.86b)$$

The normalized voltage wave amplitudes are given by

$$y^+ = \frac{V + Z_L}{2y/Z_L} \quad (4.87_Q)$$

$$\bar{y}^- = \frac{V - Z_L I}{2y/Z_L} \quad (4.87_6)$$

Consider now a generator with internal impedance Z_g connected complex load impedance Z_L as shown in Fig. 4.246. By analogy with (do we will choose the power wave amplitudes T'' and 2^*Y to be

$$T'' = \frac{V + Z_L I}{Z_L} \quad (4.88_Q)$$

$$2^*Y = \frac{V - Z_L I}{2} \quad (4.88_6)$$

where V and I are the actual voltage and current in the circuit. When the load impedance Z_L is equal to the complex conjugate Z^* of the generator impedance, then

$$V = \frac{KZL}{z_g + Z_L} = \frac{V_g Z^*}{z_g + Z^*} \quad I = \frac{v_g}{z_g + Z_L} = \frac{v_g}{z_g + Z^*}$$

Upon using these expressions in (4.88), we obtain

$$\begin{aligned} M &= \frac{V_g}{2} Z^* - \frac{V_g Z^* - Z^*}{2 Z_g + Z^*} \\ &= \frac{V_g Z^* - Z^*}{2 Z_g + Z^*} \end{aligned}$$

Thus the definitions chosen for T'' and 2^*Y have the desired property that when the load impedance is conjugate matched to the generator impedance, for maximum power transfer, the amplitude of the reflected power wave is zero. The power delivered to the load is given by

$$1 \left| \frac{\%}{z_g + Z_L} \right| \quad IV / \quad |V|$$

where $Z_g = Z_L = R_g + jX_g$. This also represents the available power *hot* the source and can be expressed in the following way:

$$F_{ava} = \frac{|r + P| |Y^*|^2}{2R_g} \approx 2 \sim$$

where the normalized voltage wave amplitude has been chosen jo b\&
 $y + Z_L/R_g$

When the load impedance is not conjugate matched to the generator impedance, we will have a reflected power wave with amplitude r

ralized $j_{oa}(j)$ reflection coefficient, which we designate as V'_L , is given by

$$\begin{array}{cccc} V \sim & V - IZ^* & V/I - Z^* & Z_L - Z^* \\ L' & < y+ & V + IZ_g & V/I + Z_g & Z_L + Z_g \\ & & & & y' \end{array}$$

When Z is real, we obtain the usual load reflection coefficient. For the nonmatched case the power delivered to the load is given by

$$P \wedge R e W^* = -R e z_+;$$

$$SR_g|Z_L + Z_g|^2 = MP_{ava} \quad (4.90)$$

where the impedance mismatch factor $M = AR$, $R_g/Z + Z_g|^2$. The impedance mismatch factor M is always less than or equal to one. For $Z = Z^*$ we have $M = 1$. The mismatch factor determines the fraction of the available power that is delivered to the load. By using the definition for TJ and some algebra, we can readily show that

$$M = 1 - |V_L|^2 \quad (4.91)$$

Consequently, we can write

$$\wedge = (1 - HII^2) P_{ava} \quad (4.92a)$$

which is analogous to (4.85). Upon using $V = Z_L V_g / (Z_g + Z_L)$ and $/ = V_g / (Z + Z_L)$ in (4.88a), we find that $\wedge^* = V/2$; so we can also write

$$\begin{array}{ccc} 1 & & < y+ \\ & & (I - |r_L|^2) = (I - ir[(^2)p_{inc}] \end{array} \quad (4.926)$$

where P_{inc} is the power incident from the power wave launched by the generator.

The new set of variables \wedge^* and $T\sim$, which are linearly related to V and $/$, form a convenient pair of new variables for analyzing a circuit having complex generator and load impedances. If we have a two-port network as shown in Fig. 4.24c, we can introduce incident and scattered power waves on both the input and output sides. The power wave normalized amplitudes are linearly related to the normal voltages and currents on the input and output sides as follows:

$$\begin{array}{ccc} \bar{\wedge} = \frac{V_+ + I_x Z_x}{2} & \wedge = \frac{V_- - LZ^*}{2/i?} \end{array}$$

$$y+ = \frac{2}{2jR_2} \quad " \quad \frac{2}{2jR_2}$$

$-e Z_1$ and Z_2 , with real parts $i?$, and R_2 , are the complex generator

impedances on the input and output sides. The two-port network will choose the incident power wave amplitudes to be linearly related to the remaining two variables by the generalized scattering-matrix representation

$$\begin{aligned} \hat{A} &= \hat{A}_1 + \hat{A}_2 \\ (4.930) \quad (4.936) \end{aligned}$$

The generalized scattering-matrix parameters S_{11}^* cannot be measured directly. However, they can be expressed in terms of the normal two-port scattering-matrix parameters S_{11} . In practice, we would embed the two-port network into a transmission line with characteristic impedance Z_0 (usually 50 ohms) and measure the S^A . We can linearly relate V_t^+, V_L^-, V_2^+, V^A to V_1 and V_2 . The power wave amplitudes are linearly related to the total voltages and currents. By setting up these linear relationships and expressing V_1 and V_2 in terms of V^A and V_J , we can, after a number of matrix manipulations, show that

$$|S^A| = [D^*y \cdot [S] - [r^*]]([t] - [HfS])^{-1} \quad (4.94)$$

where $[D]$ is a diagonal matrix with elements

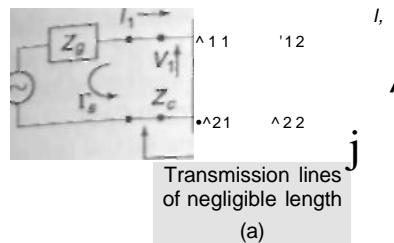
$$\begin{aligned} B_H &= [i - T_r l - T_t J_l - |T_f|] \quad * = 1,2 \\ Z_r &= Z_s \\ R_1 &= 2, + Z_s \end{aligned}$$

and $[T]$ is also a diagonal matrix with elements $f_M = I^n$. The unit matrix is $[U]$.

We will now examine the network shown in Fig. 4.25a. The two-port network can be described in terms of the normal scattering-matrix parameters S_{11} or in terms of the generalized scattering-matrix parameters. In the network shown in Fig. 4.25a, we do not have a generator on the output side. The termination is the load impedance Z_L . Also we have labeled the source impedance on the input side as Z_s and the source voltage as V_g . These changes correspond to setting $V_{g2} = 0$, $V_{gl} = V_g$, $Z_l = Z_s$, $Z_2 = Z_L$, $t = t$, $V_{g2} = 0$ the output current is $J_2 = -V_2/Z_L$. Consequently,

$$n = \frac{v_2 + hZ_L}{2JT7} = 0$$

so there is no power wave reflected from the load. The power derived by the load is



$$I_1 = -I_s$$



FIGURE 4.25
(a) a two-port network terminated in a complex load impedance; (b) Thevenin equivalent circuit.

The transducer gain of a two-port network, which could represent a microwave amplifier, is defined by the relation

$$G = \frac{\text{power delivered to load}}{\text{available power from source}} \quad (4.96)$$

Since the available power is simply $|T|^2/2$, we find that

$$G = |y_{21}|^2 \quad (4.97)$$

From (4.94) we obtain the following expressions for the generalized scattering-matrix parameters:

$$\bullet * u - \frac{i}{wi - r^*} \left[(\langle u - CXI - T_L S_{22} \rangle + S_{12} S_{21} r_L) \right] \quad (4.98a)$$

$$5^* = \frac{i}{li - rjii - rj} \sim l^r \cdot 2^{1/2} \quad (4.98b)$$

$$S_{21} = \frac{S_{21}}{\wedge_{12}} \quad (4.98c)$$

$$S_{22} = j \wedge [(S_{22} - r_L^*) (i - r_s s_u) + S_{12} S_{21} r_3] \quad (4.98d)$$

where

$$w = (i - r_s s_u) (i - r_L S_{22}) - S_{12} S_{21} T J_L$$

Note that for a reciprocal two-port network with $S_{12} = S_{21}$ that $S_{x2} = y_2$

also. We can now express the transducer gain in the form

$$G = i^2_{21} i^2 = \frac{(i - ir)(i - irj^2)is_{21}i^2}{(i - r_s M)(i - r_t s_{22}) - s_{12}s_{21}r_s r_t j^2} \quad (4.99)$$

In microwave amplifier design a more useful expression for the power gain G_p , which is defined by

$$G_p \gg \frac{\text{power delivered to load}}{\text{input power to amplifier}} \quad (4.100)$$

The input power is given by

$$P_{in} = (i - 1^n l)^2 \quad (4.101)$$

so we obtain

$$G_p = \frac{G}{1 - 1^n l^2} = \frac{W^2}{1 - l^n l^2} \quad (4.102)$$

Another gain expression is the available power gain, which is given by

$$G_a = \frac{\text{maximum available load power}}{\text{available input power}} \quad (4.103)$$

The maximum power that can be delivered to the load is the load power when the load impedance is made equal to the complex conjugate of the amplifier output impedance Z_{out} . We will show that the actual power delivered to the load is given by

$$P_L = (i - L^* s_{22})^2 P_{L,ava} \quad (4.104)$$

where $P_{L,ava}$ is the available load power. This relationship is of the same form as (4.101) which relates the input power to the power available at the source. When we use this expression in (4.103), we get

$$P_{in} = P_{L,ava}$$

$$\frac{1 - 1^n M l^2}{1 - |s_{22}|^2} = G = \frac{W^2}{1 - l^n f} \quad (4.106)$$

The above expressions for the various gains associated with an ^{linear**} ^{^triaxial*} two-port network are all expressed in terms of the parameters of the generalized scattering matrix and are the main reason for introducing the concept of power ^{w.a.}

We now return to the derivation of (4.104). We can replace the two-port network by a Thevenin equivalent network consisting of a voltage source V_{in} in series with the input port and a load Z_L connected to the output port. The input current i_{in} is given by

stage generator with open-circuit voltage V_T and a series impedance Z_T as shown in Fig. 4.256. The voltage V_T is the voltage across the amplifier terminals when Z_L is removed. The impedance Z_T is the impedance PPQ looking into the output terminals with the generator short-circuited. It is equal to the amplifier output impedance Z_{out} .

From the Thevenin equivalent circuit, it is readily seen that the power delivered to the load is given by

$$P_i = o I' / JX = \frac{W_T^2 R_L}{2(Z_T + Z_L)}$$

The maximum available load power equals P_L when we choose $Z_L = Z_T$ and is

$$P_{L \text{ ava}} = \frac{i V_R^2}{8 R_T}$$

Thus, in general, we can express P_L in the form

$$P_L = \frac{|V_T|^2}{8R_T} \frac{AR_L R_T}{|Z_T + Z_L|} = \frac{M}{8R_T}$$

where $M = AR_L R_T / |Z_T + Z_L|^2$ is the impedance mismatch factor.

If we can show that $M = 1 - |y_{22}|^2$, then we can express G_a as

$$G_a = \frac{P_{L \text{ ava}}}{MP_{L \text{ ava}}} = \frac{PL}{M} = \frac{G}{M} = \frac{G}{1 - |y_{22}|^2}$$

which is the expression we are trying to establish. When $V_g = 0$ we have $i_1 = 0$; so $y_{22} = 0$. The scattering-matrix relations now give

$$y_{11} = y_{12} y_{22}^{-1} = \frac{y_2 - hzt}{IJRT_L}$$

which gives us

$$-L_{22} = \frac{y_2 - hzt}{V_2 + i_z Z_L} = \frac{y_2 / i_z - hzt}{V_2 / J_2 + Z_L} = \frac{Z_T - Z_L}{Z_T + Z_L}$$

The Thevenin impedance Z_T is thus found to be

$$Z_T = S_{22}^h Z_L$$



We now expand the expression $1 - |S''_{22}|^2$ to obtain

$$\begin{aligned}
 & \frac{|Z_T|^2 - (Z_T - Z_t)(Z_L - Z_t)}{|Z_T + Z_t|^2} \\
 & \frac{(Z_T + Z_L)(Z_T^* + Z_L^*) - (Z_T - Z_L)(Z_T^* - Z_L^*)}{|Z_T + Z_L|^2} \\
 & \frac{4R_L R_T}{|Z_T + Z_L|^2} = M
 \end{aligned}$$

which completes the proof.

As a final observation we note that when $Z_t = Z_L$ then $M = 1$ and $G_Q = G$. Also if $Z_g = Z^*$, then $|Z_g|^2 = 0$ and $G_p = G_g = G$. Thus, with conjugate impedance matching at both the input and output, all three gains become equal.

*4.12 EXCITATION OF WAVEGUIDES

The preceding sections have dealt with the circuit aspects of passive microwave junctions. In order to complete the picture, it is necessary to consider also the equivalent-circuit representations for typical sources that are used to excite waves in a waveguide or transmission line. This particular aspect of waveguide theory is somewhat specialized, and it is not possible to give a complete analysis in this text without departing too far from the main theme. However, we shall present certain aspects of the excitation problem that provide a basis for choosing appropriate equivalent circuits and generators for representing typical sources and, in addition, make it possible to solve a number of coupling problems of engineering importance. The theory is, for the most part, developed by considering specific examples.

Probe Coupling in a Rectangular Waveguide

Figure 4.26 illustrates a typical coaxial-line-waveguide probe coupling—short-circuit position ℓ and probe length d can be adjusted to maximize power transfer from the coaxial line into the waveguide—center conductor of the coaxial line extends into the waveguide to form an electric probe. Any waveguide mode that has a nonzero electric field at the probe will excite currents on the probe. By reciprocity, when the probe current is produced by a TEM wave incident from the coaxial line, the waveguide modes will be excited.^t It is thus easy to see that, for maxi-

^tThis reciprocity principle is very useful for determining what modes a given probe can excite.*

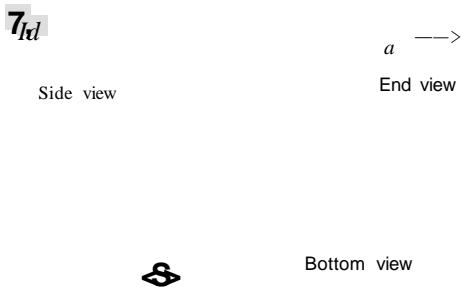


FIGURE 4.26
Coaxial-line probe coupling to a waveguide.

coupling to the dominant TE_{1n} mode in a rectangular guide, the probe should extend into the guide through the center of the broad face so as to coincide with the position of maximum electric field for the TE_{10} mode. The evanescent modes that are also excited are localized fields that store reactive energy. These give the junction its reactive properties. The section of short-circuited waveguide provides an adjustable reactance that may be used to tune out the probe reactance. The probe reactance can be evaluated by determining the amplitudes of the evanescent modes that are excited and computing the net reactive energy stored in these nonpropagating modes.¹ Since the details are rather lengthy, we shall evaluate only the amplitude of the radiated TE_{10} mode.

The current *on* the probe must be zero at the end of the probe. For a thin probe a sinusoidal standing-wave current distribution is a reasonable approximation to make for the probe current. Thus let the probe current be considered as an infinitely thin filamentary current of the form

$$I \sim /_0 \sin k_0(d - y) \quad 0 < y < d \quad x = - \quad z = 0 \quad (4.106)$$

We wish to determine the amplitude of the TE_{10} mode excited by this current. A general technique for accomplishing this is a mathematical formulation of the reciprocity principle invoked earlier to determine which waveguide modes a given source will excite. The required results are derived below.

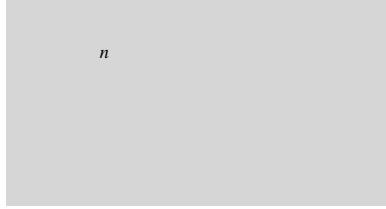
Figure 4.27 illustrates an infinitely long waveguide in which a current source J is located in the region between z_1 and z_2 . The field radiated by this source may be expressed as an infinite sum of waveguide modes as

¹For example, R. E. Collin, "Field Theory of Guided Waves," 2nd. ed., chap. 7, IEEE Press, Piscataway, N.J., 1991.

E'' , H'' E^* , H^*

FIGURE 4.27
A current source in a wavem Jin

follows:

	$z > z_n$	(4.107a)
	$z > z_2$	(4.107b)
	$z < z_2$	(4.107c)
$\mathbf{H} = \mathbf{E} \mathbf{c}_{,,} - (\mathbf{h}_n + \mathbf{h}_s) e^{Jt_n i}$	$z < z_2$	(4.107d)

In (4.107) n is a general summation index and implies a summation over all possible TE and TM modes. The unknown amplitudes C_n may be determined by an application of the Lorentz reciprocity formula (2.135). For the volume V , choose that bounded by the waveguide walls and cross-sectional planes located at z_2 and z_n in Fig. 4.27. Let the field $E_p H_j$, to be used in the Lorentz reciprocity formula, be the field radiated by the current source. This field is given by (4.107). For the field E_2, H_2 , choose the n th waveguide mode E_n, H_n ; that is,

$$\mathbf{E}_2 = \mathbf{E}_n = (\mathbf{e}_n + \mathbf{e}_{,,}) e^\lambda \quad \mathbf{I}$$

$$H_2 = H_n = (-h_{,,} + h_{,n}) e^{\lambda^2}$$

Equation (2.135) gives

$$\oint_S (E_{,,} \times H_{,,} - E_n \times H_n) \cdot dS = \int_V J \cdot E \cdot dV$$

since the field E_2, H_2 is a source-free solution ($J_2 = 0$) within V . The surface integral is zero over the waveguide walls by virtue of the boundary condition $\mathbf{n} \times \mathbf{E}_n = \mathbf{n} \times \mathbf{E}_{,,} = 0$. Since the modes are orthogonal, i.e.,

$$\int E_n^* \times H_{,,} \cdot dS = 0 \quad n \neq m$$

all the terms except the n th in the expansion of E_n, H_n vanish when

integrated over the waveguide cross section S_0 . Thus we have

$$\begin{aligned} & f C_n^+[(e_{,,} + e_{zn}) X (-h_{,,} + h_{,,}) - (e_{,,} - e_{,,}) X (h_{,,} + h_{,,})] \bullet a, dS \\ & -fC;[(e_{,,} - e_M)X(-h_n + 1iJ \\ & - (e_{,,} - e_J)X(-h_{,,} + h_j)] - a_z dS \\ & = -2Cj e_n X h_n \bullet a, dS = J E_n - J dV \end{aligned}$$

since the integral over the cross section at z_x vanishes identically. Hence C^* is given by

$$C^* > -\bullet_5 - / E - J dV^r = -\wedge - / (e_{,,} - e_{,,}) - J e^* \times r f V \quad (4.108_o)$$

If E^+ , H^\wedge is chosen for the field E_2 , $H_{,,}$, we obtain

$$C \approx -y f_y K^- * < W = -y j_v (e_n + e_{,,}) \bullet J e^{\wedge} dV \quad (4.1086)$$

$$\text{where } P_n = 2/\pi e_{,,} X h_{,,} \bullet a, dS \quad (4.108c)$$

and S_0 is a cross-sectional surface of the waveguide. The normalization constant $P_{,,}$ depends on the choice of expressions used for $e_{,,}$ and $h_{,,}$, the latter being arbitrary.

The above results are now applied to the probe problem introduced earlier. For the TE_{10} mode with fields given by

$$E_y = e_y e^{\wedge j \theta z} = \sin \frac{\theta}{a} e^{\wedge *} \quad (4.109_a)$$

$$H_x = h_x e^{-j \theta z} = -Y_w \sin \frac{\theta}{a} e^{\wedge j P}, \quad (4.1096)$$

we have

$$P_{10} = 2 \int_0^a \int_0^b Y_w \sin^2 \frac{\theta}{a} dx dy = ab Y_w \quad (4.110)$$

where Y_w is the wave admittance for the TE_{10} mode and θ is the propagation constant.

The probe in the short-circuited guide is equivalent to the original Probe plus its image at $z = -2l$ placed in an infinite guide, as in Fig. 4.28. we assume that the field radiated into the $z > 0$ region is

$$E := C^+ \sin \theta e^{-\theta z} \quad (4.111)$$

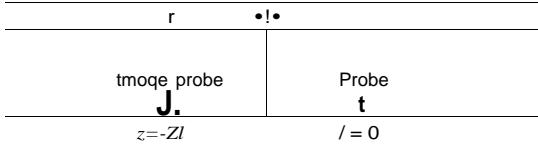


FIGURE 4.28
Probe and its image.

then application of (4.108a) gives

$$\mathbf{C}^+ = - \frac{abY_0}{\pi} \int_0^\infty I_0 \sin k_0(d-y) dy - \int_0^\infty I_0 \sin k_0(d+y) e^{-jw} dy \quad (4.112)$$

since

$$\begin{aligned} \sin \frac{rrx}{a} e^{k_0^2} &= 1 && \text{at } z = 0, z = \frac{a}{2} \\ \therefore 10 & & & \text{at } z = -\frac{a}{4}, z = -2I \end{aligned}$$

We have assumed that the current on the probe can be replaced by a line current along the probe axis and with a density given by (4.106). The volume integrals in (4.108) are consequently replaced by line integrals taken along the probe axis. Note also that the direction of the current in the image probe is reversed. This is necessary so that the fields radiated by the probe and its image will give a zero tangential electric field at the short-circuit position.

The total transverse field of the TE₁₀ mode radiated by the probe is thus, for $z > 0$,

$$E_y = \frac{UZ}{abk} e^{-jw} (1 - \cos fe_0 d) \sin \frac{\pi x}{a} e^{-k_0^2} \quad (4.113a)$$

$$H_x = -Y_w E_y \quad (4.H36)$$

The total radiated power is given by

$$\begin{aligned} P &= T \int_{-L/2}^{L/2} \int_0^a \int_0^{\pi/2} V < M, \\ &= \frac{I^2 \pi}{A a b k_0^2} e^{-2j f''} - l \ln |l - \cos k_0 d| \end{aligned} \quad (4.U^*)$$

At the base of the probe antenna ($y = 0$), the total coaxial-line current is, from (4.106),

$$I_s = I_0 \sin k_0 d$$

» ζ the input impedance seen from the coaxial line, referred to the base, be
 $\zeta = i\omega_0 + i\% -$ The complex Poynting vector theorem then gives [Eq. (4.14)]

$$Z_{in} = R_0 - jX = \frac{F}{UR} + \frac{2jmIW_m - W_M}{UR}$$

where P is the power radiated into the guide and W_e is the reactive energy stored in the vicinity of the probe owing to the excitation of nonpropagating (evanescent) modes. Since P has been evaluated and is given by (4.114), we can compute the input resistance. We obtain

$$R_0 = \frac{I^2 \pi^2 k_0 d}{2abk\%} \frac{Z_{in}}{1 - e^{-\gamma^2 \tan^2 \theta}} \quad (4.115)$$

upon using the identities $1 - \cos 2\theta = 2 \sin^2 \theta$ and $\sin 2\theta = 2 \sin \theta \cos \theta$. This input resistance is called the radiation resistance of the probe. Note that its value can be varied by varying the parameters λ and d , that is, the short-circuit position and probe length. Varying these parameters thus enables an optimum amount of power transfer to be achieved by adjusting f_0 to equal the characteristic impedance of the coaxial line and introducing a suitable reactance to tune out the reactance jX . Suitable techniques for reactance cancellation are discussed in the next chapter.

Radiation from Linear Current Elements

Figures 2.29a and 6 illustrate linear current elements in a waveguide. For the case of the transverse current element, (4.108) shows that

$$\mathbf{C} = \mathbf{C}_0 = -J_t \mathbf{J}_0 dl \quad (4.116)$$

where J_t is the line current density in amperes. This result may be interpreted to mean that a transverse-current element is equivalent to a shunt voltage source connected across an equivalent transmission line representing the waveguide when only a single mode, say the $n = 1$ mode, propagates. The reason for this is that the transverse current radiates a field with transverse electric field components that are equal on adjacent



FIGURE 4.29

Linear current elements in a waveguide.

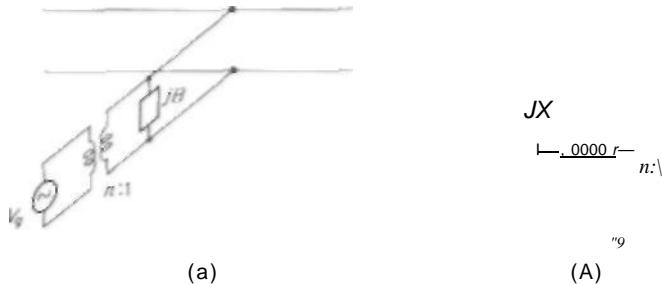


FIGURE 4.30

Equivalent circuits for current sources in a waveguide, (a) Transverse current source; (&) - current source.

sides of the current source, and this is equivalent to continuity of the equivalent voltage across the source region. The transverse magnetic field is discontinuous across the source region, and thus the equivalent current is also discontinuous across the equivalent voltage generator. Figure 4.30 illustrates the equivalent circuit for this type of source for the dominant propagating mode. The ideal transformer provides a means of adjusting the coupling between the voltage generator and the transmission line so that the same amount of power is radiated as in the waveguide. The shunt susceptance jB represents the net reactive energy stored in the field of the evanescent modes that are excited.

For an axial current located at $z = 0$, (4.108) gives

$$C := \sum_n jJ_n e^{jn\theta} dl$$

$$C \sim = \sum_n -jI_n e^{-jn\theta} dl$$

If the current is a symmetrical function of z between $-l < z < l$ then, is not a function of z , we have

$$C > -C_n = yf^* i \cdot e_{nn} \cos p_n z dz \quad (4.117)$$

For this case the radiated transverse magnetic field is continuous across source but the transverse electric field is discontinuous. The source is equivalent to a voltage generator connected in series with an equiangular transmission line, as illustrated in Fig. 4.306.

A linear current element may be viewed as an equivalent oscillating electric dipole. From Maxwell's equation we have

$$\nabla \times H = j(\omega E + J) = j(\epsilon_0 E + j\mu_0 H + j\sigma P + J)$$

and hence J enters into the field equations in the same manner as



FIGURE 4.31
A current loop in a waveguide.

polarization current jwP . Thus J may be considered equivalent to an electric dipole P given by

$$\mathbf{J}$$

$$J \ll \infty$$

Radiation from Current Loops

Figure 4.31 illustrates a linear current loop in a waveguide. The amplitude of the n th radiated mode is given by

$$C := -\hat{\mathbf{E}} \cdot \mathbf{r} \times \mathbf{I} d\ell$$

where $\mathbf{T}/$ is the vector current flowing around the contour C . \mathbf{r} is a unit vector along C . By Stokes' law we obtain

$$C := -\oint \mathbf{E} \cdot d\mathbf{l} = -\int \mathbf{E} \cdot \mathbf{r} d\ell$$

But $\nabla \times \mathbf{E} \sim -jwB \sim -jw n_0 H \sim$, and hence

$$C = -\int \mathbf{E} \cdot \mathbf{r} d\ell = -\int B d\ell = -f v x E \cdot dS \quad (4.118a)$$

Similarly,

$$C = -f B \cdot dS \quad (4.118b)$$

It is seen that the excitation amplitude of the n th mode is proportional to the total magnetic flux of this mode passing through the loop.

If the current loop is so small that the field B_n of the n th mode may be considered constant over the area of the loop, we obtain

$$C = -B \cdot dS = -\frac{joI}{2\pi} \cdot B \cdot S_0$$

Now $/S_0$ is the magnetic dipole moment M of the loop, where S_0 is the vector area of the loop; so we obtain

$$C = -B \cdot T \cdot M \quad (4.119a)$$

and similarly

/at

(4.1)

Radiation from a small current loop may be considered to be magnetic dipole radiation, as these equations show. For an axial magnetic dipole (transverse current loop), the equivalent source is a shunt-connected voltage source, whereas a transverse magnetic dipole is equivalent to a series connected voltage source.

*4.13 WAVEGUIDE COUPLING BY APERTURES^t

The foregoing formulation of the radiation from currents in a waveguide in terms of radiation from equivalent electric and magnetic dipoles is applicable to the coupling of waveguides by small apertures, or holes, in a common wall. To a first approximation a small aperture in a conducting wall is equivalent to an electric dipole normal to the aperture and having strength proportional to the normal component of the exciting electric field, plus a magnetic dipole in the plane of the aperture and having a strength proportional to the exciting tangential magnetic field. The constants of proportionality are parameters that depend on the aperture size and shape. These constants are called the electric and magnetic polarizabilities of the aperture and characterize the coupling or radiating properties of the aperture. A qualitative argument to demonstrate the physical reasonableness of these properties of an aperture is given below.

Figure 4.32a illustrates the normal electric field of strength E at a conducting surface without an aperture. When an aperture is cut in the screen, the electric field lines fringe through the aperture in the manner indicated in Fig. 4.32b. But this field distribution is essentially that produced by an equivalent electric dipole as shown in Fig. 4.32c. Note that the dipole is oriented normal to the aperture.

In a similar manner the tangential magnetic field lines shown in Fig. 4.32d will fringe through the aperture as in Fig. 4.32e. These fringing field lines are equivalent to those produced by a magnetic dipole located in the plane of the aperture.

In Bethe's original theory the dipole moments are determined by the field in the waveguide in the absence of the aperture. Thus, for a circular aperture of radius $r_0 \ll A_0$, the dipole moments are related to the

^tThe theory was originally developed by H. A. Bethe, Theory of Diffraction by Small Apertures, *Phys. Rev.*, vol. 66, pp. 163-182, 1944.

[#]For a derivation of these results, see Collin, *loc. cit.*

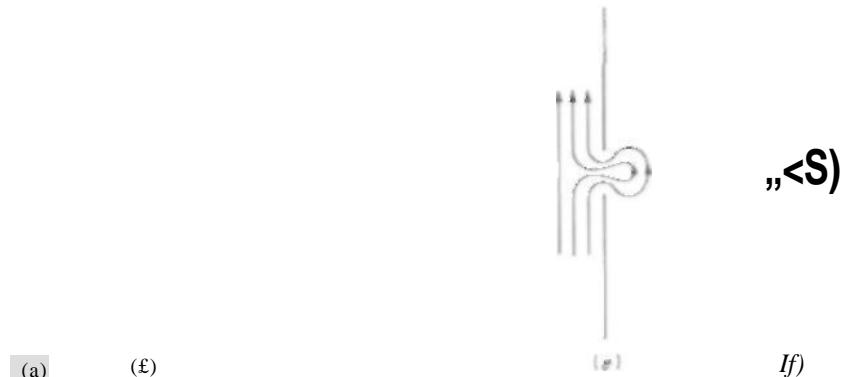


FIGURE 4.32
Aperture in a conducting wall.

fields as follows:

$$-\epsilon_0 a_f (\mathbf{n} \cdot \mathbf{E}) \mathbf{n} \quad (4.120a)$$

$$\mathbf{M} = -a_m \mathbf{H}, \quad (4.120b)$$

where $\mathbf{n} \cdot \mathbf{E}$ is the normal electric field and \mathbf{H} is the tangential magnetic field at the center of the aperture. The electric polarizability a_e is given by

$$a_e = -|\mathbf{r}_0|^3 \quad (4.121a)$$

and the magnetic polarizability a_m is given by

$$\mathbf{a}_m = \mathbf{a}_e \quad (4.121b)$$

The presence of an aperture also perturbs the field on the incident side of the screen. This perturbed field is that radiated by equivalent dipoles which are the negative of those given by (4.120) and located on the input side of the screen. It is important to note that when the aperture is replaced by equivalent electric and magnetic dipoles, the field radiated by these is computed by assuming that the aperture is now closed by a conducting wall. The equivalent dipoles correctly account for the field coupled through the aperture in the conducting screen.

Bethe's theory does not lead to an equivalent circuit for the aperture that includes a conductance to represent power coupled through the aperture. The reason for this is that the field assumed to excite the dipoles is chosen as the unperturbed incident field in the waveguide. In actual fact one should use the sum of the incident field and excited field as the polarizing field. Since the excited field is small, the correction to the dipole moments is so small. However, by including the excited dominant modes (the propagating modes) as part of the polarizing field, we will obtain the needed conductance element in the equivalent circuit. The dominant-mode fields react back on the dipoles to account for the

radiation of power by the dipoles. Thus, in place of (4.120), the following expressions are used for the dipole strengths: t_{tg}

For radiation into the output waveguide,

$$P = -6_0 \epsilon_e [n \cdot E^{\wedge} + n \cdot E_{Sr} - n \cdot E_{2r}]n$$

$$\mathbf{M} = -a_m [H_{gl} \quad H_{tr} \quad H_{2r}]n, \quad \text{H-1226}$$

For the radiation into the input waveguide,

$$P = e_0 a_e [n \cdot E_{g1} + n \cdot E_{1r} - n \cdot E_{2r}]n \quad (4.122c)$$

$$\mathbf{M} = a_m [H_{g1} + H_{1r} - H_{2r}]n \quad (4.122d)$$

where the generator fields E_{R1}, H_{g1} are the dominant-mode fields in the input waveguide in the absence of the aperture, E_{1r}, H_{1r} are the dominant-mode fields radiated by the dipoles in the input waveguide, and E_{2r}, H_{2r} are the dominant-mode fields radiated by the dipoles in the output waveguide. The unit vector n is normal to the aperture and directed from the input waveguide to the output waveguide. The subscript t denotes the tangential component of the magnetic fields. The field resulting from the aperture is determined by closing the aperture by an electric perfectly conducting surface and calculating the fields radiated by the dipoles given above and located at the center of the circular aperture.

The theory is readily extended to include noncircular apertures. However, the procedure outlined above is restricted to circular apertures in a very thin common wall between two waveguides. There is considerable attenuation in the coupling through an aperture in a thick wall and in many practical applications this attenuation must be taken into account.¹

The examples discussed next will illustrate the application of small-aperture coupling theory to an aperture in a transverse wall and an aperture in the broad wall between two identical rectangular waveguides.

Aperture in a Transverse Wall

Figure 4.33a illustrates a small circular aperture in a transverse wall & rectangular waveguide. To determine the exciting generator field, assume that the aperture is closed. A TE_{10} mode incident from $z < 0$ is reflected on the conducting wall at $z = 0$ to produce a standing-wave field in the region

¹R. E. Collin, *he. cit.*

²See, for example, N. A. McDonald, Electric and Magnetic Coupling Through Small Apertures in Shield Walls of Any Thickness, *IEEE Trans.*, vol. MTT-20, pp. 689-695, 1972.



(a)

tM

CD/M

X

ZH

le)

FIGURE 4.33
Aperture in a transverse
waveguide wall.

$z < 0$. This field is

$$E_v = C(e^{-\beta z} - e^{-\beta a}) \frac{TTX}{a} \quad (4.123a)$$

$$H_x = -CY_w(e^{j\beta z} + e^{-j\beta z}) \sin \frac{TTX}{a} \quad (4.1236)$$

plus a z component of magnetic field which is not required to be known for the present problem.

The normal electric field at the aperture is zero; so no induced electric dipole is produced. The tangential magnetic field at the center of the aperture is, from (4.1236),

$$H_x = -2CY_w$$

and hence an induced x -directed magnetic dipole M is produced.

In order to determine the total polarizing field using (4.1226) and (4.122d), we must find the fields H_{1r} and H_{2r} radiated into guides 1 and 2 by a magnetic dipole Ma_x . The field radiated into the region $z > 0$ is that radiated by the magnetic dipole M , as illustrated in Fig. 4.336. This dipole is equivalent to a half circular current loop in the yz plane as illustrated. To find the field radiated by this dipole in the presence of the conducting transverse wall, image theory may be used. Since the image of the half

circular current loop in the transverse wall is the other half of the r loop, the image of M is another magnetic dipole of moment M . The effective magnetic dipole moment M_{eff} in the transverse wall is equivalent to removing the wall and doubling the strength of the dipole, as depicted in Fig. 4.33c. If the field radiated into the region $z > 0$ is

$$\mathbf{B}_v^+ = Ae^{-\gamma z} \sin \frac{irx}{a} = Ae_v e^{-\gamma z}$$

$$/\hat{\mathbf{z}} = -AY^0 e^{-\gamma z} \sin \frac{TX}{a} = Ah_x e^{-\gamma z}$$

application of formula (4.119a) gives

$$\mathbf{A} = -\hat{\mathbf{z}} H_z (2M)$$

since the field B_z is $-ti^2 h^0 = Y_0^0 \sin(7rx/a)$ in the present case. The constant P_{10} is given by

$$\text{?io} = -2 / \int e_y h_x dx dy$$

$$= 2 \int_0^a \int_0^{cb} \sin \frac{TX}{a} dx dy = ab Y_w$$

Hence we obtain

$$\mathbf{A} = \frac{-2}{ao} \hat{\mathbf{z}} M = \frac{-2}{ab} \hat{\mathbf{z}} M \quad (4.124)$$

The presence of the aperture causes a field to be scattered into the region $z < 0$ also. For radiation into this region, the effective magnetic dipole moment is the negative of that used to obtain (4.124). Application of (4.119) now gives

$$E_v = A \sin \frac{Jfi}{a} \quad z < 0$$

for the radiated field in the input waveguide and where A is given by (4.124). As expected, the magnetic dipole M_a , acts as a shunt source. The $\hat{\mathbf{z}}$ component of the radiation reaction fields H_{1r} and H_{2r} , at the center of the aperture, are

and the generator field is

$$H_{gl} = -2CY_w$$

Since M represents the dipole strength for radiation into guide 2, we use these fields in (4.1226) to obtain

$$M = -a, -2CY_{,,} \frac{J4k_0Z_0}{abZ_{,,}} \cdot M$$

which can be solved for M to give

$$M = \frac{2a, YjC}{\frac{j4k_0Z_0}{abZ_{,,}}} \quad (4.125)$$

We can now complete the evaluation of the constant A by using this expression for M in (4.124); thus

$$Jk_0Z_0$$

$$A = \frac{4jk_0Z_0}{i + \frac{abZ_{ul}}{abZ_{ul}}} \quad (4.126)$$

The total electric field in the input waveguide is

$$E_y = [Or^{**} + (A - C)e^\lambda] \sin \frac{TTX}{A-C}$$

so the input reflection coefficient T is given by

$$A-C$$

The input normalized admittance is

When we substitute for A from (4.126) into the expression for Y , we find that

$$IL = \frac{2 - A/C}{A/C} = 1 - \frac{3ab}{8r_0^3/3} \quad (4.127)$$

The equivalent circuit of the aperture, as seen from the input waveguide, is a normalized shunt conductance of unit value plus a shunt inductive susceptance. The conductance term is called the radiation conductance and accounts for the power coupled, i.e., radiated, into the output waveguide. * The amplitude of the transmitted electric field is A which is given by (4.126). The transmission coefficient is A/C . From the equivalent circuit the transmission coefficient is $1 + T = 2/(1 + F_{in})$ which gives the same result. The aperture is equivalent to an inductive susceptance connected across the transmission line. The conductance term represents the output transmission line terminated in a matched load.

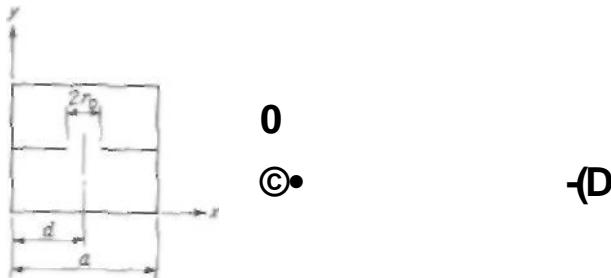


FIGURE 4.34
Aperture in a broad wall separating two waveguides.

Aperture in Broad Wall of a Waveguide

Figure 4.34 illustrates a circular aperture of radius r_0 placed in the broad wall separating two rectangular waveguides. The incident field is a TE_{01} mode in the lower guide, and is given by

$$E_{\bar{z}} = C \sin \frac{-\alpha z}{a} \quad (4.128a)$$

$$H_x = -C K \cdot \sin \frac{-\alpha z}{a} e^{-j P *} \quad (4.128b)$$

$$H_z = j \frac{TTY_{01}}{lia} C \cos \frac{TTX}{a} e^{-\alpha z *} \quad (4.128c)$$

At the center of the aperture located at $z = 0$, $x = d$, the exciting field is

$$E_v = C \sin \frac{vd}{a} \quad (4.129a)$$

$$H' = C Y J - a \cdot \sin \frac{vd}{a} + j \alpha \cos \frac{vd}{a} \frac{ird}{fc} \quad (4.129b)$$

These expressions show that there will be a y-directed electric dipole and z-directed magnetic dipoles excited in the aperture.

Let the fields E^+ , E_{f0} , B_{f0} , and B_{f0} be chosen as

$$E_{f0}^+ = a_v \sin \frac{vd}{a} e^{-\alpha z *}$$

$$E_{f0} = a_v \sin \frac{TTX}{a} jfc$$

$$M_0 H_{f0} = -\rightarrow o Y_w \frac{TTX}{a} sm - j \frac{TTX}{a} cos - \frac{TTX}{a} m$$

$$M_0 H_{f0} - B_{f0} = \rightarrow o \frac{TTX}{a} \sin - j \frac{TTX}{a} \cos - \frac{TTX}{a} m$$

Also let the dominant-mode field radiated by the electric dipole in the upper guide be

$$\begin{aligned} & \bullet A \& n & z > 0 \\ & " \quad A_2 E n & z < 0 \\ & l \quad A \quad ^\wedge & z > 0 \\ & \backslash A_2 H r_0 & z < 0 \end{aligned}$$

whereas that radiated by the magnetic dipoles is

$$\begin{aligned} & ; A_3 E ; _0 & z > 0 \\ & A_3 E \{ _0 & z < 0 \\ & M_3 H ; _0 & Z > 0 \\ & \backslash A_4 H r_0 & Z < 0 \end{aligned}$$

The electric dipole P is equivalent to an electric current given by jwP . Since the dipole is oriented in the transverse plane, (4.108) gives

$$A_r = A_2 = - \frac{1}{10} J O J P \cdot a_y \sin \frac{\text{Trd}}{a} = \frac{j o > P_y}{a o \Delta u s} \sin \frac{v d}{a} \quad (4.130)$$

Note that no integration is necessary since P is an infinitesimal dipole of total strength P_y . The constant P_{10} is equal to abY_w .

The field radiated by the magnetic dipoles may be found by using (4.119). Thus

$$\begin{aligned} & . \quad J a U f X w I \quad . \quad T ? d \quad , \quad 77 \quad v d \quad \backslash \\ & = \frac{j a \gg l i_0 I}{a b} \sin \frac{v d}{a} + j \frac{v M_z}{a} \cos \frac{\text{Trd}}{a} \quad (4.131) \end{aligned}$$

Similarly, it is found that

$$A_4 = \frac{j o u_a I}{a b} \sin \frac{rrd}{a} + j \frac{v M_z}{p a} \cos \frac{v d}{a} \quad (4.132)$$

With the above expressions for the amplitudes, the total field radiated by the upper waveguide is readily evaluated. It is given by

$$(A_1 + A_3)E_{1-0} \quad z > 0 \quad (4.133a)$$

$$(A_2 + A_4)E_{r0} \quad z < 0 \quad (4.133b)$$

$$(A_1 + A_3)H r_0 \quad z > 0 \quad (4.133c)$$

$$(A_2 + A_4)H f_0 \quad z < 0 \quad (4.133d)$$

that the electric dipole and the magnetic dipole Af. radiate the same in

both directions but the magnetic dipole M_x does not. By correctly choosing the aperture position d , it is possible to obtain zero radiation in the direction; that is, $A_2 + A_4$ can be made to vanish. We will return to this property later.

For radiation into the lower waveguide, the sign of the dipole reversed. Since the mode functions are the same, this has the effect of changing the sign of the amplitude constants. By using the expression for the mode functions given earlier and the derived expressions for the amplitudes, we find that the radiation reaction fields are given by

$$E \propto r v - \hat{A}^2 r y \sim 10^{-} (\hat{A}^2 K v a),$$

$$\frac{vd}{ab Y_w} \frac{2W/J_{0,0} VM}{\sin' \frac{vd}{la^2 b}} \cos \frac{vd}{a} \sin \frac{vd}{a} \quad (4.134a)$$

$$\Rightarrow i \propto -H_{2rx} = -2H_{2rx} = -(A_1 + A_3)H_{10} a_t - (A_2 + A_4)Hf_0 a_x$$

$$\frac{2J13M_x}{ab} \frac{-d}{\sin' \frac{vd}{a}} \quad (4.134b)$$

$$ff_M - H_{2rt} = -2H_{2rt} = -(A_1 + A_3)Hf_0 a_t - (A_2 + A_4)Hr_0 a_x,$$

$$\frac{2a > v P_y}{pa^2 b} \frac{vd}{a} \cos \frac{vd}{a} + \frac{2j o p_j Y_w v^2 M_x}{P^2 a^3 b} \cos \frac{vd}{a} \quad (4.134c)$$

Note that in evaluating the reaction fields at the center of the aperture, we take one-half of the field at $z = 0^-$ plus one-half of the field at $z = 0^+$, that is, on adjacent sides of the center of the aperture. The above equations show that there is interaction between the two dipoles P_y and M_x .

The above expressions for the reaction fields along with the generator fields given by (4.129) can now be used in (4.122a) and (4.122b) to write the following set of linear equations which will determine the dipole strengths

$$\begin{aligned} P_y &= -a, \epsilon_0 C \sin \frac{vd}{a} - \frac{2jcoP}{abY} \sim \frac{vd}{a} \sin^2 \frac{vd}{a} + \frac{2(o(j_0 v M_x)}{pa^2 b} \sin \frac{vd}{a} \cos \frac{vd}{a} \\ M_x &= -CY, \sin \frac{vd}{a} - \frac{2jpM_x}{ab} \sim \frac{vd}{a} \sin' \frac{vd}{a} \\ M_x &= -a \frac{jvY_w \hat{A}^2}{Pa} C \cos \frac{vd}{a} - \frac{2oJV_P V}{Pa^2 b} \sim \frac{vd}{a} \sin \frac{vd}{a} \cos \frac{vd}{a} + \frac{2jo > f_0 Y_w v^2 M_x}{P^2 a^3 b} \cos \frac{vd}{a} \end{aligned}$$

We can solve these equations for the dipole strengths which are found

given by

$$P_y = \frac{e_0 a_e \sin}{r \cdot \lambda^c} \quad (4.135a)$$

$$M_* = \frac{\langle m \rangle, \sin \frac{rrd}{pa_m}}{2 \text{ irrf } C} \quad <41356>$$

$$M_z = \frac{j a_c w Y_0 \cos \frac{\bullet rrd}{\ll_0 a A}}{\langle \ll_0 a A \rangle^o C} \quad (4.135c)$$

where

$$A = 1 + \frac{1}{2} j a_c \frac{\text{An sm}^z \frac{rrd}{\tilde{f}_{ab}} + 2 J_{\text{irrf}}^*}{\tilde{f}_{ab}} \frac{2 \cos^2 \frac{rrd}{pa^3 b}}{2 \text{ irrf}}$$

We can use these expressions to find the amplitudes A_1, \dots, A_6 for the field in the upper waveguide. Thus for the electric field we have

$$\begin{aligned} (A_1 + A_3) E_0 & \quad z > 0 \\ '(A_2 + A_4) E_{in} & \quad z < 0 \end{aligned}$$

where

$$\begin{aligned} A_1 + A_3 &= \frac{j a_c k Z_{sm} \sim \frac{rrd}{\sqrt{3} a^2 A}}{\sqrt{3} a^2 A} \frac{i \pi^2}{C} + \frac{j a_m \sqrt{a} \frac{rrd}{\lambda^2} \cos^2 \frac{rrd}{0 a A}}{0 a A C} \\ &+ \frac{j a_m p w a^2 \frac{rrd}{z r^2}}{a o J_1 + 2 j \ll_0 \sin^2 \frac{rrd}{a}} T T C \end{aligned} \quad (4.136)$$

The amplitude $A_2 + A_4$ is given by this expression with the sign of the last term changed from positive to negative. If we now set $A_2 + A_4$ equal to zero, we find that this can occur for

$$\sin \frac{vd}{a} = -7 \frac{A_n}{\sqrt{6} a} \quad (4.137)$$

thus an aperture position exists such that there is no radiation through Port 4. Power entering port 1 in Fig. 4.34 is coupled into ports 2 and 3 only. If the incident field were through port 2, we would find that there is no power coupled into port 3. A four-port network with these properties is a directional coupler about which more is said in Chap. 6.

In the lower guide the aperture dipoles will radiate dominant fields with amplitudes that are the negative of those for the modes "upper guide since the sign of the dipoles is reversed. Thus the reflected transmitted electric fields are

$$E = (C - A_1 - A_3) E_{10} \quad z > 0 \quad (4.13g)$$

$$E = -(A_2 + A_4) E_{f0} \quad z < 0 \quad (4.13h)$$

This shows that when the aperture is positioned so that there is no one coupled into port 4, then there is also no reflected power in port 1.

We will define the aperture susceptance jB and aperture reactance "f by the expressions

$$2a_e k l \quad i r d \quad -2a_m (7r/af) \quad , d$$

$$2a_n p \quad . \quad rrd$$

In terms of these parameters, the expressions for the amplitudes $A_1 + A_3$ and $A_2 + A_4$ can be written as

$$\begin{aligned} & A_1 + A_3 = -\mathbf{r} \cdot \mathbf{r} \cdot \mathbf{7i}^c + \mathbf{m} \cdot \mathbf{f}^c \quad \text{4.140b} \\ & A_2 + A_4 = \frac{jB/2}{1+jB} ; C - \frac{jX/2}{1+jX} \quad (4.140d) \end{aligned}$$

The condition $A_1 + A_3 = 0$ is met by setting $X = B$. Since a_e is negative, we can obtain a negative value of S . Thus the amplitude $A_1 + A_3$ can also be made small by setting $X = -B$. However, in this case $A_2 + A_4$ does not vanish exactly. In order for $A_2 + A_4$ to equal zero, we would require $X + B = 0$ and $XB = 0$ which has only the trivial solution. By using $X = -B$ we obtain an imperfect directional coupler.

PROBLEMS

- 4.1. For TM modes in a waveguide, show that the line integral of the transverse electric field between any two points on the boundary is zero.

Hint: Note that $\mathbf{V}_e \cdot d\mathbf{l} = (de/dl)dl = \text{directional derivative of } e$, along the path. Integrate this and use the boundary conditions for e_z . As an alternative, note that there is no axial magnetic flux, so that the line integral around a closed path in the transverse plane must vanish.

- 4.2. For TE modes show that the line integral of the transverse electric field between two points located on the guide boundary depends on the path integration chosen.

Hint: Note that, because there is an axial magnetic field, the line integral around a closed path does not vanish.

- 4.3. An obstacle located at $z = 0$ excites evanescent H modes that decay exponentially away from the obstacle in the positive z direction. Integrate the cor

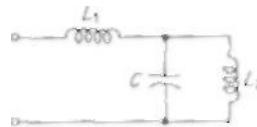
Poynting vector over cross-sectional planes at $z = 0$ and $z = w$ and the guide walls for the ramth evanescent H mode, and show that there is no power transmitted into the region $z > 0$. Show also that the reactive energy stored in the nonpropagating H mode in the region $z > 0$ is predominantly magnetic.

Hint: Note from (2.59) that the total inward flux of the complex Poynting vector equals the power loss (which is to be taken equal to zero in this problem.) $\text{plus } 2jw(W_m - W_e)$.

- 4.4. Repeat Prob. 4.3 for the case of an E mode and show that nonpropagating E modes store predominantly electric energy.
- 4.5. For the circuits illustrated in Fig. P4.5, verify that the slope of the reactance function is given by (4.25).

(a) 

(a)



(b)

FIGURE P4.5

- *4.6. For the N-port junction choose an excitation such that all $I_{ij} = 0$ except I_{11} , thus $V_j = Z_{jj}I_j$ for all j . Show that all Z_{ij} must have real parts that are even functions of ID and imaginary parts that are odd functions of W .
- 4.7. Generalize the result (4.25) to show that for a lossless iV-port junction

$$[i^*L^{8Z} \quad \mathbf{u}] = \sum_{n=1}^N \sum_{m=1}^{A?} \mathbf{n}^{HZ} \frac{-/m}{So} = 4j(W; + W_m)$$

- 4.8. Verify that (4.51) and (4.44) are equal.

- 4.9. Show that a length l of transmission line of characteristic impedance Z_c is equivalent to a T network with parameters

$$z_n = 2_{22} = \sim jZ_c \cot pi \quad Z_n = -jZ_c \operatorname{esc} pi$$

- 4.10. Let Z_n^A , Z_n^B , Z_n^C , Z_n^D be the input impedance of a T network when terminals 2 are short-circuited, when terminals 1 are short-circuited, when terminals 2 are open-circuited, and when terminals 1 are open-circuited, respectively. In terms of these impedances show that the parameters of the T network are given by

$$\mathbf{Zu} = \mathbf{Zl} \quad z_{22} = z_{oc} \quad zh = (zl - zl)zl = \{zl - zl\}zi$$

Use these relations to verify the equations for the circuits of Figs. 4.13d and e.

- 4.11. For the microwave circuit shown in Fig. P4.11, evaluate the power transmitted to the load Z_L . Find the standing-wave ratio in the two transmission-line

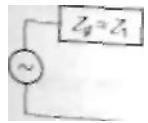
 f^* $Zt = Z<$ $/W = %$ $\mathbf{Z}_c = \mathbf{Z}_s$ Z_i $\underline{\text{fil}} \cdots \underline{\text{A}}$

FIGURE P4.11

sections. Assume $Z_L = 2Z_U$, $X_x = X_2 = Z_{\parallel}$, $V_p = 5V$ (peak). Check your answers using TLINE.

- 4.12.** FOT the microwave junction shown in Fig. P4.12, the equivalent-T parameters are $Z_{n1} = j2$, $Z_{L2} = j/2$, $Z_{22} = -j0.25$. Find the param^* for the alternative equivalent circuit illustrated. ** for

$$\& = 1$$

01 n: \

FIGURE P4.12

- 4.13.** For the three-port junction illustrated in Fig. P4.13, compute the power delivered to the loads $Z_1 = 50$ ft and $Z_2 = 100$ ft. Assume that $V_p = 10Q$ peak.

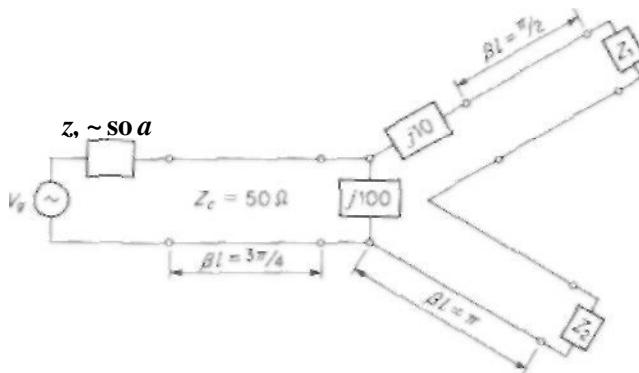


FIGURE P4.13

- 4.14.** For the transmission-line circuit shown in Fig. P4.14, find (a) the load reflection coefficient, (b) the impedance seen by the generator, (c) the VSWR on



FIGURE P4.14

transmission line, (d) the fraction of the input power delivered to the following parameters apply:

$$50 \text{ ft} \quad Z_g = 50 \text{ ft} \quad fil'J = \pi/4 \quad Z_L = 40 \text{ ft}$$

- 4.15.** For the circuit shown in Fig. P4.15, find the VSWR on both transmissij* and the load impedance at the generator terminals. The following P&

liP*

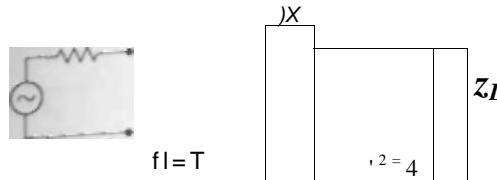


FIGURE P4.15

apply:

$$I = - \quad l_2 = - \quad Z_c = 50 \text{ n} \quad Z_g = 50 \text{ ft} \quad Z_L = 25 + j25 \quad /X = /25$$

- 4.16.** For the transmission-line circuit shown in Fig. P4.16, find the required value of Z that will match the 20-ft load resistance to the generator. The generator internal resistance $R_g = 60$ (i. Find the VSWR on the transmission line. Is R_L matched to the transmission line?



$\Sigma = \$$

FIGURE P4.16

- 4.17.** For the transmission-line circuit shown in Fig. P4.17, find the VSWR on each

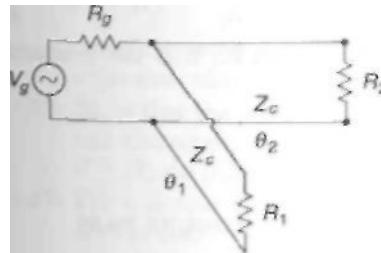


FIGURE P4.17

transmission line and the relative powers delivered to R_1 and R_2 . The following parameters apply:

$$\mathbf{0i} = IT \quad \mathbf{e}_2 = - \quad R_1 = 25 \text{ n} \quad R_2 = 75 \text{ n} \quad z_c = 50 \text{ n} \quad R_g = 50 \text{ n}$$

- ¹⁸ for the transmission-line circuit shown in Fig. P4.18, find the open-circuit voltage V_x and the Thevenin equivalent impedance. Use these results to find the power delivered to R_L . Assume that $pi = TT/4$, $R_g = 75$ ft, $Z_c = 50$ ft, $L = 30$ n.

$$\begin{array}{c} R_g \\ | - \text{WV} \end{array}$$

&

$$pl=E$$

FIGURE P4.18

- 4.19. Consider the junction of two transmission lines with characteristic impedances Z_1 and Z_2 as illustrated in Fig. P4.19. When the usual transmission-line currents I_1 and I_2

-?12

$$S_2,$$

$$5, , ^\wedge 5 2$$

FIGURE P4.19

voltages and currents are used, show that the scattering-matrix parameters are given by

$$\begin{matrix} 2 & 1 \\ Z_2 + Z_1, & 2Z_1, \\ S_{11} = & S_{21} = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \end{matrix}$$

The normalized voltages V_1 and V_2^+ are given by $V_1 = /F^\wedge VY$, $V_2^+ = /f \#$, where the unprimed quantities are the usual transmission-line voltages. When normalized voltages are used, show that the scattering-matrix parameters are

$$\begin{matrix} 0 & Z_2 - Z_1, & "19 = "S1 = \frac{2V/Z^A_2}{Z_2 + Z_1} \\ ^\wedge 11 = & ^\wedge 22 = Z_2 + Z_1, & \end{matrix}$$

- 4.20. For the circuit shown in Fig. P4.20. find the scattering-matrix parameters.

$$Z_C$$

$$jX_2$$

$$Z_c = 50 \Omega$$

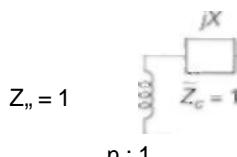
FIGURE P4.20

When $jX_1 = j25$, $jX_2 = j100$ verify that

$$|IS_u|^a + IS_v / = 1$$

$$S_{11}^{-1} S_{22}^T S_{12}^{12} - 0$$

- 4.21. Find Z^* , Z_L and Z_R for the circuit shown in Fig. P4.21.



** FIGURE P4.21

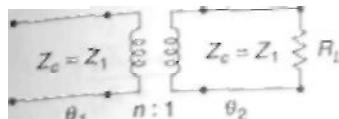


FIGURE P4.22

- 4.22. For the circuit shown in Fig. P4.22, find the input reflection coefficient Γ_{in} .
- 4.23- For the circuit shown in Fig. P4.23, find the expressions for Z_{in} , r_{in} , and the input VSWR. Find the value of jX that will minimize the VSWR.

Hint: The VSWR will be minimum when $|r_{in}|^2$ is a minimum. Why?

FIGURE P4.23

- 4.24. The field of a TE_U mode in a rectangular guide of width a and height b is derived from

$$h_r = C \cos \frac{irx}{a} \cos \frac{vy}{b}$$

Determine the expressions for the equivalent-transmission-line voltage V^* and current I' for the two cases (1) when $Z_c - Z_u$ = wave impedance of the TE_U mode, (2) when $Z_c = 1$.

- 4.25. Apply the complex Poynting vector theorem to show that, for a one-port microwave termination, the reflection coefficient V satisfies the relation

when the wave amplitudes are normalized, so that $W^* = H^*$, that is, the equivalent characteristic impedance is unity.

- 4.26. Show that the $j/4f$ parameters for a section of transmission line of length l and characteristic impedance Z_c are given by $s = \cos pl$, $\ell = jZ_c \sin fl$, $\alpha = jY_c \sin jl$.
- 4.27. For a section of transmission line of length l , show that the wave-amplitude transmission matrix is a diagonal matrix with elements

$$A_{11} = A_{22} = e^{-jl} \quad A_{12} = A_{21} = 0$$

- 8- Consider the junction of two transmission lines as in Prob. 4.19. Using conventional transmission-line voltages, show that the [A]-matrix parameters describing the junction are $A_{11} = A_{22} = (Z_1 + Z_2)/2Z_2$, $A_{12} = A_{21} = (Z_2 - Z_1)/2Z_2$. When normalized wave amplitudes (voltages) are used, show that $A_{11} = A_{22} = (Z_1 + Z_2)/(2Z_1Z_2)$, $A_{12} = A_{21} = (Z_2 - Z_1)/(2Z_1Z_2)$.
- find the [Aj]-matrix parameters for a shunt susceptance jB connected across a transmission line of unit characteristic impedance. Repeat for a reactance jX connected in series with the line.
 - Show that when normalized voltages are used the scattering-matrix parameters of a two-port junction are given in terms of the equivalent-T-network parameters

ters by

$$\begin{aligned} A &= 1 + Z_u - Z_{22} \\ ^{11} &\sim A + 1 + Z_u + Z_{22} \\ S|2 - S_{21} &= \frac{2Z_{12}}{A + 1 + Z_u + Z_{22}} \\ A &= 1 + Z_{22} - Z_j j \\ ^{22} &= \frac{A + 1 + Z_n + Z_{22}}{A + 1 + Z_n + Z_{22}} \end{aligned}$$

where

$$A = Z_n Z_{22} - Z_{12}$$

- 4.31. Show that the T-network parameters are related to the scattering-matrix parameters as follows:

$$Z_{11} = \frac{(1 + S_n)(1 - S_{22}) + Sf_2}{w}$$

$$Z_{22} = \frac{2S_{12}}{w}$$

where

$$W = (1 - S_{11})(1 - S_{22}) - Sf_g$$

- 4.32. For a discontinuity in a waveguide, the following scattering-matrix parameters were measured:

$$^{11} = 5 + Ja \quad ^{12} = Ja \quad ^{22} = a \sim J \sim S$$

Find the parameters of an equivalent T network that will represent the discontinuity (Prob. 4.31).

- 4.33. For an ϵ -plane step (Fig. 4.6), the following were measured:

$$^{11} = 3 + j \quad ^{22} \sim 3 + j$$

An equivalent circuit of the form illustrated in Fig. P4.33 is to be J^1 . Represent the junction. Determine the susceptance jB and the ideal transformer turns ratio $n:l$ from the above given data.

JB

$v.n$

FIGURE P4.33

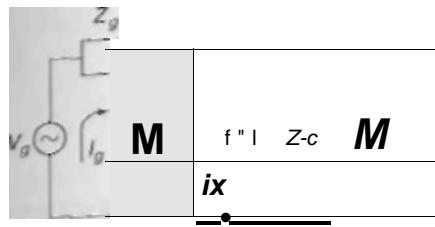


FIGURE P4.34

- *4.34. For the circuit illustrated in Fig. P4.34 construct a signal flow graph relating the variables V_g , I_g , V_1 , I_1 , V_2 , and I_2 . Use signal flow graph analysis to find the voltage V_2 across Z_L .

Hint: See Prob. 4.9 for the T-network parameters of a transmission line.

- *4.35. For the circuit illustrated in Fig. P4.35, construct a signal flow graph. Use signal flow graph reduction to derive an expression for the load voltage V_3 . Use the S^*S^* chain matrix to write relationships between the variables.

Bint: See Prob. 4.26.

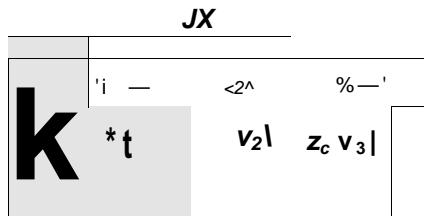


FIGURE P4.35

- *4.36. For the circuit illustrated in Fig. P4.36, construct a signal flow graph using the scattering-matrix relationships between the variables. From a reduction of the signal flow graph find the load voltage V_L .

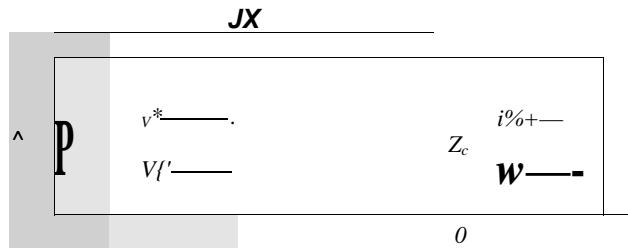
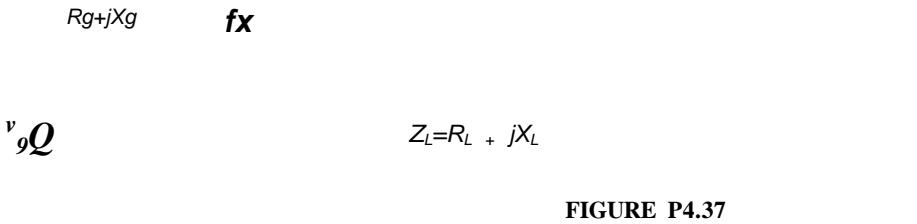


FIGURE P4.36

- For the transmission-line circuit shown in Fig. P4.37, find the generalized scattering-matrix parameters for the series reactance jX . Use these results to derive an expression for the power delivered to Z_L . Verify your answer using a more conventional method of analysis.

**FIGURE P4.37**

***4.38.** Repeat Prob. 4.37 with jX replaced by a shunt element jX_s .

***4.39.** Find the TE_{10} field radiated by the current loop illustrated in Fig. p. Consider the loop area to be so small that (4.119) is applicable. The area of the loop equals S_0 .

$$\begin{array}{c} a/2 \\ / = 0 \end{array}$$

FIGURE P4.39

***4.40.** Find the TE_{10} field radiated by the current loop of Prob. 4.39 if a short circuit is placed at $z = -\frac{a}{2}$.

***4.41.** A linear constant current I extends across the center of a rectangular waveguide at $z = a/2, z = 0$. Show that the total radiated electric field is

$$E_v = \frac{-jwfi_g I}{a} \sum_{n=1}^{\infty} \sin \frac{nv}{2} \sin \frac{nir}{a} e^{-y_n |z|}$$

where

$$y_n = J_0 n = 2 \pi n / a$$

REFERENCES

1. Kerns, D. M.: Basis of Application of Network Equations to Waveguide Problems, *J. Nail. Bur. Std.*, vol. 42, pp. 515-540, 1949.
2. Marcuvitz, N. (ed.): "Waveguide Handbook," McGraw-Hill Book Company, New York, 1951.
3. Montgomery, C. G., R. H. Dicke, and E. M. Purcell (eds.): "Principles of Microwaves," McGraw-Hill Book Company, New York, 1948.
4. Pannenborg, A. E.: On the Scattering Matrix of Symmetrical Waveguide Junctions, *Res. Rept.*, vol. 7, pp. 131-157, 1952.

CHAPTER 5

IMPEDANCE TRANSFORMATION AND MATCHING

In this chapter we are concerned with the important problem of impedance matching, such as the matching of an arbitrary load impedance to a given transmission line or the matching of two lines with different characteristic impedances. Methods of impedance matching to obtain maximum power transfer are presented, along with broadband design methods for quarter-wave transformers and tapered transmission-line impedance transformers. To facilitate the development of the theory, the Smith chart, a graphical aid for the solution of many transmission-line and waveguide impedance problems, is described first.

In a computer-oriented age the reader may very well question why one should be interested in a graphical aid, such as the Smith chart, to solve an impedance-matching problem. There are two basic reasons why the microwave engineer needs to be familiar with the Smith chart. One reason is

^{at} Using the Smith chart to solve an impedance-matching problem shows m a very vivid way how adding reactive elements moves the impedance >mt around and this provides considerable insight into the impedance-matching problem. The second reason is that the Smith chart is widely used

¹ he industry to display the performance of a microwave circuit in terms of nput impedance versus frequency, VSWR or reflection coefficient versus ^u^ncy, the frequency variation of scattering-matrix parameters, etc.

in microwave amplifier design the Smith chart is indispensable as a aid to show how gain, noise figure, stability, and input and output cning are interrelated and how these operating characteristics depend

on the load and source impedances. Without the aid of the Smith chart, intuitive understanding of microwave amplifier design would be much more difficult to acquire.

5.1 SMITH CHART

In Sec. 3.6 it was shown that a load impedance Z_L was transformed into an impedance

$$Z_{TM} = \frac{Z_L + jZ_c \tan \phi}{Z_c + jZ_L \tan \phi} \quad (5.1)$$

when viewed through a length ℓ of transmission line with characteristic impedance Z_r . This formula is valid for any waveguiding system with phase constant $\beta\ell$, provided the impedances are properly interpreted in terms of suitably defined equivalent voltages and currents. Alternatively, the reflection coefficient $T(Z)$, a distance ℓ from the termination, is uniquely given by

$$r(\ell) = \frac{Z_r e^{j\beta\ell} - Z_L}{Z_r e^{j\beta\ell} + Z_L} \quad (5.2)$$

with Z_{mil} given by (5.1). The reflection coefficient is a physical quantity that can be measured, and the normalized impedances Z_n/Z_c and Z_L/Z_c may therefore be appropriately defined in terms of the reflection coefficient r at any point ℓ on the line and the reflection coefficient Y_L of the load, as follows:

$$\begin{aligned} Z_m &= \frac{Z_c}{1 - r(\ell)} & i - r(\ell) &= i - y e^{-j\beta\ell} \\ \sum_I z_i &= I j t^* & I &< i - r(\ell) \end{aligned} \quad (J)$$

The Smith chart is a graphical representation of the impedance-transformation property of a length of transmission line as given by (5.2). Clearly, it would be impractical to plot all values of Z_m in a rectangular coordinate impedance plane, with one coordinate representing the real part, or resistance, and the other coordinate representing reactance, since this would require a semiinfinite sheet of paper, yet on the other hand, all values of the reflection coefficient lie within a unit circle in the reflection-coefficient plane since $|r| < 1$. Furthermore, each value specifies a value of normalized input impedance by means of (5.3a). There is a one-to-one correspondence between reflection coefficient r and normalized input impedance. Instead of plotting contours of constant values of the reflection coefficient, contours of constant values of input resistance and reactance are plotted on the reflection-coefficient plane. For a given value of reflection coefficient, the corresponding input impedance can be read off the chart.

rectly from the plot. In addition, a movement a distance d along the line corresponds to a change in the reflection coefficient by a factor e^{-2pd} only. This is represented by a simple rotation through an angle $2/3d$; so the corresponding impedance point moves on a constant radius circle through this angle to its new value. The chart thus enables the transformation of impedance along a transmission line to be evaluated graphically in an efficient and straightforward manner. A more detailed description of the chart and its use is given below. In addition, a number of matching problems are solved with the aid of the Smith chart in later sections of this chapter.

Let the reflection coefficient T be expressed in polar form as

$$Y = pe^{j\theta} \quad (5.4)$$

where $p = |Y|$ and $\theta = ZT = LT_U - 2/3$. Let the normalized input impedance be

$$Z_m = \frac{1 + T}{1 + pe^{j\theta}}$$

From (5.5) it is readily found that in the reflection-coefficient plane (p, θ plane), the contours of constant R and constant X are given by (Prob. 5.1)

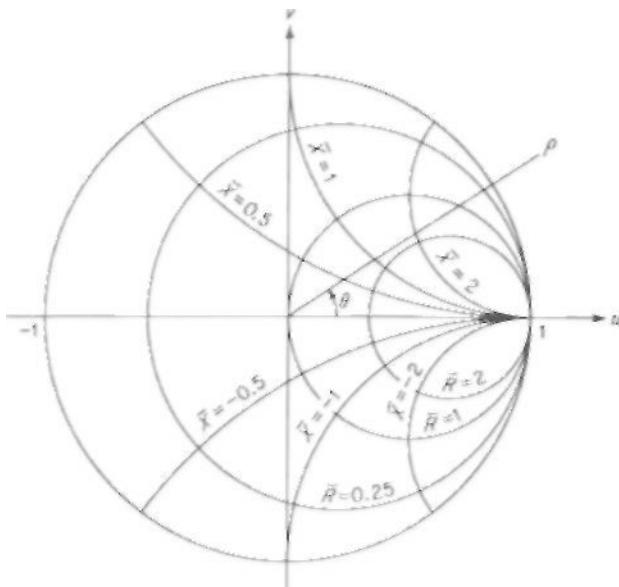
$$u - \frac{R}{\bar{K} + 1] + v^2} = \frac{1}{(S+1)^2} \quad (5.6a)$$

$$(\ll -1 \gg + (* - jLj - i j) \quad (5.6b)$$

where $u = p \cos \theta$ and $v = p \sin \theta$ and are rectangular coordinates in the p, θ plane. The above constant R and constant X contours are circles and plot as illustrated in Fig. 5.1.

For convenience in using the chart, a scale giving the angular rotation $2/3I = 4\pi I/\lambda$ in terms of wavelength λ is attached along the circumference of the chart. Note that moving away from the load (toward the generator) corresponds to going around the chart in a clockwise direction, as illustrated in Fig. 5.2. A complete revolution around the chart is made in going a distance $I = \lambda/2$ along the transmission line. At these intervals the input impedance repeats itself. The origin for the angular scale is arbitrarily chosen at the left side of the circle.

To illustrate the use of the chart, let a line be terminated in a load impedance $Z_L = 0.5 + j0.5$. This point is located in Fig. 5.2 and labeled P_1 . At a distance $l = 0.2\lambda$ away, the corresponding input impedance may be found as follows: A constant-radius circle through P_1 is constructed. The new impedance point P_2 lies on this circle at an angle $2/3l = 0.8$ rad. The new input impedance is

**FIGURE 5.1**

Constant R and X circle in the reflection-coefficient plane.

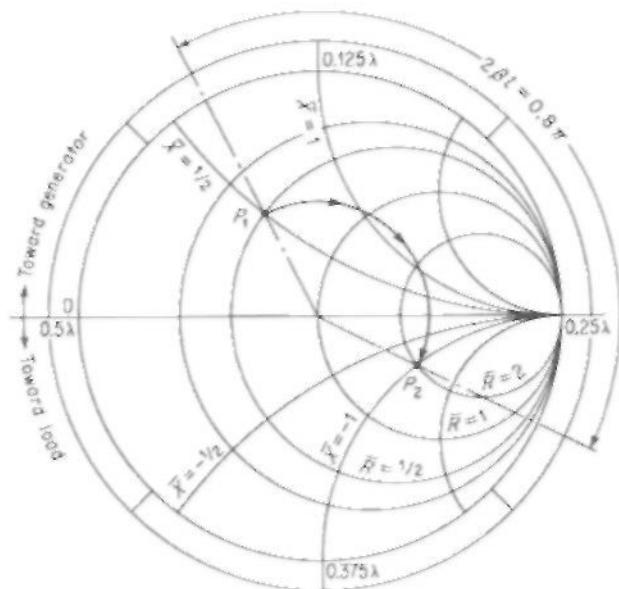


FIGURE 5.2
The Smith chart.

rad in a clockwise direction from P_y . This angular rotation is ^{re} ^{gctio} out by adding $0.2A$ to the wavelength reading obtained from the ^e _{j,grjce} of the radius vector through P_x and the angular scale at the circuit _{jzpj} of the chart. From the chart it is found that the new value of n° impedance is

$$R_2 + jX_2 = 2 - j1.04$$

If we begin at a point P_y , where the impedance is $R_x + jX_h$ and move a constant-radius circle an amount $A/4$ to arrive at a point diametrically opposite, \parallel changes into $-I\backslash$ ($2/3$ changes by TT), and we obtain an impedance

$$R_2 + jX_2 = \frac{\mathbf{i} - \mathbf{r}}{\mathbf{i} + \mathbf{r}} = \frac{\mathbf{i}}{R_I + jX_I} = G, \quad \cdot jB_t$$

Thus the input normalized admittance $G, +jB_t$ corresponding to a given input impedance $R_x + jX_I$ may be found from the value of impedance at a point diametrically across from the first impedance point, provided R_2 and jX_2 are interpreted as^the input conductance and susceptance. To clarify this, note that $R_2 + jX_2$ at P_2 is the normalized input impedance at point P_2 and equals the normalized input admittance at point P_y at a distance $= A/4$ away.

The Smith chart *may be used to* find the transformation of admittances equally well. All that is required is_to interpret the constant resistance and reactance contours (constant R and jX contours) as constant conductance G and susceptance jB contours. Note Jhat a positive X corresponds to an inductive reactance but a positive B corresponds to a capacitive susceptance.

In order to facilitate the use of the Smith chart in situations where it is necessary to convert back and forth between impedances and admittances, the impedance-admittance chart is used. This chart has a second Smith chart, rotated by 180° , superimposed on the regular Smith chart. Thus one set of circles gives impedance values and the second set of rotated circles, usually shown in a different color, give the corresponding admittance values directly. For the conventional Smith chart, the admittance

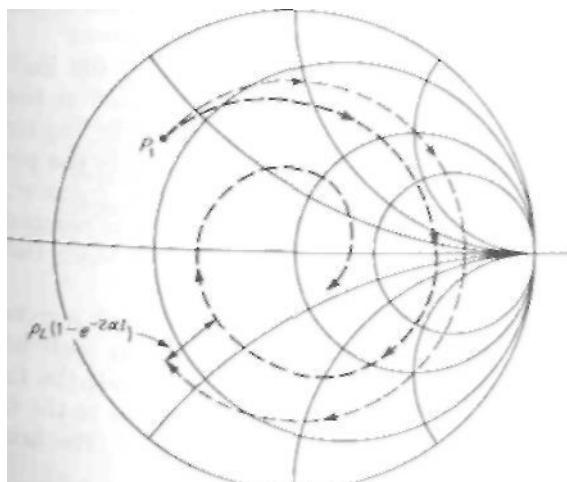


FIGURE 5.3
Inward spiraling of the impedance point on a Smith chart for a lossy transmission line.

values are obtained by rotating the impedance points by 180° impedance-admittance chart, the second set of rotated circles are unnecessary to rotate the impedance point in order to determine the corresponding value of the admittance.

On a lossy line the reflection coefficient at any point is given by

$$\rho_L = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (5.7)$$

As we move from the load toward the generator, $\rho_L = \sqrt{\epsilon} e^{-2\mu L}$ continues to decrease, and hence we move along a spiral that eventually reaches the center, as in Fig. 5.3. In practice, we move on a constant ρ_L circle first through the angle $2/3\pi$ and then move in radially until we are a distance $\rho_L e^{-2\mu L}$ from the center. Many practical charts have convenient scales attached to them, so that the amount of inward spiraling is readily obtained. Note that the center of the chart represents a matched condition ($\rho_L = 0$).

5.2 IMPEDANCE MATCHING WITH REACTIVE ELEMENTS

When a given load is to be connected to a generator by means of a transmission line or waveguide many wavelengths long, it is preferable to match the load and generator to the transmission line or waveguide at each end of the line. There are several reasons for doing this, perhaps the most important one being the great reduction in frequency sensitivity of the match. Although the transformed load impedance as seen from the generator end of the transmission line can be matched to the generator for maximum power transfer, a small change in the operating frequency will change the electrical length ϕ of a long line by an appreciable fraction of π rad, and hence greatly modify the effective load impedance seen at the generator end and thus modify the matching requirements as well. To avoid this frequency sensitivity of the matching requirements, the load and generator should be individually matched to the transmission line or waveguide.

Another disadvantage of not matching the load to the transmission line is that when a matching network is used at the generator end of the line there may be a large standing-wave field along the transmission line if the original load is badly mismatched. This reduces the power-handling capacity of the system since, for a given power transfer, the maximum strength before dielectric breakdown occurs is reached sooner. In addition greater transmission losses are also incurred when there is a standing current along the line.

The techniques that may be used to match a given load impedance to a transmission line or waveguide may equally well be used to mate a generator to the line. Hence it suffices to limit the following discussion to that of matching an arbitrary load impedance to the transmission line. For convenience, normalized impedances are used. The first matching tecB

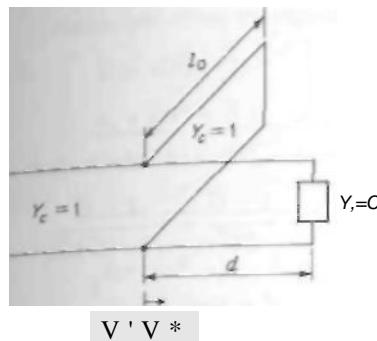


FIGURE 5.4
Single-shunt-stub matching network.

discussed employs short-circuited (or open-circuited) sections of transmission lines as reactive elements, and is referred to as *stub* matching. However, the principles involved are general in nature and may be applied to any waveguiding system by substituting suitable shunt or series reactive elements for the transmission-line stubs. A description of some typical reactive elements that may be used is given in a later section of this chapter.

Single-Stub Matching

Case 1 Shunt stub. Consider a line terminated in a pure conductive load of normalized admittance $Y_L = G$, as in Fig. 5.4. At some point a distance d from the load, the normalized input admittance will be $Y_m = 1 + jB$. At this point we can connect a stub with normalized input susceptance $-jB$ across the line to yield a resultant

that is, to arrive at a matched condition. The stub should be connected at the smallest value of d that will give $Y_m = 1 + jB$ in order to keep the frequency sensitivity as small as possible. The stub may be either an open-circuited or a short-circuited section of line, the latter being the most commonly used version for two-wire lines, coaxial lines, and waveguides because of ease in adjustment and better mechanical rigidity. In a microstrip circuit an open-circuited stub would be preferred since it does not require a connection to the ground plane.

To find the position d , we must solve the equation

$$\hat{Y}_{in} = 1 + jB = \frac{Y_L + jt}{1 + jY_L'} \quad t = \tan pd$$

If we assume that $Y_L = G$ is pure real, we require

$$(l+jB)(l+JGt) = &+jt$$

or by equating real and imaginary parts, we obtain

$$\begin{aligned} 1 - BGt &= G \\ j(B - Gt) &= jt \end{aligned} \quad (5.8a)$$

Equation (5.86) gives $B = (1 - G)t$, and substitution of this into $(^{\wedge} \S_a)$ yield,

$$\frac{1 - G}{EG} \sim \frac{1 - G}{(i - G)U_t}$$

or

$$i^l = \tan^2 Bd = \frac{1}{G}$$

Replacing $\tan^2 Bd$ by $(1 - \cos^2 Bd)/\cos^2 Bd$ finally gives

$$d = \frac{A}{VTTg} / \frac{1}{G} \quad (5.9)$$

where $B = 2TT/A$. Note that two principal values of d are possible, depending on which sign is chosen for the square root. An alternative relation is obtained if we replace $2\cos^2 Bd$ by $1 + \cos 2Bd$; thus

$$\frac{2G}{1 - \frac{1}{\sqrt{1 + G}}}$$

and

$$\frac{G - 1}{1 + \frac{1}{\sqrt{1 + G}}}$$

which gives

$$\frac{A}{u} \frac{G - 1}{\cos \frac{4TT}{A} d_L} \quad (5.10)$$

If d_L is a solution of (5.10), then $A/2 - d_L$ is another principal solution, since $\pm d_x \pm nA/2$ are all solutions of (5.10).

The value of the input susceptance jB is given by

$$\beta = f(i - B) t - \sim \quad (5.11)$$

since $\tan^2 Bd = 1/G$. The required length l_0 of a short-circuited stub to { an input susceptance $-jB$ is found from the relation

$$Y_{in} = -JB = -j \cot Bl_0$$

and (5.11); thus

$$\cot 3/0 = \frac{1 - G}{VG}$$

$$\text{or } l_0 = \frac{A}{2TT} \tan^{-1} \frac{1 - G}{G} \quad (5.12')$$

where the sign of \sqrt{G} must be chosen to give the correct sign for B in

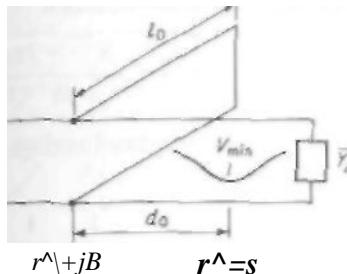


FIGURE 5.5
Location of stub relative to a voltage minimum.

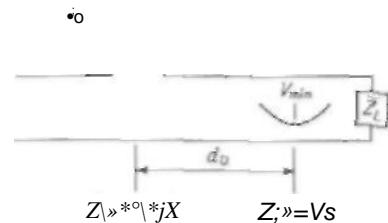


FIGURE 5.6
The series stub.

$0 < d < A/4$, the positive square root should be used, whereas if the other solution, $A/4 < d < A/2$, is chosen, the negative square root must be used.

A similar analysis may be carried out when Y_L is complex, but it becomes *more* involved. The following procedure is usually followed instead. First locate a position of a voltage minimum from the load. At this point the reflection coefficient is a negative real quantity and the input admittance is pure real and given by

$$Y_m = \frac{1 - T}{\mathbf{i} + \mathbf{r}} = \frac{1 + p}{\mathbf{i} - p} = S \quad (5.13)$$

where S is the standing-wave ratio on the line. Let d_0 be the distance from this voltage-minimum point to the point where $F_{in} = 1 + jB$, as in Fig. 5.5. The equations to be solved for the stub position d_0 and stub length l , are the same as given earlier, but with S replacing G . Hence

$$\frac{A}{*''} = \frac{S - 1}{STT} \quad (5.14a)$$

$$A = vS \quad (5.14b)$$

The position of the stub from the load is readily computed by finding the distance from the load to the V'' position and adding this to d_0 . Note that a stub position $d_0 \pm A/2$ is also a suitable one. Thus, if $d_0 - A/2$ is still on the generator side of the load, the stub should be placed at this point instead of at "o in order to reduce the frequency sensitivity of the match.

Case 2 Series stub. At a position of a voltage minimum, $Z_{in} = S'$. At some position d_0 From this point, $Z^A = I + jX$. By connecting a stub with a normalized input reactance of $-jX$ in series with the line at this point, the resultant input impedance is reduced to unity and a matched condition is obtained. This series stub-matching network is illustrated in Fig. 5.6.

To find d_0 , we must solve the equation

$$\frac{S^{-1} + j \tan Bdr}{1 + S} = \tan/3rf_0$$

This is the same equation as considered earlier, with X , S^{-1} replaced by $\tan Bdr$, and thus the solutions are

$$d_0 = \frac{\lambda}{4\pi} \cos^{-1} \frac{S^{-1} - 1}{S^{-1} + 1} = \frac{\lambda}{4\pi} \cos^{-1} \frac{1 - S}{1 + S}$$

$$= \pm \frac{\lambda}{2} + \frac{\lambda}{4\pi} \cos^{-1} \frac{S - 1}{S + 1} \quad (5.15Q)$$

$$\bar{X} = \left(1 - \frac{1}{S}\right) \tan \beta d_0 = \frac{S - 1}{\sqrt{S}} \quad (5.156)$$

where the sign of JS must be chosen to yield the correct sign for $\tan(3d)$ — that is, for $0 < d_0 < A/4$, use $+JiS$, and for $A/4 < d_0 < A/2$, use $-js$.

The required stub length l_0 is determined from the relation

$$j \tan /3/l_0 = -jX$$

and hence, from (5.156), we obtain

$$\frac{A}{-2} \stackrel{A}{?} \stackrel{,I-S}{\text{ton}} \stackrel{,I-S}{\sim} \mathbf{T} \quad (5.16)$$

The shunt stub is most commonly used for coaxial lines because it is easy to construct a shunt stub for a coaxial line, whereas a series stub is difficult to build. A disadvantage with a single-stub-matching system is that every "load" requires a new stub position. The use of two stubs spaced by a fixed amount and located a fixed distance from the load may be used to overcome this disadvantage. However, a double-stub-matching system of this type will not match all possible values of load admittance. The theory of double-stub matching is presented in the next section.

DOUBLE-STUB MATCHING NETWORK

The double-stub tuner, or matching network, is illustrated schematically in Fig. 5.7. We may transform the formalized load admittance Y_L into an equivalent load admittance $Y_L = G_L + jB_L$ at the plane aa and treat the problem illustrated in Fig. 5.76 without loss in generality.

Let the point P_K on the Smith chart in Fig. 5.8 represent Y_L . The stub adds a susceptance jB_l which moves P_K along a constant-conductance circle to point P_2 in Fig. 5.8. At the plane bb just on the right-hand side of the second stub, the input admittance is $Y_b = G_b + jB_b$, and is obtained by moving along a constant-radius circle from P_2 to P_3 through an angle $\theta = 2/3rf = 47rd/A$ rad in a clockwise sense. The point P_3 must lie on the $G = 1$ circle if the addition of a susceptance jB_b contributed by the sec-

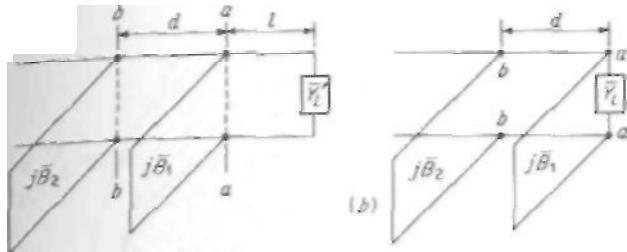


FIGURE 5.7
The double-stub tuner.

stub is to move point P_3 into the center of the chart (matched condition) along the $G = 1$ circle.

From the description just given, it is clear that the first stub must add a *susceptance of just the right amount, so that after the admittance at plane cm is transformed through a length of line d , we end up at a point on the $G = 1$ circle*. The required value of susceptance jB_1 to be contributed by the first stub may be obtained by rotating the $G = 1$ circle through an angle $-f$. The intersection of the rotated $G = 1$ circle and the G_L circle determines the point P_2 , and hence jB_1 as illustrated in Fig. 5.9. A point P'_2 would also be suitable; the location P'_3 then corresponds to the admittance just to the right of the second stub.

From Fig. 5.9 it is clear that for all values of Y_L that lie within the $G = G_0$ circle, a match cannot be obtained since all values $G > G_0$ will not intersect the rotated $G = 1$ circle. The conductance circle $G = G_0$ is tangent

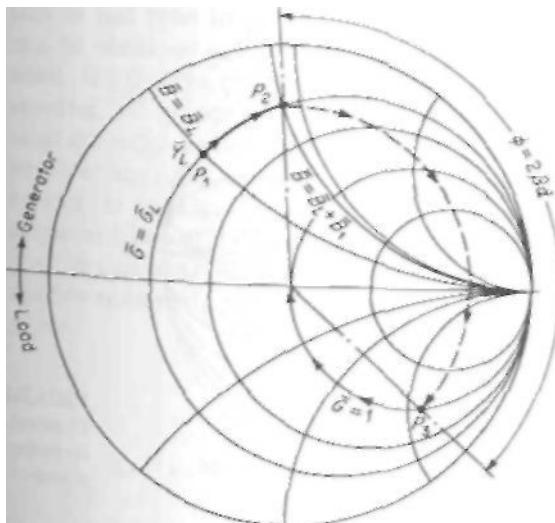


FIGURE 5.8
Graphical representation of the operation of a double-stub tuner.

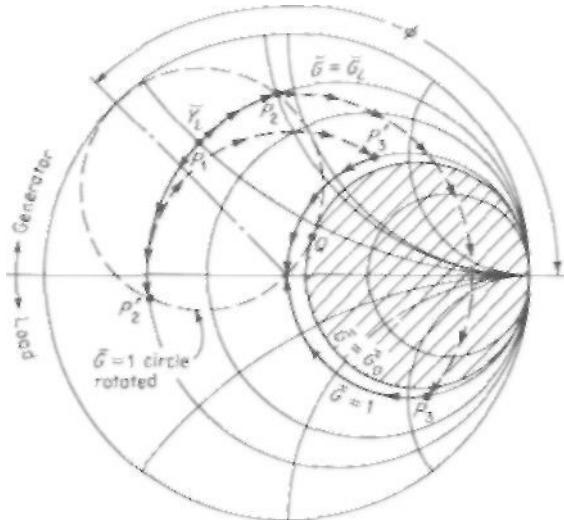


FIGURE 5.9
Graphical determination of required susceptance for the first stub in a double-stub tuner.

to the rotated $G = 1$ circle at the point Q . It is easy to see that the smaller the distance d , the larger the range of load admittances that may be matched (see Fig. 5.10 for the case of $d = A/8$, $4\theta = \pi/2$). Also note that G_0 will always be greater than unity; so all loads with $G_L < 1$ can be matched.

At plane aa in Fig J5.7 we have $y_{aa} = G_L + jB_L$. Just to the left of the first stub we have $Y_a = G_L + jB_L + jB_y$. Just to the right of the second stub

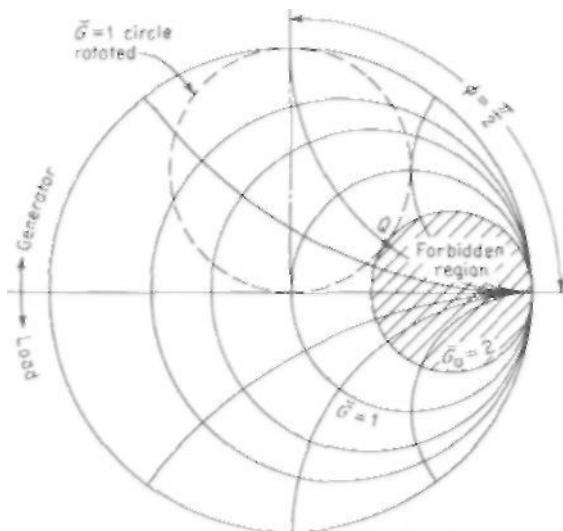


FIGURE 5.10
Illustration of range of load impedance which can be matched when $d = A/8$.

we have

$$F^* \sim \frac{G_L + jB_L + jB_l + jt}{1 + j^*(G_L + jB_L + jB_l)} \quad t = \tan pd \quad (5.17)$$

Since Y_b must equal $1 + jB$, (5.17) gives, upon equating the real part to unity,

$$= 0 \quad (5.18a)$$

or

$$\bar{G}_L = \frac{1 + t^2}{2\epsilon^2} \cdot \frac{1 \pm \sqrt{1 - \frac{U\sqrt{-B_h t - B_x t}}{(1 + t^2)^2}}}{1 \pm \sqrt{1 - \frac{U\sqrt{-B_h t - B_x t}}{(1 + t^2)^2}}} \quad (5.186)$$

In (5.186) we note that the term under the radical sign equals one minus a positive quantity. Since G_L must be real, the term under the radical sign must be positive, or zero. Hence the value of the square-root term lies between zero and one. The corresponding limits on G_L are

$$0 < \bar{G}_L < \frac{1 + t^2}{\sin^2 lid}$$

For any given choice of d , the whole range of load admittances outside the circle $G_0 = \epsilon^2 fid$ may be matched. As an example, for $d = A/8$, $\sqrt{3}d = TT/4$, $\angle f = 2(Zd = IT/2)$, and all values of load admittance outside the $G_0 = \frac{1}{2}$ circle may be matched, as shown in Fig. 5.10. For $d = A/4$, all values of Y_L with $G_L < 1$ may be matched.

Although the theory predicts that virtually all load impedances may be matched by choosing d near zero, or $A/2$, so that $\epsilon^2 fid$ becomes infinite, this is not true in practice. The maximum value of stub susceptance that can be obtained is limited by the finite attenuation of the transmission line used. If jp were replaced by $jp + a$, it would be found that, even with $A/2$ spacing, all values of load admittance could not be matched. In addition, stub spacings near $A/2$ lead to very frequency-sensitive matching networks, so that in practice spacings of $A/8$ or $3A/8$ are preferred. The larger spacing is used at the higher frequencies, where the wavelength is too small to permit use of $A/8$ spacing.

A complete analytical solution to the double-stub matching network is readily obtained. Solution of (5.18a) for the susceptance B , of the first stub gives

$$\frac{n_{ts}}{n_L} = \frac{l \pm \sqrt{(l + t^*) G_L^2 G T t^*}}{t^*} \quad (5.20)$$

where B_L , G_L , and $t = \tan pd$ are all known. Equating the imaginary part

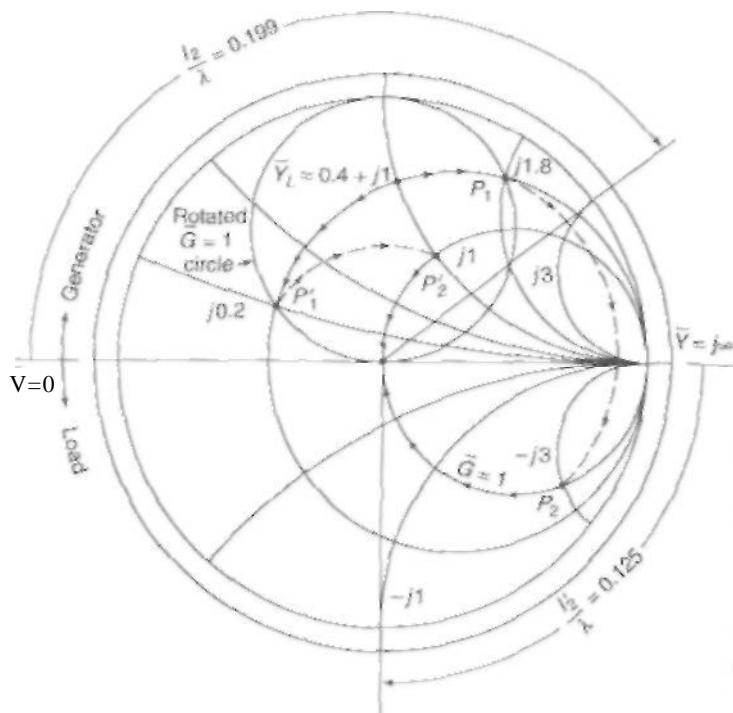


FIGURE 5.11
Illustration of the design of a double-stub matching network.

of (5.17) to jB gives

$$s = \frac{l\Lambda_z IiLiIi^t)(B\Lambda + B_i + t)}{(l-Bj-Sjty+Gf^2c^2)} GU$$

Substituting for B_x into this equation yields

$$\frac{B}{B} + M(1 + \Lambda^2)^{1/2} - G_t \quad (5.2D)$$

The upper and lower signs in (5.20) and (5.21) go together. The susceptance Λ of the second stub must be chosen as $-jB$ in order to provide a mate condition.

Example 5.1. We want to design a double-stub-matching system to ^{ma} normalized load admittance $i_{ij} = 0.4 + jl$. The stubs are spaced $A/\lambda = 47^\circ$. We first construct the $G = 1$ circle rotated by an angle $I\lambda d = 47^\circ$ in the counterclockwise direction as shown in Fig. 5.11. Next we find the point Y_L at the intersection of the $G = 0.4$ and $B = 1$ contours. The point labeled Y_L in Fig. 5.11. We can move Y_L so that it intersects

rotated $G = 1$ circle by moving it to P_+ or P_- along the $G = 0.4$ contour. At $p_{(>)} jB = jl.8$, so we need to add $jB_x = y'(1.8 - 1) = j'0.8$ to P_L to get to P_\pm . At $p_- jB = j0.2$, so to move Y_L to P_L , we need a stub with susceptance $jB_l = #0.2 - 1) = -j0.8$.

The next step is to rotate P_+ and P'_- by $\pi/2$ rad in the clockwise direction. The rotated points lie on the $G = 1$ circle at P_2 and P_2' . At $P_2 jB = -j3$, so we need a second stub with $jB_2 = -jB - j3$ in order to move P_2 into the origin along the $G = 1$ contour. Similarly, in order to move P_2' to the origin, we need a second stub with $jB_2 = -jB = -jl$ since jB at P_2 equals jl .

We can also use the Smith chart to find the stub lengths. For the first solution where we needed a stub with $jB_2 = j3$, we will assume that we use an open-circuited stub. We draw a radius vector from the origin through the point where the $jB = jZ$ contour cuts the outer boundary of the Smith chart as shown in Fig. 5.11. When we move along an open-circuited or short-circuited stub, we will be moving on the $p = 1$ circle⁰ or outer boundary of the Smith chart. At the open-circuit end of the stub, $Y = 0$ and this is the point on the left-hand side of the Smith chart as shown. We begin at this point and move on the $G = 0$ or $p = 1$ circle toward the generator (clockwise) until we get to the $j3$ point. From the normalized distance scale on the outer boundary of the Smith chart (not shown), we find that the required stub length is $l_2/A = 0.199$.

For the second solution where we need $jB_2 = -jl$, we will assume that we use a short-circuited stub. At the short-circuit position $Y = jx$ and this is the point on the right-hand side of the chart. We begin at this point and move on the $G = 0$ circle until we get to $-jl$ as shown. This gives $l_2/A = 0.125$ for the stub length.

The reader can verify that for stub 1 the required length is, for an open-circuited stub,

$$\frac{h}{A} = 0.107 \quad l_2/A = 0.393$$

while for a short-circuited¹ stub,

$$\frac{h}{A} = 0.143 \quad l_2/A = 0.357$$

If series stubs are used, then we work with the normalized load impedance and the normalized reactances of the stubs. A point to keep in mind when designing the length of a stub to give a specified reactance is that at a short-circuit $Z = 0$, while at an open-circuit $Z = j\infty$. Thus, for an open-circuited stub, we begin at the right-hand side of the Smith chart and move clockwise on the $p = 1$ circle until we arrive at the desired jX point. For a shunt stub that was open-circuited, we start at the left-hand-side edge of the Smith chart.

LE-STUB TUNER

The disadvantage of not being able to match all load admittances with a double-stub tuner may be overcome by using a triple-stub tuner, as illustrated in Fig. 5.12. Stub 1 provides a susceptance jB_l such that $Y_L + jB_l$

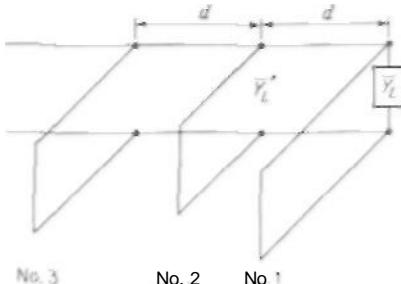


FIGURE 5.12
Triple-stub tuner.

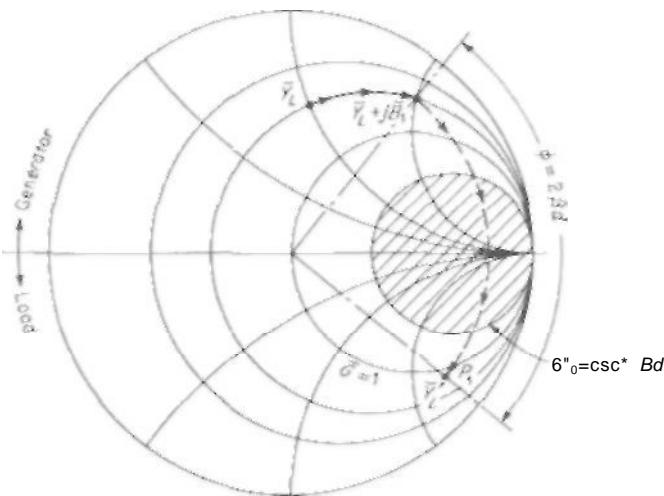


FIGURE 5.13
Transformation of $Y^$ into $Y\ell$.

transforms to some new admittance Y'_L just to the right of stub 2. Stubs 2 and 3 provide a conventional double-stub tuner for matching Y'_L to the line. These two stubs will match all values of Y_L for which $G' < \text{esc}^2 pd$. Thus the function of stub 1 is to ensure that a susceptance jB_x is added such that the transformed admittance Y'_L has a G'_L less than esc^2 . To find a suitable value of jB_1 , we note that, after moving a distance d " Y_L , the admittance $Y_L + jB_1$ must transform into a point P_1 outside the circle $G_0 = \text{csc}^2/3d$, as in Fig. 5.13. If we rotate the $G_0 = \text{csc}^2/3d$ circle through an angle $-\angle \ell = -2(3d)$, we can readily see at once the susceptances jB_x that may be added to $Y^$ to keep the resulting Y'_L on the $G_0 = \text{esc}^2(3d)$ circle. The procedure is illustrated in Fig. 5.14. In example Y_L falls within the rotated $G_0 = \text{esc}^2(3d)$ circle, and susceptance jB_x must be added to move the resultant load $Y_L + jB_x$ or some point beyond P_1 or P_2 , say to P_f or P_g , which provides for a margin.

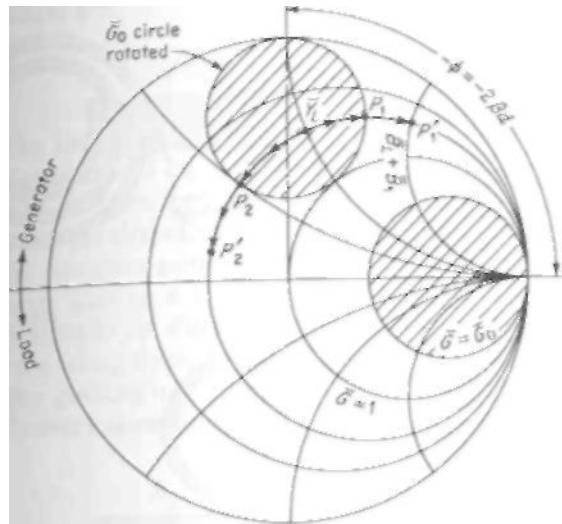


FIGURE 5.14
Graphical solution for jB_t for a triple-stub tuner.

safety. The resultant load when transformed to the position of stub 2 will lie outside the circle $G_0 = \text{esc}^2 / 3d$, and hence can be matched by stubs 2 and 3.

A triple-stub tuner can match all values of load admittances. It may be considered to be two double-stub tuners in series, i.e.,

$$jB_2 = jB'_2 + jB_2$$

so part of jB_2 is associated with each end stub of the two double-stub tuners even though only one physical stub is present in the center. The triple-stub-matching system has more free variables so it can be optimized to increase the bandwidth.

ELEM^{AN}⁰¹³ MATCHING WITH LUMPED

Lumped-parameter elements such as conventional inductors and capacitors are not compatible with coaxial transmission lines and waveguides and consequently are generally not used with these structures for impedance matching. However, the widespread use of microstrip transmission lines and the minaturization of inductors and capacitors have brought lumped-parameter elements into prominence for impedance-matching purposes. In a microstrip-line, series connections of capacitors and inductors are easily made. The shunt connection of a capacitor or inductor is somewhat more difficult since a connection through the substrate to the ground plane must

be made. In the frequency range up to several gigahertz, the use of lumped capacitors and inductors for impedance-matching results in a more compact circuit.

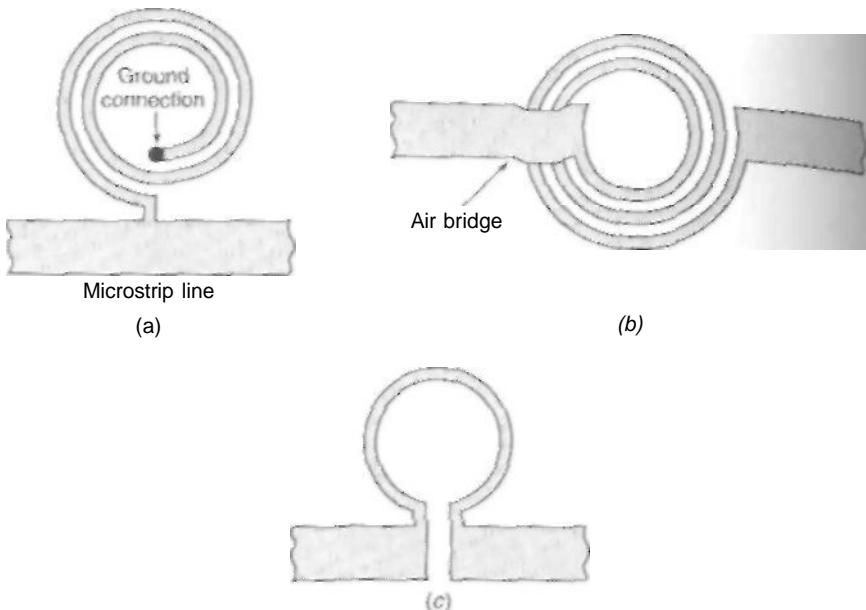


FIGURE 5.15

(a) A shunt-connected spiral inductor; (b) a series-connected spiral inductor; (c) a single loop inductor.

For a microstrip circuit the most common form of inductor is the spiral inductor shown in Fig. 5.15a. For a shunt connection the center of the spiral can be connected to the ground plane. For a series connection an air bridge has to be used as shown in Fig. 5.15b. If only a small inductive reactance is required, a one-turn loop as shown in Fig. 5.15c may be used. If a spiral inductor is to behave as a lumped inductor, its total length must be a small fraction of a wavelength. An estimate of the inductive reactance can be obtained by considering the inductor to be a length L of a transmission line. By using the formulas for the characteristic impedance and propagation constant, namely,

$$T = \gamma \cdot \quad p = cojw$$

we obtain

$$aiL = f\pi Z_c = k_0 Z_{c0} Sl/m$$

where Z_{c0} is the characteristic impedance with air as the dielectric. As can be seen from this expression, a high-impedance line (narrow width) should be used in an inductor. As an example, if we have a 100-f_i 1-m²/v in 1-mm-thick substrate with an effective dielectric constant of 4 at 2 Gnz>

get

$$wL = \frac{y}{A_0} \times 100 = \frac{1}{15} \times 200 = 83.8 \text{ ft/cm}$$

The inductance L is 6.67 nH/cm. A three-quarter-turn loop with a mean diameter of 1 cm would have an inductance of about 16 nH. Unfortunately, a single-turn loop this long would not function as an ideal inductor because every printed-circuit inductor has distributed capacitance associated with it. We can gain some insight into the length restriction by treating the single-turn loop as a short length of a high-impedance transmission line. With reference to Fig. 5.15c, let Z_0 be the characteristic impedance of the microstrip lines, let Z_c be the characteristic impedance of the transmission line making up the inductor, and let Bl be the electrical length of the loop. Transmission-line theory gives

$$\begin{aligned} Z_{in} &= \frac{Z_0(Z_c + jZ_0t) + jt(Z_c - Z_0)}{Z_c + jZ_0t} \\ &= \frac{jt(Z_c - Z^*)}{Z^* + jZ_0t} \end{aligned}$$

where $t = \tan Bl$.

Since we normally would use a large value of Z_c and keep Bl small, we will assume that $Z^*t^2 \ll Z^2$ in which case we get

$$Z_{in} = Z_0 + Z_0t \approx 1 - \frac{Z_0}{Z_c} | + jZ_0t$$

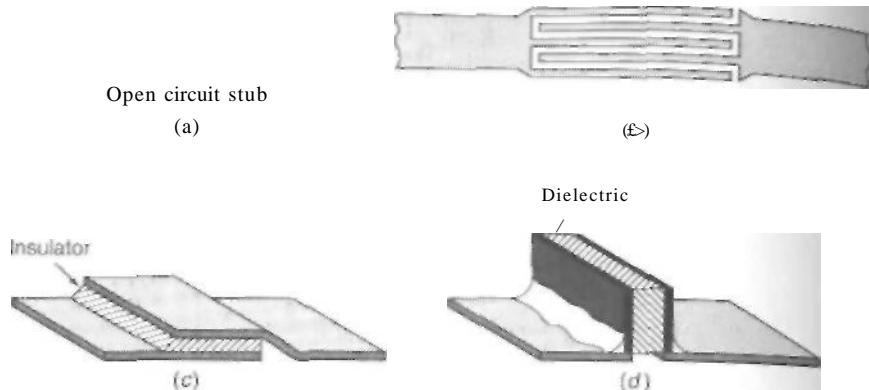
For a short-circuited line,

$$jX_L \approx jZ_c \tan Bl = jZ_c Bl$$

Hence

$$Z_{in} = Z_0 + Z_0t^* \quad \left(\begin{array}{l} Z_c^2 \\ Z_c \end{array} \right) \quad l \approx \sqrt{| + jX_c|l}$$

Thus the effective inductance is reduced by the factor $1 - Z^*/Zf$. A more serious departure from an ideal inductor behavior is the change in the real part of Z_m . For an ideal series-connected inductor, we would have $Z^* = o^+ J^* L^- \propto n$ order that the actual Z_m should approximate this ideal result, we need to make the factor $t \approx 1 - Z^*/Z^*$ of the order of 0.05. In a typical situation where $Z_0 = 50$ ft and $Z_c = 100$ ft, this requires that Bl be no larger than 0.26. The corresponding maximum value for the length I is $\sqrt{15/2} = 7.5$ cm. Thus I should be no greater than 0.3 cm. This leads to a maximum usable inductance of only 2 nH. We can, of course, use a larger w for l but the change in the resistive part of Z_{in} away from the ideal

**FIGURE 5.16**

(a) A short open-circuited stub; (b) an interdigital capacitor; (c) a metal-insulator-metal (MIM) capacitor; (d) a chip capacitor soldered across a microstrip-line gap.

value Z_0 must then be taken into account in the design of a circuit requiring a series inductance.

The spiral inductor can provide larger values of inductance. For example, a five-turn spiral inductor approximately 1.4 mm in diameter, with a conductor width of 0.06 mm and spacing 0.038 mm, has an inductance of 25 nH at 2 GHz.t

A short open-circuited stub as shown in Fig. 5.16a will function as a lumped capacitor connected in shunt across a microstrip line. The distributed capacitance of a microstrip line is typically in the range $0.2y^{\infty}$ to $y/f7$ pF/cm; so a short stub is suitable for providing a shunt capacitance up to about 1 pF. At 4 GHz a 1-pF capacitor has a reactance of about 40 ft. The interdigital capacitor shown in Fig. 5.16b can provide a series capacitance up to several picofarads depending on the number of fingers used and their length. For monolithic microwave integrated circuits, the metal-insulator metal (MTM) capacitor shown in Fig. 5.16c is generally used. A capacitance up to 20 pF or more can be obtained since the insulator thickness can be very small. For example, an insulator 1 mm by 1 mm and $10/xm$ thick having a dielectric constant of 10 provides a capacitor with a capacitance about 9 pF. For hybrid microstrip circuits, the chip capacitor illustrated in Fig. 5.16d¹ is used. It is soldered in place and can provide a capacitance up to 100 pF or more.

^tD. A. Daly, S. P. Knight, M. Caulton, and R. Ekholdt, Lumped Elements in f^{LiCTO} , Integrated Circuits. *IEEE Trans.*, vol. MTT-15, pp. 713-721, December, 1967.

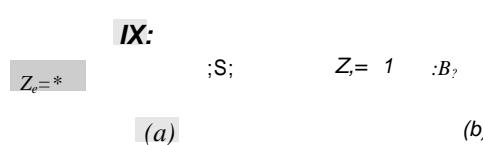


FIGURE 5.17
Two basic lumped-element matching circuits, (a) Circuit used when $G_L < 1$; (b) circuit used when $R_L < 1$.

The basic lumped-element impedance-matching circuit is a circuit consisting of a parallel-connected and a series-connected reactive element as shown in Fig. 5.17. The topology of the circuit is such that it is commonly referred to as an L matching network (compare with the designation of T and FI for the impedance and admittance networks for a two-port network). For the circuit shown in Fig. 5.17a, the shunt element jB_L is connected in parallel with the load and the series element jX_2 is connected in series. This configuration can be used to match any load admittance Y_L having a normalized conductance $G_L < 1$. The circuit shown in Fig. 5.17b can be used to match any load impedance having a normalized load resistance R_L less than 1. For some load impedances both G_L and R_L are less than 1 and either circuit may be used. When the normalized admittance lies inside the $G = 1$ circle on the Smith chart, the corresponding normalized load impedance, which is the reflection of the normalized load admittance through the origin, will lie outside the $R = 1$ circle and hence can be matched using the circuit shown in Fig. 5.17b. When Z_L lies inside the $R = 1$ circle, Y_L will lie outside the $G = 1$ circle and can be matched using the circuit shown in Fig. 5.17a. When both Z_L and Y_L lie outside the $R = 1$ and $G = 1$ circles, either matching circuit can be used.

The required values of the matching elements are easily found using the Smith chart. The procedure to be followed is described below. From this the reader will easily understand the rationale that underlies the use of the Smith chart to solve the matching problem using lumped reactance elements. Since the Smith chart procedure uses normalized immittance parameters, the first preliminary step is to determine the normalized values of Y_L and Z_L by dividing by the characteristic admittance Y_e or characteristic impedance Z_c , respectively, of the input transmission line.t

Case 1. The circuit in Fig. 5.17a is used. A match can be obtained only if $G_L \leq 1$. If $G_L > 1$ use the circuit in Fig. 5.17b (see Case 2).

With reference to Fig. 5.18.

1. Construct the $G = 1$ circle rotated by 180° .
2. From the point 7_t add jB_L to move along a constant-conductance circle until the rotated $G = 1$ circle is intersected. There are two possible solutions.

term immittance is used to designate either an impedance or admittance.

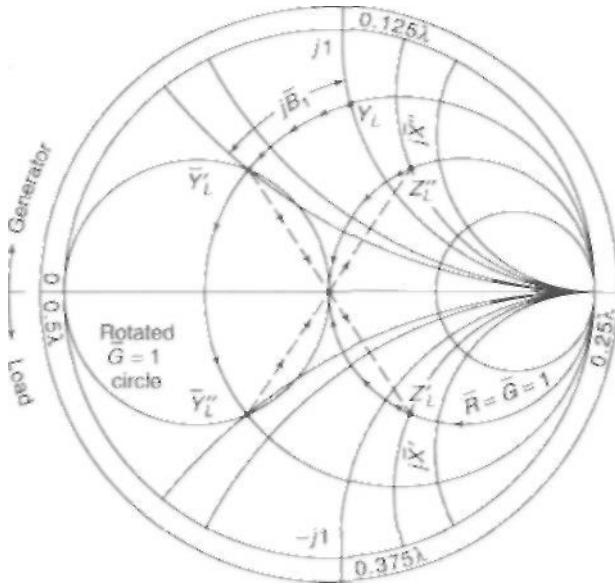


FIGURE 5.18

Illustration of steps followed in designing the matching circuit in Fig. 5.1 la (Case 1).

The new values of Y_L are Y'_L and Y''_L . Note that if Y_L is inside the $G = 1$ circle, it cannot be moved to intersect the rotated $G = 1$ circle because adding jB only moves Y_L on a constant-conductance circle.

3. Reflect Y'_L and Y''_L through the origin to get the corresponding impedance values Z'_L and Z''_L . These lie on the $f_t = 1$ circle since we made Y'_L and Y''_L lie on the rotated $G = 1$ circle.
4. Since Z'_L and Z''_L lie on the $f_t = 1$ circle, these impedance points can be moved into the origin by subtracting a reactance jX' or jX'' . The origin represents a matched condition. Hence the required value of jX_2 is either $-jX'$ or $-jX''$. The required value of jB^A is $j(B' - B_L)$ or $j(B'' - B_L)$.

The greatest bandwidth is obtained when the reactive elements are as small as possible so that the circuit Q is as low as possible. We will return to this point after we provide the steps to be followed using the other matching circuit.

Case 2. The circuit in Fig. 5.176 is used. A match can be obtained only if $R_L < 1$. If $R_L > 1$ the circuit in Fig. 5.17a must be used (Case 1).

1. Construct the $f_t = 1$ circle rotated by 180° as shown in Fig. 5.19-
The motion is along a constant resistance circle. There are two solutions & L ZZ-
2. Add jX_1 to Z_L to move Z_u to intersect with the rotated $f_t = 1$ circle.
3. Reflect Z'_L and Z''_L through the origin to obtain Y'_L and Y''_L .

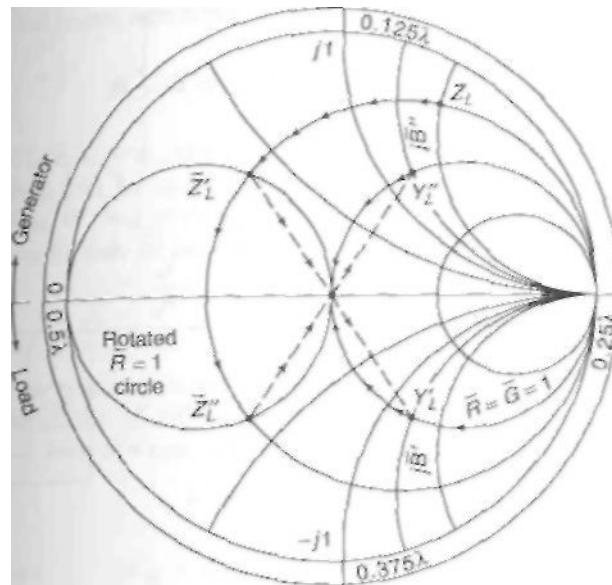


FIGURE 5.19
Illustration of steps followed in designing the matching circuit in Fig. 5Mb (Case 2),

4. The admittance points Y_L and Y_L'' can be moved into the origin along the $G = 1$ circle by subtracting jB' and jB'' , respectively. Hence the required value of jB_2 is $-jB'$ or $\sim jB''$. The required value of jX , is $j(X' - X_L)$ or $j(X'' - X_L)$.

Analytic solutions for the required values of the matching elements are readily derived. The analytic solutions are given in Probs. 5.17 and 5.18 but the reader has to supply the derivations.

Circuit Q and Bandwidth

When a complex load impedance has been matched to a transmission line with characteristic impedance Z_c , the input impedance looking toward the load equals Z_c . Thus the reactive elements present in Z_L and the matching network make up a resonant circuit that is loaded by R_L and Z_c . A resonant circuit has a quality factor, or Q , that can be evaluated from the general definition

$$Q = \frac{\text{average stored electric and magnetic energy}}{\text{power loss}} \quad (5.22)$$

At resonance the average stored electric energy in the capacitors equals the average stored magnetic energy in the inductors. Hence Q can be expressed as

$$Q = \frac{2i\omega W_c - 2t\omega W_m}{\text{power loss}} \quad (5.23)$$

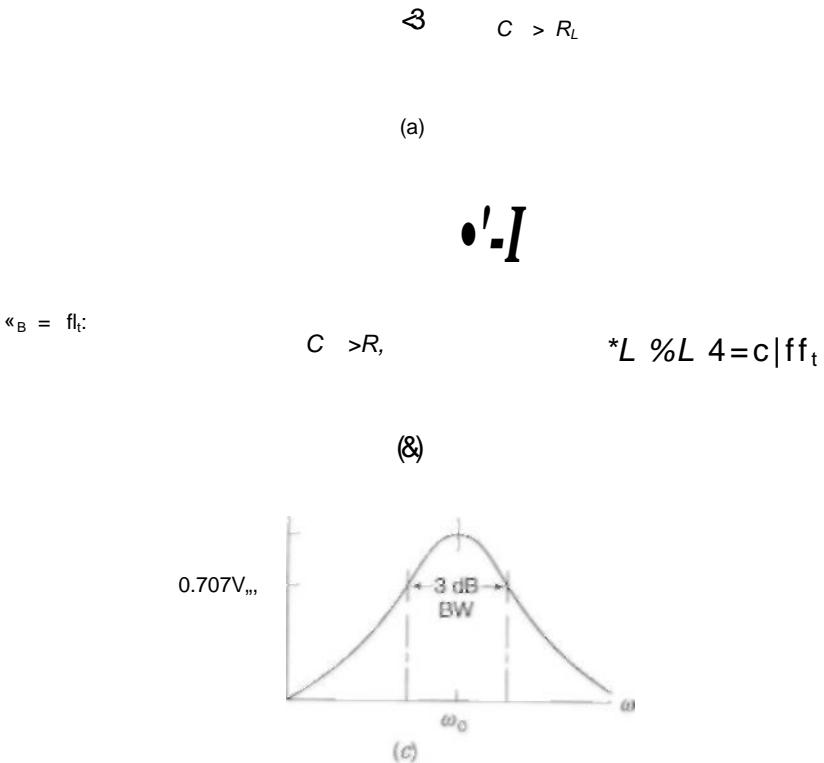


FIGURE 5.20

(a) Parallel LCR resonant circuit; (b) loaded resonant circuit; (c) frequency-response curve showing half-power bandwidth.

The bandwidth of the circuit is the frequency band over which one-half or more of the maximum power is delivered to the load. This bandwidth is called the half-power or 3-dB bandwidth and is inversely proportional to the loaded Q of the circuit.

In order to clarify the above concepts, we will analyze the parallel resonant circuit shown in Fig. 5.20. For the circuit in Fig. 5.20a, the voltage across R_L is given by

$$V = Y_m \quad G_L + j\omega C - j\omega/L \quad (5.24)$$

where $G_L = 1/R_L$. The resonant frequency of the circuit is given by

The input admittance can be expressed in the form

$$\mathbf{y}_{fa} = \mathbf{GL} + > \mathbf{C}^{\text{(OJ}^2 - \text{ft)}} \wedge \mathbf{r}$$

When $\omega_0 = \infty_0$ then $V = I_g R_L$, which is the maximum load voltage that can be obtained. When $j\omega C(u^2 - \omega_0^2)/\omega^2 = y''G_z$, we see from (5.24) that $|V| = / / v^2 G_x$, and hence the power in i_L is only one-half that at resonance. In terms of the circuit elements, the Q is given by

$$\frac{R_L}{Q} = \frac{T^2 R L O_o C}{i_0 L} = \frac{u_0 C}{O_0} \quad (5.26)$$

The reader can verify that this is the same Q as would be found using the definition (5.23) (see Prob. 5.19). When $Q > 10$ the frequency ω is close to ω_0 , over the useful band of operation; so we have

$$\begin{aligned} \omega^2 - C\partial'Q &= (OJ - \omega_0)(\omega + \omega_0) \approx 2\omega(\omega - \omega_0) \\ \text{and } Y_{ii} &\approx G_L + j\omega C \frac{2\omega(\omega - \omega_0)}{IO^2} = G_i + 7\omega_0 C - \frac{2(\omega - \omega_0)}{UJ_i} \\ &= G_j + 2JQ \wedge \backslash \end{aligned} \quad (5.27)$$

where $A\omega = \omega - \omega_0$. From this expression it is readily seen that the 3-dB fractional bandwidth is given by

$$2Q \frac{Aw}{\omega_0} = 2Q \frac{BW}{2} = 1$$

or $BW = 1/Q$ since the half-power points occur when $2jQ A\omega/\omega_0 = \pm j$.

When the circuit is connected to a matched source as shown in Fig. 5.206, it is clear that

$$V = \frac{S}{2G_L(l + 2jQ_L u / \omega_0)} \quad (5.28)$$

where $Q_L = Q/2$ is the loaded Q of the circuit. Consequently, for the loaded circuit, the bandwidth is twice as great. A typical response curve showing the bandwidth is given in Fig. 5.20c.

In an impedance-matching problem, there are generally two solutions available. If we want a narrowband design, we should choose the solution that gives the largest loaded Q . On the other hand, when we want a broadband match, we should choose the circuit with the lowest loaded Q .

load impedance plus matching network will contain either one capacitor and two inductors or one inductor and two capacitors. Thus, when using —3) to evaluate the circuit Q , the energy stored in the single reactive element or that stored in the two opposite reactive elements must be

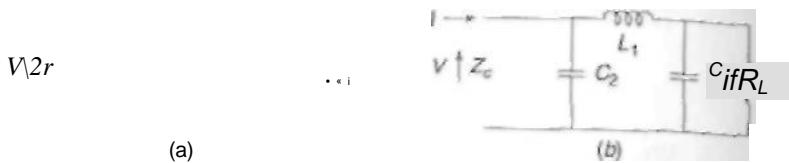


FIGURE 5.21
Two matching networks.

evaluated. We will illustrate the evaluation of circuit Q by means of examples that show the steps involved. By making use of the facts that resonance $Z_{in} = Z_c$ and that $P_{in} = P_L$, the evaluation of the circuit Q can be reduced to a few simple steps.

For the circuit shown in Fig. 5.21a, we will base the evaluation of the circuit Q on the energy stored in the capacitor C_L . The power dissipated in R_L is given by

$$PL = \frac{IVV}{2R_L}$$

The average electric energy stored in C_L is

Hence from (5.23) we get

$$Q = \frac{2u >_0 W_e}{UO^C L^R L} = 5$$

The loaded Q is $Q_L = Q/2$ since the system is matched at w_0 . The operating bandwidth is determined by the loaded Q .

For the circuit in Fig. 5.21b, we note that

$$P_{in} = \frac{1}{2} \frac{|V|^S}{Z_c} = PL = \frac{IVV}{2R_L}$$

since $Z^A = Z_c$ at $w = w_0$. The energy stored in C_2 is

$$w_{e2} = \{iv^2\}c_2$$

while that stored in C_L is

$$W_{ei} = |vJc_L - |c_L^A|vf|$$

of VI-

where we used the equality of P_{in} and P_L to express $|V_L I$ in terms of v .

The circuit Q is given by

$$Q = \frac{2\pi_0(W_{e2} + W_{eL})}{PL} = \frac{\pi_0(C_2 + (R_L/Z_C)C_L)V}{W_L^A/R_L} = \pi_0 C_2 Z_f + c_0 C_L R_L$$

This circuit Q is larger than that for the circuit in Fig. 5.21a. In general, the smallest circuit Q is obtained when the matching circuit contains two similar reactive elements that are opposite to the reactive element in the load i.e., if the load contains an inductive element, the matching circuit should be made up of capacitive elements only, and vice versa.

For circuits with a loaded Q of 5 or more, the frequency response is very nearly the same as that of the parallel resonant circuit shown in Fig. 5.20, over the useful operating frequency band. Thus the 3-dB fractional bandwidth is $\Delta f/f_c$. For circuits with a low Q , the frequency behavior is different but similar.

The discussion above has been based on the assumption that ideal lossless inductors and capacitors are used. As a general rule, lumped capacitors have very little loss and the assumption of negligible loss does not introduce significant error. However, lumped inductors do not have negligible loss, and in high- Q circuits inductor loss should be taken into account. The Q of an inductor is the ratio of the reactance ωL to the series resistance R of the inductor. Lumped inductors may have Q values in the range of 25 to several hundred. If the loaded circuit Q is less than about one-fifth of the inductor Q , then neglecting the loss in the inductor will not produce serious error. When several inductors are used in a matching circuit, the losses in the inductors should be taken into account since these losses can account for a significant fraction of the total power delivered to a circuit.

For microwave circuits the useful bandwidth is often much smaller than the 3-dB bandwidth. It is not uncommon to require an impedance match providing a VSWR of no more than 2, or even less than 1.5 in critical applications. A VSWR of 2 corresponds to a reflection coefficient of 0.333. With this mismatch a fraction of 0.889 of the incident power is delivered to the load.

The degree of mismatch is usually described in terms of the input VSWR or in terms of the return loss. The return loss is the ratio of the reflected power to the incident power, expressed in decibels; thus

$$\text{Return loss} = RL = -20 \log p = -20 \log \frac{\text{VSWR} - 1}{\text{VSWR} + 1}$$

A well-matched system will have a return loss of 15 dB or more. A VSWR equal to 2 gives a return loss of 9.54 dB, while a VSWR of 1.5 gives a return loss of 13.98 dB. Thus a return loss of 15 dB, corresponding to a VSWR of 2, does indeed represent a well-matched system.

In order to get a feeling for how narrow the bandwidth of using this criterion, let us assume that the loaded Q equals 1.43, we have $p = 0.178$. We can use

$$Y_{in} = Y_e \left(1 + \frac{A_{co}}{2jQ} \right)$$

$$P = \frac{Y_o - Y_e}{Y_o + Y_e} = \frac{2Q A_{co}/co_0}{1 + 2Q A_{co}/co_0}$$

where $Q = 2Q_L$. Hence the fractional bandwidth is

$$\frac{2\Delta f}{f_0 Q} = \frac{2p}{Q} = \frac{p}{Q_L} = 0.0356$$

This is much smaller than the 3-dB bandwidth which equals $1/Q$ or 0.9

5.6 DESIGN OF COMPLEX IMPEDANCE TERMINATIONS

In microwave amplifier design, we often require a complex load impedance for the output and a complex source impedance for the input even though the final input and output connections are to 50- Ω lines. In a typical amplifier circuit as shown in Fig. 5.22a, the function of the input and output matching networks is to transform the 50- Ω line impedances into the required complex source impedance Z_s and load impedance Z_L . In a broadband amplifier design, the design of the matching networks can be quite complex. In narrowband amplifier design, elementary matching networks can be used.

The matching problem for an amplifier is the reverse of that of matching a complex load impedance Z_L to a transmission line. The first step is to choose a network topology to be used. In Figs. 5.226 to d we illustrate three possible networks that could be used. Many other choices can be made. We will illustrate the use of the Smith chart to design matching networks of the type shown in Fig. 5.22. The methods used can be extended to more complex networks.

Design procedure for the circuit in Fig. 5.226. The matching network in Fig. 5.226 consists of a transmission-line stub placed at a distance l_{st} from the input. The stub may be either an open- or short-circuited or a lumped susceptance since its function is to provide a normalized admittance jB_i connected across the transmission line at a distance l_{st} from the input. The following procedure should be used to determine A and B :

- Locate the required value of Z_L on the Smith chart and reflect Z_L through the origin. This procedure is illustrated for the case where $Z_L = 0.4 - j0.2$.

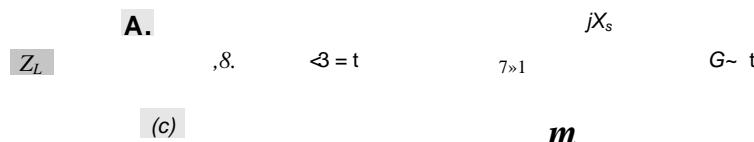
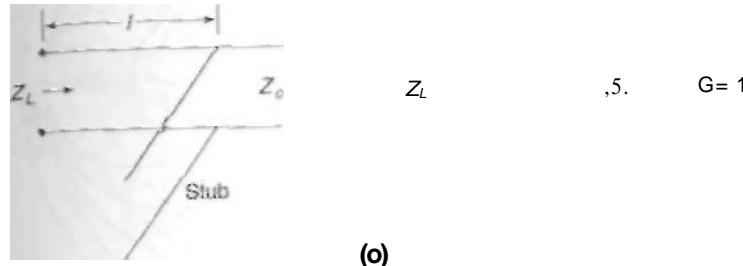
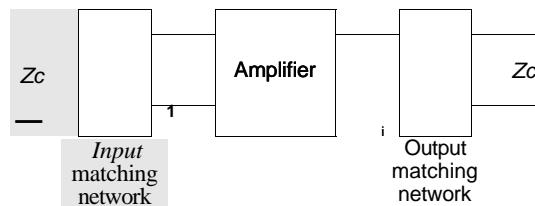


FIGURE 5.22

(a) Microwave amplifier circuit; (b) transmission-line matching network; (c) and (d) alternative matching networks.

2. Rotate the point Y_{ij} on a constant-radius circle until it intersects the $G = 1$ circle. There are two possible solutions. From the angle of rotation the normalized length $|l|$ is found. Note that the rotation is in a *counterclockwise direction since we are moving toward the load*.
3. When we move to the right of the point where jB_x is connected, Y_m decreases by jB_l to become $Y_{jn} \sim 1 + jB - jB_v$. But to the right of jB_v we require $Y^* = G = 1$ so $B_x = B$. Thus the values of jB on the $G = 1$ circle, where the rotated Y_L intersected the circle, gives the required values of jB_x for the two possible solutions.

Note that if we begin at $Y_m = 1$ and move left across jB_U then Y_m changes to $1 + jB_x = 1 + jB$. A rotation in the clockwise direction through an angle $2/3\pi$ then makes $J^* = Y_L$. Thus using the reverse of the usual matching procedure results in the desired value of Y_L .

For our specific example, $l = 0.0626A$ or $0.4375A$ and the corresponding values required for jB_x are jI or $-j/l$, respectively, as shown in Fig. 5.23.

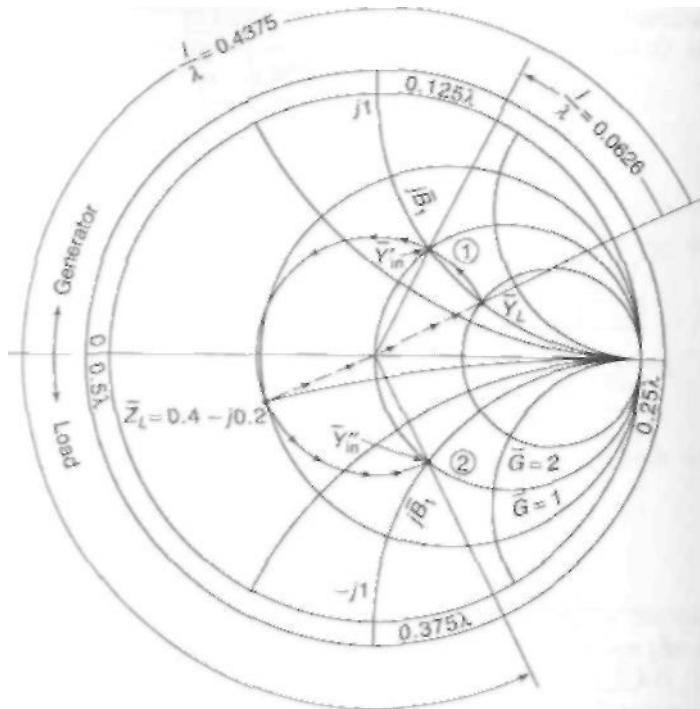
**FIGURE 5.23**

Illustration of design procedure for the matching network in Fig. 5.226.

The analytic solution for the above matching network is

$$B_L = \pm \frac{Bl + Gl + I - 2G_L}{G} \quad \frac{R'i + Xl + l^2}{R} = \frac{! \times l^*}{2R_L} \quad (5.29a)$$

$$\tan pi = \frac{G_L - I}{B_L + Bf_l} \quad \frac{Rl + Xt - R_L}{X_L - B_l R_L} \quad (5.296)$$

Design procedure for the circuit in Fig. 5.22c. The circuit shown in Fig. 5.22c can be used only if $R_L < 1$. The following steps should be followed to design this circuit using a Smith chart:

1. Construct the $R = 1$ circle rotated by 180° .
2. Locate the desired impedance Z_L on the Smith chart.
3. Move Z_L in a counterclockwise direction (toward the load) on the constant circle until it intersects the rotated $R = 1$ circle as shown in Fig. 5.24. There are two possible solutions. The required value for the reflection coefficient jX_L is given by $jX_L = jX_L - jX$, where jX equals jX' or jX'' .

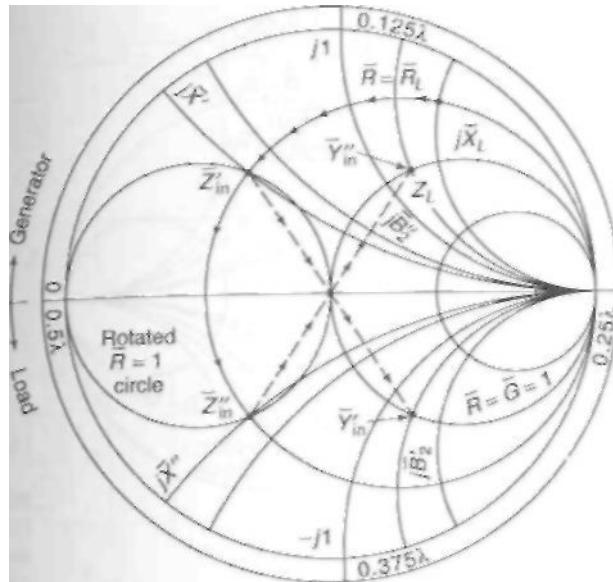


FIGURE 5.24
Illustration of design procedure for the matching network in Fig. 5.22c.

value obtained where Z_L intersects the rotated $R = 1$ circle. The two possible values of jX are always equal in magnitude but opposite in sign.

4. Reflect the two new values of Z^{\wedge} through the jprigin to obtain the corresponding values for F_{in} . This gives $V_{in} = 1 + jB$. As we move to the right of the element jB_2 , we obtain $Y_{in} - jB_2$, which must be equal to 1. Hence the required value of jB_2 is equal to jB .

The reactance jX_x and susceptance jB_2 can be realized using lumped elements or transmission-line stubs.

Design procedure for the circuit in Fig. 5.22f. The matching network shown in Fig. 5.22d can be used only when $G_L < 1$ (can you explain why?). The network can be designed using the Smith chart by following the steps given below.

1. Construct the $G = 1$ circle rotated by 180° .
2. Locate the desired load admittance point Y_L (reflection of Z_L through the origin).
3. Move Y_L on the $G = G_L$ circle in a counterclockwise direction until it intersects the rotated $G = 1$ circle. There are two possible intersections. The required value of jB_L is given by $jB_x = j(B_L - B)$, where jB equals jB' or jB'' , as shown in Fig. 5.25.
4. Reflect the admittances Y'_L and Y''_L through the origin to get the corresponding impedances Z'_m and Z''_m , that lie on the $R = 1$ circle.
5. When we subtract jX_2 we must get $Z_m = 1$, so $jX_2 = jX$, where jX equals jX' or jX'' .

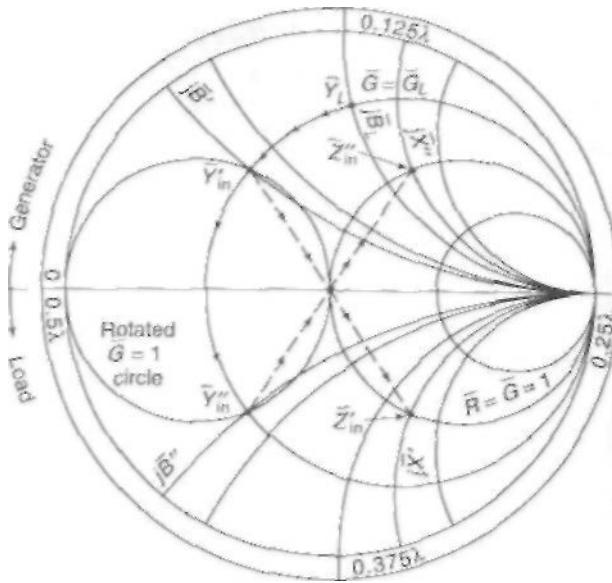


FIGURE 5.25
Illustration of design procedure for the matching network in Fig. 5.22d.

5.7 INVARIANT PROPERTY OF IMPEDANCE MISMATCH FACTOR

Consider the network shown in Fig. 5.26a. The T network between the load impedance Z_L and the source will be assumed to be lossless. The impedance looking into the network toward the load is Z_m . For this network the input current is

$$I = \frac{\mathbf{V}_s}{\mathbf{Z}_s + \mathbf{z}_{in}}$$

and the input power is

$$P_{in} = \frac{R_m}{2(\mathbf{Z}_s + \mathbf{z}_m)} R_m = gU_1 I_x$$

If $\mathbf{z}_{io} = Z^*$ then $R_m = R_s$, $I_x = X_s$, and we get maximum power transferred from the source to the load resistance R_L . This power is

$$P_{in} = \frac{IJKT_fita}{2(R_{in} + R_s)} \frac{1.1V}{2(4R_s)} = P_{in}$$

(5-30)

and is called the available power from the source. When $R_{in} < R_s$ we have an impedance mismatch and $P_m < P_{ava}$. We can express P_m as

$$P_m = \frac{1}{2} \frac{WfR_s}{|Z_s + Z|^2} \frac{1}{2} \frac{IVJ}{AR_{in}} \frac{*R_s R_m}{M} = *_{ava} M$$

(5-31)

where $M = 4R_{in}R_s/Z_{in} + Z^2$ is called the impedance mismatch factor.

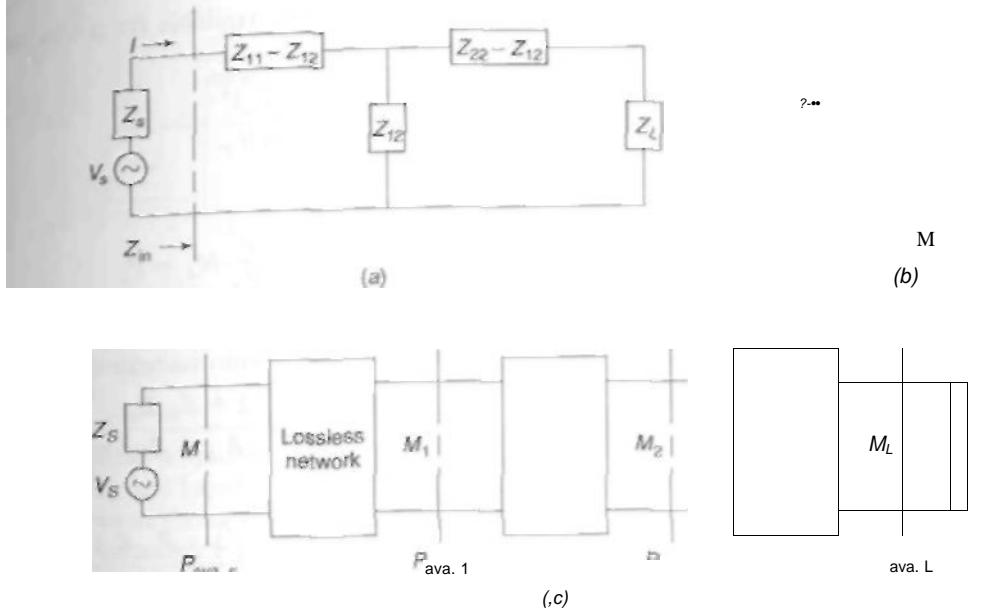


FIGURE 5.26

(a) A T matching network; (b) Thevenin equivalent network; (c) a cascade connection of lossless networks for which $P_{ava,i} = P_{ava,1}$ for $i = 2, 3, \dots, L$.

We can use the Thevenin equivalent circuit shown in Fig. 5.266 to calculate the power delivered to R_L , which is P_L and must equal P_{in} since the network with elements Z_u, Z_{12}, Z_{22} is lossless.

From circuit theory we find that

$$V_{00} = \frac{A + Z_{22}Z_s}{Z_u + Z_s} \quad \text{if } Z_u \neq -Z_s$$

where $A = Z_u Z_{22} - Z_s^2 = X_f z - X_u X_{22}$, upon using $Z_u = jX_u$, $Z_{22} = jX_{22}$, $Z_{12} = jX_{12}$. The available power from the network is

$$\frac{UVJ}{2 \cdot 4fL} \quad \text{where } Z_T = R_T + jX_T$$

The power delivered to R_L is

$$P_L = \frac{i IV J^2}{2 \cdot 4i?_r} \frac{4R_T R_L}{|Z_T + Z_J|^2} \quad \text{if } iv; \quad \frac{i}{2 \cdot 4i2} M_L = P_L$$

We will now show that the power available from the network equals that available from the source, i.e.,

$$\frac{W_J}{8fL} = \frac{W_J}{8fL}$$

We must then have

$$\frac{iV}{8R_s} = \frac{p_{in}}{8R_L} = \frac{IK}{8R_L} = \frac{iR^2 M_L p_L}{8R_L} = \frac{A}{Z_s}$$

which gives $M = M_L$, that is, the impedance mismatch at the output equals that at the input plane. The Thevenin resistance is given by

$$Z_s = \frac{Z_u + Z_s}{2}$$

We can express Z_T in the form

$$Z_T = \frac{(A + Z_s Z_u Z_f + Z_s^*)}{(Z_n + ZM Z_f i + Z_s)}$$

The numerator equals

$$(X_2 - X_n X_{22} - X_u X_s + j X_{22} R_s)(-j X_u - j X_s + S)$$

and has a real part which is readily found to be simply $X_f R_a$. Hence $R_T = X_{22}^* R_s / Z_n + Z_f$. We now get

$$\frac{|Y|}{8R_s} = p$$

Since the available power at the output plane is the same as that from the source, the impedance mismatch $M_L = M$.

If we have several lossless networks (reciprocal) between the source and Z_L as shown in Fig. 5.26c, we can repeat the analysis using Thevenin equivalent circuits at each stage to find that at any plane the available power and impedance mismatch are invariant quantities.

If the networks are not reciprocal, i.e., are linear active networks, the impedance mismatch is not invariant, e.g., an amplifier supports more output power than is fed in at the input.

The invariance of the impedance mismatch factor in a lossless network places an important constraint on interstage matching networks in a stage amplifier. The design of the interstage matching network must take this constraint into account.

In Fig. 5.27 we show a multistage microwave amplifier. The gains of the stages at microwave frequencies are often not absolutely stable, so it is not possible to use conjugate impedance matching at the output of each stage. For the first stage let Z_{jn} be the impedance looking into the stage.

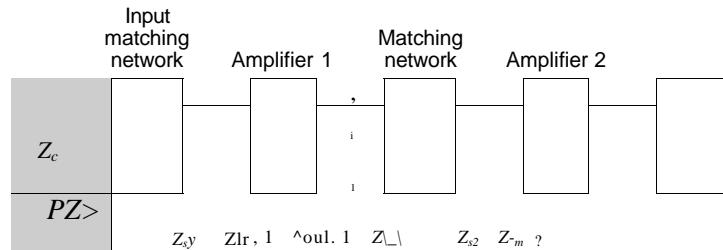


FIGURE 5.27
A multistage microwave amplifier.

first amplifier and let Z_{s1} be the source impedance presented to this stage by the input matching network (the networks are called matching networks even though they do not produce conjugate impedance matching). The input impedance mismatch is

$$M_x,$$

This same impedance mismatch must exist between the input transmission line and the input to the matching network. Hence on the input line we will have a reflection coefficient p that satisfies the constraint

$$1 - p^2 = M_x,$$

Thus $p = (1 - M_x)^{1/2}$ and the input VSWR will be

$$\text{VSWR} = \frac{1 + \sqrt{1 - M_x}}{1 - \sqrt{1 - M_x}} \quad (5.32)$$

Consequently, an input VSWR greater than unity is unavoidable when conjugate impedance matching cannot be used. A resultant problem in microwave amplifier design is to find an optimum load impedance that will result in an input impedance Z_{inA} that is close to the complex conjugate of the optimum source impedance Z_{s1} so as to minimize the input VSWR. The optimum source impedance is constrained by the requirement for a low noise figure for the input stage.

For the interstage network the matching network must transform the input impedance Z_{jn2} into the required load impedance Z_{l1} for stage 1 and simultaneously transform the output impedance Z_{oul1} of stage 1 into the required source impedance Z_{s2} for stage 2. The design of the matching network is constrained by the condition that the impedance mismatch be

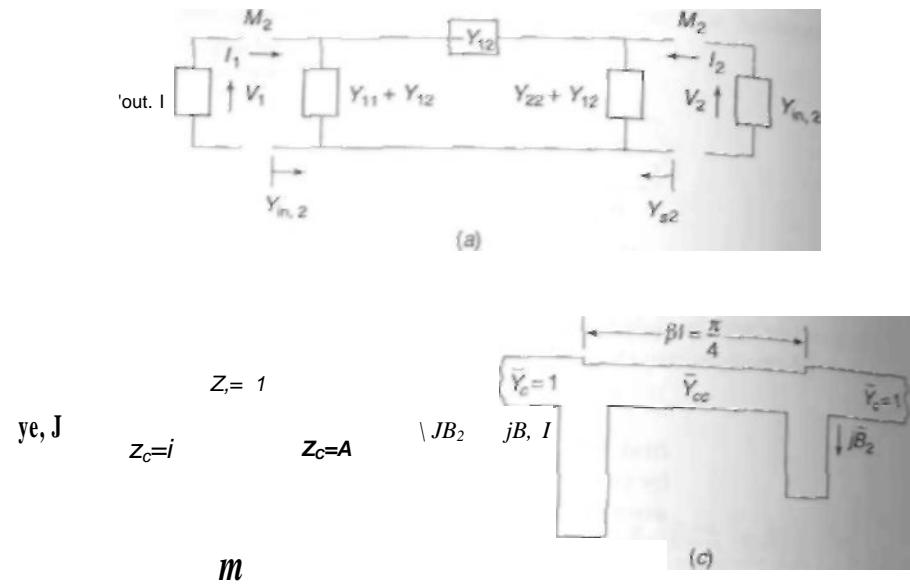


FIGURE 5.28

(a) A II matching network; (b) and (c) transmission-line matching networks.

constant; thus

$$\begin{array}{ll} 4 * z.i *_{out,1} & 4fli,..8-B.2 \\ I Z^+ + ZOUMI^2 & 12*,,, \quad s2* \end{array} \quad (5.33)$$

The design of microwave amplifiers is described in detail in Chap. 1. As part of the design procedure, the four impedances Z_{lb} , Z_{mlb} , Z^+ , Z_s^* are determined so that the impedance mismatch constraint given in (5.33) is satisfied. Hence we can assume that we know the admittances $Y_{out,i}$, $Y_{in,2}$, and Y_{s2} . In terms of these self-consistent quantities, we can determine the parameters of the U matching network illustrated in Fig. 5.28a. By conventional circuit analysis, we readily find that

$$y.i - ?u " \bullet \frac{y^2}{r!2} \\ '22 + *_{in,2}$$

$$* v9 ~ Y9.9 = \frac{y_{12}^2}{*11 + *_{out,1}}$$

For a lossless network $Y_u = jB_{lb}$, $Y_{12} = jB_{l2}$, and $Y_n = J^{Bz^2}$: After 9 number of algebraic steps, the following pair of equations

can be derived:

$$(G_{s2} + SL.)Bn - (G_{out}, + G_{IA})B_{22} = G_L B_{in,2} + B_L G_{m2} - B_{s2} G_{outA} - G_{s2} B_{uuIA} \quad (5.34a)$$

$$(B_{s2} + B_{in,2})B_n - (B_{IA} + B_{imIA})B_{22} = G_{s2} G_{oot}, - G_L G_{m2} + B_L B_{m,2} - B_{s2} B_{01111} \quad (5.34b)$$

After B_n and B_{22} have been found, we use

$$B_{\backslash 2} = BJK^*.i + (Bn + B_{uuIA})(B_{22} - B_{s2}) \quad (5.35)$$

to find B_{12} . The expressions for Y_u and Y_{s2} generate four equations which would generally not allow a FI network to be determined unless (5.33) was also satisfied.

Many matching networks can now be designed by relating their parameters to those of the TI network. Two possible transmission-line networks are shown in Figs. 5.286 and c. The parameters of the network in Fig. 5.28 are given by

$$\csc/\alpha = |S_{12}| \quad (5.36a)$$

$$jB_x = j(B_{u,} + \cot pi) \quad (5.36b)$$

$$jB_2 = j(B_{22} + \cot pi) \quad (5.36c)$$

where jB_l and jB_u are the stub susceptances. This network can only be used if $|Bi_2| > 1$ since $\csc pi > 1$. When $|Bi_2| < 1$ the network shown in Fig. 5.28c can be realized. Its parameters are given by

$$jB_u = j(B_u + B_{i2}) \quad (5.37a)$$

$$jB_2 = j(B_{22} + B_{i2}) \quad (5.37b)$$

$$T^* = B_{l2} \quad (5.37c)$$

where T^* is the characteristic admittance of the center one-quarter-wavelength-long section.

WAVEGUIDE REACTIVE ELEMENTS

In the place of transmission-line stubs, any other element that acts as a shunt susceptance may be used for the purpose of matching an arbitrary load impedance to a waveguide or transmission line. A number of such reactive elements for use in rectangular waveguides supporting the dominant TE_{10} propagating mode are described in this section.¹ The formulas given for the normalized susceptance of these elements are approximate

tDetailed

information on susceptance values and equivalent circuits are given in N. Marcuvitz (ed.), "Waveguide Handbook," McGraw-Hill Book Company, New York, 1951.

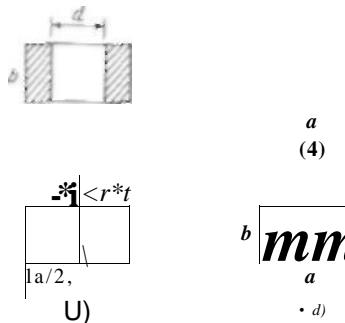


FIGURE 5.29
Shunt inductive elements, (a) Symmetrical d ($>$) asymmetrical diaphragm; (c) thin circular apertures.

ones, with accuracies of the order of 10 percent or better. The derivation of these formulas requires the detailed solution of boundary-value problems and is outside the scope of this text.

Shunt Inductive Elements

Figure 5.29 illustrates a number of rectangular waveguide elements that act as shunt inductive susceptances for the TE_{00} mode. These consist of thin metallic windows extending across the narrow dimension of the guide as in Figs. 5.29a and b, a very thin cylindrical post as in Fig. 5.29c, and a small circular aperture as in Fig. 5.29rf. When a TE_U mode is incident on any of these discontinuities, evanescent TE_{n0} modes are excited in order to provide a total field that will satisfy the required boundary condition of a vanishing tangential electric field on the obstacle. These nonpropagating modes store predominantly magnetic energy and give the obstacle its inductive characteristics.

Approximate values for the normalized inductive susceptance of these obstacles are:

For Fig. 5.29a.

$$\frac{B}{B_0} = \frac{2\pi}{pa} \cot^2 \frac{-d}{2a} \frac{I}{I_0} + \frac{ay_3 - 377}{4\pi^2} \frac{rrd}{a} \quad (5.38)$$

where $p = [\frac{2\pi}{\lambda} - (\frac{\pi}{\lambda})^2]^* \approx 1.86$ and $y_3 = [(3ir/a)^2 - A^2]^{1/2}$

For Fig. 5.29b,

$$\frac{B}{B_0} = \frac{2\pi}{3a} \cot^2 \frac{ird}{2a} \frac{I}{I_0} + \frac{\pi^2}{2a} \frac{r^2 d^2}{A^2} \quad (5.39)$$

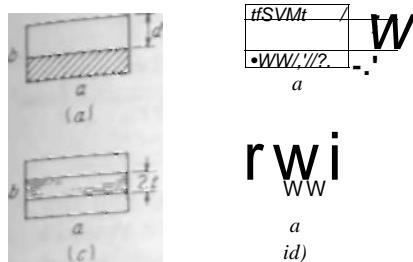


FIGURE 5.30
Shunt capacitive elements, (a) Asymmetrical capacitive diaphragm; (b) symmetrical diaphragm; (c) capacitive nid; (d) capacitive post.

For the thin inductive post of Fig. 5.29c,

$$4 - \ln \frac{a}{\pi t} - 1 + 2 \left(\frac{a}{\pi t} \right)^2 \sum_{n=3,5,\dots}^{\infty} \left(\frac{\pi}{a \gamma_n} - \frac{1}{n} \right) \sin^2 \frac{n \pi t}{a} \quad 15.40)$$

where $y_{nr} = [(n\pi/a)^2 - A5]^{1/2}$ and t is the post radius. For the small centered circular aperture of Fig. 5.29d,

$$B = \frac{Sab}{8fR} \quad (5.41)$$

Shunt Capacitive Elements

Typical shunt capacitive elements that may be used for matching purposes are illustrated in Fig. 5.30. Those consist of thin metal septa extending across the broad dimension of the guide to form capacitive diaphragms as in Figs. 5.30a and 6, a thin circular rod extending across the guide as in Fig. 5.30c, and a short thin circular post extending into the guide as in Fig. 5.30d. The post illustrated in Fig. 5.30a" behaves more like an *LC* series network connected across a transmission line. When the depth of penetration is between 0.76 and 0.96, it becomes resonant and acts almost like an ideal short circuit. For lengths greater than this resonant length, the post is equivalent to a shunt inductive susceptance. Actually, for a post of finite thickness, the equivalent circuit is a T network, but for small-diameter posts, the series elements in this T network are negligible (for post diameters less than about 0.05a).

Approximate expressions for the normalized susceptance of the obstacles illustrated in Fig. 5.30 are:

For the asymmetrical diaphragm of Fig. 5.30a,

$$\mathbf{S} = \mathbf{I}^\Delta \text{ In esc } \frac{\tau_d d}{26} + \frac{1}{|by_l|} \cos' \frac{-d}{26} \quad (5.42)$$

where $\Delta = [\kappa - (TT/O)^2]^{-1/2}$ and $y_l = [\sqrt{bf} - 0"]^{-1/2}$.

For the symmetrical diaphragm of Fig. 5.306,

$$S = \frac{vd}{2b} \ln \frac{I2\pi v}{b^2} + \frac{1}{4} \cos^4 \frac{\pi d}{2b} \quad (5)$$

where $y_2 = \{[2ir/b]^2 - ft^2\}^{1/2}$.

For the capacitive rod of Fig. 5.30c and the post illustrated in Fig. 5.30d, no simple approximate formulas are available. Analytical expressions for the T-network parameters for the capacitive rod are given by Leutheusser, but are not reproduced here.

The inductive and capacitive properties of the waveguide obstacles described above are determined by the electric and magnetic energy in the evanescent modes excited by the obstacle. For example, the inductive obstacles discussed only excite TE evanescent modes and these modes store more magnetic energy than electric energy. The evanescent modes are nonpropagating modes so they represent the local field in the vicinity of the obstacles. The evanescent modes are excited with amplitudes such that the combination of all of the evanescent modes with the incident, reflected, and transmitted dominant mode, produces a total field with a zero tangential electric field at the surface of the obstacle.

There is no way to assign a unique inductance or capacitance to these obstacles since there is no unique way to define the voltage, current, or characteristic impedance for a waveguide. Nevertheless, these obstacles do produce a reflected wave with a reflection coefficient that is correctly determined in terms of the equivalent normalized reactance and susceptance of the obstacle as connected into the equivalent-transmission-line model of the waveguide.

Waveguide Stub Tuners

An approximate equivalent of a single-stub matching network is the screw tuner illustrated in Fig. 5.31a. This consists of a variable-depth screw mounted on a sliding carriage free to move longitudinally along the guide over a distance of at least a half guide wavelength. The screw penetrates into the guide through a centered narrow slot in the broad wall. This slot is cut along the current flow lines so that it has a negligible perturbing effect on the internal field. Since the position of the screw is adjustable over at least a half guide wavelength, its penetration need to be so great that it will behave as an inductive element; this can be obtained with a shunt capacitive susceptance in all cases, of single-stub-matching theory will verify.

tOp. cit., chap. 2. This text contains many excellent and instructive derivations of circuit parameters for a variety of waveguide structures.

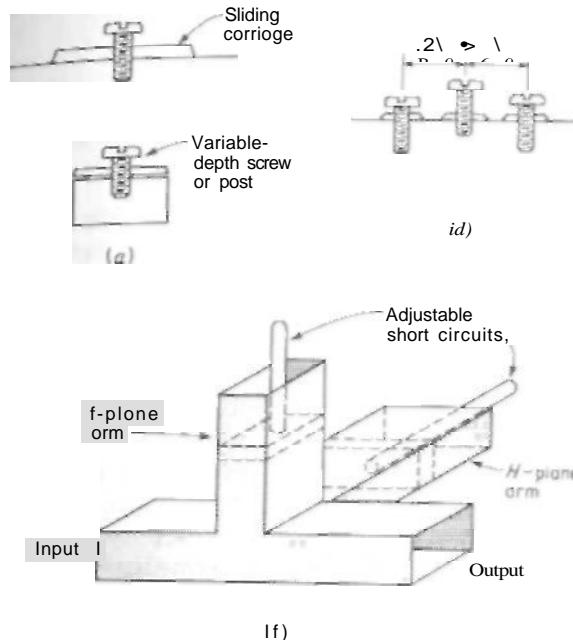


FIGURE 5.31
Waveguide tuners, (a) Sliding-screw tuner; (b) triple-screw tuner; (c) E-H tuner.

Three variable-depth screws spaced a fixed distance of about $3A_g/8$ apart as in Fig. 5.31b are essentially equivalent to a triple-stub tuner and can match a large variety of loads even though the range of susceptance values obtainable from a single screw is limited.

Circuits that are physically more like the actual short-circuited transmission-line stub are also possible. Figure 5.31c illustrates combined E-plane and f-plane stubs, referred to as an E-H tuner. The positions of the sliding short circuits in the E- and H-plane arms are variable, so that a wide range of load impedances may be matched. The equivalent circuit of either an E-plane or H-plane junction is, however, much more elaborate than a simple shunt- or series-connected transmission line because the junctions are of the order of a wavelength in size and hence produce a very complicated field structure in their vicinity. Nevertheless, since no power now is possible through the arms terminated in the short circuits, these do still provide adjustable reactance elements that may be used for matching purposes.

WAVE TRANSFORMERS

quarter-wave transformers are primarily used as intermediate matching sections when it is desired to connect two waveguiding systems of different characteristic impedance. Examples are the connection of two transmission lines with different characteristic impedances, connection of an empty wave-

$$Z_C = Z >$$

$$Z_r = Z, \quad Z_L$$

V_{in}

~?

FIGURE 5.32

H A quarter-wave transformer.

guide to a waveguide partially or completely filled with dielectric of two guides of different width, height, or both, and the match^{11**1,011} of a dielectric medium such as a microwave lens to free space. If a match^{"8} of a narrow band of frequencies suffices, a single-section transformer ^{<^*x*} used. To obtain a good match over a broad band of frequencies, two three or even more intermediate quarter-wave sections are commonly used. The optimum design of such multisection quarter-wave transformers is * presented in this section.

The essential principle involved in a quarter-wave transformer is readily explained by considering the problem of matching a transmission line of characteristic impedance Z , to a pure resistive load impedance Z_L , as illustrated in Fig. 5.32. If an intermediate section of transmission line with a characteristic impedance Z_x , and a quarter wavelength long is connected between the main line and the load, the effective load impedance presented to the main line is

$$\frac{Z_L + jZ_x \tan(0A/4)}{Z_L + jZ_x \tan(\pi/4)} = \frac{Z_L}{Z_L}$$

If Z_L is chosen equal to $\sqrt{Z_x Z_L}$, then $Z = Z$, and the load is matched to the main line. In other words, the intermediate section of transmission line of length $A/4$ transforms the load impedance Z_L into an impedance Z , hence acts as an ideal transformer of turns ratio $\sqrt{Z_x/Z_L}$. A perfect match is obtained only at that frequency for which the transformer is a quarter wavelength (or $\lambda/4$) long. /

Let θ be the electrical length of the transformer at the frequency ω that is, $p\theta = \pi/2$, where the phase constant p has been a function of frequency. For a TEM wave in an air-filled line, $p = 2\pi/\lambda$ at any frequency the input impedance presented to the main line is

$$\frac{Z_L + jZ_x t}{Z_L + jZ_x t} \quad (5.45)$$

where $t = \tan \theta = \tan \pi/2$. Consequently, the reflection coefficient

$$\frac{\frac{Z_L - Z_x}{Z_L + Z_x}}{\frac{Z_L + Z_x}{Z_L + Z_x} + j\frac{Z_x}{Z_L}} = \frac{Z_x(Z_L - Z_x) + j(Z_L + Z_x)Z_x}{Z_x(Z_L + Z_x) + j(Z_L + Z_x)Z_x} \quad (5.46)$$

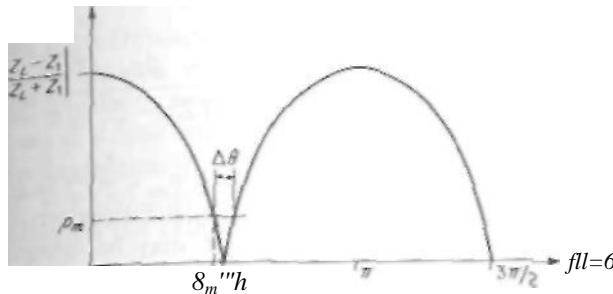


FIGURE 5.33
Bandwidth characteristic
For a single-section quarter-
wave transformer.

The latter form is obtained by using the relation $Z_l = Z_x Z_s$. The magnitude of F, denoted by p , is readily evaluated and is given by

$$P = \frac{|Z_L - Z_x|}{\sqrt{[(Z_L + Z_x)^2 + U^2 Z_x Z_L]}} \quad (5.47)$$

$$\frac{1}{1 + 2V^2 Z T} \sec \theta$$

For θ near $\pi/2$, this equation is well approximated by

$$P = \frac{|Z_L - Z_x|}{2Z^2} \sec \theta \quad (5.48)$$

A plot of p versus θ is given in Fig. 5.33, and this is essentially a plot of p versus frequency. The variation of p with frequency, or θ , is periodic because of the periodic variation of the input impedance with frequency; i.e., the impedance repeats its value every time the electrical length of the transformer changes by πr . If p_m is the maximum value of reflection coefficient that can be tolerated, the useful bandwidth provided by the transformer is that corresponding to the range $A\theta$ in Fig. 5.33. Because of the rapidly increasing values of p on either side of $\theta = \pi/2$, the useful bandwidth is small. The value of θ at the edge of the useful passband may be found from (5.47) by equating p to p_m ; thus

$$\theta_m = \cos^{-1} \frac{2p_m J Z |Z_L|}{(2L - 2I > I/W)} \quad (5.49)$$

In the case of a TEM wave, $\theta = \pi/3 = (f/f_0 Y ir/2)$, where f_0 is the frequency for which $\theta = \pi/2$. In this case the bandwidth is given by

$$A/I = 2(f_0 - f_m) = 2(\theta_0 - X)$$

and the fractional bandwidth is given by

$$\frac{A''}{\omega_0} = 2 - \frac{4}{\pi^2} \cos^{-1} (Z_1 / Z_2)^{1/2} \quad (5.50)$$

where that solution of (5.49) that gives $\omega_m < \pi/2$ is to be chosen.

Although there are a number of instances when the bandwidth provided by a single-section transformer may be adequate, there are a number of situations in which much greater bandwidths must be obtained. The required increase in bandwidth can be obtained by using quarter-wave transformers. The approximate theory of these multisection transformers is discussed first, in order to develop a theory that has application in the design of other microwave devices, such as directivity couplers and antenna arrays. This is followed by a discussion and presentation of results obtainable from a more exact analysis.

It should be noted that in the previous discussion it was assumed that the characteristic impedances Z_1 and Z_2 were independent of frequency. For transmission lines this is a good approximation, but for waveguides the wave impedance varies with frequency, and this complicates the analysis considerably. In addition, for both transmission lines and waveguides, there are reactive fields excited at the junctions of the different sections, brought about because of the change in geometrical cross section necessary to achieve the required characteristic impedances. These junction effects can often be represented by a pure shunt susceptance at each junction. The susceptive elements will also vary the performance of any practical transformer from the predicted performance based on an ideal model when junction effects are neglected. In spite of all these limitations, only the theory for ideal transformers is developed here; i.e., junction effects are neglected. Thus the theory given will only be indicative of the performance that can be obtained in the nonideal case.

S. B. Cohn. Optimum Design of Stepped Transmission Line Transformers. *IRE Trans. on Antennas and Propagation*, MTT-3, pp. 16-21, April, 1955. This paper presents an approximate theory for stepped transmission line transformers, together with a method of accounting for the reactances introduced. For typical application to waveguide transformers, see: R. E. Collin and J. Brown, The Design of Quarter-Wave Matching Layers, *Proc. IEE*, vol. 103, pt. C, pp. 153-158, March, 1956.

L. Young, Optimum Quarter-Wave Transformers, *IRE Trans.*, vol. MTT-10, September, 1960; also Inhomogeneous Quarter-Wave Transformers of Two Sections, *IRE Trans.*, vol. MTT-11, pp. 645-649, November, 1960.

E. S. Hensperger, Broad-Band Stepped Transformers from Rectangular Waveguide, *IRE Trans.*, vol. MTT-6, pp. 311-314, July, 1958.

ORY OF SMALL REFLECTIONS

As a preliminary to the approximate analysis of multisection quarter-wave transformers, some results pertaining to the overall reflection coefficient arising from several small reflecting obstacles are required. Consider the case of a load impedance Z_L connected to a transmission line of characteristic impedance Z , through an intermediate section of line of electrical length $\delta l = 0$ and characteristic impedance Z_2 , as illustrated in Fig. 5.34. For each junction the reflection and transmission coefficients are

$$\begin{matrix} Z_r \sim Z, \\ ?i \quad \mathbf{z}_a + \mathbf{z}_r, \end{matrix} \quad \mathbf{r}_2 = -\mathbf{r},$$

$$T_{2l} = \mathbf{i} + \mathbf{r} = \frac{2Z}{\mathbf{z}_a + \mathbf{z}_r}, \quad T_{v2} = \mathbf{i} + \mathbf{r}_2 = \frac{2Z}{Z_2}$$

$$\begin{matrix} Z_L - Z, \\ \mathbf{r}_8 = \mathbf{z}_L + \mathbf{z}_2 \end{matrix}$$

A wave of unit amplitude is incident, and the total reflected wave has a complex amplitude V equal to the total reflection coefficient. When the incident wave strikes the first junction, a partial reflected wave of amplitude T_1 is produced. A transmitted wave of amplitude T_{2l} is then incident on the second junction. A portion of this is reflected to give a wave of amplitude $r_3 r_{21} e^{-j\delta l}$ incident from the right on the first junction. A portion $T_v 2 T_{2l} r_3 e^{-2J\delta l}$ is transmitted, and a portion $I^2 / r^2 T^2 e^{2J\delta l}$ is reflected back toward Z_L . Figure 5.35 illustrates the first few of the infinite number of multiply reflected waves that occur. The total reflected wave of amplitude 1 is the sum of all the partial waves transmitted past the first junction toward the left. This sum is given by

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_1 + T_{12} T_{2l} r_3 e^{-2J\delta l} + r_{12} r_{21} r_3 r_{21} e^{-4J\delta l} + \dots \\ &= \mathbf{F}_1 + \sum_{n=0}^{\infty} T_{12} T_{2l} r_3 e^{-W_n} \mathbf{r}_{21}^n T_3^n e^{4nJ\delta l} \end{aligned}$$

$0l = 6$

$2l - r_1 - 5 C_r$

FIGURE 5.34
A microwave circuit with two reflecting junctions.

$$1^{\wedge}r, * - * * - \\ r^*r, r_l r l^* - * - \\ r_2, r_2 r_5 f - ^{\wedge} \bullet$$

$$/\bullet_a r |r|^{*-w}$$

FIGURE 5.35

Multiple reflection of waves far from two reflecting junction*

This geometric series is readily summed to give [note that $F_{n, o, r}^x \ll (1 - r)^{-1}$]

$$r = i \backslash \frac{T_{12} T_{21} r_3 e^{-W}}{i - r_2 r_3 e^{-2\Lambda}}$$

Replacing T_{12} by $1 + f_2 = 1 - I^\wedge$ and T_{21} by $1 + I^\wedge$ gives

$$r = \frac{r, + r_3 e^{-270}}{1 + r^\wedge e^{-1\Lambda}} \quad (5.51)$$

If $|r_j$ and $|T_3|$ are both small compared with unity, an excellent approximation to T is

$$r = r, - r_3 e^{-2\Lambda} \quad (5.52)$$

This result states that, for small reflections, the resultant reflection coefficient is just that obtained by taking only first-order reflections into account. This is the result that will be used to obtain a first-order theory of multisection quarter-wave transformers. As an indication of the accuracy of the approximate formula, note that if $I_f, I = |T_3| = 0.2$, the error does not exceed 4 percent.

5.11 APPROXIMATE THEORY FOR MULTISECTION QUARTER-WAVE TRANSFORMERS

Figure 5.36 illustrates an AT-section quarter-wave transformer. At the junction the reflection coefficient is

$$\frac{z_1 z_0}{z_{I+Z_0}} = P_0 \quad (D.530)$$

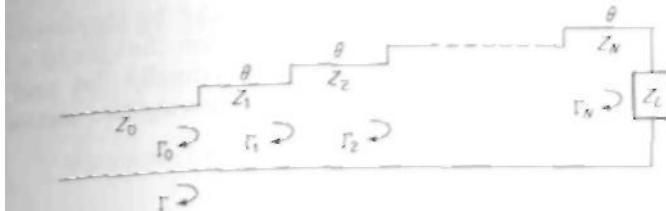


FIGURE 5.36
A multisection quarter-wave transformer.

Similarly, at the n th junction, the reflection coefficient is

$$\frac{rs}{Z_{n+l} + Z_n} = P_n \quad (5.53b)$$

The last reflection coefficient is

$$r_{nn} = \frac{Z_n - Z_L}{Z_L + Z_n} = p_n \quad (5.53c)$$

Note that Z_0 is a characteristic impedance, and not necessarily equal to $(f_0/e_0)^{1/2}$ here. Each section has the same electrical length $\pi f l = \pi/2$, and will be a quarter wave long at the matching frequency f_0 . The load Z_L is assumed to be a pure resistance, and may be greater or smaller than Z_0 . In this analysis it is chosen greater, so that all $p_n = p_{nn}$, where p_{nn} is the magnitude of r_{nn} . If Z_n is smaller than Z_{nn} , all p_n are negative real numbers and the only modification required in the theory is replacing all p_{nn} by $-p_{nn}$.

For a first approximation the total reflection coefficient is the sum of the first-order reflected waves only. This is given by

$$r = p_0 + p_1 e^{-2j\pi f l} + p_2 e^{-4j\pi f l} + p_N e^{-2j\pi f N l} \quad (5.54)$$

where $e^{-2jn\pi f l}$ accounts for the phase retardation introduced because of the different distances the various partial waves must travel.

At this point it is expedient to assume that the transformer is symmetrical, so that $p_0 = p_N$, $p_1 = p_{N-1}$, $p_2 = p_{N-2}$, etc. In this case (5.54) becomes

$$T = e^{-j\pi f N l} [p_0 (e^{j\pi f l} + e^{-j\pi f l}) + p_1 (e^{j\pi f l} + e^{-j\pi f l}) + \dots] \quad (5.55)$$

where the last term is $p_{(N-1)/2} (e^{j\pi f l} + e^{-j\pi f l})$ for N odd and $p_{N/2}$ for N even.

It is thus seen that for a symmetrical transformer the reflection coefficient is given by a Fourier cosine series:

$$T = 2e^{j\pi f N l} [p_0 \cos(N\pi f l) + p_1 \cos((N-2)\pi f l) + \dots + p_n \cos((N-2n)\pi f l) + \dots] \quad (5.56)$$

The last term $\propto P(N-i)/2 \cos(i\pi f l)$ for N odd and $\propto p_{N/2}$ for N even. It will be apparent that by a proper choice of the reflection coefficients \propto and hence the Z_{nn} , a variety of passband characteristics can be obtained. The series is a cosine series, the periodic function that it defines is

periodic over the interval - corresponding to the frequency ω which the length of each transformer section changes by a half wavelength. The specification of $\Delta\omega$ to obtain a maximally flat and an even passband characteristic is given in the following two sections.

5.12 BINOMIAL TRANSFORMER

A maximally flat passband characteristic is obtained if $p \triangleq |f|$ and the $N-1$ derivatives with respect to frequency (or ft) vanish at the match frequency ω_m , where $ft \sim TT/2$. Such a characteristic is obtained if we choose

for which

$$p = |\text{ri} = |A2^{\frac{N}{2}}(\cos\alpha)^{\frac{N}{2}}| \quad (5.58)$$

When $\alpha = 0$ or π , we have $V = (Z_L - Z_0)/(Z_L + Z_0)$, and from (5.57a) we obtain $T = A2^{\frac{N}{2}}$. Thus the constant A is given by

$$A = 2 \sqrt{\frac{C^*}{Z_0}} \quad (5.58)$$

However, if we use the theory of small reflections, then the constant A should be chosen in a different way, which we will explain shortly.

Expanding (5.57a) by the binomial expansion gives

$$r = 2^{-\frac{N}{2}} \sim \sim d - e^{-W} = 2^{-\frac{N}{2}} | \wedge \frac{C}{2} > \dots \quad (5.59)$$

where the binomial coefficients are given by

$$\dots = \frac{N(N-1)(N-2)\dots(2N+L)}{(2N+L)!} (iV)^\frac{N}{2} \dots \quad (5.60)$$

Note that $C^* = C\frac{N}{2}$, $C_m^v = 1$, $C_m^A = N = C\frac{N}{2}$ etc. Comparing with (5.54) shows that we must choose

since $C_m^v = C_m^A$.

To obtain a simple solution for the characteristic impedances it is convenient to make a further approximation. Since we have already indicated that all p_n are to be small, we can use the result

$$Z_{m+1} = Z_m + \frac{Z_p}{2}$$

Thus we have

$$\ln \frac{Z_L}{Z_0} = 2 - \frac{\lambda^2}{4} \ln \left| \frac{Z_L + Z_0}{Z_L - Z_0} \right| \quad (5.62)$$

where we have also used the approximation

$$Z_0 = Z_h + Z_n = S | Z_L + Z_0 | \quad Z_L + Z_0,$$

When we use the theory of small reflections (5.55) gives, for $\theta = 0$,

$$H_O = P_0 + P_1 + P_2 + \dots + \frac{1}{n!} v$$

When we use (5.62) to evaluate the characteristic impedances, and also use $A = 2^{r(N+l)} M Z_L / Z_0$, we obtain

$$2^N A = P(0) = - \ln \frac{Z_0}{Z_1} + \ln \frac{Z_1}{Z_2} + \dots + \ln \frac{Z_N}{Z_0}$$

$$= \frac{1}{2} \ln \frac{Z_L}{Z_0} - \frac{1}{2} \ln \frac{Z_0}{Z_L} = \frac{1}{2} \ln \frac{Z_L}{Z_0}$$

Thus the approximations introduced above for finding the characteristic impedances should be used along with (5.63) to evaluate H_O . This will make the quarter-wave transformer designs using the theory of small reflections self-consistent. In place of (5.58) we use the expression given above for A .

Equation (5.62) gives the solution for the logarithm of the impedances, and since these are proportional to the binomial coefficients, the transformer is called a binomial transformer. Since the theory is approximate, the range of Z_L is restricted to about

$$0.5Z_0 < Z_L < 2Z_0$$

for accurate results.

As an example, consider a two-section transformer. From (5.62) obtain

$$\ln \frac{Z_2}{Z_0} = \frac{1}{4} \ln \frac{Z_L}{T} \quad \text{or} \quad Z_2 = Z_0 \cdot Z_L / 4$$

and

$$\ln \frac{Z_1}{Z_0} = \frac{1}{2} \ln \frac{Z_L}{Z_0} \quad \text{or} \quad Z_1 = Z_0 \cdot Z_L / \sqrt{2}$$

since $C_0 = 1$ and $Cf = 2$. Although the approximate theory was used, it turns out that the above values for Z_1 and Z_2 for the special case of the

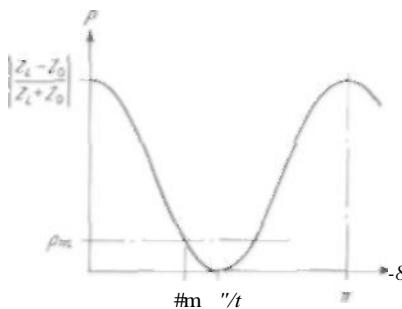


FIGURE 5.37
Passband characteristic for a maximally flat transformer.

two-section transformer are the correct nonapproximate solutions a res 1 that gives an indication of the accuracy of the approximate theory.

The type of passband characteristic obtained with a maximally flat transformer is illustrated in Fig. 5.37. Let p_m be the maximum value of p that can be tolerated. The angle θ_m that gives $p = p_m$ is given by

$$\cos^{-1} \frac{2p_m}{m(Z_L/Z_0)} = \theta_m \quad (5.64)$$

as obtained from (5.576). In the case of transmission-line sections, $6 \text{irf}/2f_0$, and hence the fractional bandwidth is given by

$$\frac{\Delta f}{f_0} = \frac{4}{\pi} \cos^{-1} \left| \frac{2p_m}{m(Z_L/Z_0)} \right| \quad (5.65)$$

since $\theta_m = \text{irf}_m/2f_0$. Note that in (5.65) the solution to the inverse cosine function is chosen so that $\theta_m < \pi/2$. By comparing Figs. 5.33 and 5.37 it is clear that a multisection maximally flat transformer can provide a much greater useful bandwidth than a single-section transformer.

5.13 CHEBYSHEV TRANSFORMER

Instead of a maximally flat passband characteristic, an equally useful characteristic is one that may permit p to vary between 0 and p_m , in an oscillatory manner over the passband. A transformer designed to yield an equal-ripple characteristic as illustrated in Fig. 5.38 is of this type. It provides a considerable increase in bandwidth over the binomial former design. The equal-ripple characteristic is obtained by using Chebyshev polynomials, and hence the name "Chebyshev transformer". It is possible to have p vanish at as many frequencies in the passband as there are transformer sections. Chebyshev polynomials may be used in the design, it is necessary to consider the basic properties of these polynomials first.

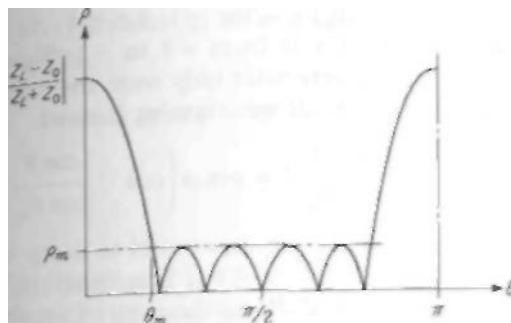


FIGURE 5.38
Equal-ripple characteristic obtained from a Chebyshev transformer.

The Chebyshev polynomial of degree n , denoted by $T_n(x)$, is an n -th-degree polynomial in x . The first four polynomials and the recurrence relation are

$$T_0(x) = 1$$

$$T_1(x) = 2x^0 - 1$$

$$T_2(x) = 4x^2 - 3x$$

$$T_3(x) = 8x^4 - 8x^2 + 1$$

$$T_n(x) = 2xT_{n-1} - T_{n-2}$$

The polynomials T_n oscillate between ± 1 for x in the range $|x| < 1$ and increase in magnitude indefinitely for x outside this range. Figure 5.39 gives a sketch of the first four polynomials. If x is replaced by $\cos \theta$, we have

$$r_{\infty}(\cos \theta) = \cos n\theta \quad (5.66)$$

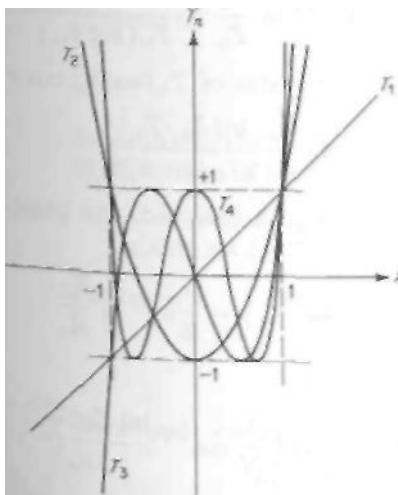


FIGURE 5.39
Chebyshev polynomials.

which clearly shows that $|T_I| < 1$ for $-1 < x < 1$. As θ varies from 0 to π , we wish to have the equal-ripple characteristic only over the range $0 \leq x \leq p$. We cannot use $T_n(\theta)$ directly. If we consider instead

$$\frac{\cos \theta}{\cos \theta_m} = \cos n \frac{1}{J} \cos \theta \frac{\cos \theta}{\cos \theta_m J}$$

we see that the argument will become equal to unity when $\theta = Q$ be less than unity for $\theta_m < \theta < \pi$. This function will therefore contain the equal-ripple oscillations of $T_n(\theta)$ to the desired passband.

The function given in (5.67) is an n th-degree polynomial in variable $\cos \theta / \cos \theta_m$. Since $(\cos \theta)^n$ can be expanded into a series of terms such as $\cos \theta, \cos 2\theta, \dots, \cos n\theta$, it follows that (5.67) is a series of the form (5.56). Hence we may choose

$$\begin{aligned} p &= 2e^{-jNB} [P_0 \cos NO + P_1 \cos(N-2)\theta + \dots \\ &\quad + P_{n-1} \cos(N-2n)\theta + \dots] \\ &= Ae^{-jNB} (\sec \theta_m \cos \theta) \end{aligned} \quad (5.68)$$

where A is a constant to be determined. When $\theta = 0$, we have

$$r = \frac{Z_r - Z_r}{Z_r + Z_r} = \frac{1}{A \sec \theta_m} \sim \frac{Z_r}{A \sec \theta_m}$$

and so

$$A = \frac{\ln(Z_r / Z_0)}{2T_A(\sec \theta_m)}$$

Consequently, we have

$$r = \frac{L - j \tan \theta}{Z_0} \sim \frac{j \ln(\sec \theta_m)}{2T_A(\sec \theta_m)} \quad (5.69)$$

In the passband the maximum value of $|T_{jv}(\sec \theta, \cos \theta)|$ is unity, and hence

$$P_m = \frac{j \ln(Z_r / Z_0)}{2T_w(\sec \theta_m)} \quad (5.70a)$$

If the passband, and hence θ_m , is specified, the passband tolerance P_m is fixed, and vice versa. From (5.70a) we obtain

$$T_w(\sec \theta_m) = \frac{1}{-p} \ln \frac{Z_r}{Z_0}$$

or by using (5.67) for $\cos \theta = 1$,

$$\sec B_m = \cos \left[-\frac{\ln(Z_r / Z_0)}{2p} \right], \quad (5.70b)$$

which gives θ_m in terms of the passband tolerance on p , that is, P_m .

In order to solve (5.68) for the unknown $p_{,,}$, we need the following results:

$$\begin{aligned}
 (\cos 0)^n &= 2^{-n} e^{-n^2/(l_m + e^{2j\theta})} = 2^{-n} e^{-n^2/l} \sum_{m=0}^{\infty} C_m e^{j2mt} \\
 &= 2^{-n+1} [C_0 \cos \theta + C_1 \cos(\theta - 2\pi) + \dots \\
 &\quad + C_n \cos(n - 2m)\theta + \dots] \quad (5.71)
 \end{aligned}$$

The last term in (5.71) is kC_n^n for n even and $C_{(n-1)/2} \cos \theta$ for n odd. Using (5.71) and the earlier expression for $T_{,,}(x)$, we can obtain the following:

$$T_x(\sec \theta_m \cos \theta) = \sec B_m \cos \theta \quad (5.72a)$$

$$r_2(\sec \theta_m \cos \theta) = 2(\sec \theta_m \cos \theta) = \sec^2 \theta_m + \cos 2\theta_m - 1 \quad (5.72b)$$

$$T_3(\sec \theta_m \cos \theta) = \sec^3 \theta_m (\cos 3\theta_m + 3\cos \theta_m) - 3 \sec \theta_m \cos \theta \quad (5.72c)$$

$$T_4(\sec \theta_m \cos \theta) = \sec^4 \theta_m (\cos 4\theta_m + 4\cos 2\theta_m + 3) - 4 \sec^2 \theta_m (\cos 2\theta_m + 1) \quad (5.72d)$$

These results are sufficient for designing transformers up to four sections in length. A greater number of sections would rarely be required in practice.

Example 5.2 Design of a two-section Chebyshev transformer. As an example, consider the design of a two-section transformer to match a line with $Z_0 = 1$ to a line or load with $Z_L = 2$. Let the maximum tolerable value of p be $P_m = 0.05$. Using (5.70a), we obtain

$$T_2(\sec \theta_m) = 2 \sec^2 \theta_m - 1 = \frac{2 \sec^2 \theta_m - 1}{3(0.05)} = 6.67$$

and hence $\sec \theta_{,,} = 1.96$, and $\theta_{,,} = 1.04$. Thus the fractional bandwidth that is obtained is

$$\frac{A_0}{TT/2} = \frac{A_0}{f_0} = \frac{4/\omega}{7T/2} = \frac{1.04}{1.04} = 0.675$$

From (5.68), (5.69), and (5.72c), we obtain [refer to the remarks following (5.56) as regards the last term in the cosine series for p]

$$\begin{aligned}
 2p_0 \cos 2\theta_m + p_{,,} &= r_2(\sec \theta_m \cos \theta) \\
 &= p_{,,} \sec^2 \theta_m \cos 2\theta_m + i \geq m \sec^2 \theta_m - 1
 \end{aligned}$$

and hence

$$p_0 = i p_m \sec^2 \theta_m = p_2 = 0.099$$

$$P_{,,} = p_m (\sec^2 \theta_m - 1) = 0.148$$

The impedances Z_1 and Z_2 are given by

$$Z_1 = e^{2j\theta_m} Z_0 = 1.219 \quad Z_2 = e^{2j\theta_m} Z_L = 1.639$$

A plot of the passband characteristic is given in Fig. 5.40. As a check on the

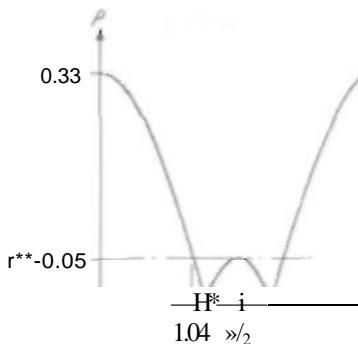


FIGURE 5.40

Passband characteristic for a two-section transformer with $r \approx 0.05$, $P \wedge T \wedge$

accuracy, we calculated Z_{in} for $H = 0$ using

$$Z_{in} = \frac{Z^2}{ip^2} = 1.1063$$

From Z_m we obtain

$$P = \frac{Z}{1+r} = 0.05047$$

which is within 1 percent of the design value.

*5.14 CHEBYSHEV TRANSFORMER (EXACT RESULTS)

An exact theory for a multisection transformer having an equal-ripple passband characteristic has also been developed (see the references at the end of this chapter). Since the analysis is rather long, only the final results for the two- and three-section transformers are given here.

In the exact theory of multisection ideal transformers, it is convenient to introduce the power loss ratio P_R , which is defined as the reflected power (incident power) divided by the actual power delivered to the load. If the incident power is P_i , the reflected power is $p^2 P_i$, and the power to the load is $(1 - p^2)P_i$. Hence

$$\Pr.R = \frac{P_i}{(1-p^2)P_i} = \frac{1}{1-p^2} \quad (5.730)$$

and

$$P = \frac{P_R}{LR} = \frac{1}{LR} \quad ,5-73*1$$

If T is the overall transmission coefficient, then $|T| \approx 1 \sim p^2$ obtained.

For any transformer an expression for Z_{in} is readily obtained from this P , and hence P_R , can be computed. When this is

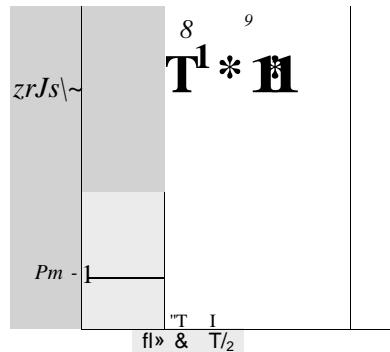


FIGURE 5.41
Passband characteristic for a two-section Chebychev transformer.

that P_{LR} can be expressed in the form

$$P_{LR} = 1 + Q_2^*(\cos \theta) \quad (5.74)$$

where $Q_{2N}(\cos \theta, f)$ is an even polynomial of degree $2AT$ in $\cos \theta$, with coefficients that are functions of the various impedances $Z_{..}$. To obtain an equal-ripple characteristic, P_{LR} is now specified to be

$$P_{LR} = 1 + /e^{2T}(\sec 0, \cos 0) \quad (5.75)$$

where k^2 is the passband tolerance on P_{LR} ; that is, the maximum value of P_{LR} in the passband is $1 + k^2$, since T^1 has a maximum value of unity. By equating (5.74) and (5.75), algebraic equations that can be solved for the various characteristic impedances are obtained.

Figure 5.41 is a plot of p versus θ for a two-section transformer. For this transformer

$$P_{LR} = 1 + \frac{|Z_L - Z_0|}{*Z_L Z_a} \frac{(\sec^2 \theta_t \cos^2 \theta - 1)}{\tan^4 \theta_z} \quad (5.76)$$

where θ_z is the value of θ at the lower zero where p vanishes. The maximum value of P_{LR} in the passband is

$$(Z_L \sim Z_0)^A$$

and hence

$$1 + k^2 \quad (5.77)$$

where $k^2 = \cot^4 \theta_z (Z_L^A - Z_0^2)^2 / 4Z_L Z_0$. The required values of Z_1 and Z_2

are given by

$$\frac{1}{Z^2} - \frac{1}{Z^a} \frac{(Z_L - Z_0 Y)}{AZl \tan^4 B_t} + \frac{1}{Z^2} \approx 2 \sim \tan^2 i$$

$$Z = \frac{1}{Z}$$

(5.786)

The value of $B_{,,}$ is given by

$$\theta_m = \cos^{-1} \sqrt{2} \cos 0,$$

and

$$\frac{\pm f}{f_u} = 2 \frac{4}{W} \cos^{-1} \sqrt{2} \cos \theta_m$$

(5.79a)

(5.796)

provided $2 A \ll 1 = A \cdot 1 / f_u$, [if not, (5.796) gives $2 A f / f_u$]. If the bandwidth is specified, then B_2 , and from (5.79a) $B_{,,}$, are fixed. Equation (5.77) then specifies $\theta_{,,}$. On the other hand, if $\theta_{,,}$ is given, the bandwidth is determined.

In the limit as B , approaches $\pi/2$, the two zeros of p coalesce to give a maximally flat transformer. From (5.78) it is found that, for this case (compare with the approximate theory),

$$Z_x = Z I Z^\Delta \quad (5.80a)$$

$$Z_2 = 2^{1/2} v < \quad (5.80b)$$

For the maximally flat transformer, the value of $\theta_{,,}$ at the point where $P = P_m$ is given by

$$B_m = \cos^{-1} \cot \theta_{,,} \quad (5.81)$$

where $\theta_{,,}$ is the previously defined quantity for the Chebyshev transformer. Equations (5.79) and (5.81) provide a comparison of the relative bandwidth obtainable from the Chebyshev transformer and the maximally flat transformer. This comparison is illustrated in Fig. 5.42 for $N = 2$ and j shows that the Chebyshev transformer can give bandwidths that are considerably greater for the same maximum tolerable value p_m .

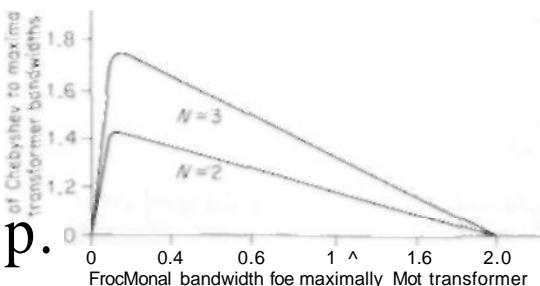


FIGURE 5.42 $\frac{P_m}{P_{max}}$ vs. f/f_m for Chebyshev and n^{**} transformers.

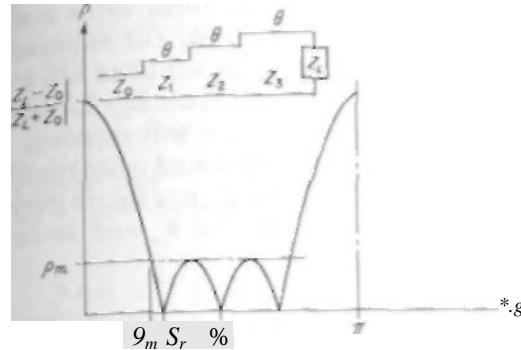


FIGURE 5.43
Passband characteristic for three-section Chebyshev transformer.

Figure 5.43 illustrates the passband characteristic for a three-section Chebyshev transformer. The power loss ratio is given by

$$\frac{P_m}{Z_L Z_0} = \frac{(Z_L - Z_0 f (\sec^2 \theta_0, \cos^2 \theta_0 - 1)^2 \cos^2 \theta)}{\tan^4 f L} \quad (5.82)$$

The passband tolerance A^2 is given by

$$A^2 = \frac{(Z_L Z_0)^2 I}{4 Z_L Z_0} \frac{2 \cos \theta_0}{I^{3/3} \tan^2 \theta_0} \quad (5.83)$$

from which P_m may be found by using (5.77). Again the general result that specifying k^2 determines the bandwidth, and vice versa, holds. The value of θ_m is given by

$$\theta_m = \cos^{-1} \frac{1}{3} \cos \theta_0 \quad (5.84a)$$

and for transmission lines for which $A f/f_m = 2 A f / f_m$,

$$\frac{Af}{f_m} = \frac{2(71/2 - ti_m)}{\pi/2} = 2 \frac{4}{\pi} \cos \frac{\sqrt{7}}{\sqrt{3}} \cos B_z \quad (5.84*)$$

The characteristic impedance Z_l is determined by solving

$$Z_l = \frac{Z_0}{\tan^2 \theta_0} \frac{z l}{2} \frac{Z_L}{Z_l}^{1/2} \frac{Z_0}{Z_l} \sim \sqrt{f J} z r z i \quad (5.85a)$$

and Z_2 and Z_3 are given by

$$Z_l = (Z_h Z_0)^{1/2} \quad (5.856)$$

$$Z_a = \frac{Z_l}{\sqrt{f J}} \quad (5.85c)$$

When θ_Z approaches $\pi/2$, all three zeros coalesce at $Z_x = Z_L/Z_I$. The required value of Z_x is obtained from (5.85a) by equating the left-hand side to zero. That is, $Z_x = Z_L/Z_I \approx 1$, where $i < a < 1$. With Z_L/Z_I near unity, a solution for Z_x can be found quite readily by a trial-and-error process (note that the equation for Z_x is a fourth-degree equation).

$$\cos \frac{1}{3} \sqrt{3} / \left(\sin S J^{2/3} \right)^{1/2} \text{cosfl}$$
(5.86a)

$$T = \frac{A^1}{\cos \theta} v^4$$
(5.86b)

The Chebyshev transformer represents an optimum design in that no other design can give a greater bandwidth with a smaller passband tolerance. If it is assumed that some choice, other than (5.75), for the polynomial Q_{2N} in (5.74) can give a smaller passband tolerance for the same bandwidth, it will be found that a plot of the polynomial Q_{2N} will intersect the polynomial T in at least $N + 1$ points. Since the polynomials are even in $\cos \theta$, they have at most $N + 1$ coefficients. Thus Q_{2N} must be equal to T since they have $N + 1$ points in common. But this equality contradicts the original assumption that Q_{2N} could yield a better result, and hence proves that the Chebyshev transformer is an optimum one.

5.15 FILTER DESIGN BASED ON QUARTER-WAVE-TRANSFORMER PROTOTYPE CIRCUIT

A very interesting filter design based on the theory of multisection wave transformers was given by Young. The quarter-wave transformer is a bandpass filter but would normally not be used as a filter since its input and output impedances are very different. In a multisection quarter-wave transformer, the impedances increase monotonically from input to output. In most filter applications we desire equal input and output impedances. Young showed that every other impedance step in a multisection quarter-wave transformer could be replaced with an $ABCD$ step. By alternating between a step up in impedance level and a step down in impedance level, we can end up with a final output impedance equal to that at the input. The filter design based on this concept is described below.



FIGURE 5.44

(a) An impedance step; (b) an equivalent junction when $Z'_{1'}/Z_1 = Z_2/Z_1$.

Consider the two junctions shown in Fig. 5.44. The first junction is a simple impedance step from Z_1 to Z_2 . The second junction is also an impedance step from Z_1 to Z_2 , but, in addition, it has an ideal transmission-line section of electrical length $T/2$ on either side. Furthermore, we assume that this electrical length does not vary with frequency. Obviously, we have introduced a nonphysical element and the reader may rightfully question whether anything useful can come from introducing such nonphysical elements. We will show that in the final filter configuration these nonphysical transmission lines can be eliminated. Thus their introduction is only to facilitate the development of the theory for the filter design. The two junctions will be fully equivalent if the scattering-matrix parameters S_u and S_{22} are the same as S_u and S_{22} . If the output is terminated in a matched impedance Z_2 , we get

$$\frac{Z_2 - Z_1}{Z_2 + Z_1} = \frac{(Z_2/Z_1) - 1}{(Z_2/Z_1) + 1}$$

for the *first junction*. Similarly, [or a matched termination on the input side, we get

$$** \quad \frac{i - (Z_2/Z_1)}{i + (Z_2/Z_1)} = S_{11}$$

For the second junction we use the quarter-wave-transformer formula to evaluate Z_{in} with the output matched. This gives

$$7 \quad (*a*) \\ '2$$

from which we get

$$11 \quad \frac{(Z_1)^2/Z_2 \sim Z_1}{(Z_1 f/Z_2 + Z_1)}$$

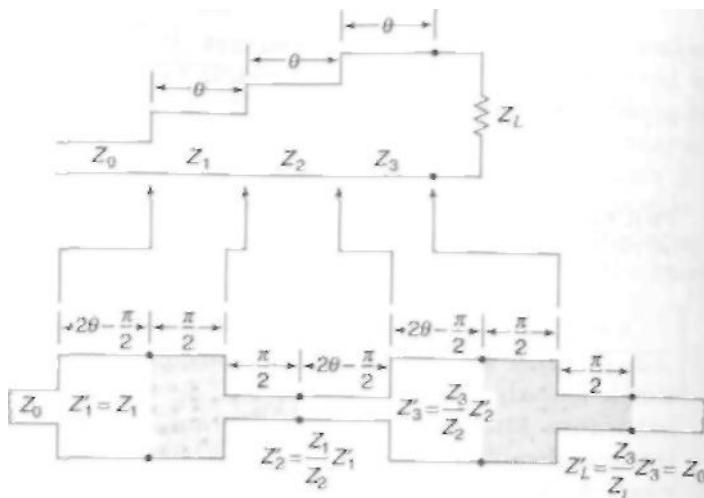


FIGURE 5.45

Microstrip filter design obtained from a multisection quarter-wave-transibrmer prototype circuit by replacing every other impedance step by the equivalent junction shown in Fig

In a similar way we easily find that $S_{22} = -S'_n$. By comparing the expressions for S_u and S'_n , we now conclude that the two junctions are equivalent if

$$\mathbf{Z}_u^{\wedge} = \mathbf{Z}_v^{\wedge} \quad (5.87)$$

This is precisely the property we are interested in because if $Z_2/Z_x > s_a 8t$ up in impedance, then Z'_2/Z'_1 must be a step down in impedance level.

Our next step is now to replace every other impedance step »n multisection quarter-wave transformer by this new

equivalent junction, w» assume that the multisection quarter-wave transformer we use as type circuit has already been designed to give a desired power loss ra "». The filter design procedure is illustrated in Fig. 5.45. In order that each section line have an electrical length $20 - \sim 2$. Thus, when $\sim = .Y_n$ and $TT/2 = 77/2$. In other words, we use a physical line of electrical length $TT/2$ and a nonphysical line of electrical length $-TT/2$ for reasons fully explained later. For the first junction in the filter, we

$$\mathbf{Z}_u^{\wedge} = \mathbf{Z}_x^{\wedge} \quad 5.88^\circ$$

which is the same as in the quarter-wave transformer. The second**

is replaced by the *new* equivalent junction; so we make

$$\sim - \sim \quad \% \sim \bullet = - + \quad (5.886)$$

This new junction incorporates the two nonphysical transmission lines of electrical length $TT/2$. The next junction is an impedance step like that in the quarter-wave transformer; so

$$-f = ^\wedge \quad Z'_0 = -\sim = -\$Z_Z \quad (5.88c)$$

The next junction is again the new equivalent junction which requires

This procedure is continued until we have worked our way through all the impedance steps in the quarter-wave-transformer prototype circuit.

The filter illustrated in Fig. 5.45 is a three-section filter. If this filter is examined, it will be seen that each section contains a transmission line of electrical length $20 - -/2 + -/2 = 20$, so the nonphysical lines have been eliminated. We chose each transmission line to have an electrical length of $20 - IT/2$ specifically for the purpose of eliminating the nonphysical lines that are part of the new junctions that replace every other impedance step in the quarter-wave-transformer prototype circuit. However, this means that the power loss ratio of the filter is obtained by replacing 0 by $26 - IT/2$ in the expression for the power loss ratio for the quarter-wave transformer. If the quarter-wave transformer is a Chebyshev transformer with power loss ratio given by (5.75), then the power loss ratio for the filter is given by

$$\begin{aligned} P_{\text{loss}} &= 1 + k^2 T^* \sec 0_{\text{center}} |2f| - \\ &= 1 + \&^2 T^* [\sec 0_{\text{center}} \sin 20] \end{aligned} \quad (5.89)$$

in which $\sin 20$ replaces $\cos 0$ as the frequency-dependent variable. At the center frequency $28 = TT$, so each filter section is one half-wavelength long. For this reason the filter is called a half-wave filter. The band edges which occur at $\cos 0 = \pm \cos 8_m$ for the quarter-wave transformer now occur when $\sec 0_{\text{center}} \sin 20 = \pm 1$ or

$$\sin 20 = \pm \cos 0_{\text{center}} = \pm \sin | - \pm 0_m$$

which gives

$$pi = 20 = - \pm e_{\text{center}} \pm nv$$

The upper band edge will be at $(3! = 3w/2 - \delta_m)$ and the lower will be at $/3! = TT/2 + 0$. Consequently, the fractional vJ nd <&> $MPD/IT = 1 - 29_{m/v}$, which is one-half that of the quarter - ^{d 1 h} former prototype circuit. Thus the prototype circuit should be'd^{11**} have a bandwidth twice as large as that required for the filter t this change the frequency response of the filter is the same $\approx u^A$ ^{that}, prototype circuit.

As a final point we note that, in a quarter-wave transform odd number of sections, the power loss ratio is unity at th_n ^{*} ^{**} ^{***} fft. Since

$$\begin{array}{cccc} z_L & z\%^\Delta & zi \\ \gamma_m & \gamma^2 & \gamma^2 & \gamma^0 \end{array}$$

for an odd number of sections, the final impedance Z'_L in the filter equal* Z_{Δ} . Hence the filter operates between impedance levels of Z_Q and Z last nonphysical transmission line that appears in the output line (see Fk 5.45) for an odd number of sections can be deleted since this is a matched line of arbitrary length.

For an even number of sections, Z_{in} does not equal Z_0 at the center frequency because the power loss ratio does not equal unity; it equals $1 + *^3$ for this case. This means that the last impedance element in the filter, which is given by

$$\sim, \quad y, \quad z^2 I^z 3 \quad z_{N-L} ? \quad - y$$

will be different from Z_0 . By using (5.77) for p at the center frequew can solve for Z_{in} to get

$$z = Z_0 f^\Delta = \left(*^+ / r T \Delta \right)^2 \Delta \quad (5.90)$$

Since k is normally very small, the output line for the filter has i stic impedance not quite equal to that of the input line. ^{^j- ora}

In Table 5.1 we list the required values of the Impecia^{vgjaef} Chebyshev quarter-wave transformer with three sections for se of the passband tolerance k^2 . We will use this table to design ^{^n t} filter in the example that follows. More extensive tables are e literature.t

See, for example, G. L. Matthaei, L. Young, M T * - " - 5
Impedance Matching Networks, and Coupling Structures," Artec Mass. 1980.

Chebyshev quarter-wave-transformer data

Z_{in}/Z_0	A/V/o-0.2		$\sqrt{f/f_o}$	0.4	A f/f,, = 0.6	
	Z_{in}/Z_0	k^*	Z_{in}/Z_0	f_c^2	$Z_{\text{in}}/Z_{\text{out}}$	f_e^*
2	1.09247	1.19×10^{-7}	1.09908	$7.89 > 10^{-6}$	1.1083	9.57×10^{-6}
4	1.1474	5.35×10^{-7}	1.20746	3.55×10^{-5}	1.23087	4.31×10^{-4}
10	1.349	1.92×10^{-7}	1.37482	1.28×10^{-4}	1.4212	1.55×10^{-3}
20	1.48359	4.29×10^{-7}	1.52371	2.85×10^{-4}	1.60023	3.45×10^{-3}
100	1.87411	2.33×10^{-6}	1.975	1.55×10^{-3}	2.17928	1.87×10^{-2}

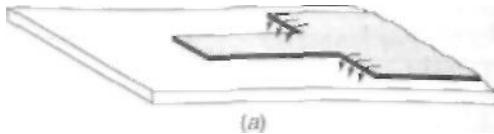
7. $\sqrt{Z_{\text{in}}/Z_0} = \sqrt{k^*} = \sqrt{f_e^*/f_c^2}$

Example 5.3 Filter design. We want to design a bandpass filter with a fractional bandwidth of 0.2 and having a VSWR of 1.02 or less in the passband. From the given VSWR we find $\sqrt{f/f_o} = (1.02 - 1.00)/2.02 = 9.9 \times 10^{-3}$. By using (5.77) we get $k' = t>f_o/(l - p\%) = 9.8 \times 10^{-5}$. The quarter-wave-transformer bandwidth must be chosen as $2 \times 0.2 = 0.4$. From Table 5.1 we find that, for $Z_{\text{in}}/Z_0 = 10$ and $Hff = 0.4$, $k^* = 1.28 \times 10^{-4}$. This value of k' would give a maximum value of VSWR equal to 1.023 in the passband. We will accept this value since an entry for $k' = 9.8 \times 10^{-5}$ is not given. From Table 5.1 we get $Z_{\text{in}}/Z_0 = 1.37482$. We will use 50-S1 input and output lines. Thus $\sqrt{Z_{\text{in}}/Z_0} = 1.37482 \times 50 = 68.74$ SI. The required values of Z_2 and Z_{out} are $Z_2 = \sqrt{W_1} = \sqrt{50} = 158.1$ (1). $Z_{\text{out}} = 500 \times 50/68.74 \Rightarrow 363.69$ (1). This completes the design of the quarter-wave-transformer prototype circuit. For the filter we need $Z_1 = Z_{\text{in}} = 68.74$ SI, and by using (5.886) and (5.88c) $Z_2 = 29.89$ il. $Z_{\text{out}} = 68.74$ SI. These impedance values are readily realized for a microstrip filter.

The maximum out-of-band attenuation occurs at $\theta = 0$ and is the same as what is obtained when the input line is connected directly to a load $Z_L = 10Z_0$. The reflection coefficient will be $9/11$ so 66.9 percent of the incident power is reflected. The transmission through the filter is reduced by $-10 \log(1 - p^2) = 4.8$ dB. This is a small attenuation and shows that the particular filter configuration used here will generally not have a large out-of-band attenuation. If we use $Z_{\text{in}}/Z_{\text{out}} = 100$, we would obtain an attenuation of 14 dB. But in this case the passband tolerance would be larger. The required passband tolerance could be achieved by using a filter with more sections. However, there are better filter configurations to use when large out-of-band attenuation is needed (see Chap. 8).

C \wedge C pen n sa C tro P n a CitanCe a n d L e n \wedge h

* an abrupt step in the width of a microstrip line, there will be an additional fringing electric field from the open-circuited portions of the wider strip as shown pictorially in Fig. 5.46a. The effect of this fringing field can be modeled as a shunt capacitance at the junction. The equivalent circuit of the step is shown in Fig. 5.46b and consists of an ideal impedance



(a)

-Of

$$S_{11} \approx 2^{-z} < c^*$$

(b)

$$\mathbf{au} - ? \quad \mathbf{y}^* \quad \mathbf{v}^* \quad \mathbf{C} \quad \mathbf{H}_{22}$$

(c)

FIGURE 5.46

(a) Fringing electric field at a step change in width for a microstrip line; (b) equivalent circuit for a step-change-in-width junction; (c) an alternative equivalent circuit.

step and a shunt capacitive susceptance. The susceptance is small but does modify the performance of the filter if not properly compensated. The junction capacitance can be compensated for by changing the length of each filter section. The equivalent circuit in Fig. 5.466 will be shown to be nearly equivalent to the circuit in Fig. 5.46c. The latter consists of the ideal impedance step plus two short lengths of transmission line. The electrical lengths of these transmission lines are denoted by θ_0 and θ_{-z} . The two electrical lengths are of opposite sign.

For the equivalent circuit in Fig. 5.466, we have

$$S_{11} = \frac{Y_l - Y_2 - jB}{Y_l + Y_2 + jB} \quad \frac{Y_x - Y_t}{Y_t + Y_{2l}} \frac{1 - \frac{jB}{jB}}{\frac{Y_x - Y_t}{jB} + \frac{Y_{2l}}{Y_l + Y_2}}$$

Since $B/(Y_x + Y_2)$ is very small, we can use the binomial expansion to obtain

$$\frac{B}{Y_l + Y_2} \quad \frac{\mathbf{B}}{Y_x - Y_2} \quad \frac{B}{Y_l T Y_{2l}} \\ \frac{* \mathbf{i} - Y_2}{\mathbf{r} + Y_l} \quad 1 - \frac{JB}{Y_l + Y_2} \quad \frac{B}{Y_l + Y_2} \\ \frac{\mathbf{r} - Y_2}{Y_l + Y_2} \quad 1 - \frac{-2jBY_l}{Y_l - Y_2}$$



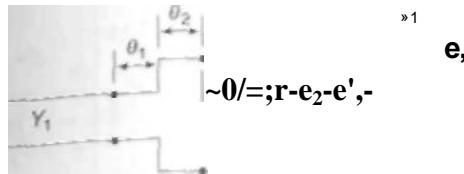


FIGURE 5.47
Length compensation of a filter section.

For the circuit in Fig. 5.46c, we have

$$Y_1 - Y_2 \quad , \quad Y_r - Y, \quad (5.92)$$

when θ_1 is small. In a similar way we obtain

$$\frac{Y_2 - Y_1}{Y_{22}} = \frac{1}{F_2 + F_1} - \frac{2jBY_2}{2} \quad (5.93)$$

$$S_{22} = \frac{Y_1 - Y_2}{Y_1 + Y_2} R (1 - 2^{\lambda}) \quad (5.94)$$

A comparison of S_u and S_{22} with S' , and S_{22} shows that the two junctions are equivalent, to the order of approximations used, if we choose

$$\frac{BY}{Y_1} = \frac{Y_2 - Y_1}{Y_2} \quad (5.95a)$$

$$\bullet \rightarrow \frac{BY}{Y_2} = \frac{Y_1 - Y_2}{Y_1} \quad (5.95b)$$

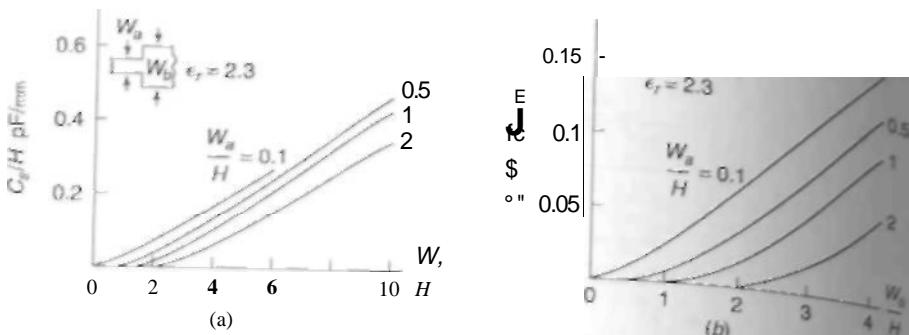
In the half-wave filter the length of each section is now changed to compensate for the junction capacitance as shown in Fig. 5.47. The required section length is 2ft including θ_1 and θ_2 . Hence the physical length l is chosen to make

$$l/ = 20 - \theta_2 - \theta_1 = \pi - \theta_2 - \theta_1 \quad (5.96)$$

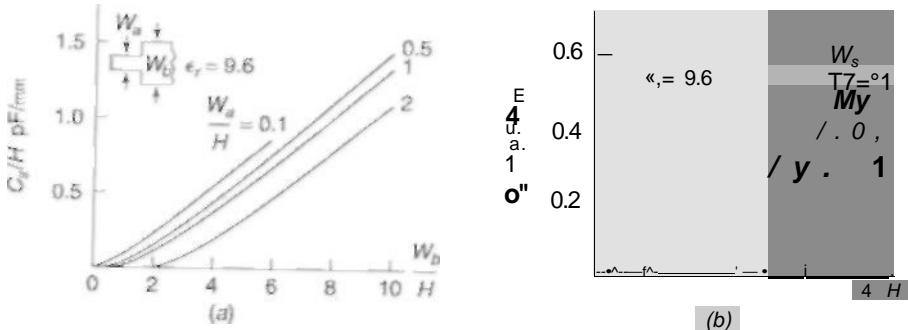
at the center frequency. Since θ_1 and θ_2 , etc., vary with frequency like $B = coC$, compensation is obtained at all frequencies. The parameters θ_2 and θ_1 are small so the change in physical length of each section is small but important enough to take into account in the design of a filter.

Similar junction capacitance effects also occur in some quarter-wave-transformer realizations and can be compensated for by changing the Physical length of each section by a small amount.

The application of the above method of compensation will be illustrated in Example 5.4, but first we need data for the junction capacitance. A number of investigations of junction equivalent circuits for microstrip discontinuities have been carried out, but surprisingly very little data for design purposes are available. Gupta and Gopinath have evaluated the

**FIGURE 5.48**

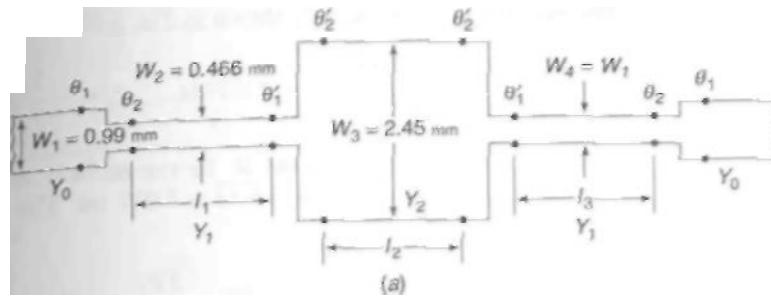
(a) Junction capacitance C_f for a step change in width for a substrate with $\epsilon_r = 2.3$.
 (b) Enlarged portion of curves in (a).

**FIGURE 5.49**

(a) Junction capacitance C_f for a step change in width for a substrate with $\epsilon_r = 9.6$.
 (b) Enlarged portion of curves in (a).

shunt capacitance at a step in width for a microstrip line for substrates with dielectric constants of 2.3, 4, 9.6, and 15.1. Their data give the capacitance normalized by the distributed capacitance per meter of a wider microstrip line. We have used these data to evaluate the capacitance for the two cases where the dielectric constant of the substrate is 2.3 and 9.6. The junction capacitance C_s divided by the substrate thickness H is shown in Figs. 5.48 and 5.49.

Example 5.4 Microstrip half-wave filter. The filter described in Sec. 5.3 is to be built using microstrip construction. The substrate is a



IM

FIGURE 5.50

(a) Three-section half-wave filter using microstrip construction; CM scaled drawing of filter.

a dielectric constant of 9.6 and is 1 mm thick. By using the computer program MSTP, we find after a few iterations the following required widths and the effective dielectric constants of each section:

$$\text{For } Z_c = 50 \text{ il} \quad W = 0.99 \text{ mm} \quad e_e = 6.49$$

$$\text{For } Z_c = 68.74 \text{ JI} \quad W = 0.466 \text{ mm} \quad e_e = 6.18$$

$$\text{For } Z_c = 29.89 \text{ (1)} \quad W = 2.45 \text{ mm} \quad e_e = 7.07$$

With reference to Fig. 5.49, we identify W_a/H and W_b/H for the junction between the 50- and 68.74-SI lines to be

$$H = 0.466 \quad H = 0.99$$

From Fig. 5.496 we estimate CJH to be 0.03 and since $H = 1 \text{ mm}$, $C_{ss} = 0.03 \text{ pF}$. The center frequency of the filter passband is to be 4 GHz. Hence $B = ioC_s = 7.5 \times 10^{-4} \text{ S}$. The characteristic admittances y_{ss} and Y_t are (see Fig. 5.50)

$$* \bullet - \gg - * * \quad * \mathbf{k} - \frac{1}{68.74} = 0.01455$$

We now use (5.95) to obtain

$$e_s, \quad \frac{BY_0}{Y^2 - V^s} = 0.08 \quad 6.2 = \frac{\gg}{-0.058} = -0.058$$

For the junction between the 68.74-fl line and the 29.89-\$! line, we have

$$H = 0.466 \quad \mathbf{k} = \bullet$$

From Fig. 5.496 we estimate C_s to be 0.18 pF which gives $B = 4.52 \times 10^{-3} \text{ S}$.

We now determine $D \setminus$ and $f f_z$ shown in Fig. 5.50.

$$BY \quad Y$$

where $Y_3 = 1/29.89 = 0.03346$ S. In the 68.74-fl section wavelength is $A_{fl}/JT_e = 7.5/\underline{16A8} = 3.017$ cm. The required length of this section is $\hat{\wedge}^{** le"e<> /}$

$$l = \frac{A}{2^{fl/v}} \sim n^2 \sim ffl = "2V^A = 0.52A = 1.57 \text{ cm}$$

For the 29.89-12 line, we have $e_c = 7.07$ and a wavelength $7.5/\sqrt{107} = 2.821$ cm. The required length l_t is given by

$$h = r - (ir - 20') = 0.447A = 1.261 \text{ cm}$$

Note that the length corrections are about 4 and 10 percent. If the filter were designed for 2-GHz operation, the length corrections would be half as large or about 2 and 5 percent. Below 1 GHz, compensation for the junction susceptances could be ignored. A proportionally scaled drawing of the foil pattern for the filter is shown in Fig. 5.506.

5.16 TAPERED TRANSMISSION LINES

In a multisection quarter-wave transformer used to match two transmission lines with different characteristic impedances, the change in impedance is obtained in a number of discrete steps. An alternative is to use a taper transition which has a characteristic impedance that varies continuously! a smooth fashion from the impedance of one line to that of the other line. transition, or matching section, of this type is referred to as a P& transmission line. An approximate theory of tapered transmission analogous to the approximate theory presented earlier for mul transformers, is readily developed. This approximate theory is pf • below. A following section gives a derivation of the exact differenti gives a brief evaluation of the validity of the approximate theory. ^

Figure 5.51a illustrates schematically a tapered transmission ^ ^ ^ to match a line with normalized impedance unity to a load with impedance Z_L (assumed to be a pure resistive load). The normalized impedance Z which is a function of the distance z , taper. Figure 5.516 illustrates an approximation to the contmuo^ erentis considering it to be made up of a number of sections of one length dz and for which the impedance changes by different from section to section.

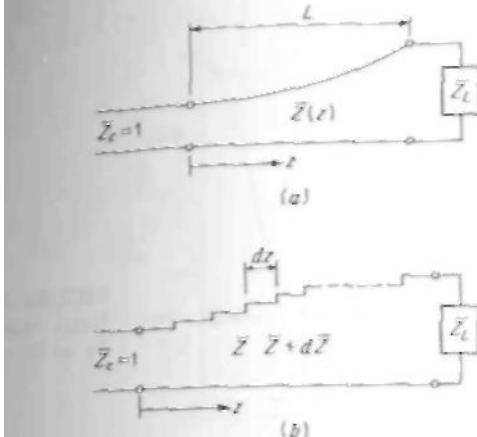


FIGURE 5.51
Tapered-transmissinn-line matching section.

The step change dZ in impedance at z produces a differential reflection coefficient

$$\begin{aligned} dr_0 = & \frac{Z + dZ - Z}{Z + dZ + Z} \frac{dZ}{2Z} - \wedge d(\ln Z) \\ & - \frac{1}{2} \frac{d}{dz} (\ln Z) c k \end{aligned} \quad (5.97)$$

At the input to the taper, the contribution to the input reflection coefficient *From this step* is

$$dI_{in} = e^{\bullet} + m \frac{1}{2} \frac{d}{dz} (\ln Z) r f z$$

If it is assumed that the total reflection coefficient can be computed by summing up all the individual contributions, as was done in the approximate theory of the multisection quarter-wave transformer, the input reflection coefficient is given by

$$* - i r * \sim S O - *) * \quad (5.98)$$

where L is the total taper length. If the variation in Z with z is known, I_r may be readily evaluated from the above. A problem of much greater practical importance is the synthesis problem, where $Z(z)$ is to be determined to give I_r , the desired characteristics as a function of frequency. Before taking up the synthesis problem, two examples of practical taper designs are presented.

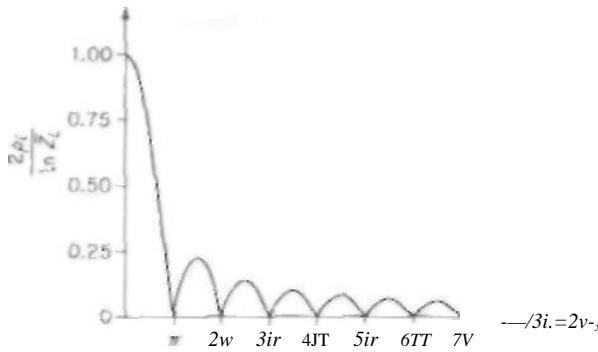


FIGURE 5.52
Input reflection
for an exponential taper

Exponential Taper

The exponential taper is one for which $\ln Z$ varies linearly, and hence Z varies exponentially, from unity to $\ln Z_L$; that is,

$$\ln Z = -\ln Z_L \quad (5.99a)$$

$$z = \epsilon(z/L)n^2, \quad (5.99b)$$

Substituting (5.99) into (5.98) gives

$$r = 2^{1-i/m} Z, \quad r^e dz \sim r^{1+iL} \tan Zi - r^{5100} \quad (5.100)$$

where it has been assumed that we are dealing with a transmission line which ($i = k = 2ir/A$ and is not a function of z). A plot of $p_r = (H_r/I)$ versus z is given in Fig. 5.52. For a fixed length of taper, this is a plot of P_r as a function of frequency since $k = 2\pi f/\lambda$. Note that when L is greater than $A/2$, the reflection coefficient is quite small, the first minor lobe being about 22 percent of the major-lobe maximum.

Taper with Triangular Distribution

If $d(\ln Z)/dz$ is chosen as a triangular function of the form

$$\frac{d(\ln Z)}{dz} = \begin{cases} \frac{4}{L} \ln Z, & 0 < z < j \\ \frac{4}{L} (L-z) \ln Z_L, & j < z < L \end{cases} \quad (5.10D)$$

a matching section with more desirable properties is obtained. Integ^{ati}

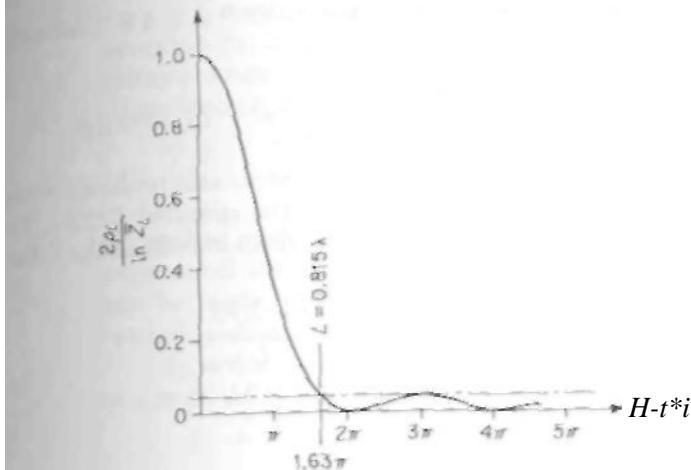


FIGURE 5.53
Input reflection coefficient for a taper with a triangular distribution of reflections.

(5.101) gives

$$Z = \frac{e^{jz/fcr^*lbffc}}{e^{(4j/I, -2s^2/-id-1)\ln Z}}, \quad \frac{L}{2} < z < L \quad (5.102)$$

Substituting (5.101) into (5.98) and performing the straightforward integration give

$$I = -e^{-' \ll M n Z}, \quad \frac{\sin(3L/2)}{0^2} \quad (5.103)$$

A plot of p , versus $/3L$ is given in Fig. 5.53. Note that, by comparison with the exponential taper, this taper has a first minor-lobe maximum which is less than 5 percent of the major-lobe peak. However, this small value of reflection coefficient occurs for a taper length of about $3A/2$, or for a length twice that for the exponential taper. If Z_L is considerably greater than unity, this latter taper will be preferable because of the much smaller values of p , obtained for all frequencies, such that the taper length is greater than 0.815λ , which corresponds to the lower edge of the passband in Fig. 5.53.

SYNTHESIS OF TRANSMISSION LINE TAPERS

Equation (5.98) is repeated here for convenience:

$$r_{(2/3)} = \frac{1}{2J_0} \int_{f_V}^{\infty} \frac{d(f_n Z)}{dz} \quad (5.104)$$

This equation may be interpreted as the Fourier transform of a function

$d(\ln Z)/dz$, which is zero outside the range $0 < z < L_f$. As such inversion formula gives

$$\frac{1}{2} \frac{d(\ln Z)}{dz} = \frac{1}{2} \frac{-\infty}{W}.$$

This formula, in principle, solves the synthesis problem since "t required value of $c/(\ln Z)/dz$ to yield the specified $r_{-}(2/3)$. To " " \wedge the discussion follow it will help enenpnienf to intmrlnm 4L- /• i> " " the ized variables:

$$\hat{\gamma} = L/2$$

$$\begin{matrix} 0 & L & 2L \\ TT & A & \end{matrix} \quad (5.106)$$

In this case (5.104) becomes

$$\bullet \frac{1}{2^e} \frac{H_f^*}{J J} \frac{d(\ln Z)}{dp} dp \quad (5.107)$$

Now define $g(p)$ to be

$$d(\ln Z) \quad (5.108a)$$

and $F(\zeta)$ by

$$F(M) = \int e^{-jmu} g(p) dp \quad (5.108b)$$

Thus

$$F_\zeta = \pm e^{jL} F(u) \quad (5.109)$$

The Fourier transform pair (5.104) and (5.105) now may be expressed as

$$Flu) - f_j - \int g(p) dp \quad (5.110)$$

$$1 / \bullet . , \quad r T 1106$$

The synthesis problem may now be stated as follows: Specify a function-coefncient characteristic $F(u)$ that will give the desired $\wedge \wedge j \wedge a$ mance and yet be such that the $g(p)$ computed by (5.110) is identically zero outside the range $|p| > \cdot$. This f_j^* ent from corresponds to the physical requirement that $d(\ln Z)/dz$ be di

The Fourier transform relation was first pointed out by F. Bolinder, *Fourier Transform*, *Proc. IRE*, vol. 38, p. 1950.

pro only in the range $0 < z < L$. Obviously, any arbitrary $F(u)$ cannot be specified, for in general this would lead to a $g(p)$ that exists over the whole infinite range $-\infty < p < \infty$. For example, if $F(u)$ were chosen to be equal to unity for $-1 < u < 1$ and zero otherwise, then (5.1106) would give

$$g(p) = \frac{\sin p}{ip} \quad x < p < \infty$$

To realize such a $g(p)$ would require an infinitely long taper, clearly an impractical solution. Before further progress with the synthesis problem can be made, restrictions to be imposed on $F(u)$, to obtain a physically realizable solution, must be deduced.

In order to derive suitable restrictions on $F(u)$, let $g(p)$ be expanded in a complex Fourier series as follows:

$$g(p) = \sum_{n=-\infty}^{\infty} a_n e^{jn\pi p} \quad -\pi < p < \pi \quad (5.111)$$

where the a_n are as yet unspecified coefficients subject to the restriction $a_n = a_{-n}$, so that g will be a real function. Substitution into (5.110c) now gives

$$\begin{aligned} F(u) &= 2\pi \sum_{n=-\infty}^{\infty} \frac{\sin \pi(U-n)}{\pi(U-n)} \\ &= 2\pi \sum_{n=-\infty}^{\infty} a_n (-1)^n \frac{\sin \pi u}{\pi(u-n)} \quad \text{and} \quad \sum_{n=-\infty}^{\infty} a_n (-1)^n = 1 \end{aligned} \quad (5.112)$$

The coefficients a_n can be related to $F(u = n)$, for when u equals an integer n ,

$$\lim_{n \rightarrow \infty} \frac{\sin \pi(U-n)}{\pi(U-n)} = 1$$

and $\lim_{n \rightarrow -\infty} \frac{\sin \pi(U-n)}{\pi(U-n)} = 0$

Thus $F(n) = 2a_n$ or $a_n = F(n)/2\pi$; so

$$F(n) = \frac{1}{2} \sum_{n=-\infty}^{\infty} a_n = \frac{1}{2} \sum_{n=-\infty}^{\infty} \frac{\sin \pi(U-n)}{\pi(U-n)} \quad (5.113)$$

This result is a statement of the well-known sampling theorem used in communication theory and states that $F(u)$ is uniquely reconstructed from a knowledge of the sample values of $F(u)$ at $u = n$,

$$n = 0, \pm 1, \pm 2, \dots$$

by means of the interpolation formula (5.113). One possible way to specify $F(u)$ is now seen to be a relaxation on the specification of $F(u)$ at all integer values of u only. This, however, is not a satisfactory solution, because we have no a priori knowledge to specify $F(u)$ at the integer values of u only, the resultant $F(u)$ will be an acceptable reflection-coefficient characteristic for values of u , even though it can be realized by a $g(p)$ given by $a_n = F(n)/2ir$.

We should like to obtain greater flexibility in the choice of $F(u)$. To see how this may be accomplished, let it be assumed that all a_i for $i > N$ are zero. In this case

$$g(P) = \sum_{n=0}^N a_n e^{jn\theta} \sin vu - jv \quad , \quad u$$

$$F(u) = \sum_{n=0}^N a_n \frac{1}{u^n - 1} \quad (5.144)$$

The series in (5.1146) can be recognized as the partial-fraction expansion of a function of the form

$$Q(u)$$

where $Q(u)$ is, apart from the restriction $Q(-u) = Q^*(u)$ so that $a_n = a_{-n}$, an arbitrary polynomial of degree $2N$ in u , and the denominator is the product of the N terms $(u^2 - 1), (u^2 - 4)$, etc. Using the partial-fraction-expansion formula, we have

$$\begin{aligned} Q(u) &= \sum_{m=0}^{\infty} \frac{Q(m)}{(u^2 - m^2)^{n+1}} \\ \mathbf{n}^{\mathbf{x}} \ll \mathbf{i}^{\mathbf{*}} &= \frac{Q(0)}{u^2} + \frac{Q'(0)}{u^3} + \dots + \frac{Q^{(N)}(0)}{u^{2N+1}} \\ &= \frac{Q(0)}{u^2} + \frac{Q'(0)}{u^3} + \dots + \frac{Q^{(N)}(0)}{u^{2N+1}} \\ &= \frac{Q(0)}{u^2} + \frac{Q'(0)}{u^3} + \dots + \frac{Q^{(N)}(0)}{u^{2N+1}} \quad (5.145) \end{aligned}$$

where the prime means omission of the term $m = 0$. This is clear. same form as the series in (5.1146), with

$$\left(\dots \right) = \frac{Q(m)}{2m^2 n_{m+1} \dots n_1} \quad (5.146)$$

$$\underline{Q(Q)} \quad [5.n^{6b}]$$

The expression for $F(u)$ now can be written as

$$F(u) = \frac{\sin u}{2v} \frac{Q(u)}{n v^2}, \quad (5.117)$$

•here $Q(u)$ is an arbitrary polynomial of degree $2N$ in u , subject to the restriction $Q(-u) = Q^*(u)$. This result states that the first $2N$ zeros of in TTU which are canceled by the denominator in (5.117), can be replaced by $2N$ new arbitrarily located zeros by proper choice of $Q(u)$. If $g(p)$ were a constant (exponential taper), we should have $F(u)$ proportional to $(\sin u)/iru$. But with $2N + 1$ coefficients available in the expansion of $g(p)$, we are at liberty to rearrange $2N$ of the zeros of $(\sin TTU)/TTU$ to obtain a more desirable $F(u)$. We have now reduced the synthesis problem to one of specifying an arbitrary polynomial $Q(u)$. To illustrate the theory, two examples are discussed below.

A qualitative insight into how $Q(u)$ should be specified may be obtained by imagining that $F(u)$ is a rubber band stretched horizontally at some height above the fL or u axis. The zeros of $F(u)$ may then be thought of as points at which this rubber band is pinned down to the it axis. If the band is pinned down at a number of closely spaced points, it will nod rise much above the u axis in the regions between. The corresponding reflection coefficient will then also be small in this region. At a double zero the band is pinned down in such a fashion that its slope is zero at the point as well. This results in a less rapid increase in the height with distance away from the point. The number of zeros available (the points at which the band may be pinned down) is fixed and equal to those in the $\sin TTU$ function. The polynomial $Q(u)$ permits only a relocation of these zeros.

With reference to Fig. 5.52, which illustrates the characteristic for an exponential taper, let the zero at $/3Z_1 = IT$ be moved to $2-$ to form a double zero at this point. Likewise, let the zero at 377 be moved to form a double zero at $4-$, etc. The function $Q(u)$ that will provide this shift in every other zero so as to produce double zeros at $u = \pm 2, \pm 4, \pm 6, \pm 8$, etc., can be chosen as

$$n = 1$$

From (5.117) we obtain

$$\sin^n u - 4^n$$

m t n i s c ae now wish to let N go to infinity. However, the products do not converge; so we must modify the expression for $F(u)$ to the following:

$$f(u) = C \frac{\sin run Z_{-1}(l-u^2/4n^2)^2}{N}, \quad .2.$$

This modification is permissible since Q contains an arbitrary multiplier. All constants can be incorporated in the one constant following infinite-product representations for the sine functions

$$\begin{aligned} \sin TTU &= \prod_{n=1}^{\infty} (1 - \frac{u^2}{n^2}) \\ -TTU &= \prod_{n=1}^{\infty} (1 + \frac{u^2}{n^2}) \\ \sin(7T\pi/2) &= \prod_{n=1}^{\infty} (1 - \frac{u^2}{(n+1)^2}) \end{aligned}$$

it is readily seen that as $N \rightarrow \infty$ we obtain

$$W \approx C \frac{\sin(-u/2)}{TTU/2}^{1/2} \quad (5.118)$$

This is the reflection-coefficient characteristic for the taper with a triangular $d(\ln Z)/dz$ function discussed earlier. In the present case we have arrived at this solution for a taper with equally spaced double zeros by a direct-synthesis procedure. As Fig. 5.53 shows, the specification of double zeros holds the values of $F(u)$ (that is, p_n) to much smaller values in the region between zeros.

As shown earlier, the coefficients a_n in the Fourier series expansion of $g(p)$ are given by

$$a_n = \frac{1}{2\pi} \int_0^{\pi} g(p) e^{-jnpr} dp = \frac{C}{2\pi} \frac{\sin(mr/2)}{mr/2} \quad (5.119)$$

from (5.118). The reader may readily verify that the expansion (5.114a) $g(p)$, with the above coefficients, is a triangular wave. To fix the constant (C) we integrate (5.108a) to obtain

$$\int_0^{\pi} g(p) dp = \int_0^{\pi} \frac{d(\ln Z)}{f-pr} dp = \ln \frac{Z(\pi)}{Z(0)} = \ln \frac{Z_L}{Z_0}$$

But from (5.111) we have

$$\int_0^{\pi} g(p) dp = 2ir\omega_0 \quad (5.120)$$

and hence

$$a = \frac{C}{2\pi} \ln \frac{Z_L}{Z_0}$$

From (5.119) we now find that

$$a = \frac{1}{2\pi} \ln \frac{Z_L}{Z_0} = \frac{C}{2\pi}$$

so $C = \ln Z_L/Z_0$. With this value of C , the reflection coefficient to the $F(u)$ in (5.118) is easily verified to be the same as that in (5.103).

As a second example, consider the synthesis of a tapered horn having a triple zero at $u = \pm 2$. This can be accomplished by

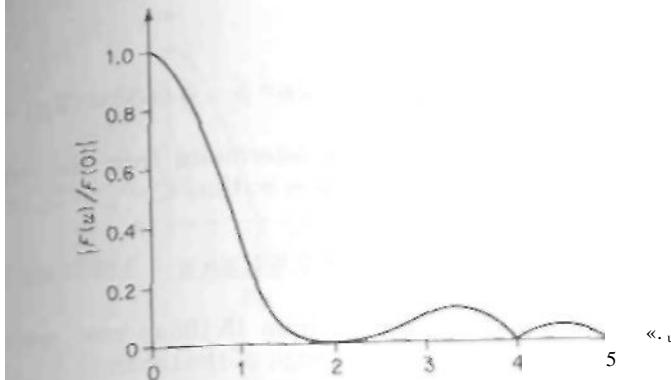


FIGURE 5.54
Reflection-coefficient characteristic for a taper with a triple zero at $|\ll| = 2$.

zeros at ± 1 and ± 3 into the points $u = \pm 2$. The resultant *rejection* coefficient should remain very small for a considerable region around $|u| = 2$. In the present case $N = 3$, and we choose

$$Q(\ll) = C(V - 4)^{1/2}$$

Thus
$$F(u) = \frac{\sin \pi u}{2\pi \ll} \cdot \frac{(\ll^2 - 4r^2)}{(z^{1/2} - 1)(u^2 - 4)(u^2 - 9)}$$

A plot of $|F(\ll)/F(0)|$ is given in Fig. 5.54. As anticipated, $F(u)$ remains small in a considerable region around the point $|\ll| = 2$. Since the zeros at $1 \leq |u| \leq 3$ have been removed, $F(u)$ reaches a relatively large value at this point. However, for a range of frequencies around which $L = A$, this taper represents a very good design.

The coefficients in the expansion for $g(p)$ are given by

$$o_0 = a_{-0} = \frac{1}{F(l)} = \frac{0.316}{-\ln Z_L}$$

$$a_{-1} = a_{-3} = 0$$

$$a_1 = a_{-1} = -F(3) = -\frac{0.098}{2\pi r} - \frac{1}{nZ_L}$$

$$a_{-2} = a_{-3} = 0 \quad \ll > 3$$

Hence

$$S(P) = \frac{d(\ln Z)}{dp} = \frac{nZ_L}{2\pi} (o_0 + 2a_1 \cos p + 2a_3 \cos 3p)$$

$$= \frac{1}{2\pi} (1 + 0.632 \cos p - 0.196 \cos 3p)$$

Integrating gives

$$\ln \bar{Z} = \frac{\ln Z}{\sim(P + 0.632 \sin p - 0.0653 \sin 3p) + c}$$

The constant of integration C is determined from the requirement that $\ln Z = 0$ at $p = 0$, or $\ln Z_0$ at $p = 0$; thus $C = A \ln Z_L$ and

$$\ln Z = \frac{\ln Z_0}{(P + 0.632 \sin p - 0.0653 \sin 3p)}$$

Replacing p by $2rr(z - L/2)/L$ from (5.106a) now specifies $\ln Z$ as a function of g and completes the design of the taper.

The foregoing synthesis procedure must be used with some caution stemming from the approximations involved. The theory is valid as long as $(\ln Z)/dz$ is small; that is, $\ln Z$ must be a slowly varying function of z so that the reflection coefficient everywhere along the taper be small, that is, $|g'(z)| \ll 1$. This means that $|g(p)|$, and hence all $|g_{mn}|$, must not be permitted to assume excessively large values. Consequently, $|F(n)|$ must not be permitted to become too large. If too many zeros are closely grouped together around a particular value of u , then outside this range $F(u)$ may become excessively large and the accuracy of the theory will suffer. Such "supermatched" designs must be avoided.

*5.18 CHEBYSHEV TAPER

If the number of sections in a Chebyshev transformer is increased indefinitely, with the overall length L kept fixed, we obtain a Chebyshev taper. This taper has equal-amplitude minor lobes and is an optimum design in the sense that it gives the smallest minor-lobe amplitudes for a fixed tap length, and conversely, for a specified minor-lobe amplitude it has the shortest length. As such it is a good taper by which to judge how far other tapers depart from an optimum design. It has been shown that, in the limit as the number of sections in a Chebyshev transformer goes to infinity, the reflection coefficient becomes

$$2 \cosh \frac{i}{2} \theta_0 L$$

where θ_0 is the value of θ at the lower edge of the passband, as shown in Fig. 5.55. As i increases from zero to $/3_0$, the magnitude of g decreases to a final value of $(\ln Z_L)/(2 \cosh \theta_0 L)$, since in

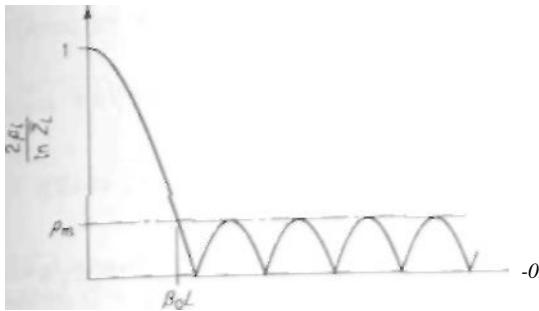


FIGURE 5.55
Reflection-coefficient characteristic for a Chebyshev taper.

$\cos L \sqrt{fp^2 - P_0} = \cos n \Delta VP? > - u^2$. Beyond this point the function in the numerator is the cosine function that oscillates between +1 and produces the equal-amplitude minor lobes. The major-lobe to minor-lobe amplitude ratio equals $\cosh j\beta_0 L$. Hence, if this is specified so as to keep p , less than or equal to some maximum value p_m in the passband, the taper length L is fixed for a given choice of the frequency of the lower edge of the passband which determines β_0 . We have

$$\cosh j\beta_0 L = \frac{\ln Z_L}{*P_m} \quad (5.122)$$

Conversely, if β_0 and the taper length L are given, the passband tolerance p_m is fixed.

The theory given earlier may be used to determine the function $g(p)$ that will produce the reflection coefficient given by (5.121). Introducing the u variable again, we find that the function $F(u)$ is

$$F(u) = (\ln Z_L) - \frac{\cos -Ju^2 - u^2}{\cosh 77 U_n} \quad (5.123)$$

where $-u = pL$, $-u_d = f_i L$. The function $\cos TT - Ju^2 - u^2$ can be expressed in infinite-product form as

$$\cos TT - Ju^2 - u^2 = \cosh TTU_n \prod_{n=1}^{\infty} [1 - \frac{1}{(n + 1)^2}]$$

and this is the limiting value of the polynomial $Q(u)$ in (5.117) as $N \rightarrow \infty$. The $\sin -u$ term has been canceled by the infinite product $\prod_{n=1}^{\infty} C_n \sim u^2/n^2$. In the first example presented on taper synthesis. However, we do not need this product expansion since $TTU_n = F(n)$ in any case. From (5.123) we have

$$\frac{1}{*P_m} = \frac{\ln Z_L \cos ir - Ju^2 - u^2}{\cosh 77 U_n} \quad (5.124)$$

Thus

$$\#(p) = 9 \frac{\ln Z_L}{ZT} (\cosh iru_0 + 2 \sum_{n=1}^{\infty} \frac{\cos ir^n * 3^n}{n} \cos np - ul \cos fa$$

$$+ 2 \cos Tr^7 4 - ul \cos 2p + \dots)$$

$$\frac{[aZ_L]}{2TT \cosh TT} (\cosh iru_0 + 2 \sum_{n=1}^{\infty} \frac{\cos ir^n * 3^n}{n} \cos np - ul \cos fa) \quad (5.55)$$

Integrating with respect to p gives

$$\ln Z = \frac{\ln Z_L}{2, \cosh TM_0} \int_{p=0}^{p=8hTM} + 2 \sum_{n=1}^{\infty} \frac{\cos Tr-Jn^2 - ul}{n} \sin np + C \quad (5.126)$$

where C is a constant of integration. To render this result more suitable for computation, it is expedient to add and subtract a similar series: i.e.,

$$\ln Z = \frac{\ln Z_L}{2v} + \frac{\ln Z_h}{\cosh u_j ir} - \frac{\cos nv}{2} \sin np$$

$$\frac{\ln Z_L}{TT \cosh U_0 T} \sum_{n=1}^{\infty} \frac{\cos -yn^2 - u\% - \cos nir}{n} \sin np + C$$

The second series converges rapidly because $\cos vyn^2 - u\%$ approach $\cos nIT$ as n becomes large. The first series may be recognized as Fourier sine series for the sawtooth function Sip ,

$$S(p) = \begin{cases} \frac{P}{2TT} & \text{if } -TT < p < TT \\ 0 & \text{otherwise} \end{cases}$$

$$p = \pm ir$$

5.56.

and the periodic continuation of this function as illustrated in

Sip)

-ZIT

FIGURE 5.56
Sawtooth funcWW-

Hence we obtain

$$\begin{aligned} \ln Z &= \frac{P \cdot 1^\circ Z /}{217} + S(P) \\ &+ \frac{\operatorname{Ir} Z_L}{77 \cos ih TTU_{..}} \cdot \frac{x}{L} \cdot \frac{\cos Txyn^2 - K| - \cos \gg TT}{n} \sin np + C \end{aligned}$$

At $p = rr$, we have $\ln Z = \ln Z_t$, and since $S(p)$ and $\sin np$ are zero at this point, $C + |\ln Z_L| = \ln Z_{..}$ or $C = |\ln Z_{..}|$. Our final result is

$$\begin{aligned} \ln Z &= \frac{p - i}{2JT} + \frac{p}{2T7} \cosh \frac{i}{\operatorname{cosh} 7ru_0 n} \ln Z_{..} \\ &\quad \ln Z_{..} * \frac{\cos irJn^2 - u'i - \cos /? rr}{77 \cosh 7ru_0 n} \operatorname{sinrap} \quad (5.128) \end{aligned}$$

for $-77 < p < TT$. An interesting feature of the above result is that $\ln Z$ changes in a stepwise fashion from 0 to $(\ln Z_{..})/(2\cosh TT/..)$ as p changes from $-IT - e$ to $-TT + e$, where $e \ll 1$. Likewise, at the other end of the taper, $\ln Z$ changes abruptly from a value $\ln Z_L - (\ln Z_L)/(2\cosh -u_a)$ to $\ln Z_{..}$ as the point $p = TT$ is reached. This means that the optimum taper has a step change in impedance at each end. The physical basis for this is readily understood by noting that when the frequency is very high, so that the taper is many wavelengths long, the reflection from the smooth part of the taper vanishes. Thus, in order still to maintain equal-amplitude minor lobes, the two-step changes in impedance must be provided to give a reflection coefficient

$$f_{..} = \frac{\ln Z_{..} \cos jiL}{2 \cosh (jiL)} \quad \text{for } P_H^* \bullet P_n^H$$

As an indication of the superiority of the Chebyshev taper, computations show that it is 27 percent shorter than the taper with $d(\ln Z)/dz$ a triangular function, for the same passband tolerance and lower cutoff frequency. If the tapers are made the same length, the Chebyshev taper provides a major- to minor-lobe ratio of 84 as compared with 21 for the taper with a triangular distribution.

C₀E⁹PFICnf^T E Q U A T I O N F O R T H E REFLECTION

The basic equation (5.98) for the input reflection coefficient T , was derived by neglecting all multiple reflections between individual differential sections. The exact equation, derived below, enables an estimate of the range of validity of the approximate theory to be made. First of all, the differential

equation describing the total reflection coefficient f at any point line, according to the approximate theory, is derived for later co

With referenceto Fig. 5.57, let $d\backslash$, be the reflection coeffic-^{11^50,1} from the change dZ in characteristic impedance in the interv $1 \wedge \wedge^g$. This differential reflection was shown earlier to be given by [see te a $^{3t} \ast$]

$$1 d(\ln Z)$$

The total reflection coefficient at z is the sum of all differential tions rfr_0 from z to L and is

$$1 ,/, .., d(\ln Z)$$

where u is a dummy variable that measures the distance from the point $2 = 0$ toward the load end. The phase angle of the reflected wave arising at u is $2(3(z - u))$ relative to the forward propagating wave at z .

Differentiating Hz with respect to z gives

$$\begin{aligned} rfr & \frac{2jp}{dz} ,_L M & d(\ln Z) & 1 .., d(\ln Z) \\ & 2 /, du & 2 du \\ & = 2 \wedge r - - \backslash & & (5-129) \\ & 2 dz \end{aligned}$$

This is the approximate differential equation for the total reflection coefficient at any point z along the taper.

To find the exact differential equation for T , let \bar{Z}_{in} be the input impedance at z and $Z_m + dZ_{in}$ be the input impedance at $z + dz$. We have

$$Z_{in} + dZ_{in} + jZ \tan O dz$$

$$\frac{Z_m + dZ_m + jZpdz}{Z + j(Z_{in} + dZ_{in})fSdz}$$

$$\left(\frac{Z_m + dZ_m + jZpdz}{Z + j(Z_{in} + dZ_{in})fSdz} \right)^2$$

$$= \frac{dZ_m + jZpdz}{Z + j(Z_{in} + dZ_{in})fSdz}$$

$$= \frac{dZ_m + jZpdz}{Z + j(Z_{in} + dZ_{in})fSdz}$$

$$= \frac{dZ_m + jZpdz}{Z + j(Z_{in} + dZ_{in})fSdz}$$

upon replacing the tangent by its argument $/? dz$, neglecting

orodu^

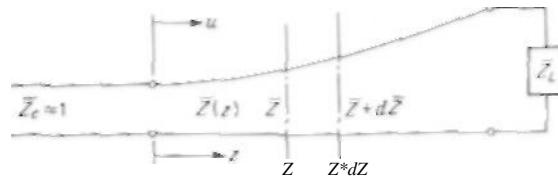


FIGURE 5.57
Tapered transmission line.

differential terms, and replacing the denominator by

$$\frac{[Z+j(Z_{in}pdz)]}{dz} \sim J \mathcal{T}^{**}$$

The above gives

$$\frac{dZ}{dz} = -J P^{Z \wedge -Z}$$

Now we also have

$$Z_{in} = \frac{i + r}{i - r} = \bar{Z}$$

and hence

$$\frac{dZ_{in}}{dz} = \frac{r}{i - r} \frac{dz}{dz} + \frac{2Z}{(i - r)^2} \frac{dr}{dz}$$

Combining these two equations for dZ_{in}/dz and replacing Z_{in} by $Z(1 + T)/(1 - r)$ finally give

$$\frac{dV}{dz} = \frac{1}{f^2} \frac{rf(\ln Z)}{\ln Z} \quad (5.130)$$

upon using the relation $Z^{-x} dZ/dz = d(\ln Z)/dz$. If we compare (5.130) with (5.129), we see that the approximate equation differs only by the factor $1 - T^2$ that multiplies the $d\ln Z/dz$ term in the exact equation. If $|IT^2| < K$ everywhere along the line, the approximate equation would be expected to yield good results.

The exact equation (5.130) is called a Riccati equation. It is a nonlinear equation because of the term in f^2 and does not have a known general solution. This equation can be integrated only in certain special cases (one such case is for the exponential taper, as in Prob. 5.48). However, the practical difficulty in applying the exact theory does not stem so much from the lack of a general solution of (5.130), since numerical integration or iteration schemes can always be employed, as from being unable to specify what the characteristic impedance Z is along a general taper. If the taper is very gradual, then $Z(z)$ can be taken as the characteristic impedance of a uniform line having the same cross-sectional dimensions as the taper has at the plane z . But for such gradual tapers the approximate theory is valid and

gives good results, so that the more complicated Riccati equation is required. For more rapidly varying tapers, the field structure $T(z)$ is not ^{not} ^{on a taper} perturbed to such an extent that no simple method of solving the new boundary-value problem will be available to determine $Z(z)$, and this solution will also provide the reflection coefficient along the taper. Thus one concludes the inability to specify $Z(z)$, except for the case of gradual tapers in the exact equation of minor importance in practice.

In the case of waveguide tapers, f_1 is a function of the cross-section geometry and hence a function of z along the taper. In order to handle the approximate equation (5.129), in this case, an auxiliary variable θ is introduced as follows:

$$\theta = (\sim 2p(z)dz) \quad d\theta = 2(idz)$$

and hence

$$\frac{d}{dz} = \frac{d}{d\theta} \frac{dd}{dz} = 26 \frac{d}{d\theta}$$

We now have

$$dY \quad dY \quad d(\ln Z)$$

$$dd^J \quad 2 \quad dti$$

This is readily integrated to give (multiply both sides by e^{Jt} and note that $dVe^{-Jt}/d\theta = -JTe^{-Jt} + e^{-Jt}rfr/d\theta$)

$$, \quad 1 \quad HI \quad d(\ln Z) Z$$

where

$$\theta_Q = (\sim 2p dz)$$

If $Z = Z_L$ at $z=L$, then $T(0) = 0$, and since $T(0) = T_0$, we have

in terms of the new variable θ , the problem is formally ^{formally} ^{hard} ^{to} ^{solve} ^{the} ^{problem} ^{is} ^{essentially} ^{the} ^{same} ^{as} ^{considered} ^{earlier} ⁱⁿ ^{the} ^{last} ^{step}, i.e., expressing $Z(H)$ as a function of H in or ^{or} ^{the} ^{shape} ^{of} ^{the} ^{taper}.

- 5.1. Show that the $R = \text{constant}$ and $X = \text{constant}$ contours in the reflection-coefficient plane are circles given by (5.6a) and (5.66).
- 5.2. In the circuit shown in Fig. P5.2, what is the smallest value of d that will make the resistive part of Z_m equal to 50 SI at the plane t_1 ? Find the required value of jX to make Z_m equal to 50 SI at the plane t_2 . Use a Smith chart to find d and jX .



FIGURE P5.2

- 5.3. On a certain line terminated in a normalized load impedance $Z_L >$ it was found that the standing-wave ratio S was equal to 2 and that a voltage minimum occurred at $A/4$ from the load. What is Z_t ? Find the position and length of a single shunt stub that will match the load to the line. Use a Smith chart.
- 5.4. What are the length and position required for a series stub to match the load specified in Prob. 5.3? Use a Smith chart.
- 5.5. A normalized load $Z_L = 2$ terminates a transmission line. Does a voltage maximum or minimum occur at the load? Find the position and length of a shunt stub that will match the load to the line at a frequency f , where the wavelength is A . With the stub parameters and the load fixed, let the wavelength be increased to $1.1A$. What is the new value of stub susceptance and the standing-wave ratio on the line? If at the wavelength A , the stub is placed $A/2$ farther toward the generator from its original position, what is the standing-wave ratio on the line when the wavelength is increased to $1.1A$? This illustrates the greater frequency sensitivity of the match when the stub is placed farther from the load.
- 5.6. A double stub with spacing $0.25A$ is used to match a normalized load admittance $Y_L = 0.5 + j\frac{1}{2}$. Find the required stub susceptances by using a Smith chart.

Answer:

$$\text{Solution 1: } f_i = -0.5 \quad L_z = 0.176A \quad B_z = 1 \quad L > = 0.375A$$

$$\text{Solution 2: } B_1 = -1.5 \quad L_x = 0.0936A \quad B_2 = -1 \quad L_2 = 0.125A$$

- 5.7. Can a normalized load admittance $Y_L = 2.5 + j\frac{1}{2}$ be matched with a double-stub tuner with stub spacing $A/10$?
- 5.8. What is the minimum stub spacing in a double-stub tuner that will permit a normalized load admittance $Y_L = 1.5 + j\frac{1}{2}$ to be matched if the spacing is restricted to be greater than $A/4$? If the spacing must be less than $A/4$, what is the maximum stub spacing that can be used?

- 5.9. Show that, for a double-stub tuner with $d = A/4$, the required values of susceptances are given by

$$B >= -B_L \pm [G_L(l-G_L)]^{1/2}$$

$$\mathbf{B}_{\pm} = \frac{1}{l} \begin{pmatrix} -G_L & 1/2 \\ 1/2 & -G_L \end{pmatrix} \mathbf{G}_L$$

Hint: Take the limit of (5.20) and (5.21) as $\tan \phi$ becomes infinite.

- 5.10. A normalized load admittance $Y_L = 0.8 + j0.5$ is to be matched to a transmission line using a double-stub tuner. Use a Smith chart to find the required stub susceptances and lengths. Assume that short-circuited spaced by a distance $0.4A$ are used.

Answer:

Solution 1j $B_x = -3.171 \quad i_1 = 0.0486A \quad B_2 = -2.994 \quad L_2 = 0.0513A$

Solution 2: $B_x = -0.582 \quad L_1 = 0.166A \quad B_z = 0.2416 \quad L_2 = 0.288A$

- 5.11. A line with an attenuation constant $\alpha = 0.01$ Np/m is used as a short-circuited stub. Using the formula $Y_{ss} = Y_c \coth(\alpha + jB)l$, find the maximum value of normalized susceptance this stub can give. If $a = 0.02$, what is the maximum value of stub susceptance that can be obtained? Use $I = 1$ m.

- 5.12. Consider the series-shunt matching circuit illustrated in Fig. P5.12. What values of series stub reactance and shunt stub susceptance will match an arbitrary load $Y_L = G_L + jB_L$ a distance d away?

Hint: Note that Y_m at the position of the shunt stub equals $(jY - Z'_m Y')$, where Z'_m is the input impedance just to the right of the series stub. Next impose the condition $r'_m = I + jB$ to find the required value of jX .

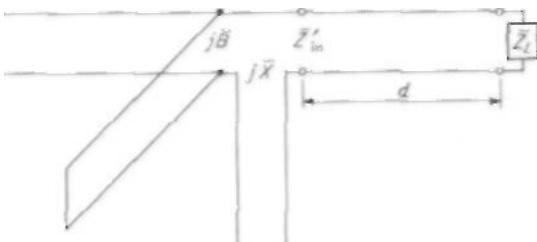


FIGURE P5.12

- 5.13. Develop a graphical method, using the Smith chart, to solve Problem 5.12.

Hint: V_m must lie on the $G = 1$ circle. But Y_m is obtained by going 180° around the chart from the Z_m point. Hence to make Z_m lie on the $G = 2$ circle, choose jX to make Z_m lie on the n through 180° .

- 5.14. A horn antenna is fed from a rectangular waveguide in which E_{TE010} mode propagates. At the junction a reflection coefficient $\rho = 0.5$, the normalized susceptance and the spacing in guide waveguide « junction of an inductive diaphragm that will match the horn.

$-JBt$

FIGURE P5.14

- 5.15.** A load impedance $40 \Omega j10$ is to be matched to a 50- Ω line using lumped elements. Use a Smith chart to find the parameters of a suitable matching circuit. Evaluate the loaded Q of the circuit. Two solutions are possible; choose the one with the lowest Q . Check your answers using LCMATCH.
- 5.16.** Use a Smith chart to design a lumped-parameter matching network that will match a load admittance $Y_L = 0.04 + j0.08$ to a 50- Ω transmission line. Find both solutions and evaluate the loaded Q for each. Check your answers using LCMATCH.
- 5.17.** For the matching circuit shown in Fig. 5.17«, show that the required values of jB_y and JX_c are given by

$$0x -j[-B_t \pm j/G]l -G_L))$$

$$Mi - \pm J$$

Hint: Note that $Z_m = JX_c + (JB_X + Y_h)$ must equal 1Ω .

- 5.18.** For the matching circuit shown in Fig. 5.176. show that the required values of jX_c and JB_y are given by

$$\wedge - / (- . X i i / B ^ 1 - R i))$$

$$JB_X = \pm j \frac{I - M l}{R L}$$

- 5.19.** Show that the definition (5.23) for Q gives the expressions in (5.26).
- 5.20.** A load admittance $G_L + jB_L = 0.02 + j0.05$ is to be matched to a 50- Ω transmission line using a double-stub tuner with short-circuited shunt stubs spaced by 0.18A. The stubs have a characteristic impedance of 40 Ω . Use the computer program STUBMACH to find the required stub lengths for the two possible solutions. Use the frequency scan option with a normalized frequency step of 0.02 and five steps below and above the matching frequency to evaluate the frequency sensitivity of the two solutions. Plot the return loss as a function of normalized frequency and find the bandwidth of each circuit for which the return loss is less than -20 dB.
- 5.21.** Repeat Prob. 5.20 but use open-circuited stubs. Compare the bandwidths obtained with those when short-circuited stubs are used.

- 5.22. A load impedance $Z_L = R_L + j\omega L = R_L + jX_L = 30 + j00 \Omega$ is to be matched to a 50- Ω transmission line using open-circuited series stubs apart. The stubs have a characteristic impedance of 75Ω . Use *STUBMACH* to find the required stub reactances. Carry out a frequency scan to determine the bandwidth over which VSWR remains below 1.5.

Answer: The two bandwidths are 0.045 and 0.087.

- 5.23 (a) For Prob. 5.22 replace the first stub by a short-circuited shunt. Does this improve the bandwidth? (b) Replace the second stub by a short-circuited shunt but keep the original first stub. How does this affect the bandwidth?

Answer: (a) BWs are 0.01 and 0.017. (b) BWs are 0.03 and 0.07.

- 5.24 A load impedance $Z_L = 5 + j50 \Omega$ is to be matched to a 50- Ω transmission line. Design a lumped-element matching network that uses a capacitor "in series with the load as the first element plus a second shunt element. Use computer program *LCMATCH*. Evaluate the return loss as a function of the normalized frequency using normalized frequency steps of 0.01. For the two solutions plot the return loss as a function of the normalized frequency. Find the 3-dB bandwidth for the two circuits. Evaluate the loaded Q_L for the two circuits and compare the bandwidth with that given by VQ_L .

Answer: $Q_L = 5$ and 6.5. The bandwidths are 0.21 and 0.16 and are close to the respective VQ_L values.

- 5.25. Repeat Prob. 5.24 but use a shunt element for the first element in the matching network (option b in the computer program *LCMATCH*).

- 5.26. A load admittance $Y_L = 0.002 + j0.1 \text{ S}$ is to be matched to a 50- Ω transmission line. Use a matching network with the first element being a shunt element. Evaluate the return loss as a function of the normalized frequency and find the 3-dB bandwidth of the two circuits. Evaluate the loaded Q_L and compare the bandwidth with that given by VQ_L . Use *LCMATCH*.

- 5.27. Repeat Prob. 5.26 but use a series element as the first element.

- 5.28. A load admittance $G_L = 0.002 - j0.05 \text{ S}$ is to be matched to a 50- Ω transmission line. Find solutions for four different matching networks that can be used. Evaluate the loaded Q_L of each and identify the network with the greatest bandwidth. Use *LCMATCH*.

- 5.29. A load impedance $Z_L = 80 \angle 10^\circ \Omega$ is to be matched to a 40- Ω transmission line. Find the elements for two possible matching networks and the corresponding bandwidths for which the VSWR remains below 1.5. Use *LCMATCH*.

- 5.30. A microwave amplifier requires a load termination $Z_L = 30 + j50 \Omega$. Use a Smith chart to design a matching network of the form $s_{ho} * s_{on} \text{ in series with } s_{in}$. The transmission line is an open-circuited one. The transmission coefficient is 50 JL .

- 5.31. Use a Smith chart to design a lumped-element network of the form $s_{ho} * s_{on} \text{ in series with } s_{in}$ that will provide a load $Z_L = 30 + j50 \Omega$ for an amplifier with a characteristic impedance of 50 Ω .

- 5.32. What are the required length and impedance of a quarter-wavelength transmission line that will match a 100- Ω load to a 50- Ω line? What is the frequency band of operation over which the reflection coefficient remains less than 0.1?

- 5.33. Design a two-section binomial transformer to match the load given in Prob. 5.32. What bandwidth is obtained for $p_m = 0.1$?
- 5.34. Design a three-section binomial transformer to match a 100-12 load to a oO-il line. The maximum VSWR that can be tolerated is 1.1. What bandwidth can be obtained? Plot $/>$ versus f .
- 5.35. Design a two-section Chebyshev transformer to match a 100-ii load to a 50-11 line. The maximum voltage standing-wave ratio (VSWR) that can be tolerated is 1.2. What bandwidth is obtained? Plot p versus f . Use the approximate theory.
- *g.36. Use the exact theory to design a two-section Chebyshev transformer to match a normalized load $Z_s = 5$ to a transmission line. The required fractional bandwidth is 0.6. What is the resultant value of $p_{,,}$?
- *g.37. Design a three-section Chebyshev transformer (exact theory) to match the load specified in Prob. 5.36. The same bandwidth is required. Compute $/>$, and note the improvement obtained.
- *g.38. Let $x_{,,}$ be a zero of $T_s(x)$; that is, $T^{\wedge}ixJ = 0$. Let the corresponding value of $\cos \theta$ be $\cos \theta_2$. Note that $T_{IV}(\sec H_m \cos \theta) = 0$ when $\sec f l_{,,} \cos " = \sec \theta_m \cos H$. Use the result

$$T_v(\cos />) = \cos Ntl>$$

to compute the zeros of ' $.,(v)$ '. Note that the zeros occur when $\cos 2\langle ft = 0$, or $\langle i \rangle = \pi/4, 3\pi/4$, and that $x_{,,} = \cos \langle />..$ Using these results together with the relation $\sec \theta_m = x_s \sec \langle ,,$ show that, for a two-section transformer,

$${}^P LR = 1 + k^2 Ti(sect)_m \cos(\theta)$$

reduces to (5.76). Note that when $H = 0$, $p^2 = (Z_L - Z_0)/(Z_L + Z_{,,})^2$, and hence

$$\begin{aligned} k^2 = & (Z_L - Z_0) \\ & 4Z_L Z_0 Ti(sect),,, \end{aligned}$$

- 5.39. For a particular application it is desired to obtain a reflection-coefficient characteristic

The normalized load impedance equals 2. What must the value of C be? Use the approximate theory for an JV-section transformer to design a four-section transformer that will approximate the above specified $/>,.$ Expand p_t in a Fourier cosine series to determine the coefficients $i>,,.$ Plot the resultant p , versus θ and compare with the specified characteristic. Show that the approximation to the specified $/>$ is a least-mean-square-error approximation. If the number of sections in the transformer were increased indefinitely, could the specified p , be realized?

- 5.40. An empty rectangular waveguide is to be matched to a dielectric-filled rectangular guide by means of an intermediate quarter-wave transformer. Find the length and dielectric constant of the matching section. Use $a = 2.5$ cm, $f = 10,000$ MHz, $e_{r2} = 2.56$. Plot $/;$, versus frequency f . Note that the appropriate impedances to use here are the wave impedances for the TE_{10} mode. ϵ_{r2} is the dielectric constant in the output guide.



FIGURE P5.40

- 5.41.** Design a two-section binomial transformer to match the empty dielectric-filled guide in Prob. 5.40. Use intermediate sections with constants ϵ_n and ϵ_{n+1} . Plot the input reflection coefficient as a function of frequency.
- 5.42.** Find the required impedances for a three-section half-wave filter having a fractional bandwidth of 0.3 and a passband tolerance $k^2 = 1.87 \times 10^{-3}$. The input and output transmission lines have a characteristic impedance of 50Ω .
- 5.43.** Find the required impedances for a two-section half-wave filter having a fractional bandwidth of 0.2 and a passband tolerance $k^2 = 10^{-2}$. The input transmission line has a characteristic impedance of 50Ω . What is the required characteristic impedance of the output transmission line? You will need to design the quarter-wave-transformer prototype circuit first.
- 5.44.** The filter described in Prob. 5.42 is to be built using microstrip construction on an alumina substrate ($\epsilon_r = 9.6$) 1 mm thick. The center frequency is 2 GHz. Determine the line widths and lengths including compensation for junction capacitance. Use the MSTP program.
- 5.45.** Obtain an expression for p , for a taper which has $d(\ln Z)/dz = C \sin(THZ/L)$. Determine C so that $\ln Z = \ln Z_0$ at $z = L$ and equals zero at $z = 0$. Plot p , versus pL .
- *5.46.** Design a taper that has double zeros at $0L = \pm 2\pi$ and $\pm 3\pi$. Plot F_{iu} versus u for this taper. Determine $\ln Z$ as a function of z . Achieve the design by moving the zeros at $\pm v$ and $\pm 4\pi$ into the points at $\pm 2\pi$, $\pm 3\pi$.
- *5.47.** Design a taper with single zeros at $j8L = \pm \pi$, $\pm 1.25\pi$, $\pm 1.5\pi$, $\pm 1.75\pi$, $\pm 2\pi$, $\pm 3\pi$, $\pm 4\pi$ into the specified points. Plot the resultant F_{iu} . Determine the expansion for $\ln Z$ as a function of z for the taper. Note that the close spacing of zeros keeps Δz small in the region $-2\pi < z < 2\pi$, but that removal of the zeros at $\pm 3\pi$ and $\pm 4\pi$ lets Δz take a large value in this region.
- *5.48.** Show that, for the exponential taper, the exact solution of the input reflection coefficient is

$$\frac{A \sin(BL/2)}{B \cos(BL/2) + 2jp \sin(BL/T)}$$

where $A = (\ln Z_L)/L$, $B = \sqrt{4/3^2 - A^2}$.

- *5.49.** Show that the approximate differential equation for the input along a slowly varying tapered line is $(Z_{in} = Z$ at all points)

$$dZ/dx = -2MZ - Z_{in}$$

Integrate this to obtain the input impedance to an exponential line \approx

Integrate this to obtain the input impedance to an exponentially tapered line $Z_{in} = Z_0 e^{-(x-x_0)/L}$

Quarter-wave transformers

1. Collin, R. E.: Theory and Design of Wide Band Multisection Quarter-Wave Transformers. *Proc. IRE*, vol. 43, pp. 179-185. February. 1955.
2. Cohn, S. B.: Optimum Design of Stepped Transmission Line Transformers. *IRE Trans.*, vol. MTT-3, pp. 16-21, April, 1955.
3. Riblet, H. J.: General Synthesis of Quarter-Wave Impedance Transformers, *IRE Trans.*, vol. MTT-5, pp. 36-48, January, 1957.
4. Young, L.: Tables for Cascaded Homogeneous Quarter-Wave Transformers, *IRE Trans.*, vol. MTT-7, pp. 233-237, April, 1959. See also *IRE Trans.*, vol. MTT-8, pp. 243-244, for corrections.
5. Solymar, L.: Some Notes on the Optimum Design of Stepped Transmission Line Transformers, *IRE Trans.*, vol. MTT-6, pp. 374-378. October. 1958.

Tapered lines

6. Bolinder, F.: Fourier Transforms in the Theory of Inhomogeneous Transmission Lines. *Proc. IRE*, vol. 38, p. 1354, November. 1950.
7. Klopfenstein, R. W.: A Transmission Line Taper of Improved Design, *Proc. IRE*, vol. 44, pp. 31-35, January. 1956.
8. Collin, R. E.: The Optimum Tapered Transmission Line Matching Section, *Proc. IRE*, vol. 44, pp. 539-548, April, 1956.
9. Bolinder, F.: Fourier Transforms in the Theory of Inhomogeneous Transmission Lines, *Trans. Roy. Inst. Techno!., Söckholm*, no. 48, 1951.
10. Matsumaru, K.: Reflection Coefficient of B-Plane Tapered Waveguides, *IRE Trans.*, vol. MTT-6, pp. 143-149, April, 1958.
11. Johnson, R. C.: Design of Linear Double Tapers in Rectangular Waveguides, *IRE Trans.*, vol. MTT-7, pp. 374-378. July, 1959.

CHAPTER 6

PASSIVE MICROWAVE DEVICES

A large variety of passive microwave circuit components or devices have been developed for use in both the laboratory and in microwave communication and radar systems. In this chapter we will describe the basic operating principles for a number of the most commonly used devices, such as attenuators, phase shifters, directional couplers, power dividers, hybrid junctions, and ferrite devices such as isolators or gyrators and circulator. The physical form that these devices take depends on whether waveguide or coaxial transmission line, strip line, microstrip, etc., is used as the transmission medium. The development of transistors and switching diodes over applications in the microwave frequency band has resulted in the development of a number of electronically controlled devices such as attenuators and phase shifters. Some of these electronically controlled circuit elements are also examined in this chapter.

6.1 TERMINATIONS

Two types of waveguide and transmission-line terminations are in use. One is the matched load, and the other is a variable short circuit that produces an adjustable reactive load. These terminations are often used in the laboratory when measuring the impedance

or scattering parameters of a microwave circuit element. The matched load is a termination that absorbs all the incident power, and hence is a short circuit. The circuit is a termination that reflects all the incident power. 1 <>



la:



(b)

FIGURE 6.1

Matched loads for a waveguide, (a) Lossy wedge: (b) tapered resistive card.

reflected wave is varied by changing the position of the short circuit, and this is equivalent to changing the reactance of the termination.

The usual matched load for a waveguide is a tapered wedge or slab of lossy material inserted into the guide, as illustrated in Fig. 6.1. Since the material is lossy, the incident power is absorbed. Reflections are avoided by tapering the *lossy* material into a wedge. Thus the termination may be viewed as a lossy tapered transmission line. An overall length of one or more wavelengths is usually sufficient to provide a matched load with an input standing-wave ratio of 1.01 or smaller.

Circuit

The simplest form of adjustable short circuit for use in a waveguide is a sliding block of copper *or some other good conductor* that makes a snug fit in the guide, as illustrated in Fig. 6.2. The position of the block is varied by means of a micrometer drive. This simple form of adjustable short circuit, however, is not very satisfactory in its electrical performance. The erratic-contact between the sliding block and the waveguide walls causes the equivalent electrical short-circuit position to deviate in a random fashion from the physical short-circuit position which is the front face of the sliding block. In addition, some power leakage past the block may occur, and this results in a reflection coefficient less than unity. These problems may be overcome by using a choke-type plunger, as discussed below.

The choke-type plunger is an example of the use of impedance transformation properties of a quarter-wave transformer. Consider, for example, a load impedance Z , that is approximately zero. If this impedance is

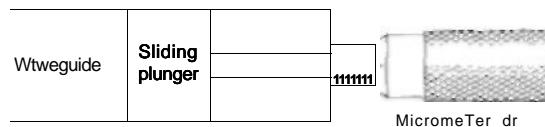


FIGURE 6.2

A simple short circuit for a waveguide.

$\times \gg / \parallel$ \times / \ll

FIGURE 6.3

Two-section quarter-wave transformer

viewed through a two-section quarter-wave transformer, as in Fig. 6.3, the input impedance seen at the input is

$$Z' = \frac{Z_1}{Z_2} \quad 6.1,$$

If Z_2 is chosen much greater than Z_1 , the new impedance Z' will approximate a short circuit by a factor $(Z_1/Z_2)^2$ better than Z_1 does. This essentially the principle used in choke-type plungers. The improvement factor of course deteriorates when the frequency is changed, so that the transformer sections are no longer a quarter wave long. However, by proper design, a bandwidth of 10 percent or more can be achieved. In very critical applications more than two sections may be used. The foregoing theory is now applied to a choke-type plunger for use in a rectangular guide.

For the TE₀₁ dominant mode in a rectangular guide, the surface currents on the interior wall flow up and down along the side walls and back across and in the axial direction on the broad walls. In the simple type of short circuit illustrated in Fig. 6.2, the axial current must flow across the gap between the upper and lower waveguide walls and vertically across the front face of the sliding block. The currents that flow along the side wall flow in the vertical direction and do not need to cross the gap between the waveguide walls and the front face of the plunger. Consequently, the performance arises only from the axial currents flowing on the upper and lower walls. To avoid this erratic behavior, the plunger may be made in the form illustrated in Fig. 6.4. The width of the plunger is uniform and slightly less than the interior guide width. However, the height of the plunger is made nonuniform. The front section is a quarter guide wavelength less than the guide height b by an amount 26% . The gap g_t is made as small as possible consistent with the requirement that the front section touch the upper and lower waveguide walls. The second section is a quarter guide wavelength long, but with the gap $b\%$ made

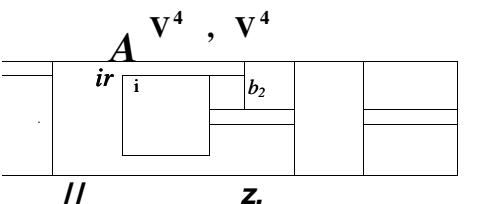


FIGURE 6.4
Choke-type adjuster
circuit (side view)

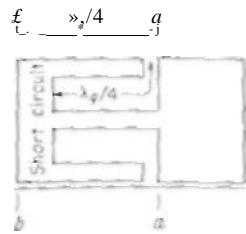


FIGURE 6.5
Alternative choke-type plunger design

possible consistent with maintaining the mechanical strength of the plunger. The final back section makes a sliding fit in the guide. The quarter-wave sections have equivalent characteristic impedances proportional to $26_1/6$ and $26_2/6$ relative to that of the input guide. Thus application of (6.11) gives

K

for the normalized input impedance. By good mechanical design 6_2 can be made about 10 times greater than b_x , and hence an improvement in performance by a factor of 100 over the non-choke-type short circuit may be achieved.

A somewhat different design, as illustrated in Fig. 6.5, is also frequently used. In this plunger a two-section folded quarter-wave transformer is used. The inner line transforms the short-circuit impedance to an ideal open circuit at the plane *aa*. At this point, i.e., at an open-circuit or infinite impedance point, the axial current is zero. Hence there is no current present to flow across the gap between the waveguide wall and the plunger at the contact point *aa*. The next, or outer, quarter-wave transformer transforms the open-circuit impedance at *aa* into a short-circuit impedance at the front end of the plunger, i.e., at plane *bb* in Fig. 6.5. Short-circuit plungers of this type give very satisfactory performance.

The above application of quarter-wave transformers is also used in the construction of *choke* joints for joining two waveguide sections together, in rotary joints, for plungers used to tune cavity resonators, etc.t

ATTENUATORS

Attenuators may be of the fixed or the variable type. The first is used only if a fixed amount of attenuation is to be provided. For bridge setups used to measure transmission coefficients, the variable attenuator is used. There are many ways of constructing a variable attenuator; only one type, the

rSee, for example. G. L. Kagan (ed.). "Microwave Transmission Circuits." McGraw-Hill Book Company, New York. 1948.

rotary attenuator, is considered in detail. A simple form of consists of a thin tapered resistive card, of the type used for mat t whose depth of penetration into the waveguide is adjustable Th^{at*} inserted into the guide through a longitudinal slot cut in the ce f* Card •» broad wall of a rectangular guide. An attenuator of this form ha^c complicated attenuation variation with depth of insertion and fr

A better precision type of attenuator utilizes an adjustable I^{enCy}, waveguide operated below its cutoff frequency. The disadvantages of this type of attenuator is that the output is attenuated by reducing the between the input and output guides, and not by absorption of the power. As a result, a high degree of attenuation corresponds to a refl coefficient near unity in the input guide, and this is often undesirable

Perhaps the most satisfactory precision attenuator developed fo AI • the rotary attenuator, which we now examine in some detail. The components of this instrument consist of two rectangular-to-circular wa\ guide tapered transitions, together with an intermediate section of circular waveguide that is free to rotate, as in Fig. 6.6. A thin tapered resistive card is placed at the output end of each transition section and oriented parallel to the broad walls of the rectangular guide. A similar resistive card is located in the intermediate circular-guide section. The incoming TE₀ mode in the rectangular guide is transformed into the TE_n mode in the circular guide with negligible reflection by means of the tapered transition. The polarization of the TE_n mode is such that the electric field is perpendicular to the thin res card in the transition section. As such, this resistive card has a negligible effect on the TE_n mode. Since the resistive card in the center section can be rotated, its orientation relative to the electric field of the incoming TE_n mode can be varied so that the amount by which this mode is attenuated: adjustable.

With reference to Fig. 6.7. let the center resistive card be orient'd at an angle θ relative to the direction of the electric field polarization of TE_n mode. The TE₀₀ mode polarized in the x direction may be decomp* into the sum of two TE₀₀ modes polarized along the u and v direct

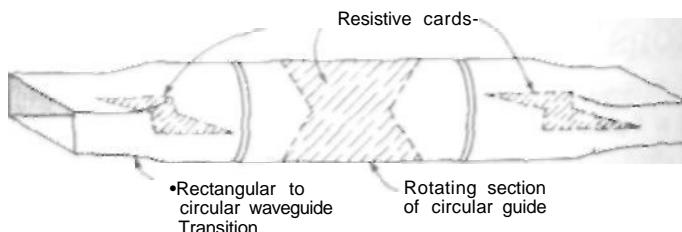
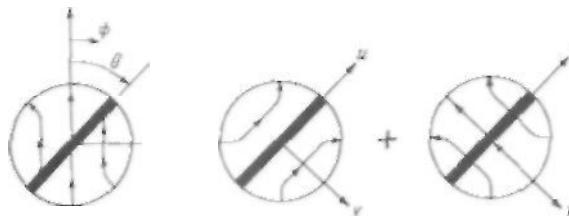


FIGURE 6.6
Basic construction of a rotary attenuator.



FICURK 6.7
Decomposition of TE_{11} mode into two orthogonally polarized modes.

illustrated in Fig. 6.7. That portion which is parallel to the resistive slab will be absorbed, whereas the portion which is polarized perpendicular to the slab will be transmitted. However, upon entering the transition section, the transmitted mode is again not perpendicular to the resistive card in this section; so some additional attenuation will occur.

To derive the expression for the dependence of the attenuation on the rotation angle θ , consider the analytic expression for the TE_{11} -mode electric field. For polarization in the x direction the electric field is given by (Table 3.6)

$$\mathbf{E} = \frac{J_1(p_i r/a)}{r} a_r \cos \theta \cdot \frac{P'u}{a} \hat{\mathbf{j}} + \frac{P'u^r}{a} \sin \theta \hat{\mathbf{i}} \quad (6.2)$$

apart from the propagation factor e^{jklz} . Since

$$\sin \theta = \sin(45^\circ - 0^\circ + 8^\circ) = \cos 8^\circ \sin(45^\circ - 0^\circ) + \sin 8^\circ \cos(45^\circ - 0^\circ)$$

and similarly

$$\cos \theta = \cos(45^\circ - 0^\circ + 0^\circ) = \cos 0^\circ \cos(45^\circ - 0^\circ) - \sin 0^\circ \sin(45^\circ - 0^\circ)$$

the above expression may be written as

$$\begin{aligned} \mathbf{E} = & \cos \theta \frac{J_1(p_i r/a)}{r} a_r \cos(45^\circ - 0^\circ) \frac{P'u}{a} \hat{\mathbf{j}} + J_1(p_i r/a) \sin(45^\circ - 0^\circ) \\ & - \sin \theta \frac{J_1(p_i r/a)}{r} a_r \sin(45^\circ - 0^\circ) \hat{\mathbf{i}} + J_1(p_i r/a) \cos(45^\circ - 0^\circ) \hat{\mathbf{k}} \end{aligned}$$

which is equivalent to referring the angle variable ϕ to a new origin at 8° . The first term in brackets in (6.3) is a TE_{11} mode polarized with the electric field along the u axis as in Fig. 6.7, and the second term in brackets is a TE_{11} mode polarized along the v axis. Since the first part is completely absorbed, only the portion multiplied by $\sin \theta$ is transmitted into the output transition section. If we assume that the resistive card in this section is parallel to the y axis, only the component of the transmitted field which is polarized along the v axis is transmitted. We have, at the input to the

transition section,

$$E = -\sin fl \frac{J}{a_r} \sin(\phi^* - d) + \frac{p'_{xl}}{A^{**}} \cos(\theta - \phi_0)$$

$$= \sin^2 \theta \frac{A}{a_r \cos \theta} \sin \phi$$

$$= \sin \theta \cos \phi - a_r \sin \phi + 1/3 \cos 4\phi$$

of which the first part is a TE₀₁ mode polarized along the x direction transmitted. Comparing (6.4) with (6.2) shows that the transmitted amplitude is reduced by a factor $\sin^2 H$ from the amplitude of the incident field H . The attenuation produced is given in decibels by

$$\alpha = -20 \log(\sin^2 \theta) = -40 \log(\sin fl) \quad /g_{5j}$$

A notable feature is that the attenuation depends only on the angle of rotation θ , a result that is verified in practice to a high degree of approximation.

Electronically Controlled Attenuators

For applications in various microwave systems, it is desirable to have an attenuator whose attenuation can be controlled by the application of a suitable signal, such as a dc voltage or a bias current. Two devices that are suitable for use in an electronically controlled attenuator are the PIN diode and a field-effect transistor. These devices can be used as variable resistors whose resistance is controlled by the applied signal.

The basic attenuator network is a symmetrical resistive T or Δ network as shown in Figs. 6.8a and 6.8b. The resistor values R_1 and R_2 are chosen so that when the attenuator is terminated in a resistance equal to the transmission-line characteristic impedance Z_c , the input is $Z_m = Z_c$, and to provide an output voltage reduction by

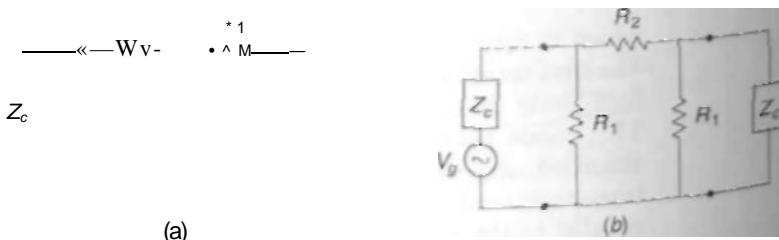


FIGURE 6.8
The two basic attenuator networks.

factor K . For the T network we have

$$B_m = \frac{R_2(R_l + Z_r)}{R_l + R_2 + Z_r}$$

For $R_{in} = Z_r$ we get the equation

$$R_l(R_l + 2R_2) = Z^*$$

When $R_m = Z_c$ the Thevenin impedance seen by the load equals Z , also. The Thevenin open-circuited voltage across R_2 is

$$V_{TM} = \frac{f_2}{R_l + R_2 + Z_c}$$

The power delivered to the load is

$$P_d = \frac{\frac{V_{TH}}{2Z}^2}{R_l + R_2 + Z_c} = \frac{f_2^2}{8Z}$$

The available power is IV^2C/SZ , so that the power attenuation K^2 is given by

$$K^2 = \frac{P_d}{IV^2C/SZ} = \frac{f_2^2}{R_l + R_2 + Z_c}$$

When K has been specified, and also requiring that $R_m = Z_r$, we have two equations that are readily solved for the required values of f_2 and R_2 . Thus we find that

$$R_x = \frac{1 - K}{2K} Z_r \quad (6.6a)$$

$$R_2 = \frac{1 - K^2}{2K^2} Z_r \quad (6.6b)$$

For 10-dB attenuation in a 50- Ω system, we require $K = 0.1$ which results in $R_x = 25.97$ ft and $R_2 = 35.14$ ft. For a 3-dB attenuator, we require $R_x = 8.58$ ft and $R_2 = 141.4$ ft. Thus, in order to produce a wide range of attenuation values, R_x and R_2 must be capable of being varied over a substantial range of values. For the II network, P_2 and R_2 are given by

$$P_2 = \frac{1 + K}{1 - K} P_1$$

$$R_2 = \frac{1 - K^2}{2K^2} Z_r \quad (6.76)$$

The P/iV diode differs from a conventional PN junction diode by having a thin layer of intrinsic semiconductor material between the usual Positive and Negative doped regions. The addition of the intrinsic region

reduces the junction capacitance, since the P and N regions apart. It also makes the forward conductivity of the diode a linear function of the diode bias current. In the forward bias charge carriers are injected into the intrinsic layer and thus the current is proportional to the injected charge, which in turn is proportional to the bias current. The equivalent circuit of the PIN diode is shown in Fig. 6.96. The resistance R_i and capacitance C_i represent the intrinsic resistance R_i , the junction capacitance C_j and the diffusion capacitance C_d represent the PN diode characteristics. The resistance R_p is the bulk semiconductor regions. The capacitance C_f is the stray capacitance between the terminals of the chip and can usually be neglected. For packaged diodes the equivalent circuit also includes the lead inductance and the package capacitance C_p .

Under the reverse bias conditions, the diffusion capacitance is negligible and the junction resistance R_j is very large. Also the parallel impedance of R_i and C_i is negligible relative to the impedance of C_d and R_p , so the equivalent circuit reduces to that shown in Fig. 6.96. A typical value

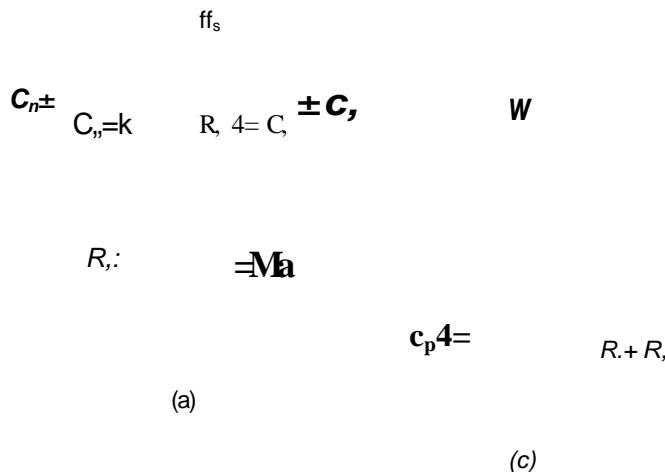


FIGURE 6.9

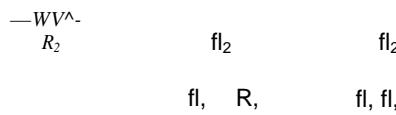
(a) Equivalent microwave circuit of a PIN diode; (b) equivalent circuit for reverse bias conditions; (c) equivalent circuit for forward bias conditions.

for the impedance of the diode in the off state is predominantly a capacitive reactance of $-j40$ to $-j400 \text{ f}\Omega$. For forward bias conditions, the diffusion capacitance C_d is large and provides a low-reactance shunt across the junction. Also C_s is small, so that $R_s + R_t$ becomes the controlling element for the diode impedance. The equivalent circuit under forward bias conditions is shown in Fig. 6.9c. In the fully on state, the resistance R_s of the intrinsic layer is less than 1 Ω in a typical diode. The resistance $/R_t$ varies inversely with the bias current.

PIN diodes in the forward biased state can be used for the resistors R_1 and R_2 in the attenuator networks shown in Fig. 6.8. The presence of reactive elements in the equivalent *PIN* diode circuit will result in some inevitable input and output mismatch. The dynamic range of an attenuator using three *PIN* diodes as variable resistors is also limited to usually 10 dB or less.

A more versatile attenuator can be built by using chip resistors to construct a cascade connection of Fl networks and using *PIN* diodes as switches to short-circuit the series arms for low attenuation or to be in the open state for one or more sections for various amounts of attenuation. An illustration of this type of attenuator is shown in Fig. 6.10. A basic section producing 2 dB of attenuation in a 50- Ω system would have $R_x = 436.2 \text{ }\Omega$ and $\#2 = 11.61 \text{ Cl}$. In the fully forward biased on state, the *PIN* diode short-circuits R_x , and, since the parallel combination of $/R_t$, with a second $/R_t$, is a resistance of 218.1 Ω , the loading effect on a 50- Ω line is relatively small, so that attenuation is also quite small.

A field-effect transistor with zero voltage applied between the drain and source has a channel whose resistance can be controlled by the gate voltage. Thus FETs can be used as the variable resistors in one of the basic attenuator circuits shown in Fig. 6.8. In Fig. 6.11 we show a photograph of a commercial electronically controlled attenuator built as a MMIC circuit. This attenuator will operate over the frequency range from dc to 26 GHz and can produce a maximum attenuation greater than 30 dB. The return loss is greater than 10 dB and the insertion loss is 1.5 to 2.5 dB. This particular device is suitable for use in a microstrip circuit. The attenuator uses six MESFET devices as variable resistors in a basic T attenuation circuit.

**FIGURE 6.10**

A switchable attenuator network. *PIN* diodes are used to switch R_t in and out of the circuit.

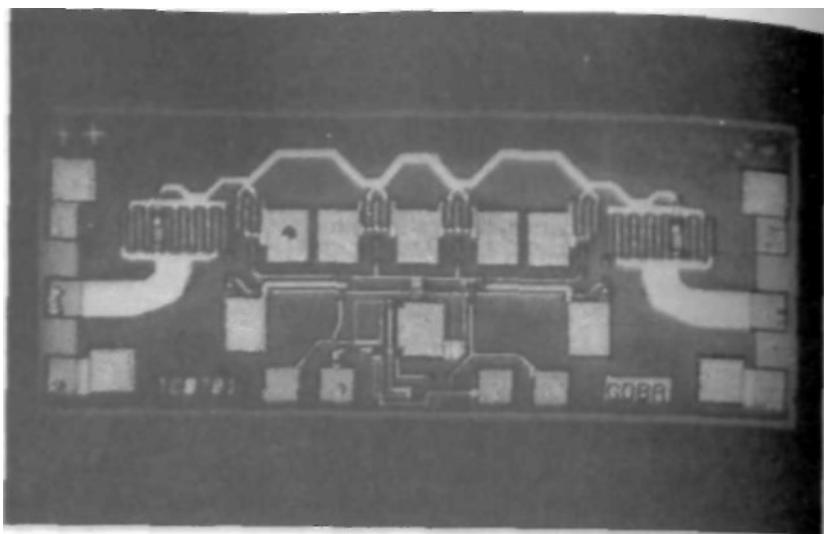


FIGURE 6.J1

A photograph of a MMIC electronically controlled attenuator. (Photograph courtesy of fo»
Moskaluk. Hewlett-Packard Company.)

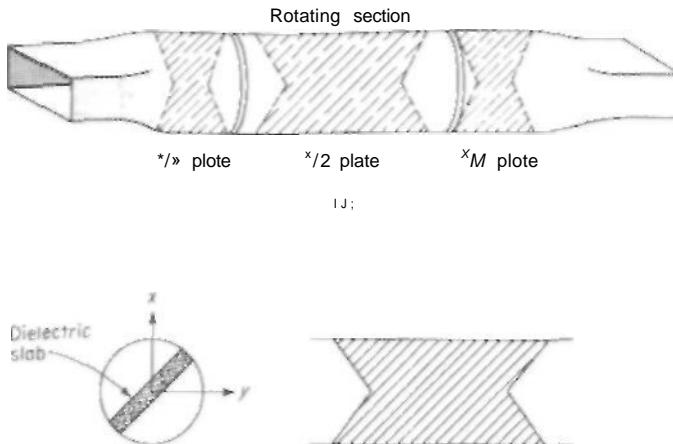
6.3 PHASE SHIFTERS

A phase shifter is an instrument that produces an adjustable change in the phase angle of the wave transmitted through it. Ideally, it should perfectly matched to the input and output lines and should produce zero attenuation. These requirements can be met to within a reasonable deg of approximation. There are a variety of designs for phase shifters mechanically adjustable type. The rotary phase shifter is the best in class and is the only one we will discuss.

Electronically controlled phase shifters using *PIN* diodes as swwj have become popular for use in phased-array antennas. In a pnas .^ antenna, there are many radiating elements, such as printed-circui ^ and the radiated beam can be scanned in direction by varying ⁱe J^ excitation of each element in the array. In a large array this requ , ^ be phase shifters; so a design that is small electronically controlled, techn0J- economically produced in large quantities using integrated cir nts. ogy is desirable. The *PIN* diode phase shifters meet these requir

Rotary Phase Shifter

The *rotary phase shifter* is a precision instrument that is v ^{J in}
microwave measurements. Its basic construction is similar ^{inly use o t ol}.

**FIGURE 6.12**(a) Rotary phase shifter; $l >$ quarter-wave plate.

rotary attenuator, except that the center resistive card is replaced by a half-wave plate and the two outer resistive cards are replaced by quarter-wave plates. The quarter-wave plates convert a linearly polarized TE_n mode into a circularly polarized mode, and vice versa. The half-wave plate produces a phase shift equal to twice the angle θ through which it is rotated. The analysis of the principles of operation is given below.

A circularly polarized field is a field with x and y components of electric field that are equal in magnitude but 90° apart in time phase.^f A quarter-wave plate is a device that will produce a circularly polarized wave when a linearly polarized wave is incident upon it. Figure 6.12 illustrates the basic components of the rotary phase shifter. The quarter-wave plate may be constructed from a slab of dielectric material, as illustrated in Fig. 6.12b. When the TE_n mode is polarized parallel to the slab, the propagation constant β_3 is greater than for the case when the mode is polarized perpendicular to the slab; that is, $\beta_3 > \beta_2$, where β_2 is the propagation constant for perpendicular polarization. The length l of a quarter-wave plate is chosen to obtain a differential phase change $(\theta_1 - i\beta_3 l)$ equal to 90° . The ends of the dielectric slab are tapered to reduce reflections to a negligible value. The half-wave plate is similar in construction, except that its length is increased to produce a differential phase change of 180° .

^fCircularly polarized fields are discussed in greater detail in Sec. 6.7.

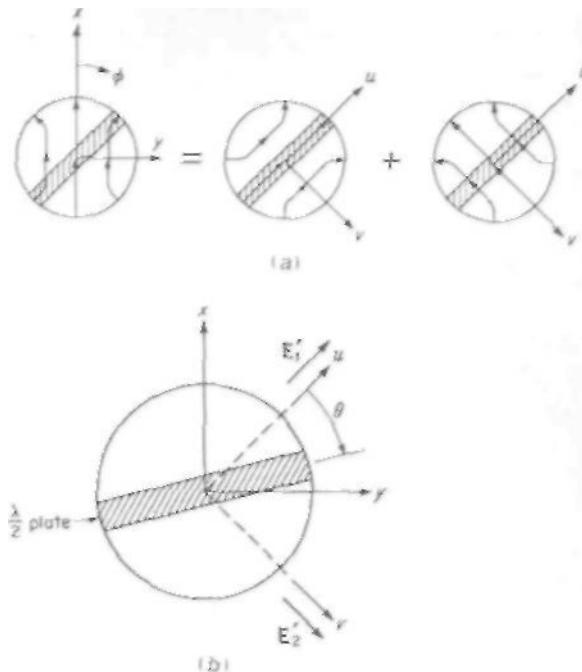


FIGURE 6.13

la) Decomposition of incident TE_n mode; (b) oriental; half-wave plate.

In the rotary phase shifter, the quarter-wave plates are oriented at an angle of 45° relative to the broad wall of the rectangular guide. The incoming TE_n mode may be decomposed into two modes polarized parallel and perpendicular to the quarter-wave plate, as illustrated in Fig. 6.13a. The incident mode is assumed given by (6.2) as

$$E = \frac{J}{a_r} \cos \phi \left(E_1 + E_2 \right)$$

If we replace $\cos \phi$ by

$$\cos(\phi - \pi/4) = -\cos(45^\circ) - \sin(45^\circ)$$

and $\sin \phi$ by

$$\frac{\sqrt{2}}{2} \left(\cos(\phi - \pi/4) + \sin(\phi - \pi/4) \right)$$

the above expression for the incident field may be written as

$$E = E_1 + E_2$$

where

$$E_1 = \frac{J_l}{2} \frac{I}{a_r \cos \phi} - \frac{T}{2} j - \frac{P_{11}}{2} J_{a,ft} \sin \phi$$

(6.8o)

(6-8°)

$$E_2 = \frac{J_l}{2} \frac{3}{a_r \sin \phi} - \frac{P_{11}}{2} J_{a,ft} \sin \phi$$

The field E_1 is polarized parallel to the slab, and E_2 is polarized perpendicular to the slab. After propagation through the quarter-wave plate, these fields become

$$E' = E_1 e^{-jJM} \quad (6.9c)$$

$$E'_2 = E_2 e^{-jw} = E_2 e^{-jis|t_e-juh-P|t} = JE e^{-J'''} \quad (6.9d)$$

since $(f_2 - i_x)l = -IT/I$. The resultant field consists of two orthogonally polarized TE_U modes of equal amplitude and 90° apart in time phase, and hence constitutes a circularly polarized field.

Consider next the action of the half-wave plate on the above circularly polarized field. Let the half-wave plate be rotated by an angle θ past the quarter-wave plate, as in Fig. 6.136. The field E' , + E'_2 may be expressed in terms of TE_U modes polarized parallel and perpendicular to the half-wave plate by changing the origin of the angle variable $\langle f \rangle$ to $7\pi/4 + \theta$; that is, we use the relations

$$\cos \langle f \rangle = \cos \text{rf} \quad = \cos ft \cos Ufr - \theta - j$$

$$- \sin \theta \sin j d \rangle - 9$$

and

$$\sin \langle f \rangle = \sin \theta \cos \langle f \rangle - \theta - \frac{\pi}{4} + \cos \theta \sin \phi - \theta -$$

Thus we obtain

$$E_i = \frac{Pn}{r} a_f \cos \left(\frac{I}{4} \right) - \theta - \frac{\pi}{4}$$

$$- \frac{Pn}{a} J J a_{,,} \sin U \sim \theta \sim \frac{\pi}{4} - \sin \theta - a_f \sin l r f - \theta -$$

$$+ \frac{Pn}{a} \cos U \sim Q - \frac{\pi}{4} \quad (6.10a)$$

$$E' = -j & \cos \theta \frac{I}{y} a_r \sin \theta - \frac{\pi}{4}$$

$$+ \frac{Pn}{a} \cos \left(\frac{I}{4} \right) - \theta - \frac{\pi}{4} + \sin \theta \cos \left(\frac{I}{4} \right) - \theta -$$

$$- \frac{Pn}{a} j a_f \sin \left(\frac{I}{4} \right) - \theta - \frac{\pi}{4} \quad (6.10b)$$

The field polarized parallel to the half-wave plate has an r component varying as $\cos(\phi - \theta - \pi/4)$, whereas the perpendicularly polarized mode has an r component of electric field varying as $\sin(\phi - \theta - \pi/4)$. Hence,

from (6.10), we obtain

$$E'_1 + E'_2 = Ej' + Eg$$

where

(6.1j)

$$-Jsintf) - a_r \cos \angle f - 6 - i)$$

$$- 8 - ?)$$

12

->S|' - 7«

(6.116)

$$\mathbf{B} \ll -\frac{iV2}{r} e^{-jPii-jo} - a_r \sin I \angle fi - 0 \frac{1}{4} + \frac{P_u}{a} j; a_c \cos \left(U - 6 - \frac{\pi}{4} \right)$$

(6.1k)

After propagating through the half-wave plate of length $2I$, this field becomes

$$E_3 = E'; e^{-2\pi i l}$$

$$E_{13} = E \angle e - W = Erge^{-2\pi i l} = -EJ$$

since $2(\hat{n}_2 - p_t)l = -TT$.

This new field may now be decomposed once again into two modes polarized parallel and perpendicular to the quarter-wave plate in output guide. If we assume that this plate is parallel to the input quarter wave plate, we can obtain the required decomposition by referring the i variable $\angle j$ to $ir/4$ as the origin. If we follow the procedure used earlier, obtain

$$E_3 * \ll, - E_a t K$$

, 613 < "

where

$$E' = \frac{V2}{2} e^{-3j0|i-r|} \frac{Ji}{r} a. \cos " 4J a$$

, 6-13*)

$$E'4 = \frac{j\sqrt{2}}{2} i - Wi' - m \frac{Ji}{r} a. \sin U - \frac{1}{4}; + \frac{-V}{a} a. \cos U - \frac{4}{J} Jate, t^{he}$$

Finally, after propagating through the second quarter output field becomes

, g

$$E_0 = E\% + ET_4$$

where

$$E''_s = tfcr*** \quad (6.146)$$

$$E; = E'je^{ - " ***} = jE_4e^{ - " i'l} \quad (6.14c)$$

When the fields E^\wedge and E^j are combined, we obtain

$$E_n = e^{-4JW} - W - 1 a_f \cos \angle b - - J, a,, \sin \angle \phi \quad (6.15)$$

which is again a linearly polarized TE_U mode having the same direction of polarization as the incident field given by (6.2) and (6.8). Note, however, that the phase has been changed by an amount $4\pi/3 + 2ft$. Thus rotation of the half-wave plate through an angle θ changes the phase of the transmitted wave by an amount 2θ .

I

This simple dependence of the phase change on a mechanical rotation is the chief advantage of the rotary phase-shifter.

Besides dielectric slabs, the circular guide may be loaded with metallic fins or rods to produce 90° and 180° differential phase-shift sections. These methods are discussed in a paper by Fox.⁺

Electronically Controlled Phase Shifters

There are several basic designs that are used to build digital-type electronically controlled phase shifters. In all designs the *PIN* diodes are used to switch circuit elements in and out of the transmission path. Each switching operation adds or subtracts a finite phase-shift increment such as $+11.25^\circ$, $\pm 22.5^\circ$, $+45^\circ$, $\pm 90^\circ$, etc. By using a cascade connection of several phase shifters, the full range 0° to 180° can be covered with a resolution equal to the smallest phase increment that is made available.

The simplest phase-shifter design uses *PIN* diode switches to switch one of two alternate transmission-line lengths into the transmission path as shown in Fig. 6.14. The bias currents can be supplied by connections at the midpoint of a half-wave open-circuited stub. The first quarter-wave section is a low-impedance stub and transforms the open-circuit impedance to a short circuit or low impedance at the point where the bias line is connected. The next quarter-wave section uses a high-impedance line and transforms the low impedance of the midpoint into a high impedance which produces negligible loading of the main transmission line. The dc return for the bias current can be obtained by connecting the input and output lines to the ground plane through short-circuited high-impedance quarter-wave-line sections as shown in Fig. 6.14.

If the two transmission-line sections have lengths l_x and l_2 , then the incremental phase change produced when line 2 is switched in to replace line 1 is $\Delta\phi = /3(l_2 - l_1)$, where $/3$ is the propagation phase constant. This

⁺A. G. Fox, An Adjustable Waveguide Phase Changer, *Proc. IRE*, vol. 35, pp. 1489-1498, December, 1947.

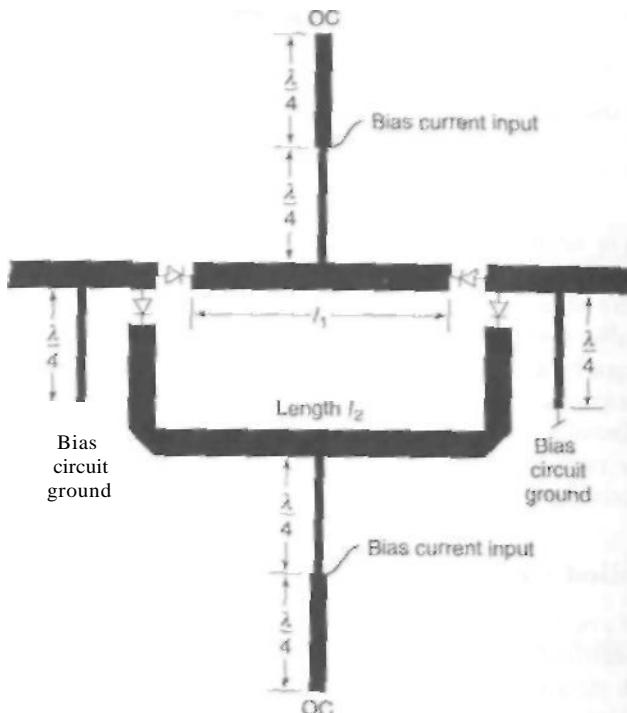


FIGURE 6.14
Incremental-line-type phase shifter.

type of phase shifter produces an incremental change in phase that depends on the frequency since p' is a function of frequency. If the dispersion of line is small, i.e., phase velocity is independent of frequency, then the increment in signal time delay is a constant and is given by $A = \frac{1}{v_p} = \frac{l_1}{v_p} = \frac{l_1}{c/\sqrt{\epsilon_r}} = \frac{l_1}{c} \sqrt{\epsilon_r}$. When broadband signals are to be radiated from a phased array, the use of different phases for the excitation of the elements in the array will generally result in some signal distortion. The phase shifters are true time-delay devices.

Since the PIN diodes are not ideal switches and have a finite resistance and some series lead inductance, the use of several phase shifters in series will result in a significant overall insertion loss.

A second type of phase shifter is shown in Fig. 6.15. In this type, susceptances of equal value jB are switched in shunt with the transmission line when the diodes are on. The stubs are used to obtain the required susceptances and are usually short-circuited at the ends to provide a dc return for the bias current. The stubs are spaced apart; so the phase shift is $\pi jB = B$ when the diodes are off. The network will still present an input and output impedance equal to the connecting transmission lines.

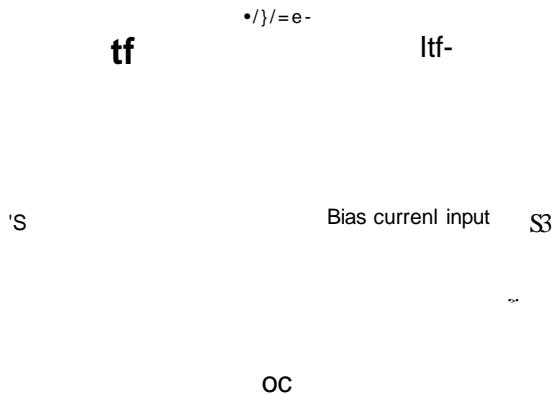


FIGURE 6.15
A phase shifter using switched reactive elements.

We can analyze this circuit in a straightforward way using wave-amplitude transmission matrices (Sec. 4.9). The wave amplitudes at various points in the circuit are designated by V_f , V_+ , V_2^* , $V.J$, etc., as shown in Fig. 6.15. The plus superscript designates the amplitudes of waves propagating to the right, while the minus superscript is used for the amplitudes of the waves propagating to the left. The wave-amplitude transmission matrices for a normalized shunt susceptance jB and a section of transmission line of electrical length θ are

$$[A_1] = \begin{matrix} B & & \\ & & B \\ B & & B \end{matrix} \quad [A_2] = \begin{matrix} J \\ & & \\ & & J \\ l \sim J_2 \end{matrix} \quad (1)$$

We can now write

$$\left| \begin{array}{c} \boxed{ } \\ \boxed{ } \\ \boxed{ } \end{array} \right| = AA \left| \begin{array}{c} \boxed{ } \\ \boxed{ } \\ \boxed{ } \end{array} \right| = [A_1][A_2] \left| \begin{array}{c} \boxed{ } \\ \boxed{ } \\ \boxed{ } \end{array} \right| = [A_1][A_2][A_3] \quad vr$$

$$\left| \begin{array}{cc} A_u & A_v \\ " 2 1 & A_2 2 \end{array} \right| \quad \left| \begin{array}{c} v \\ v \end{array} \right|$$

If we choose $V_+ = 0$, then $V \sim = V_f / A_n$; so the transmission coefficient from port 1 to port 4 is $T_u = A_u^x$. By multiplying out the matrix product shown above, it is readily found that

$$T_u = \frac{B^2}{1 + 7 \cdot e^{J\theta} - 1} e^{J\theta} + \frac{B^2}{e^{-J\theta}}$$

After some algebra we find that $|T_M| = 1$ provided we choose

$$\tan \theta = \frac{B}{B}$$

(6.16)

When we use this relationship to replace $B/2$ in the expression for obtain

$$T_u = -e^{j\theta} = e^{-j(T_0 - \theta)} \quad (6.17)$$

Thus the change in phase when the diodes are switched on is $(jr - 0) - e = 77 - 2^\circ$. For $B = 2$, $\theta = \pi/4$ and $A_0 = \pi/2$ or 90° , $B = 1$, $(V = 1.107 \text{ rad and } A_0 = 53.14^\circ)$. Thus, by choosing B appropriately, a wide range of phase shifts can be obtained.

To a first approximation the diode impedance in the on state is a small series inductive reactance. This should be included as part of the susceptance that is switched into the line. In the off state the diode impedance due to the shunt capacitance. Thus the shunt capacitance must be small if the phase shifter is to work properly. Since the off-state impedance is not always that large, an alternative design that allows some compensation for both the on-state and off-state diode impedance can be used. This alternate design is shown in Fig. 6.16. In this design two PIN diodes are used to open-circuit or short-circuit the tap points P and P' of two stubs that are connected in shunt to the main transmission line. The stubs are open-circuited at the remote end if shunt switches are used.

Since the condition (6.16) for a matched circuit cannot be simultaneously met for two different susceptances, this last design requires that the two susceptances jB_1 and jB_2 , that load the main line when the diodes are on and off be small. It is preferable to choose $jB_1 = -jB_2$ so as to maximize the phase difference in the transmission coefficient for the two states. Spacing the stubs $\lambda/4$ apart, the reflections from the two stubs almost cancel. The stub lengths d and tap point distance d , can be chosen such that $B_1 = -B_2$, including the diode reactances in the on and off states.

When ($I = \pi/2$) the use of wave transmission matrices shows

$$\mathbf{r}_{14} = \frac{B^2}{B^2 + 1} \quad \mathbf{B}^2$$

$$= -/(1+B)^2$$

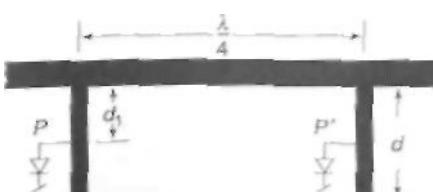


FIGURE 6.16
A phase shifter using two PIN diodes spaced $\lambda/4$ apart. P and P' are the tap points when the diodes are respectively. The bias circuit

when $B \ll 1$. Hence $|T_{u1}|^2 = d_+ B^2$ which is close to one. The phase angle of T_u is $-tr/2 - \tan^{-1} B = -jr/2 - B$. The change in phase between the states when $B = B_+$ and $-B_+$ is thus $2B_+$. A value of B_+ as large as 0.2 would not produce a significant mismatch and would result in a change in phase $\langle \rangle = 0.4 = 22.92^\circ$ between the two states. This type of phase shifter is limited to relatively small phase shifts between states in order to keep the input VSWR small.

A variety of other phase shifter circuits are also possible. A good summary of various circuits that have been analyzed and the performance that can be obtained can be found in the literature.

6.4 DIRECTIONAL COUPLERS

A directional coupler is a four-port microwave junction with the properties discussed below. With reference to Fig. 6.17, which is a schematic illustration of a directional coupler, the ideal directional coupler has the property that a wave incident in port 1 couples power into ports 2 and 3 but not into port 4. Similarly, power incident in port 4 couples into ports 2 and 3 but not into port 1. Thus ports 1 and 4 are uncoupled. For waves incident in port 2 or 3, the power is coupled into ports 1 and 4 only, so that ports 2 and 3 are also uncoupled. In addition, all four ports are matched. That is, if three ports are terminated in matched loads, the fourth port appears terminated in a matched load, and an incident wave in this port suffers no reflection.

Directional couplers are widely used in impedance bridges for microwave measurements and for power monitoring. For example, if a radar transmitter is connected to port 1, the antenna to port 2, a microwave crystal detector to port 3, and a matched load to port 4, the power received in port 3 is proportional to the power flowing from the transmitter to the antenna in the *forward* direction only. Since the reflected wave from the antenna, if it exists, is not coupled into port 3, the detector monitors the power output of the transmitter.

If the coupler is designed for 3-dB coupling, then it splits the input power in port 2 into equal powers in ports 2 and 3. Thus a 3-dB *directional* coupler serves as a power divider. Directional couplers with 3-dB coupling are also called hybrid junctions and are widely used in microwave mixers and as input, and output couplers in balanced microwave amplifier circuits. There are many available designs and configurations for directional couplers, hybrid junctions, and power dividers. Space limitations will allow us to only examine some of these.

t1. Bah) and P. Bhartia, "Microwave Solid State Circuit Design," -John Wiley & Sons. Inc., New York, 1988.

K. Chang (ed.), "Handbook of Microwave and Optical Components, Microwave Solid State Components," vol. 2. John Wiley & Sons, Inc., New York, 1990.

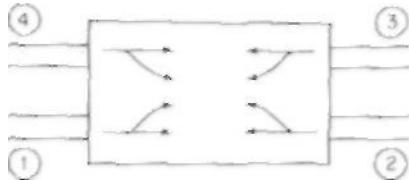


FIGURE 6.17

A directional coupler. Arrows indicate direction of power flow.

©
©



•n FIGURE 6.18

\wedge directional coupler.

Directional couplers using waveguides usually consist of two rectangular waveguides with coupling apertures located in a common wall as illustrated in Fig. 6.18. Since these devices are required to operate over a band of frequencies, it is not possible to obtain ideal performance over the whole frequency band. The performance of a directional coupler is measured by two parameters, the *coupling* and the *directivity*. Let P_i be the incident power in port 1, and let P_f be the coupled power in the forward direction in arm 3. The coupling in decibels is then given by

$$C = 10 \log \frac{P_f}{P_i} \quad (6.18)$$

Ideally, the power P_h coupled in the backward direction in arm 4 should be zero. The extent to which this is achieved is measured by the directivity D , which is defined as

$$D = 10 \log \frac{P_f}{P_h} \quad (6.19)$$

The directivity is a measure of how well the power can be coupled in the desired direction in the second waveguide. deduced

A number of properties of the ideal directional coupler may be deduced from the symmetry and unitary properties of its scattering matrix. The least stringent definition of a directional coupler as illustrated in Fig. 6.17 is that it is a four-port junction with

$$S_{14} = S_{23} = 0$$

$S_{11} = S_{22} = 0$, that is, ports 1 and 2 matched, and the coupling coefficients S_{13} , S_{14} , S_{23} , and S_{24} are not equal to zero. The scattering matrix is

A very good overall survey of types of directional couplers, hybrid junctions, and dividers can be found in K. Chang, "Handbook of Microwave Components," vol. 1, John Wiley & Sons, Inc., New York, 1990.

the form

$$\begin{aligned} \{S\} = & \begin{array}{cccc} 0 & 0(2) & 0[3] \\ S_{12} & 0 & 0 & \mathbf{S}_{24} \\ \mathbf{S}_{24} & \mathbf{0} & S_{33} & \mathbf{S}_{33} \\ 0 & S_{24} & 0 \cdot 3 & 1 \cdot 4 \cdot 1 \end{array} \end{aligned}$$

If we form the product of row 1 with the complex conjugate of row 3, and also the product of row 2 with the conjugate of row 4, we obtain

because of the unitary nature of the scattering matrix. Since S_{13} and S_{34} are assumed to be nonzero, these equations show that $S_{33} = S_{44} = 0$; that is, all four ports are matched. Thus the scattering matrix becomes

$$\{S\} = \begin{vmatrix} \mathbf{0} & S_{12} & S_{13} & \mathbf{0} \\ 0[2] & \mathbf{0} & \mathbf{0} & S_{2A} \\ S_{13} & \mathbf{0} & \mathbf{0} & S_{34} \\ \mathbf{0} & S_{24} & S_{34} & \mathbf{0} \end{vmatrix} \quad (6.20)$$

If we take the product of row 1 with the conjugate of row 4, and similarly row 2 with the conjugate of row 3, we now find that

$$S_{12}S_{24} + S_{13}S_{34} = 0 \quad S_{12}S_{24}^* + S_{13}S_{34}^* = 0$$

If we note that $I - S^* S I = |S_{12}| I S_{24} I$, these equations are seen to give

$$|S_{12}| I S_{24} I = |S_{13}| I S_{34} I \quad (6.21a)$$

$$I S_{12} I |S_{13}| = I S_{24} I |S_{34}| \quad (6.21b)$$

When we divide the first equation by the second equation, we obtain

$$\begin{aligned} & I S_{13} I \\ & |S_{13}| I S_{24} I \\ \text{or} \quad & |S_{13}| I = |S_{24}| I \end{aligned} \quad (6.22a)$$

thus the coupling between ports 1 and 3 is the same as that between ports 2 and 4. Use of (6.22a) in (6.21a) also gives

$$|S| I = |S|^* I \quad (6.22b)$$

so that the coupling between ports 1 and 2 equals that between ports 3 and 4 also.

The product of the first row with its conjugate equals unity, so that

$$|S_{12}|^2 + |S_{13}|^2 = 1 \quad (6.23a)$$

similarly,

$$I S_{12} I^2 + I S_{24} I^2 = 1 \quad (6.23b)$$

By choosing the terminal plane in arm 1 properly, we can adjust the phase

angle of S_{12} so that $S_{2,2}$ is real [see (4.53)]. Thus let S_{12} be a positive number C . Similarly, by choosing the terminal plane in arm 3 \hat{z} ^{postiv%} can make S_{13} a positive imaginary quantity jC_2 , where $C_2 > \text{real}$ ^{oper, y, we} _{aiKj}

$$|C_{11} + C_{22}| = 1 \quad (6.24)$$

We can choose the reference plane in arm 4 so as to make S_{33} equal to C , by virtue of (6.226). It is now necessary for $S_{2A} \neq 0$ since $S_{12}S_{2A} + S_{32}S_{3A} = 0$, as given earlier. Thus the simplest form of the scattering matrix of an ideal directional coupler is

$$[S] = \begin{vmatrix} 0 & C & JC_2 & 0 \\ C & 0 & 0 & JC_2 \\ jC_2 & 0 & 0 & C \\ 0 & JC_2 & C & 0 \end{vmatrix} \quad (6.25)$$

where $C_2 = (1 - Cf)^{1/2}$ from (6.24).

It may also be shown from the unitary properties of the scattering matrix of a lossless reciprocal four-port junction that if all four ports are matched, the device must be a directional coupler.

Directional-Coupler Designs

There are a great variety of ways of constructing directional couplers. Some of the more common aperture-coupled types are described below. Their design is based on the small-aperture-coupling theory presented in Sec. 4.13. This theory was originally developed by Bethe.¹

Bethe-Hole Coupler

The Bethe-hole directional coupler consists of two rectangular waveguides coupled by means of a small circular aperture located in the common broad wall. To achieve directional coupling, the axis of the aperture must be at an angle θ , as illustrated in Fig. 6.19a. A vertical aperture design consists of a similar arrangement, with $\theta = 0$, but an angle ϕ as in Fig. 6.19b.

The theory for the coupler in Fig. 6.19b was given in Sec. 4.13. An incident TE₁₀ mode in guide 1, with an amplitude A_1 , propagates along the electric dipole in the aperture plus a tangential magnetic dipole p and in the same direction as the magnetic field of the incident

¹C. G. Montgomery, R. H. Dicke, and E. M. Purcell, "Principles of Microwaves," McGraw-Hill Book Company, New York, 1948.

²H. A. Bethe, Theory of Diffraction by Small Holes, *Phys. Rev.*, vol. 66, pp.

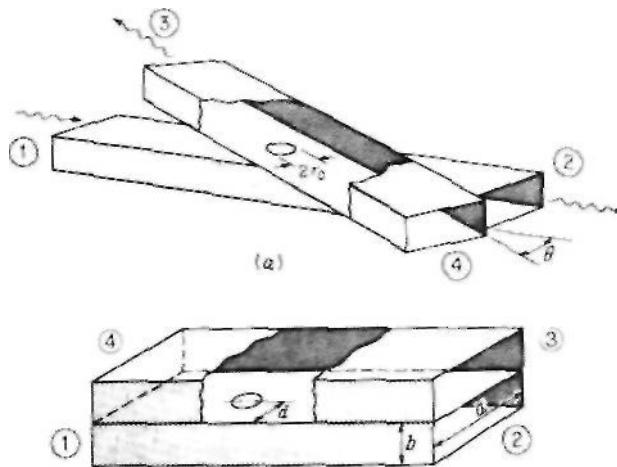


FIGURE 6.19
Bethe-hole directional coupler.

upper guide the normal electric dipole and the axial component of the magnetic dipole radiate symmetrically in both directions. The transverse component of the magnetic dipole radiates antisymmetrically. By varying the angle θ or the aperture position d , the amplitude of the fields coupled into ports 3 and 4 can be controlled. For the directional coupler shown in Fig. 6.19 a, the optimum value for the angle θ is given by

$$\cos \theta = \frac{2A^*}{2A^*_0} \quad (6.26)$$

This choice for θ will minimize the field coupled into port 4. Since the coupling is not zero, a perfect directional coupler is not obtained. A detailed analysis shows that the coupling and directivity that can be obtained are given by

$$C = 20 \log \frac{1}{X \cos \theta} \quad (6.27a)$$

$$D = C + 20 \log \frac{2 \cos \theta}{1 + \cos \theta} \quad (6.27b)$$

where $X = 16\pi r_0^2 / 3a^2 c A_{\text{eff}}$.

The directional coupler shown in Fig. 6.19b was analyzed in Sec. 4.13. When the spacing d from the side wall is chosen to satisfy the following

relation:

$$\sin \frac{vd}{9} / 6a \quad (6.28)$$

there will be zero power coupled into port 4. The coupling into port 4 with $1 + X^2$ replaced by $-J_1 + Y^2$ and $A_x + 4$, interchanged with $I + J$. Directional-coupler characteristics are also obtained by choosing $\sin \frac{vd}{8d}$ to satisfy the equation

$$\sin \frac{vd}{9} = \frac{y/2(K\% - a^2)}{1 + X^2} \quad (6.29)$$

When $A_0 = /2a$ the aperture will be located at the center, that is $d = 0$. For the above condition the field coupled into port 3 is minimized and that coupled into port 4 is a maximum. The coupling and directivity that can be achieved are given by [see (4.139) and (4.140) and let $X = -B$]

$$C = 20 \log \frac{|C|}{A_2 + A_4} = 20 \log \frac{1}{g} - \frac{+X^2}{1 + X^2} \quad (6.30a)$$

$$D = 20 \log \frac{A_2 + A_4}{A_1 + A_a} = 20 \log X^{-1} \quad (6.30b)$$

where $X = (167rro/3a6A)^{\sin^2(\text{irt}/a)}$. These formulas apply only at the design frequency.

The above results are based on the assumption that the guide wall in which the aperture is located is infinitely thin. For the normal thickness used in a waveguide wall, the coupling will be 1 to 2 dB smaller.

Example 6.1. Design of a Bethe-hole directional coupler. We design a directional coupler based on (6.28) and (6.30a). The waveguide dimensions are $a = 0.9$ in., $b = 0.4$ in. The center frequency is 9×10^9 Hz, which $0 = [\pi^2 - (7\pi/a)^2]^{1/2} = 1.29$ rad/cm. From (6.28) we obtain $\sin \frac{vd}{9} = 3.333/2.286/6$ which gives $d = 0.464$ cm. Thus the center of the aperture is located at 0.464 cm from the side of the waveguide. We require a 30-dB coupling. From (6.30a) this requires that $(1 + X^2) = 31.623$ or $X = 0.03164$. The required aperture radius is given by

$$\frac{3obA_gX}{16\pi \sin^2(7\pi d/a)} = 0.392 \text{ cm}$$

This is already a large aperture; so clearly it would be difficult to couple of 20 dB. The variation of the coupling and directivity as a function of frequency can be obtained by using the calculated values to

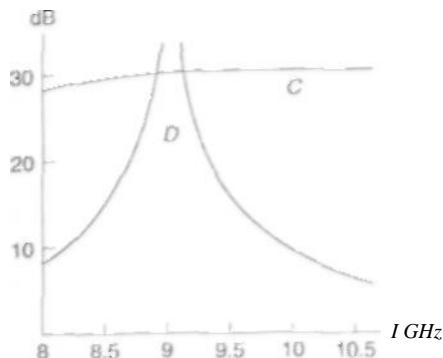


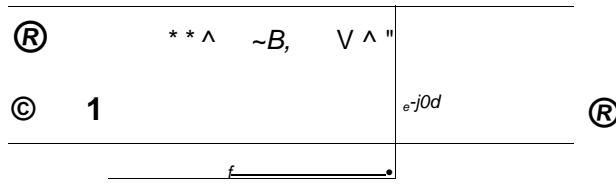
FIGURE 6.20

The coupling and *directivity* of the JSe! he-hole directional coupler as a function of frequency for the coupler in Example 6.1.

expressions For $A^+ + A_3$ and $A_- + A_4$ given by (4.140) and using these in (6.30). In Fig. 6.20 we show a plot of coupling C and directivity D as a Function of Frequency. The coupling remains almost constant. It increases From 28.46 dB at 8 GHz to 30.96 dB at 10.5 GHz. The directivity, however, drops rapidly as the frequency changes From 9 GHz. From these results we can see that the Bethe-hole coupler is a narrowband device. By using a second aperture on the opposite side of the center line the amplitudes oF the coupled waves are doubled. This will increase the coupling by 6 dB without affecting the directivity. Note that increased coupling is measured by a decrease in the coupling C when expressed in decibels by (6.18).

Hole Couplers

Two-hole couplers consist of two rectangular waveguides coupled by two identical apertures spaced a quarter guide wavelength $\lambda_g/4$ apart as in Fig. 6.21. The aperture may, in general, have directive properties, i.e., radiate a field with different amplitudes in the forward and reverse directions. With a wave of unit amplitude incident at port 1, let the field coupled into the second guide have an amplitude B_f in the forward direction and B_h in the backward direction. Since B_f and B_h are the amplitudes of the coupled fields for an incident wave of unit amplitude, they are called the aperture-coupling coefficients. If only a small amount of the incident power is coupled by the first aperture, the amplitude of the incident wave is essentially unity

FIGURE 6.21
Two-hole directional coupler.

at the second aperture also. Thus this aperture couples the same power into the second guide. Note, however, that because of the increase of path length, the phase of the field coupled by the second aperture is relative to that coupled by the first aperture. The total forward wave in the upper guide at the plane bb is $2B_f e^{-jpd}$ and the total backward wave in the plane aa is $B_b(l + e^{-2jpd})$. Hence, since the forward-path currents in the two guides are always the same, the forward waves in the two guides are always in phase. The backward waves will add out of phase whenever $2pd = n\pi$, $n = 1, 3, 5, \dots$. In particular, a value of $d = A_g/4$ will result in cancellation of the backward wave. The coupling is given by

$$C = -20 \log 2|B_f| \quad (6.31Q)$$

and the directivity is given by

$$\begin{aligned} D &= 20 \log \frac{21B_f}{BJI_1 + e^{-2jpd}} = 20 \log \frac{|B_f|}{B_h |\cos(3d)|} \\ &= 20 \log \frac{B_f}{B_h} + 20 \log |\sec(3d)| \end{aligned} \quad (6.316)$$

The directivity is the sum of the inherent directivity of the single aperture plus a directivity associated with the array (in this case a two-element array only). Since B_f and B_h are the aperture-coupling parameters and are generally slowly varying functions of frequency, the coupling C is not particularly frequency-sensitive. However, the directivity is a sensitive function of frequency because of the sensitivity of the array factor $\sec(3d)$.

Schwinger Reversed-Phase Coupler

The Schwinger reversed-phase coupler is designed to interchange the frequency sensitivity of the coupling C and directivity D . This is accomplished by making one aperture radiate a field which is the negative of that radiated by the other. With reference to Fig. 6.21, let the first aperture radiate B_f , and the second aperture $-B_h$. At plane bb in the PP^r_{noj} the total field is now $B_b - B_f = 0$ under all conditions. Hence port a is coupled to port 1. At the plane aa the total field is, after accounting for phase change due to propagation,

$$B_b - B_f e^{-2\pi d} = e^{j\pi d} B_b 2j \sin(id)$$

Thus the coupling between ports 1 and 4 is

$$C = -20 \log 2|B_b \sin(id)| \quad p \backslash s$$

and is a maximum for $d = A_g/4$. For this coupler the coupling coefficient is theoretically infinite and independent of frequency, whereas the directivity is quite frequency-sensitive, although not as frequency-sensitive as the directivity D given by (6.31*), since $\sin(id)$ varies more slowly than $\cos(3d)$. Actually, in practice, the directivity D is

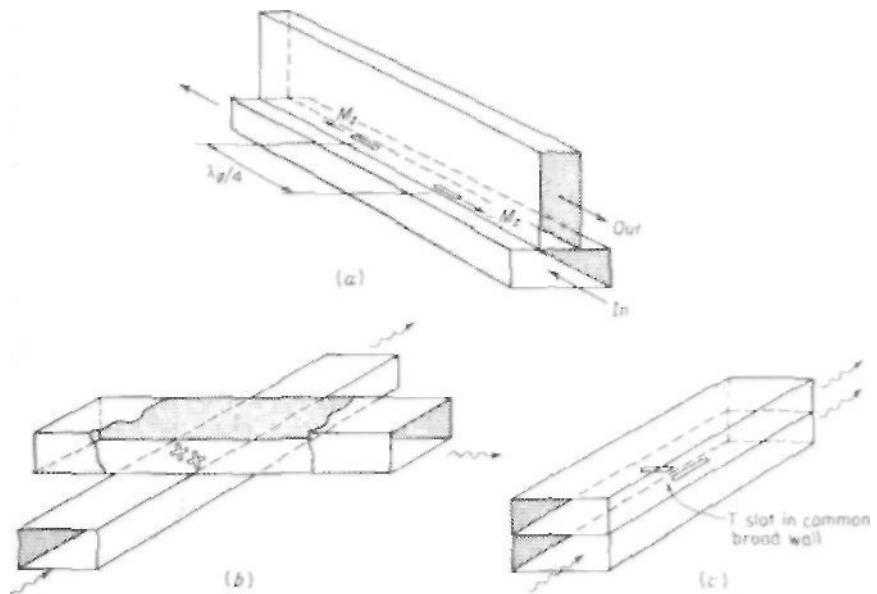


FIGURE 6.22

(a) Schwinger reversed-phase coupler; (b) Moreno crossed-guide coupler; (c) Riblet T-slot coupler.

in the foregoing discussion, it was assumed that the same incident field was present at each aperture, and each aperture radiated the same field into the upper guide. Because of interaction effects between the two apertures, the assumption of equal-amplitude fields coupled by both apertures is an approximation valid for a small amount of coupling only.

Figure 6.22a illustrates a typical reversed-phase coupler. The TE₀₀ mode has a zero normal electric field and transverse magnetic field at the narrow wall, and hence the coupling to this mode in the upper guide is through the induced axial magnetic dipole moment of the aperture only. In the lower guide the axial magnetic field of the TE₁₀ mode is of opposite sign on the two sides of the center, so that induced dipoles M_z and $-M_z$ are produced. These dipoles radiate symmetrically in both directions, but are phase-reversed to obtain the desired reversed-phase directional coupler.

The other double-aperture-coupled directional couplers in common use are the Moreno crossed-guide coupler and the Riblet T-slot coupler, illustrated in Fig. 6.22 also. Design nomograms for these couplers, as well as for the Schwinger reversed-phase coupler, are given in a paper by Anderson.^t

^tT. N. Anderson, Directional Coupler Design Nomograms, *Microwave J.*, vol. 2, pp. 34-38, May 1959.

Multielement Couplers

To achieve good directivity over a band of frequencies, course apertures may be used. The theory and design of such couplers are given for multisection quarter-wave transformers in Ch 6.2.3 illustrates an $N + 1$ element coupler with all aperture to d . If we assume that the total power coupled is relatively small with the incident power, the incident wave can be considered essentially the same amplitude A at each coupling aperture and an additional phase change. Let the apertures having coupling coefficients $n \sim 0, 1, 2, \dots, N$, in the forward direction, and D_n , $n = 0, 1, 2, \dots, N$, in the reverse direction. At the position of the N th aperture the total wave in the upper guide is

$$B_r = A e^{-\frac{N d}{Z}} \sum_{n=0}^N C_n \quad (6.33)$$

At the plane of the first aperture, the total backward wave has an amplitude

$$B_b = A Z \sum_{n=0}^N D_n e^{-\frac{n^2 \pi^2}{Z}} \quad (6.34)$$

The coupling and directivity are given by

$$C = -20 \log \left| \sum_{n=0}^N D_n e^{-\frac{n^2 \pi^2}{Z}} \right| \quad (6.35a)$$

$$D = -20 \log \left| \sum_{n=0}^N D_n e^{-\frac{n^2 \pi^2}{Z}} \right| = -C - 20 \log \left| \sum_{n=0}^N D_n \right| \quad (6.35b)$$

In a multiaperture directional coupler the required coupling from each aperture is small so the aperture radii are then also small. The corresponding aperture susceptance B and reactance X are then also small. The amplitudes of the waves coupled in the direction of ports 3 and 4 in Fig. 6.23, which are given by (4.140), can be approximated by

$$A_3 = j \frac{IB}{2} + \frac{X}{2} A$$

$$A_4 = j \frac{B}{2} - \frac{X}{2} A$$

$$D_0 A \quad C_0 A \quad D_1 A \quad C_1 A$$

$$D_N A \quad C_N A$$

$$V_-$$

FIGURE 6.23
A multielement directional coupler.

In terms of the notation being used in this section, we can thus write

$$C_{,,} = j \frac{B}{2} + X \quad (6.36a)$$

$$D_{,,} = J \frac{B}{2} - X \quad (6.366)$$

By using these expressions we obtain the advantage that for fixed aperture offsets both $C_{,,}$ and $D_{,,}$ are constants multiplied by the radius $r_{,,}$ cubed for the nth aperture. Hence we can express $C_{,,}$ and $D_{,,}$ in the form

$$C_{,,} = T_f r^* = T_f d_{,,} \quad (6.37a)$$

$$D_{,,} = 7V_n? = T_h d_{,,} \quad (6.376)$$

where $d_{,,} = \text{r.f.}$ Both T_f and T_h depend on the frequency.

We can now express the directivity D in the form

$$D = -C - 20 \log |T_f| - 20 \log \left| \sum_{i=0}^N \mathbf{d}_i \mathbf{e}^{-j2\pi f_i d_i} \right| \quad (6.38a)$$

and express C as

$$C = -20 \log |T_h| - 20 \log \left| \sum_{r=1}^{N/2} \mathbf{Z}_r \mathbf{d}_r \right| \quad (6.386)$$

In the expression for D the term $20 \log |T_f/T_h|$ gives the intrinsic directivity of a single aperture. This directivity is usually small except possibly at the design frequency. Thus, if we are to achieve a broadband design, we must design the array factor

$$F = \sum_{r=0}^M \mathbf{d}_r \mathbf{e}^{-j2\pi f_r d_r} \quad (6.39)$$

to give a high directivity over the frequency band of interest.

In order to obtain an equal-ripple characteristic in the passband, the array factor F is made proportional to a Chebyshev polynomial. If we choose a symmetrical array with $d_v = d_{v-1}$, $d_v = d_{v-2}$, etc., we obtain [Eq. (5.56)]

$$F = \sum_{n=-M}^M 2d_{,,} \cos(7V_n - 2n)pd \quad (6.40)$$

where $M = (N - 1)/2$ for N odd and $N/2$ for N even. Note that for N odd there are an even number of apertures, since the first aperture has been labeled the zeroth aperture. In (6.40) the M th term is d_M for N even and $2d_M \cos(N - 2M)\sin(\theta)$ for N odd. To obtain a Chebyshev-type response, we

now choose

$$F = \sum_{n=0}^M 2d_n \cos(N - 2n)\theta - K|T_N(\sec d_m \cos \theta)| \quad (6.41)$$

as in (5.68). In this equation $\theta = (id)$ and $\sec \theta_m$ is the value of $\sec \theta$ at the upper and lower edges of the passband. At the center of the passband, $\theta = \pi/2$, corresponding to a spacing $d = \lambda/4$. The constant $K^2 l_s^2 \Delta S_b \tan \theta_m$ give the desired value of coupling C in the center of the band, when $\theta = \pi/2$. Thus we obtain

$$\begin{aligned} C &= -20 \log \left| \sum_{n=0}^N y_n \right| \\ &= -20 \log K |T_f| |T_N(\sec \theta_m)| \end{aligned} \quad (6.42)$$

since $E \ll U = K|T_N(\sec \theta_m)|$ from (6.41).

J

If we use (6.38a), (6.41), and (6.42), the expression for directivity may be written as

$$\begin{aligned} D &= 20[\log K |T_f| |T_N(\sec \theta_m)| - \log |T_b| - \log |T_f/T_b(\sec \theta_m \cos \theta)|] \\ &= 20 \log \frac{T_f(\sec 0)}{T_N(\sec \theta_m \cos \theta)} \end{aligned} \quad (6.43)$$

Since T_f/T_b is a function of frequency, D will not have a Chebyshev-type behavior. However, the departure from a true Chebyshev behavior will usually be small since T_f/T_b gives very little directivity, except perhaps near the center of the band. For a conservative design we choose the minimum value of D on the basis that T_f/T_b contributes negligible directivity. Certainly, for a broadband design, this will be the case at the edges of the passband. The minimum value D_m of directivity in the passband as contributed by the array factor F occurs when

$$T_f(\sec 0, \cos 0) = 1$$

Hence let D_m be defined as

$$D_m = 20 \log i_{AR}(\sec 0_m) \quad (6.44)$$

This equation shows that if we specify D_m , then $\sec \theta_m$ is fixed, which in turn fixes the bandwidth, and vice versa. Thus we must specify either θ_m or $\sec \theta_m$, and the other is fixed. We may then solve (6.42) for the coupling C in terms of the given value of the coupling C at the center.

From (6.41) the coefficients d_n are found.

three-hole

We will outline the procedure by considering the design of a three-hole Chebyshev directional coupler utilizing offset circular apertures in a common broad wall between two rectangular waveguides as shown in Fig. 6.23. For the n th aperture the field coupled in the forward air gap is given by (6.36a). The field coupled in the n th aperture, where C_n is given by (6.36a), is

direction is given by D_n in (6.36)). By using (4.139) we get

$$\mathbf{C}_{,,} = \frac{2/e^2}{3/3a_6} \sin^2 \frac{\theta}{a} + \frac{0 \sin^2 \frac{JTX_0}{a}}{3ao} + \frac{77^2 \cos^2 \frac{-Xn}{3a}}{a} \quad (6.45a)$$

$$\mathbf{O}_{,,} = J \frac{2/e^2}{3/3a_6} \sin^2 \frac{\theta}{a} - \frac{77^2 \cos^2 \frac{-Xn}{3a}}{a} \quad (6.456)$$

We have used x_0 for the aperture position so as not to confuse this with the aperture spacing d. The factors multiplying r,f are T_f and T_h , respectively, and x_0 is the aperture offset measured from the waveguide side wall.

For a three-hole coupler,

$$T_N(\sec e_{,,}, \cos \theta_{,,}) - T_2(\sec \theta_{,,}, \cos \theta_{,,}) = 2 \sec^2 Q_{,,} \cos^2 0 - 1$$

From a specification of $D_{,,,}$ we can solve for d_m using

$$D_{,,} = 20 \log 17^2 (\sec 0_m) |$$

which for $n = 2$ gives

$$\sec H_{,,} = \frac{(10^{D_{,,}/20} + 1)^{1/2}}{2}$$

In general we can let $\sec 0_{ra} = \cosh \phi$ and use $(\cosh j) = \cosh Af$; thus

$$\cosh Nf = 10^{(E_{,,}/20)} = y \quad (6.46a)$$

$$\phi = \frac{1}{N} \cosh^{-1}(10^{E_{,,}/20}) = T7 \ln(V + \sqrt{y^2 - 1}) \quad (6.466)$$

$$Q_{,,} = \frac{1}{\cosh \phi} \quad (6.46c)$$

There will be two solutions for $\theta_{,,}$ one less than $77/2$ and one greater than $77/2$. These two values determine the values of fid = θ_m at the lower and upper edges of the passband.

When θ_m has been found, then if we specify the desired coupling $C = C_0$ at the center of the band we must have, from (6.42) with $\theta = 77/2$, $C_0 = -20 \log K|T_f| ITy \sec \theta_m |$ or

$$K = \frac{JQI - CO/20}{77^2/2 - 77^2/2} = \frac{1}{T_f |T_f| D_{20} \cos n}. \quad (6.47)$$

where $|T_f|$ is found from (6.45a) at the center of the band where fid = $TT/2$. After we have found the constant K , we express $T_N(\sec \theta_m \cos \theta)$ as a

Fourier series. For $N = 2$ we have

$$T_c(\sec e_m \cos \theta) = \sec^2 0_{m,} \cos 2\theta + \sec^2 B_m - i$$

From (6.41) we can then determine the $d_{..}$. For our specific p_{case} we

$$2d_0 \cos 2fld = 2d_0 \cos 2\theta = K \sec^2 B_m \cos 2\theta \\ d_{..} = K(\sec^2 0_m - 1)$$

Thus $d_Q = d_2 = (X/2)\sec^2 0_m$. Since $d_{..} = r_{..}^{-3}$ we get

$$r_0 \equiv *v = | \frac{K}{\sec^2 fl} |^{1/3}$$

$$r_{..} = (i \sec^2 0_m - if)^{1/3}$$

This completes the design of the directional coupler. The coupling and directivity as a function of frequency can be evaluated using (6.42) and requires the evaluation of T_b , T_b , and $7/(sec 0_m \cos \theta)$ at each frequency of interest. The following numerical example illustrates the performance that can be achieved.

Example 6.2 Three-Hole Chebyshev directional-coupler design. A three-hole directional coupler with a coupling of 20 dB and a minimum array factor directivity of 30 dB will be designed. The waveguide dimensions are $a = 0.9$ in, $b = 0.4$ in. The design frequency is 9 GHz. At 9 GHz the waveguide TE_{0,0} propagation constant is

$$\kappa = 0_{0,0} = \sqrt{\epsilon_r} = 1.29 \text{ rad/cm}$$

Thus the aperture spacing $< i$ is $\sim 2/(i) = 1.218$ cm. By using (6.46) we obtain

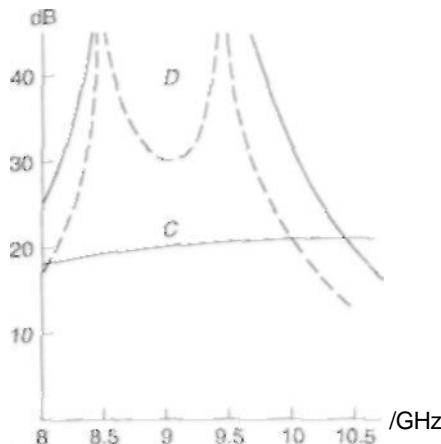
$$< b = \sqrt{a^2 + b^2} = 2.07338$$

and $o_m = 1.3206$ and $T_b = 1.3206 = 1.821$.

The values of $/3$ at the lower and upper band edges are given $(2fl_{..}/7r)0_0$ or 1.0845 and 1.4954. The corresponding values of $*o$ $(1.0845^2 + TTVO^2)^{1/2} = 1.7506$ and $(1.4954^2 + Tr'^2/a^2)^{1/2} = 2.03097$. From Δo we obtain the frequencies at the band edges and these are 8.358 GHz. The fractional bandwidth is $\Delta f/f_u = 0.149$.

The next step is the evaluation of K using (6.47). The parameter P_{param} is the magnitude of the coefficient of rf , in (6.45a) at the frequency ω_0 . We will choose an aperture offset given by (6.28), that is, $x_{ll} = 0.14954$ found then find that $K = 6.0244 \times 10^{-3}$. The aperture radii are now r_0 and are $r_0 = r_2 = 0.3663$ cm and $r_1 = 0.4518$ cm.

The coupling and directivity at any frequency is given by $-A_j^2$. In figure 6.24 we show the overall performance of the directional coupler. The $D(f)$ curve shows the directivity contributed by the array factor F_d . The $D(f)$ curve is greater because we chose the aperture offset such that the aperture directivity contributes significantly at 9 GHz; thus the aperture directivity contributes in a significant way.

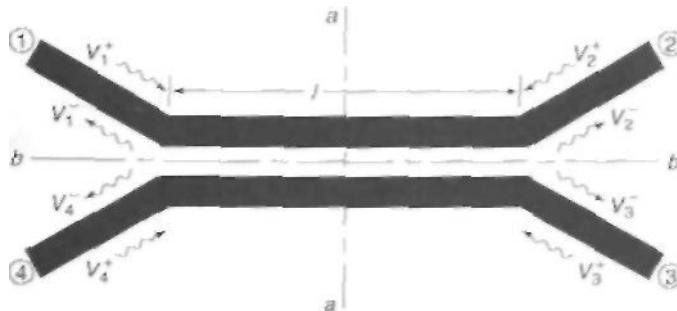
**FIGURE 6.24**

Performance of a three-hole Chebyshev directional coupler with $C = 20$ dB and $D_m = 30$ dBi. The dashed curve shows the directivity contributed by the array factor and is never less than 30 dB in the passband between 8.358 and 9.697 GHz.

overall directivity. If we used a second *set of apertures* spaced a distance x_0 from the opposite side wall, the coupling would be increased by 6 dB without changing the directivity. We could thus obtain a nominal coupling of 14 dB over the frequency band.

Coupled-Line Directional Couplers

Aperture-type directional couplers are not suitable for microstrip or strip-line construction. For planar-transmission-line structures, coupled transmission lines are frequently used for building directional couplers. In Fig. 6.25 we illustrate a microstrip directional coupler that involves two coupled microstrip lines. In practice, the printed circuit board would be housed in a shielded box and coaxial-transmission-line connectors would be bonded to each microstrip line. The analysis of the coupled-line directional coupler is readily carried out by taking advantage of the fourfold symmetry of the

**FIGURE 6.25**

A microstrip coupled-line directional coupler.

structure. We can choose excitations so that the symmetry plane corresponds to an electric wall (short circuit) or a magnetic wall (open circuit) and also so that the symmetry plane bb corresponds to an electric wall or a magnetic wall. When bb corresponds to an electric wall the propagation on the coupled line is the odd mode which has a characteristic impedance Z_o and propagation constant β_o . When bb corresponds to a magnetic wall, the mode of propagation is the even mode with characteristic impedance Z_e and propagation constant β_e . The constants are different because the effective dielectric constants for the modes are different.

We will consider the following four different excitations:

(a)

$$V+ = V_- = V_0 = V_4 = V^+$$

This case corresponds to both symmetry planes aa and 66 being magnet walls. For this case we only need to analyze the equivalent circuit of one-quarter of the structure as shown in Fig. 6.26a.

(b)

$$V_- = V_+ = V^* \quad V_- t = V_+ = -V^*$$

For this case the plane aa is an electric wall and the plane 66 is a magnetic wall. The equivalent circuit is shown in Fig. 6.26b.

(c)

$$V_- = -V_+ = V^* \quad V_+ = -V_- = V^+$$

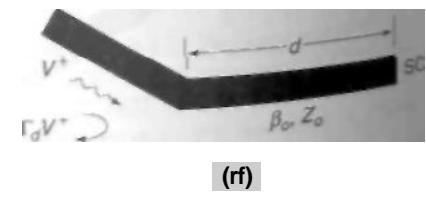
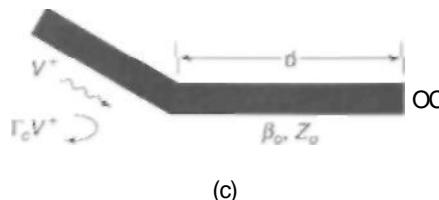
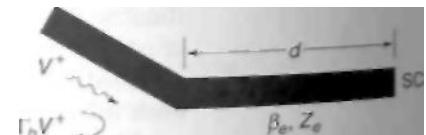
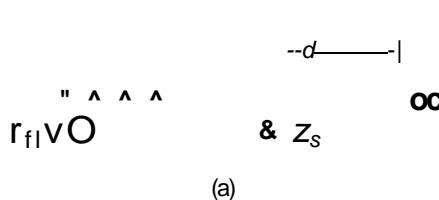


FIGURE 6.26

Equivalent circuit for one-quarter of the coupled-line directional coupler. aa and 66 are magnetic walls, (a) aa is an electric wall and 66 is a magnetic wall and 66 is an electric wall, (d) aa and 66 are both electric

when the P_{plan}^{**} is ∞ .

For this case the plane aa is a magnetic wall and the plane bb is an electric wall. The equivalent circuit for this case is shown in Fig. 6.26c.

$$(d) \quad v + \sim Vf = V \quad | \mathcal{E} = \sim i \mathcal{E} - \sim y +$$

For this case both symmetry planes correspond to electric walls and the equivalent circuit is that shown in Fig. 6.26d.

For case (a.) it is readily seen from the equivalent circuit in Fig. 6.26c that

$$\begin{aligned} Z_m &= -jZ_e \cot j\beta_e d \\ &= \frac{Z_{-n}}{Z_m + Z_{(+)}} = \frac{-jZ_e \cot l_e d}{Z_m + Z_{(+)}} = -jZ_e \cot M + Z_c \end{aligned}$$

where $Z_{(+)}$ is the characteristic impedance of the input microstrip line. From symmetry considerations we have

$$Vf = r_v v \quad V_2 = r_0 v \quad V_3 = r_v v^* \quad V_4 = r_v v$$

for this case. For case (6.) we have

$$\begin{aligned} &jZ_e \tan Brf - Z_c \\ &* \bullet - * - * - \bullet'' = j Z_c \wedge A + z, \end{aligned}$$

and

$$v_r = r_v v^+ \quad v_2 = -r^A \quad v_3 = -v_t y \quad v_r = v y$$

For case (c) we have

$$\begin{aligned} Z_{-n} &= -jZ_0 \cot fij \\ &\sim jZ_c \cot p_d d + Z_c \\ v_r &= r_e v^* \quad v_r = r_v v - y_3 = -r_v v^+ \quad v_r = -J_z v \end{aligned}$$

For the last case, namely case (d), we have

$$\begin{aligned} Z_m jZ_0 \tan ftd & \quad I_{rf} - J_{zotanM} + Z_r \\ *T = C_r y_r &= -r_{ff} v^* \quad V_3 = r_v v \bullet K_r = -r_v v' \end{aligned}$$

We now superimpose these four solutions. The superposition of the four incident waves at each port gives $V_j^* = 4V'$, $V'' = Vf = Y? = 0$. The only port 1 is excited. The superposition of the reflected waves gives

$$v_r - (r_a + r_c + r_e + r_f) V^+ = i(r_a + r_v + r_e + r_f) v \quad (6.48a)$$

$$V_a - fcr \cdot i t + r_v - r^A v_r \quad (6.48*)$$

$$v_r = 7(r_0 - r_6 - r_v) + r_d V_f \quad (6.48a)$$

$$\wedge -iO_i + i v_r - r_j v_z \quad (6.48a)$$

With some algebraic steps it is easy to show that

$$\mathbf{r}_s + \mathbf{r}_n = \frac{2(Z_e - Z_n t_o)}{jZ_c [Z_s \sim Z_n t_o] \sim Z_c^2 c e Z_n t_o}$$

where $t_{ss} = \tan \theta_0 d$ and $t_{nn} = \tan \beta_f d$. We also readily find that

$$\begin{aligned} \mathbf{r}_s + \mathbf{r}_n &= \frac{2(Z_c \wedge -Z_e Z_n, f_s)}{jZ_c r (Z_o - Z_n t_o) - ZX - Z_j \epsilon} \\ &\quad (6.496) \end{aligned}$$

For the ideal directional coupler, the two propagation constants should be equal, that is, $(t_{ss} = p_s = 0)$, in which case $t_e = 1$. For a coupled-line coupler, we can approach this ideal situation by placing a dielectric layer h made from the substrate material on top of the conducting strips. This surrounds the conductors with dielectric and will make the propagation constants equal to that for TEM waves in the solid dielectric. The use of a dielectric overlay will change the even- and odd-mode characteristic impedances but they will remain different. When $r_n = t_n$ we see from (6.49) that, provided we choose the line dimensions to make

$$Z_s = Z_l \quad (6.50)$$

$I_s + I_j = 0$ and $\Gamma_s + \Gamma_j = 0$. An examination of (6.48) now shows that $V_1 = 0$ and $V_3 = 0$. Thus there is no reflection in port 1 and no power coupled into port 3 at any frequency. Thus the four-port junction is a directional coupler. When $t_e = t_n = t$ and we make $Z_e Z_n = Z_s$, then we also have $r_u + l_A - \Gamma_s - Y_d = 2(|s| + V_h)$. The latter is given by

$$\begin{aligned} &2(Z_c^2 - Z_s^2) \sin 2\beta_f d \\ &/T_a + T_b \sim j Z_c Z_s (i - t^*) - (z_s^? + z_j^?)t \end{aligned}$$

This expression can be reduced to the following form:

$$\begin{aligned} \Gamma_s + W &= -2(Z_s - Z_n) \sin 2\beta_f d \\ &/j2Z_s \cos 2\beta_f d \sim (Z_s + Z_n) \omega x 2\beta_f d \end{aligned}$$

The transmission coefficient into port 4 is $\Gamma(r_a + T_A)$ and is the scattering matrix parameter $S_{41} = S_{ss}$; thus

$$\begin{aligned} S_{41} &= \frac{S_t^Z \wedge i \sin 2\beta_f d}{2Z_s \cos 2\beta_f d + j \sin 2\beta_f d} \\ &\quad (6.51) \\ &\quad Z_s + Z_n \end{aligned}$$

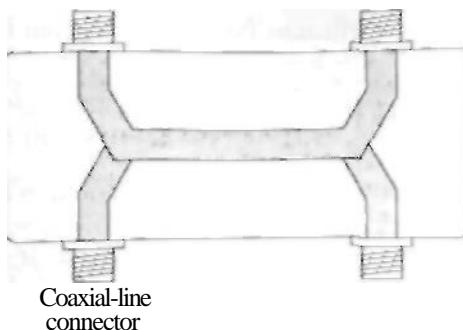


FIGURE 6.27
A strip-line coupled-line directional coupler

The parameter $(Z_e - Z_{ss})/(Z_e + Z_J)$ is the voltage-coupling parameter c . For this coupler the coupling is given by (note that $2c = I$)

$$\begin{aligned} C &= 20 \log \frac{[(1 - c^2)\cos^2/3] + \sin-p/}{C \sin fil} \\ &= 20 \log \frac{[1 - c^2 \cos^2/J?]^{1/2}}{c \sin jil} \end{aligned} \quad (6.52)$$

The maximum coupling occurs when $2fid = pi = TT/2$ which corresponds to a coupled-line one-quarter wavelength long. The maximum coupling is $20 \log(1/c)$. In a microstrip coupled line it is not practical to obtain a ratio of Z_e/Z_a greater than 2; so the maximum coupling is limited to a value of 9.5 dB or perhaps up to 8 dB with careful design. By using broadside coupled strips in a strip-line configuration as shown in Fig. 6.27, a coupling of 3 dB is readily achieved. Figure 6.28 shows a plot of $|S_{41}|$ as a function of $2fid = fil$ for a coupled-line directional coupler.

The directivity of this type of coupler is given by $V^{\wedge}V-J$ and is infinite since $V_{f\sim}$ is zero at all frequencies.

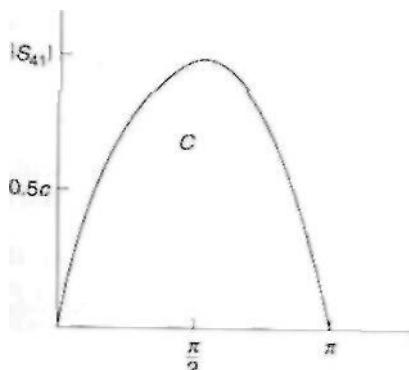


FIGURE 6.28
The variation of $|S_{41}|$ as a function of pi for coupled-line directional coupler.

The transmission coefficient into port 2 is given by

$$S_{2i} = \frac{V_1 - c^2 \cos \theta / +j \sin \theta}{V_1 + c^2 \cos \theta / +j \sin \theta} \quad (6.5g)$$

The symmetry of the structure requires $S_u = S_{22} = S_{3;i} = S_{4;i} = n_o$
 $S_{21} = S_{34} = S_{43} = S_{14} = S_{23} = S_{32}$ and $S_{24} = S_{34} = g$. Thus all scattering-matrix parameters are known. If the propagation constants are not equal we can still use the coupled-line directional coupler, but it will not have infinite directivity. The performance can be improved for the case ($i_e = t$) by adding a small shunt capacitor between the two coupled lines at the input and output to increase the bandwidth.

Branch-Line Directional Coupler

The branch-line directional coupler shown in Fig. 6.29 is readily fabricated using microstrip construction and can be designed for 3-dB coupling without any difficulty. The analysis of this coupler is also readily carried out using the fourfold symmetry that is inherent in the structure. With proper excitation the symmetry planes aa and bb can be made to correspond to either electric or magnetic walls. If excitations are chosen the same as for cases (a) to (d) in the coupled-line coupler, then the equivalent circuits for the four excitations are those shown in Figs. 6.30a to d. From these equivalent circuits we readily obtain

$$\begin{aligned} & m \quad \underline{Y_r \sim jY_{t1} - jY_{t2}} \\ & \circ \quad \underline{Y_r + jY_{t1} + jY_{t2}} \end{aligned} \quad (6.54Q)$$

$$h = \frac{\underline{Y_A + jY_{t1} - jY_{t2}tih}}{\underline{YJi - jYi + jY_2t_1t_2}} \quad (6.546)$$

$$r = \frac{\underline{Y_{t2} - jY_{t1}t_2 + jY_2}}{\underline{Y_{rt2} + jY_{t1}t_2 - jY_2}} \quad (6.54C)$$

$$r_{rl} = \frac{\underline{Y_{ct1t2} + jY_{lt2} + jY_{lt1}}}{\underline{Y_{ct1t2} - jY_{lt2} - jY_{lt1}}} \quad (6.54rfl)$$

where $t_x = \tan \theta_0 = \tan \theta_{jt1}$, and $t_2 = \tan \theta_2 = \tan p_2 d_2$ and A_{Jtbe} are the characteristic admittance of the input line, the through line,

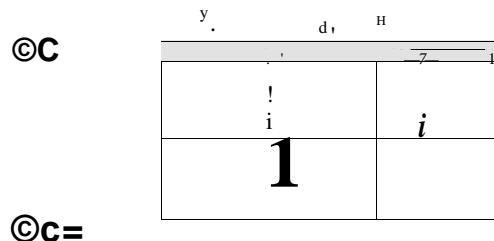


FIGURE 6.29
A branch-line directional coupler.

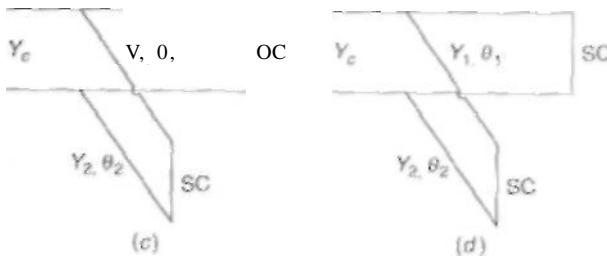
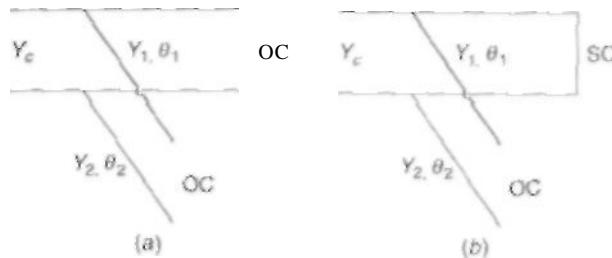


FIGURE 6.30
The equivalent circuit for one quarter section when (a) aa and bb are magnetic walls, (b) aa is an electric wall and bb is a magnetic wall, (c) aa is a magnetic wall and bb is an electric wall, (d) aa and bb are both electric walls.

branch line as shown in Fig.6.29. The relations (6.48) apply to the branch-line coupler also; so we have

$$S_{\infty} = \langle^j(r_a + i; + r_c + r_j) \quad (6.55a)$$

$$S_{12} = S_{21} = i(c r_i - r_{ft} + r_c - r_4) \quad (6.556)$$

$$S_{13} = S_{31} = i(r_a - r_6 - r_c + r_0) \quad (6.550)$$

$$S_{\infty} = S_u = H r_a + r_i - r_e - r_0 \quad (6.55*0)$$

If we choose $\epsilon_s = \epsilon_2 = 1$ so that the through lines and branch lines are one-quarter wavelength lines and also choose $K^2 - Kf = Y^*$, then we find that $S_{\infty} = S_{22} = S_{33} = S_{44} = 0$, $S_u = 0$, $S_{31} = i(r_o - i;)$ and $S_{zr} =$

$i(r_a + T_r)$. These latter expressions give

$$\frac{c}{31} = \frac{11}{v} \quad (6.56_a)$$

$$S_{zI} = -J_{\frac{y}{I}} f \quad (8.566,$$

for $t_x = I_2 = 1$. A 3-dB coupler is obtained if we choose $Y_1 = Y_2 = \sqrt{2} Y_c$, a condition that is easily satisfied. A 3-dB directional coupler with the two outputs 90° out of phase is also called a 90° hybrid. The coupling and directivity at any frequency are given by

$$C = 20 \log \frac{|I_{11}|}{|I_{12}|} = 20 \log \frac{\mathbf{i}_1 + \mathbf{r}_{1f} - \mathbf{r}_6 - \mathbf{g}}{\mathbf{i}_1} \quad (6.57Q,$$

$$D = 20 \log \left| \frac{\mathbf{r}_{11}}{\mathbf{r}_{12}} \right| = 20 \log \frac{\mathbf{r}_{11} + \mathbf{r}_d - \mathbf{r}_b - \mathbf{r}_1}{\mathbf{r}_{11} + \mathbf{n}_1 - \mathbf{r}_1} \quad (6.576)$$

The branch-line directional coupler is a relatively narrowband device. However, by cascading several sections a broadband coupler can be obtained by appropriate choices for Y_1 and Y_2 for each section. Design formulas are available in the literature.

Lange Directional Coupler^:

The last directional coupler that we will discuss is the Lange coupler shown in Fig. 6.31. This coupler uses several coupled lines in order to obtain larger coupling than what is possible in the simpler coupled-line coupler discussed earlier. The design of the coupler is such that wire connections between some of the lines are needed as shown in Fig. 6.31. This is the main shortcoming of the Lange coupler since such wire connections are readily made in an MMIC circuit. The outstanding features of the coupler are its compact size and very broadband characteristics. The coupler is often used as an input coupler in balanced microwave circuits. For this application it is designed as a 3-dB coupler and the output signals are in phase quadrature, so that it is a 90° hybrid junction. Design formulas for a Lange coupler have been developed by "re-

tG. L. Matthaei, L. Young, and E. M. T. Jones, "Microwave Filters, Networks and Coupling Structures," Artech House Books, Dedham, Mass.,

±J. Lange, Interdigitated Stripline Quadrature Hybrid, *IEEE Trans.*,

1150-1151, December, 1969. rrr'6 PP 80~,

§A. Presser, Interdigitated Microstrip Coupler Design, *IEEE Trans.*, vol. N October, 1978.

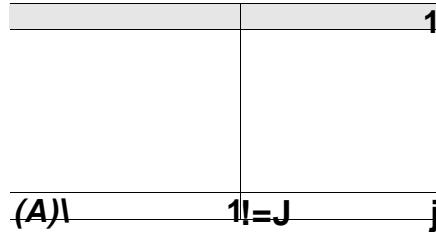


FIGURE 6.31
Tilt' Lange directional coupler.

6.5 HYBRID JUNCTIONS

Magic T

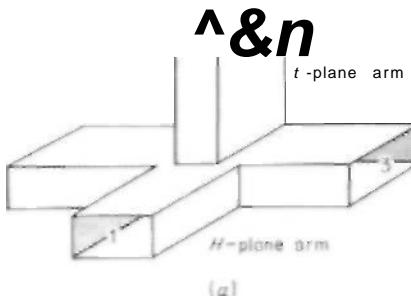
A waveguide hybrid junction, known as a magic T, is illustrated in Fig. 6.32. When a TE_{10} mode is incident in port 1, the electric field within the junction is like that sketched in Fig. 6.32b. This electric field has even symmetry about the midplane and hence cannot excite the TE_{10} mode in arm 4 since this mode must have an electric field with odd symmetry (shown dashed in Fig. 6.32b). Thus there is no coupling between ports 1 and 4. The coupling between ports 1 and 2, and 1 and 3, is clearly the same, as may be seen from the symmetry involved.

For a TE_{10} mode incident in arm 4, the electric field within the junction is sketched in Fig. 6.32c. Symmetry again shows that there is no coupling into port 1 (this is required by reciprocity as well). The coupling from port 4 into ports 2 and 3 is equal in magnitude but 180° out of phase. The scattering matrix of this hybrid T thus has the form

$$[S] = \begin{vmatrix} S_{11} & ^{12} & S_{12} & 0 \\ ^{12} & S_{22} & S_{21}^* & -S_{24} \\ S_{12} & S_{21} & S_{33} & -S_{24} \\ 0 & -S_{24} & -S_{24} & S_{44} \end{vmatrix}$$

since $S_{13} = S_{13}$, $S_{00} = -S_{34}$, from symmetry.

Matching elements that do not destroy the symmetry of the junction may be placed in the E-plane and // -plane arms so as to make $S_{11} = S_{44} = 0$. For a lossless structure we may then show that the unitary properties of the scattering matrix require that $S_{12} = S_{31} = 0$, so that all ports are matched. In addition, $S_{23} = 0$; so ports 2 and 3 as well as ports 1 and 4 are uncoupled. The hybrid T now becomes a directional coupler with 3-dB coupling, and is often called a *magic T*, even though there is nothing magic about its

**FIGURE 6.32**

<a> Hybrid-T junction; (**M d f c**
trie field pattern for wave b
5 den in P^{or}, 1; **c**) electric field
pattern for wave incident g
port 4.

operation. The magic T is commonly used in waveguide balanced mixers and in bridge networks.

With $S'_{nn} = S_{44} = 0$, the scattering matrix becomes

$$[S] = \begin{bmatrix} \mathbf{0} & S_{12} & ^{12} & \mathbf{0} \\ S_{VZ} & S_{22} & ^{23} & *^{24} \\ S_{L2} & ^{23} & S_{33} & -*^{24} \\ \mathbf{0} & S_{24} & -S_{24} & \end{bmatrix}$$

The product of the second row with its conjugate gives

$$|S_{12}|^2 + |S_{22}|^2 + |S_{33}|^2 + |S_{24}|^2 = 1 \quad (6.58a)$$

and the similar expression for row 3 is

$$|S_{13}|^2 + |S_{23}|^2 + |S_{33}|^2 = 1 \quad (6.58b)$$

If we subtract these two equations, we obtain

$$|S_{12}|^2 - |S_{23}|^2 = 0 \quad (6.58c)$$

so $|S_{22}|^2 = |S_{33}|^2$. From rows 1 and 4 we have

$$2|S_{12}|^2 = 1 \quad \text{or} \quad |S_{12}|^2 = \frac{1}{2}$$

$$2|S_{23}|^2 = 1 \quad \text{or} \quad |S_{23}|^2 = \frac{1}{2}$$

$$\mathbf{IS}, \mathbf{J} =$$

and thus

$$|S_{12}I - IS24I - \frac{ft}{z}| \quad (6.59)$$

Use of this relation in (6.58a) gives

$$I + |S_{23}f + ISJ^I| = I$$

or $ISw_i + IS23I^2 = 0$. This sum can equal zero only if both S_{TZ} and S^{TM} vanish. From the relation (6.58) it follows that S_{ext}^* equals zero also.

The reduced form of the scattering matrix becomes

$$[\ast] - \begin{vmatrix} 0 & S_{12} & S_{12} & \mathbf{0} \\ S_U & \mathbf{0} & \mathbf{0} & \\ S_{Ia} & \mathbf{0} & \mathbf{0} & S_{T24} \\ \mathbf{0} & S_{24} & -\ast \ll & 11 \end{vmatrix}$$

By proper choice of terminal planes in arms 1 and 4, we can make both S_{Vi} and $S_{2..}$, real. Thus the scattering matrix of a magic T can be exhibited in the form

$$\begin{matrix} \mathbf{ft} & \left| \begin{matrix} \mathbf{0} & 1 & 1 & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & 1 \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 \\ \mathbf{0} & \mathbf{I} & -1 & \mathbf{0} \end{matrix} \right. \end{matrix} \quad 16.60)$$

upon using the relations (6.59).

The branch-line coupler designed for 3-dB coupling is a 90° hybrid junction. The magic T is a 0° or 180° hybrid junction since the two outputs are in phase if port 1 is the input port and are 180° out of phase if port 3 is the input port. A 180° hybrid junction that is readily made using microstrip construction is shown in Fig. 6.33. To understand its operation, consider a wave incident in port 1. This wave splits equally into two waves traveling around the ring circuit in opposite directions. The two waves will arrive in phase at ports 2 and 4 and out of phase at port 3. Thus ports 1 and 3 are uncoupled. Similarly, ports 2 and 4 are uncoupled since the two paths coupling these ports differ in length by $A/2$.

A quantitative analysis of the hybrid ring is readily carried out. Each input line has a characteristic impedance Z , and the ring has a characteristic impedance Z_r . We will let θ_r be the electrical length of the ring between ports 1 and 2, 2 and 3 and 3 and 4, while $\theta_r = 3ft$, represents the electrical length of the ring between ports 1 and 4. We can choose excitations such that the symmetry plane aa is either an electric wall or a magnetic wall. This will allow us to characterize one-half of the structure in terms of

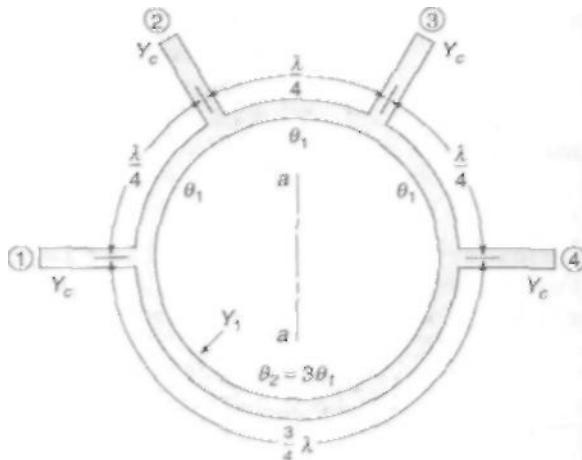


FIGURE 6.33
The hybrid ring ("rat-race"),

scattering matrices for a two-port junction, ports 1 and 2. From the symmetry properties the complete four-port scattering matrix can be constructed.

Let ports 1 and 2 have incident waves V_{in} and V_{out}^T . Also let ports 4 and 3 have incident waves V^* and V_A^+ , respectively. For this symmetrical excitation the electric field must be a maximum on the symmetry plane aa and the magnetic field must be zero. Thus aa can be replaced by a magnetic wall, i.e., the ring is open circuited on the plane aa . The equivalent circuit of half of the structure is shown in Fig. 6.34a. It consists of two input lines with characteristic impedance Z_v and an interconnecting line of electric length $\lambda/2$ and characteristic impedance Z_v . The input port 1 is shunted by a susceptance $jB_1 = jY_1 \tan 30^\circ/2$ due to the open-circuited section of the



$$\begin{array}{ll} V_i & V_o \\ V_i & V_o \\ V_i & V_o \\ V_i & V_o \end{array}$$

FIGURE 6.34
one-half
(a) Equivalent circuit of hybrid ring for $0 \leq \theta \leq \pi/2$
ditions on the symmetry plane
equivalent circuit $J \rightarrow p \wedge J J$

two sides for one-half of
the structure.

hybrid ring having an electrical length $30_0/2$. The output port 2 is shunted by a susceptance $jB_2 = jY_l \tan 0_0/2$ due to the open-circuited section of the ring of electrical length $0_0/2$. We will let the two-port scattering-matrix parameters under open-circuit conditions on the plane aa be S_{ff} . Thus we can write

$$\left| \begin{array}{c} \mathbf{V} \\ \wedge \end{array} \right| = \left| \begin{array}{cc} \mathbf{S} \mathbf{n} & \stackrel{\text{co-c}}{\circ} \\ \circ_2^{\text{oc}} & \circ_2^{\text{oc}} \end{array} \right| \left| \begin{array}{c} \mathbf{v} \\ \mathbf{y} \end{array} \right| \quad (6.61a)$$

derations

$$\left| \begin{array}{c} \wedge^4 \\ \mathbf{n} \end{array} \right| = \left| \begin{array}{cc} \stackrel{\text{CIK}'1}{\circ_1^{\text{12}}} & \mathbf{v} \\ \stackrel{\text{CM}}{\circ_2^{\text{2I}}} & S_{22} \end{array} \right| \left| \begin{array}{c} \mathbf{v} \\ \mathbf{v} \end{array} \right| \quad (6.616)$$

since port 4 is similar to port 1 and port 3 is similar to port 2.

Let us now change the incident waves at ports 4 and 3 to $-V\ell$ and $-V^a$. We now have an antisymmetrical excitation and the symmetry plane is an electric wall or short circuit. For this case the equivalent circuit is the same as that in Fig. 6.34a, except that the shunting susceptances are due to short-circuited transmission-line sections, so jB_l is replaced by $jB_d = -jY_l \cot 30_0/2$ and jB_r is replaced by $jB_4 = -1/T, \cot 0_0/2$. For this case we will designate the two-port scattering-matrix parameters under short-circuit conditions on the plane aa by $\&S$. For this odd excitation we can write

$$\left| \begin{array}{c} \mathbf{S}_{12}^{\text{sc}} \\ \mathbf{v}_u, \end{array} \right|_{322} \quad (6.62a)$$

and

$$\left| \begin{array}{cc} \mathbf{S}_{21}^{\text{sc}} & \stackrel{\text{-V}}{\circ_1^{\text{12}}} \\ S_{22} & \mathbf{-VS} \end{array} \right| \quad (6.626)$$

Let us now superimpose the two solutions which then gives $Vf = 2V_u$, $V\ell = 2V\ell$, $VQ = Vf = 0$. The superposition of the scattered waves gives

$$\left| \begin{array}{c} \mathbf{K} \\ \cdot \mathbf{V} \end{array} \right| = \left| \begin{array}{c} 1 \\ " 2 \end{array} \right| \left| \begin{array}{c} 8\ell + S\% \\ \circ_2^{\text{oc}} \end{array} \right| \left| \begin{array}{c} \mathbf{S}_{22}^{\text{sc}} \\ \circ_2^{\text{oc}} \end{array} \right| \quad (6.63a)$$

$$\left| \begin{array}{c} \mathbf{V} \\ \mathbf{4-1} \end{array} \right| = \left| \begin{array}{c} 1 \\ " 2 \end{array} \right| \left| \begin{array}{c} \circ_2^{\text{oc}} - \circ_2^{\text{esc}} \\ \circ_2^{\text{oc}} \end{array} \right| \left| \begin{array}{c} \mathbf{S}_{12}^{\text{sc}} \\ \circ_2^{\text{oc}} \end{array} \right| \quad (6.636)$$

$$\left| \begin{array}{c} \mathbf{V} \\ \cdot \mathbf{V} \end{array} \right| = \left| \begin{array}{c} 1 \\ " 2 \end{array} \right| \left| \begin{array}{c} \circ_2^{\text{oc}} - \circ_2^{\text{esc}} \\ \circ_2^{\text{oc}} \end{array} \right| \left| \begin{array}{c} \mathbf{v} \\ \mathbf{v} \end{array} \right| \quad (6.636)$$

From these equations we can identify the following eight four-port scattering-matrix elements:

$$\&u = |(S_n + S_u)J| \quad S_{21} = a(\wedge i_2 + S_{21})$$

$$S_{21} = s(S_{21} + S_{ff}) \quad S_{22} = a(Sa_2 + S^a)$$

$$\circ_2^{\text{ai}} = -1 / Q_{\text{oc}} - Q_{\text{sc}} \quad S_{32} = s(\wedge 22 - S_{at})$$

$$41 = 2V^a \circ_1^{\text{11}} \quad \wedge II \quad 42 = \frac{1}{2V^a \circ_1^{\text{12}}} \quad \frac{C_{\text{sc}}}{\circ_1^{\text{12}}}$$

From symmetry considerations $S_M = S_u$, $S_{33} = S_{2a}$, $S_{ff} = S^{\text{TM}}$. The re-

maining five elements follow from reciprocity, namely, $S_{11} = S_{22} = 5$, $S_{12} = S_{21}$, four-port hybrid can be characterized in terms of two sets "of scattering-matrix parameters.

In order to obtain the two-port scattering-matrix parameters for the circuit shown in Fig. 6.346. We assume incident and scattered waves exist at ports 1 and 2. On the interconnecting line we assume of a forward and backward propagating voltage wave given by

$$V^+ e^{-jkz} + V^- e^{jkz}$$

with associated current waves

$$I^+ Y_r e^{-jkz} - I^- Y_r e^{jkz}$$

The continuity of the voltage at ports 1 and 2 gives

$$\mathbf{V}^+ \mathbf{Y}_r + \mathbf{V}_r = \mathbf{V}^- + \mathbf{V}_- \quad (6.64Q)$$

$$V^+ + V^- = V_r + J^* > + V_r - * \quad (6.64G)$$

The continuity of the current at each port gives

$$(V^+ - V^-)Y_C = (V^+ - V^-)7, + (V^+ + V^-)jB_1 \quad (6.64c)$$

$$(V^+ e^{-jkz} - V^- e^{jkz}) Y_r = -(V_2^+ - V_2^-)Y_C + (V_2^+ + V_2^-)jB_2 \quad (6.64rf)$$

We can solve the first two equations for V^+ and V^- . We then substitute these solutions into the last two equations and solve for V_f and V_2 in terms of V_f and V_2 . This will give us the two-port parameters S^* . After carrying out the algebra, we obtain

$$S_{11} = jI^* Y_C \sim Y_i + B_1^2 / 2 - Y_i (S_1 + B_2) \cot B_x + jY_c (B_2 - B_1) \quad (6.65a)$$

$$\ll - * \gg \ll . \gg = \frac{2}{\pi} T_r Y_1 C S C A \quad (6.65b)$$

$$S_{21} = jX(B_1, -B_2) \quad (6.65c)$$

$$\text{where } A = 7C^2 - B_1 B_2 + Y_i (B_1 + B_2) \cot 0, + 7^*$$

$$+ Y_r (B_1 + B_2 - 2 \tan 0,)$$

$$30, \quad 0i$$

$$B_1 = 7, \tan 0, \quad B_2 = 7, \tan 0,$$

The short-circuit two-port scattering-matrix parameters are obtained by replacing B_1 by $B_3 = -7, \cot(30, / 2)$ and B_2 by $B_4 = -Y_i * 1^* . * \ell$

At the center of the frequency band of interest, we have $\ell = 1$ for which $B_1 = -Y_i$, $B_2 = Y_i$, $B_3 = 7, \cot 30, / 2$, and $B_4 = -7, \cot 30, / 2$. For these conditions the expressions for the scattering-matrix parameters simplify to

readily find that

$$\gg n - \frac{Y^2 - 2Yj^2 + 2jY_r Y_i}{Y^2 + 2Y^2} = S_{22}$$

$$S_{22}^{oc} = \frac{y_r^2 - 2Yf - 2jY_r Y_i}{Y^2 + 2Y^2} = S?$$

$$S_{22}^{oc} = \frac{y_r^2 + 2Fj^2}{Y^2 + 2Y^2}$$

From these expressions we get

$$" a \sim Sjn \quad S_{33} = S_{44} = \frac{y_r^2 - 27^2}{Y_c^2 + 2Y^2} \quad (6.66a)$$

$$S_{11} = S_{42} = 0 \quad (6.66b)$$

$$S_{12} \sim -S_{34} = -S_{41} = S_{32} = \frac{-2iy_r T}{y_r^2 + 2y^2} \quad (6.66c)$$

We see that port 3 is uncoupled to port 1 and ports 2 and 4 are uncoupled. Also ports 2 and 4 are coupled to port 2 but the port 2 output is 180° out of phase with the port 4 output since $S_{21} = -S_{41}$. All ports will be matched if we choose $Y^2 = 2Y_r^2$ or

When this latter condition is imposed, then $|S_{21}| = |S_{41}| = 1/2$ and we obtain a 3-dB directional coupler or 180° hybrid junction. The four-port scattering matrix reduces to

$$\begin{matrix} & \begin{matrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{matrix} \\ \begin{matrix} 1 & j \\ V2 \end{matrix} & \end{matrix} \quad (6.67)$$

which is the same as that for the magic-T hybrid junction, apart from a different numbering of the ports and the choice of input terminal reference planes.

At frequencies away from the frequency at which $\theta_I = ir/2$, the performance of the hybrid ring can be found by evaluating the various scattering-matrix parameters of interest as a function of the normalized frequency variable $2B_r/ir$. In Fig. 6.35 we show a plot of the couplings $|S_{12}|$, $|S_{U1}|$, and $|S_{23}|$, the isolation $|S_{13}|$, and the input reflection coefficients $|S_{U1}|$ and $|S_{22}|$ as a function of $2\chi/v$.

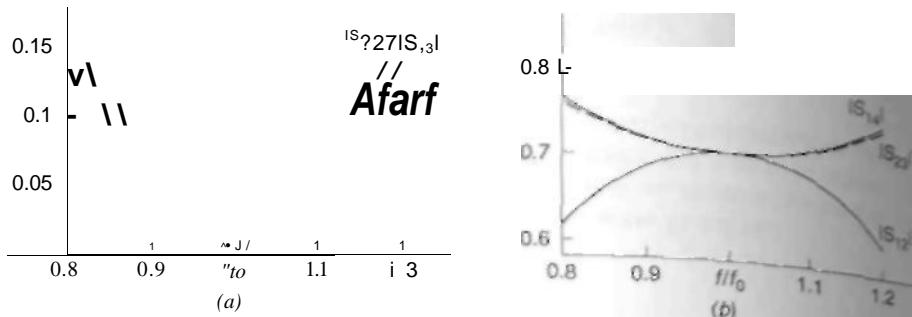


FIGURE 6.35

Hybrid-ring performance as a function of reflection coefficient and iso-coefficients.

port coupling coefficient; (b) magnitude of coupling

6.6 POWER DIVIDERS

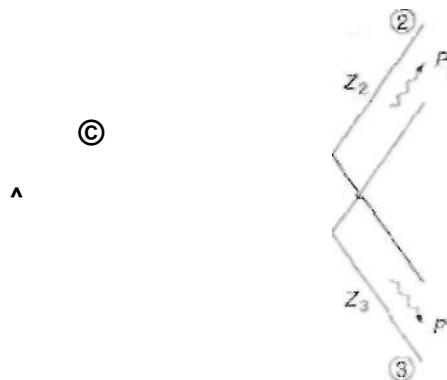
Power dividers are used to divide the input power into a number of smaller amounts of power for exciting the radiating elements in an array antenna. They are also used in balanced power amplifiers both as power dividers and power combiners.

A fundamental property of a lossless reciprocal three-port junction is that not all three ports can be simultaneously matched. If we assume that all three ports can be matched, then $S_{11} = S_{22} = S_{33}$ and the scattering matrix has the form

$$[S] = \begin{vmatrix} 0 & S_{12} & S_{13} \\ 0 & 0 & S_{23} \\ 0 & S_{23} & 0 \end{vmatrix}$$

For a lossless junction the scattering matrix is a unitary matrix. This means that the sum of the products of the elements in any row with the conjugate of the elements in another row is zero. For the discussion this would require $S_{11}S_{11}^* = S_{22}S_{22}^* = S_{33}S_{33}^* = 1$. These conditions will hold only if two of the transmission coefficients S_{12}, S_{23} and S_{13} are zero, in which case we do not have a functioning three-port junction. If we want to use a lossless three-port junction to split or divide power P_1 into fractions $aP_1 = P_2$ and $(1-a)P_1 = P_3$ at ports 2 and 3, we can choose Z_2 and Z_3 so that the input port 1 is matched and a power split is obtained. If the input is matched, then $P_1 = P_2 + P_3$ because of the parallel connection of all three ports.

$$P_1 = P_2 + P_3$$



FfGUK 6.36

A lossless three-port junction used as a power divider.

For an impedance match we require $Y_x = Y_s$ and, in order to obtain the desired power division, we require

$$Y_s$$

$$Y_s = 1 - \alpha$$

For example, if we want to split the input power so that $P_2 = P_i/3$ and $P_a = 2P_i/3$, then $Y_2/Y_A = 1/3$ or $Y_3 = 2Y_s$. Consequently, in order to make port 1 matched, $Y_s = Y_2 + Y_3 = 3Y_2$; so $Z_2 = 3Z_s$, and $Z_3 = 1.5Z_s$. This type of lossless power divider will not have matched output ports and since S_{13} will not be zero it also does not have isolation between the output ports. If there is a shunt susceptance at the junction, such as would occur from excitation of evanescent modes in a waveguide T or Y junction, the input port can still be matched by placing a suitable shunt-compensating susceptance at an appropriate position in the input line. It is desirable to have $S_{23} = 0$, so that reflected power at port 2 does not couple into port 3, and vice versa.

Wilkinson developed an N-way power divider that would split the input power into output power at N ports and that would also provide isolation between the output powers.^t A unique feature of the Wilkinson power divider is the use of resistors connected between the various output ports. When the output ports are terminated in the correct load impedance, there is no current in the resistors; so they do not absorb any power. If one port is matched, then the reflected power from that port is partly absorbed by the resistor network and partly returned to the input, but no power is coupled into the other output ports as long as they remain properly terminated. Many different versions, including broadband designs, of the Wilkinson power divider have been developed. The multisection broadband

^tE. Wilkinson. An N -Way Hybrid Power Divider. *IEEE Trans.*, vol. MTT-8, pp. 116-118 1960.

design for a two-way power divider was developed by Cohn and used for a number of practical designs.

The basic Wilkinson power divider is illustrated in Fig. 6.69. It consists of two quarter-wave sections with characteristic impedance Z_3 connected in parallel with the input line, which has a characteristic impedance $Z_{L,i}$. A resistor R is connected between ports 2 and 3. If we want to split the input power P_i into output powers P_2 and P_3 , and if $P_2 = K^2 P_i$ and also maintain zero current in R when ports 2 and 3 are terminated, then the output voltage V_2 at port 2 must equal the output voltage V_3 at port 3. In order to obtain the desired power ratio we require $K^2 |V_2|^2 / Z_{L2} = |V_3|^2 / Z_{L3}$ for $V_2 = V_3$. Hence we need

$$K \cdot Z_{L2} = Z_{L3} \quad (6.68)$$

For matched output terminations the resistor R has no effect on the operation of the circuit. In order to obtain a *matched* input at port 1 we require $Y_m = Y_r$. By using the transforming properties of the quarter-wave sections, we have

$$\frac{Z_{L3}}{Z_{L2}} + \frac{1}{K} = -1 \quad (6.69a)$$

From the above two equations we obtain

$$(K \cdot Z_{L2} + Z_{L3}) = \frac{Z_{L3}}{Z_{L2}} \quad (6.69b)$$

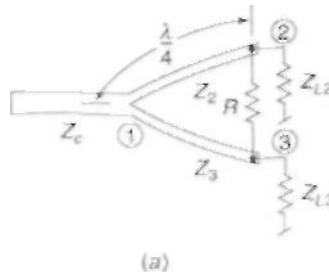
upon eliminating Z_{L2} using (6.68).

At port 1 the load impedance Z_{L2} is transformed into an admittance $y_m = 2/z$ and Z_{L3} is transformed into an admittance $y_{w,3} = z$. The power delivered to Z_{L2} and Z_{L3} is the same as that delivered to $Z_{L,i}$ and $y_{in,3}$, respectively. Thus, in order to get the desired power ratio we require that

$$i|V|^2 \cdot 3 = K \cdot V \cdot Y_{m,2}$$

which gives, upon using (6.68),

$$Z_{L2} = K \cdot Z_{L3}$$



$$\frac{G}{-VW}$$

$$Y_{33}, *Y_{*23} \quad Y_{22} *Y_{x23} \quad \boxed{\text{IMI}}$$

(o)

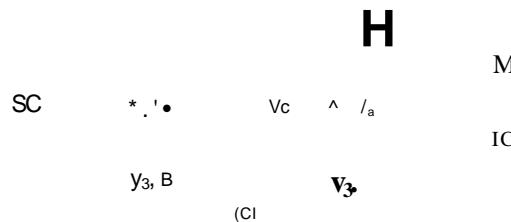


FIGURE 6.37
 (a) The Wilkinson power divider;
 (b) the equivalent circuit with port 1 terminated and ports 2 and 3 excited, (c) equivalent circuit between ports 2 and 3 with port 3 short-circuited and R removed.

At any frequency Y_{L2} transforms into

$$Y_{L2} + jY_2 \tan \theta \\ *ta.2 \quad Y_{2y_2} + jY_{L2} \tan e$$

and y_{L3} transforms into

$$\mathbf{n}, \mathbf{8} = *3 \frac{Y_{L3} + jY_3 U_m \theta}{Y_3 + jY_{L3} \tan 0} \quad m \quad K \cdot iy \quad \text{in, 2}$$

by virtue of (6.68) and (6.70). Thus the input current on line 3 will be K'

larger than the input current on line 2. Since the line 2 and line 3 are identical, apart from the relative impedance levels the I_{L2}^* will also be K^2 larger than the load current in Y_{L2} . Hence the ratio of the load voltages will be

$$\frac{V_{L2}}{V_{L3}} = \frac{h_{21}^2}{h_{11}^2} = K^2$$

Since the load voltages are equal, there is no current in R at any frequency as long as ports 2 and 3 are terminated in their matched load impedances. When there is no current in R , the input admittance at port 1 is $Y_{in} = (1 + K^2)Y_m$. The input reflection coefficient is given by

$$r_{in} = \frac{T_{in} - Y_{in}}{T_{in} + Y_{in}}$$

$$= \frac{Y_m^2 - (1 + K^2)^2 Y_m}{Y_m^2 + (1 + K^2)^2 Y_m + 2.7(1 + K^2)Y_m \tan 0}$$

This result is obtained by substituting the expression given earlier for Y at any frequency. Equation (6.71) gives the interesting result that input port 1 will be matched at all frequencies if we choose

$$y_{in} = T_{in} F \quad (6.72)$$

In order to analyze the coupling between ports 2 and 3, we will terminate port 1 in a load admittance Y_L . We now have a two-port network. Apart from the resistor R , we can represent the network between ports 2 and 3 by a π network as shown in Fig. 6.376. The resistor R is a shunt conductance $G = 1/R$ in parallel with $-Y_{23}$. From this network representation it is clear that ports 2 and 3 will be uncoupled if $G = Y_L$, so the admittance between ports 2 and 3 vanishes. The circuit equations removed are

$$\begin{matrix} p_z \\ W \end{matrix} \mid T \wedge Y & \& \begin{matrix} irvz \\ ba \end{matrix} \mid JW$$

For $V_3 = 0$, that is, port 3 short-circuited, we have

$$Y_n \sim v_2 \quad v_3 = 0$$

Thus we can evaluate K_{23} by finding the short-circuit current i_{23} when port 2 is excited. The transmission-line circuit to be analyzed is Fig. 6.37c.

Let the voltage and current waves on line 2 be

and those on line 3 be

The terminal conditions at port 2 are

$$\mathbf{V}_2^+ + \mathbf{V}_2^- = \mathbf{V}_2 \quad (6.73a)$$

$$(6.736)$$

At port 1 the terminal conditions are

$$V_1^+ + V_1^- = V_1^+ + V_1^- \quad (6.73c)$$

$$(Vje^{-j\theta} - V_a e^{j\theta})Y_2 - (V_1^+ + V_1^-)Y_t + (V_1^- - V_h)Y_3 \quad (6.73d)$$

At port 3 we have

$$V^h e^{-j\theta} + V_h' e^{j\theta} = 0 \quad (6.73e)$$

$$\{V^h e^{-j\theta} - V_h' e^{j\theta}\}Y_3 = -7_a \quad (6.73/-)$$

where $\theta = f_2 l_2 = f_3 l_3$ and l_2 and l_3 are the line lengths. From the last two equations we get $I_3 = -2V^h e^{-j\theta}$. By adding and subtracting (6.73c) multiplied by Y_2 to (6.73d), we can solve for V_1^* in terms of either V_j or V_\sim . We can combine the two solutions to get V_1^* in terms of $V^* + V_a = V_2$ and then solve for $Y_{13} = -(2V^* + V_a)/V_2$. The results are

$$Y_{13} = \frac{7_a (1 - \cos 2\theta)}{7_a (1 - \cos 2\theta) - j(Y_2 + Y_3) \sin 2\theta} \quad (6.74)$$

When $\theta = 77^\circ/2$ we get $Y_{13} = Y_2 Y_3 / Y_f$. In order to uncouple ports 2 and 3, we thus require

$$G = 7_a = -\hat{\mathbf{Y}} \quad (6.75a)$$

$$R = \frac{Z}{Z} \quad (6.756)$$

From (6.68), (6.696), (6.70), and (6.756), we can express all •
in terms of R in the following form:

$$Z_{,}^{\wedge} K \{ R Z_C$$

$$Z_3 = -W <$$

$$K^* R$$

$$" \times 2 \quad K^2 + 1$$

$$R$$

$$Z_{,,} = K^2 + 1$$

Impedance

<6.76Q)

'6.766,

16.76c,

(6.7W)

The resistance R can be chosen arbitrarily. If we require port 1 matched at all frequencies, then R is determined by the condition given by (6.72).

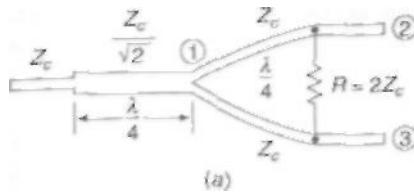
We can also solve for I_2 in terms of V_1 , from (6.73). The ratio V_2/V_1 for $V_3 = 0$ gives the parameter $Y_{2,}$. We find that

$$Y_m = Y_1 \frac{Y_3 - Y_2 + (Y_a + Y_2) \cos 2\theta + jY_e \sin 2\theta}{Y_1(\cos 2\theta - I) + j(Y_2 + Y_3 \sin 2\theta)} \quad (6.77)$$

The parameter Y_{33} is obtained by replacing Y_2 by Y_3 and Y_3 by y_2 in the above equation. The parameters Y_{22} and Y_{33} are needed in order to determine the port 2 and port 3 reflection coefficients as a function of θ or frequency.

A 3-dB power divider is obtained by choosing $K^2 = 1$. If we also specify that $Z_{,2} = Z_{/3} = Z_{<}$, we find from (6.76c) that $R = 22$. From (6.76a) we get $Z_2 = Z_3 = J_2 Z_V$ for this case. A 3-dB power divider that has port 1 matched at all frequencies requires, from (6.72), $Y_2 = Y_e^*/Z$. From (6.76a) and (6.76c) we obtain $R = 4Z_f$ and $Z_w = A t 8" 2Z_{,2}$. For this case the output lines have a characteristic impedance t that of the input line. We could choose $R = 2Z_{,2}$, $Z_{,,2} = Z_{L3} = Z_c$ and a quarter-wave transformer with impedance $ZJ v2$ to transform the input impedance, which now equals $Z_{,2}/2$, into an impedance A . 6.38a illustrates this type of 3-dB power divider.

For the unequal power division case if we choose $Y_1 = Y_2 = Y_3$ (6.72), the input port 1 is matched at all frequencies. From (6.76a) that $J_2 = (1 + K^2)/K^2$. By using (6.76c) we get $Z_{,2} = <^*$ and hence there is no discontinuity at the junction of line 2 and its load. Also $Z_w = Z_3$, so that there is no discontinuity at the junction of line 2 and its load. The only disadvantage with this design is that impedances $Z_{,2} = (1 + K^2)Z_f$ and $Z_{,,2} = (1 + K^2)Z_f$ are different from $Z_{,2}$. Output quarter-wave transformers can be added Z_{L2} and Z_{L3} into the impedance $Z_{,2}$. The required transforms V^A , J_{1+ff^*} and $(1 + K^2)/K^2$ are different. Thus the m -



(a)

$$\frac{Z_c}{4}$$

$\frac{Z_c}{4}$

$\frac{Z_c}{4}$

(b)

 Z_c **FIGURE 6.38**

(a) A 3-dB power divider with an input quarter-wave transformer.¹ (b) a broadband unequal power divider.

function of frequency will be different for the two output transformers. In order to equalize the transformation ratios and thereby obtain similar characteristics for ports 2 and 3, the output impedances should be chosen according to the relations

$$Z_{L2} = KZ_t \quad (6.78a)$$

$$Z_{L3} = -\frac{K}{K} \quad (6.78b)$$

We now require

$$R = \frac{1 + K^2}{K} \quad (6.78c)$$

$$Z_2 = -Z_c JK \sqrt{1 + K^2} \quad (6.78d)$$

$$V = \frac{1}{K} \quad (6.78e)$$

The required output quarter-wave transformers that will transform Z_{L2} and Z_{L3} into Z_e will have characteristic impedances given by $i/Z_{L2}Z_c = 1/KZ_c$ and $yjZ_{L3}Z_c = Z_e/|K|$. A power divider designed on this basis will have a bandwidth approaching one octave.

A modified design that incorporates an input quarter-wave transformer as shown in Fig. 6.38b will give a significant increase in performance. The design equations are arrived at by noting that when ports 2

¹L. I. Parad and R. L. Moynihan. Split-Tee Power Divider, *IEEE Trans.*, vol. MTT-13, pp. 91-95, 1965.

and 3 are terminated in matched loads we can connect these parallel admittances at the location of the resistor R . The characteristic admittance of J^{poles} is $Y_p = (1 + K^2)Y_L$. The terminating load admittance is $Y_t = (1 + K^2)Y_{L2} = Y_t$. The structure shown in Fig. 6.386 can be used as a two-section quarter-wave transformer. Thus, for a maximum reflection coefficient characteristic (see Sec. 5.12), we should choose $Z_t = Z_p^* Z_s^* S^*$

$$Z_t = Z_p^* Z_s^* S^*$$

where Z_t is the characteristic impedance of the input transformer

$$(1 + K^2)Y_t = Z_t^2 / Z_p^* = \frac{1}{1 + K^2}$$

With $Z_{L2} = KZ_C$ and $Z_{A;i} = Z_{A;i}/K$, we have $y_t = Q + \#^2 y_c / \#_h$ we must choose

$$Z_t = \frac{K}{1 + K} \quad (6.79a)$$

$$Z_p = f f^{3/4} (1 + K)^{1/4} Z_e \quad (6.79b)$$

$$Z_s = \frac{(H - JT')''}{K} \quad (6.79c)$$

$$tf = \frac{1 + K^2}{K} \quad (6.79a')$$

6.7 MICROWAVE PROPAGATION IN FERRITES

The development of ferrite materials suitable for use at microwave frequencies has resulted in a large number of microwave devices. A number of them have nonreciprocal electrical properties; i.e., the transmission coefficient through the device is not the same for different directions of propagation.¹¹ An understanding of the operation of ferrite devices may be based on the basic nature of microwave propagation in an infinite unbounded medium. In this section we consider plane-wave propagation in an infinite ferrite medium with a static biasing magnetic field present. It will be found that the natural modes of propagation in the direction of B , are left and right circularly polarized waves and the modes have different propagation constants. In addition, we find that the permeability of the ferrite is not a single scalar quantity, but a tensor, which can be represented as a matrix.

Ferrites are ceramic-like materials with specific resistivity ρ and conductivity σ that can be as much as 10^4 greater than that of metals and with dielectric constant around 10 to 15 or greater. Ferrites are made by sintering metallic oxides and have the general chemical composition $M_{2-x}O_4$, where M is a divalent metal such as manganese, magnesium, iron,

nickel, cadmium, etc., or a mixture of these. Relative permeabilities of several thousand are common. The magnetic properties of ferrites arise mainly from the magnetic dipole moment associated with the electron spin. By treating the spinning electron as a gyroscopic top, a classical picture of the magnetization process and, in particular, the anisotropic magnetic properties may be obtained.

The electron has a number of intrinsic properties such as a charge of $-e = -1.602 \times 10^{-19}$ C, a mass $w = 9.107 \times 10^{-31}$ kg, an angular momentum P equal in magnitude to hh , or 0.527×10^{-31} J • s (ft is Planck's constant divided by 2), and a magnetic dipole moment m equal to one Bohr magneton, that is, $m = eh/2w = 9.27 \times 10^{-29}$ A • m². For the electron, the angular momentum P and magnetic dipole moment m are antiparallel. The ratio of the magnetic moment to the angular momentum is called the gyromagnetic ratio γ ; that is,

$$\gamma = \frac{m}{P} \quad (6.80)$$

If an electron is located in a uniform static magnetic field B_0 , a torque T given by

$$T = m \times B_n = -\gamma P \times B_0 \quad (6.81)$$

will be exerted on the dipole moment. This torque will cause the dipole axis to precess about an axis parallel to B_0 , as illustrated in Fig. 6.39. The equation of motion is obtained from the condition that the rate of change of angular momentum is equal to the torque and hence is

$$\frac{dP}{dt} = T = -\gamma P \times B_0 = w_t \times P \quad (6.82a)$$

$$\text{or } \gamma PB_0 \sin \phi = w_t P \sin \phi = mB_0 \sin \phi \quad (6.82b)$$

where w_t is the vector-precession angular velocity directed along B_0 , and ϕ

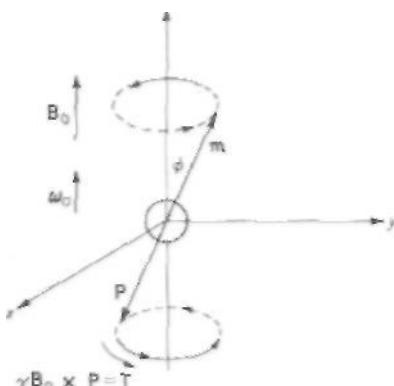


FIGURE 6.39
Free precession of spinning electron.

is the angle between \mathbf{m} and \mathbf{B}_0 . For free precession the angle θ is given by

$$\omega_0 = \gamma B_0$$

and is independent of the angle ϕ . The angular velocity at the Larmor frequency.

If a small *ac magnetic field* is superimposed on the static field \mathbf{B}_0 , the magnetic dipole moment will undergo a forced precession. Of interest is the case where the *ac magnetic field* is circularly polarized in the plane perpendicular to \mathbf{B}_0 . A circularly polarized field results when the *x* and *y* components of the *ac field* are equal in magnitude and 90° in time phase. Thus let the *ac magnetic field* be given by the phase

$$\mathbf{B}_t = B_x(a_x + JB_y) \quad (6.83a)$$

If we assume B_t to be real, the physical field is given by

$$\mathbf{B}_r = \Re(a, vja_y)e^{j\omega t} = B_r(a_r \cos \omega t - a_v \sin \omega t) \quad (6.83d)$$

The resultant field has a constant magnitude B_r , but the orientation of the field in space changes or rotates with time. At time t the resultant field vector makes an angle

$$\tan^i = -\tan \tan \text{tot} = -\text{tot}$$

with the *z* axis and hence rotates at the rate ω_0 , as in Fig. 6.40a. It is this rotation of the field vector in space that results in the field being called circularly polarized. If the above *ac magnetic field* is that of a wave propagating in the *z* direction, it is said to be left circularly polarized. If the direction of rotation is clockwise, looking in the direction of propagation, the wave is called right circularly polarized. The latter type of wave would be an *ac magnetic field* given by

$$\mathbf{B}_r = B_r(\mathbf{a}_I - \mathbf{a}_R) \quad (6.84)$$

With a left circularly polarized *ac magnetic field* superimposed on the static field $\mathbf{B}_0 = B_0 \mathbf{a}_z$, the resultant total field \mathbf{B}_r is inclined at

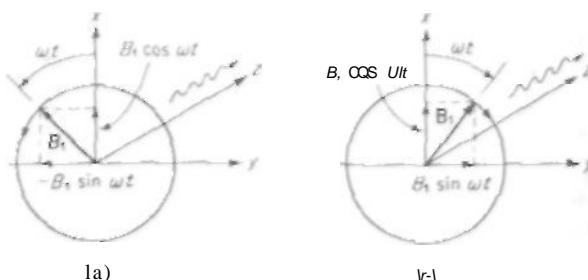


FIGURE 6.40. Circularly polarized waves-right, or postive-polarization-

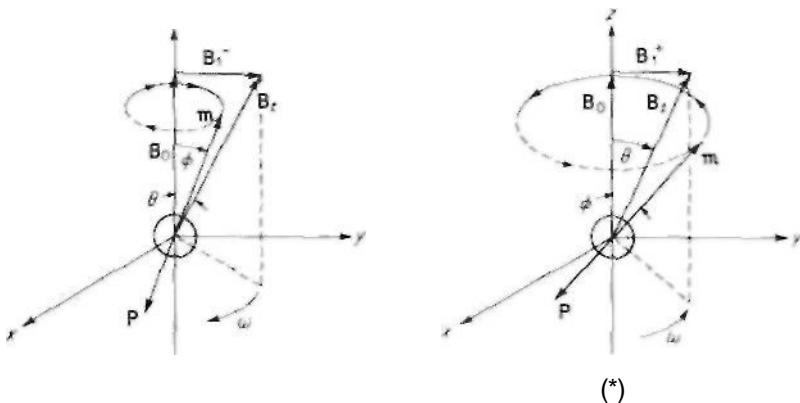


FIGURE 6.41
Forced precession of spinning electron.

$\delta = \tan^{-1}(B_t/B_0)$ with the z axis and rotates at a rate $-\omega_0$ about the z axis, as illustrated in Fig. 6.41a. Under steady-state conditions the magnetic dipole axis will be forced to precess about the z axis at the same rate. Thus the precession angle ϕ will have to be less than δ , as in Fig. 6.41a, in order to obtain a torque to cause precession in a counterclockwise direction. The equation of motion (6.82) gives

$$\mathbf{T} = \mathbf{m} \times \mathbf{B}_t = -y \mathbf{P} \times \mathbf{B}_t = \frac{d\mathbf{P}}{dt} = -\omega_0 \mathbf{a}_x \times \mathbf{P}$$

or

$$-yPB_t \sin(\theta - \phi) = -\omega_0 P \sin \delta$$

Expanding $\sin(\theta - \phi)$, replacing $B_t \sin \theta$ by B_t , $B_t \cos \delta$ by B_0 , and solving for $\tan \phi$ gives

$$\tan \phi = \frac{yB_t}{yB_0 + \omega_0} = \frac{yB_t}{to_{\phi} + to} \quad (6.85)$$

The component of \mathbf{m} which rotates in synchronism with \mathbf{B}_t in the xy plane for the left (also called negative) circularly polarized ac field is $m_Q = m \sin \phi$, where $m_Q = m \cos \phi$ is the z -directed component of \mathbf{m} . For B_t very small compared with B_0 , the angle ϕ is small, so that m_Q is approximately equal to m . Using (6.85) gives

$$m_Q = m_0 \tan \phi = \frac{ym_0 B_t}{\omega_0 + to} \quad (6.86)$$

If we have a right (or positive) circularly polarized ac field superimposed on the static field B_0 , the forced precession is in a clockwise sense about the z axis. A torque giving precession in this direction is obtained only if the angle ϕ is greater than the angle δ , as in Fig. 6.41b. In this case

the equation of motion (6.82) gives

$$yB \sin(\theta) = w \sin \phi$$

from which we obtain

$$\tan \phi = \frac{yB}{w_0}$$

The component of magnetization in the xy plane rotating in synchrotron with the positive circularly polarized ac field is

$$m' = m_0 \tan \phi = \frac{v / \omega_0}{5}, \quad (6.88)$$

The foregoing discussion has pointed out the essential features of the motion of a single spinning electron in a magnetic field consisting of a static field along the z axis and a small circularly polarized ac field in the x plane. A ferrite material may be regarded as a collection of IV effective spinning electrons per unit volume. Since the spacing between electrons is of atomic dimensions, we may regard the density of magnetic dipoles per unit volume as a smeared-out continuous distribution from a macroscopic viewpoint. The total magnetic dipole moment per unit volume is $M = Nm$. When the static field B_0 is large enough to cause saturation of the magnetization in the ferrite, $M = M_s$. In a saturated ferrite all the spins are very tightly coupled, so that the whole sample acts essentially as a large single magnetic dipole. The magnetization M , produces a contribution to the total internal B field according to the relation $B = M_0(H_0 + M_s)$. The torque acting on M , is due only to the field $B_{\perp} = \mu_0 H_0$, since the cross product $/j, M_v$ with M_s is zero and hence does not contribute to the torque. Thus, the equation of motion for the magnetization, the field producing the torque is $M \propto H$, where H , is the total static plus ac magnetic field intensity U in the ferrite medium. That is,

$$\frac{dM}{dt} = -y(M \times B) = -77i_m M^x (H + M) = -TM_0 M_x H, \quad t \neq 0$$

If the magnetic field intensity in the ferrite is $H_0 + H_f$, by analogy with expressions analogous to (6.86) and (6.88), but with m_0 replace j , Nm_0 . The total ac magnetic field in the xy plane is the $M_0 H_f$ plus the contribution from the ac magnetization. Thus the total positive and negative circular polarization are [we are using $B = M_0(H + M)$]

$$B' = M_0 M_z + B_r = M_0 (M^x + H_f) = u_0 \left(1 + \frac{u_0 y M_h}{i \omega_0} \right) H_f; \quad (6.89)$$

$$B = M_0 M_z + B_f = M_0 \left(1 + \frac{MQVM_0}{te_n + t} \right) H_f$$

where $M_0 = M_0 \cos \theta$, $B_r = M_0 \frac{a_T}{r} + H_0$, $T = \frac{1}{2} i^2 \pi^2 \cdot M_0$ in (6.89a), and H_0 in (6.89c) equals $H_0 (a_T + j a_y)$, as seen from (6.84) and (6.83a). The quantity M may be replaced by the saturation magnetization M_s in the ferrite since the static field B_r is usually large enough to cause saturation.

If we assume that $B_r \ll B_0$, so that $M_r = M_s$, the effective permeabilities for positive and negative circularly polarized ac fields are seen to be given by

$$\mu_r = M_s J_1 + \dots * \quad (6.90a)$$

$$\mu_r = M_s J_1 + \frac{1}{I_w + \frac{1}{w_0 + \frac{1}{r}} /} \quad (6.90b)$$

Plane circularly polarized TEM waves propagating in the direction of the static field B_r will have propagation constants

$$\beta_r = w_v + k_r \quad (6.91a)$$

$$p_r = \sqrt{\epsilon_r} \omega_r \quad (6.91b)$$

where ϵ_r is the dielectric permittivity of the ferrite. The significance of the inequality of $\beta_r < p_r$ and (i_r) is discussed later. The results expressed by (6.91) are also derived in an alternative way later.

If small-signal conditions $B_r \ll B_0$ are not assumed, we cannot replace M_0 by M_s . In place of (6.85) and (6.87), which give solutions for $\tan \theta$, we can solve for $\sin \theta$ to obtain, respectively,

$$\sin \theta = \frac{-j}{V \sqrt{1 + \tan^2 0}} = \frac{r}{V (r B_r)^2 + (\epsilon_r + \epsilon_w)} \quad (6.92a)$$

$$\sin \theta = \frac{-y B_r}{V (y B_r)^2 + (w_0 - w)} \quad (6.92b)$$

The magnetizations M^* and M_\sim are given by

$$M^* = M_s \sin \theta = \frac{V a_r M_v /}{2} \quad (6.93a)$$

$$M_\sim = \frac{V a_r M_v /}{2} \quad (6.93b)$$

It is seen that the ac magnetization depends nonlinearly on the ac field strength $/$, and hence, under large-signal conditions, x_+ and x_- will be functions of the applied ac field strength. The nonlinear behavior of ferrites under large-signal conditions results in the generation of harmonics of the

fundamental frequency ω . For this reason ferrites may be used generators.¹

For Af it is clear that, if $B_x < K B_0$, that is, $\beta \ll \sqrt{D}$ then (6.936) is well approximated by

$$M = \frac{y \rho x_0 M_i H_x}{\tan + CO}$$

Similarly, if ω is not too close to the resonant frequency ω_0 we see that (6.93a) becomes

$$AT = \frac{y MoM_{ss}}{\omega_n - a}$$

These latter values of M' and M_∞ lead directly to the expressions (6.901) for n_4 and α . In any practical ferrite medium, damping effects are always present, so that M' will remain finite and small compared with M_∞ even when $a = \omega_0$. Thus, for small-signal conditions, we can assume that $M' = M_s$ in an actual ferrite medium. Damping effects are discussed in more detail later.

It will be instructive to study the propagation of a plane wave in an unbounded ferrite medium by solving Maxwell's equations directly, together with the equation of motion for the magnetization. This analysis will illustrate the general technique of linearization to be applied in the small-signal analysis of propagation through a medium such as a ferrite. However, it will not give as clear an insight into the physical reason why x_{ss} and α are different, as the analysis above did. That is, basically, $M + M'$ differ because the precession angle θ must be greater than the angle θ in one case and less than θ in the other case, and hence the projection of the magnetic dipole moment onto the xy plane is different in the two cases.

Consider an infinite unbounded ferrite medium with an applied magnetic field $B_0 = \hat{x}_0 H_0 = B_0 a_z$. Let the magnetization in $J_{JJJJ_{n_8}}$ M_s per unit volume when no time-varying magnetic field is applied. A time-varying magnetic field $\hat{x}_0 H$ is also applied, a tiny-varying component of magnetization will be produced. The equation of motion for magnetization per unit volume is similar to that for a single electron, hence we have

$$\frac{dt}{dt} = -YMO(M_s X H_0 + M_s X \hat{x}_0 H + / X H_0 + /^5) > f\$4i$$

If small-signal conditions are assumed, i.e..

$$|jg| \ll |MJ| \quad \text{and} \quad \ll \text{Pol}$$

the nonlinear term $je \times \mathbf{H}$ in (6.94) may be dropped. We then obtain for the equation of motion the linearized equation

$$\frac{d\mathbf{M}}{dt} = -y(M_0 M_s - X^A + Jt \times B_0) \quad (6.95)$$

since $M_s \times B_0 = 0$, because the saturation magnetization is in the same direction as the applied static field.

Let the time dependence be $e^{i\omega t}$, and let J and H be represented by the phasors \mathbf{M} and \mathbf{H} . From (6.95) we obtain

$$jcoM + yM \times B_0 = wM + w_0M \times a_2 = -y^A M_s \times H$$

where $w_0 = yf_0H_0 \sim yB_a$. In component form we have

$$,iioM_x + w_0M_v = yM_s ix_0H_y$$

$$ju>M_y - u),,M_x = -yM_s f_j u H_x$$

$$jioM_z = 0$$

The solution of these equations for M_x , M_v , and M_z gives

$$M_x = \frac{W_0 Y d o M_a H_x + joj y f_0 M_s R}{R} \quad (6.96a)$$

$$M_y = \frac{-D_0 y u_0 M_x H_y - jwy i J_Q M_s H_x}{R} \quad (6.96b)$$

$$M_z = 0 \quad (6.96c)$$

In the solution of Maxwell's equations it is convenient not to have to deal explicitly with the magnetization. The magnetization may be eliminated by introducing the magnetic susceptibility and permeability. In the scalar case this is done by means of the relations $M = \chi \mathbf{H}$, $B = H_Q(M + H) = n_0(l + x, \mathbf{H})H = f_i H$. For a ferrite similar relations may be used, but χ_m and n will not be scalar quantities. In matrix form (6.96) gives

$$\begin{array}{c|ccc|c} M_x & X_{xx} & X_{xy} & \mathbf{0}^T & \mathbf{K} \\ M_y & *_{vx} & X_{yy} & \mathbf{0} & H_Y \\ M_z & \mathbf{0} & \mathbf{0} & \mathbf{0} & H_z \end{array} \quad (6.97)$$

where

$$= \chi_{vv} \frac{w_0 \gamma v \cdot \mu_0 M_s}{j i \gg y n_0 M_s}$$

$$\begin{array}{c} \text{A V J} \\ \text{<#?, - at}^2 \end{array}$$

The matrix with the parameters x_{xx} , x_{xy} , $*_{yy}$, $*_{yz}$, i_n (6.97), μ_0 , ϵ_r , B and H fields is

$$\mathbf{B} = \mathbf{M}_{\text{eff}}(\mathbf{H} + \mathbf{M})$$

or

$$= Mo \begin{vmatrix} 1 + X_{xx} & X_{xy} & 0 \\ X_{yx} & 1 + X_{yy} & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{matrix} Hy \\ Hx \\ Hz \end{matrix} \quad (6.98)$$

The matrix relating the components of H to B in (6.98) is the permeability tensor for the ferrite. It will be denoted by a boldface p , with an overbar \bar{p} .

$$\begin{matrix} 1 + X_{xx} & X_{xy} & 0 \\ X_{yx} & 1 + X_{yy} & 0 \\ 0 & 0 & 1 \end{matrix}$$

In shorthand notation the matrix equation (6.98) will be written as

$$\mathbf{B} = \mathbf{H} \bar{\mathbf{p}} \mathbf{H} \quad (6.99)$$

In the literature the minus sign in the equation of motion is often deleted, and this amounts to replacing y by $-y$ in the equations used in this.

Losses present in a ferrite may be accounted for in a phenomenological way by introducing into the equation of motion a damping term that produce a torque tending to reduce the precession angle θ . The following modified form of the equation has often been used in practice:

$$\frac{d\mathbf{H}}{dt} = -y Mo(\mathbf{M}, +*) \mathbf{X} (\mathbf{H}_0 + \mathbf{H}) - \overline{a} \mathbf{M} \mathbf{X} \mathbf{j} - \frac{dJf}{dt} \quad (6.1001)$$

where a is a dimensionless damping constant. With a small-sign the elements of the susceptibility matrix are now found to be

$$\begin{aligned} X_{xx} &= X_{yy} = X - JX - X \\ X_{xy} &= -X_{yx} - J(K - jK'') - jK \end{aligned} \quad , 6.101^{**}$$

Go * * * *

R. F. Soohoo, "Theory and Application of Ferrites," chap. 5. Prentice-Hall, N.J., 1960.

where

$$\begin{aligned} & \langle j_0 a \rangle_m (o > o - t o^2) + \wedge \cdot w p f e T^2 f l^2 \\ * " & [\wedge - w^2 (1 + \langle \rangle^2)]^2 + 4 \wedge > V \\ a) w,, f t [a] f, & + w^2 (1 + a')] \end{aligned}$$

r =

$$\begin{aligned} K' &= \frac{\wedge M_B [M J - f^2 (1 + \langle \rangle^2) j]}{[w^2 - a^2 (1 + \langle \rangle^2)]^* + 4 w 5 \gg V} \\ K'' &= \frac{2 w^2 W |, W,,| \langle \rangle}{[t \frac{g}{\pi} \sim a^2 (1 + o^2)] + -4 f i > g w V} \end{aligned}$$

and

$$w,, = n_0 y M_s$$

As we derived the permeability tensor to use in the constitutive equation relating B and H , the only remaining task is to find solutions for Maxwell's equations in the form

$$\nabla \times E = -y' w B = -y' w j_l \cdot H \quad (6.102a)$$

$$\nabla \times H = j \langle \rangle E \quad (6.102b)$$

$$\backslash " B = \nabla \cdot E = 0 \quad (6.102c)$$

For a TEM wave propagating in the z direction, i.e., along the direction of B_0 , solutions are readily found. Let the electric field be given by

$$E = E_0 e^{imz}$$

where $E,,$ is a constant vector in the xy plane. Equation (6.102a) gives

$$\nabla \times E = -E_0 X V e^{imz} = \wedge' 0 E,, X a, e^{imz} = -j c o p. \cdot H$$

Let the solution for H be $H_{fl} e^{-j*te}$, where H_0 is also a constant vector in the xy plane. From (6.102b) we obtain

$$j p n,, X a_2 = y W E_0$$

If this equation is substituted into the equation

$$j p E,, X a, = -j o j J L \cdot H,,$$

so as to eliminate E_{fl} , we obtain

$$j p \frac{M}{r} (H_n X a,) X a, = -j w p. \cdot H_0$$

Expanding the left-hand side gives (note that $a. \cdot H,, = 0$)

$$p^2 H_{is} = c o^2 e J L \cdot H_0 \quad (6.103a)$$

This equation may be written in the following matrix form-

$$\begin{matrix} p^2 - u >^2 e n_0(l+x) & -j o >^2 e f i_0 K \\ j(o^2 e^{\wedge} n_0 K) & P^* - u >^2 e n_0(l+x) \end{matrix} \quad \text{ox} \quad (6.1036)$$

For a nontrivial solution for H_{α} , the determinant must condition yields the following eigenvalue equation for the α constant 0:

$$\lambda^2 - \langle \nabla, (I - x) \mathbf{f} - \langle \mathbf{u} \mathbf{V} \mathbf{V}_0 \mathbf{f} \rangle^2 = 0$$

or

$$\lambda^2 = w^2 e n_0(l+x) \pm \lambda^2 e n_0 K \quad (6.104)$$

If we substitute for x and K and assume a lossless ferrite so $X'' = K'' = 0$, we readily find that the two solutions for λ^2 are

$$\lambda^2 = p \ell = w^2 e_{M+} \quad (6.105Q)$$

$$\lambda^2 = f i t = u >^2 e f i_- \quad (6.105E)$$

where x and n_- are given by (6.90).

For each eigenvalue or solution for λ^2 , the ratio of H_{0t} to U_{α} is determined. For the solution λ^2 , the first equation in the pair of equations (6.1036) gives

$$[f i t - \langle \nabla, (I + x) \rangle] H_{0t} - j o > h n_0 K H_{Qy} = 0$$

or

$$H_a^{(f)} = j$$

But this condition means that $H_{\alpha} = H_0(a, -/a_v)$, or is a positive circularly polarized wave. Similarly, the solution $f i t$ gives

$$\mathbf{H}_{By} = -J \quad (6.1066)$$

which specifies a negative circularly polarized wave. Therefore it is seen that the natural modes of propagation along the direction of the static ferrite are circularly polarized TEM waves. If directions other than along B_0 were considered, it would be found that again, two modes of propagation, but these are no longer $circ^U$ g_o the plane of polarization rotates, since $0.$ and j_i are not eq^U $otation$. It is a nonreciprocal one, and is called Faraday rotation discussed in the following section.

FARADAY ROTATION

applied

Consider an infinite lossless ferrite medium with a static $j_{ar,jzed} \circ^*$ along the z direction. Let a plane TEM wave that is linearly P

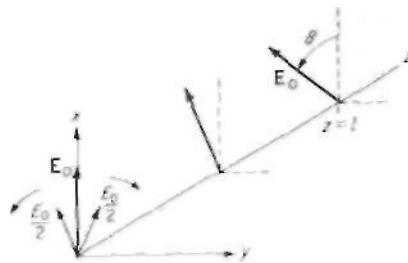


FIGURE 6.42
Faraday rotation.

the x axis at $z \sim 0$ be propagating in the z direction, as in Fig. 6.42. We shall show that the plane of polarization of this wave rotates as it propagates (Faraday rotation). The linearly polarized wave may be decomposed into the sum of a left and right circularly polarized wave as follows:

$$\mathbf{E} = a_x E_0 - (a_r + a_l) \mathbf{y} + (a_r - a_l) \mathbf{j} \times \mathbf{y} = 0 \quad (6.107)$$

The component waves propagate with different phase constants $f_{l,r}$ and $|I_{l,r}|$, and hence the wave at $z = L$ becomes

$$\begin{aligned} \mathbf{E} &= (\mathbf{a}_v - \mathbf{j}\omega) e^{iJ_z l} + (a_r^* - a_l^*) \mathbf{y}^\wedge \mathbf{T}^e \sim''' \\ &= \mathbf{a}_r e^{-iJ_z P_l} e^{iJ_z l} + e^{-iJ_z P_l} \mathbf{a}_l^\wedge \mathbf{B}_r, \quad \text{if } l > r, r < l \\ &\quad \mathbf{B}_r = J_1 H''' - f_a (e^{-iH'''/2} + iW'' - i' \rightarrow i^*) \\ &\quad + J B_r (e^{-iH''' - iH'''^2} - \cos((i - P_r)z) - a_v \sin((p_r - 0)z)) \\ &= E_r e^{-iH''' - iH'''^2} a_r \cos((i - P_r)z) - a_v \sin((p_r - 0)z) \quad (6.108) \end{aligned}$$

This resultant wave is a linearly polarized wave that has undergone a phase delay of $(f_{l,r} + p_r)/2$. The new plane of polarization makes an angle

$$\theta = \tan^{-1} f = \tan^{-1} \left(-\frac{\tan(i - P_r)}{-(H''' - p_r)} \right) \quad (6.109)$$

with respect to the \wedge axis. When $o_i < m_0$, that is, below the ferrite resonant frequency, θ_r is greater than i and the plane of polarization rotates counterclockwise, looking in the direction of propagation (Fig. 6.42). The rate of rotation is $(p_r - P_r)/2$ rad/m. Rotation of 100° or more per centimeter is typical in ferrites at a frequency of 10,000 MHz/s.

If the direction of propagation is reversed, the plane of polarization continues to rotate in the same direction. Thus, if we consider the propagation of the wave described by (6.108) back to the plane $z = 0$, the original

polarization direction is not restored; instead the wave will at $z = 0$ be polarized at an angle 2θ relative to the x axis. This result is derived by noting that if the component circularly polarized upwards in the linearly polarized wave in (6.108) are propagated from $z = 0$, they undergo additional phase delays of amount $\rho / \tan H$ and become, at $z = 0$,

$$\mathbf{E} = (\dots \cdot ; . ,) \mathbf{E} \cdot -' \cdot (\cdot , - ; . , ! \ll - \ll . , \quad (6ii0)$$

By analogy with (6.108) it is now clear that the new direction of nnls. at $z = 0$ makes an angle

$$2\theta = -(\theta / 3)l$$

with respect to the x axis. Thus Faraday rotation is a nonreciprocal effect

A practical ferrite medium has finite losses, and this will have significant influence on the propagation. The propagation constants $y = JP++\ll$ and $y = j/i_+ + u_-$ for circularly polarized waves will have unequal attenuation constants as well as unequal phase constants. When Josses are present, the propagation constants are given by (6.104) if p is replaced by $-y^2$; thus

$$y_+ = > v f c o 7 (1 + x'' - JX'' + K' - jK'')^{in} \quad (6.111a)$$

$$y_- = > v^0 7 (1 + x' - JX'' - K' + jK'')^{l/2} \quad (6.111b)$$

where x', X'', K', K'' are given after (6.101). The solutions of (6.111) are

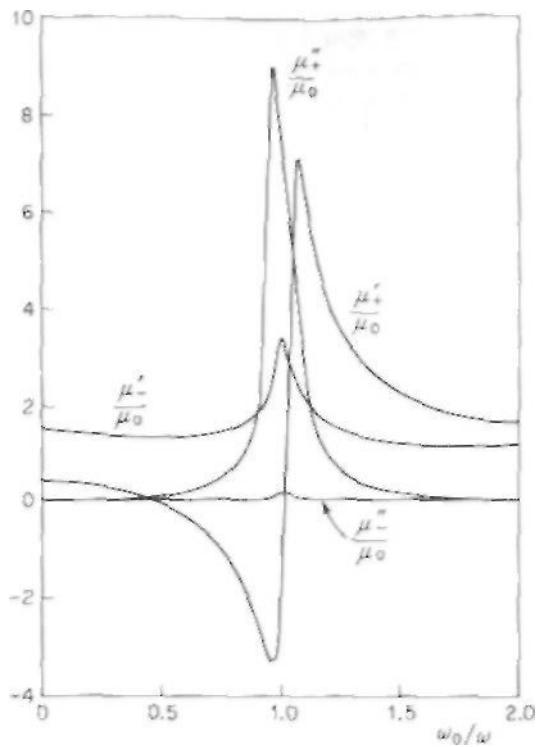
$$\frac{V_6}{V_6} = \frac{1 + *' \pm K' + \sqrt{(1 + x' \pm K'f + (X'' \pm K'')^2)}}{2} <^6 \quad (6.112)$$

$$a_+ = a > V_0^* \frac{V'' \pm g''}{2/3_f} = \frac{\ll V_+}{2/3} \quad (6.112b)$$

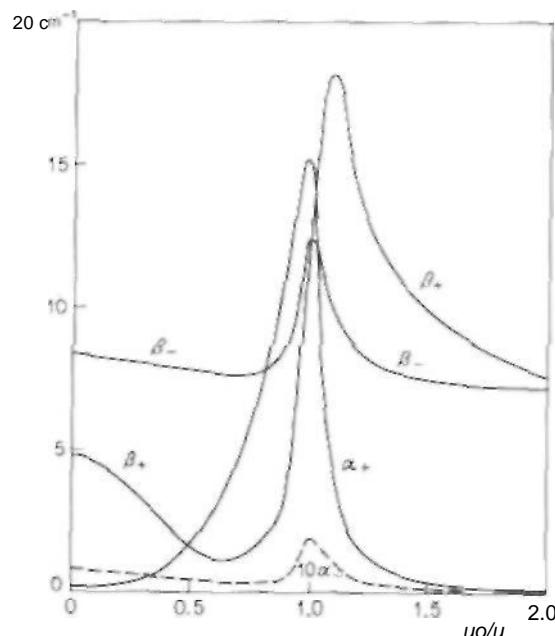
The permeabilities for circularly polarized waves are

$$M_+ = M W M W o (1 + *' - \bullet / *'' \pm K' + \bullet / *'')^>$$

The values of u_+ , u_- as given by (6.101) and (6.113) and the reflection factors $/3.$ and \ll_+ are plotted in Figs. 6.43 and 6.44. The ferrite with parameters $w_{ss} = 2\pi X 5.6 \times 10^9$, $\ll = 0.05$, as f'''_{Hz} at a frequency of 10,000 MHz. Note that $u_>_o$ equals 2×2.8 , applied field r_0 and that 4×10^{-1} Oe is a field strength of $Q(200) G$. value of w_{ss} chosen corresponds to a saturation magnetization or $n_0 M_s$ equal to 0.2 Wb/m^2 . The curves in Fig. 6.44 show that \ll is always very small but that \ll_+ is large in the vicinity of frequency $\ll_+ = co$. For i_0 considerably above co , the attenuation becomes small, but in this region i_+ and $j i_-$ do not differ greatly. The rate of Faraday rotation would be small. At low values of $\ll_+,$

**FIGURE 5.43**

Real and imaginary components of permeability for circularly polarized waves in a ferrite as a function of ui/m_0 for $w/2^* \ll 10$ GHz, $">_{\mu}/2fr = 5.6$ GHz. $\ll = 0.05$.

**FIGURE 6.44**

Propagation and attenuation constants for circularly polarized waves in a ferrite, with parameters given in Fig. 6.43 $U = 10e_{\mu}$). Note that $10a_-$ is plotted since a_- is very small.

6.9 MICROWAVE DEVICES EMPLOYING FARADAY ROTATION

Gyrator

A gyrator is defined as a two-port device that has a relative phase shift of 180° for transmission from port 1 to port 2 as compared to the phase shift for transmission from port 2 to port 1. A gyrator may be obtained by employing the nonreciprocal property of Faraday rotation. Figure 6.45 illustrates a typical microwave gyrator. It consists of a rectangular guide with a 90° twist connected to a circular guide, which in turn connects to another rectangular guide at the other end. The two rectangular guides have the same orientation at the input ports. The circular guide contains a thin cylindrical rod of ferrite with the ends tapered to reduce reflections. A static axial magnetic field is applied so as to produce 90° Faraday rotation of the TE₁₁ dominant mode in the circular guide. Consider a wave propagating from left to right. In passing through the twist the plane of polarization is rotated by 90° in a counterclockwise direction. If the ferrite produces an additional 90° of rotation, the total angle of rotation will be 180° , as indicated in Fig. 6.45. For a wave propagating from right to left the Faraday rotation is still 90° in the same sense. However, in passing through the twist, the next 90° of rotation is in a direction to cancel Faraday rotation. Thus, for transmission from port 2 to port 1, there

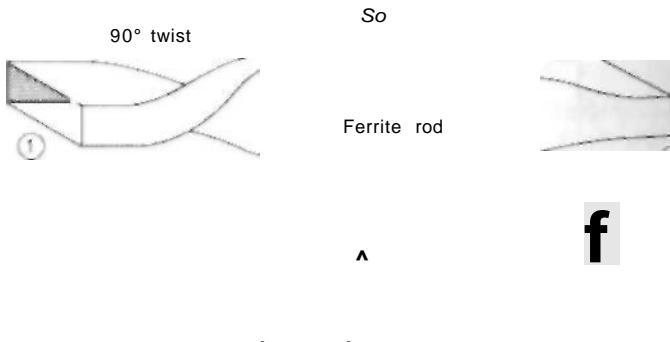
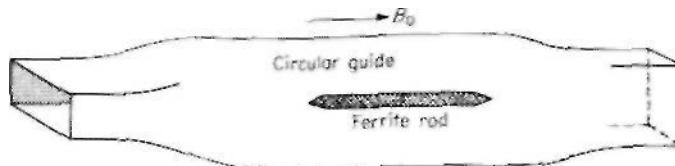


FIGURE 6.45
A microwave gyrator.



ki

▲
▲

FIGURE 6.46
A gyrator without a twist section.

net rotation of the plane of polarization. The 180° rotation for transmission from port 1 to port 2 is equivalent to an additional 180° of phase shift since it reverses the polarization of the field. It is apparent, then, that the device just described satisfies the definition of a gyrator.

If the inconvenience of having the input and output rectangular guides oriented at 90° can be tolerated, a gyrator without a 90° twist section can be built. With reference to Fig. 6.46, it is seen that if the ferrite produces 90° of rotation and the output guide is rotated by 90° relative to the input guide, the emerging wave will have the right polarization to propagate in the output guide. When propagation is from port 2 to port 1, the wave arriving in guide 1 will have its polarization changed by 180° , as shown in Fig. 6.46. Hence a differential phase shift of 180° is again produced.

The solution for wave propagation in a circular guide with a longitudinal magnetized cylinder placed in the center can be carried out exactly. However, the solution requires a great deal of algebraic manipulation, and it is very laborious to compute numerical values from the resultant transcendental equations for the propagation constants. The solution does verify that Faraday rotation takes place as would be expected, by analogy with propagation in an infinite ferrite medium.

tA. A. Th. M. van Trier. Guided Electromagnetic Waves in Anisotropic Media, *Appl. S&R.* vol. B3, p. 305, 1953.

M. L. Kales, Modes in Waveguides That Contain Ferrites, *J- Appl. Phys.*, vol. 24, p. 604, 1953.

The isolator, or uniline, is a device that permits unattenuated transmission from port 1 to port 2 but provides very high attenuation for transmission in the reverse direction. The isolator is often used to couple a microwave generator to a load network. It has the great advantage that all the available power can be delivered to the load and yet reflections from the load do not get transmitted back to the generator output terminals. Consequently the generator sees a matched load, and effects such as power output variation and frequency pulling (change in frequency), with variations in the load impedance, are avoided.

The isolator is similar to the gyrator in construction except that it employs a 45° twist section and 45° of Faraday rotation. In addition, thin resistive cards are inserted in the input and output guides to absorb the field that is polarized, with the electric vector parallel to the wide side of the guide, as shown in Fig. 6.47. The operation is as follows: A wave propagating from port 1 to port 2 has its polarization rotated 45° counterclockwise by the twist section and 45° clockwise by the Faraday rotator. It will emerge at port 2 with the correct polarization to propagate in the output guide. A wave propagating from port 2 to port 1 will have its plane of polarization rotated by 90° and will enter the guide at port 1 with the electric field parallel to the resistance card, and hence be absorbed. Without the resistance card, the wave would be reflected from port 1 because of the incorrect

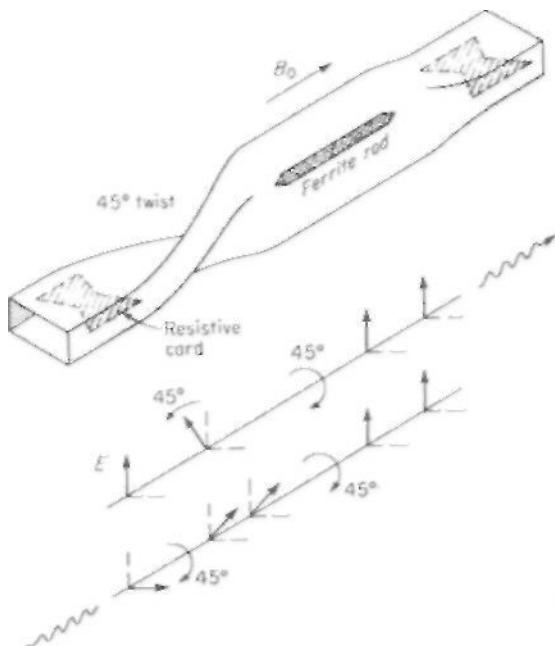


FIGURE 6.47
A Faraday-rotator isolator.

polarization, which cannot propagate in the guide constituting port 1. However, multiple reflections within the isolator will lead to transmission in both directions, and this makes it necessary to use resistance cards in both the input and output guides for satisfactory performance. Typical performance figures for an isolator are forward transmission loss of less than 1 dB, reverse attenuation of 20 to 30 dB, and bandwidth of operation approaching 10 percent.

Isolator

If the curves in Fig. 6.44 for the propagation constants of circularly polarized waves in an infinite ferrite medium are examined, it will be seen that the attenuation constant for negative circular polarization is always very small whereas that for positive circular polarization is very large in the vicinity of the resonance point $w_0 = \alpha a$. This property may be used as the basis for a resonance isolator by using a negative circularly polarized wave for transmission in the low-loss direction and a positive circularly polarized wave for transmission in the reverse direction. In the latter case the wave is rapidly absorbed or attenuated.

The condition for circular polarization is an inherent property of the dominant TE_{10} mode in a rectangular guide at two positions within the guide. The TE_{ul} -mode fields are

$$\begin{aligned} K_r &= \sin \frac{\theta}{a} e^{j\phi} \\ H_v &= \pm 70 \sin \frac{\theta}{a} e^{j\phi} \\ H_z &= -\frac{c}{a} \cos \frac{\theta}{a} e^{\pm j\phi} \end{aligned}$$

Since H_x and H_z differ in phase by 90° , circular polarization occurs when $|H_x| = |H_z|$, or when $\phi = \pm \frac{\pi}{2}$, where

$$\tan \frac{\theta}{a} = \pm \frac{c}{Pa} = \pm \frac{\theta}{2a} \quad (6.114)$$

For the solution in the range $0 < \theta < \pi/2$, the ratio of H_x to H_z is

$$\hat{\theta} = +j \quad (6.115a)$$

and the solution occurring for $\pi/2 < \theta < \pi$ gives

$$\hat{\theta} = -j \quad (6.115b)$$

With respect to the y axis, the solution given by (6.115a) corresponds to a negative circularly polarized field for propagation in the $+z$ direction and to

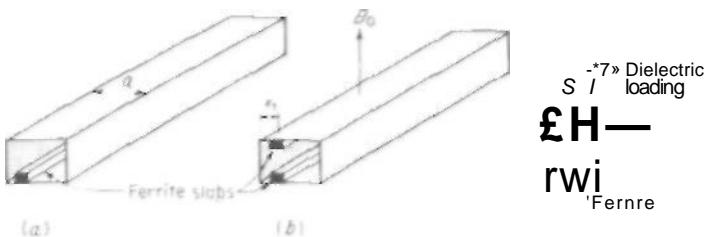


FIGURE 6.48
Rectangular-waveguide resonance isolators.

positive circular polarization for propagation in the $-z$ direction. If v , solution given by (6.3 ISA) is considered, the direction of polarization reversed.

The above property of the TE_{00} mode is utilized in the resonance isolator by locating a thin ferrite slab (or two slabs, as in Fig. 6.48b) in a rectangular guide at a position where the RF magnetic field is circularly polarized. The ferrite is magnetized by a static field applied in the y direction, as in Fig. 6.48. Since the sense of the circular polarization depends on the direction of propagation, as (6.115) shows, it follows that, for propagation in one direction, the magnetic field is negative circularly polarized and suffers little attenuation, whereas in the reverse direction the field is positive circularly polarized and rapidly attenuated. By proper design the forward loss can be kept to under 0.5 dB/in at $A_{\text{eff}} = 3 \text{ cm}^2$, and the reverse loss can be as high as 6 to 10 dB/in or even more. Dielectric loading, as illustrated in Fig. 6.48c, gives an improved reverse-to-forward attenuation ratio.

6.10 CIRCULATORS

A circulator is a multiport device that has the property (Fig. 6.4) that a wave incident in port 1 is coupled into port 2 only, a wave incident in port 2 is coupled into port 3 only, and so on. The ideal circulator is also a Δ device; i.e., with all ports except one terminated in matched loads, the Q_f impedance of the remaining port is equal to the characteristic impedance $q's o$ of its input line, and hence presents a matched load. t^{*o}

A four-port circulator may be constructed from two Δ product hybrid junctions and a gyrator as shown in Fig. 6.50. The gyrator $g u$ provides an additional phase shift of 180° for propagation in the directio n u b in Fig. 6.50. For propagation from d to a , and also from c to t^{*o}
the electrical path lengths are equal

Consider a wave incident in port 1. This wave is $h y b'$ equal-amplitude in-phase waves propagating in the side arm

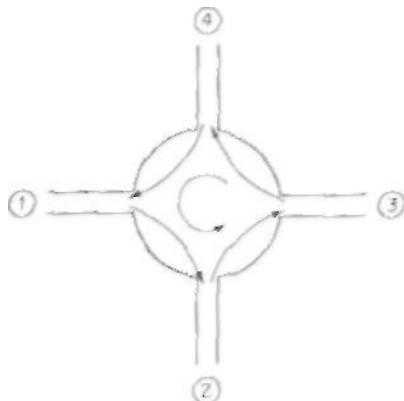


FIGURE 6.49
Schematic diagram for a four-port circulator.

junction. The waves will arrive at points *a* and *c* in phase, and hence will emerge from port 2. A wave incident in port 2 will be split into two waves, one arriving at *d* with a phase $\Delta\phi$ and the other arriving at *b* with a phase $\Delta\phi + \pi$ because of the presence of the gyrator. These partial waves have the right phase relationship to combine and emerge from port 3 in the hybrid junction. A wave incident in port 3 is split into two equal-amplitude waves, differing in phase by 180° , and hence will arrive at the other hybrid junction with the correct phase to combine and emerge from port 4. In a similar manner a wave incident in port 4 will split into two equal waves 180° apart in phase. But now the gyrator will restore phase equality, so that the waves will combine and emerge from port 1. Consequently, the microwave device illustrated in Fig. 6.50 has the required circulating transmission property.

A more compact form of four-port circulator may be constructed by employing 3-dB side-hole directional couplers and rectangular-waveguide nonreciprocal phase shifters. The nonreciprocal phase shifter will be described first. It consists of a thin slab of ferrite placed in a rectangular guide at a point where the ac magnetic field of the TE₀₁ mode is circularly polarized, as in Fig. 6.51. A biasing field S_y is applied in the *y* direction.

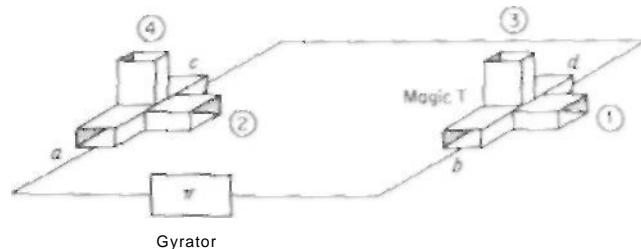


FIGURE 6.50
A four-port circulator.

-Ferrieie slab

-/ FIGURE 6.51
A nonreciprocal phas,

Since the ac magnetic field is right circularly polarized at x direction of propagation and left circularly polarized for the opposite⁶ direction of propagation, the perturbing effect of the ferrite slab is different*! the two directions of propagation. Consequently, the propagation constant β_3 for forward propagation is different from the propagation constant β_4 for reverse propagation. By choosing the length of slab so that $(\ell - \#) = \pi r/2$, a differential phase shift of 90° for the two directions of propagation can be achieved.

A four-port circulator utilizing two 90° nonreciprocal phase shifters is illustrated in Fig. 6.52. The phase shifters are biased, with oppositely directed static fields—an arrangement easily achieved in practice with permanent magnets, as shown in Fig. 6.52. One guide is loaded with a dielectric insert to provide an additional 90° of reciprocal phase shift. The coupling holes are arranged to provide 3 dB of coupling. The wave coupled

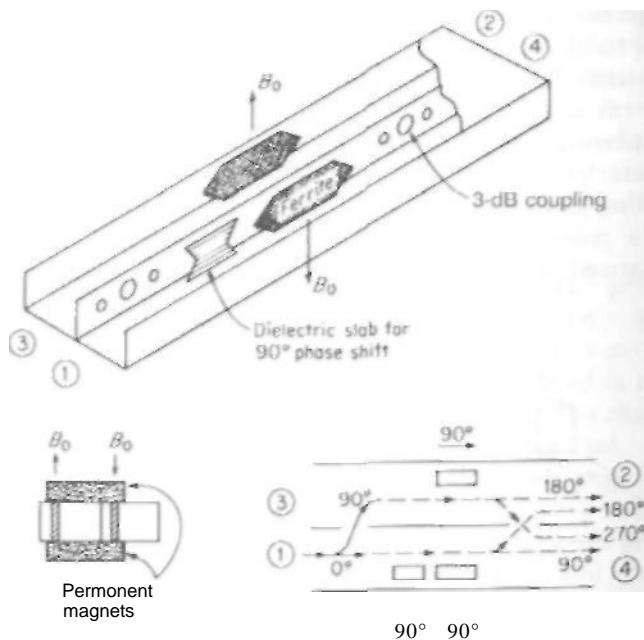


FIGURE 6.52
A compact form of four-port circulator.

through the apertures suffers a 90° change in phase, and this phase change is important in the operation of the circulator.

Consider a wave in port 1. This wave is split into two waves by the first 3-dB coupler, the wave in the upper guide undergoing a 90° phase change because of the transmission properties of an aperture. The wave in the upper guide will arrive at the second coupler with a relative phase of 180° , and the wave in the lower guide with a relative phase of 90° . The second coupler splits these waves in the manner illustrated in Fig. 6.52. It is seen that the resultant waves are out of phase in port 4 but in phase at port 2. Thus *transmission is from port 1 to port 2*. A similar analysis will verify that a wave incident in port 2 emerges at port 3, or, in general, that the sequence $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow$ is followed.

-Port Circulator

Carlin has shown that any lossless, matched, nonreciprocal three-port microwave junction is a perfect three-port circulator.^t This theorem is readily proved from the properties of the scattering matrix. A perfectly matched three-port junction has a scattering matrix of the form

$$\begin{matrix} & \mathbf{O} & \mathbf{S}_v, \mathbf{S} \mathbf{v} \\ \mathbf{S} & = & S_{21} & \mathbf{O} & S_{23} \\ & & S^* & S_{12} & \mathbf{O} \end{matrix}$$

(6.116)

For a nonreciprocal junction the scattering matrix is no longer symmetrical; that is, $S_{uv} \neq S_{vu}$. However, if the junction is lossless, conservation of power still requires that the $|S|$ matrix be unitary. Thus (4.63a) will hold for any lossless microwave junction independently of whether or not the junction is reciprocal. Applying the unitary condition to (6.116) gives

$$\begin{aligned} S_{12}S_{21}^* + S_{13}S_{31}^* &= 1 \\ S_{21}S_{12}^* + S_{23}S_{32}^* &= 1 \\ S_{13}S_{31}^* + S_{23}S_{32}^* &= 1 \\ S_{12}S_{23} - S_{13}S_{32} &= 0 \end{aligned}$$

Let us assume that $S_{21} \neq 0$. The fourth of the above equations then gives $S_{31} = 0$. The third equation now requires $|S_{32}| = 1$, and thus $S_{32} = 0$ from the fourth equation, $|S_{13}| = 1$ from the first equation, and $S_{13} = 0$ from the fourth equation again. Thus we see that $|S_{21}| = 1$ also from the second equation, so that

$$|\mathbf{S}_{21}| = |\mathbf{S}_{13}| = |\mathbf{S}_{32}| = 1$$

$$S_{12}S_{21}^* = S_{13}S_{31}^* = 1$$

^tH. J. Carlin, Principles of Gyrator Networks, *Polytech. Inst. Brooklyn. Microwave Res. Inst. Symp. Ser.*, vol. 4, p. 175, 1955.

Consequently, there is perfect transmission from port 1 into p_{or} , port 2 into port 3, and from port 3 into port 1. There is zero transmission in any other direction. The resultant scattering matrix of any matched three-port junction must then have the form

$$[S] = \begin{vmatrix} 0 & S_{12} & S_{13} \\ 0 & 0 & 0 \\ S_{21} & 0 & 0 \end{vmatrix} \quad (6-U7)$$

If the locations of the terminal planes in the three input lines are chosen, the phase angles of S_{13} , S_{21} , and S_{32} can be made zero and if

$$S_{13} = S_{21} = S_{32} = i$$

Practical realizations for three-port circulators usually involve the symmetrical junction (Y junction) of three identical waveguides or "gtn line" type of transmission lines, together with an axially magnetized ferrite rod or disks placed at the center. Figure 6.53 illustrates both a waveguide version and a balanced strip-line version of the three-port circulator. The ferrite rod or disks are magnetized by a static B_0 field applied along the axis and give the junction the required nonreciprocal property. By placing

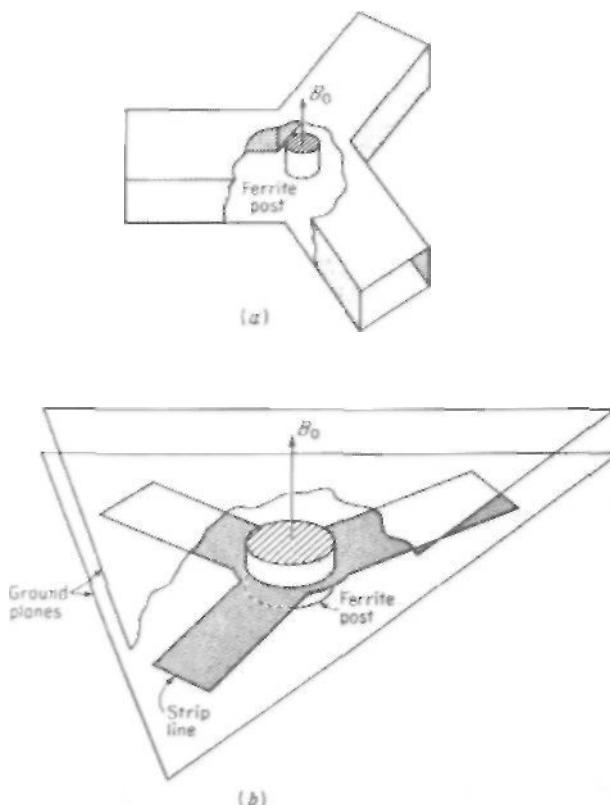


FIGURE 6.53
Three-port circulator
Waveguide vertf* line circulator

suitable tuning elements in each arm (these can be identical in each arm because of the threefold symmetry involved) the junction can be matched; that is, S_u , S_v , and S_{33} can be made zero. The analysis given above then shows that the *junction must*, of necessity, be a perfect circulator if all losses are negligible. Losses are, of course, always present, and this limits the performance that can be achieved. Typical characteristics that can be obtained are insertion loss of less than 1 dB. that is, $|S_{13}|$, $|S_{21}|$, $|S_{32}|$ greater than 0.89, isolation from 30 to 50 dB, and input reflection coefficients less than 0.2. The isolation that can be obtained corresponds to values of $|S_{11}|$, $|S_{22}|$, and $|S_{33}|$ in the range 0.01 to 0.03.

The junction circulator is an essential component used to separate the input and output ports in negative resistance amplifiers (see Sec. 11.5). Circulators are also used to couple a transmitter and receiver to a common antenna. Circulators ranging from miniature units that can handle a few watts of power up to units that can handle 100 or more kilowatts of average power are commercially available. Bandwidths of more than one octave have been obtained.

of Three-Port Circulator

The field analysis of a three-port circulator is based on the assumption that the electric field in the ferrite disk has only a single component \mathbf{E}_z in the axial direction. The applied dc magnetic field is also in the z direction. The permeability tensor then has the same form as in rectangular coordinates; so we can write

$$\mathbf{B}_r, \quad \mathbf{M}_r \quad [6.118]$$

where $i_r = 1 + A^*$ and $JK \sim X_{xy}$ in order to obtain a good circulator, the losses in the ferrite must be very small. If we neglect losses

$$\begin{aligned} M_r &= \frac{WQ}{(On - \epsilon)} + (O_0 U)_r \\ K &= \frac{WW}{2} \\ &\quad W_n - W \end{aligned}$$

Equation (6.118) is easily derived from (6.98) by expressing H_x , H_y , B_x and ϕ_V in terms of the r and $\langle l \rangle$ components of the fields.

From Maxwell's equations we obtain

$$\begin{aligned} \mathbf{V} \times \mathbf{E} &= -\mathbf{a}, \quad \mathbf{X} \mathbf{V}\mathbf{f} = \mathbf{a} - \frac{AE}{rH} - \mathbf{a} \cdot \mathbf{H} \mathbf{7} \\ &= -jcon \cdot \mathbf{H} = -j(on_0 a_r(tL_r H_r + JKH_{11})) \\ &\quad - J \ll M \ll J - jKH_r + n_r H_+ \end{aligned} \quad (6.119)$$

We can solve this equation for H_r and H^\wedge and use the results in the equation

$$\nabla \times H = \omega \epsilon_0 E, a,$$

to obtain

$$\frac{1}{r} \nabla^\wedge \cdot H = \frac{1}{r^2} \frac{\partial^2 E}{\partial r^2} + \frac{M}{r} \cdot \nabla \times H = 0 \quad (6.120)$$

where $A^2 = a V_0 \sim K^2 e / f_i r$. This equation has the same form as H_o for the axial electric field for TM modes in a circular waveguide but replacing k_c . The general solution for E_z is, by analogy, of the form

$$E_z = \sum_{m=0}^{\infty} (J_m(r) \cos(m\theta) + J_m(r) \sin(m\theta)) \cdot \frac{e^{imz}}{r} \quad (6.121)$$

where a_{m0} and b_{m0} are amplitude constants and J_m is the Bessel function of order m . The corresponding solution for the component of the magnetic field is

$$H_r = \sum_{m=0}^{\infty} (J'_m(kr) \cos(m\theta) - J_m(kr) \sin(m\theta)) \cdot \frac{e^{imz}}{r} \quad (6.122)$$

where $J'_m(kr) \sim dJ_m(kr)/dk_r$. We can interpret this solution as waves that circulate around the ferrite disk in the $\pm\theta$ directions. The terms multiplied by $a_{m0} e^{imz}$ are waves circulating in the $+z$ direction and the terms multiplied by $b_{m0} e^{-imz}$ are waves circulating in the $-z$ direction. Since H_{r0} is different for the two sets of waves, the ferrite disk clearly will exhibit nonreciprocal properties, an essential requirement for a circulator.

Consider now the circulator configuration shown in Fig. 6.54. Each of the three microstrip input ports are 120° apart and extend over a coupling angle α . In the field analysis of the circulator, it is assumed that H_a is constant over each coupling region and is zero along the remainder boundary of the ferrite disk. Thus at $r = a$ we let

$$H_r = i \begin{cases} \mathbf{ff}, & \theta > \alpha \\ \mathbf{H}_2, & 2\pi - \alpha < \theta < \alpha \\ \mathbf{H}_3, & \theta < -\alpha \end{cases} \quad (6.123)$$

The assumption of a constant magnetic field H_{r0} over θ is a first-order approximation to the magnetic field of the circulator. At $r = a$ the solution for H_{r0} given in Fig. 6.122 is

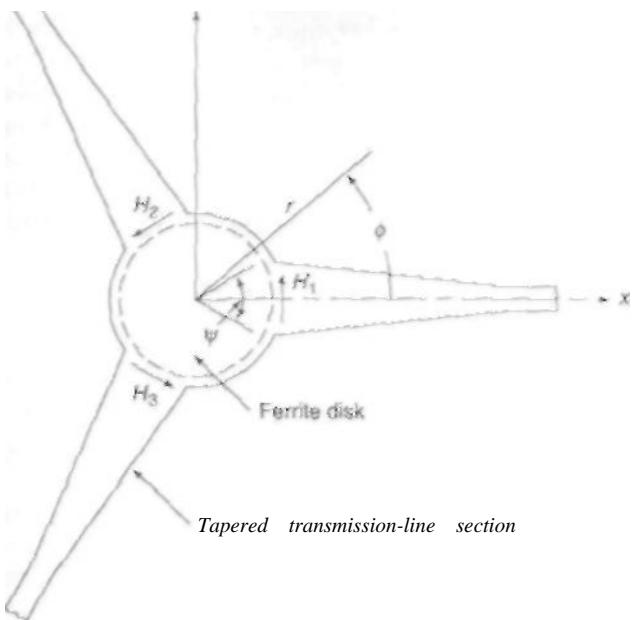


FIGURE 6.54
Three-port circulator showing coupling angle ϕ and tapered transmission lines for matching.

Fourier series; so the amplitude coefficients a_{α} and b_{α} are readily found by equating (6.122) to (6.123). When the coefficients a_{α} and b_{α} have been found by Fourier analysis, the electric field given by (6.121) can be evaluated at $r = a$.

In order to obtain a circulator, the condition $E_z = 0$ at $d = 4ir/3$ is imposed so that there will be no coupling into port 3. This condition establishes a constraint on the ferrite disk radius a , the magnetization M_s in the ferrite, and the coupling angle ϕ . In practice, these parameters are optimized so as to obtain the best circulator performance possible, over as broad a band of frequencies as possible. It is found that the optimum coupling angle ϕ is given approximately by

$$\phi = \frac{2\pi r}{1.84V_3} \frac{K}{M_r \ll M_s} \quad (6.124)$$

tH. Bosma, On Stripline Y-Circulators at UHK, *IRE Trans.*, vol. MTT-12, pp. 61-72, 1964.

V. S. Wu and F. J. Rosenbaum, Wide-Band Operation of Microstrip Circulators, *IEEE Trans.*, vol MTT-22, pp. 849-856. 1974.

This angle is typically around 30° to 50° and results in a strin-l' that corresponds to a low-impedance transmission line. Conse " ^ ^ ^ h pered transmission-line sections are often used to transfonnv,^{1^1} *** impedance into the 50-ii impedance of the input lines. The taper H ^{'o*} are illustrated in Fig. 6.54. Satisfactory circulator designs have bee hy taking as few as 3 to 6 modes into account, that is, n up to 6 ^{reved}

In the above account we have outlined the principal featur field analysis of circulators. The references at the end of the [chaDt.fr](#) J chapter provfc many of the details that we have omitted.

6.11 OTHER FERRITE DEVICES

The devices utilizing ferrites for their operation described in the preced" sections represent only a small number of the large variety of devices th-! have been developed. In addition to the above, there are other forms of isolators, both reciprocal and nonreciprocal phase shifters, electronically controlled (by varying the current in the electromagnet that supplies the static biasing field) phase shifters and modulators, electronic switches and power limiters, etc. The nonlinear property of ferrites for high signal levels has also been used in harmonic generators, frequency mixers, and parametric amplifiers. A discussion of these devices, together with design considerations, performance data, and references to the original literature, contained in the book by Lax and Button, listed in the references at the end of this chapter. The recent article by Rodrique gives a good survey of the present status of ferrite devices.t

PROBLEMS

- 6.1. Determine the values of $/?$ and R_0 , in the T and II attenuator netwl shown in Fig. 6.8 in order to obtain 6 dB of attenuation.
- 6.2. Derive the equations (6.7) for the II attenuator network.
- 6.3. For the electronically controlled attenuator shown in the photograpfi in 6.10, the return loss is 10 dB. Calculate the input VSWR (see Sec..
- 6.4. In the rotary phase shifter, show that if the output quarter-wav^P^^ transition, and rectangular guide are rotated by an angle B_v ^a phase change of 0, is produced in the transmitted wave.
- 6.5. For the phase shifter shown in Fig. 6.15. determine the length ⁱⁿ _{incremen-} wavelength) for the two short-circuited stubs in order to obtain tal phase change of 45° ^{,es aci l*}
- 6.6. For the phase shifter shown in Fig. 6.16. assume that the PW. ^{ghu,,(on t<} idea) switches. When the diodes are on, the susceptance placed i

main line is $jB_x = -j \cot pd_v$. When the diodes are off, the susceptance loading the main line is $jB_{2-} = j \tan(3d)$. Show that by choosing $d_t = A_s/4$ I S and $d = A_s/2 - \delta$, If, $= -B_2$.

- 6.7. For the phase shifter shown in Fig. 6.15, the transmission coefficient $T_M = \sim / [1 + jB - B^2/2]$, where jB is the susceptance of the stub. By using $1 - / j - = \underline{T} \underline{I} \underline{J} \underline{R}$, find the magnitude p of the reflection coefficient and the input VSWR when $B = 0.2$ and when $B = 0.4$.

Answer: VSWR = 2.04. 2.273

- 6.8. For a Bethe-hole directional coupler with the two guides aligned ($\theta = 0$) and a centered aperture, why does not (6.28) give a useful solution for A_0 as a function of a whereas (6.29) does?

- 6.9. Design a Bethe-hole directional coupler with a centered circular aperture. The waveguide size is $a = 0.9$ in. $b = 0.4$ in. The center frequency is 9.8 GHz. The required coupling is 30 dB. Find the aperture radius and the frequency band over which the directivity remains greater than 20 dB.

- 6.10. Design a Bethe-hole coupler based on (6.29) and (6.30). Assume $a = 0.9$ in., $b = 0.4$ in., $f = 9$ GHz, and $C = 30$ dB. Find the aperture position d and radius r . Evaluate C and D as a function of frequency over the band 8.5 to 10.5 GHz and compare the performance with that shown in Fig. 6.20.

- B.11. Figure P6.11 illustrates two rectangular guides coupled by circular apertures in a common side wall. A TE₀₁ mode of unit amplitude radiates a field of amplitude $j^{r^2 - rr/a} Habi$ in both directions in the other guide. Design a five-hole directional coupler of the binomial type. The coupling required is 30 dB at a frequency of 10 GHz. The guide width a is 2.5 cm. and the height $b = 2.2$ cm. Find the required aperture radii and the frequency band over which the directivity D remains greater than 40 dB.

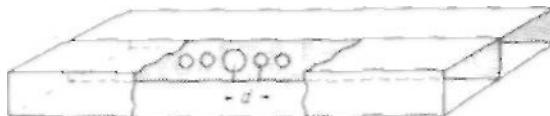


FIGURE P6.11

- 6.12. For the coupler described in Prob. 6.11, find the aperture radii to give a Chebyshev coupler. The minimum value of directivity required is $D_m = 50$ dB. Find the corresponding frequency band. How much greater bandwidth is obtained as compared with that of the binomial coupler of Prob. 6.11?

- 6.13. Design a three-hole Chebyshev directional coupler using rectangular waveguides with $a = 0.9$ in., $b = 0.4$ in. The center frequency is 9 GHz. The coupling and minimum directivity D_m contributed by the array factor is 20 dB and 30 dB at 9 GHz. The apertures are located at $x_{\text{ap}} = a/4$. Plot C and D as a function of frequency and compare the performance with that shown in Fig. 6.24.

- 6.14. Repeat the design problem given in Prob. 6.13 but instead of using $D_m = 30$ dB, the requirement is that the fractional bandwidth $A_0/0_{(0)} = 0.1$, where $B_0 = 1.29$ at 9 GHz and $A_0 = \#_8 - p_1$ with $/S_2 = (S_1 + 0.05/\#_u)$ and $/\#_1 = /?_1 - 0.050$,. What is the resultant value of D_{min} for this design?

- 6.15.** Design a three-hole Chebyshev directional coupler using cent the common broad wall between two rectangular waveguides (*f aperlur es i_n* center frequency is 9.5 GHz. The waveguide width equals $Otf^{-6,23,1}$, "be height equals 0.4 in. The required coupling is 30 dB and th^m ^nd th^e directivity is 30 dB. Find the aperture radii, spacing, and the band⁶ ?'niini ni awWth of the coupler. Why is this not a good design?
- 6.16.** Design a broadside-coupled strip-line directional coupler with 3 HR See Figs. 6.27 and P6.16 for details. The ground-plane spacing "S2 CQUpng input and output lines have a characteristic impedance of $5n$ n p.^{1"} ^e required strip width W and spacing S . Find W so as to obtain $Z = \sqrt{\mu_0 \epsilon_0}$ the input and output lines. For this calculation assume that th^K ? centered between the ground planes. In the actual coupler the 1^c upper lines are not centered, but since the spacing S is small the chamT["] ^ will also be small. The computer program CSTPL should be used to obt["] i needed line parameters. Find the length of the coupled section when["] frequency is 4 GHz. Assume that the strips have a negligible thickness



FIGURE P6.16

- 6.17.** A branch-line coupler of the type shown in Fig. 6.29 and having 6-dB coupling at a frequency of 5 GHz is needed. The input and output transmission lines have $Z_c = 50 \Omega$. Determine the characteristic impedances of the through lines and branch lines and their lengths in terms of wavelength. The coupler is to be built using microstrip lines on a substrate with a dielectric constant of 1 and 1 mm thick. Determine the widths of all transmission lines and the lengths of the through lines and branch lines. Use the computer program MSTP to determine the microstrip line parameters. Use the computer program MSTPD to determine the effective dielectric constants at 5 GHz and modify the transmission-line lengths so as to take into account the dispersion in the effective dielectric constant at 5 GHz.
- 6.18.** The hybrid ring shown in Fig. 6.33 is constructed using microstrip lines on 1-mm-thick substrate with a dielectric constant of 2.3. The input and output transmission lines have a characteristic impedance of 50Ω . Determine the width of the transmission lines and of the ring as well as the size of the ring. The frequency of operation is 2 GHz. Use the computer program MSTP.
- 6.19.** Determine the impedances and the value of the resistor R in the divider shown in Fig. 6.38. The power in port 2 is one-half of that in port 3.
- 6.20.** Use the equation of motion (6.94) to study second-harmonic generation in a ferrite. Assume that $M = M_0 \sin(\omega t) + M_2 \sin(2\omega t)$ for the ac field. The magnetic field is $H_s e^{j\omega t} a_x$. Neglect the third-harmonic term and show that

$$2j < oM_2 = y \gg_n H_s a_x, X M_0, + y p_0 t_0^*, * ***$$

Thus $2JMM_{2C} = W_u H_{M_{1v}}$. For M_{1v} , take the small-signal solution * , //, and use the value of X_{Vx} ^a* resonance for a lossy ferrite to show that

$$M_{as} = \underline{\hspace{2cm}}$$

Note that, for good efficiency, a must be small (small damping), so that the precession angle will be large at resonance.

- 6.21.** From the unitary properties of the scattering matrix for a lossless nonreciprocal two-port microwave junction, show that it is not possible to have S_{21} zero while S_{12} is finite. Thus a lossless one-way transmission device cannot be built.

6.22. Continue the argument in the text to verify that the transmission sequence $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ is followed in the circulator illustrated in Fig. 6.52.

6.23. Show that the scattering matrix for an ideal lossless JV-port circulator can be put into the form

$$[S] = \begin{vmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & \cdots & 1 & 0 \end{vmatrix}$$

by choosing proper terminal-plane locations in each port.

- 6.24.** Show that, for TEM-wave propagation in a direction perpendicular to B_0 , in an infinite ferrite medium, the two solutions are linearly polarized waves with propagation constants

$$Y^{-1} j \omega \sqrt{\mu_0 \epsilon}$$

Hint: Consider propagation along x , and in one case assume E to have only a y component, and for the other case assume E to have a z component

1. Montgomery, C. G., R. H. Dicke, and E. M. Purcell: "Principles of Microwave Circuits," McGraw-Hill Book Company, New York, 1948.
 2. Ragan, G. L. (ed.): "Microwave Transmission Circuits." McGraw-Hill Book Company, New York, 1948.
 3. Marcuvitz, N. (ed.): "Waveguide Handbook." McGraw-Hill Book Company, New York, 1951.
 4. Harvey, A. F.: "Microwave Engineering." Academic Press, Inc., New York, 1963. An outstanding handbook, offering a comprehensive survey of the whole microwave field together with an extensive bibliography covering the international literature.
 5. Bahl, I., and P. Bhartia: "Microwave Circuit Design." John Wiley & Sons, Inc., New York 1988.
 6. Chang; K. (ed.): "Handbook of Microwave and Optical Components. Microwave Passive and Antenna Components," vol. 1, John Wiley & Sons, Inc., New York, 1989.
 7. D. M. Pozar, "Microwave Engineering." Addison-Wesley Publishing Company, Reading Mass., (990).

Microwave ferrites

8. Lax, B., and K. J. Button: "Microwave Ferrites and Ferrimagnetics" Mef Company. New York, 1962. A very complete treatment of the physical properties of ferrites, electromagnetic propagation in ferrites, measurement techniques of many ferrite devices. Includes a 24-page bibliography.
9. Soohoo, R. F.: "Theory and Application of Ferrites," Prentice-Hall Inc., N.J., 1960.
10. Clarricoats, P: "Microwave Ferrites." John Wiley & Sons, Inc., New York, IQRI Consultants Bureau, New York, 1963.
11. Roberts, J.: "High Frequency Application of Ferrites," D. Van Nostrand Princeton, N.J., 1960.

Field theory of circulators

12. Bosma, H.: On Stripline Y-Circulators at UHF. *IRE Trans.*, vol. MTT-12, p. 61-72.
13. Fay, C. E., and R. L. Comstock: Operation of the Ferrite Junction Circulator /*BPi Trans.*, vol. MTT-13, pp. 15-27, 1965.
14. Wu, Y. S., and F. J. Rosenbaum: Wide-Band Operation of Microstrip Circulators *IEEE Trans.*, vol. MTT-22, pp. 849-856, 1974.
15. Schloemann, E., and R. E. Blight: Broad-Band Stripline Circulators Based on YIG and Li-Ferrite Single Crystals. *IEEE Trans.*, vol. MTT-34, pp. 1394-1400, 1986.

CHAPTER 7

ELECTROMAGNETIC RESONATORS

RESONANT CIRCUITS

Resonant circuits are of great importance for oscillator circuits, tuned amplifiers, frequency filter networks, wavemeters for measuring frequency, etc., at all frequencies from a few hertz up to and including light frequencies. Electric resonant circuits have many features in common, and it will be worthwhile to review some of these by using a conventional lumped-parameter *RLC* parallel network as an example. Figure 7.1 illustrates a typical low-frequency resonant circuit. The resistance *R* is usually only an equivalent resistance that accounts for the power loss in the inductor *L* and capacitor *C* and possibly the power extracted from the resonant system by some external load coupled to the resonant circuit.

At resonance the input impedance is pure real and equal to *R*. This implies that the average energies stored in the electric and magnetic fields are equal, since from (2.60)

$$Z_{in} = \frac{P_r + 2ja(W_m - W_e)}{2^{11}} \quad (7.1)$$

This equation is valid for any one-port circuit provided a suitably defined equivalent terminal current i is used. Thus resonance always occurs when $W^e = W_g$, if we define resonance to be that condition which corresponds to a pure resistive input impedance. In the present case the time-average energy stored in the electric field in the capacitor is

$$W_c = |VV^*C|$$

I-

L a ^ FIGURE 7.1
Lumped-parameter resonant circuit.

and that stored in the magnetic field around the inductor is

$$W_m = \frac{1}{4} \frac{LI}{A} J^2 = \frac{1}{4} \frac{\epsilon}{w^2 L} V^2$$

The resonant frequency w_0 is now found by equating W and

$$\mathbf{Q}_0 = (\mathbf{L} \mathbf{0})^{-\frac{1}{2}} \quad (7.2)$$

An important parameter specifying the frequency selectivity, and performance in general, of a resonant circuit is the quality factor, or Q . A very general definition of Q that is applicable to all resonant systems is

$$Q = \frac{w(\text{time-average energy stored in system})}{\text{energy loss per second in system}} \quad (7.3)$$

At resonance $W_m = W$, and since the peak value of electric energy stored in the capacitor is $2W_c$, and occurs when the energy stored in the inductor is zero, and vice versa, the average energy W stored in the circuit is

$$W = W_{m,} + W_c = 2W_c = kCW^* \quad (7.4)$$

The power loss is GW^* and is the energy loss per second. Hence, for the circuit of Fig. 7.1,

$$\begin{array}{ccc} \frac{1}{j\omega C} & & R \\ G & & \\ & & \frac{1}{j\omega L} \end{array} \quad (7.5)$$

since $\omega^2 LC = 1$ at resonance and $G = R$.

In the vicinity of resonance, say $\omega = \omega_0 + Aw$, the input impedance can be expressed in a relatively simple form. We have

$$Z_{in} = \left(\frac{1}{R} + \frac{1}{j\omega L} + j\omega C \right)^{-1} = \frac{1}{|R|} + j\omega_0 C + j\Delta\omega C + \frac{1 - AW/G > 0}{j\omega_0 L}$$

where the approximation $1/(w_0 + Aw) = (1 - Aw/w_0)/\omega_0^2$

Since $j\omega_0 C + 1/j\omega_0 L = 0$, we obtain

$$Z_{in} = \frac{(A'JRL)}{A'^2 RL + j2RAa} = \frac{R}{1 + j2Q/ba > / < *o}$$

A plot of Z_{in} as a function of Aw/w_0 is given in Fig. 7.2, and is a typical resonance curve. When $|Z_m|$ has fallen to 0.707 of its maximum value, the phase is 45° if $\omega < \omega_0$ and -45° if $\omega > \omega_0$. From (7.6) the

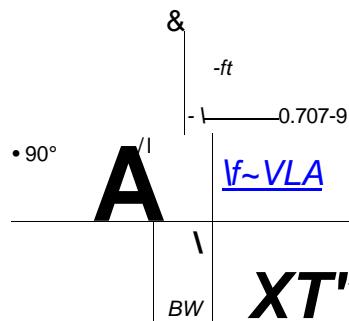


FIGURE 7.2
 $Z_{\text{,,}}$ for a parallel resonant circuit.

value of A_w is found to be given by

$$\frac{A \text{ to}}{2Q} = 1$$

or

$$A_w = 2Q$$

The fractional bandwidth BW between the $0.707ft$ points is twice this; hence

$$Q = \frac{1}{2Au >} \frac{1}{BW} \quad [7.7]$$

This relation provides an alternative definition of the Q ; that is, the Q is equal to the fractional bandwidth between the points where $|Z_{\text{,,I}}$ is equal to 0.707 of its maximum value (for a series resonant circuit this definition applies to $|Y_m|$).

If the resistor R in Fig. 7.1 represents the losses in the resonant circuit only, the Q given by (7.5) is called the *unloaded Q*. If the resonant circuit is coupled to an external load that absorbs a certain amount of power, this loading effect can be represented by an additional resistor R_L in parallel with R . The total resistance is now less, and consequently the new Q is also smaller. The Q , called the *loaded Q* and denoted Q_{L} , is

$$Q_L = \frac{RR_L/(R + R_L)}{(oL)}$$

The external Q , denoted Q_e , is defined to be the Q that would result if the resonant circuit were loss-free and only the loading by the external load were present. Thus

Use of these definitions shows that

1 J_ 1

Another parameter of importance in connection with a resonant circuit is the damping factor δ . This parameter measures the rate at which the oscillations would decay if the driving source were removed. For a high-Q circuit, δ may be evaluated in terms of the Q , using a perturbation technique. With losses present, the energy stored in the resonant circuit will decay at a rate proportional to the average energy present at any time (since $P \propto W^*$ and $W \propto VV^*$, we have P_aW), so that

$$\begin{aligned} dW \\ \text{---} &= -2SW \quad (7.9) \\ \text{at} \\ w &= W_n e^{\lambda t} \quad (7.10) \end{aligned}$$

where W_u is the average energy present at $t = 0$. But the rate of decrease of W must equal the power loss, so that

$$\frac{dW}{dt} = {}_2 S W = P,$$

Consequently,

$$S = \frac{P}{2W} \sim \frac{\lambda f_i - JL}{2O} \quad (7.11)$$

upon using (7.3). The damping factor is seen to be inversely proportional to the Q . In place of (7.10) we now have

$$W \equiv W_0 e^{m/Q}$$

In (7.12) Q must be replaced by Q , if an $\text{d}^1 \text{t}^{\text{e}}$ circuit. The damping factor δ is also a measure of how fast oscillations in the resonant circuit can build up upon application.

S o TM n microwave systems sections of transmission fines^{J*}
 sures called cavities are used as resonators in place of the $\frac{f}{g}$
 circuit. The reason for this is that lumped J_s be $\ll J^*$ losses
 from both conductor loss and radiation loss $\frac{\text{ave}}{\text{ca}}$
 in "wave frequencies. In calculating the impedance $\frac{s}{e_n}$, The J
 it is sometimes convenient to assume there a e^P impedance f can
 be evaluated separately, and in terms of $i_h^* \times T M M f_{\text{onant}}$ fire****
 be modified to account for losses by replacing the $\frac{=}{w}$, W
 an equivalent complex resonant frequency $w, U + //$

that (7.6) can be written as

$$\frac{w_0 R / 2Q}{j[w - u >_o (l + j/2Q)]} \quad (7.13)$$

which shows that when losses are present this is equivalent to having a complex resonant frequency $w_0(l + j/2Q)$. Equation (7.13) neglects the small change in resonant frequency that occurs when small losses are present.

7.2 TRANSMISSION-LINE RESONANT CIRCUITS

Series Resonance; Short-Circuited Line

At high frequencies, usually in the range 100 to 1,000 MHz, short-circuited or open-circuited sections of transmission lines are commonly used to replace the usual lumped *LC* resonant circuit. It is therefore of interest to consider the order of magnitude of *Q* and impedance that can be obtained. It will be assumed that air-filled lines are used, so that the only losses are those due to the series resistance *R* of the line. This is usually the case in practice since a dielectric-filled line has some shunt conductance loss and hence would result in a lower *Q*.

Consider a short-circuited line of length *l*, parameters *R,L,C* per unit length, as in Fig. 7.3. Let $l = A_0/2$ at $f = f_{0s}$ that is, at $\omega_0 = \omega_{0s}$. For f near f_0 , say $f = f_0 + Af$, $\text{fil} = lirfl/c = T(U/W_0) = \pi + 77 \text{ rad/sec}$, since at ω_0 , $\text{fil} = v$. The input impedance is given by

$$Z_{in} = Z_v \tanh(jliI + al) = Z_c \frac{\tanh al + i/fit}{1 + j \tan j5l \tanh \alpha/l}$$

But $\tanh al = al$ since we are assuming small losses, so that $al < 1$. Also $\tan fit = \tan(\omega_0 t + Aa)/(\omega_0) = \tan IT \text{ rad/sec} = \sim Au/\omega_0$ since $|u|/\omega_0$ is

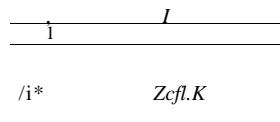


FIGURE 7.3
Short-circuited transmission-line resonator.

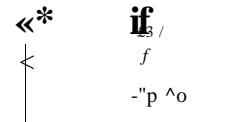


FIGURE 7.4
A series resonant circuit

small. Hence

$$\frac{aI}{Z_m} = \frac{1}{\omega^2 C + R_Q l + jL_0 I} \quad (7.14)$$

since the second term in the denominator is *very small*. Now

$a = |RY_r| = (R/2)JC/L$, and $0/ = co_0 l C l = TT$; SO $TT/w_0 = / \wedge c$ and the expression for Z_m becomes

$$Z_m = \frac{1}{fC} - \frac{|R_Q|}{2} \frac{1}{fL_0} + j \frac{1}{w_0} = \frac{1}{2} f_i + i \frac{1}{L_0} A_a \quad (7.15)$$

It is of interest to compare (7.15) with a series $R_Q L_0 C$ circuit illustrated in Fig. 7.4. For this circuit

$$Z_{lr} = S_0 + jL_0 l - jD \sim L_0 C_0$$

If we let $oig = 1/L_0 C_0$, then

$$Z_{in} = f_o + jw Z_{lr}$$

Now $w^2 - oig = (at - w_0)(u + w_0) = AOJ 2w$ if $a - co_0 = Aw$ is small.

$$Z_{in} = f_o + jw L_0 = \frac{2(oAw)}{2} = f_o + j' 2 L_0 A_c \quad (7.16)$$

By comparison with (7.15), we see that in the vicinity of the frequency I which $/ = A_0/2$, the short-circuited line behaves as a series resonant arc with resistance $R_Q = \wedge Rl$ and inductance $L_0 = \wedge Ll$. We note that \wedge are the total resistance and inductance of the line; so we might wonder why the factors \wedge arise. These enter because the current on the short-circuited line is a half sinusoid, and hence the effective circuit parameters no longer one-half of the total line quantities. \wedge

The Q of the short-circuited line may be defined as for the a *Fig. 7.4:*

$$Q = \frac{a >_0 L_0}{R_f} = \frac{<_0 Q L}{R} = \frac{\mathcal{Z}}{2a} \quad (7.17)$$

As an alternative, the general definition (7.3) may be used, $\wedge \wedge s$ to evaluate the Q of the short-circuited line from this definition j

an approximate method valid for high-Q (i.e., low-loss) systems in order to illustrate a method of great utility in connection with many microwave devices (see Sec. 3.8, where this perturbation method is discussed). For small losses the energy stored in the system is, to first order, the same as if *no losses* were present. For a loss-free short-circuited line, the current on the line is a pure standing wave

$$I = I_0 \cos(\omega z)$$

where z measures distance along the line from the short toward the input end. In a length dz , the energy stored in the magnetic field is

$$dW_m = |H^*Ldz = |I^*L \sin^2 iizdz$$

The total time-average stored energy in the magnetic field is

$$W_m = \frac{1}{4} X_L \int_0^L \cos^2 iiz dz = \frac{1}{2} L I_0^2 \quad (16)$$

The energy stored in the electric field, i.e., in the line capacitance, is equal to W_m at resonance; so the time-average energy stored in the system is

$$W = W_m, \text{ if } f = \frac{1}{2} L I_0^2$$

To a first approximation the losses do not modify the current distribution along the line. Hence the power loss is given by

$$P = \sim r R n^* dz = \sim i s r \partial os^* z dz = \sim m^2$$

Thus, at $\omega = \omega_0$,

$$Q = \frac{\omega_0 A_u L I^*/8}{K \pi A_0 / 8} = \sim R \quad (7.18)$$

which checks with the earlier result. Typical values of Q range from several hundred up to about 10,000. As contrasted with low-frequency lumped-parameter circuits, the practical values of Q are very much higher for microwave resonators. It should be noted that in the above analysis the losses in the short circuit have been neglected. This does not introduce appreciable error if the length L is considerably greater than the conductor spacing.

Line

By means of an analysis similar to that used earlier, it is readily verified that an open-circuited transmission line is equivalent to a series resonant circuit in the vicinity of the frequency for which it is an odd multiple of a

$?t$

FIGURE 7.5

ST C_Q resonator.

quarter wavelength long. The equivalent relations are (Fig. 7.5)

$$\begin{array}{c} / - \text{A} \circ \\ 4'' \end{array} \quad * \quad (7.19a)$$

$$7 \quad (\quad , \quad Aw \quad w_{\text{in}} \quad I \quad (7.196)$$

$$Ro = 2^{Rl} \quad (7.19c,$$

$$L_Q = -LI \quad (7.19rf)$$

$$<-(L_o C_o) \quad (7.190)$$

Antiresonance

Short-circuited transmission lines behave as parallel resonant circuits in the frequency range where they are close to an odd multiple of a quarter wavelength long. The same property is true of open-circuited lines that are 1 multiple of a half wavelength long. When they behave as parallel resonant circuits, they are said to be antiresonant. $\wedge \wedge \wedge *$

The case of parallel resonance is best analyzed on an admittance. With reference to Fig. 7.6, let $/$ equal $A_0/4$ at $w = w_0$. Then

$$III = w_0 f L Cl = w_0 f l \sim Cl + A_0 f L Cl$$

n.

r <. oc.B

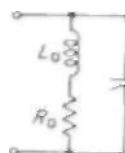


FIGURE 7.6

Short-circuited transmission line.

at co , and the input admittance is given by

$$\begin{aligned} Y_{in} &\sim Y_c \coth(a + j\tilde{f}l) \\ Y_c &= \frac{1 + 7 \tan pi \tanh al}{\tanh al + j \tan /3/} \\ &\quad 1 - jal/(\pm wLjLC) \\ al &\quad -j/(Acoy/LC) \end{aligned}$$

since

$$\begin{aligned} \tan(w_0/\sqrt{LCZ} + AWLC/) &= \tan I^{\frac{\pi}{2}} + Ao>/\sqrt{LC}/ \\ &= -(Awv/LC)^{-1} \end{aligned}$$

and $\tanh al \sim al$. A further approximation yields

$$\begin{aligned} Y_m &= Y_c \frac{j \ln jLCl + al}{I + j \ln al^2/LC} \\ &= Y_c(jAa>yfLCl + al) \\ &= \frac{RC}{R_n + jcoL_n} - 1 + jAcoCl \end{aligned} \tag{7.20}$$

after replacing Y_c by $\frac{1}{\sqrt{LC}}$ and a by $i.R/2$.

For the parallel resonant circuit of Fig. 7.6, we have

$$\begin{aligned} &\frac{1}{K_0 + jcoL_0} \\ &\frac{jcoC_0(R_0 + jcoL_0) + 1}{R_n + jcoL_n} \\ &JCOQRQ = co LQCQ + 1 \end{aligned}$$

since we assume $R_a \ll wL_0$. If we define w_0 by $\omega_0^2 = 1/L_0C_0$, then

$$C_0 i?_0 = co_n L_0 C_0 - co L_0 C_0$$

$$f_0 = \frac{(a>_0 - OJ)(<U_0 + co)}{J^L o^C i}$$

$$* \quad ^\wedge o T^1 + . / C_0 2 A W \tag{7.21}$$

Comparison with (7.20) shows that the short-circuited line in the quarter wavelength long is equivalent to a parallel resonator or

$$\frac{C_n R_Q}{\sim l} = \frac{RC}{CI} = \frac{R}{2C_0}$$

The Q of the circuit is given by

$$Q = \frac{a^2}{2i_0} \cdot \frac{1}{R} = \frac{2\sigma}{R} \quad (7.22)$$

Although sections of transmission lines behave as simple two-parameter resonant circuits in the vicinity of a particular resonance frequency, they are in reality a much more complicated network having infinite number of resonance and antiresonance frequencies. The resonance frequencies occur approximately when the short-circuited line is a multiple of a half wavelength long, that is, $f_s = \pi c/2l$, and the antiresonance frequencies (parallel resonance) when the line is an odd multiple of a quarter wavelength long, that is, $f_a = (2n + 1)c/4l$, where n is an integer. Thus the exact equivalent circuit would consist of an infinite number of resonant circuits coupled together. However, in practice, the frequency range of interest is normally such that a simple single-resonant-frequency circuit represents the transmission-line resonator with adequate accuracy.

7.3 MICROSTRIP RESONATORS

In microwave circuits that use microstrip transmission lines, an open-circuited section of microstrip line may be used as a simple resonator. In Fig. 7.7a we show a microstrip resonator that is capacitively coupled to the input microstrip line. In Fig. 7.7b we show a similar resonator that is capacitively coupled to the side of the microstrip line. An alternative resonator config-

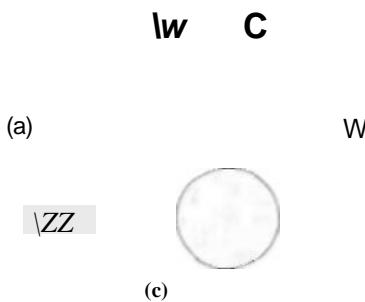
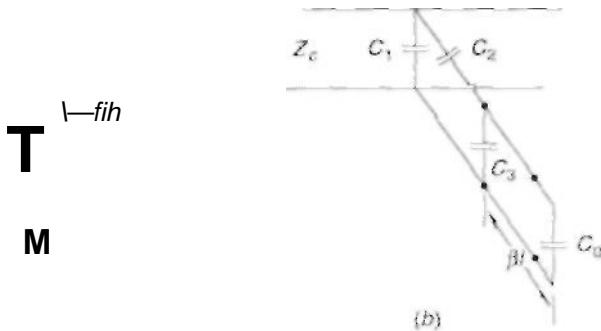


FIGURE 7.7

consisting of a single approximatively one-half waveguide section and a circular disk or resonator coupled to a microstrip line, the same resonator but coupled to the side of the microstrip line.

**FIGURE 7.8**

The equivalent circuit of the resonators shown in Figs. 7.7o and b.

ration is the circular disk that is capacitively coupled to a microstrip line as shown in Fig. 7.7c. The equivalent circuit for the resonators shown in Figs. 7.7a and 6 is a transmission line of length l with a capacitor C_c , at the open-circuit end and a capacitive II network that represents the coupling region as shown in Fig. 7.8. The fringing electric field at the open-circuit end results in additional charge on the microstrip line near the open-circuit end and this is modeled by the capacitor C_{ci} . The capacitor C_{ci} makes the electrical length of the resonator appear to be longer than the physical length l , because the line length has to be increased beyond $A/2$ to provide the additional inductance that will compensate for the additional capacitance. Resonators of the type shown in Figs. 7.7a and b are commonly used in microstrip filters (see Chap. 8).

In order to design a resonator of the type shown in Fig. 7.7a, the open-circuit capacitance C_c , and the values of C_{ci} , C_2 , and C_3 , in the coupling network must be known. These parameters have been evaluated by Silvester and Benedek using the assumption of static potential fields.^t The results are accurate for frequencies up to several gigahertz for microstrip lines with substrates 1 mm thick and even at higher frequencies for thinner substrates. In Fig. 7.9 we show typical values of the open-circuit capacitance C_0/W , where W is the width of the microstrip line. In Fig. 7.10 we show typical results for C_{ci} , C_2 , and C_3 for various microstrip parameters including the gap spacing s .

^tP. Silvester and P. Benedek, Equivalent Capacitances of Microstrip Open Circuit^{*}: IEEE Trans., vol. MTT-2G, pp. 511-516. 1972.

P. Benedek and P. Silvester, Equivalent Capacitances for Microstrip Gaps and Steps, IEEE Trans., vol. MTT-20, pp. 729-733, 1972.

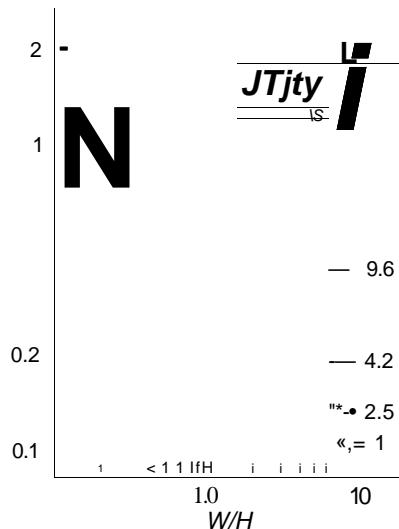


FIGURE 7.9

Microstrip-line open-circuited capacitance.
(Based on Fig. 6 of Silvester and Benedek.)

The resonator and coupling circuit shown in Fig. 7.8a can be represented by the equivalent lumped-parameter circuit shown in Fig. 7.11a, in the vicinity of the resonant frequency. In these circuits the capacitance C , and inductance L , are those contributed by the transmission line and are given by

$$c_s = |ci| \quad (7.230)$$

$$(7.236)$$

where C and L are the distributed capacitance and inductance per unit length for the microstrip line and I is the length of the resonator. The factor $\sqrt{2}$ arises because the voltage and current standing waves on the line are sinusoidal, so that the stored electric and magnetic energies are each one-half from what one would have for constant voltage and current amplitudes, and this makes the effective capacitance and inductance each one-half of those for the equivalent lumped-parameter model. The total loss resistance R represents the total losses in the resonator.

For the circuit in Fig. 7.11a, the input admittance at the port is given by

$$Y_{IM} = j \frac{C_0 + j Y_e t a x i p l}{C_0 - w C_0 \tan \phi} \quad (7.237)$$

as reference to Fig. 7.8a clearly shows. For convenience we denote C_0 as wC_V , C_1 as wC_0 , and C_2 as wC_S . The input admittance is given by

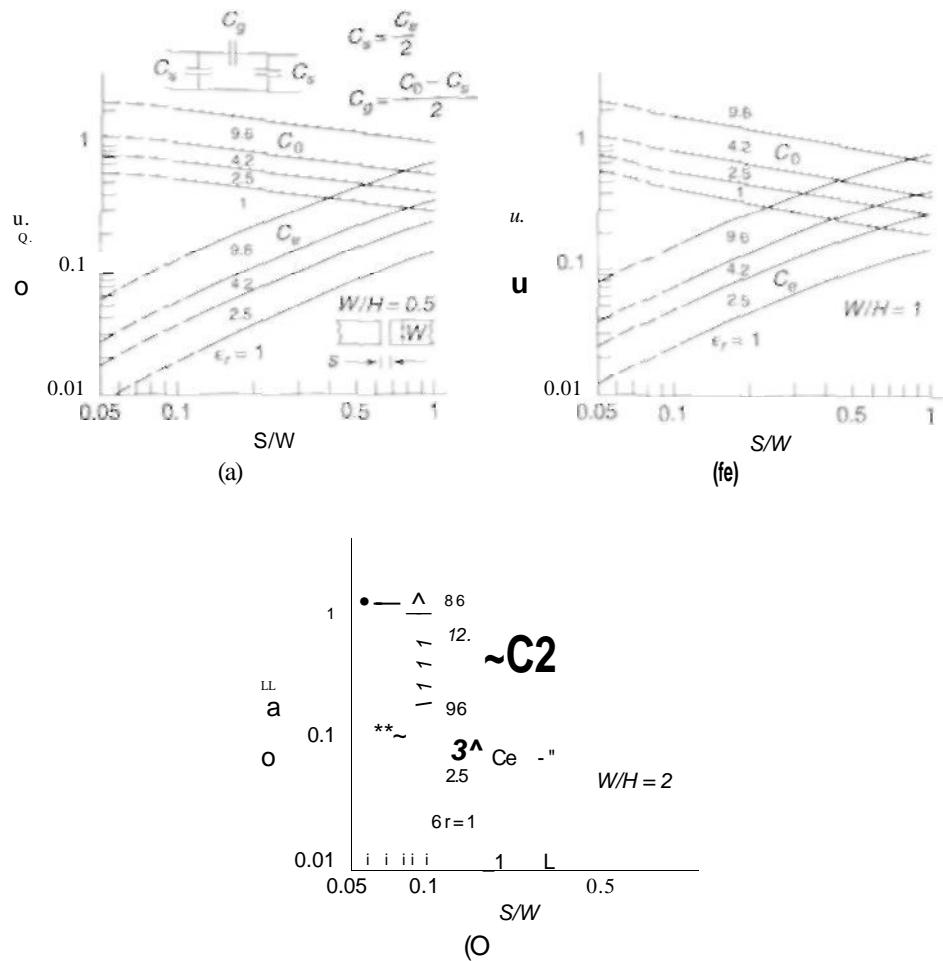


FIGURE 7.10

The Il-network capacitances for a microstrip-line gap. 1 Based on Fig. 7 of Benedek and Silvester.)

across C , is given by

$$\begin{aligned}
 & Y^{\wedge} - jB, \quad \frac{JB_{\wedge}U}{JB_2 + Y_{;n}} \\
 & \underline{j(B, \frac{B_2 Y_{;n} - B_1 B_2}{JB_2 + Y_{;n}})} \quad (7.24)
 \end{aligned}$$

In a gap coupled microstrip resonator $B_t \ll S_2$ so it is operated as a series resonant circuit. Series resonance occurs when the imaginary part of the

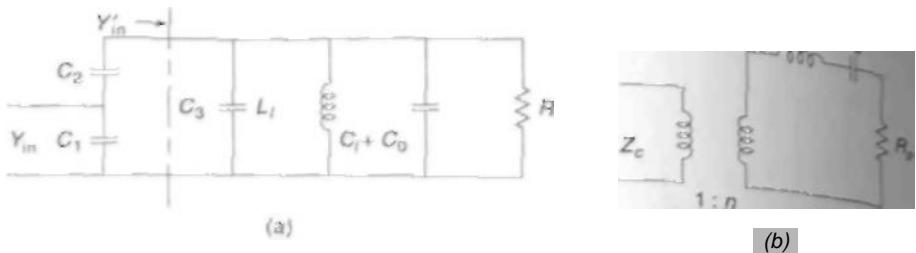


FIGURE 7.11

(a) Equivalent lumped-parameter circuit for the resonator and coupling network
7.8a; (b) a simplified equivalent circuit incorporating an ideal transformer

?*

denominator vanishes. The condition for resonance is thus

$$Y_{in} - \frac{1}{G + jB_m} = J^B m \quad \sim J^B 2 \quad (7.25)$$

When we substitute for Y_{in} , we obtain an equation for $\tan pi$ which is

$$\tan pi = \frac{Y_C(B_2 + B_0 + B_{.3})}{B_0(B_2 + B_{.3}) - Y_c^2} \quad (7.261)$$

Since B_2 , $B_{.3}$, and B_0 are all very small, pi will be somewhat smaller than $\pi/2$ since the expression for $\tan pi$ is small and negative. The resonant length is shorter than one-half wavelength because of the capacitive loading at the input and output ends. When we use (7.25) in the numerator in (7.24) and assume that we can approximate $G + jB_m$ by $-jB_2$ for all values of ω in the vicinity of the resonant frequency, we are able to express (7.24) in the form

$$Y_{in} = \frac{B_2}{G + jB_m + B_2}$$

C f HI

When we compare this expression with that for the input admittance circuit in Fig. 7.116, we see that

$$Z_{in} = M \frac{j\mu L_1}{R_s + jS1G + \frac{B_0}{n}} \quad (7.262)$$

so the ideal transformer turns ratio is

$$n = \frac{B_2}{G}$$

where R_s can be chosen arbitrarily.

Thus we have

$$F_{in} = \frac{1}{G(C_2 + C_3 + C_0) + jY_c \tan pi + G}$$

In (7.28) we have used the approximation $Y_{-n} = j\omega C_{-i} + j\omega C_0 + jY_r \tan \pi i$ + G which is obtained by replacing the denominator factor $Y_e - \omega C_0 \tan \pi i$ by Y_c since $\omega^2 C_0 \tan \pi i$ is a second-order small term near resonance because both $j\omega C_0$ and $\tan \pi i$ are small. The term $jY_c \tan \pi i$ can be represented by a parallel connection of C_t and L , in the vicinity of resonance as discussed earlier for transmission-line resonators. Thus, in the equivalent series resonant circuit shown in Fig. 7.11b, the total inductance is the sum of C_t , C_0 , C_3 , and C_2 , that is,

$$\frac{1}{n^2} = \frac{1}{C_2 + C_3 + C_0 + t?} \quad (7.29)$$

When the dispersion of the microstrip line can be neglected, the Q is given by (7.17) as

$$Q = \frac{P}{R} = \frac{1}{R} * \ll_6 C_{fl} \quad (7.30)$$

where R is the equivalent parallel resistance that will account for conductor and dielectric losses in the microstrip line. From the examples considered in Chap. 3, we found that typical values of attenuation for a microstrip line was in the range 0.5×10^{-2} to 1.5×10^{-2} Np/wavelength. Hence typical values of Q obtained from (7.30) would be 200 to 600. There is some additional loss caused by radiation of power from the open-circuited end. Measured values of Q for microstrip-line resonators are usually in the range 100 up to several hundred.

When $K_{in} = Y_r$ at resonance, all of the incident power will be coupled into the resonator. Since $Y_m = n^2/R_r$ at resonance, the required value of n^2 for critical coupling is

$$n^2 = Y_c R_r$$

The parameter

$$\Omega t Y_e = \frac{1}{n^2} Y_c = n Y_c = \frac{1}{n} U C \frac{1}{n} l = 2i\omega_0 C$$

so $R = 2Q/\Omega Y_c$, as may be seen by using (7.30) for Q . Thus

for critical coupling. If $Q = 157$, then $B_{-2} = y_c/10$. For a high- Q resonator, only a small coupling capacitance is required to obtain critical coupling.

The loading of the resonator by the external transmission line is equivalent to a resistance $n^2 Z_r$ connected in series with R_c . Thus the

external Q_e is given by

$$\begin{array}{ccc} a > 0L_e & OQCX & QY_C G \\ n^* Z_c & \mathbf{Bf} & \mathbf{Bt} \end{array}$$

The coupling parameter K is given by the ratio Q/Q_e and [18]

$$K = \frac{Bl}{Y_G} \quad \frac{2QBI}{2QBI} \quad (7.326)$$

as found by using $R \equiv 2Q/TW_{\text{c}}$ given above [W.31]. Since $Y_{\text{JT}}/2Q \gg iR/2Q = G$ our earlier assumption that the numerator in (7.24) could be approximated by Bf is justified. From Fig. 7 it can be found that for the typical values of C_2 that are required C and C are negligible. A parallel resonant circuit coupled by a small series connect capacitor functions as a series resonant circuit.

For a transmission line without dispersion, that is, p is a linear function of w , the Q of the transmission line is given by $p/2a$. If the propagation constant p is not a linear function of a , the transmission line is said to be dispersive. For a microstrip line the effective dielectric constant ϵ_e is a function of o ; so $j = |J|e_e(o)$ is not a linear function of w . In the frequency range where e_e is changing quite rapidly with frequency, a different expression from that given above must be used for the Q . The power flow along the transmission line is given by Wv_g , where W is the average stored energy per unit length and v_g is the group velocity which is given by $(dp/dv)_{\perp}$. Thus $W = P/u_g = P(dp/d\omega)$ and the Q is given by

$$\frac{IW}{2aP} \quad \frac{coP(dp/dw)}{2a} \quad \frac{w(dp/d\omega)}{2a}$$

When there is no dispersion, $\beta = w/v_p$, where v_p is a frequency-independent phase velocity. For this case the general expression for Q reduces to $P/2a$. When there is significant dispersion, the capacitance C , in the equivalent circuit of the resonator is also a function of ω because e_e is.

Circular Disk Resonator

The circular disk or planar radial resonator is shown in Fig. 7.12. An approximate analysis of this resonator is readily carried out by taking the outer boundary at $r = a$ as a perfect open-circuit boundary ^{independently}, upon which $\mathbf{n}^* \times \mathbf{H} = 0$. The field in the resonator will not be zero ^{and} will have an axial electric field E_r . Thus the modes in the ^{resonator} with d like those of TM modes in a circular waveguide at cutoff ^{subject} to the boundary being a magnetic wall instead of an electric wall. The expression for the axial electric field is

$$E_z = CJ_{nn}(kr)\cos n\phi$$

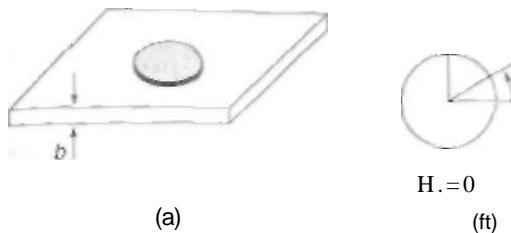


FIGURE 7.12
The circular disk resonator.

where C is an amplitude constant, J_{nr} is the Bessel function of order re , and $k = JTrk/\epsilon_r$ with ϵ_r being the dielectric constant of the substrate material. Since this is the only electric field component present, the magnetic field is given by

$$\nabla \times a_z E_x = -a_z, \quad \nabla \times V E_x = -j \omega \mu_0 H_z = -jk^2 YL$$

or

$$H_z = -\frac{J_{nr}}{\epsilon_r} \nabla \times V E_x = \frac{\omega \mu_0 Z_0}{r} \left[i r + a_z \frac{a_{nr}}{r} \right] \hat{y} \quad (7.34)$$

$$= -\frac{\omega^2 - k^2 - a_z^2}{\epsilon_r} \hat{y}$$

where $J'_{nr}(kr) = dJ_{nr}(kr)/d(kr)$. In order for 72° to vanish at the boundary $r = a$, we require $J'_{nr}(ka) = 0$. The roots of this equation are given in Table 3.7. The smallest root is 1.841 corresponding to the use of the junction $J_x(kr)$. The resonant frequency is obtained by equating $ka = \sqrt{\epsilon_r} k_0 a$ to 1.841; thus

$$\omega_{R0} = \frac{1.841c}{\sqrt{\epsilon_r} r_0} \quad (7.35a)$$

where $c = 3 \times 10^{10}$ cm/s is the speed of light.^t Since $n = 1$, 1.841 is the first root, and there is no dependence on z , this dominant mode is designated as the TM_{00} mode. As an alternative we can solve for the radius a which is given by

$$a = \frac{1.841}{\sqrt{\epsilon_r}} \quad (7.35b)$$

For a resonator at a frequency of 4 GHz and using alumina with $\epsilon_r = 9.7$ as

^tMore accurate expressions for the resonant frequencies are given in I. Wolff and N. Knoppik, Rectangular and Circular Microstrip Disk Capacitors and Resonators, *IEEE Trans.*, vol. MTT-22, pp. 857-864, 1974.

the substrate material], we require $a = 0.7056$ cm which relatively compact resonator.

The Q of the resonator is readily evaluated. The stored $\frac{oU_{\text{electr}}}{c e_{\text{ner}}}$ is given by

$$\begin{aligned} W_e - |Cf| &= -JlJJjJtW \cos^2 r dr dA dz \\ &= |C|^2 \frac{7\pi^2 \epsilon_r b^6}{k^2 a^2} J_l(ka) + 11 - \frac{1}{k^2 a^2 l} J_l(ka) \end{aligned} \quad (7.36)$$

since $J_l(ka) = 0$. In this expression ϵ_r is the *real part* of the dielectric constant and b is the substrate thickness. The stored magnetic energy W equals V^2 , at resonance.

The power loss in the dielectric is given by

$$P_{n<=T} / \langle E_f r dr dA \rangle = \frac{2\pi e}{T} W_t \quad (7.37)$$

The power loss in the conducting disk and ground plane is given by

$$\begin{aligned} &2 \cdot \frac{\epsilon_0}{\epsilon_r} \cdot \frac{d}{dr} \\ &= \frac{R}{a} \frac{j^2}{\omega} \int_0^a (\langle H \rangle + \langle H \rangle) r dr d\theta \end{aligned}$$

since the current density is given by $a, X H$ on the ground plane and $j/a, X H$ on the disk. After substituting from (7.34) for the magnetic field we obtain

$$P_c = -i \epsilon_0 r j g r k^2 f(kr) + \sim J f(kr) r dr$$

after integrating over 4π which gave a factor of IT . We can integrate term by parts to obtain

$$\begin{aligned} \int_0^a \frac{j \partial J_x(kr)}{dr} \frac{dJ_x(kr)}{dr} r dr &= \int_0^a r J_x(kr) \frac{dJ_x(kr)}{dr} dr \\ &\quad - \int_0^a J_x(kr) \frac{dJ_x(kr)}{dr} dr \end{aligned}$$

take note oft**

Since $J'_x(ka) = 0$ the integrated term is zero. We now have differential equation

$$\frac{d}{dr} r \frac{dJ_x}{dr} + \frac{I}{k^2 r} J_x = 0$$

satisfied by $J_x(kr)$. By using this equation to express the second derivative of J_i in terms of $J_{,,}$, the expression for P_{lr} reduces to

$$,R,,ir$$

$$-**-\mathbf{SM}^*\mathbf{V-W}^{\text{TM}}*^f$$

The Q is given by

$$\frac{2a>W_f}{Q} \quad \frac{k_0 b}{p_{ld} + p_h} \quad 2i?,,,y_0 + 6';fe_0 6/c; \quad (7.38)$$

which is a surprisingly simple result in that there is no dependence on the Bessei function or the radius of the resonator. In (7.38) the substrate thickness is given by the parameter 6.

As an example we will consider a disk resonator operating at 4 GHz. The substrate will be chosen as alumina with $e'_r = 9.7$ and $e'' = 0.0002$ and having a thickness $6 = 1$ mm. For this resonator $k_0 b = (2\pi/75) = 0.08377$:

$$Q = \frac{2 X 16.44 \times 10^{-3}/120}{2 X 8.377 X 10^{-3}} + 2 X 8.377 X 10^{-3} \approx 942$$

The dominant loss mechanism is conductor loss. The Q of this resonator is higher than that of a transmission-line type since the conductor losses are smaller. In practice, the Q can be expected to be less because of radiation loss and extra loss from surface roughness on the disk. There will also be a much larger azimuthal current density $J_{,,}$ at the edge than is obtained using the approximate theory given above. For an infinitely thin disk, the current eL should exhibit the usual edge singularity and be proportional to $(a^2 - r^2)^{-1/2}$ as the edge is approached. There will also be some current flowing on the top surface of the disk and this will also increase the conductor losses.

The actual resonant frequency of the disk resonator is lower than that predicted by the simple approximate theory used above. More accurate formulas have been obtained that will give the resonant frequency to within about 1 percent. The reader is referred to the literature for a more comprehensive treatment.!

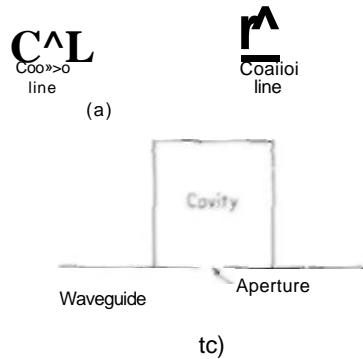


FIGURE 7.13

Cavity-coupling methods, (a) Loop coupling. . . , probe coupling; (c) aperture coupling.

In addition to the circular disk, other shapes such as ellipses, rines triangles, and squares, have been considered for use as resonators, as well as coupled sections of microstrip lines.t

7.4 MICROWAVE CAVITIES

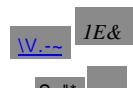
At frequencies above 1,000 MHz, transmission-line resonators have relatively low values of Q, and so it becomes preferable to use metallic enclosures, or cavities, instead. A cavity can be considered as a volume enclosed by a conducting surface and within which an electromagnetic field can be excited. The electric and magnetic energies are stored in the volume of cavity. The finite conducting walls give rise to power loss and thus equivalent to some effective resistance. The fields in the cavity may be excited, or coupled to an external circuit, by means of small coaxial probes or loops. Alternatively, the cavity may be coupled to a wave by means of a small aperture in a common wall. These coupling methods are illustrated in Fig. 7.13. Before considering the effects of coupling on the performance, the field solutions in rectangular and cylindrical cavities will be presented.

Rectangular Cavity

Figure 7.14 illustrates a rectangular cavity of height b , width a , and wavelength λ . It may be considered to be a section of rectangular waveguide in a short circuit at $z = d$. If d equals a multiple of a half wavelength, the field solutions in rectangular and cylindrical cavities will be presented.

tj. lielszajn and D. S. James, Planar Triangular Resonators with *
Trans., vol. MTT-26, pp. 95-100. 1978.

I. Bahl and P. Bhartia. "Microwave Solid State Circuit Design," John Wiley & Sons, New York. 1988.



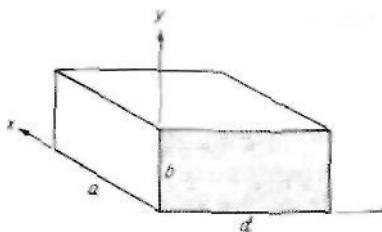


FIGURE 7.14
A rectangular cavity.

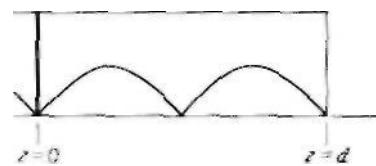


FIGURE 7.15
Standing-wave pattern in a short-circuited waveguide.

at the frequency f , the resultant standing-wave pattern is such that the x and y components of electric field are zero at $z = 0$. Consequently, a short circuit can be placed at $z = 0$ as well, as in Fig. 7.15. The resultant structure is a rectangular cavity. This description of a cavity also shows that the field solution may be obtained directly from the corresponding waveguide solutions. For the nm th TE or TM mode, the propagation constant is given by

$$\star \bullet - \star \ll - (\mathbf{T}) - \mathbf{I} \mathbf{T}) \quad <7.39a>$$

where $k_0 = 2\pi f_0/c$. We require $p_{nm} d = lir$, where I is an integer in order for the cavity to be a multiple of a half guide wavelength long. Thus, when d is specified, ω_{nm} is given by

$$\omega_{nm} / = 1, 2, \dots \quad (7.396)$$

However, this relation is consistent with the earlier one only for certain discrete values of k_{nl} . Only if $k_{nl} = k_{nm}$, where k_{nm} is given by

$$k_{nl} = \frac{\pi}{a} \sqrt{\left(\frac{n}{a}\right)^2 + \left(\frac{m}{b}\right)^2} \quad (7.40)$$

will (7.39a) and (7.39b) be satisfied. These particular values of k_0 give the resonant frequencies of the cavity; i.e.,

$$\frac{ck_{nl}}{2\pi} = c \frac{l}{2d} \quad \frac{n}{a} \quad \frac{m}{b} \quad (7.41)$$

where c is the velocity of light. Note that there is a triply infinite number of resonant frequencies corresponding to different field distributions. Also note that there is more than one field solution for a given resonant frequency since (7.41) holds for both TE and TM modes. In addition, because of a lack of a preferential coordinate, in the case of a rectangular cavity, field solutions corresponding to TE and TM modes with respect to the x and y

axes could also be constructed, and these would have the same frequency. However, these latter modes are just a linear combination of TE and a TM mode with respect to the z axis and therefore do not form a new solution.

To illustrate the method of solution for the fields in a rectangular cavity and the evaluation of the unloaded Q , the TE_{101} mode is detailed. If $b < a < d$, this will be the mode with the lowest resonance frequency and corresponds to the TE_{101} mode in a rectangular waveguide. Mode subscripts nml pertain to the number of half-sinusoid variations along the x , y , and z axes, respectively. From (3.206), the field solution for a TE_{101} mode is

$$B_z = (A + e^{-\mu_0 z}) + A' V^{-1} J_0 \cos \frac{\pi x}{a}$$

$$H_x = \frac{1}{\pi} (A' e^{jPz} - A e^{-jPz}) \sin \frac{\pi x}{a}$$

$$E_y = -j \frac{k(Z, a)}{\pi} (A e^{-jPz} + A' e^{jPz}) \sin \frac{\pi x}{a}$$

where A' and A'' are amplitude constants for the modes propagating in the $+z$ and $-z$ directions, respectively. To make E_y vanish at $z = 0, d$, we must choose $A'' = -A'$ so that

$$A + e^{-rt} + A' e^{jPz} = -2jA^H \sin p_{101} z$$

and also choose $p_{101} = ir/d$. The corresponding value of k_0 is thus

$$K_0 = \sqrt{1 + \frac{r^2}{d^2}} \quad (7.42)$$

and this determines the resonant frequency. The solution for the field now be expressed as follows:

$$\mathbf{E} = \frac{-2A^H k_0 Z_{101}}{\pi} \sin \frac{\pi x}{a} \sin \frac{\pi z}{d} \quad (43a)$$

$$H_x = \frac{2jA^H}{d} \sin \frac{\pi x}{a} \cos \frac{\pi z}{d} \quad (7.43c)$$

$$H_z = -2jA^H \cos \frac{\pi x}{a} \sin \frac{\pi z}{d}$$

Note that the magnetic field is $\pm 90^\circ$ out of phase relative to the electric field. This is always the case in a lossless cavity and corresponds to the current being $\pm 90^\circ$ out of phase with each other within the circuit.

At resonance the time-average electric and magnetic energy stored in the cavity are equal. The average stored electric energy is given by

$$W_e = \frac{\text{En}}{4} \int_Q \int_0^{\infty} \int_0^{2\pi} E^2 E^* dx dy dz$$

$$= \frac{6f}{4 \cdot 2 \cdot a^3 M A \cdot Z_{01} \cdot \alpha_0 \cdot \mu_0} \cdot AT \quad (7.44)$$

The reader may readily verify that

$$W_m = \frac{1}{4} \int_0^{\infty} \int_0^{\pi/2} \int_0^{2\pi} j^d (H_x H_z + H_y H_r) dx dy dz = W_e \quad (7.45)$$

In order to determine the cavity Q, the losses caused by the finite conductivity of the cavity walls must be evaluated. For small losses the surface currents are essentially those associated with the loss-free field solutions (7.43). Thus the surface current is given by

$$\mathbf{J}_s = \mathbf{n} \times \mathbf{H}$$

where \mathbf{n} is a unit normal to the surface and directed into the cavity. Hence the power loss in the walls is given by

$$P_l = \int_{\text{walls}} \mathbf{J}_s \cdot \mathbf{J}_s^* dS = \int_{\text{walls}} \mathbf{H}_{\perp}^2 dS < 7.46 >$$

where $R_m = I/\&S_S$ is the resistive part of the surface impedance exhibited by the conducting wall having a conductivity a and for which the skin depth is $\delta_S = (2\pi f \mu_0 \sigma)^{1/2}$. In (7.46) H_{\perp} is the tangential magnetic field at the surface of the cavity walls. Substituting from (7.43) into (7.46), a straightforward calculation gives

$$P_l = \frac{2a^3 b + 2d^3 b + ad^2 + da^2}{A \cdot TR_m} \quad (7.47)$$

With the use of (7.3), we find that the Q is given by

$$Q = \frac{wW}{P_r} = \frac{2a>W_e}{P_r} = \frac{<\mathbf{J}_s \mathbf{J}_s^* dS>}{2\pi R_{\perp}^2 (2a^3 b + 2d^3 b + a^2 d + d^2 a)}$$

$$= \frac{(ft_{101} ac \Omega^2 Z_n)}{2i\pi^2 R_{\perp}^2 (2a^3 b + 2d^3 b + a^2 d + d^2 a)} \quad (7.48)$$

upon replacing $wZ_D \epsilon_0 = w/\alpha_0 \mu_0$ by k_m .

As a typical example, consider a copper cavity ($a = 5.8 \times 10^{-3}$ S/m) with $c = 6 = rf = 3$ cm. The resonant frequency is found to be 7,070 MHz. The surface resistance $R_{\perp} = 0.022$ ohms, and the Q comes out equal to 12,700. The damping factor $\gamma = \omega/2Q$ equals 1.74×10^6 Np/s, or about 2.5×10^{-4} Np/cycle of oscillation. Because of the high value of Q , 4,000

cycles of free oscillation can occur before the amplitude has $\frac{1}{\sqrt{e}}$ factor.

If the cavity is filled with a lossy dielectric material with $\epsilon = \epsilon' - j\epsilon''$, the time-average electric energy stored in the cavity is given by

$$W_e = \frac{\epsilon}{4} \int_V |E|^2 dV \quad (7.49c)$$

The lossy dielectric has an effective conductivity $\omega\epsilon''$, and hence the loss in the dielectric is

$$P_d = U J \cdot E \wedge y = \frac{1}{V} \int_V |E|^2 dV \quad (7.49d)$$

If Q_d is the Q when a lossy dielectric is present but the walls are perfectly conducting, then

$$Q_d = \frac{2\pi W}{I_d} \quad (7.50)$$

When wall losses are also present, the net Q is Q' and given by

where Q is the quality factor when lossy walls are present and $\epsilon'' = 0$. Q is given by (7.48), with ϵ_0 replaced by ϵ' . Also note that, for a cavity filled with dielectric, the resonant frequency is given by

$$\frac{f_{res}}{\epsilon'} = \frac{1}{2\pi} \sqrt{\frac{\epsilon_0}{\epsilon'}} \frac{c}{2\pi} \frac{a}{n} \quad (7.52)$$

Cylindrical Cavity

The cylindrical cavity is a section of circular waveguide of length l , radius a , with short circuiting plates at each end, as in Fig. 7.16. It is very commonly used in practice for wavemeters to frequency, because of the high Q and wide frequency range of operation.

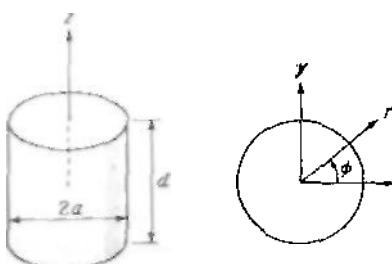


FIGURE 7.16
Cylindrical cavity.

provides. A high Q is necessary in a frequency meter in order to obtain a high degree of resolution or accuracy in the measurement of an unknown frequency. When the cavity is tuned to the frequency of the unknown source, it absorbs a maximum of power from the input line. A crystal detector coupled to the input line can be used to indicate this dip in power at resonance.

The fields in the cylindrical cavity may be determined from the corresponding waveguide solutions. The lowest resonant mode is the TE_{U1} mode corresponding to the dominant TE₁₁ mode in the circular guide. This mode is examined in detail below. Using the field solutions tabulated in Table 3.6 and combining a forward and backward propagating TE_U mode, we have

$$H_r = J_0 \frac{Pn^r}{a} \cos \angle b(A^+ e^{-Jli})^* + A^- e^{Jli\kappa} \quad (7.53a)$$

$$H_t = \frac{P'u}{P'n} J_0 \frac{iP'i^r}{a} \cos \angle (A^* e^{-JPnm} - A' e^{Jnm\kappa}) \quad (7.53b)$$

$$H_z = \frac{JPna^2}{(Pn)r} \frac{P'n^r}{a} \sin \angle (A^* e^{-Jnm\kappa} - A' e^{-Jnm\kappa}) \quad (7.53c)$$

$$E_r = \frac{A \cdot Z_0 O}{(P'u)} \frac{1}{J_0} \left(\frac{1}{r} \right) \sin d \angle (A^* e^{-Jnm\kappa} + A' e^{Jnm\kappa}) \quad (7.53d)$$

$$E^* = \frac{E_r}{Pn} \frac{v}{a} I \quad (7.53e)$$

$$\mathbf{f} = 0 \quad (7.53/-)$$

where $P'_u = 1.841$. To make E_r and E^* vanish at $z = 0$, d , we must choose $A^- = -A^*$. The factor $A^* e^{-Jnm\kappa} + A' e^{Jnm\kappa}$ becomes $-2jA' \sin f_m z$. /?, must be chosen equal to $-rr/d$ to make $\sin f_m z$ vanish. The resonant frequency is determined from the relation

$$*_{\text{o}} = *_{\text{L}} \frac{V}{a}^{1/2} \quad (7.54)$$

To find the Q of the TE_{U1} mode, the time-average energy W stored in the cavity and the power loss in the walls must be calculated. The general expressions are

$$IT - MIT - \$ f r f (\langle E_f + l E_f \rangle r d) d \rangle dr dz$$

$$P_w = \frac{1}{2} \int_{\text{walls}} |m_{tan}|^2 ds$$

These integrals may be evaluated by substituting the expressions given

earlier for the fields. The final result obtained for the Q is

$$Q = \frac{\left[1 - \left(\frac{n}{p'_{nm}}\right)^2\right] \left[\left(p'_{nm}\right)^2 + \left(\frac{l\pi a}{d}\right)^2\right]^{3/2}}{2\pi \left[\left(p'_{nm}\right)^2 + \frac{2a}{d} \left(\frac{l\pi a}{d}\right)^2 + \left(1 - \frac{2a}{d}\right) \left(\frac{nl\pi a}{p'_{nm} d}\right)^2\right]}$$

where $I = n = m = 1$ for the TE_{JH} mode. For the $TE_{m/m}$ mode, also given by (7.55), with the appropriate values of n , m , l and p'_{nm} inserted. Note that all terms on the right are independent of frequency hence the Q varies as $1/S$, for any given cavity and thus decreases*⁰

An analysis similar to the above may be carried out for the other $TE_{m/m}$ and $TM_{m/m}$ modes to obtain expressions for the fields. For the $TE_{m/m}$ modes it is necessary only to replace $\cos \phi$ and $\sin \phi$ by $\cos mj$, and $\sin nj$, J , by J_m , p'_n by p'_{nm} and β_n by $l\pi a/d$ in (7.53). Of particular interest is the TE_{011} mode for wavemeters because its Q is two to three times that of the TE_{11} mode. Another advantage of the TE_{011} mode is that $H_{it} = 0$, and hence there are no axial currents. This means that the end plate of the cavity can be free to move to adjust the cavity length d for tuning purposes without introducing any significant loss since no currents flow across the gap; i.e., the gap between the circular end plate and the cylinder wall is parallel to the current flow lines. However, the TE_{011} mode is not the dominant mode; so care must be exercised to choose a coupling scheme that does not excite the other possible modes that could resonate within the frequency tuning range of the cavity.

To determine what modes can resonate for a given value of $2a/d$ at a frequency, it is convenient to construct a mode chart. For any given mcx we have

$$\begin{aligned} k_{nmI} &= \frac{IT}{21/2} \\ \text{or } (2a)_{m,m}^2 &= cx_{m,m} + \frac{21/2}{d} \end{aligned} \quad (7.56)$$

where $x_{lm} = p_{lm}$ for TE modes and p_{lm} for TM modes. Figure 7.17 gives a plot of $(2a)_{m,m}^2$ against $(2a/d)^2$ for several modes and constant IT and c . Examination of this chart shows over what range of $2a/d$ only a single mode can resonate (in the case of degeneracy, two modes resonate at the same frequency). For example, for $IT = 10^8$ and $c = 3$, only the TE_{011} and TM_{011} modes can resonate in the dashed rectangle corresponding to $(2a)^2$ between 16.3×10^8 and 20.4×10^8 and d , the dashed rectangle in Fig. 7.17. If the TM_{011} mode is

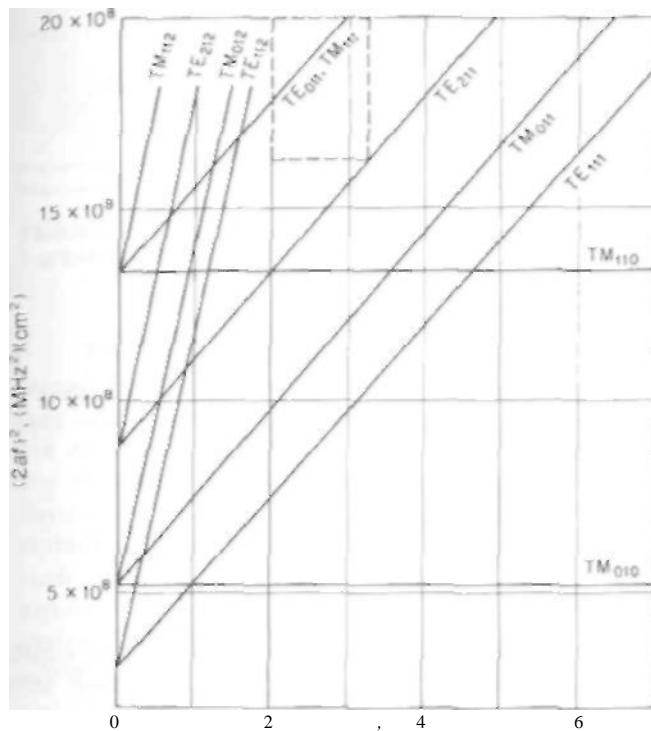


FIGURE 7.17
Mode chart for a circular cylindrical cavity.

frequency range at least can be tuned without spurious resonances from other modes occurring.

For the TM modes the Q can be evaluated to give

$$\frac{f}{\omega} = \frac{2 - (1 + 2a/d)}{2TT(1 + a/d)} \quad (7.57)$$

Figure 7.18 gives a plot of $Q8/\omega_0$ against $2a/d$ for several modes. Note the considerably higher value of Q obtained for the TE_{011} mode relative to that for the TE_{in} mode. Optimum Q occurs for $d = 2a$. At $A_0 = 3 \text{ cm}$, $S_s/A_0 = 2.2 \times 10^{-5}$, and hence, from Fig. 7.18, it is apparent that typical values of Q range from 10,000 to 40,000 or more. At $A_0 = 10 \text{ cm}$, the corresponding values of Q would be $\sqrt{10}/3$ greater.

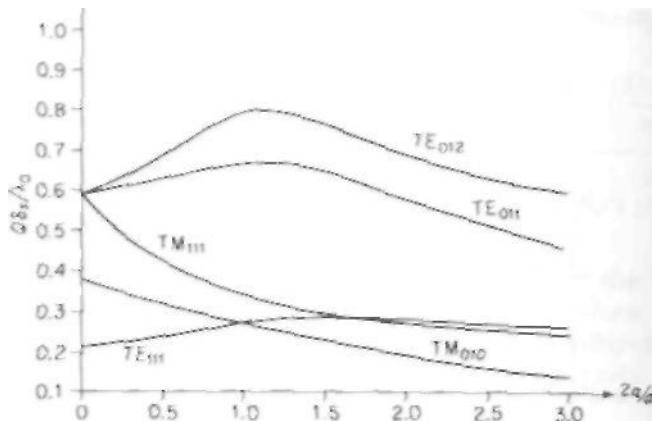


FIGURE 7.18
 Q for circular-cylindrical-cavities modes.

7.5 DIELECTRIC RESONATORS

A dielectric resonator consists of a high dielectric constant cylinder, sphere, or parallelopiped as shown in Fig. 7.19. These structures will support resonant modes similar to those in metallic cavities but with the field extending beyond the boundaries into the surrounding air region. The linear dimension of the resonator is of order $A_0/\sqrt{e_r}$, where A_0 is the free-space wavelength and e_r is the dielectric constant. For $e_r = 100$ the resonator is approximately one-tenth the size of a metallic cavity, and for this reason it is a very attractive alternative to other types of resonators; use in microstrip circuits such as filters and oscillator circuits. Since required size of a dielectric resonator in order to resonate is very small relative to the free-space wavelength, the electromagnetic field outside the resonator is quasistatic and very little radiation takes place. Thus dielectric resonators will exhibit a relatively high Q if a low-loss high dielectric constant material is used. Typical values for the unloaded Q of a dielectric resonator range from around 100 to several hundred. If the resonator is enclosed in a shielding box so that there is no energy loss due to radiation, an unloaded Q approaching the intrinsic Q of the material can be obtained.



FIGURE 7.19
Basic geometrical shapes used in dielectric resonators. (a) Cylinder, (b) sphere*, (c) cube.

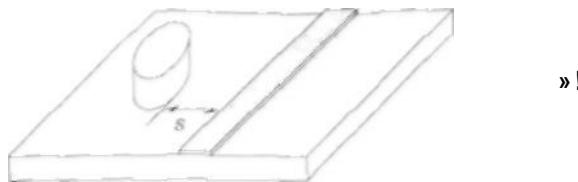


FIGURE 7.20
A cylindrical dielectric resonator coupled to a microstrip line.

The resonant frequency of a dielectric resonator depends on the dimensions of the resonator and its dielectric constant. Since these parameters change with the temperature, it is necessary to use materials that have low coefficients of expansion and relatively temperature-independent dielectric constants. In recent years a number of ceramic compounds have been developed that are suitable for use in dielectric resonators and can provide a stability of a few parts per million per degree change in temperature. One such material is barium tetratitanate which has a dielectric constant of around 40 and a loss tangent of about 0.0005. The intrinsic Q of the material is the reciprocal of the loss tangent and equals 2.000 for barium tetratitanate.

The dielectric cylinder with a height H approximately equal to the radius R and mounted on the substrate as shown in Fig. 7.20 is a popular choice for a dielectric resonator. The coupling between the resonator and a microstrip line is adjusted by adjusting the spacing S . The dominant resonant mode is a TE mode having an azimuthal electric field E_{θ} , that is independent of the angle θ . The field has a dependence on the axial coordinate z but does not change with z rapidly enough to exhibit a standing-wave pattern along z . For this reason the mode is designated as the TE_{0U} mode with the subscript 8 signifying less than one-half cyclic variation with z .

Dielectric resonators in the shape of a cylinder or a parallelopiped do not have simple analytic solutions for the various modes of oscillation. On the other hand, there are simple solutions for the fields in a spherical resonator and also for a hemisphere mounted on a ground plane. The hemisphere on a ground plane shown in Fig. 7.21 is quite similar to the dielectric cylinder, so that an understanding of the field solution for the hemispherical resonator will provide good insight into the properties of a



FIGURE 7.21
The hemispherical dielectric resonator on a ground plane.

cylindrical resonator as well. For this reason we will examine for TE modes in a hemispherical resonator.

Solutions to Maxwell's equations in spherical coordinates separated into TE and TM modes with respect to the radial coordinate r electric field for TE modes is given by

$$\mathbf{E} = \mathbf{V} \times a_{knl} \hat{\mathbf{r}} = \mathbf{M}$$

and the electric field for TM modes is given by

$$\mathbf{E} = -\mathbf{V} \times \mathbf{V}_M \mathbf{r}^* = \mathbf{N}$$

(7.5g6)

where $\psi(r, \theta, \phi)$ is a solution of the scalar Helmholtz equation in spherical coordinates, i.e.,

$$(\nabla^2 + k^2)\psi = 0$$

(7.58c)

The M and N functions have the useful property that

$$\mathbf{V} \times \mathbf{M} = kN \quad (7.59a)$$

$$\mathbf{V} \times \mathbf{N} = kM \quad (7.59b)$$

Thus, if the electric field is described by an M function, the magnetic field given by $-j\omega \mu_0 H = \mathbf{V} \times \mathbf{E}$ will be described by an N function. For circularly symmetric fields, the solutions for $E_{\theta}(r, \theta)$ are

$$E_{\theta}(r, \theta) = z_{nl}(kr) P_n(\cos \theta)$$

where $z_{nl}(kr)$ is a spherical Bessel function and $P_n(\cos \theta)$ is the Legendre polynomial. The spherical Bessel functions are related to the cylinder Bessel functions $Z_n + iJ_n(kr)$ as follows:

$$z_{nl}(kr) = \frac{1}{2kr} J_{n+1/2}(kr)$$

When the origin is included, the function

$$\frac{i}{2kr} \mathbf{H}_{nl}(kr)$$

herical¹

is used. In order to represent outward propagating waves, Hankel function

$$h_{nl}(kr) \sim \frac{i}{2kr} \mathbf{H}_{nl}(kr)$$

al**5

is used. The Bessel functions of order $n + 1/2$ and hence the $J_{n+1/2}(kr)$ functions can be expressed in terms of simple sine and cosine

example,

$$J_0(x) = \frac{\sin x}{x} \quad (7.61a)$$

$$M^* = \begin{pmatrix} \sin x & \cos x \end{pmatrix} \quad (7.616)$$

$$(7.61c)$$

$$\begin{array}{ccccccccc} 1 & * & . & . & . & . & . & . & . \\ & & & & & & & & \end{array} \quad (7.61c/)$$

$$2n + 1 \quad (7.61*)$$

The first few Legendre polynomials are

$$P_0(\cos\theta) = 1 \quad (7.62a)$$

$$P_1(\cos\theta) = \cos\theta \quad (7.62b)$$

$$P_2(\cos\theta) = [\cos^2\theta - \frac{1}{3}] = \frac{1}{2}\cos 2\theta + \frac{1}{2} \quad (7.62c)$$

For circularly symmetric modes, we can use (7.60) in (7.58) to obtain

$$M_{rr} = -k_z z_n(kr)*e_r, \quad (7.63a)$$

$$N_{rr} = \frac{s(s-f)}{r} P_{n+1}(r) \quad (7.63b) \quad (MM)$$

We will use the above solutions to construct the solution for the dominant TE₀₁₂ mode for the hemispherical resonator. The electric field has an E_{rr} component only. The ground plane is located at $\theta = \pi/2$, so that E_{rr} must vanish at $\theta = \pi/2$. The M₁ mode does not vanish at $\theta = \pi/2$ but the M₂ mode does; so the first mode that can exist for the hemispherical resonator on a ground plane is the TE₀₁₂ mode. A suitable solution for the electric field is

$$E_{rr} = -A_x/r e_r - j_2(fer)a_{rf}, \quad r < a \quad (7.64a)$$

$$E_{rr} = -A_0 k_n r \frac{dP_0}{dr} (kr) n_{rr}, \quad r > a \quad (7.64b)$$

where $k = \sqrt{k_x^2 + k_y^2}$ inside the sphere and A_x and A_0 are amplitude constants. The solution for the magnetic field is given by

$$"oA>$$

$$"oA<$$

$$"QAO$$

and hence is given by

$$\mathbf{H} = \frac{JBk}{k_0 Z_0 r} \cdot P_2 j_2(kr) \mathbf{a}_r$$

$$A_r \frac{jk}{k_0 Z_0 r} \frac{dP_2 d(krj, (kr))}{de} \mathbf{l}_e \quad r < a \quad (7.65a)$$

$$K = \frac{A_2 \sim P_2 h(k_a r) a_r}{Z^a r} + \frac{j}{Z^a r} \frac{dP_2}{dO} \frac{d(k_0 r h(k_0 r))}{d(k_0 r)} *_{\ll} \quad r > a \quad (7.656)$$

At $r = a$ the tangential fields \mathbf{E}_{\perp} , and H_e must be continuous across the boundary. Thus we must have

$$A_x k j_2(ka) = A_2 k_o h^2_2(k_o a) \quad (7.66a)$$

$$\frac{d(krj, (kr))}{d(kr)} = A_2 k_o \frac{d(k_a r h^*(k_0 r))}{d(k_0 r)} \quad (7.666)$$

All the other factors in the expressions for the fields are the same and cancel. By using (7.61e) we find that

$$j_2(kr) = \frac{k^3 r^3}{kr} - \frac{\sin kr}{k^2 r} - \frac{J_z}{k^2 r} J \cos kr \quad (7.67a)$$

$$h^*(k_0 r) = \frac{37}{k l r^3} e^{-i \pi f t} \quad (7.676)$$

With the aid of these relations, the boundary conditions can be expressed as

$$A_x k \cdot \mathbf{J} \cdot \mathbf{a}^{S \text{in} k a} \cdot \mathbf{k} \cdot V^C \cdot \mathbf{k} \cdot \mathbf{a}^{S \text{in} k a} = 0 \quad (7.68a)$$

$$A_2 k_{\perp} \frac{8j}{k^2 a^3} - \frac{A}{k_0 a} = 0 \quad (7.68a)$$

$$A_x k \frac{3}{ka} - \mathbf{a}^3 \cdot \mathbf{J} \cdot \mathbf{a}^6 \cdot \mathbf{T}^6 \cdot \mathbf{a}^6 = 0 \quad (7.68b)$$

$$- A_2 k_u \frac{6j}{k^2 a^3} + \frac{6}{K a^2} \frac{Sj}{k_0 a} - W e^{-i \pi f o^a} = 0 \quad (7.68b)$$

In order for this homogeneous set of equations to have a solution \mathbf{A}_2 , the determinant must vanish. By equating the determinant to zero we obtain the eigenvalue equation for the resonant wave number n^u .

$k = fco - v^s$ Th's equation is

$$(r^2 + jr)x - 1) \sin x - [(ij^2 - I)^* + jvx^2 - \epsilon x^3] \cos x = 0 \quad (7.69)$$

where we have let $e_r = 17^2$ and $k_0a = ka/r = x/77$. Since the roots of (7.69) are complex, the solution for x is not straightforward to carry out. An iterative method can be used to advantage. Since the dielectric constant is normally quite large, the approximate solution of (7.69) is obtained by keeping only the terms multiplied by r^2 . Thus we have

$$\sin x - x \cos x = 0$$

which has the solution $x = 4.4934 \sim x_0$. If we denote the eigenvalue equation (7.69) as $f(x)$, then a Taylor series expansion about the point x_0 gives

$$\frac{df}{dx} \left(x_0 \right)$$

We can equate df/dx to zero and then get the following improved solution

$$A \cdot t_0$$

$$df/dx,$$

$$= *o + \frac{(jvx'o - 1^*o \sim *o) \cos x_0 - (jvx_0 - 1) \sin x_0}{(tjXo + JVXZ + j-q - |*| - jc_0) \sin x_0 - (717^*o - *o) \cos x_0} \quad (7.70)$$

This second approximation can be called x_0 and used in (7.70) to obtain a third-order approximation. However, the second-order solution is usually quite accurate as the following example illustrates.

If the dielectric constant $e_r = 49$, then the second-order solution obtained using (7.70) is $ka = x = x_0 - 0.02619 4 J0.019497 \approx 4.4672 + j0.019497$. The complex resonant wave number is given by $k_0 = x/r/a$. Since we have chosen n as real, the resonant frequency is complex only because the oscillating resonator loses energy by radiation. The radiation Q is given by

$$* - iteS \quad Re ka \quad <7,71$$

and for our example equals 114.56. If the dielectric loss tangent were no greater than 0.001, the dielectric losses would reduce the Q by about 10 percent only; so the radiation loss is the major loss. By enclosing the resonator in a shielding box, a much higher unloaded Q can be obtained. However, when a shielding box is used, care must be exercised to avoid exciting cavity modes in the box.

For a dielectric constant of 49 and a resonant frequency of 10 GHz, the radius of the hemispherical resonator is $a = 4.4672/\epsilon k_0 = (4.4672/2\pi 7T7)A_0 = 0.1016A_0$ or 0.3047 cm. Thus the resonator is very compact. The parameter $k_0a = 0.638$ is quite small, so that the electromag-

netic field outside the resonator is approximately quasistatic Th A₂ can be expressed in terms of A, using (7.66a) and approxiny U^{const} by 4.4672, that is, neglecting the imaginary part. Outside the re "A * o * * electric and magnetic fields are given by ^nator the

$$\mathbf{E} \cdot \mathbf{k} = -r A_2 f_o \sin 2\theta \quad * 2r^3 \quad k^2 r^2 \quad > \sim jkgr$$

-a)

$$\mathbf{H}_r = j A_2 k_l Y_l (3 \cos 2\theta + 1) \quad k^2 r < , 3, 3 \quad /, 2, 2 \quad \bullet -JK \quad (7.726)$$

$$\mathbf{H}_{ll} = j A_2 k_l Y_l \sin 2\theta \quad , 3, 3 \quad <, 2, 2 \quad + \quad . \sim / * qr \quad (7.72c)$$

The terms in \mathbf{E}_{rr} and \mathbf{H}_H that vary like $1/k_r r$ represent the radiation field and account for the power loss due to radiation.

If a cover plate at a height c above the ground plane is installed, the radiation field will be suppressed provided $c < A_0/2$. Since \mathbf{E}_r must now vanish at $z = 0$ and $z = c$, the field outside the resonator can be expressed in terms of cylindrical waves. The lowest-order wave will have the form

$$\frac{irz}{c} H_0^1 P \quad \mathbf{17}$$

where p is the radial coordinate $T/X^2 + y^2$ and H_0^1 is the cylinder Hankel function. When p is large

$$m \quad (7) \quad 77_{pp} / kf \quad (TT/CY) \quad , \quad j p / f H_0^1 - i/c)^2 \quad -1/2$$

This is a nonpropagating wave when $k_0 < v/c$ or $c < A_0/2$. The cover will not have much influence on the field close to the resonator since terms that vary like $1/k_r r$ and higher inverse powers of $k_r r$ will be small at the location of the cover plate when c is several times larger. Thus when a cover plate is used, we can approximate the fields by dropping the radiation terms. When the radiation from the resonator is suppressed, the unloaded Q will be much larger since the losses on the $^{\text{losses}}$ $^{\text{on}} \text{fa}$ those arising from the loss in the dielectric and from the current $^{\text{nat., r.}}$

ground plane that are associated with the resonating mode of the microstrip line.

The magnetic field from the resonator will link the nearby microstrip line as illustrated in Fig. 7.22. As a result, an oscillating magnetic field excites propagating waves on the microstrip line. These waves flow of power and produce the external loading on the resonator. The resonator can be represented by a parallel resonant circuit coupled to the microstrip line by means of an ideal transformer with a P_{in} value as shown in Fig. 7.23. The resistor R must have a value

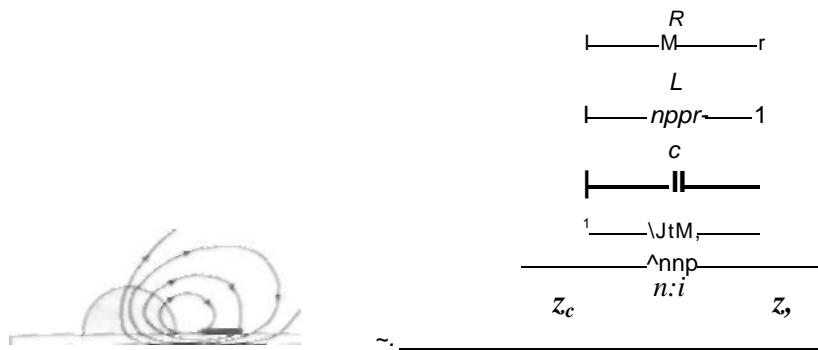


FIGURE 7.22
Magnetic coupling between a dielectric resonator and a microstrip line.

FIGURE 7.23
Equivalent circuit for a dielectric resonator coupled to a microstrip line.

$10CR$ equal to the unloaded Q , and L must have a value such that $w^2LC = 1$, at the resonant frequency. These two conditions do not establish the impedance level or absolute value of R . We can choose R arbitrarily as long as the turns ratio $n:l$ is adjusted to give the correct coupling between the resonator and the microstrip line. At resonance a series resistance n^2R is coupled into the transmission line, and this makes the input reflection coefficient equal to

$$(rc^2i? + Z_t) - Z, \quad n^2R \\ = (n^2R + Z_e) + Z_e = n^2R + 2Z_C$$

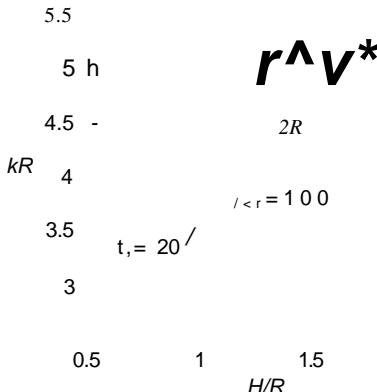
when the output line is terminated in a matched load. By measuring T the parameter

$$\frac{n^2R}{Z} = \frac{2T}{i - r}$$

is readily determined and this provides the additional information that is needed to specify the equivalent circuit. In some cases the turns ratio can also be evaluated analytically but in general it is preferable to measure it.

A number of papers have provided analysis of the cylindrical dielectric resonator. In Fig. 7.24 we show the dominant TE_{012} -mode eigenvalue kR as a function of the ratio of the resonator height H to the radius R for a cylindrical resonator mounted on a ground plane. When the dielectric constant is greater than 20, the eigenvalue kR is essentially independent of the dielectric constant. This independence was also found for the hemi-

+Y. Komatsu and Y. Murakami, Coupling Coefficient Between Microstrip Line and dielectric Resonator, *IEEE Trans.*, vol. MTT-31, pp. 34-40, 1983.

**FIGURE 7.24**

The eigenvalue kR for the TE₀₁₁ mode of a cylindrical resonator on a ground plane, JF on Fig. 3 in R. E. Collin and D. J. Ulmer, "Boundary Element Method for Dielectric Resonators and Waveguides," *Radiation Science*, Vol. 22, pp. 1155-1167, 1987.)

spherical resonator since, to a first approximation, (7.69) gave $ka = 4.4934$ independent of ϵ_r . For a given cylindrical resonator, there exists an equivalent hemispherical resonator that will have the same resonant frequency. If we let $kR = x_c$ for the dielectric resonator and let $ka = x_s$ for the hemispherical resonator, then since k is the same we must have

$$\frac{I}{R} = \frac{x_s}{x_c}$$

From Fig. 7.24 we find that a dielectric resonator with radius R equal to a must have a height H equal to $0.71a$ in order to have the same resonant frequency as a hemispherical resonator with the same radius. The inset in Fig. 7.24 shows the two equivalent resonator cross sections and clearly they are very similar.

When a dielectric resonator is mounted on the substrate material instead of on a ground plane, the resonant frequency will change by a small amount. The electromagnetic field around the resonator will also change but will remain similar to that for a resonator on a ground plane. The quasistatic field outside the resonator is similar to that of a n^{th} order quadrupole. For a cylindrical or spherical resonator in free space

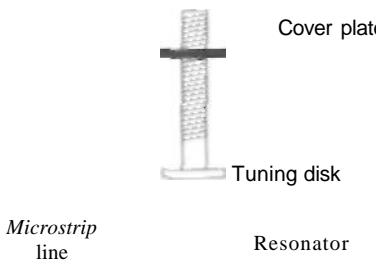


FIGURE 7.25 A method for tuning a dielectric resonator. The position of the metal disk at the end face of the resonator is varied.

dominant mode has a local field similar to that of an oscillating magnetic dipole. The image of a magnetic dipole in a ground plane, along with the original magnetic dipole, creates a magnetic quadrupole. Thus, when a resonator is mounted close to a ground plane, there is a significant change in the field around the resonator.

The *resonant frequency* of a dielectric resonator in a shielded enclosure *can be* changed by a small amount by varying the position of a small metal disk relative to the end face of the resonator as shown in Fig. 7.25. When dielectric resonators are used as part of a filter, some form of tuning for adjusting the resonant frequency is normally required.

7.6 EQUIVALENT CIRCUITS FOR CAVITIES

In this section the equivalent circuits of cavities coupled to transmission lines and waveguides are examined. A complete treatment is not given because of the complexity of the problem in general (see Sec. 7.9 for further details). Instead, a number of specific examples are examined in order to indicate the type of results that are obtained.

Aperture-Coupled Cavity

As an example of an aperture-coupled cavity, consider the rectangular cavity coupled to a rectangular guide by means of a small centered circular hole in the transverse wall at $z = 0$, as illustrated in Fig. 7.26. As indicated earlier, a small circular aperture in a transverse wall behaves as a shunt inductive susceptance with a normalized value given by (5.41) as

$$\frac{3ab}{BL =_{TR} p} epr_0 \quad (7.73)$$

where a is the guide width, b is the guide height, r_0 is the aperture radius, and $p = [kg - (Tr/aff)]^{1/2}$ is the propagation constant for the TE_{10} mode. The equivalent circuit of the aperture-coupled cavity is thus a short-circuited transmission line of length d shunted by a normalized susceptance B .

To analyze this coupled cavity, we shall assume initially that there are no losses. The modifications required when small losses are present are given later. The cavity will exhibit an infinite number of resonances, and the input impedance Z_m will have an infinite number of zeros interlaced by an infinite number of poles, this being the general behavior of a distributed-parameter one-port microwave network. If we are interested in a resonance corresponding to a high value of Z_{in} , infinite in the case of no loss, we should examine the nature of Z_{in} in the vicinity of one of its poles. This case corresponds to a parallel resonant circuit,



FIGURE 7.26

(a) Aperture-coupled
'- cavity; (WequSv^TSf
mission-line circuit.

The input impedance is given by the parallel impedance of jX_L and $-jX_L$ and is

$$Z_{in} = \frac{-X_L \tan pd}{jX_L + j \tan pd} \quad (7.74)$$

where $jX_L = (-jB_L)$. The jmtiresonances occur when the denominator vanishes, i.e., at the poles of Z_{in} , or when

$$X_h = -\tan pd = \frac{8r20d}{Zabd} \quad (7.75)$$

To solve this equation for the values of P that yield resonances, graphical methods are convenient. By plotting the two sides of (7.75) as functions of pd , the points of intersection yield the solutions for pd . When p is known, the resonant frequency may be found from the relation

$$\frac{w}{2} = f = \frac{c}{2} \sqrt{\beta^2 + \left(\frac{1}{a}\right)^2}^{1/2} \quad (7.76)$$

This graphical solution is illustrated in Fig. 7.27. Note that there are an infinite number of solutions. Normally, X_L is very small, so that the value of pd for the fundamental mode is approximately equal to π . The higher-order modes occur for $pd = P, d = (n - 1)frr$ when n , an integer, is 1

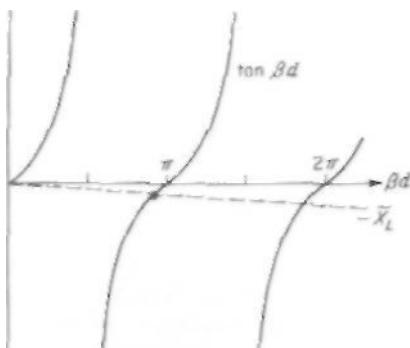


FIGURE 7.27
Graphical solution for resonant frequency of an aperture-coupled cavity.

The value of (i for the first mode will be denoted by $/3.$, and the corresponding value of w by w_t , as determined by putting $0 = 0$, in (7.76).

An infinite number of equivalent lumped-parameter networks can be used to represent Z_{in} in the vicinity of ω_0 . Usually, the simplest possible network is used. This equivalent network must be chosen so that its input impedance Z equals Z_{in} at w_0 . Likewise, for small variations $\Delta\omega$ about ω_0 , the two impedances must be equal. A general procedure for specifying this equivalence is obtained by expanding the impedance functions in a power series in $w - w_0 = \Delta\omega$ about w_0 , and equating these series term by term. Since Z_{in} has a pole at $w \sim w_0$, the Taylor expansion cannot be applied directly to Z_{in} . However, it may be applied to $(w - w_0)Z_{in}(w)$ to give

$$(w - w_0)Z_{in}(w) = \lim_{\omega \rightarrow \omega_0} (w - w_0)Z_{in}(\omega) + \frac{d}{d\omega} Z_{in}(\omega_0) (w - w_0) + \frac{1}{2!} \frac{d^2}{d\omega^2} Z_{in}(\omega_0) (w - w_0)^2 + \dots$$

We now obtain

$$Z_{in}(w) = \frac{U_{rn.}}{w_0 - w} + \frac{3 \sim i^w \sim "M^z, i}{a_{in}} + \dots \quad (7.77)$$

A similar expansion of Z gives (note that Z must have a pole at w_0 also)

$$Z(\omega) = \frac{U_{rn.}}{\omega_0 - \omega} + \frac{(w_0 - \omega)Z(w_0)}{\omega_0 - \omega} + \frac{d}{d\omega} Z(w_0) \frac{(\omega - \omega_0)^2}{2!} + \dots \quad (7.78)$$

Expansions of this type are called Laurent series expansions, and the coefficient of the $(\omega - \omega_0)^{-1}$ term is called the residue at the pole ω_0 . These two series must be made equal term for term up to the highest power in $w_0 - \omega_0 = Aw$ required to represent Z_{in} with sufficient accuracy in the frequency range of interest. For a microwave cavity, the Q is usually so high, and the frequency range dw/w , of interest is approximately the range between the two points where $|Z_m|$ equals 0.707 of its maximum. This latter fractional frequency band is equal to $1/Q$, and hence Aw/w is so small that normally only the first term in the expansion (7.77) would be required to represent Z_{in} with sufficient accuracy in the vicinity of w_0 . In the present case a simple parallel LC circuit would be sufficient to represent $Z_{in}(w)$ around ω_0 .

In order to specify the values of L and C , we must evaluate the first terms in (7.77) and (7.78). For the LC circuit we have

$$Z = \frac{j\omega L}{1 - \omega^2 LC}$$

We now choose wf $LC = 1$ to produce a pole at $u >_I$ for Z . Henc

$$ui'i - to^2$$

and $\frac{i}{hm} \frac{1}{(io - o >_x)Z} = \frac{-jao)lL}{<o + a), <a,} - Jo > lL$

Thus we have

$$Z(co) = \frac{1}{2(cu - oij)}$$

for $a >$ near W_j .

To evaluate the behavior of Z_{in} near w , we can place w equal to u in the numerator in (7.74). The denominator is first expanded in a Taylor series in p about f_i , to give

$$\begin{aligned} X_L(i) + \tan pd &= X_{L1} + \tan/3jrf \\ &\quad \frac{dX_L}{dp} \frac{d \tan pd}{dp} \boxed{iP \sim P >} \\ &\quad + \frac{ds < x^2 p_i d / (p - p_x)}{Pi} \end{aligned}$$

since $X_L(p_i) = X_{L1} = -\tan / \wedge d$ and $dX/dp = (l/p)X_L$. Next we expand p in terms of $o >$ about w , to give

$$P = P_X + \frac{d\mathbf{f}}{dw}$$

If we denote dp/dot at \ll_t by p , we see that Z_m can be expressed as

$$\frac{JX_{L1} \tan Pid}{[X_{L1} + \wedge, d(1 + \tan^2 M)] / (3' i / \wedge.) (\wedge^r \wedge)}$$

upon replacing $\sec^2 p_i d$ by $1 + \tan^2 / 3, d$. Replacing $\tan p_i d$ by $-A \wedge^r$ gives

$\ll .$

Normalizing, $X_{L1} \ll 1$, and since $/3, d \sim ir$, we can make further approximations to obtain (we shall verify that $X_{L1} \ll 1$ later)

$$y2 \quad (7.80)$$

$$\text{in } J p \backslash d(o > - \wedge)$$

Comparison with (7.79) shows that we must choose

$$\begin{aligned} & \frac{\langle o | L }{2} \quad Xf_A \\ & \quad p | d \\ & \circ r \quad L = \wedge l^{2Xf}, \end{aligned} \quad (7.81)$$

The capacitance C is determined by the condition $\langle u | L C = 1$ given earlier.

Up to this point we have neglected the losses in the cavity. For a high-Q cavity these may be accounted for simply by replacing the resonant frequency w , by a complex resonant frequency $w(1 + j/2Q)$, as indicated in Sec. 7.1. That is, the natural response of a lossy cavity is proportional to $e^{-\langle j^* f \rangle}$ where $g = w/2Q$. This is equivalent to having a complex resonant frequency $Wjd + j/2Q$. The field in the cavity is, apart from *some local fringing because of the presence of the aperture*, a TE i_{01} mode. Its Q was evaluated in Sec. 7.4, and is given by (7.48). For the lossy case we then have

$$/3,d(w - OJ, - jw - j/2Q)$$

At resonance ($\omega = \omega_0$), we now obtain a pure resistive impedance R_m given by

$$R_m Z_{in} = -\wedge j^{2XhQ} \quad (7.83)$$

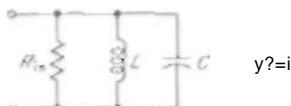
If we want the cavity to be matched to the waveguide at resonance, we must choose the aperture reactance X_u so that $R_{in} = 1$; that is,

This matched condition is referred to as critical coupling. If R_{in} is greater than the characteristic impedance of the input line (unity in the case of normalized impedances), the cavity is said to be overcoupled, whereas if R_{in} is smaller, the cavity is undercoupled. If R_m is the normalized input resistance at resonance for a parallel resonant cavity, then R_m is defined as the coupling parameter K . In the case of series resonance, the coupling parameter equals the input normalized conductance at resonance.

For the rectangular cavity discussed in Sec. 7.4, with $a = b = d = 3$ cm, we found $f_c = 7,070$ MHz and $Q = 12,700$. For this cavity

$$p | = -\wedge \gg 4.7 \times 10^{-11} \text{ s/cm}$$

and (7.84) gives $X_{IA} \sim 0.0157$ for critical coupling. The corresponding



f& ZZC

FIGURE 7.28
Equivalent circuit for an ϵ_r coupled cavity.

aperture radius r_0 from (7.73) is found to be 0.37 cm. Note that X so that our earlier approximation in neglecting X_{LI} compared with u is justified. Also note that a solution of (7.84) for the required value of ω^2 give critical coupling must, in general, be carried out simultaneously with the solution of (7.75) for the resonant frequency ω_0 . However, for a thin cavity, to, may be approximated with negligible error by the frequency corresponding to $f_{id} = -\omega_0$ in (7.84). This was done in the above calculation.⁵

For the lossy cavity the equivalent circuit must include a resistance R_{int} in parallel with L and C as illustrated in Fig. 7.28a. The reader may readily verify that the input impedance Z now becomes, for $w \ll a$

$$Z = \frac{i\omega L}{2(w - w_0 - JW/2Q)} \quad (7.85)$$

where $Q = R_{int}/u > 1$.

Since the cavity is coupled to an input waveguide, the cavity terminals are loaded by an impedance equal to the impedance seen looking toward the generator from the aperture plane. If the generator is matched to the waveguide, a normalized resistance of unity is connected across the cavity terminals, as in Fig. 7.286. The external Q_e is given by (Sec. 7.1)

$$Q_e = \frac{w_0 L}{R} \quad (7.86a)$$

and the loaded Q_L by

$$|Q|^2 = \frac{1}{(1 + R^2 L^2)^{1/2}} \quad (7.866)$$

The loaded and unloaded quality factors are related as follows:

$$\frac{1}{Q} = \frac{1}{Q_e} + \frac{1}{R} = \frac{1}{Q} + \frac{1}{R} \quad (7.87)$$

or

$$Q = (1+K)Q_L$$

In general, the coupling parameter K may be defined as

(7.88)

For a parallel resonant circuit this is seen to give

$$\frac{B_{in}}{oi_x L} \frac{Vjl'}{1}$$

which agrees with the earlier definition, ^likewise, for a series resonant circuit with normalized input conductance G_{in} , the unloaded and loaded Q's are given by $Q = IO^2 G^2$ $Q_{in} = wL$. and hence $K = G_m = Q/Q_L$ again. The coupling parameter is a measure of the degree of coupling between the cavity and the input waveguide or transmission line.

The external Q , Q_e , is sometimes called the radiation Q . The reason for this is that the cavity may be considered to radiate power through the aperture into the input waveguide. This power loss by radiation through the aperture is equal to the power lost in the normalized unit resistance connected across the *resonator* terminals in the equivalent *circuit* illustrated in Fig. 7.286.

For the rectangular cavity under discussion, the next-higher frequency at which a resonance occurs corresponds to a series-type resonance at which $|Z_m|$ is a minimum. In the loss-free case, $Z_{in} = 0$, and from (7.74) this is seen to correspond to $\tan jid = 0$, or $fid = jr$. At a series resonance, F_{in} has a pole, and consequently an analysis similar to that presented for Z_m is applied to Y_{in} . It is readily found that in the vicinity of $m = \frac{\pi}{2d}$ where $u > 2$ is the value of w that makes $j\beta d = ir$,

$$, n = \frac{1}{\sim p.A \cdot \omega^2 \cdot jco_z / 2QJ}$$

In this case the aperture has no effect, as would be expected, because the standing-wave pattern in the cavity is now such that the transverse electric field is zero at the aperture plane. The input admittance near w_2 is just that of a short-circuited guide near a half guide wavelength long. This resonance is not of any practical interest since it corresponds to a very loosely coupled cavity.

Cavity

Figure 7.29 illustrates a cavity that is coupled to a coaxial line by means of a small loop. Since the loop is very small, the current in the loop can be considered to be constant. Any mode in the cavity that has a magnetic field

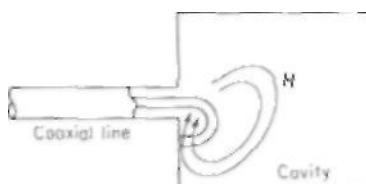


FIGURE 7.29
Loop-coupled cavity.

\mathbb{M} ,

> •

..,

 $\backslash^M_{\ll}, \dots, C_n$ \bullet^R_{\ll} FIGURE 7.30

Equivalent circuit of loop-coupled cavity.

with flux lines that thread through the loop will be coupled by the loop. However, at any particular frequency ω_0 , only that mode which is resonant at this frequency will be excited with an appreciable amplitude. The fields excited in the cavity by the current flowing in the loop can be found by solving for a vector potential arising from the current I . From the vector potential the magnetic field, and hence the flux passing through the loop, may be found. For a unit current let the magnetic flux of the n th mode that threads through the loop be ϕ_{n0} . This is then equal to the mutual inductance M_{n0} between the coupling loop and the n th mode. Each mode presents an impedance equivalent to that of a series LC circuit to the coupling loop. Thus a suitable equivalent circuit is an infinite number of series LC circuits coupled by mutual inductance to the input coaxial line, as illustrated in Fig. 7.30. The input impedance is thus of the form

$$Z_m = \frac{\omega^2 M_{n0}^2 C_{n0}}{L_0 + j \sum_{n=1}^{\infty} \frac{1}{\omega_n^2 - \omega^2} \frac{1}{L_n C_{n0}} - j \frac{h}{2L_0} \frac{1}{2i_B}} \quad (7.89)$$

where L_0 is the self-inductance of the coupling loop. If we define ω_n as resonance frequencies ω_n , by $\omega_n^2 L_n C_{n0} = 1$ and the unloaded Q of the mode by $Q_{n0} = \omega_n L_{n0} / R_{n0}$ we can rewrite (7.89) as

$$\frac{\omega^2 M_{n0}^2 C_{n0}}{L_0 + j \sum_{n=1}^{\infty} \frac{1}{\omega_n^2 - \omega^2} \frac{1}{L_n C_{n0}} - j \frac{h}{2L_0} \frac{1}{2i_B}} \quad (7.90)$$

If $\omega = \omega_{n0}$, then all terms in the series in (7.90) except the $n=0$ term are small. Thus, in the vicinity of the n th resonance,

$$Z_m = \frac{1}{L_{n0}(\omega^2 - \omega_{n0}^2 + j\omega_{n0}/Q_{n0})} \quad (7.91)$$

$$\sim \frac{j}{2L_{n0}} \frac{M_{n0}^2}{\omega_{n0}^2 - \omega^2 - j\omega_{n0}/Q_{n0}}$$

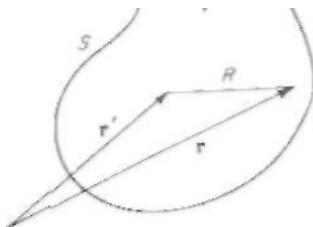


FIGURE 7.31
Illustration for Helmholtz's theorem.

The equivalent circuit now reduces to a single *LCR* series circuit mutually coupled to the input line. For efficient excitation of a given mode, the loop should be located at a point where this mode provides a maximum flux linkage.

The preceding results represent a formal solution to the loop-coupled cavity. In order actually to specify the circuit parameters, the boundary-value problem for the fields excited in a cavity by a given loop current must be solved. Also the *Q*'s of the various cavity modes must be determined. For simple cavity shapes these calculations can be carried out with reasonable accuracy. However, they are too lengthy to include here.

*7.7 FIELD EXPANSION IN A GENERAL CAVITY

Considerable insight into the general properties of an electromagnetic cavity can be obtained by examining the problem of expanding an arbitrary electromagnetic field into a complete set of modes in a cavity of unspecified shape. Such general modal expansions are required in order to determine the fields excited in a cavity by an arbitrary source. These expansions are also required in the evaluation of the input impedance or admittance of cavities coupled to external transmission lines or waveguides.

A fundamental theorem which is basic to general cavity theory is Helmholtz's theorem.^t This theorem states that in a volume *V* bounded by a closed surface *S* as in Fig. 7.31 a general vector field *P*(*x*,*y*,*z*) is given by

$$\begin{aligned} \mathbf{P}(*, y, z) = \mathbf{V} & - \frac{\nabla}{4\pi R^3} \int_V \mathbf{P}(\mathbf{r}') dV + \oint_S \frac{\mathbf{P}(\mathbf{r}') \cdot \mathbf{n}}{4\pi R^3} dS' \\ & \mathbf{V} \times \hat{h} \int_V \frac{\mathbf{V} \times \mathbf{P}(\mathbf{r}')}{4\pi R^3} dV + \oint_S \frac{\mathbf{P}(\mathbf{r}') \times \mathbf{n}}{4\pi R^3} dS' \quad (7.92) \end{aligned}$$

where $R = |\mathbf{r} - \mathbf{r}'|$. Thus the volume sources for *P* are given by $-\nabla \cdot \mathbf{P}$

^tR- '5. Collin, "Field Theory of Guided Waves," 2nd ed., Mathematical Appendix, IEEE Press, Piscataway, N. J., 1991

and $\nabla' \times P$, and the surface sources are given by $P \cdot n$ and $P v$ only if $\nabla \cdot P = 0$, $n \cdot P = 0$, can P be derived entirely from the^{*} suitable vector potential. Also, only if $n \times P = 0$, $\nabla \times P = Q$ derived from the gradient of a scalar potential, as (7.92) shows IF $\nabla \cdot P = 0$ and $\nabla \times P = 0$ in V , then P is said to be a source-free field. In this case P can always be derived from the gradient of a scalar n ^t but this potential must be multivalued if $n \times P$ does not equal zero situations. This statement will be clarified later.

In setting up a suitable set of modes in which to expand a vector P inside a given volume V , it is necessary to know the boundary conditions^{*} that must be imposed on these modes in order that a unique set of modes may be obtained. This uniqueness is needed so that when a solution for the field has been obtained we are assured of the uniqueness of that solution. Since the electromagnetic field satisfies Helmholtz's equation, we are concerned with obtaining a unique solution to this equation.

Consider $\nabla^2 A_{\text{ext}} + k^2 A_n = 0$. We wish to determine the type of boundary conditions that must be imposed on A_{ext} such that the solution to this equation is unique for a specific eigenvalue k_n . Assume that a second solution B_{ext} exists such that $\nabla^2 B_{\text{ext}} + k^2 B_{\text{ext}} = 0$ also. The difference solution $C_{\text{ext}} = A_{\text{ext}} - B_{\text{ext}}$ satisfies $\nabla^2 C_{\text{ext}} + k^2 C_{\text{ext}} = 0$ also. We shall require of B_{ext} the conditions $\nabla \cdot B_{\text{ext}} = \nabla \cdot A_{\text{ext}}$, $\nabla \times B_{\text{ext}} = \nabla \times A_{\text{ext}}$, in V , so that both A_{ext} and B_{ext} have the same volume sources. Then

$$\nabla \cdot C_{\text{ext}} = \nabla \times C_{\text{ext}} = 0$$

in V . Consider

$$\begin{aligned} \int_V f c_j^* / |C| J^a W - \int_V c_{\text{ext}} - \nabla^2 C_{\text{ext}} dV \\ = \int_V j(C_n \cdot \nabla \times V \times C_{\text{ext}} - C_{\text{ext}} \cdot W \times C_{\text{ext}}) dV \end{aligned}$$

Using the relations

$$\nabla \cdot (C_{\text{ext}} \times V \times C_{\text{ext}}) = IV \times C_{\text{ext}}|^2 - C_{\text{ext}} \cdot V \times V \times C_{\text{ext}}$$

$$\text{and } \nabla \cdot (C_{\text{ext}} \cdot V \cdot C_{\text{ext}}) = IV \cdot C_{\text{ext}}|^2 + C_{\text{ext}} \cdot V V \cdot C_{\text{ext}}$$

we obtain

$$\begin{aligned} k^2 \int_V |C_{\text{ext}}|^2 dV = & \int_V (IV \times C_{\text{ext}} \cdot dV + \int_V |V \cdot C_{\text{ext}}|^2 dV \\ & - \int_V (C_{\text{ext}} \cdot V - C_{\text{ext}} \cdot C_{\text{ext}} + C_{\text{ext}} \cdot V \times V \times C_{\text{ext}})^* dV) \end{aligned} \quad (7.93)$$

The volume integrals on the right vanish, and to make $\int_V |C_{\text{ext}}|^2 dV$ vanish which implies $C_{\text{ext}} = 0$, we must make the two surfaces boundary This may be accomplished by making A_{ext} and B_{ext} satisfy

conditions

$$\hat{n} \cdot A_{\parallel} = n \times A_{\parallel} = 0 \quad \text{on } S \quad (7.94a)$$

$$\text{or} \quad n \cdot A_{\parallel} = n \times V \times A_{\parallel} = 0 \quad \text{on } S \quad (7.946)$$

and similarly for B_{\parallel} . These are the boundary conditions that will be used for the electric- and magnetic-type modes, respectively. Other possibilities are $n \cdot A_{\parallel} = n \times A_{\parallel} = 0$ on S or $V \cdot A_{\parallel} = n \times V \times A_{\parallel} = 0$ on S . The imposition of these boundary conditions gives a unique solution. Note that two conditions must be specified for a vector function.

Cavity Field Expansions in Terms of Short-Circuit Modes

The short-circuit modes are those corresponding to the field solutions inside ideal (perfectly conducting walls), totally closed cavities. For these modes the electric field modes E_{\parallel} satisfy the boundary condition $n \times E_{\parallel} = 0$ on S , and the magnetic field modes satisfy the boundary condition $n \cdot H_{\parallel} = 0$ on S . In addition, as (7.93) and (7.94) show, we must impose the additional boundary conditions $V \cdot E_{\parallel} = 0$ and $n \times V \times H_{\parallel} = 0$ on S in order to obtain unique solutions.

Sometimes it is convenient to consider a cavity as having a surface part of which is perfectly conducting and part of which acts as a perfect open circuit. In this case the boundary conditions for E_{\parallel} and H_{\parallel} are interchanged on the perfect open-circuit portion of the boundary. In practice, it is difficult to solve the cavity problem when the same boundary conditions do not apply on the whole surface. For this reason we restrict our analysis to the use of short-circuit modes, since these are readily found for the common types of cavities encountered.

There are three basic types of cavities to be considered, as illustrated in Fig. 7.32. Type 1 is a simply connected volume with a single surface, whereas type 2 is a simply connected volume with a multiple (double) surface. Finally, type 3 is a multiply connected volume with a single surface. Examples of the latter type are toroids and short-circuited coaxial lines.

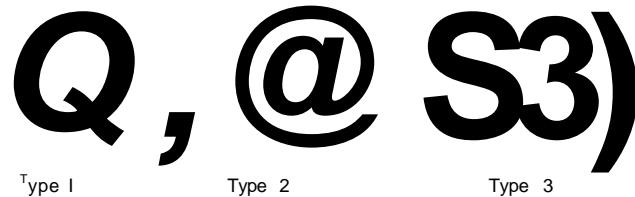


FIGURE 7.32
Basic types of cavities.

Consideration of these basic cavity types is of importance in connection with the existence of zero frequency modes, as will be seen.

Electric Field Expansion

In general, both $\nabla \times E$ and $\nabla \cdot E$ are nonzero; so we require two modes: a solenoidal set with zero divergence and nonzero curl, and an irrotational set with zero curl but nonzero divergence. The solenoidal modes are defined by

$$\begin{aligned} \nabla^2 E_{\text{in}} + \mu^* E_{\text{in}} &= 0 && \text{in } V \\ n \times E_{\text{in}} &= 0 && \text{on } S \\ V \cdot E_{\text{in}} &= 0 && \text{in } V \text{ and on } S \end{aligned} \quad (7.95)$$

and the irrotational modes by $E_{\text{irr}} = \nabla \times F_{\text{irr}}$, where

$$\begin{aligned} \nabla^2 b_{\text{irr}} + \kappa^* b_{\text{irr}} &= 0 && \text{in } V \\ \nabla \cdot b_{\text{irr}} &= \text{const on } S && \text{that is, } n \times V \times b_{\text{irr}} = 0 \text{ on } S \\ V \times F_{\text{irr}} &= 0 && \text{in } V \end{aligned} \quad (7.96)$$

where $k_{\text{irr}}, \kappa_{\text{irr}}$ are eigenvalues for the problem. The constant is chosen as zero, except for the $n = 0$ mode, which has $b_{\text{irr}} = 0$. The above definition for F_{irr} is chosen so that

$$\int_V dV = - \frac{1}{\kappa^*} \int_S \nabla \times b_{\text{irr}} \cdot dS = - \frac{j}{\kappa^*} \int_V (\nabla \cdot \nabla b_{\text{irr}} - \nabla^2 b_{\text{irr}}) dV$$

Note that only one boundary condition is imposed on the F_{irr} , since scalar functions b_{irr} are uniquely determined if either $\nabla \times b_{\text{irr}}$ or $d b_{\text{irr}}/dn$ is specified on S . The modes are assumed normalized, so that

$$\int_V f_{\text{irr}} dV - \int_V j W_m \cdot F_{\text{irr}} dV - \int_V E_{\text{irr}} \cdot E_{\text{irr}} dV = 1 \quad (7.97)$$

For $n = 0$ the eigenvalues κ_0, k_0 are assumed to be the zero eigenvalues. The corresponding eigenfunctions E_0, F_0 are the zero-frequency modes. We shall show that E_0 cannot be distinguished from F_0 ; so only the E_0 will be retained. We have $\nabla^2 E_0 = 0 = VV \cdot E_0 - V \times V \times E_0$, or $V \cdot E_0 = 0$, since $V \times E_0 = 0$ by hypothesis. Thus

$$-\int_V E_0 \cdot V \times V \times E_0 dV = 0$$

$$\int_V (V \cdot E_0 \times V \times E_0 - V \times E_0 \cdot V \times E_0) dV = 0 \quad \text{ft}$$

This gives $\int_V |V \times E_0|^2 dV = \mathbf{n} \cdot E_0 \times V \times E_0 dS = 0$ since $\mathbf{n} \times E_0 = 0$. Thus we must have $V \times E_0 = 0$, which implies $E_0 = V f$ where f is a scalar function which is constant on the surface S . This latter relation holds for F_0 as well, so that E_0 can be discarded as long as F_0 is retained. The zero-frequency mode F_0 will be denoted by

$$\begin{aligned} F_0 &= V \Delta D_0 \\ V \times F_0 &= 0 \\ n \times F_0 &= 0 \quad \text{on } S \quad \text{that is, } \Delta D_0 = \text{const on } S \\ V^2 \Delta D_0 &= 0 \quad \text{since } l_Q = 0 \end{aligned} \tag{7.98}$$

For cavity types 1 and 3, a solution for ΔD_0 other than a constant does not exist. Hence the F_0 mode is present only for the type 2 cavity, with ΔD_0 having different constant values on S_1 and S_2 . This mode is just the static electric field that may exist between two conducting bodies at different potentials.

Orthogonality Properties

Nondegenerate E_{mn} and F_{mn} modes are orthogonal among themselves and with each other. Consider

$$\int_V (E_m V^2 E_{mn} - E_{mn} V^* E_m) dV$$

$$= \int_V (E_{mn} \cdot V \times V \times E_m - E_{mn} \cdot V \times V \times E_{mn}) dV$$

$$= \int_V [V \times E_{mn} \cdot V \times E_m - V \times E_m \cdot V \times E_{mn}] dV$$

$$- V \cdot (E_{mn} \times V \times E_{mn} - E_m \times V \times E_{mn}) dV$$

$$= \int_S (\mathbf{f}(\mathbf{n} \times E_m - V \times E_{mn} - \mathbf{n} \times E_{mn} - V \times E_m) dS) = 0$$

Hence, if $k_m^2 \neq k_n^2$,

$$\int_V (E_{mn} - E_m) dV = S_{nm} \tag{7.99}$$

where the Kronecker delta $S_{nm} = 0$ if $n \neq m$ and equals unity for $n = m$.

For the modes F_m we show first of all that the Φ_n are orthogonal. We have

$$\int_V (\nabla \cdot F_m) dV = (I_f - I_l) f_m^* a$$

so that

$$\int_V F_m dV \sim 0 \quad \text{when } n \neq m$$

Next consider

$$\begin{aligned} \int_V (\nabla \cdot F_n) dV &= \int_S \nabla \cdot F_n dS = \int_S (\nabla^* \cdot V^*) dS = \int_V (V^* \cdot \nabla^* F_n + \nabla^* \cdot V^* F_n) dV \\ &= \int_V (F_n \cdot \nabla V^* - V^* \cdot \nabla F_n) dV = 0 \end{aligned}$$

Since the latter integral is zero, we have

(7.100)

for n and m not both equal to zero. If $m = 0$, $n \neq 0$, the proof of orthogonality still holds. However, the normalization for the F_0 mode will be chosen as

$$\int_V F_0 \cdot F_0 dV = \int_V |F_0|^2 dV = 1 \quad (7.102)$$

To show that the E_{nn} and F_m modes are orthogonal, consider

$$\begin{aligned} \int_V E_{nn} \cdot F_m dV &= (E_{nn} \cdot V) \cdot F_m = (V \cdot E_{nn}) \cdot F_m = (V \cdot E_{nn}) \cdot F_m \\ &= 0 \end{aligned}$$

since $V \cdot F_m = 0$. We now obtain

$$\begin{aligned} \int_V J W_m \cdot E_{nn} dV &= \int_S \nabla \cdot (J W_m) \cdot E_{nn} dS \\ &= \int_S J \cdot \nabla W_m \cdot E_{nn} dS = 0 \end{aligned} \quad (7.102)$$

For $n = m = 0$, the modes F_0 and E_{nn} are not orthogonal, $J \neq 0$. Sets identical. Since the F_m are orthogonal to the E_{nn} , it follows that E_{nn} are needed, including F_0 in general, to expand an arbitrary elec

Magnetic Field Expansion

To expand an arbitrary magnetic field (we consider the possibility of $\mathbf{n} \cdot \mathbf{H} \neq 0$ over a portion of the surface to exist), we shall set up a dual set of modes analogous to those used to expand the electric field. The solenoidal modes are defined by

$$\begin{aligned} \nabla^* \mathbf{H}_{\text{so}} + \nabla \times (\nabla \times \mathbf{H}_{\text{so}}) &= 0 \\ \nabla \cdot \mathbf{H}_{\text{so}} &= 0 \\ \mathbf{n} \cdot \mathbf{H}_{\text{so}} &= 0 \quad \text{on } S \\ \mathbf{n} \times \nabla \times \mathbf{H}_{\text{so}} &= 0 \quad \text{on } S \end{aligned} \quad (7.103)$$

and the irrotational modes are defined by

$$P_{\text{irr}} G_{\text{irr}} = \nabla \times \mathbf{G}_{\text{irr}}$$

$$\begin{aligned} \frac{\#_n}{(1/2)} &= 0 \quad \text{on } S \\ \nabla \times \mathbf{G}_{\text{irr}} &= 0 \end{aligned} \quad (7.104)$$

When $P_{\text{irr}} = 0$, we have (assuming this occurs for $n = 0$) $\nabla \times \mathbf{G} = 0$, $\nabla^* \mathbf{G} = 0$, $\mathbf{n} \cdot \mathbf{G}_{\text{irr}} = 0$. Helmholtz's theorem now states that \mathbf{G}_0 can be derived from the curl of a vector potential, say $\mathbf{G}_0 = \nabla^* \times \mathbf{A}_0$, where $\nabla \times \nabla \times \mathbf{A}_0 = 0$. Since \mathbf{G}_0 has also been assumed to be given by $G_{\text{irr}} = V_i / \#_0$, it follows that $i/\#_0$ must be multivalued. The function \mathbf{G}_u corresponding to a static magnetic field can exist in the type 3 cavity only. For example, in a short-circuited coaxial line,

$$\mathbf{H} = \frac{i a_{fi}}{2r} = \nabla \frac{Id}{2ir} = \frac{a_{\text{irr}} d \ln r}{r} = \frac{Ia_{\text{irr}}}{2r} = -\nabla \times \frac{Ia_{\text{irr}}}{2v} \ln r$$

Note that the scalar potential $i\theta/2\pi$ is multivalued. If $k_{\text{irr}} = 0$ for $n = 0$, then (7.103) gives $\nabla \times \nabla^* \times \mathbf{H}_{\text{irr}} = 0$. A volume integral of $\mathbf{H}_{(1)} \cdot \nabla \times \nabla \times \mathbf{H}_{\text{irr}}$, similar to that employed earlier in connection with \mathbf{F}_0 and \mathbf{E}_0 shows that $\nabla \times \mathbf{H}_0 = 0$. Hence \mathbf{H}_{irr} has the same properties as \mathbf{G}_{irr} ; that is, $\nabla \times \mathbf{H}_{\text{irr}} = \nabla \cdot \mathbf{H}_0 = 0$, $\mathbf{n} \cdot \mathbf{H}_0 = 0$ on S . Therefore we shall not retain the \mathbf{H}_{irr} mode, but keep the \mathbf{G}_{irr} mode for the zero-frequency mode.

Orthogonality Properties

We assume normalization according to the following:

$$\int_V (\mathbf{H}_{\text{so}} \cdot \mathbf{H}_{\text{so}}) dV = \int_V (G_{\text{so}} - G_n) dV = \int_V l dV = 1 \quad (7.105a)$$

$$\int_V (G_l - G_0) dV = \int_V |V - l|^{-2} dV = 1 \quad (7.105b)$$

By methods paralleling that used for the electric-type modes the orthogonality properties may be derived:

$$(H_n \cdot H_m) dV = (\nabla H_n \cdot G_m) dV = (G_n \cdot G_m) dV = \delta_{nm} dV = \delta_{nm}$$

for $n \neq m$. For $n = m$ we also have

(7.106,

Relationship Between E_{nn} and H_{nn} Modes

The eigenvalues for both the E_{nn}, H_{nn} modes were designated as k_n because they are, in fact, equal. Furthermore, we can show that

$$\nabla \times E_{nn} = k_n n_n \quad \nabla \times H_{nn} = f e_{nn} \quad (7.108)$$

The curl of the first relation gives

$$\nabla \times \nabla \times E_{nn} = \nabla V \cdot E_{nn} - V^2 E_{nn} = \nabla^2 V \times H_{nn} = f c^2 E_{nn} = -V^2 E_{nn}$$

by using the second relation. Similarly,

$$\nabla \times \nabla \times H_{nn} = -V^2 H_{nn} = k_n V \times E_{nn} = \nabla^2 H_{nn}$$

Hence (7.108) is consistent with the Helmholtz equation of which the E_{nn}, H_{nn} are solutions. Furthermore, the boundary condition on E_{nn} is $\mathbf{n} \cdot E_{nn} = 0$ on S . This implies $\mathbf{n} \times V \times H_{nn} = 0$ on S , which is the boundary condition that has been imposed on the H_{nn} functions. Also, we have $\mathbf{n} \cdot H_{nn} = 0$ on S , which implies $\mathbf{n} \cdot V \times E_{nn} = 0$ on S . Now

$$\begin{aligned} \mathbf{n} \cdot \nabla^2 \times E_{nn} &= \mathbf{n} \cdot V \cdot \frac{-n}{dn} \times X(E_{nn}) - \mathbf{n} \cdot \mathbf{n} \cdot E_{nn} J \\ &\quad + V \cdot \mathbf{n} \cdot I \times \mathbf{n} \cdot \mathbf{n} \cdot E_{nn} + n \frac{\partial}{\partial n} \times E_{nn} = \mathbf{n} \cdot V \times \mathbf{B}_{nn} \end{aligned}$$

since only the term $V \cdot \mathbf{n} \cdot E_{nn}$ is in the direction of \mathbf{n} . Here : t' \wedge_Q components tangent to the surface S . Since $E_{nn} = 0$ on S , $\mathbf{n} \cdot E_{nn} = 0$ on S and is consistent with the boundary condition $\mathbf{n} \cdot W_{nn} = 0$. (7.108) is the only possible relation between the E_{nn}, H_{nn} modes. A integral of

$$\mathbf{v} \cdot \mathbf{E}_{nn} \times \mathbf{r} \times \mathbf{E}_{nn} = -\mathbf{E}_{nn} \cdot \mathbf{r} \times \mathbf{r} \times \mathbf{E}_{nn} + iV \times \mathbf{E}_{nn} \mathbf{J}^8 = -\mathbf{M}_{nn}^{1*1,1}$$

shows that the normalization of H_{nn} is consistent with the norm $\|H_{nn}\|^2 = \int_S \mathbf{H}_{nn} \cdot \mathbf{H}_{nn} dS$ the E_{nn} also. It should now be apparent that the E_{nn}, F_{nn} and H_{nn} have the properties that enable them to represent the electric a

fields, respectively. An arbitrary field would require an infinite sum of these modes for its expansion.

OSCILLATIONS IN A SOURCE-FREE CAVITY

Consider a type 1 cavity with perfectly conducting walls and free of all currents and charges. We wish to determine the possible modes of oscillation. Let the fields in the cavity be expressed in terms of infinite series of the form

$$\mathbf{E} = \sum_n I_n(\omega) \mathbf{E}_{n\alpha} + \sum_n F_n(0) \mathbf{F}_{n\alpha} \quad (7.109a)$$

$$\mathbf{H} = \sum_n \mathbf{M}_n^* \mathbf{H}_{n\alpha} + \sum_n \mathbf{G}_n^* \mathbf{G}_{n\alpha} \quad (7.109b)$$

where $e_{n\alpha}$, f_n , h_n , and g_n are amplitude factors that are functions of time. Since $\mathbf{n} \times \mathbf{E} = \mathbf{n} \cdot \mathbf{H} = 0$ on S and the mode functions satisfy similar boundary conditions, the series are uniformly convergent and may be differentiated term by term. Maxwell's curl equations thus give

$$\nabla \times \mathbf{E} = \sum_n e_n V \times \mathbf{E}_{n\alpha} = \sum_n e_n k_n H_n = -u \mathbf{E} \wedge \mathbf{H}_{n\alpha} - u \mathbf{E} \wedge \mathbf{G}_{n\alpha}$$

$$\nabla \times \mathbf{H} = \sum_n b_n V \times \mathbf{H}_{n\alpha} = \sum_n h_n k_n E_n = e \mathbf{E} \wedge \mathbf{E}_{n\alpha} + e \mathbf{E} \wedge \mathbf{F}_{n\alpha}$$

If the first equation is scalar multiplied by $\mathbf{H}_{n\alpha}$ and $\mathbf{G}_{n\alpha}$ in turn and integrated over the cavity volume, we obtain

$$e_{n\alpha} k_{n\alpha} = -I_n \text{ at} \quad (7.110a)$$

$$b_n = 0 \quad (7.110b)$$

by virtue of the orthogonality properties (7.106). Similarly, when the second equation is scalar multiplied by $\mathbf{E}_{n\alpha}$ and $\mathbf{F}_{n\alpha}$ in turn and integrated over the cavity volume, we obtain

$$(7.111a)$$

$$at \quad (7.111b)$$

From (7.110a) and (7.111a) we obtain

$$+ \frac{k^2}{fie} = 0$$

The solution for e_{nn} is

and from (7.111a) the solution for h_{nn} is then

$$, \quad JK \quad .TT \quad . rr . \\ "M \quad V M \quad V M \quad <1126)$$

where $\omega_{nn} = A \cdot e$ is the resonant frequency for the n th mode

In the absence of volume sources, the F_{nn} and G_{nn} for $n \neq Q$ exist. However, the zero-frequency modes F_0 and G_0 may exist in types 2 and 3, respectively. These modes are independent of the E^{nn} modes. The n th free oscillation in the cavity is given by

$$E = c_s E_{nn} = E_{nn} e^{-\gamma t} \quad (7A13aj)$$

$$H = jy \cdot H_n e^{\gamma t} \quad (7.1m)$$

These results are valid if the material in the cavity is lossy as well, provided e and n are taken as complex quantities. In this case OJ_{nn} is complex, with the imaginary part representing a damping of the mode.

Cavity with Lossy Walls

Consider a cavity with finite conducting walls on which

$$n \times E = Z_{nn} n \times J_s = Z_{nn} H_n, \quad (7.1H)$$

where H_n is the tangential magnetic field and n is a unit outward directed normal. The surface impedance $Z_m = (1 + j)/aS$. Let the fields be expanded as follows:

$$Ee^{nn} = Z_{nn} E_n e^{\gamma t} + f_m F_n^{***} \quad <7U5a)$$

$$He^{nn} = f_m n \cdot H_n y^{nn} + E g_n G_n e^{\gamma t} \quad (7-115*)$$

where e_{nn} , f_m , h_{nn} , and g_n are amplitude constants independent of time. We have assumed a time variation e^{Jnn} so that the concept of a impedance Z_{nn} can be applied to account for the finite conductivity.

Walls

In the present case $n \times E$ and $n \cdot H$ do not vanish on the boundary conditions in which E and H are expanded satisfy the boundary conditions $n \times E_{nn} = n \times F_{nn} = 0$ on S and $n \cdot H_{nn} = n \cdot G_{nn} = 0$ on b' . If expansions for E and H will not be uniformly convergent at the termin. To Consequently, the curl of (7.115a) cannot be evaluated term by term. To overcome this difficulty, we use the divergence theorem

integration by parts) to obtain

$$\begin{aligned} \int_V (\nabla \cdot \mathbf{E}) \mathbf{H} - \mathbf{E} \cdot \nabla \mathbf{H} dV &= \int_V (\nabla \times \mathbf{E}) \cdot \mathbf{H} dV - \int_V \mathbf{E} \times \mathbf{H} \cdot \nabla dV \\ &= tpnXE \cdot \mathbf{H}_n dS \end{aligned}$$

Replacing $\nabla \times \mathbf{H}_{,,}$ by $k_n E_{,,}$ and $\nabla \times \mathbf{E}$ by $\sim ja > nH$ and using the expansions (7.115) and the orthogonal properties of the eigenfunctions now give

$$-jtofxh_{,,} - k_n e_n = \int_S \mathbf{X} \cdot \mathbf{E} \cdot \mathbf{H}_{,,} dS = Z_m(f) H_{,,} \cdot \mathbf{H}_{,,} dS \quad (7.116a)$$

Note that $\mathbf{H}_{,,} \cdot \mathbf{H}_{,,} = \mathbf{H}_{,,} \cdot \mathbf{H}$, since $\mathbf{n} \cdot \mathbf{H}_{,,} = 0$. Similarly, we find

$$\begin{aligned} (\nabla \cdot \mathbf{E}_{,,} \mathbf{X} \cdot \mathbf{H}) dV &= \int_V (\nabla \times \mathbf{E} \cdot \mathbf{H} - \nabla \times \mathbf{H} \cdot \mathbf{E}_{,,}) dV \\ &= \int_S \mathbf{n} \times \mathbf{E} \cdot \mathbf{H} dS = 0 \end{aligned}$$

Replacing $\nabla \times \mathbf{E}_{,,}$ by $k_n H_{,,}$ and $\nabla \times \mathbf{H}$ by $jweE$ and using the expansion (7.115) yield

$$juee_{,,} = k_n h_{,,} \quad (7.1166)$$

This result is the same as that obtained by taking the curl of (7.115b) term by term. However, (7.116a) cannot be obtained by taking the curl of (7.115a) term by term because of its nonuniform convergence.

From the two relations (7.116) separate expressions for e_n and $h_{,,}$ may be obtained. For $h_{,,}$ we find

$$K = \int_R^R \int_n^R \int_S Z_m(bH \cdot \mathbf{H}_{,,} dS) \quad (7.117)$$

where $k^2 = w^2 ixe$.

Let us now assume that the field is essentially that of the nth mode and that this mode is not degenerate, i.e., no other mode has the same eigenvalue $k_{,,}$. Then we have $\mathbf{H} = / *_{,,} \mathbf{H}_{,,}$, and the surface integral becomes

$$h_n Z_m j_{,,} H_{,,} \cdot \mathbf{H}_n dS = h_n Z_m j_{,,} |H_n| dS$$

since $H_{,,}$ is real. The power loss in the walls for the nth mode is

$$P = \int_2^2 \int_S ? \sim 6l H_n l^2 dS$$

and the average stored magnetic energy is

$$W_{,,} = - / \int_4 J_y H J - dV = \frac{1}{4}$$

The Q for the nth mode is

$$Q_n = \frac{2wW_h}{2P}$$

and thus

$$\text{SIH } I^2 dS = \frac{2P}{4c\omega W_m} > M_a$$

For (7.117) we now get

$$= \frac{(1 + J)R}{(1 - J)R}$$

a relation that can hold only if

$$k^* - k_a^* = Q, \quad 1$$

We thus find that, for the cavity with lossy walls, the resonant frequency differs from the no-loss resonant frequency ω_n by a factor $1 - 1/2Q$. In addition, a damping constant $\zeta = \omega_n/2Q$ is introduced. In terms of w and ζ , (7.119) gives

$$J_1 - J_1 - \{l - k\} - k \quad (12)$$

Degenerate Modes

The volume orthogonality of the G_{m1} and H_{m1} modes hold even if $p \gg 1$ as an examination of the method used in the proof will show. However, the proof of the volume orthogonality of the H_{m2} modes breaks down. In this case

$$\int_V H_{m2} H_{m2} dV$$

may or may not vanish. If the integral does not vanish, the two modes will be coupled together, since the average magnetic energy stored in the modes will contain a nonzero interaction term arising from the integral.

In addition to volume coupling between degenerate modes, it is also possible to have coupling arising from finite wall losses. The surface coupling is examined below.

If two modes $H_{,,}$ and $H_{,,,}$ are degenerate, so that $k_n = k_{,,,}$ and if in addition

$$\int_S H_{,,} \cdot H_m dS \neq 0 \quad (7.121)$$

then the power loss associated with these two modes will contain a cross-interaction term arising from the above integral. In this case the two modes are said to be coupled together by the finite surface impedance of the cavity walls. It is not possible to have just one of these modes present since the presence of one mode automatically couples the other mode. However, for most practical cavities such coupling does not exist. Nevertheless, the possibility of mode coupling must be kept in mind since, if it exists, both modes must be included in any calculation of energy stored, power loss, and Q .

In the case where k_n and k_m are not equal, the surface integral may be shown to vanish for rectangular and cylindrical cavities. Although a general proof is not available, we should anticipate that this is a general property of nondegenerate modes.

The problem of coupled degenerate modes may be circumvented by introducing new modes that are linear combinations of the old degenerate modes in such a fashion that they are uncoupled. If $H_{,,}$ and $H_{,,,}$ are degenerate coupled modes, choose new modes

$$H'_{,,} = c_1 H_{,,} + c_z H_{,,,} \quad (7.122a)$$

$$H'_{,,,} = d_1 H_{,,} + d_2 H_m \quad (7.122b)$$

with c_1 and d_1 chosen so that

$$\int_S H'_{,,} \cdot H_{,,,} dS = 0 \quad (7.123a)$$

$$\int_V H'_{,,} \cdot H'_{,,,} dV = 0 \quad (7.123b)$$

$$\int_V |H'_{,,}|^2 dV = \int_V n'_{,,} dV = 1 \quad (7.123c)$$

These new modes are uncoupled and can exist independently of each other in the lossy cavity. For these new uncoupled modes the Q may be evaluated for each mode individually, since the cross-coupling term in the expression for power loss has been made equal to zero and, similarly, the cross-coupling term in the expression for stored magnetic energy has been made equal to zero. If more than two modes are degenerate, a similar procedure may be applied to find a new set of uncoupled normalized modes. In a general

discussion we may therefore assume that all the degenerate been chosen so that they are uncoupled. [^] have

*7.9 EXCITATION OF CAVITIES

In this section we consider the application of the modal expand field in a cavity to the problem of finding the field excited by omen t^o [^] electric dipoles, which may represent a current loop or probe resntr In addition, a small aperture may be described in terms of eo ['] electric and magnetic dipoles as well. Thus the theory to be develop be sufficiently general to treat the three common methods of coudr ^{*} cavity to an external waveguide or coaxial transmission line.

Let a cavity contain infinitesimal electric and magnetic dipoles

$$\mathbf{P} \ll \mathbf{P}_0 \delta(\mathbf{r} - \mathbf{r}_0) \quad (7.124Q)$$

$$\mathbf{M} = \mathbf{M}_0 \delta(\mathbf{r} - \mathbf{r}_0) \quad (7.1246)$$

at a point whose position is defined by the vector \mathbf{r}_0 .

The three-dimensional delta function $\delta(\mathbf{r} - \mathbf{r}_0)$ symbolizes that the dipoles are localized at the point $\mathbf{r} = \mathbf{r}_0$. This delta function is defined in such a manner that, for an arbitrary vector \mathbf{A} which is continuous at \mathbf{r}_0 , we have

$$\int_V \mathbf{A}(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_0) dV = \mathbf{A}(\mathbf{r}_0)$$

when the point \mathbf{r}_0 is included in the volume V (Sec. 2.11).

By analogy with the following equations governing polarization u material bodies,

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \quad \mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$$

it is seen that Maxwell's equations become

$$\begin{aligned} \nabla \times \mathbf{E} &= -j\omega \mathbf{B} = -j\omega \mu_0 \mathbf{H} - j\omega \epsilon_0 M_0 \delta(\mathbf{r} - \mathbf{r}_0) \\ \nabla \times \mathbf{H} &= j\omega \mathbf{D} = j\omega \epsilon_0 \mathbf{E} + j\omega \mathbf{P}, \end{aligned} \quad n.X2ob$$

where a time factor $e^{j\omega t}$ has been suppressed.

We now use the general expansion (7.115) for the fields to

$$\mathbf{E} = \mathbf{L} \cdot \mathbf{e}_0 \mathbf{E}_0 + \mathbf{I} f_n V, \quad n.l26t$$

^{..e}

/7.126")

n.l26t»

To find the expansion coefficients $e_{,,}$, $h_{,,}$, we follow the derivation leading to (7.116) but note that $\nabla \times \mathbf{E}$ and $\nabla \times \mathbf{H}$ are replaced by the right-hand sides of (7.125). Thus

$$\begin{aligned}
 & \int_V ((\nabla \times \mathbf{E}) \cdot \mathbf{H} - j \frac{\partial}{\partial y} \mathbf{X} \cdot \mathbf{H}_{,,} \cdot \mathbf{E}) dV \\
 &= Z_m \int_S \mathbf{H} \cdot \mathbf{H}_{,,} dS \\
 &= \int_V [-j \omega_0 f t_0 B - j \alpha v_0 M_Q 8(r - r_0) \mathbf{J} - \mathbf{H}_{,,} \cdot \mathbf{dV} - k_{,,} \mathbf{J} \mathbf{E}_{,,} \cdot \mathbf{E} dV \\
 &\quad + j \int_S \mathbf{H}_{,,} \cdot \mathbf{H}_{,,} dS] \\
 &= -j(\omega_0 h_n - k_{,,} e_{,,} - J(OM_0 M_0 \cdot \mathbf{H}_{,,}(\mathbf{r}_0))) \quad (7.127a)
 \end{aligned}$$

and similarly,

$$k_{,,} h_n = \int_S \mathbf{G}_{,,} \cdot \mathbf{G}_{,,} dS \quad (7.1276)$$

To obtain an equation for $g_{,,}$, consider

$$\begin{aligned}
 \int_V \mathbf{f} \mathbf{V} \cdot \mathbf{E} \times \mathbf{G}_{,,} dV &= \int_V (\mathbf{f} \mathbf{G}_{,,} \cdot \mathbf{V} \times \mathbf{E} - \mathbf{E} \cdot \mathbf{V} \times \mathbf{G}_{,,}) dV \\
 &= \int_S 6n \mathbf{X} \mathbf{E} \cdot \mathbf{G}_{,,} dS = Z_{,,} \int_S (\mathbf{f} \mathbf{H} \cdot \mathbf{G}_{,,}) dS
 \end{aligned}$$

Using (7.125a) for $\nabla \times \mathbf{E}$, the expansion (7.1266), and the orthogonal properties of the modes $\mathbf{H}_{,,}$ and $\mathbf{G}_{,,}$ now gives

$$\int_S \mathbf{f} \mathbf{V} \cdot \mathbf{E} \times \mathbf{G}_{,,} dS = -Z_m \int_S \mathbf{H} \cdot \mathbf{G}_{,,} dS \quad (7.128)$$

since $\nabla \times \mathbf{G}_{,,}$ is zero.

In a similar fashion, use of the relation

$$\begin{aligned}
 \int_V \mathbf{f} \mathbf{V} \cdot \mathbf{F}_{,,} \times \mathbf{H} dV &= \int_V ((\mathbf{H} \cdot \nabla \times \mathbf{F}_{,,} - \mathbf{F}_{,,} \cdot \nabla \times \mathbf{H}) \cdot \mathbf{V}) dV \\
 &= - \int_V \mathbf{F}_{,,} \cdot \nabla \times \mathbf{H} dV = 6n \int_V \mathbf{F}_{,,} \cdot \mathbf{H} dV = 0
 \end{aligned}$$

together with (7.1256) and (7.126a) yields

$$j \omega_0 f_n = -j c_0 P_0 F_n(r_0) \quad (7.129)$$

We now have the following equations for the expansion $\sum_m \sum_n e_m g_n$:

$$\begin{aligned} jweoe_n &= k_n h_n - y' w P_0 \cdot E_{n\perp}(r) \\ h_{mn} &= \frac{1}{W^2 k l} \left\{ J_{mk} n P_0 + E_{mn} + \Delta M_{mn} H_{mn} - J^{mn} \frac{4}{\pi} n \cdot H_{mn} ds \right\} \end{aligned} \quad (7.130)$$

$$\begin{aligned} > * / . . . * * * . . . & \\ / \bullet P dU - J_{0if} i_0 M_0 \cdot G_{mn} - Z_m(pU \cdot G_{mn}) ds & \end{aligned} \quad (7.130c)$$

where $\Delta M_{mn} = w^2 M_0 \epsilon_0$ and (7.130a) has been used in order to eliminate t and obtain (7.1306).

In many practical problems dealing with cavities, the above equations may be simplified. Usually, ΔM_{mn} is very nearly equal to a particular resonant frequency w_{mn} . As (7.1306) shows, all coefficients h_{mn} , $m \neq n$, will then be small compared with h_{mn} . Thus all the coefficients e_m , $m \neq n$, will also be small compared with e_{mn} , and the field is predominantly that described by the H_{mn}, E_{mn} mode. In the surface integrals, which represent small perturbations from the loss-free solution, we may approximate H by $A_{mn} H_{mn}$, without appreciable error. In addition, in the equation for g_n , we may neglect the surface-integral term to a first approximation. We may then use the relation

$$0H_{mn} \cdot H_{mn} ds = \frac{2P_m}{Q_m R_m} \ll M_{mn}$$

derived earlier if the H_{mn} modes have been chosen so that they are uncoupled. In place of (7.130) we now obtain the following simplified equations

$$\begin{aligned} *SU \frac{1 - J}{Q_m} - f_t h_{mn} &= -jw k_n P_0 \cdot E_{mn} - k^2 M_0 H_{mn} \\ h_{mn} &= M_{mn} P_{mn} \cdot E_{mn} + k l U_m \cdot H_{mn} \end{aligned} \quad (7.131a)$$

$$k \frac{1 - k^2}{Q_m}$$

$$jw k_i i_0 M_0 \cdot H_{mn} + P_0 \cdot E_{mn} \quad (7.131*)$$

$$e_{mn} = -$$

$$k l - 4n -$$

$$(7.1)$$

$$\langle \epsilon / \rangle = -P_0 \cdot F$$

$$8\epsilon = -M_0 \cdot G_{mn}$$

Since

Equations (7.131a) and (7.131*) may be used for all h_m and ΔM_{mn} may be assumed to be equal or nearly equal to w_{mn} , the denominator

replaced by $k\% - k^2$, for $m \neq n$. For n we may factor the denominator to give

$$\begin{aligned} '_{,,} &= *_{0,} 1 + i^{-}; \quad !_{,,} + k_0 l H I^{-}; \\ &= -2k_{,,} - k_{,,} 1 - \frac{1}{2Q_{,,}} - j \end{aligned}$$

since $Q_{,,}$ is very large.

Usually, we are primarily interested in the strength of excitation of the resonant mode $E_{,,}, H_{,,}$. Its excitation coefficients are given by (7.131o) and (7.1316). The coefficients $'_{,,}$ and $g_{,,}$ describe the local field that exists around the dipole sources. This field is a quasistatic field in its configuration. The excitation of fields in cavities by volume distributions of currents may be solved by the same method outlined above. As an example, consider a cavity with a volume distribution of current $J(r_h)e^{''''''''}$. A differential element of current may be considered equivalent to an electric dipole P with a moment given by $JJOJ$. Thus, from a current element $J(r_{,,}) < 5(r - r_0)$ located at the point $r_{,,}$, the amplitudes of the n th mode are given by (7.131) divided by jto . When the current varies with the frequency $\omega_{,,}$ of the n th resonant mode, only this mode is excited with a large amplitude. The amplitude of the electric field of the n th mode due to the volume distribution of current is found by superposition, i.e., adding up the contributions from each current element. Thus we have

$$\begin{aligned} e_n &= -Jw/x_f (vJ(r_0) - E_j r^c/V_{,,}) \\ &\quad k'i - ki \quad i + \end{aligned} \tag{7.132}$$

where the integration is taken over the volume of the current distribution. A similar expression holds for the amplitude constant $'_{,,}$.

The fields excited in cavities by volume distributions of currents may also be solved in terms of a vector potential function. A specific application of the use of the vector potential is given in Chap. 9 in connection with the klystron tube; so we do not consider this method here.

CAVITY PERTURBATION THEORY

The resonant frequency of a cavity can be varied over a small range by inserting a small adjustable screw into the cavity as shown, for example, in Fig. 7.33a. A small obstacle, such as a dielectric sphere, placed in a cavity will also change the resonant frequency and the Q of the cavity if the dielectric is lossy. Small obstacles, when acted upon by electric and magnetic fields that are uniform over the volume occupied by the obstacle, can be

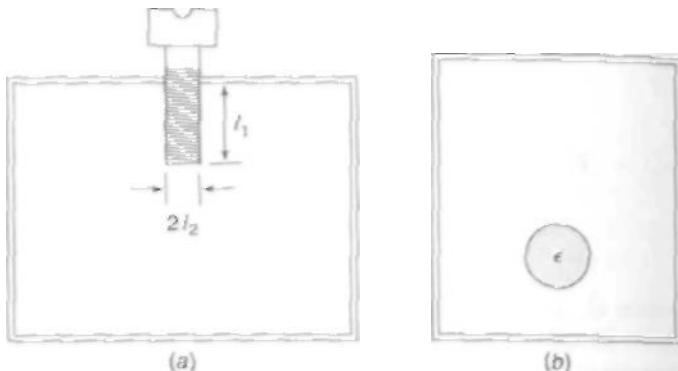


FIGURE 7.33

(a) An adjustable screw used to tune a cavity; (b) a small dielectric sphere placed in a cavity

characterized in terms of induced electric and magnetic dipole moments. For example, a dielectric sphere of radius a_e and with a dielectric constant ϵ will have a total electric dipole moment P given by

$$P = \frac{4\pi a_e^3}{3} \epsilon_0 E_0 = \alpha_e \epsilon_0 E_0 \quad (7.133)$$

where E_0 is the electric field at the location of the dielectric sphere when it is absent. The parameter α_e is called the electric polarizability of the dielectric sphere. For other obstacles we can write, in a similar way,

$$P = \alpha_e U A \quad (7.134c)$$

$$M = \alpha_m H_0 \quad (7.134d)$$

where M is the induced magnetic dipole moment and α_m is the magnetic polarizability. In Table 7.1 we list the polarizabilities for a number of obstacles that have the shape of prolate or oblate spheroids or their degenerate forms. The polarizabilities depend on whether the field acts along major axis or minor axis.

When the dominant mode in the cavity is the n th resonant mode, the fields E_0 and H_0 in (7.134) represent the fields of the n th cavity mode. The additional field radiated by the induced dipole moments have am given by (7.131). If we assume that the field in the cavity is essential the perturbed n th mode, then the dipole moments are given by

$$P = e_n [a_{eu} 6 \epsilon_0 a_u + a_{ev} \epsilon_0 E_{nu} a_D] \quad (7.135e)$$

Let $E_{nu} \bullet P = e_{nu} A$ and $H_{nu} \bullet M = n_{nu} A$, that is,

$$,7.136o$$

$$,7.1366)$$

TABLE 7.1
Polarizabilities of common obstacles

Prolate spheroid

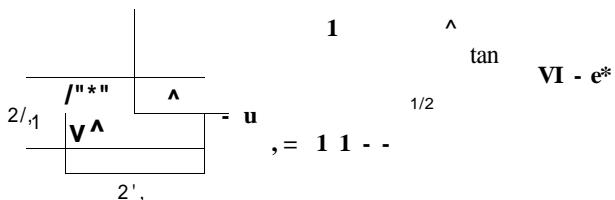
$$\Delta_p = \frac{(\epsilon_r - 1)V}{(\epsilon_r + 1)L + 1} \quad \text{for } \epsilon_r > 1$$

$$\frac{2}{2}, \quad -U \quad \frac{2(\epsilon_r - 1)V}{\epsilon_{rr} - R_f + (f_r - 1)L} \quad \frac{2(M_r - DV)}{1 + M_r - (M_r - 1)L}$$

$$Z_r = \frac{1 - e^2}{2c^3} \ln \frac{1 + e}{1 - e}$$

$$\frac{1}{t/S} \quad \mathbf{V} - \langle \mathbf{5} \rangle / \mathbf{i3}$$

Oblate spheroid



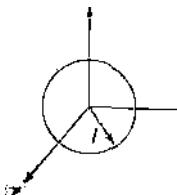
Sphere

$$\frac{4\pi r^3}{3} \epsilon_r + 2$$

2/

For conducting obstacles let $\epsilon_r \rightarrow \infty$ and set $p_r = 0$

Circular disk (metallic)



16

Equations (131a) and (1316) give

$$\begin{aligned} h_{,,} - \frac{jk_{,,} Y_0 k_n A_e e_{,,}}{k^2 - k^2 \backslash 1} &= 0 \\ \mathbf{e}_{,,} - \frac{-jk_1 Z_0 k_{,,} A_m h_n}{Q_{,,}} + \frac{k_1 \hat{\mathbf{e}}_e e_n}{Q_{,,}} &= 0 \end{aligned}$$

$$\hat{\mathbf{e}} \sim * \mathbf{S} \mathbf{1} + Q_{,,}$$

This is a homogeneous set of equations and has a solution only if the determinant vanishes. By equating the determinant to zero, we obtain an equation for \hat{k} which determines the perturbed resonant frequency of the n th mode when the obstacle is placed in the cavity. When we neglect terms in A^* , A_m^2 , and $A_e A_{,,}$, it is readily found that

$$\frac{A_{..} + A_r}{\gamma_0 - \gamma_1} = 1 \quad (7.137)$$

where $\delta = (1 - j)/Q_{,,}$. Normally A_e and $A_{,,}$ are both very small; so the approximation is a very good one in many practical cases.

As an example of the application of this formula, consider the cylindrical cavity with a *tuning screw* of length l_1 and radius I_2 as shown in Fig. 7.33a. Let the mode in the cavity be the $TM_{0,0}$ mode with electric field given by

$$E = e_{oio} \hat{o} o_{io} - e_{r} \frac{\hat{A}_0 (Poi \wedge Q)}{ifirdaJiipo}$$

where d is the length of the cavity and a is the cavity radius. We can approximate the tuning screw as one-half of a prolate spheroid. If $l_1 \gg l_2$ the induced electric dipole moment (see Table 7.1) is

when the tuning screw is located along the axis of the cavity. The $a \wedge$ occurs because only one-half of the spheroid is present. The magnetic field along the axis is zero; hence (7.137) gives

$$k_0 = k \frac{1 - \frac{1}{2} Trda^2 Ji(Po)}{assuffe}$$

when we assume that the cavity Q is infinite so that $\delta = 0$.
that $a = 2$ cm, $d = 3$ cm, $l_1 = 1$ cm, and $l_2 = 0.1$ cm. The un*

cavity resonant wave number is $k_{010} = p_0/a = 2.405/2 = 1.2025$ rad/cm. The polarizability a_{eu} is given by

$$^a\mathbf{e}\mathbf{u} = \gamma^V \mathbf{h}\mathbf{H} \quad \text{In } \frac{2\pi e^3}{\ln \frac{1+e}{1-e}}$$

where $e = (1 - f/Z^2)^{1/2}$. For our example a_{eu} equals 2.065×10^{-2} . The perturbed resonant wave number is thus given by

$$\mathbf{k}_n = \mathbf{k}_{010} \frac{1 - \frac{2.065 \times 10^{-2}}{2\pi r \times 3 \times 4 \times (0.5191)}}{} = k_{010}(1 - 0.001)$$

The perturbed resonant frequency is thus 0.1 percent lower with the tuning screw present.

By measuring the change in the resonant frequency and Q of a cavity when a lossy dielectric sphere is placed in the cavity, the complex permittivity of the dielectric material can be determined (see Prob. 7.20).

- 7.1. Show that, on a short-circuited coaxial line one-half wavelength long, the time-average stored electric and magnetic energies are equal. Use the expressions for the fields given by (3.81).
- 7.2. For the folded coaxial line, show that $b = ad$ in order for the characteristic impedance of the inner and outer lines to be the same. This is a common form of line for use in high-frequency oscillators. The effective length ℓ is about twice the physical length. At a frequency of 300 MHz, what must ℓ/l be in order that $\ell = \lambda_0/4$? If $2c = 5$ cm and $2a = 2$ cm, what must lb equal for equal characteristic impedances? For a copper line ($a = 5.8 \times 10^{-7}$ S/m), find the Q and input impedance at resonance. For d fixed, what are the optimum values of a and b that will make Q a maximum?

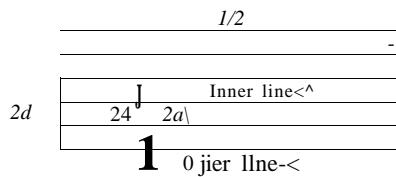


FIGURE P7.2

- 7.3. Verify that an open-circuited transmission line behaves as a series resonant circuit in the vicinity of the frequency for which it is a quarter wavelength long. Obtain an expression for the input impedance at resonance.

- 7.4. A short-circuited two-wire line is made of copper. The conduit 1 cm, the spacing is 3 cm. and the length is 40 cm. Find the "Heter frequency, the Q , and the input resistance at resonance.
- 7.5. Design a microstrip-line resonator like that shown in Fig. 7 o crostrip input line and resonator are both of width $W = 1 \text{ mm}$. The thickness is 1 mm and made from alumina with $\epsilon_r = 9.7$ and ϵ_f ^{subsl}rate equal to 2×10^{-9} . The microstrip is made from copper 0.01 mm's ^{tang} ^{Field} the required capacitance C , for critical coupling and the gap spac ^{^A} ^{Field} is the length L and Q for the resonator? The frequency of operation ^{*d} ^{What} How much shorter than one-half wavelength do you have to make $n!?$ ^{**Z} of capacitive loading at the two ends? The gap capacitance for $e = Q i^{***} 1^*$ used.
- 7.6. Design a circular disk resonator operating in the TM_{10} mode at 6 GHz. The substrate to be used has a dielectric constant $\epsilon'_r - j\epsilon'' = 6 - j0.005$ and 1 mm thick. The disk is made from copper. Determine the required radius of the disk and the Q of the resonator.
- 7.7. Find the resonant frequency and Q of a copper rectangular cavity of dimensions $a = b = d = 10 \text{ cm}$ for the TE_{111} mode.
- 7.8. A cylindrical cavity of radius $a = 2 \text{ cm}$ and a length of 6 cm is filled with a dielectric with permittivity $\epsilon = (2.5 \times 0.0001)\epsilon_0$. The cavity is made of copper. Find the resonant frequency and Q for the TE_{111} mode. Note that, in the expression for resonant frequency, the velocity of light in free space, c , must be replaced by

- 7.9. Use the results given by (4.25) and (4.26) to show that the Q of a ^{at} given by

$$\frac{\alpha x}{* 2R \text{ flu}, \quad 2(3 \angle \delta)}$$

where $Z_m = R + jX$ and $Y_m = G + jB$ are the input impedance and admittance at the terminals. Verify that these formulas give the usual results series RCL network and for a parallel RCL network.

- 7.10. A rectangular cavity of dimensions u, b, d is coupled to a ^{rec U n g ^ a in an} through a capacitive slit. The guide width is a , and the height is b . ^{an J a} equation for determining the first antiresonant frequency. Find ^f ^{slit suscep} for critical coupling. For $a = 26 = 2.5 \text{ cm}$, a ^f ^{op} copper cavity, compute the resonant frequency Q and ^{^A} ^{-j_{it}} critical coupling. Use the formula $B = (206/\pi) \text{tncsc}(-\pi/2)^*$ susceptibility. What is the loaded Q ?

$tT.$



FIGURE P7.10

- 7.11.** For a capacitive diaphragm in a rectangular guide of the dimensions given in Prob. 7.10, obtain an equivalent circuit to represent the susceptance function $B_c = (2/36/-) \ln \csc(777/26)$ correct to terms up to $Aw = w - w_l$ in the vicinity of the frequency ω_j .

Hint: Expand $1/\sqrt{1 - x}$ in a Taylor series about OJ , and choose a series LC circuit.

- 7.12.** Design a rectangular cavity of length d , height $b = 1.2$ cm, width $a = 2.5$ cm that will resonate at 10,000 MHz. The cavity is critically coupled to a rectangular guide of dimensions a by b . Specify the cavity length d and the radius of the centered circular aperture. Determine the unloaded and loaded Q 's if the cavity is made of copper.

- 7.13.** For the aperture-coupled rectangular cavity discussed in the text, let the incident power at resonance be 100 mW. The cavity is critically coupled. Evaluate the peak value of the electric field in the incident wave and in the cavity field. How does the peak amplitude of the cavity field depend on the cavity Q ? (See Prob. 7.12 for cavity dimensions.)

- 7.14.** A hemispherical resonator is to be used at 10 GHz. It is made from a dielectric with a dielectric constant of 100 and is mounted on a ground plane. Find the required radius of the hemisphere and the radiation Q_r .

- *7.15.** A cavity is excited by an impressed electric field E_{\parallel} , tangent to an aperture surface S_a cut in the cavity wall S . Use the relation $\int_V \mathbf{f} \cdot d\mathbf{V} = \mathbf{E} \times \mathbf{G}_{\parallel}$, dV to show that the amplitude g_{\parallel} is given by

$$-i\omega n g_{\parallel} = (bn \times E_{\parallel} \cdot G_{\parallel}) dS = I_{\parallel} n \times E_{\parallel} \cdot G_{\parallel} dS$$

for a cavity with perfectly conducting walls. Using the general expansion (7.1156), show that an alternative expression is

$$-y'' w M_{\parallel} g_{\parallel} = -JW \cdot \partial F^{H_{\parallel} + G_{\parallel}} / \partial V = \frac{\partial}{\partial V} \left(\frac{JW}{2\pi c} \right)^{1/2} j_{\parallel}^n \cdot H_{\parallel}^{1/2} dS$$

upon putting $H_{\parallel} \cdot V I_{\parallel} = V \cdot H_{\parallel} I_{\parallel}$, since $V \cdot H = 0$ and using the divergence theorem. This last relation shows that the G_{\parallel} modes are excited whenever $n \cdot H$ does not vanish over S .

- *7.16.** Show that the two expressions for g_{\parallel} in Prob. 7.15 are identical.

Hint: Consider

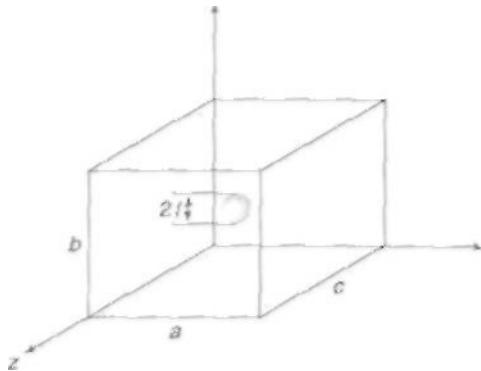
$$n \cdot V \times E_{\parallel} = (\nabla \cdot n \cdot V \times E - n \cdot E \times V)_{\parallel} = -j\omega X(j \cdot n \cdot H - n \cdot V \times E)_{\parallel}$$

and use Stokes' law to show that

$$\oint_S n \cdot V \times E_{\parallel} dS = 0$$

- *7.17.** Find the eigenfunctions E_{\parallel} , H_{\parallel} , F_{\parallel} , and G_{\parallel} for a rectangular cavity of dimensions a , b , c .

- *7.18.** Obtain a modal expansion similar to (7.126a) and (7.130) for the electromagnetic field in a cavity excited by a volume distribution of electric current $J(r)$.

**FIGURE P7.19**

- *7.19. (a) For the box cavity illustrated in Fig. P7.19, show that the *normalized* electric field mode function for the TE_{101} mode is

$$\mathbf{E}_{101} = \frac{v_x}{abc} \sin \frac{\pi x}{a} \sin \frac{\pi z}{c} \hat{\mathbf{y}},$$

- (b) Let a metal sphere of radius r be placed in the center of the cavity. Show that when $a = b = c$ and $r = a/20$ that the TE_{111} -mode resonant frequency is lowered by $TT/20$ percent.

- *7.20. A dielectric sphere of radius r and complex relative permittivity $\epsilon_r = t'_r - j\epsilon''_r$ is placed in a cavity. Show that the perturbed resonant frequency for the n th cavity mode is given by

$$k_o = k \sqrt{\frac{2Q_n}{\epsilon_r}} \left(\frac{\epsilon_r + 1}{2} \right)^{1/2} \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} \right)^{1/2} \quad \boxed{}$$

where the cavity Q due to the lossy dielectric is given by

$$Q_d = \frac{(t_o + 2)^2 + (\epsilon''_r f)^2}{2 \pi f \epsilon_r \epsilon_0},$$

REFERENCES

1. Montgomery, C. G., R. H. Dieke, and E. M. Purcell: "Principles of Microwaves," McGraw-Hill Book Company, New York, 1948. Comp^{4th} - biC ?
2. Ragan, G. L. (ed.); "Microwave Transmission Circuits," McGraw-Hill, New York, 1948.
3. Slater, J. C.: "Microwave Electronics," D. Van Nostrand Company, N.J., 1950.

4. Goubau. G.: "Electromagnetic Waveguides and Cavities," chap. 2, Pergamon Press, New York, 1962.
5. Kurokawa. K.: The Expansions of Electromagnetic Fields in Cavities, *IRE Trans.*, vol. MTT-6, pp. 178-187, April, 1958.
6. Kajfez, D., and P. Guillon: "Dielectric Resonators," Artech House Books, Dedham, Mass., 1986.
7. Van Bladel. J.: The Excitation of Dielectric Resonators of Very High Permittivity. *IEEE Trans.*, vol. MTT-23, pp. 208-217, 1975.
8. Gastine, M., Courtois, L. and J. L. Dormann: Electromagnetic Resonances of Free Dielectric Spheres, *IEEE Trans.*, vol. MTT-15, pp. 694-700, 1967.
9. Chow, K. K.: On the Solution and Field Pattern of Cylindrical Dielectric Resonators, *IEEE Trans.*, vol. MTT-14, p. 439, 1966.

CHAPTER 8

PERIODIC STRUCTURES AND FILTERS

Waveguides and transmission lines loaded at periodic intervals with identical obstacles, e.g., a reactive element such as a diaphragm, are referred to as periodic structures. The interest in waveguiding structures of this type arises from two basic properties common to all periodic structures, namely, (1) passband-stopband characteristics and (2) support of waves with phase velocities much less than the velocity of light. The passband-stopband characteristic is the existence of frequency bands throughout which a wave propagates unattenuated (except for incidental conductor losses) along the structure separated by frequency bands throughout which the wave is stopped and does *not propagate*. The former is called a passband, and the latter referred to as a stopband. The passband-stopband property is of interest for its frequency filtering aspects.

The ability of many periodic structures to support a wave having a phase velocity much less than that of light is of basic importance in traveling-wave-tube circuits. In a traveling-wave tube, efficient coupling between the electron beam and the electromagnetic field is obtained if the phase velocity is equal to the beam velocity. Since the latter is greater than 10 to 20 percent of the velocity of light, considerable absorption of the electromagnetic wave is required. Periodic structures for use in traveling-wave tubes are discussed in this chapter. The principles of operation of the tube are covered in Chap. 9.

The last part of the chapter is devoted to an introduction to filter theory. A complete treatment of all aspects of filter theory would be much too lengthy to include in this text. However

material is covered to provide a background so that the technical literature can be read without difficulty.

r APACITIVELY LOADED TRANSMISSION-

L M E - C I R C U I T ANALYSIS

To introduce a number of basic concepts, methods of analysis, and typical properties of periodic structures, we shall consider a simple example of a capacitively loaded transmission line. For a physically smooth transmission line, such as a coaxial line, the phase velocity is given by

$$v_p = (LC)^{-1/2} = (Mo^0)^{-1/2} \quad [8.11]$$

where ϵ_r is the dielectric constant of the medium surrounding the conductor. A significant reduction in phase velocity can be achieved in a smooth line only by increasing ϵ_r . This method has the great disadvantage that the cross-sectional dimensions of the line must also be reduced to avoid the propagation of higher-order modes. The phase velocity cannot be decreased by increasing the shunt capacity C per unit length because any change in the line configuration to increase C automatically decreases the series inductance L per unit length, since $LC = \lambda_0 e$. However, by removing the restriction that the line should be physically smooth, an effective increase in the shunt capacitance per unit length can be achieved without a corresponding decrease in the series inductance L . That is, lumped shunt capacitance may be added at periodic intervals without affecting the value of L . If the spacing between the added lumped capacitors is small compared with the wavelength, it may be anticipated that the line will appear to be electrically smooth, with a phase velocity

$$v_p = \frac{c}{f^{1-1/2}} \quad (8.2)$$

where C_0/d is the amount of lumped capacitance added per unit length (a capacitor C_0 added at intervals d). The following analysis will verify this conclusion.

One method of obtaining shunt capacitive loading of a coaxial transmission line is to introduce thin circular diaphragms at regular intervals, as in Fig. 8.1. The diaphragms may be machined as an integral part of the center conductor. The fringing electric field in the vicinity of the diaphragm increases the local storage of electric energy and hence may be accounted for, from a circuit viewpoint, by a shunt capacitance. The local field can be described in terms of the incident, reflected, and transmitted dominant TEM mode and a superposition of an infinite number of higher-order E modes. If the cylinder spacing $b - a$ is small compared with the wavelength, the higher-order modes are evanescent and decay to a negligible value in a distance of the order of $b - a$ away from the diaphragm in either

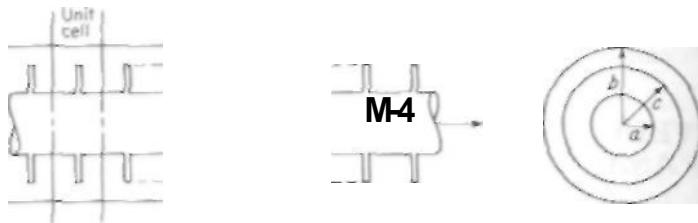


FIGURE 8.1

Capacitive loading of a coaxial line by means of thin circular diaphragms.

direction. An approximate expression for the shunt susceptance of the diaphragm is

$$a = \frac{B}{V} = \frac{8(6 - ef)}{V} \ln \frac{(6/a)}{\ln \frac{esc}{2b-a}} \quad (8.3)$$

where $Y_c = [601n(6/a)J^{-1}]$ is the characteristic admittance of an air-filled coaxial line. The expression for B is accurate for $b - a < 0.1A_0$. In this low-frequency region, B has a frequency dependence directly proportional to ω . At higher frequencies B will have a more complicated frequency dependence, although the thin diaphragm can still be represented by a shunt susceptance.

The circuit, or network, analysis of a periodic structure involves constructing an equivalent network for a single basic section or unit cell of the structure first. This is followed by an analysis to determine the voltage and current waves that may propagate along the network consisting of the cascade connection of an infinite number of the basic networks. For the structure of Fig. 8.1, an equivalent network of a basic section is a shunt normalized susceptance B with a length $d/2$ of transmission line on each side, as in Fig. 8.2a. Figure 8.26 illustrates the voltage-current relationships at the input and output of the n th section in the infinitely cascade connection.

The relationships between the input variables V_n, I_n , and the output variables V_{n+1}, I_{n+1} are readily found by using the $a/W^T J^{tra^f ren}$ matrix discussed in Sec. 4.9. The V_n and I_n are the total voltage and current amplitudes, i.e., the sum of the contributions from the incident and reflected TEM waves at the terminal plane. The circuit for a unit section is broken down into three circuits in cascade, namely, a shunt line of length $d/2$ (electrical length $\theta/2 = k_0d/2$) followed by

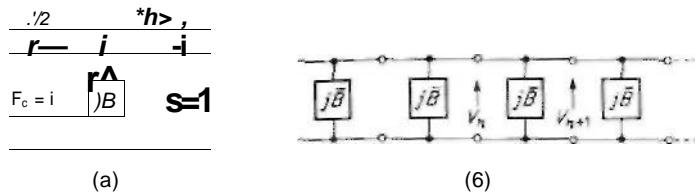


FIGURE 8.2

(c) Equivalent circuit for unit cell of loaded coaxial line; (b) cascade connection of basic unit-cell networks.

susceptance B , which in turn is followed by another length of transmission line. The 4×4 matrix for each of these individual networks is, respectively (Prob. 4.26),

$$\begin{array}{ccccc} & & & \mathbf{e} & \\ \cos- & J \sin- & & 1 & \cos- & JSin \\ & & & & & \\ & & & JB & & \\ & & & & & \\ J \sin- & \cos- & & & jsm- & \cos- \end{array}$$

The transmission matrix for the unit cell is obtained by the chain rule [see (4.75)], i.e., the product of the above three matrices, and hence we have

$$\begin{bmatrix} \mathbf{V} \\ \mathbf{h} \end{bmatrix} = \begin{vmatrix} \mathbf{e} & \mathbf{e} & 1 & \mathbf{e} \\ \cos^2 & jsm^2 & 0 & \cos- & 8 \\ \mathbf{e} & \mathbf{e} & 1 & 7 \sin- & 8 \\ j \sin- & \cos- & /sin- & 0 & 1 \\ \cdot \mathbf{0} & & & & \cdot \mathbf{I}_{3,3} \end{vmatrix} \begin{bmatrix} \mathbf{K}_{+I} \\ I_{3,3} \\ \cdot l^B - \cos 0 + smf^Q - \frac{s}{j} \cdot 1 \\ IB \\ j - \cos 8 + \sin 8 + \frac{B'}{B} \\ \cos 8 - \frac{B}{\sin 8} \end{bmatrix} \quad (8.4)$$

Note that $smf = 9$, which is always true for a symmetrical network, i.e., a symmetrical unit cell.

If the periodic structure is capable of supporting a propagating wave, it is necessary for the voltage and current at the $(n + 1)$ st terminal to be equal to the voltage and current at the n th terminal, apart from a phase delay due to a finite propagation time. Thus we assume that

$$K^*i_{n+1} = i_n \quad (8.5a)$$

$$h^* = e^{-T^*l_n} \quad (8.56)$$

where $y = jp + a$ is the propagation constant for the periodic structure. In

terms of the transmission matrix for a unit cell, we now have

$$\text{or} \quad \begin{vmatrix} V_{11} & V_{12} \\ h & f \end{vmatrix} \begin{vmatrix} V_{21} & V_{22} \\ e^{yd} & 0 \end{vmatrix} = \begin{vmatrix} K_{11} & K_{12} \\ 1 & 0 \end{vmatrix} \quad (8.6)$$

This equation is a matrix eigenvalue equation for y . A nontrivial solution for V_{11}, V_{22} , exists only if the determinant vanishes. Hence

$$j^2 - e^{2y} - tfB - \omega W + e^{2y} v - e^{2y} (sS + \omega) = 0 \quad (8.7)$$

For a reciprocal network the determinant $s^2 - \omega^2 - k^2 d^2$ of the transmission matrix equals unity (Sec. 4.9); so we obtain

$$\cosh yd = \frac{\sinh yd + Q_i}{\sinh yd - Q_i} \quad (8.8)$$

For the capacitively loaded coaxial line, (8.8), together with (8.4), yields

$$\cosh yd = \cos \theta - \frac{B}{2} \sin \theta \quad (8.9)$$

When $\cos \theta - (B/2) \sin \theta < 1$, we must have $y = jf$ and $a = 0$; that is,

$$\cos lid = \cos \theta - \frac{B}{2} \sin \theta \quad (8.10a)$$

When the right-hand side of (8.9) is greater than unity, $y = a$ and $(i - 0)$; so

$$\cosh ad = \cos \theta - \frac{B}{2} \sin \theta > 1 \quad (8.10b)$$

Finally, when the right-hand side of (8.9) is less than -1, we must have $yd = jir + a$, so that

$$\begin{aligned} \cosh yd &= \cosh(jir + ad) = -\cosh ad \\ &= \cos 0 - \frac{B}{2} \sin 1 < -1 \end{aligned} \quad (8.10c)$$

It is apparent, then, that there will be frequency bands for which unidirectional propagation can take place separated by frequency bands in which wave is attenuated. Note that propagation in both directions since y is also a solution.

A detailed study of the passband-stopband characteristic frequency ω_p Sec. 8.6. For the present we shall confine our

attention to the WJ!* then 8i* limiting value of y . When $d \ll A_0$, $\theta = k_0 d$ is small, and pa

be small. Replacing $\cos \theta$ by $1 - 0^2/2$ and $\sin \theta$ by 0 in (8.10a) gives

$$\cos \theta d - 1 \sim -g - 1 - \frac{f c^2 d^2}{B * 0 d} = -$$

Using the relations $k\% = w^2 M \ll O = a^2 LC$ and $B = B/Y_c = \sqrt{GJL/CW}$, where $wC_0 = B$, we obtain

$$\beta^2 = \ll^2 LC + \frac{w^2 LC u}{a}$$

and hence

$$f = \ll f Z | C7 + \wedge J \quad (8.11)$$

Therefore we find that, at low frequencies where $d < K A_0$, the loaded line behaves as an electrically smooth line with a shunt capacitance $C + C^2/d$ per unit length. The increase in f_i results in a reduction of the phase velocity by a factor k_0/f_i .

Another parameter of importance in connection with periodic structures is the normalized characteristic impedance Z_B presented to the voltage and current waves at the reference terminal plane, i.e., input terminals of a unit cell. An expression for Z_n may be obtained from (8.5), which may be written as

Hence

$$Z_H = \frac{V_{,,}}{I_B} = \frac{-M}{i} = \frac{v' - e'''}{e''} \quad (8.12)$$

Replacing $2e'''$ by $.rf + 0 \pm [(1/ + \&)^2 - 4]^{1/2}$ from (8.7), we obtain

$$Z_S = \frac{\% m}{\bullet s f \pm y O^2 + 4} \quad (8.13a)$$

where the upper and lower signs refer to propagation in the $+z$ and $-z$ directions, respectively. We are using the convention that the positive directions of $V_{,,}$ and $I_{,,}$ are those indicated in Fig. 8.2, independent of the direction of propagation. For a symmetrical network, $si = C^2$ and since $\&& > -\&& < \& = 1$, we have $i/2 - 1 = m\%$. In this case (8.13a) reduces to

$$Z_S = \pm \frac{V_{,,}}{\pm V^2 + 4} = \pm V_{-} \quad < 813e >$$

In general, for a lossless structure, $Z'_a = \sim(Z'_b)^*$ in the passband, since $|st + S| < 2$, as (8.8) shows.

If the unit cell is represented by a T network with parameter Z_{22} , then, by using the relations between the $st/xi't$, p - pg . the impedance parameters given in Sec. 4.9. we can also show th

"• —
rs anΛ

$$\cosh yd = 2Z_{12} \quad (8.1_4)$$

$$\bar{Z}_{,,} = \frac{*^{\wedge} L 1}{\pm A'^{\wedge}} \frac{A'^{\wedge}}{Z_{Vi} \sinh yd} \quad (8.15)$$

The waves that *may* propagate *along a periodic structure* are oft* called Bloch waves by analogy with the quantum-mechanical electron wa that may propagate through a periodic crystal lattice in a solid. It is for t] reason that we have denoted the characteristic impedance as Z for th' Bloch wave. The voltage and current at the n th terminal plane will h denoted by $V_{,,}, I_{B^*}$, for the Bloch waves from now on instead of by the quantities $V_{,,}, I_n$. The + and - signs refer to Bloch waves propagating in the +z and -z directions, respectively. We shall also adopt the convention that the positive direction of current flowjbr Bloch waves is always in the +z direction; thus $I'_b = Y^{\wedge} V^{\wedge}$ and $I_b = Y_j V_g$. However, for a symmetri- cal structure such that $-' = V$, we shall have $?_{,,} = - Y^{\wedge} = -(Z^{\wedge})^4$.

If (8.13) is used, we find that, for the loaded coaxial line,

$$\bar{Z}_B = \sqrt{\frac{B}{i}} = \sqrt{\frac{2 \sin \theta + B \cos \theta - B}{V \sin \theta + B \cos \theta + B}} \quad (8.16)$$

In the low-frequency limit, where we can replace $\sin \theta$ by

$$0 = k_0 d = i \omega jLC$$

and $\cos \theta$ by 1, we obtain

$$2_{,,} = \frac{20}{20 f 2B} \quad C + C_0/d \quad 20$$

and thus

$$Zp \sim ZffZ_{,,} - C + C_0/d$$

ically

Again we see that, in the low-frequency limit, the *loaded* $h^{\wedge} l^{\wedge} tidp' - ^1 \wedge$ smooth and the characteristic impedance is modified in $^{\wedge} \wedge$ $k_{>n} gth$ manner by the effective increase in the shunt capacitance $pe^{\wedge} \wedge$ &•

The characteristic impedance of a periodic structure is $^{\wedge} \wedge$ $(:e \backslash)$, quantity since it depends on the choice of terminal planes $^0 \wedge$ $tio 0 i$ the » the terminal planes are shifted a distance $/$ in the -

characteristic impedance becomes

$$\boxed{Z_{\infty} = \frac{Z_0 + j \tan A_n}{1 + j Z_H \tan A_n}} \quad (8.18)$$

WAVE ANALYSIS OF PERIODIC STRUCTURES

Periodic structures may be analyzed in terms of the forward- and backward-propagating waves that can exist in each unit cell with about the same facility as the network approach gives. In the wave approach the wave-amplitude transmission matrix $|A|$ discussed in Sec. 4.9 is used.

With reference to Fig. 8.3, let the amplitudes of the forward- and backward-propagating waves at the n th and $(n+1)$ st terminal plane be c^* , c_{-} , c_{+1}^* , and c_{+1} . The $e^*_{+B} e_{-H}$, are related to the c^*, c_{-} by the wave-amplitude transmission matrix as follows:

$$\begin{matrix} -AM & A_{12} \\ A_{21} & A_{22} \end{matrix} \quad (8.19)$$

The solution for a Bloch wave requires $c_{-} = e^{-y''} c^*$ and $c_{+} = e^{-y''} c^*$. Hence (8.19) becomes

$$A_{21} - e^{y''} c_{-} = 0 \quad (8.20)$$

A nontrivial solution for c_{-} , c_{+} , is obtained only if the determinant vanishes. Consequently, the eigenvalue equation for y is

$$\begin{aligned} A_{11} A_{22} - A_{12} A_{21} + e^{2y''} (A_{11} - A_{22}) &= 0 \\ \text{or} \quad \cosh yd &= \frac{A_u + A_{-v}}{A_u - A_{-v}} \end{aligned} \quad (8.21)$$

since the determinant of the transmission matrix, that is, $A_{11} A_{22} - A_{12} A_{21}$ equals 1 when normalized wave amplitudes are used.

— T——1 «&1

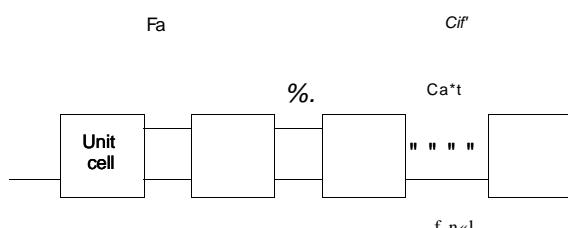


FIGURE 8.3
Wave amplitudes in a periodic structure.

The Bloch wave which can propagate in the periodic structure up from forward- and backward-propagating normal transmission waves that exist between discontinuities. When v_h or determined from (8.21), the ratio of c_∞ to c' is fixed. This ratio is the characteristic reflection coefficient Y_B . Thus the transverse $c_{\perp n}^{tL}$ field of the Bloch wave will have an amplitude

$$V_{BO} = c_0 + c_0 = c_0(1 + 1''_B)$$

at the zeroth terminal plane and an amplitude

$$v_{ft,} = <\cdot;+ c; = c; (i + v_B) = c; (i + v_B)e^{-y''d} \quad (8.22a)$$

at the n th terminal plane. The transverse magnetic field of the Bloch wave will have an amplitude

$$/_{j,} = c_0^*(1 - r_{fi})e^{-y''d} \quad (8.22b)$$

at the r th terminal plane.

The characteristic reflection coefficient may be found from the pair of equations (8.20) by eliminating e^{-yJ} by the use of (8.21). It is usually more convenient to express r_B in terms of Z_B by using the relation $Z_B = (1 + r_B)/(1 - Y_B)$. Thus we have

$$\mathbf{r} = \frac{Z_S - 1}{Z_S + 1} \quad (8.23)$$

where the + and - signs refer to Bloch waves propagating in the $+z$ and $-z$ directions, respectively.

The above wave formulation is now applied to the capacitively loaded transmission line discussed earlier. The unit cell is chosen as in Fig. 8. The wave-amplitude transmission matrices for the three sections of the unit cell are (Sec. 4.9 and Prob. 8.7)

$$\begin{matrix} 2+jB & & B \\ & 0 & \\ & , -Jkudn & \\ & & \mathbf{B} \\ & & 4+B^2 \\ & & 2(2+jB) J \end{matrix}$$

and another matrix like the first one. The $[A]$ matrix for the unit cell is obtained by multiplying the three component matrices together.

$$\begin{matrix} 2 & +jB & & B \\ & & & \\ [A] = & \begin{matrix} >/2 & 0 \\ , -J/kd & B \end{matrix} & & \begin{matrix} jk^2 \\ 0 \end{matrix} \\ & & & \\ & & & 4 + B^2 \\ & & & 2(2 + jB) \end{matrix}$$

where $\theta = k_f d$. After multiplication we obtain

$$[A] = \frac{2+jB}{-j} \cdot \frac{B}{-j} \cdot \frac{4 + B^2}{2(2 + jB)^2} \quad (8.24)$$

Making use of (8.21), we find that

$$\cosh yd = \frac{(4 + B^2)e^{j\theta} + (2 + B) V}{My^2 S} = \cos \theta, \sin \phi$$

which is the same as (8.9) obtained earlier.

« 3 PERIODIC STRUCTURES COMPOSED OF ^ SYMMETRICAL TWO-PORT NETWORKS

The capacitively loaded coaxial transmission line can be considered as made up of *symmetrical* sections by choosing terminal planes midway between each diaphragm. For other choices of terminal-plane positions the unit cell would be unsymmetrical, and its equivalent T network would then also be unsymmetrical. Other types of periodic structures are composed of intrinsically unsymmetrical unit cells such that there is no terminal-plane location that will reduce them to a symmetrical structure. Several unsymmetrical structures are illustrated in Fig. 8.4.

For nonsymmetrical structures the Bloch-wave characteristic impedance is given by (8.15), which we rewrite as

$$f W + Z \quad (8.25a)$$

$$(8.256)$$

nnnnnn

I^*
in
A.

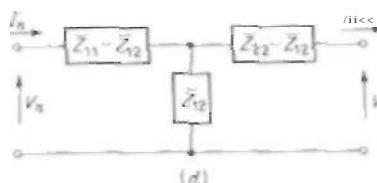


FIGURE 8.4

Periodic structures with unsymmetrical unit cells, (a, b) Rectangular waveguide loaded with thick unsymmetrical diaphragms; (c) coaxial line loaded with diaphragms and dielectric rings; (rf) equivalent T network of a unit cell.

where

$$\mathfrak{C} = \frac{Z_{12}}{Z_{21}}, \quad (8.26a)$$

$$Z = \pm Z_{12} \sinh yd = \pm JZ_{12} \sin \ell d \quad (8.266),$$

and the sign is to be chosen so that Z has a positive real part, Th_{he} constant $\$$ is given by

$$\cos fid = \frac{Z_{u+}Z_{-}}{2Z_{12}} \quad (8.26c)$$

in the propagation band. The physical length of a unit cell is d . Th quantities Zt and lid are often called the iterative parameters of the T network. A consequence of the nonsymmetry of the unit cell is that Z_B^A is different from Z_B .

Let the voltage of the Bloch wave at the zeroth terminal plane be V - where the signs + or - refer to Bloch waves propagating in the +z and -z direction, respectively. The corresponding Bloch-wave current is $Z_0^A = V/a/Z'B^-$. At the n th terminal plane the Bloch-wave voltages and currents will be

$$V_B^n = V_{B0}^n e^{f y i n t} \quad (8.27b)$$

Recall that we are taking the positive direction of current flow to be in the +z direction, independent of the direction of propagation for the Bloch waves.

If the restriction is made that the only points at which the voltages and currents will be specified are the terminal planes, the periodic structure has properties similar to any uniform transmission line or waveguide. Such transmission-line theory can be applied to study the effects of terminating a periodic structure in an arbitrary load impedance, matching sections for periodic structures, etc. These applications will be discussed in the following two sections.

8.4 TERMINATED PERIODIC STRUCTURES

Figure 8.5 illustrates a periodic structure terminated in a load impedance Z_L at the N th terminal plane. The total voltage and current at the N th terminal plane will be a superposition of an incident and reflected waves.

"sol

rx t

FIGURE 8.5
Periodic structure terminated in a load Z_L .

wave; thus

$$\frac{V}{V_{BN}} = \frac{V}{V_{B\sim e}} p^{\sim : W} + \frac{V}{V_{B0^e}} pJP^{\sim : i} \quad (8.28a)$$

$$\frac{I}{-BN} = I^* \frac{p \sim - ll^{\text{load}}}{-B\sim e} + \frac{J' u}{\sim B0^e} V^{\sim : >} \quad (8.28b)$$

(8.28c)

where $Y_B = Zg^1$. At the N th terminal plane we must have

$$I' \sim *BN \sim "IABN \sim -'III.$$

and hence

$$V \mathcal{E}_N + V_{B-A} = Z_L (Y \sim V_{iN} + ?_B V_{HN}) \quad (8.29)$$

The reflection coefficient f , of the load for Bloch waves is, from (8.29),

$$\begin{aligned} \frac{r_{BN}}{BN} &= \frac{Z_L Y S - I}{Z_L \wedge B \sim - 1} & -'B Z_i &= Z_L \\ && B Z_{\sim}, - Z'_R & \\ Z - f Z_L Z - & j & & \\ z + t z_L + z - & (& & \end{aligned} \quad (8.30)$$

For a symmetrical structure $f = 0$ and the expression for I' , reduces to the usual form.

The Bloch-wave reflection coefficient at the n th terminal plane is

$$r_n = \frac{\frac{V_{BN}}{V_{BN^e}} e^{-j(N-n)d}}{V_{BN^e}} = r_e - j x N - n) p d \quad (8.31)$$

The input impedance at the n th terminal plane is

$$\begin{aligned} Z_n &= \frac{\frac{V_L + V_E}{Jin + I_{Bn}} + V_j Wi + r_j}{V_s + V_{Bn} Y_B} \\ &\quad \frac{1 + r_n}{*B + r_n, y_B} \quad \frac{Z_B + Zf(l + r_n)}{Z_{\sim} + Z : r} \end{aligned} \quad (8.32)$$

From (8.32) we can also obtain the alternative expressions

$$ZB + ZIV_B \quad n \quad (8.33a)$$

$$r_{ss} = -\frac{(Z_s - Q - Z)}{Z_B} \quad \frac{\{z_n C\} + Z}{(8.336)}$$

If (8.31) is used to express F_{ss} in terms of Y_L and (8.30) to express F_{ss} in terms of Z_s , we find that (8.33a) gives

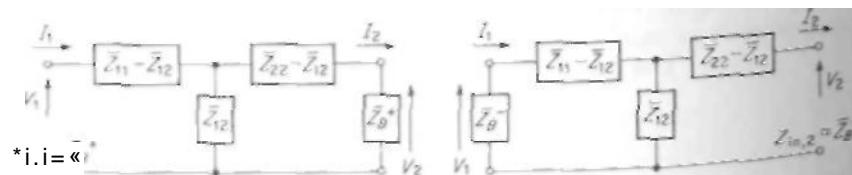
$$\begin{aligned} & \ll \quad g^+ - f + jZ \quad \tan(N-n)f/d \\ & " \quad " \quad Z + ./(Z_t - f) \tan(N - .,)/3d \quad <^{8-34} \end{aligned}$$

This equation gives the transformation of impedance along a periodic structure. It differs somewhat from the usual transmission-line formula when the unit cell is unsymmetrical, so that $f \neq 0$.

For a Bloch wave propagating in the $+z$ direction, the periodic structure must be terminated in a load $Z_s = Z_B = f + Z$ to avoid a reflected wave. Similarly, the matched-load termination for a Bloch wave propagating in the $-z$ direction is $-Z_B = Z - C$. The two characteristic Bloch-wave impedances are the iterative impedances for the T network of the unit cell. With voltages and currents chosen as in Fig. 8.6, it is readily shown that an impedance Z_B connected at terminals 2 is transformed into itself at terminals 1. Similarly, an impedance $-Z_B$ connected at terminals 1 is transformed into itself at terminals 2. It is for this reason that Z_g is called an iterative (repeating) impedance. For a lossless T network, f is pure imaginary and Z is pure real in the propagation band. Ambiguity in the sign of Z as given by (8.15) or (8.266) may be avoided by noting that, in a passband, Z is real and positive, in order to be consistent with our choice of positive direction for current. We must have positive real power transmission, and hence

$$P = Re kVf(ltf) = Rei|/\ll|^2 Z^+ = \frac{VlfZ}{l} > 0 \quad (8.35)$$

Another criterion that may be used is the one requiring the reactive part of



(61)

FIGURE 8.6
Iterative impedance properties of a T network.

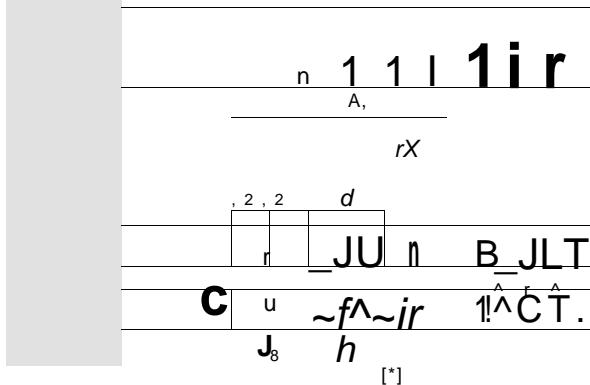


FIGURE 8.7

(a) Tapered transition matching section for a diaphragm-loaded rectangular guide; (b) quarter-wave transformer matching of a capacitively load-d coaxial line.

Z_g to have a positive derivative with respect to w (Sec. 4.3). There is also reactive power in a Bloch wave in a passband, and this is given by

$$* \text{ reactive} \quad 2'' \ll & \quad (8.36)$$

Complex power for a Bloch wave propagating in the $-z$ direction is given by $-^*V_H(I_B Y^*)$ because of our choice of direction for positive current.

MATCHING OF PERIODIC STRUCTURES

If a periodic structure is connected to a smooth transmission line or waveguide, *some means of matching* the periodic structure to the input waveguide must be provided to avoid reflection of the incident power. A situation encountered quite frequently is the one where the periodic structure is identical with the input waveguide apart from the periodic loading. One way of providing a matched transition from the unloaded to the loaded waveguide is to use a tapered intermediate section. The matching taper section is similar to the loaded waveguide except that the periodic loading is gradually reduced to zero over a distance of about a wavelength. Figure 8.7a illustrates a tapered transition in a rectangular waveguide connected to a similar guide periodically loaded with diaphragms.

Any of the matching networks discussed in Chap. 5 may also be used to match a periodically loaded guide to an unloaded guide. For example, at some distance $d'/2$ in front of the first terminal plane for the periodic structure, the characteristic admittance Y_B of the periodic structure is transformed into an admittance $1 - jB'$, so that placing a shunt susceptance jB' at this point provides a matched transition. Matching by means of a shunt susceptance may be viewed as an application of the quarter-wave transformer matching technique. The unit cell consisting of the shunt susceptance jB' plus a length $d'/2$ of transmission line (or waveguide) on either side, as in Fig. 8.7b, may be considered as part of an infinite periodic structure with a propagation phase constant f' and a normalized character-

istic impedance Z'_a . If the parameters B' and d' are chc (5'd' = 7r/2, $Z'_B = Z^{\wedge/2}$, then at the input terminal to the matok (8.34) gives so t aiciwing ^ . ^

$$7 - ' \circ = 1$$

(8.37)

Note that the matching section is a symmetrical structure and $\ell' = 0$. We also require that Z_u be real in order for (8.37) to h solution for Z'_B . When these conditions are met, we see that the section behaves essentially as a quarter-wave transformer. For svmm structures the required values of B' and d' may be found from (8.136); thus, as $\cos B'd' = \cos(\pi/2) = 0$, we have

$$= 9 >' = 0 \quad (8.38a)$$

$$\left(\frac{Z'}{Z_B} \right)^2 = r \quad (8.386)$$

For the capacitively loaded transmission line, we obtain, by using (8.4) and (8.16) applied to the matching section,

$$2 \cot f_n d' = B' \quad (8.39a)$$

$$\frac{(Z' B) I}{(Z' B) L}^2 = \frac{2 \sin k_n d' + B' \cos k_n d' - B'}{2 \sin \ell_n a'' + B' \cos \ell_n a'' + B'} = \tan^2 \frac{\pi d'}{a''} = \frac{J}{Z_B} \quad (8.396)$$

when (8.39a) is used to eliminate B' . The above results are the equivalent of those derived in Chap. 5, i.e., given by (5.8) and (5.9).

To obtain a match over a wide frequency band, more elaborate matching networks must be used since a single shunt susceptance usually does n< provide a match over a wide frequency band. Broadband matching is complicated by the fact that the characteristic impedance Z_B of a periodic structure is a function of frequency. No general technique exists for design ing broadband matching networks because of the general nature of a- Each periodic structure must be considered by itself so that the treq variation in Z_B can be incorporated into the design. For this reasi matching problem is not discussed any further.

$k_0 B$ DIAGRAM

We now turn to a detailed study of the passband-stopband characteristics earlier the capacitively loaded coaxial transmission line discussed m sections. The information contained in the eigenvalue propagation constant B in a periodic structure is usually plot (or co-B) plane. The curves of B versus k_0 show immediately the bands for propagation and also the stopbands in which no prepay

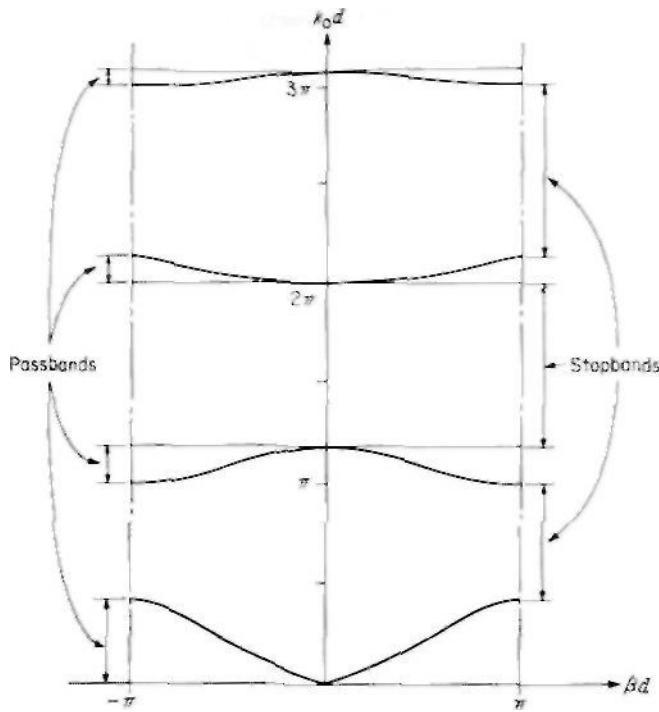


FIGURE 8.S
 k_0d - jid diagram for a capacitively loaded coaxial line, $B = 2k_0d$.

place. The resultant plot is called the k_0 - f_i diagram, or the Briilouin diagram,^t

For the capacitively loaded coaxial line, (8.9) gave

$$\cos (id = \cos k_0d - \sin k_n d) = \cos k_0d - Kk_n d \sin k_0d \quad (8.40)$$

where $B/l = wC_0/2Y_c$ has been expressed as $Kk_n d$. Curves of k_0d versus jid are sketched in Fig. 8.8 for $K = 1$, that is, for $B = 2k_n d$. A low-frequency passband exists for $0 < k_0d < 0.4167T$. This passband is followed by a stopband and further alternating passbands and stopbands. As $k_n d$ becomes large, the loading is increased, since B increases with k_0 . This has the effect of decreasing the width of the passbands in terms of frequency.

^tNamed after Briilouin, who used diagrams of this sort to illustrate the energy-band structure in periodic crystalline media.

The edges of the bands occur when the magnitude of the side of (8.40) exceeds unity. The lower edge of the first band when $0 < k_0d < v$ and

$$\cos k_0d - Kk_0d \sin k_0d = -1$$

This equation may be solved for k_0d to give

$$\begin{aligned} \cot \frac{k_0d}{2} &= Kk_0d \\ \cos \frac{k_0d}{2} &= 0 \end{aligned} \quad (8.416)$$

The corresponding principal value of k_0d is $\pi/2$, and the values of k_0d obtained from (8.41) mark the edges of all the bands for this value of f . The edges of the bands where $k_0d = 0$ are obtained by equating (8.40) to unity, in which case we obtain

$$\tan \frac{M}{2} = -Kk_0d \quad i 8420$$

$$\sin - \quad (8.426)$$

One edge of the passband always occurs when the spacing between discontinuities equals one-half wavelength in the unloaded waveguide, in the present case, when k_0d is a multiple of $\pi/2$. When the spacing between discontinuities equals one-half wavelength, they may all be lumped together, with the result that the line becomes effectively loaded at a single point by an infinite susceptance (or reactance). Clearly, power transmission along the periodic structure must reduce to zero at this frequency.

Only the principal value of k_0d is plotted in Fig. 8.8. In addition $k_0d + 2n\pi$, where n is an arbitrary integer, are solutions. These other solutions are the propagation constants of the spatial harmonics into which the Bloch wave may be expanded. The spatial harmonics are discussed in Sec. 8.8.

The k_0-f diagram for other types of periodic structures exhibits features similar to those in Fig. 8.8. For example, if the capacitive loading is replaced by inductive shunt loading, the relative locations of the passbands and stopbands are interchanged. The zero-frequency region will be a band since the shunt inductors will short-circuit the line at zero frequency.

*8.7 GROUP VELOCITY AND ENERGY FLOW

The phase velocity for a Bloch wave in a periodic structure is given by

$$v_p^0 = \frac{\omega}{k_0} \quad (8.43)$$

With reference to Fig. 8.9, it is seen that k_0d/fid is the slope of the

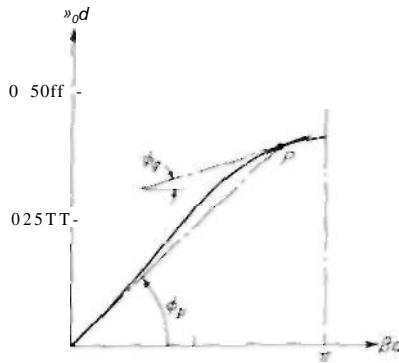


FIGURE 8.9

Enlarged drawing of first *passband* for a capacitively loaded transmission line, $B = 2k_0d$.

from the origin to a point P on the k_0d - v_g diagram. Since v_g is a function of w, the periodic structure has frequency dispersion. The group velocity v_g as given by (Sec. 3.19)

$$\frac{dv_g}{dk} = \frac{d(v_g)}{d(k_0d)} \quad (8.44)$$

is therefore different from the phase velocity. Again referring to Fig. 8.9, it is seen that the group velocity is equal to the slope of the tangent to the curve of k_0 versus $/id$ multiplied by the velocity of light c. Thus we have

$$v_g = c \tan \phi_p, \quad v_g = c \tan \phi_g$$

where ϕ_p and ϕ_g are the angles given in Fig. 8.9.

For the capacitively loaded coaxial line, use of the eigenvalue equation for 15, that is, (8.10a), enables us to obtain

$$v_g = c \frac{dk}{dft} \sim c \frac{do}{d(lid)} = c \frac{\sin fid}{\frac{B}{lk,d} T + \sqrt{1 - \sin^2 k_h d + \frac{B^2}{I} \cos^2 k_x d}} \quad (8.45)$$

This expression shows that the group velocity becomes zero when $/id = 0$ or $\pi/2$, except when k_f also equals zero. Thus, as the edges of the passbands are approached, the group velocity goes to zero.

The group velocity is also the signal velocity for any signal consisting of a sufficiently narrow band of frequencies such that $/id$ can be approximated by a linear function of w throughout the band. The signal delay r for propagation through a unit cell is given by

$$r = \frac{d}{c} \quad (8.46)$$

In Sec. 3.19 it was shown that for a waveguide, which is a dispersive medium, the velocity of energy flow in a propagating wave was equal to the group velocity. The same result will be shown to hold for a lossless periodic structure also.

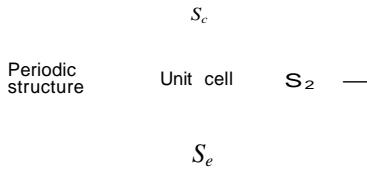


FIGURE 8.10
A unit cell of a periodic structure.

Consider a unit cell of a lossless periodic structure as in Fig. 8.10 & surface S is chosen to consist of surfaces S_1 and S_2 at the input a rf output terminal planes plus a cylindrical surface S_c surrounding the structure. If the periodic structure is enclosed by a perfectly conducting waveguide, the surface S_c coincides with the waveguide wall. If the periodic structure is an open-boundary structure, the surface S_c is that of a cylinder with infinite radius. In both cases $\mathbf{n} \times \mathbf{E}$ vanishes on S_c , so that the Poynting vector is zero over this surface. For generality we shall let the unit cell contain regions with frequency-dispersive material, i.e., material with parameters n and e that are functions of to .

In the derivation of Foster's reactance theorem in Sec. 4.3, it was shown that [see (4.24a)]

$$\begin{aligned} j) \mathbf{E} \times \frac{dU^*}{dco} + \frac{dco}{dco} \times \mathbf{H} \cdot d\mathbf{S} &= -j(\mathbf{H} \cdot \mathbf{H}) \frac{dco ii}{dco} + \mathbf{E} \cdot \mathbf{E}^* \frac{dco}{dco} \\ &= -4/(W_m + We) \end{aligned} \quad (8.47a)$$

since the latter integral is equal to four times the time-average energy stored in the volume bounded by S . Since the Poynting vector is zero on $i>_e$ and $d\mathbf{S}$ is directed inward, we have

$$\begin{aligned} \mathbf{A} \cdot \mathbf{E}, \mathbf{X} \frac{dU^*}{dco} + \frac{dco}{dco} \times \mathbf{H}, | - a, d\mathbf{S} \\ \int_{S_e}^{E_2} \mathbf{E}_2 \times \frac{dn^*}{dco} \cdot d\mathbf{S} = -4, (W_m + We) \end{aligned} \quad (8.47a)$$

where E, H are the fields at terminal plane 1 and E_2, H_2 are terminal plane 2. For a Bloch wave, $E_2 = E, e^{-pi/a}$, where pi shift through a unit cell of length a . We thus find that

$$\begin{aligned} \frac{dH^*}{da} &= \frac{aE}{dco} \quad \frac{m^*}{dco} \cdot \mathbf{F} \times \mathbf{H} \\ &+ \frac{1}{dco} \cdot \mathbf{X} \mathbf{H}, + \frac{1}{dco} \cdot \mathbf{T}_{du} \cdot \mathbf{i}^x \end{aligned}$$

Consequently, (8.476) gives (note that the integral over S_2 can be evaluated as an integral over S)

$$-2jt \frac{dp}{a \text{co}} [\text{E}^f_{\text{co}} \text{X H t - a}, dS = -4/7 \frac{dp}{a \text{to}} = -4/(W_m + W_t) \quad (8.476)$$

where $P = \text{Re } j_s E$, $X H f \bullet a$, dS is the power transmitted across a terminal plane. We now see that

$$v_g = \frac{dc o}{dp} = \frac{(W_m + W_t)/l}{(W_m + W_t)/l} \quad (8.48)$$

But the energy density $(W_m + W_t/l)$ in a unit cell multiplied by the velocity of energy flow is equal to the power P , and therefore the group velocity is the velocity of energy flow.

8.8 FLOQUET'S THEOREM AND SPATIAL HARMONICS

It has been noted that in an infinite periodic structure the field of a Bloch wave repeats at every terminal plane except for a propagation factor $e^{-y/d}$, where d is the length of a unit cell. Since the choice of location of a terminal plane within a unit cell is arbitrary, we see that the field at any point in a unit cell will take on exactly the same value at a similar point in any other unit cell except for a propagation factor $e^{-y/d}$ from one cell to the next. Thus, if the field in the unit cell between $0 < z < d$ is $E(x, y, z), THx, y, z$, the field in the unit cell located in the region $d < z < 2d$ must be

$$e^{-y/d} E(x, y, z) - d, e^{y/d} H(x, y, z) - d$$

Consequently, the field in a periodic structure is described by a solution of the form

$$E(x, y, z) = e^{-y} * E_{l>}(x, y, z) \quad (8.49a)$$

$$H(x, y, z) = e^{-y} * H_p(x, y, z) \quad (8.49b)$$

where E^\wedge and H_p are periodic functions of z with period d ; for example,

$$E_{l>}(x, y, z + nd) = E_p(x, y, z) \quad (8.49c)$$

The possibility of expressing the field in a periodic structure in the form given by (8.49) is often referred to as Floquet's theorem.^t From (8.49a) we

^tActually, Floquet's work dealt with differential equations with periodic coefficients. The case of periodic boundary conditions is an extension of that work.

see that the electric field at $z_l + d$ is related to the field at ?

$$E(x,y,z_l + d) = e^{-jWx} E(x,y,z_l)$$

follows:

which has, indeed, the correct repetitive properties of a Bloch

Any periodic function such as $E_p(x,y, z)$ may be expanded into an infinite Fourier series; thus

$$E_p(x,y,z) = \sum E_{pn}(x,y) e^{-j2\pi n z/d}$$

18.50

where E_{pn} are vector functions of x and y . Multiplying both sides by $e^{j2\pi m z/d}$ and integrating over a unit cell, i.e., from $z = 0$ to d , give

$$\int_0^d E_p(x,y,z) e^{j2\pi m z/d} dz = \sum E_{pn}(x,y) \int_0^d e^{-j2\pi n z/d} dz$$

since the exponential functions form a complete orthogonal set; i.e.

$$\int_0^d e^{-j2\pi m z/d} dz = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases}$$

The field in a periodic structure can now be represented as

$$E(x,y,z) = \sum_{n=-\infty}^{\infty} E_{pn}(x,y) e^{-j2\pi n z/d}$$

$$= \sum E_{pn}(x,y) e^{-j2\pi n z/d}$$

where $y = \sqrt{3}$ and $\beta_n = 2\pi n / d$. Each term in this expansion is called a spatial harmonic (or a Hartree harmonic) and has a propagation phase constant β_n . Some of the β_n will be negative whenever the integer n is sufficiently negative. The corresponding phase velocity of the n th spatial harmonic is

$$v_{nn} = \frac{c}{\beta_n} = \frac{c}{2\pi n / d} = \frac{d}{2\pi n} c$$

10

w

(8.53)

and will be negative whenever n is negative. The group velocity of the harmonic is

$$v_{gn} = \frac{dm}{dio} = \frac{dp}{dio} = v_g$$

(8.54)

t is see" that

and is the same for all harmonics. From the above relations it is seen that some of the spatial harmonics (approximately one-half) have group velocities that are directed in opposite directions. This made use of in the backward-wave tube oscillators. A backward wave, or reverse wave, is often used to refer to

oppositely directed phase and group velocities. The voltage and current waves can, of course, also be expanded into an infinite set of spatial harmonics (Prob. 8.10).

Although a Bloch wave *can be* expanded into an infinite set of spatial harmonics, all the spatial harmonics must be simultaneously present in order that the total field may satisfy all the boundary conditions. The eigenvalue equation for β for a periodic structure always yields solutions $f\beta_n = f\beta + 2n\pi r/d$, in addition to the fundamental solution. These other possible solutions are clearly the propagation constants of the spatial harmonics. A complete k_d -*fid* diagram thus exhibits k_d as a periodic function of (id) ; that is, the *fid* curve is continued periodically outside the range $-77 \text{ s fid} \leq \text{TT}$. The slope of the line from the origin to any point on the curve still gives the phase velocity, and the slope of the tangent to the curve gives the group velocity, when multiplied by c .

STRUCTURES FOR TRAVELING-WAVE

Traveling-wave tubes require a structure capable of supporting an electromagnetic wave with a phase velocity equal to the velocity of the electron beam. Since the latter is usually much smaller than the velocity of light, the required structure is commonly referred to as a slow-wave structure. A common type of slow-wave circuit used in traveling-wave tubes is the helix. The helix is treated in the following two sections, and hence this section is restricted to a discussion of some of the other types of slow-wave periodic structures suitable for use in traveling-wave tubes.

A periodic slow-wave structure often used for the linear magnetron tube is the vane-type structure illustrated in Fig. 8.11. It consists essentially of a corrugated plane with thick teeth. It will be instructive to apply Floquet's theorem and carry out an analysis of this structure in order to illustrate the general techniques employed. Edge effects at $* = \pm a/2$ will be neglected for simplicity; i.e., we shall treat the structure as being infinitely wide. If a is large compared with the spacing b , and this in turn is small compared with A_0 , the edge effects will not produce a significant

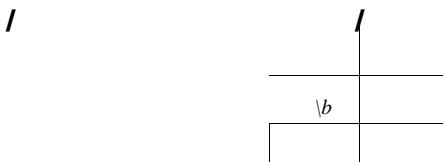


FIGURE 8.11
Vane-type, or corrugated-plane, periodic structure.

change in the characteristics of the ideal structure. For use in a man a strong axial electric field is required, and hence we shall exam' ^ possibility of having TM- or E-type modes.

For TM modes having no variation with x , the field component be expressed in terms of the single magnetic field component H present. We have

$$\mathbf{V} \times \mathbf{H} = -a_x \times Vff_x = jioe_0 E$$

and so

$$= 0 \quad bz \quad (8.55_Q,$$

$$E_z = i \cdot \hat{k}_0 \frac{\wedge}{9y} \quad (8.556)$$

The field H_z is a solution of

$$a^3 ; >^2 \quad (8.56)$$

According to Floquet's theorem, the field H_z can be expressed in the form $e^{jPz} l(y, z)$, where $l(y, z)$ is periodic in z with a period d . Hence we shall assume that

$$n_x = L f_{zz}(y) e^{-jPz}$$

where $f_{zz} = f_z + 2mr/d$ and the $f_{zz}(y)$ are functions of y to be determined. The substitution of this series into (8.56) shows that the $f_{zz}(y)$ are solutions of

$$d^2 f_{zz}(y) - 02 - * \cdot \frac{1}{2} f_{zz}(y) = 0 \quad (8.57)$$

Above the corrugations, i.e., in the region $c < y < b$, we must choose the i so that E_z will vanish on the perfectly conducting wall at $y = b$. Thus ^ require $df_z/dy = 0$ at $y = b$. Since solutions to (8.57) are $\sinh h_n y$ $\cosh h_n y$, where $h_n = (\sqrt{j} - k^2)^{1/2}$, we choose

$$f_{zz}(y) = n \cosh k_n(b-y)$$

where n is a constant. At $y = b$, this function has a zero derivative ^ the fields H_x and E_z in the region above the corrugations, we now n

$$H_x = Z \cdot n \cosh h_n(b-y) e^{-n(b-y)} \quad /< 580 I$$

$$K-J-T \quad \mathcal{L} \quad a_n h_n \sinh h_n(b-y) e^{-n(b-y)} \quad * .58 &$$

upon using (8.556).

As a next step we must obtain a suitable expansion for H_x in each corrugation, or slot. If $H_t(y, z)$ is the field in the slot extending from $* = 0$ to $z = s$ and for $0 < y < c$, then the field in the n th slot beginning at $z = nd$ will be $e^{-J\ell(nd)}H_t(y, z - nd)$ according to Floquet's theorem. Therefore we need to concentrate on one slot only. We must determine H_x so that E_v will vanish at $z = 0$ and s and also so that E_z will vanish at $y = 0$. A suitable expansion to use is

$$\begin{array}{ccc} * & & m \ 772 \\ m \ o & & \cdot s \end{array}$$

since $\langle /[\cos(mvz/s)l]/dz \rangle$ vanishes at $z = 0$ and s . If this expansion is substituted into (8.56), we find that

$$\frac{d^s g_{mrr}(y)}{dy^s} = g_{mrr}(y) = 0$$

Normally, $s \ll A_0$, so that $m/s \gg 1$, for $m \neq 0$. Thus appropriate solutions that have a zero derivative at $y = 0$ are

$$Sm(y) = 0 \cosh y$$

where $l_m = [(m/s)^2 - k^2]^{1/2}$ and m is a constant. For $m = 0$, the solution is $g_0(y) = f t, \cos fej^y$, and this part of the solution corresponds to a TEM standing wave in the slot. This mode has $E_v = 0$. In the first slot we can thus write

$$H_r = \begin{cases} E_{m=0} b_{m=0} \cosh l_m y \cos \frac{m-z}{s} & * \\ E_{m \neq 0} b_{m \neq 0} \cosh l_m y \cos \frac{m-z}{s} & m-z \end{cases} \quad (8.59a)$$

$$K_{m=0} = \frac{E_{m=0} b_{m=0} \sinh l_m y \cos \frac{m-z}{s}}{E_{m \neq 0} b_{m \neq 0}} \quad (8.59b)$$

The final step in the analysis is to determine the expansion coefficients $a_{m=0}$ and $b_{m=0}$ by imposing boundary conditions at the plane $y = c$ separating the two regions. We require the tangential electric and magnetic fields to be continuous across the gap $y = c$, $0 < z < s$. In addition, we require the tangential electric field to vanish on the upper faces of the teeth, i.e., at $y = c$ for $s \leq z < d$, or $nd + s \leq z < (n+1)d$ in general. Using (8.58) and (8.59), we see that the boundary conditions require

$$E_{m=0} a_{m=0} \cosh h_{m=0} (b_{m=0} - c) = E_{m=0} K_{m=0} \cosh l_m c \cos \frac{m-z}{s} \quad 0 < z < s \quad (8.60a)$$

$$E_{m=0} a_{m=0} \cosh h_{m=0} (b_{m=0} - c) = E_{m=0} K_{m=0} \cosh l_m c \cos \frac{m-z}{s} \quad 0 < z < s$$

$n = - =$

$$\begin{aligned} & - E_{m=0} a_{m=0} \cosh h_{m=0} (b_{m=0} - c) = E_{m=0} K_{m=0} \cosh l_m c \cos \frac{m-z}{s} \quad 0 < z < s \\ & > 0 \quad s \leq z < d \end{aligned} \quad (g 606)$$

If we multiply (8.60b) by $\sinh h_{rr}(b - c)$, we obtain

$$\int \sinh h_{rr}(b - c) e^{j2\pi Tr/d} dz$$

$$\sum_{m=0}^{\infty} K_m^* b_m \sinh h_{rr} \cos^{m-2} dz$$

f8.60c»

Now the coefficients in a Fourier series are uniquely determined only by the function which the series is to represent is specified for the co-interval over which the series is orthogonal. The functions e^{-iz} are orthogonal over the range 0 to d , and thus, since the left-hand side of (8.60c) is specified for all z over one period, we can obtain unique expressions for the a_m in terms of the b_m from (8.60c). Note that this is not true for (8.60a), which holds only in the region $0 < z < s$. If we multiply (8.60c) on both sides by e^{-irT_2} and integrate from 0 to d , we obtain (r is an integer)

$$\int da_r h_r \sinh h_r (b - c) dz$$

$$\sum_{m=0}^{\infty} \sinh h_{rr} (\frac{a_m}{r} - \frac{a_{m+1}}{r}) \cos^{r-m-2} dz$$

$$= \sum_{m=0}^{\infty} b_m j_m \sinh h_{rr} \left[\frac{j(p_m + 2irr/d)[(-l)^m e^{jmr} - 1]}{(li_m + 2Tr/d)^2 - (imr/s)^2} \right] \quad (8.61)$$

since $\int_0^d V_j^2 dz = 0$ for $n \neq r$ and equals d for $n = r$. The above represents an infinite set of equations, i.e., one for each value of r .

Although (8.60a) is not a unique equation for the a_m , it does specify the a_m uniquely in terms of the b_m . The a_m have already been expressed in terms of the b_m ; so we may regard them as known. Multiplying (8.60a) by $\cos(rT_2/s)$ and integrating from 0 to s give

$$\int_0^s a_m \cosh h_{rr} (b - c) I e^{-Jli_m z} \cos^{r-m-2} dz$$

$$= \sum_{m=0}^{\infty} a_m \cosh h_{rr} (b - c) \left[\frac{-3 - [1 - (-1)^m] e^{-M + *}}{2n_l r / df - (rW)} \right] \quad (8.62)$$

0. Th⁸ IS

where the Neumann factor $a_0 = 1$ for $r = 0$ and equals 2 for $r = 1, 2, \dots$. Equations (8.62) constitute two linear systems of equations involving the

If the solutions for the a_{mn} , as given by (8.61) are substituted into (8.62), the result is a homogeneous set of equations for the b_m . For a nontrivial solution, the determinant of this homogeneous set of equations must vanish. Setting the determinant equal to zero yields the eigenvalue equation for f_i . However, the sets of equations are of infinite order, so that, in practice, an exact solution is not possible. Therefore we shall find only a first approximation to the exact eigenvalue equation.

If the slot spacing s is small compared with A_{mn} , it seems reasonable to expect that the field in the slot can be approximated by the TEM standing-wave field alone. Thus we shall take all b_m except b_0 , equal to zero. If we lump all the constants in (8.61) together and replace r by n , the equation is of the form

$$\sum_{m=0}^{\infty} b_m R_{mn} = 0, \quad n = 0, \pm 1, \dots \quad (8.63a)$$

Likewise, (8.62) is an equation of the form

$$b_{mn} = Z_{mn} T_{mn}, \quad m = 0, 1, 2, \dots \quad (8.63b)$$

Replacing a_n by (8.63a) gives

$$b_m = L \sum_{n=0}^{\infty} b_m R_{mn} T_{mn}, \quad m = 0, 1, 2, \dots \quad (8.64)$$

The determinant of this infinite set of homogeneous equations, when equated to zero, gives the exact eigenvalue equation for f_i . When we take all b_{ni} except b_0 , equal to zero, we obtain instead

$$\begin{aligned} b_0 &= L b_0 R_{l0} T_{n0} \\ n &= -x \\ 1 - \frac{L}{Z} b_0 T_{n0} &= 0 \end{aligned} \quad (8.65)$$

for a first approximation to the eigenvalue equation. Now if ϵ are all the constants in (8.61), multiplying b_0 when the equation is solved for a_{mn} , and with r replaced by n . Likewise, T_{n0} is the constant relating b_0 to the a_{mn} in (8.62). When these values for R_{l0} and T_{n0} are substituted into (8.65), we obtain

$$\frac{1}{k_0 d \tan k_0 c} \frac{1}{H} \frac{U \ln \left(\frac{s}{2d} \right)^2}{P_{mn} s / 2} = \frac{1}{h d \tanh h} (b - c) \quad (8.66)$$

For slow waves β is much larger than k_n , and hence β , can be replaced by β , in this equation with negligible error. In this case the right-hand side is not dependent on k_0 . By evaluating the right-hand side for a range of assumed values for f_i , the corresponding value of k_a may be found by

**FIGURE 8.12**

Simplified equivalent circuit for the vane structure of Fig. 8.11.

solving (8.66). The numerical work is straightforward. Typical $r^A u_{ts}^A$ given by Hutter.

A reasonably accurate description of the dispersion curve relating $f t$ k_0 may also be obtained from a simple transmission-line analysis. The region above the corrugations is essentially a parallel-plate transmission line (strip line) with a characteristic impedance $Z_s = Z_0(6 - c)$ per unit width. The slots are short-circuited transmission-line stubs connected in series with the main line at periodic intervals d . The stubs present a reactance

$$jX = jZ_0 s \tan f c_0 c$$

to the main line. The equivalent circuit of the structure is therefore of the form shown in Fig. 8.12. This periodic circuit may be analyzed in the same way that the capacitively loaded transmission line was. It is readily found that the eigenvalue equation for p is

$$\begin{aligned} \cos f i d &= \cos k_a d - \frac{jX}{2Z_s} \sin k_n d \\ &= \cos k_Q d - \tan A_Q c \sin k_Q d \end{aligned} \quad (8.67)$$

This equation is quite accurate as long as $s \ll d$ and also much smaller than A_Q .

For frequencies such that $0 < k_Q c < ir/2$, the loading is inductive, for $77/2 < k_Q c < v$, it is capacitive, etc. A typical dispersion curve case $s = 2(6 - c)$ and $d = 0.83c$ is given in Fig. 8.13. For these the phase velocity is reduced by a factor of about 3 only. Much reduction factors are obtained by making $s/(b - c)$ larger so as to the normalized characteristic impedance of the stubs. For comparison results from (8.66) for $d = 2s$ are plotted also (broken curve), very accurately of (8.67). The first cutoff occurs approximately when a_{js0} reduce become resonant, i.e., for $k_Q c = ir/2$. Increasing c will therefore

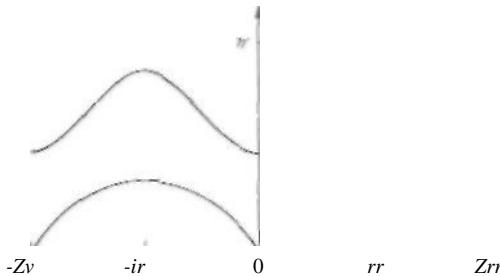


FIGURE 8.13
k₀d-lid diagram for a stub-loaded transmission line.

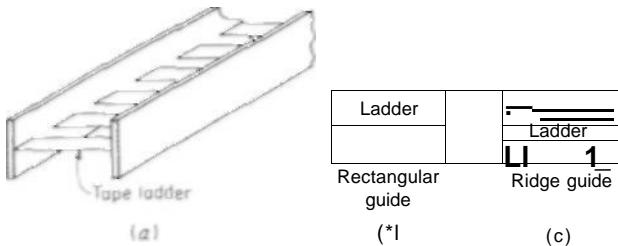
the phase velocity in the first passband, since (i_d will equal — for a smaller value of k_0 .

The foregoing analysis is typical for periodic structures that cannot be represented by simple transmission-line circuits. The essential steps to be followed are summarized:

1. Obtain suitable field expansions in each region of the periodic structure. This involves solution of the Helmholtz equation and the use of Floquet's theorem.
2. Impose appropriate boundary conditions on the fields at all common boundaries separating the different regions. In general, it will be found that both E and H modes may be required in order to satisfy the boundary conditions.
3. By Fourier analysis convert the boundary conditions into algebraic equations for the amplitude constants.
4. The system of algebraic equations can be written in the form of a homogeneous set of equations. Equating the determinant to zero gives the eigenvalue equation for (3) . Since the equations are usually infinite in order, some assumption must be made as regards the number of nonzero amplitude constants that will be *chosen*. Equating the higher-order amplitude constants to zero results in an approximate eigenvalue equation.

c Structures for Millimeter-Wave Traveling-Wave

At millimeter wavelengths a helix has too small a diameter to be a useful slow-wave structure. Various forms of tape ladder lines, interdigital tape lines, and meander tape lines are preferred. Illustrations of these structures are given in Fig. 8.14. A discussion of these structures together with typical



u r £]
(trf) (*)

FIGURE 8.14

(a-e) Tape ladder lines; (d) interdigital tape line; (e) meander tape line.

$k_0 f_i$ curves, is given in a paper by Harvey,^t which also provides references to the original analysis of these structures.

The two structures shown in Figs. 8.14 a¹ and e are complementary, or dual, structures. That is, the meander line is obtained by interchanging the open region and the conducting region in the interdigital line. For a complementary structure of this type, we can show that the field is also a dual solution and hence both structures have exactly the same $k_Q P$ dispersion curve. A detailed discussion of the dual properties is given below.

Let the interdigital line be located in the xy plane as in Fig. 8.15. Let us consider a mode of propagation having an electric field for which the transverse U and V components of field E , is an even function of x ; that is, the same on the upper and lower sides of the structure. Since $V \cdot E = 0$, we now have $dE/dz = -V \cdot E$, and hence $BE/9z$ is an even function of z . and E , must then be an odd function of z . From the curl $\nabla \times E = -j\omega H$, we can readily conclude that H , must be an odd function of z . The field structure is shown in Fig. 8.15a. The field E will vanish on the conducting surface, and is an odd function of z , it must vanish on the open part of the conductor surface, $|H|$ will equal one-half the total current density across the conductor surface, since the total change in $|H|$ across the conductor must be zero.

A dual field E', H' given by

$$E' = \pm Z_0 H \quad H' = +7_0 E$$

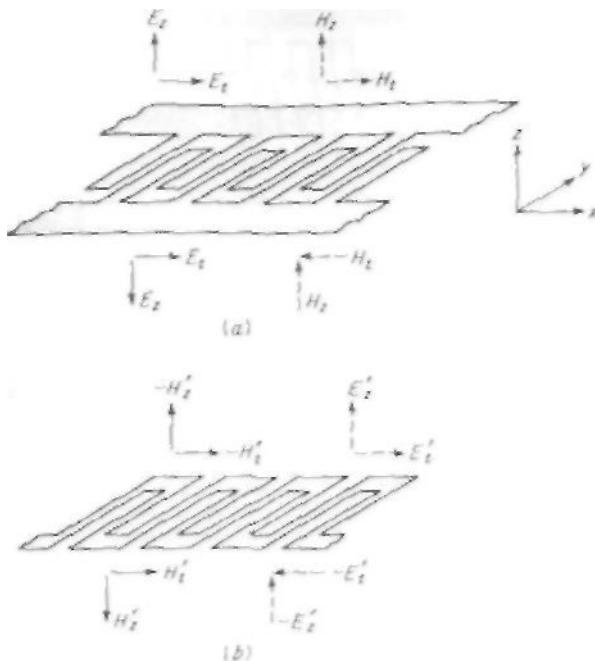


FIGURE 8.15
Illustration of dual properties of interdigital and meander lines.

is easily shown to satisfy Maxwell's equations

$$\nabla \times \mathbf{E}' = -j\omega \epsilon_0 \mathbf{H}' \quad \nabla \times \mathbf{H}' = \mu_0 \epsilon_0 \mathbf{E}'$$

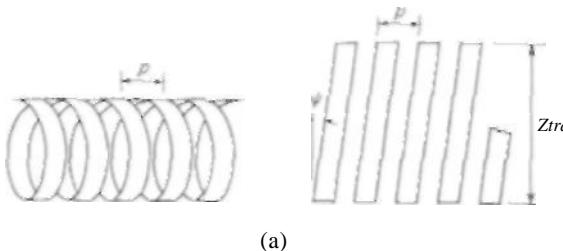
if the field \mathbf{E}, \mathbf{H} does. The dual field is a solution to the meander-line problem (Fig. 8.15b), provided we choose the dual solution

$$\mathbf{E}' = Z_0 \mathbf{H} \quad \mathbf{H}' = -Y_0 \mathbf{E} \quad (8.68a)$$

above the meander-line plane and the solution

$$\mathbf{E}' = -Z_0 \mathbf{H} \quad \mathbf{H} = Y_0 \mathbf{E} \quad (8.68b)$$

below the structure. In both regions the primed fields satisfy Maxwell's equations. The field \mathbf{E}' , will vanish on the conducting portions of the meander line since the field \mathbf{H}' , was zero in the open regions of the interdigital line. Similarly, \mathbf{H}' , vanishes over the open regions of the meander-line plane since the field \mathbf{E}' , was zero on the conducting surfaces of the interdigital line. All boundary conditions being satisfied, the solution is complete. It may now be concluded that both structures must have the same k_0-fS dispersion curve. It should be noted, however, that duality applies only if the two structures are exact complements; i.e., superimposing the two structures must result in the whole xy plane being a single conducting sheet. The sides of the interdigital line must therefore extend to $y = \pm \infty$, and the line must be infinitely long. However, in practice, the field is



-7

FIGURE 8.16
(a) A tape helix; (b) sheath

confined to the vicinity near the cuts, so that the sides do not have to extend much beyond the toothed region before they can be terminated with negligible disturbance of the field. The duality principle used above is often referred to as Babinet's principle.

8.10 SHEATH HELIX

The sheath helix is an approximate model of a tape helix. The tape helix, illustrated in Fig. 8.16a, consists of a thin ribbon, or tape, wound into a helical structure. The pitch is denoted by p , and the pitch angle by i/π . If the spacing between turns and the ribbon width are made to approach zero, the resultant structure becomes electrically smooth. At the boundary surface $r = a$, the boundary conditions for the electric field may be approximated by the conditions that the conductivity in the direction along the tape (direction of current) is infinite, whereas that in the direction perpendicular to the tape is zero. The use of these boundary conditions permits a solution for the electromagnetic field guided by the helix to be obtained with relative ease. This anisotropic conducting cylinder model of a tape helix is called a sheath helix, illustrated in Fig. 8.16b. The field solution, derived below, will show that the sheath helix supports a slow wave with a phase $v_p = c \sin i/\pi$. The wave may be considered to propagate along the conductor with a velocity c , and hence progresses along the axial a with a phase velocity $c \sin i/\pi$. The sheath-helix model is valid at low frequencies, where p is much smaller than A_0 . At higher frequencies a more realistic model must be used, and the existence of spatial harmonics becomes apparent, as shown in Sec. 8.11.

The field solution for the helix consists of both E and H fields. Along the direction of the tape, the tangential electric field must vanish.

conductivity in this direction is taken as infinite: thus

$$fi^! \cos ili + E_{il} \sin tli = E_{lb} \cos tji + E_{z2} \sin IJI = 0 \quad (8.69a)$$

where the subscripts 1 and 2 refer to the field components in the two regions $r < a$ and $r > a$. The component of electric field on the cylindrical surface $r = a$ and perpendicular to the tape must be continuous since the conductivity is taken as zero in this direction. Hence

$$E_{zl} \cos ili - E_{xi} \sin iJ/ = E_{xi} \cos iJ/ - E_{iU2} \sin iJ/ \quad (8.69e)$$

The component of H tangent to the tape must also be continuous since no current flows perpendicular to the tape; so a third boundary condition is

$$H_{zl} \sin 4i + H^j \cos iJ/ = H_{.2} \sin iJ/ + r_{.,2} \cos iJ/ \quad (8.69r)$$

Expansions for the E and H modes in the two regions $r \leq a$ may be obtained in terms of the axial field components E_z and H_z , as shown in Sees. 3.7 and 3.18. The axial fields $E_z = e,(-rf) \gg L^A$, $H_z = hAr,4) e$ are solutions of

Since we anticipate slow-wave solutions for which $fi^2 > kf_y$ the solutions involve Bessel functions with imaginary arguments, that is, $J_{\nu}(ryftf, - ii^l)$ and $Y^{\nu}ryk^2, - p^2$. In place of these functions, the modified Bessel functions $I_{\nu}(ryP^2 - k^2)$, $K_{\nu}(ryP^2 - /ef)$ are more convenient to use. These functions are related to the J_{ν} and Y_{ν} functions as follows:

$$I_{\nu}(*) = .//''•-/, ,(./*) \quad (8.70a)$$

$$K_{\nu}(x) = jj'''[J_{\nu}(jx) + jYJ jx] \quad (8.706)$$

For small values of x , the K_{ν} functions approach infinity in a logarithmic fashion, and hence only the I_{ν} functions are used in the region $r < a$. For r large, the asymptotic forms

$$h(x) \sim ihx^{ex} \quad (8.710)$$

$$K_{\nu}(x) \sim E^{\ell} \sim \quad (8.716)$$

are valid. Since we require a field that decays for large r , only the functions K_{ν} are employed in the region $r > a$. Suitable expansions for e_z and h_z in

the two regions are now seen to be

$$\begin{aligned}
 Z &= a_{,,} e^{-J''^A} I_n(hr) & r < a \\
 * & \\
 Z &= b_{,,} e^{-J''^B} K_n(hr) & r > a \\
 n = -x & \\
 h &= c_{,,} e^{-J''^C} I_n(hr) & r < a \\
 & \\
 Z &= d_n e^{-J''^D} K_n(hr) & r > a
 \end{aligned}$$

where $h = \{f\bar{r}^* - tg\}^{1/2}$ and $a_{,,}$, $b_{,,}$, $c_{,,}$, and $d_{,,}$ are unknown amplitude constants.

For the sheath-helix model it is possible to find a solution for a field that satisfies the boundary conditions (8.69) for each integer n . We are primarily interested in the solution $n = 0$, which has circular symmetry. If we make use of (3.67), (3.68), and (3.72), together with the relations

$$\begin{aligned}
 \Delta \mathbf{T} &= / \mathbf{i} (\mathbf{i} \mathbf{r}) & dK_0(hr) \\
 \mathbf{W} &= \sim \mathbf{K}(\mathbf{i} \mathbf{r})
 \end{aligned}$$

we find that the field in the two regions can be expressed as follows:

For $r < a$,

$$\begin{aligned}
 E_z &= o_0 /_0 (Ar) e^{-J''^A} & E_r &= \sim J \Delta a_0 I_0(hr) e^{-J''^A} & E_{,,} &= \sim c_0 /_0 (Ar) e^{-J''^A} \\
 H_z &= c_0 I_0(hr) e^{-J''^A} & H_r &= -c_0 Z^A r e^{-J''^A} & H^* &= \frac{!''^B}{-h} e^{-J''^B} f_t r e^{-J''^B} \\
 & & & & & \boxed{(8.72o)}
 \end{aligned}$$

For $r > a$,

$$\begin{aligned}
 E_t &= b_0 K_0(hr) e^{-J''^B} & E_r &= \sim J \pm b_0 K_0(hr) e^{-J''^B} & E_{,,} &= \sim J \Delta d_0 K_0(hr) e^{-J''^B} \\
 H_z &= d_0 K_0(hr) e^{-J''^B} & H_r &= \frac{-J A_{dK_0}(hr)}{ft} e^{-J''^B} & H_{,,} &= \sim J \Delta b_0 K_0(hr) e^{-J''^B} \\
 & & & & & \boxed{(8.72ft)}
 \end{aligned}$$

for the $rc = 0$ mode.

If the above expressions for the fields are substituted into the boundary conditions (8.69), the result is four homogeneous equations if the constants a_0 , b_0 , c_0 , and d_0 . A nontrivial solution exists if the determinant vanishes. Equating the determinant to zero results in the equation

1.0-1

0.5-

 $\sin 10^\circ$

0

0.5

1.0

FIGURE 8.17

Phase-velocity reduction factor for a sheath helix with pitch angle $i// = 10^\circ$.

eigenvalue equation for β_3 , which is

$$\frac{K_I(ha)I_{,I}(ha)}{K_O(ha)I_{,,O}(ha)} = \frac{\tan^2 i//}{(k_0a)^2} \quad (8.73)$$

For ha greater than 10, the ratio $K_I(ha)/K_O(ha)$ rapidly approaches unity. In this region (8.73) gives $h = k_0 \cot i//$, from which we obtain

$$n = (k^2 \theta + h^2)^{1/2} = ft_0 \csc i// \quad (8.74)$$

The resultant phase velocity v_p is

$$v_p = J = jc = c \sin i// \quad (8.75)$$

and is reduced by the factor $\sin i//$. A plot of v_p/c as a function of k_0a is given in Fig. 8.17 as determined by the solution of (8.73) with $i//$ equal to 10° . For k_0a greater than 0.25, the phase velocity is well approximated by (8.75). In the frequency range where $v_p = c \sin i//$, the group velocity v_g is also equal to $c \sin i//$, and there is no frequency dispersion.

SOME GENERAL PROPERTIES OF A HELIX

The tape helix consists of a thin ribbon of metal wound into a helical structure, as shown in Fig. 8.16a. A helix may also be constructed by the use of a round wire. The parameters describing the helix are the pitch p , or turn-to-turn spacing, the diameter $2a$, and the pitch angle $i//$. These parameters are given in Fig. 8.16a, which shows a developed view of a tape helix.

The helix is a periodic structure with respect to translation by a distance p along the axis and also with respect to rotation through an arbitrary angle θ , followed by a translation $pt\pi/2 - rr$ along the axial direction. In other words, an infinitely long helix translated along the z axis by a distance p or rotated by an angle B and then shifted by a distance $p8/2-n$ along z will coincide with itself.

The above periodic, or symmetry, properties of the helix 1 restrictions on the nature of the field solutions. If $E_r(r, \theta, z) = E_r(r) e^{j\theta} e^{jz/p}$ are cylindrical coordinates, is a solution for the electric field and $E_r(r, \theta, z + p/2\pi r)$ multiplied by a propagation factor $e^{-jkz/p}$ is another solution, since the point $r, \theta + \pi, z + pd/2\pi r$ is indistinguishable from the point r, θ, z . The solution $E_r(r, \theta, z)$ must be periodic in z apart from a propagation factor $e^{-jkz/p}$, must also be periodic in θ with period p . Hence E_r may be expanded in the double Fourier series

$$E_r(r, \theta, z) = \sum_{m,n} E_{mn}(r) e^{-jkz/p} J_m(\theta) e^{jn\phi} \quad (8.76)$$

where $E_{mn}(r)$ are vector functions of r corresponding to the usual amplitude constants in a Fourier series. The relationship between translation and rotation noted above requires that $e^{j\theta} E_r(r, \theta, z)$ does not change when d are replaced by $\theta + \pi, z + pB/2\pi r$. Thus, in (8.76), we require

$$\begin{aligned} e^{-jkz/p} &= g \cdot J_m(\theta) e^{jn\phi} \\ &= g \cdot J_m(\theta) e^{jn\phi} \end{aligned}$$

This condition will hold only if $m = -n$. Consequently, for a helix, the double Fourier series expansion for the electric field reduces to a single series of the form

$$E_r(r, \theta, z) = \sum_n E_n(r) e^{-jkz/p} J_n(\theta) e^{jn\phi} \quad (8.77)$$

The solution for a helix proceeds by expanding the field in the two regions $r < a$ and $r > a$ into an infinite series of E and H modes expressed in cylindrical coordinates. The boundary conditions at $r = a$ will couple the E and H modes together, so that pure E or H modes cannot exist independently. For the n th term in (8.77), the radial dependence in the region $r > a$ will accord with the modified Bessel function of the second kind, that is, $K_n(h_n r)$, of order n and with an argument

$$f(r) = \sum_{n=1}^{\infty} K_n(h_n r) e^{-jkz/p}$$

The K_n functions are asymptotic to $U/2h_n r$ for $r \gg h_n$. The field will decay exponentially as long as all h_n are real, i.e., $(0 + 2\pi n/p)^2$ greater than k^2 . When the field decays exponentially, it corresponds to a surface-wave mode guided by the helix.

At a given frequency only certain discrete values of P , λ mode possible solutions. For each value of P , corresponding to a particular mode of propagation, the field is given by a Fourier series of the form

$$E_m(r, \theta, z) = \sum_n E_{mn}(r) J_n(\theta) e^{-jkz/p} \quad (8.78)$$

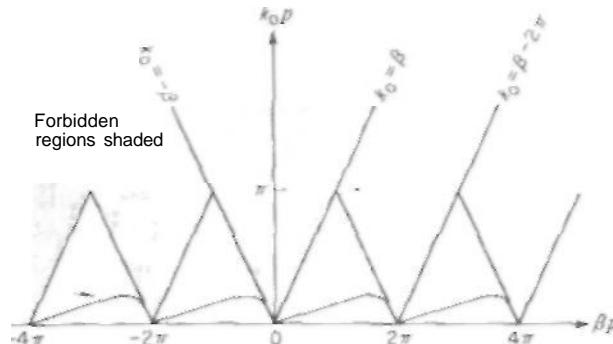


FIGURE 8.18
Illustration of allowed and forbidden regions in the $k_0 p$ - β_p diagram for a helix.

Each term in this expansion is called a spatial harmonic, and has a propagation phase constant $/3_{..} + 2mr/p$. On a $k^p p$ - β_p diagram the region above the lines $k_Q = \pm 3$ is a forbidden region, as shown in Fig. 8.18, since it corresponds to a situation where $h_0 = (f_1^z - k_0 Y^z)$ is imaginary and the $n = 0$ spatial harmonic does not decay in the radial direction. Since we also require $|/3 + 2mr/p|$ to be greater than k_0 , all possible allowed values of $/3$, corresponding to bound surface-wave modes, are further restricted to lie in the unshaded triangular regions in Fig. 8.18. The boundaries of these regions are marked by the lines

$${}^* \mathbf{o} = \pm |P \pm \frac{2\pi n}{\beta_p}|$$

where n is an integer. In the forbidden regions the propagation constant turns out to be complex rather than pure real, a feature which is different from that of a normal cutoff mode.

A first approximation to the solution for a tape helix is obtained by assuming that the current is directed along the direction of the tape only, is uniform across the width of the tape, and has a propagation factor $e^{-\alpha z}$. A typical $k_0 f_1$ - f_1 curve obtained on this basis and with $\phi = 10^\circ$ is shown in Fig. 8.18. For further results the paper by Sensiper or the book by Watkins may be consulted.t

INTRODUCTION TO MICROWAVE FILTERS

The ideal filter network is a network that provides perfect transmission for all frequencies in certain passband regions and infinite attenuation in the

tS. Sensiper, Electromagnetic Wave Propagation on Helical Structures, *Proc. IRE*, vol. 43, pp. 149-161, February, 1955.

D. A. Watkins, "Topics in Electromagnetic Theory," John Wiley & Sons, Inc. New York, 1958.

stopband regions. Such ideal characteristics cannot be obtained. The goal of filter design is to approximate the ideal requirements to an acceptable tolerance. Filters are used in all frequency ranges to^{*} w_{lt} in an nearly perfect transmission as possible for signals falling within the passband frequency ranges, together with rejection of those frequencies outside the desired frequency bands. Filters fall into ^{^ma^n} categories, namely, (1) low-pass filters that transmit all signals between frequency and some upper limit ω_0 and attenuate all frequencies above the cutoff value ω_c , (2) high-pass filters that pass all frequencies above the cutoff value ω_h and reject all frequencies below ω_t , and (3) bandpass filters that pass all frequencies in a range ω_1 to ω_2 and reject frequencies outside this range. The complement to the bandpass filter, i.e., the band-reject filter, which attenuates frequencies in the range ω_1 to ω_2 , is also of interest in certain applications.

At low frequencies the "building blocks" for filters are ideal inductors and capacitors. These elements have very simple frequency characteristic* and a very general and complete synthesis procedure has been developed for the design of filters utilizing them. It is possible to synthesize directly filters with a wide variety of prescribed frequency characteristics. The filter design problem at microwave frequencies where distributed parameter elements must be used is much more complicated, and no complete theory or synthesis procedure exists. The complex frequency behavior of microwave circuit elements makes it virtually impossible to develop a general and complete synthesis procedure. However, in spite of these added complications at microwave frequencies, a number of useful techniques have been developed for the design of microwave filters. The case of narrowband filters is particularly straightforward since many microwave elements will have frequency characteristics essentially like those of an ideal inductive or capacitive reactance over a limited frequency range. In this case a low-frequency prototype filter may be used as a model. The microwave filter is reahzed by replacing all inductors and capacitors by suitable microwave circuit elements that have similar frequency characteristics over the frequency range of interest. For this reason a good deal of the effort in microwave filter design has been based directly on the application of low-frequency synthesis techniques.

There are essentially two low-frequency filter-synthesis methods common use. These are referred to as the image-parameter method and variations thereof, such as the constant- β and *in*-derived filter design insertion-loss method. The image-parameter method provides a method having the required passband and stopband characteristics, but does not specify the exact frequency characteristics over each region. The realizable loss method begins with a complete specification of a physical frequency characteristic, and from this a suitable filter network is synthesized. The image-parameter method suffers from the snort. In order to good deal of cut-and-try procedures must often be resorted

obtain an acceptable overall frequency characteristic. For this reason the insertion-loss method is preferable and is the only method considered in detail in the following sections. The image-parameter method is only briefly outlined, in order to show its relationship to the properties of periodic networks as already discussed.

The labor involved in filter synthesis is largely obviated by the use of certain frequency transformations and element normalizations. These enable high-pass and band-pass filters operating over arbitrary frequency bands and between arbitrary resistive load terminations to be obtained from a basic low-pass filter design. The characteristics of any filter will, of course, be modified by the losses that are present in all physical network elements. To incorporate the effect of lossy elements into the synthesis procedure makes the synthesis theory a great deal more involved; so this is usually not done. At microwave frequencies losses can be kept reasonably small, to the extent that most filter designs based on the use of lossless elements do perform satisfactorily.

The aim of the following sections is to present the essential features of low-frequency filter synthesis, frequency transformations, normalized filter design, and the applications of these techniques to microwave filter design. A number of typical microwave filters are also discussed. An extensive account of all aspects of microwave filter design is beyond the scope of this text. However, a number of selected references are given where further details may be found.

8.13 IMAGE-PARAMETER METHOD OF FILTER DESIGN

Filters designed by the image-parameter method have many features in common with those of periodic structures. As noted in the previous sections, a cascade connection of lossless two-port networks behaves similar to a transmission line. For unsymmetrical networks two characteristic impedances $Z_B' = \pm Z + f$ occur, and each section has a propagation factor $e^{\pm j\phi}$. A periodic structure of this form has passband and stopband characteristics and is therefore a bandpass filter. However, the proper load termination to prevent reflections is Z_R and is complex when $f \neq 0$. Usually, a filter must operate between resistive load terminations, and it would not be possible to have matched input and output terminations in this case unless f were zero, i.e., unless symmetrical networks were used or unless matching sections were used at the input and output. For this reason the image-parameter method of filter design is based on considerations somewhat different from those which have been discussed for periodic structures.

Consider a single two-port network with parameters s_f , α , β , and γ . Let the output be terminated in a load Z_{12} , and let the input be terminated in a load Z_n , as in Fig. 8.19. For particular values of Z_1 , and Z_{12} , known as

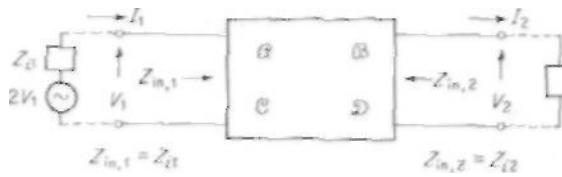


FIGURE 8.19
Image parameters for a two-port network.

the image impedances, the input impedance at port 1 equals Z_{it} and the port 2 equals Z_{it} . These impedances then provide matched terminate^{ife!!} the two-port network, and if they are real, they also provide a maximum power transfer when the generator has an internal impedance equal to the image impedance. The governing equations for the two-port network are

$$V_t = \frac{1}{K} + \frac{I_t}{SfZ_{it}} + \frac{I_t}{KZ_{it}}$$

and hence

$$\begin{aligned} V_t &= \frac{1}{K} + \frac{m_2}{SfZ_{it}} + \frac{3g}{KZ_{it}} \\ I_t &= \frac{1}{KZ_{it}} \end{aligned} \quad (8.79)$$

If we solve for V_2 and I_t in terms of V_t and I_t , we obtain

$$V_2 = \frac{1}{K}V_t - m_2 \quad I_t = -\frac{1}{K}V_t + SfZ_{it}$$

We thus have

$$7^{\text{th}} \quad \frac{1}{K}V_t - \frac{1}{K}9V_2 + SfI_t = \frac{1}{K}WZ_a + * \quad (8.80)$$

The requirement that $Z_{it} = Z_{it}$, and $Z_{it} = Z_{it}$ gives

$$Z_{it}(KZ_{it} + 0) = a/Z_{it} + \frac{1}{K} \quad Z_{it}(STZ_{it} + st) = \frac{1}{K}SfZ_{it} + *$$

A simultaneous solution of these equations gives

$$Z_{it} = \frac{1}{K} \quad (8.81a)$$

$$Z_{it} = \frac{1}{K} \quad (8.81b)$$

Also we find that $Z_{it} = 10/SS'Z_{it}$.

If a generator with internal impedance Z_n is connected at port 1, the output port 2 is terminated in a load Z_{it} , the voltage and current transfer ratios are readily found from the relations

$$I_2 = \frac{1}{K}WV_t + JTI_t = (\frac{1}{K}SfZ_a + \frac{1}{K})'i$$

where V_1 is the voltage across the network terminals at port 1 (the generator voltage is $2V_1$). Thus we find that

$$\sqrt{\frac{D}{A}} (\sqrt{A} - \sqrt{B}) \quad (8.82a)$$

$$\sim = \sqrt{\frac{D}{A}} (f\{tfy - JIm) \quad (8.82b)$$

In a similar manner the transfer constants from port 2 to port 1 are found to be [or from (8.82)]

$$\frac{V_1}{V_2} = \frac{y/F''}{\sqrt{A/A'}} (fas + 499) \quad (8.83a)$$

$$h = \frac{fgT}{\sqrt{A/A'}} (fas + JggW) \quad (8.8361)$$

The image propagation factor $e^{\sim y}$ is defined as

$$e^{\sim y} = io/9 > - 4\&W \quad (8.84a)$$

whence it is found that

$$e^{\sim y} = \frac{1}{IB} iw'9 > + vfe\%if \quad (8.846)$$

$$\text{and} \quad \cosh y = J\$@i \quad (8.84c)$$

$$\sinh y = J<M\%' \quad (8.84d)$$

The factor $(y/A/\%)^2$ is interpreted as an impedance transformation ratio and may be viewed as an ideal transformer of turns ratio $fof/\&$.

For a lossless network, sf and $\&$ are real and $.f$ and $\%$ are imaginary. In the passband of a filter, y is pure imaginary and equal to $j\beta$, and this occurs for $|jsf| < 1$, as (8.84c) shows. Also, in the passband, the image impedances are pure real, whereas in a stopband they are pure imaginary, as the following considerations show. In a passband, SS and W must be of the same sign, so that $M'i? = j<\%j'*? = -\$g<\&$ will make $\sinh y$ in (8.84a¹) pure imaginary, that is, $y = j(5)$. Thus, in (8.81), the quantity under the square root will be real and positive since $sfSt$ must be positive to give a real solution for $\cosh y$. Hence the image impedances are real in a passband.

If N two-port networks are connected in cascade and these have propagation constants y_n , $n = 1, 2, \dots, N$, and voltage transformation ratios

$$t = i / \sum_{n=1}^N T_n \quad T_1, \dots, T_N, \dots, T_N$$

and the output section is terminated in an impedance equal to its image impedance, the overall voltage transfer ratio is

$$\frac{YE}{V_{N-1}} = \frac{-T_x T_t}{T_N e^{-\frac{1}{2} \pi i \sim v_s \sim - y \cdot v}} = \frac{1}{nr..-r.} \quad (80-85)$$

provided also that the output image impedance of any one section is equal to the input image impedance of the adjacent section. With this filter network terminated in a load impedance Z_N equal to the image impedance of the output section, and with the generator at the input having an input impedance Z_{in} , the overall network is matched for maximum power transfer. The filter operates between impedance levels of Z_{in} and Z_N which provide an overall impedance-ratio change of amount

$$\frac{Z_{in}}{Z_{rl}} = \frac{Z_{in}}{Z_{rl}} = \frac{1}{nr^N} \quad (18.86)$$

If symmetrical two ports are used, $\langle / = y$ and $Z_{in} = Z_{rl}$, and both are equal to the Bloch-wave characteristic impedance Z'_{if} . For a symmetrical network no transformation or change in impedance level is obtained. The filter consisting of N symmetrical sections terminated in load impedances equal to the image impedance Z_N behaves exactly like an infinite periodic structure, with its characteristic passband and stopband features.

In the image-parameter method of filter design, the two-port parameters $s, t, i, \&, \&, \&$ are chosen to provide for the required passbands and stopbands. In addition, the image parameters are also chosen equal to the terminating impedances at the center of the passband. The shortcomings of the filter are now apparent, namely, the image impedances are functions of frequency and do not remain equal to the terminating impedances over the whole desired passband. This results in some loss in transmission (loss to mismatch) within the passband, an amount that cannot be prescribed determined before the filter has been designed. In addition, there is means available for controlling the rate at which the attenuation increases with frequency beyond the edges of the passband, apart from increasing the number of filter sections. Nevertheless, many useful microwave filters have been designed on this basis, f

tS. B. Cohn, chaps. 26 and 27 in Radio Research Laboratory Staff, "Very Low Frequency Techniques," vol. 2 McGraw-Hill Book Company, New York, 1947.

For a discussion of image-parameter methods at low frequencies, see *Communication Networks*, vol. 2, John Wiley & Sons, Inc., New York, 1930.

DESIGN BY INSERTION-LOSS METHOD

The power loss ratio of a network was defined in Sec. 5.14 as the available, or incident, power divided by the actual power delivered to the load; thus

$$\text{PLH} = \frac{1}{1 - r_r^*} - 737 \quad (8.87)$$

where T is the input reflection coefficient for a lossless network terminated in a resistive load impedance $Z_L = R_h$. The insertion loss, measured in decibels, is

$$L = 10 \log P_{LR} \quad (8.88)$$

when the terminating resistive load impedance equals the internal impedance of the generator at the input end. In general, the insertion loss is defined as the ratio of the power delivered to the load when connected directly to the generator to the power delivered when the filter is inserted.

The insertion-loss method of filter design begins by specifying the power loss ratio P_{LR} or the magnitude of the reflection coefficient $|T| \sim p$ as a function of w . A network that will give the desired power loss ratio is then synthesized. This procedure is seen to be essentially the same as was followed in the synthesis of quarter-wave transformers in Secs. 5.12 and 5.13. Indeed, the multisection quarter-wave transformer may be considered a particular type of bandpass filter. It must be kept in mind, however, that a completely arbitrary $H(w)$ as a function of w cannot be chosen since it may not correspond to a physical network. The restrictions to be imposed on $H(w)$ are known as the conditions for physical realizability, and some of these are discussed below.

For a passive network it is clear that the reflected power cannot exceed the incident power, and hence one restriction on $r(w)$ is

$$|r(w)| < 1 \quad (8.89)$$

If the normalized input impedance of the network is

$$Z(w) = R(io) + jX(co)$$

we have

$$r(w) = \frac{Z_{in} - 1}{Z_{in} + 1} = \frac{R(w) - 1 + jX(co)}{R(w) + 1 + jX(co)}$$

As shown in Sec. 4.4, $\sim R$ is an even function of 10 and X is an odd function of co . Hence

$$r(-01) = \frac{\mathbf{g(a)} - w\mathbf{X}(,,)}{R(w) + l-jX(M)} = r \wedge^{(10)}$$

and thus

$$[T(\ll)]l^* - r^*(\bullet) = ir^* - r(*>)r(-6\gg)$$

w and must therefore contain only even powers of at. Now any

W f c t i o n o f
impedance function (impedance of a network made up of resistors, and inductors) can be expressed as the ratio of two polynomials. Consequently, T can also be expressed as the ratio of two polynomials. It follows that (H/a) can then be expressed in the form

$$\frac{M(u^2)}{M(w^*)} = \frac{(a^2 - 1)^2 + X^2}{(R + if + X^2)} \quad (8.91)$$

where M and N are real and nonnegative polynomials in w². The power loss ratio can now be expressed as

$$P_{LR} = 1 - \frac{M(u^2)}{N(a^2)}, \quad (8.92)$$

The last result in (8.92) shows that $N(a^2)$ must be an even polynomial in a^2 since it equals $4/(w^2)$. Hence we write $N(u^2) = Q^2(w)$, which is clearly an even polynomial in w. If we denote $M(u)$ by the even polynomial $P(w)$ instead, we have

$$P_{LR} = 1 + \quad (8.93)$$

The conditions specified on P_{LR} up to this point are necessary conditions in order that the network may be physically realizable. It may be shown that the condition that the power loss ratio P_{LR} be expressible in the form (8.93) is also a sufficient condition for the network to be realizable. In subsequent sections we consider suitable forms for the polynomials P and Q and types of networks required to yield the corresponding power loss ratios.

8.15 SPECIFICATION OF POWER LOSS RATIO

There are virtually an unlimited number of different forms that could be specified for the power loss ratio and be realized as a physical component. However, many of these networks could be anticipated to be very complex and hence of little practical utility. The power loss ratios that have been found most useful for microwave filter design are those that give

mally flat passband response and those that give an equal-ripple, or Chebyshev, response in the passband. Such passband-response characteristics correspond to those of the binomial and Chebyshev multisection quarter-wave transformers discussed in Chap. 5. The maximally flat filter (commonly called a Butterworth filter) and the Chebyshev filter are described below for the low-pass case only. In a following section it is shown that high-pass and bandpass filter characteristics may be obtained from the low-pass filter response by suitable frequency transformations, or mappings.

ally Flat Filter Characteristic

The power loss ratio for a maximally flat low-pass filter is obtained by choosing the polynomial Q equal to unity and choosing $P(w^2)$ equal to $k^2(o_c/\omega)^{2N}$. Hence we have

$$P_{LK} = 1 + k^2 \left(\frac{\omega}{\omega_c} \right)^{2N} \quad (8.94)$$

The passband is the region from $\omega = 0$ to the cutoff value ω_c . The maximum value of P_{LR} in the passband is $1 + k^2$, and for this reason k^2 is called the passband tolerance. For $\omega > \omega_c$, the power loss ratio increases indefinitely at a rate dependent on the exponent $2N$, which in turn is related to the number of filter sections employed. A typical filter characteristic is illustrated in Fig. 8.20 for $N = 2$.

Chebyshev Filter

The power loss ratio for the equal-ripple, or Chebyshev, filter is chosen as

$$P_{LR} = 1 + k^2 T^{2N} \quad (8.95)$$

where $T_N(a)/(o_c)$ is the Chebyshev polynomial of degree N discussed in Sec.

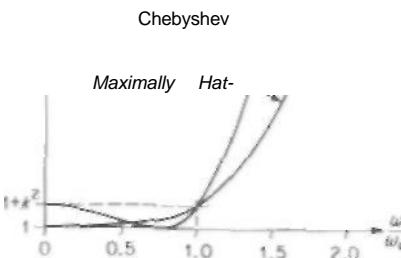


FIGURE 8.20
Low-pass-filter response for maximally flat and Chebyshev filters for $N = 2$.

5.13. Recall that

$$T_v = \cos(\omega_c t) = \cos(\omega_c t) \cos^{-1} \frac{1}{\sqrt{1 + k^2}}$$

and thus $T_N(o/\omega_c)$ oscillates between $+1$ for $\omega/\omega_c < 1$ and monotonically for ω/ω_c greater than unity. The power loss "margin" oscillates between 1 and $1 + k^2$ in the passband, equals $1 + k^2$ at $\omega = \omega_c$ frequency, and will increase monotonically for $\omega > \omega_c$. A typical curve is shown in Fig. 8.20 for $N = 2$. One particularly striking feature of the Chebyshev response curve compared with the maximally flat curve is the much greater rate of rise beyond the cutoff point. This means that the corresponding filter has a much sharper cutoff region separating the passband and stopband, which is usually a desired characteristic.

For ω_c large, the power loss ratio for the Chebyshev filter approaches

$$\text{LR}_{\text{TUT}} = \frac{k^2}{(1 + k^2)^{2N}} \quad (8.96)$$

Compared with the maximally flat response characteristic this is larger by a factor 2^{2N^2} . In fact, no other polynomial $P(co^2)$ yielding a passband tolerance of k^2 can yield a rate of increase of P_{LH} greater than that given by (8.96). Conversely, for a specified rate of increase in the power loss ratio beyond the cutoff frequency, the Chebyshev polynomial gives the smallest passband tolerance. In this sense the Chebyshev filter represents an optimum design. The proof is similar to that used in Sec. 5.14 to prove the optimum properties of the Chebyshev quarter-wave transformer.

When the power loss ratio is equal to $1 + k^2$, the magnitude of the reflection coefficient at the input is

$$P = \frac{k}{(1 + k^2)}$$

The input voltage standing-wave ratio is given by

$$S = \frac{1 + P}{1 - P} = \frac{(1 + k^2)^{1/2} + k}{(1 + k^2)^{1/2} - k}$$

If P is chosen as unity and Q is set equal to $k^2 T_u o / t$, a filter having Chebyshev behavior in the attenuation or stopband is obtained. It is also possible to choose P and Q so as to give Chebyshev both the passband and the stopband. The required network is usually too complex to be realized in a satisfactory manner with circuit elements.

SOME LOW-PASS-FILTER DESIGNS

8.16

The maximally flat and Chebyshev low-pass-filter power loss ratios discussed in the preceding section can be realized by means of a ladder network of capacitors and inductors in the form illustrated in Fig. 8.21. The load impedance is chosen equal to 1 Ω , and the generator impedance as R . The circuit in Fig. 8.21b is the dual of the circuit in Fig. 8.21a. Both circuits can be designed to give the same power loss ratio. For maximally flat or Chebyshev response in the passband, the ladder network is symmetric for an odd number of elements. This is also true for N even, in the case of the maximally flat filter. The element values are denoted by g_k and are the same in both circuits. However, the required generator impedance R' in the network of Fig. 8.21b is equal to $1/R$. For N odd both R and R' equal unity.

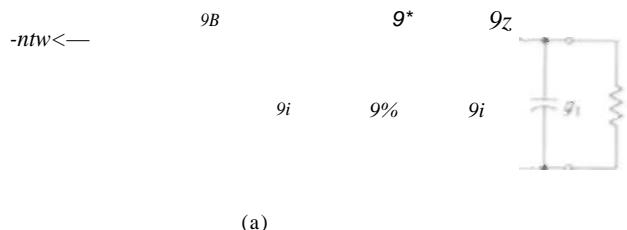
If we let Z_m be the input impedance at the plane aa in Fig. 8.21a, the reflection coefficient will be

$$\mathbf{r} = \frac{Z_m - R}{Z_m + R}$$

In terms of Z_{in} and R , the power loss ratio is readily computed to be

$$P.M = 1 + \frac{I^2 Z_{in} - \mathbf{J}^2 R^2}{2 \operatorname{ff}(Z_{in} + Z^*)^2} \quad (8.97)$$

At $t_0 = 0$, all capacitors appear as infinite impedances and all inductors as zero impedances, and hence $Z_{in} = 1$. For a maximally flat filter or a



(a)

—WP—l—npSG^—f—^555^—t—^5W^—

96 9* 9z

(A)

FIGURE 8.21
Low-pass ladder-prototype-filter networks.

Chebyshev filter with N odd, we must have unity power loss rat" 3, o, *
This requires that we choose the generator impedance equal to
Chebyshev filter with N even, we have $P_{IK} = 1 + k^2$ at $u > \frac{w}{R}$ or a
u, ^d hence

$$1 + A^2 = 1 + \frac{4}{R^2}$$

or $i? = 2k^2 + 1 - \frac{4}{R^2}(1 + k^2) = \frac{R^2 - 4k^2}{R^2}$ (8.98)

The required values of the elements g_k in the ladder network w,
obtained very readily by solving for Z_m and equating the power loss rat
given by (8.97) to the desired power loss ratio for N up to 3 or 4 As° ^
example, consider the case of $N = 2$ for the circuit in Fig. 8.21a. We'readT-
find that

$$Z_m = j\omega L + \frac{1}{1 + j\omega C}$$

where $C = g_x$ and $L = g_2$. Using (8.97) gives

$$(1 - R)^2 + \omega^2(L^2 + C^2 f_l^2 - 2LC) + *AL^2C^2$$

J \ R = 1
4-ft

To make P_{jR} equal unity at $\omega = 0$ and to obtain maximally flat response,
we must choose $R = 1$ and $L^2 + C^2 - 2LC = 0$. If we specify cutoff to
occur at $\omega = 1$ with a passband tolerance of k^2 , we also have

$$1 + k^2 = 1 + \frac{L^2 C^2}{1 + \omega^2}$$

or $LC = 2k$. hence we find $L = C = (2fc)^{1/2}$. For Chebyshev response we
equate $P, ^A$ to

$$1 + \frac{2\pi^2}{1 + \omega^2} = 1 + \frac{f_e}{f_c} - I^- I^+$$

in order to determine L and C .

For large values of N the above procedure is very laborious t^ ^n
out. In place of this direct substitution scheme general solutions a^ ^
worked out. In addition, tables of element values, i.e., values or
have been prepared by a number of people.t

tV. Belevitch, Chebyshev filters and Amplifier Networks. *Wireless Eng.*, vol. 29, pp. 10*, April, 1952.

+L. Weinberg, Network Design by Use of Modern Synthesis Techniques an *Natl. Electron. Conf.*, vol. 32, 1956.

S. Cohn. Direct Coupled Resonator Filters, *Proc. IRE*, vol. 45, pp. 187

See also "The Microwave Engineers' Handbook." Horizon House, Inc. •

For the maximally flat network with a power loss ratio

$$/\!>_L R = 1 + "^2 N \quad (8-99)$$

the element values are given by

$$R = 1 \quad (8.100a)$$

$$g_k = \frac{2k - 1}{2sm} \quad k = 1, 2, \dots, N \quad (8.1006)$$

where g_k is the value either of inductance in henries or of capacitance in farads. Each end of the filter is terminated in a 1-0 resistance.

For a Chebyshev low-pass filter with $w_c = 1$, the element values are given by

$$(1) \quad W_{odd} \quad (8.101a)$$

$$g^{N+i} \sim \sqrt{2k^2 + 1 - 2*V1 + k^2} \quad N \text{ even} \quad (8.1016)$$

When element g_N is a capacitor, $g_{N+1} = R$, but when g_N is an inductor,

$$g_k = . \quad ' \quad (8.101c)$$

where $a_k = \sin \frac{2k - 1}{v}$

$$b_{\frac{k}{2}} \sim \sin \pi r \frac{P}{2N} + \sin^2 \frac{k-r}{N}$$

$$\mathbf{p} = \frac{\mathbf{Vi} + \mathbf{k}^2 + \mathbf{i}}{\mathbf{In} / \mathbf{T} + \mathbf{A}^2 - \mathbf{1}}$$

$$2a, \\ gi \quad \sinh \frac{fi}{2N}$$

Numerical values for the g_k are given in Tables 8.1 and 8.2 for N up to 5, $a_{lc} = 1$, and a passband tolerance k^2 equal to 0.0233 (a 0.1-dB ripple in

TABLE 8.1
Values of g_k for maximally flat filter

k	N				
	2	3	4	5	
1	1.414	1.00	0.7654	0.6180	
2	1.414	2.00	1.848	1.618	
3		1.00	1.848	2.000	
4			0.7654	1.618	
5				0.6180	

TABLE 8.2
Values of g_k for Chebyshev filter with $k^2 = 0.0233$

k	N			
	2	3	4	5
1	0.8430	1.0315	1.1088	1.1468
2	0.6220	1.1474	1.3061	1.3712
3		1.0315	1.7703	1.9750
4			0.8)80	1.3712
5				1.1468

the passband) for the Chebyshev filter. More extensive tables are given in the references cited.

8.17 FREQUENCY TRANSFORMATIONS

The low-pass filter with cutoff at $w_c = 1$ and terminated in a 1- Ω load impedance may be used as a basis for the design of high-pass and bandpass filters with arbitrary resistive load termination. For this reason it is referred to as a prototype filter. For the purpose of this section, it is convenient to denote the frequency variable for the low-pass prototype filter by u . The power loss ratio may be expressed in the form

$$P_{IK} = 1 + P(OJ^2) \quad (8.102)$$

for maximally flat and Chebyshev responses.

If it is desirable to have a load termination R_L different from 1 Ω , the required filter is obtained by multiplying all other reactances and the generator resistance by a factor R' . That is, the prototype-filter reactance can be viewed as normalized with respect to R_L . The new values for the inductances and capacitances are

$$(8.103a)$$

$$(8.103b)$$

$$R' = R_L R \quad (8.103c)$$

where R' denotes the new value of R here, and not the generator "for in a" in the dual circuit of Fig. 8.216. In the discussion of frequency ".⁸ance" below, we assume a 1- Ω termination since the change in level of the filter can be made as a last step in the design P^{r o s}' w according.

If we replace the frequency variable to' by a new variable to

the power loss ratio becomes

$$P_{LR} = 1 + P(a^2) = 1 + P[f_m] \quad (8.104)$$

As a function of a or f , this power loss ratio is that of the low-pass prototype filter, but as a function of ω_c , it has a different characteristic, depending on how ω_c is chosen. A number of different frequency transformations, or mappings, are considered below.

Frequency Expansion

If it is desirable to change the cutoff frequency from unity to some other value ω_c , we choose

$$/\mathbf{H} =$$

and thus

$$\hat{\mathbf{L}} \mathbf{R} = 1 \quad (8.1051)$$

Cutoff occurs when the argument w/w_c equals unity or when $m = \omega_c$. The series reactances and shunt susceptances in the prototype filter must be replaced by new reactances and susceptances

$$jX; = j\hat{\mathbf{f}}^{-1} \mathbf{L}_k \quad / \mathbf{B} \mathbf{i} = /(-[\mathbf{i}$$

when a' is changed to w/ω_c , in order to yield the power loss ratio given by (8.105). Examination of the latter equations shows that the new values of the L_k and C_k must be

$$\mathbf{L}' = \pm \quad (8.106a)$$

$$\mathbf{Q}' = \mathbf{C}, \quad (8.1066)$$

The power loss ratios in terms of ω_c and a' are compared in Pigs. 8.22a and b.

*-Pass to High-Pass Transformation

A high-pass filter is obtained by choosing $f(w)$ equal to $-\omega_c/a'$ to yield a power loss ratio

$$P_{LR} = 1 + \frac{1}{a'} + P \quad (8.107)$$

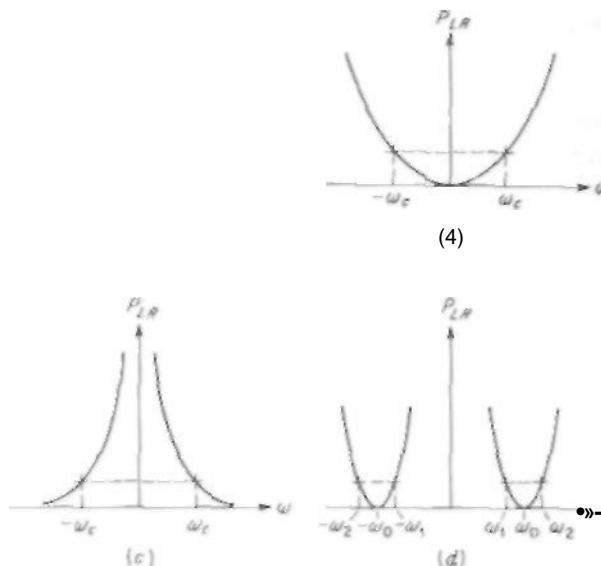


FIGURE 8.22

(a) Low-pass prototype-filter response; (b) frequency expansion; (c) low-pass-to-high-pass transformation, (d) low-pass-to-bandpass transformation.

This frequency transformation maps the point $w' = 0$ into the points $ID = \pm*$, the points $to = +1$ into $co = +w_c$, and the points $u' = \pm x$ into $co = 0$. The effect is to interchange the passband and stopband regions, as illustrated by Figs. 8.22a and c. To obtain a power loss ratio of the form (6.107), the series reactances and shunt susceptances in the new filter must be chosen as follows:

$$\&Z = \underset{co}{\sim j} \underset{i0_r}{-i^*} \quad jB'_t = \underset{co}{\sim j} \underset{to}{-c_k}$$

From these equations it is seen that all inductances L_k must be replaced by capacitances C'_k and the C_k must be replaced by inductances $L_{\setminus m}$ following manner:

$$\mathbf{P}^* = \underset{\ll L_h}{\sim} \quad (8.108a)$$

$$L_{\setminus m} = \frac{1}{c Z cl} \quad (8.108b)$$

Low-Pass to Bandpass Transformation

To obtain a bandpass filter, consider a change of variable accordin

$$u' = f(to) \quad co_0 \sim \quad (8.108)$$

$$co_2 \sim$$

$^0 \underline{\underline{H}}(z) = 5^w - H(z)$



FIGURE 8.23

Series and parallel tuned circuits.

This equation may be solved for to to give

$$\frac{(U_2 - wI)}{ft) = \frac{1}{ft) \pm \sqrt{lt^2}}$$

If we also choose $wf = o\omega_{Co}$, we obtain

$$\omega_o = \frac{to^2 - 4Q}{o^2 + yio} \quad (a); \quad -o\omega_y + 4\omega_f \quad (8.110)$$

The point $ft)' = 0$ is seen to map into the points $to = \pm to_Q$ and $to' = +1$ maps into the four points $\pm(o\omega_2 - to)/2 + (w_2 + to)/2 = \pm ft_2$ and $+i\omega_2$. Thus the prototype-filter passband between ± 1 maps into passbands extending from $ft)$, to $o\omega_2$ and $-i\omega_2$ to $-ft_2$, which represent bandpass filters with band centers at $\pm\omega_Q$ equal to the geometric mean of ω_2 and ω_1 , as in Fig. 8.22c/.

The required filter elements may be deduced by considering the frequency behavior of series and parallel connections of L and C , as in Fig. 8.23. For the series circuit we have

$$X = \frac{1}{jwC} = \frac{1}{V_c} \frac{1}{WLC} = \frac{1}{o^2 W}$$

and for the parallel circuit we have

$$B' = jtoC = \frac{m}{jML} \frac{1}{V_L} \frac{1}{WLC} = \frac{1}{tjyXC}$$

If we make yXC equal to (u_0') , we obtain

$$rn((o \cdot ft)_0) \cdot mi \cdot rc \sim (w \cdot \ll),$$

The required frequency transformation may now be seen to be obtained if we replace the reactance $jX_k = jtoL_h$ and the susceptance $jB_k = j<oC_k$ in the prototype filter by series and parallel tuned circuits such that

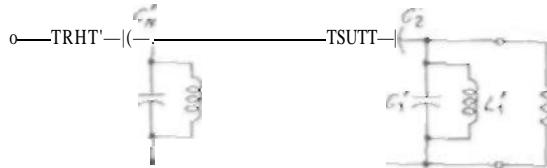


FIGURE 8.24
A bandpass-filter network.

This requires that we choose

$$\frac{u_k c_h}{w_0 L} = \frac{1}{\omega_2 - \omega_1} \quad (8.111a)$$

or $L_A' = \frac{1}{\omega_2 - \omega_x} \quad (8.1116)$

for the series circuit. For the parallel circuit we must choose

$$\frac{\hat{C}_k}{L_k} = \frac{1}{\omega_2} - \frac{1}{\omega_1 w_2} \quad (8.112a)$$

$$\frac{1}{L_k} = \omega^2 C^* = \frac{1}{w_2 - w_x}$$

The resultant filter network is illustrated in Fig. 8.24.

Period Bandpass Mapping

The general concept of frequency mapping may be applied in a variety of other ways as well. All that is necessary is to replace the reactances and susceptances $jw'C_k$ in the prototype filter by other reactances and susceptance functions having ω' replaced by a new function ω'' . In further examples we shall consider the effect of replacing all the inductances and the capacitors C_k by open-circuited stubs of length a and characteristic impedance Z_0 .

admittance Y_k . The new reactance and susceptance functions become^t

$$jX'_h = jZ_k \tan | - / = jZ_k \tan \theta = jg_k \tan \theta$$

$$jB' = jY_k \tan \theta = jg_k \tan \theta$$

This filter has a power loss ratio given by

$$P_{IM} = 1 + P(\tan^2 \alpha) \quad (8.113)$$

The frequency transformation $w' = \tan(wl/c) = \tan \theta$ maps the whole ω' axis periodically into intervals of length $-rr$ in the θ domain or of length $-c/l$ in the w domain. This filter would be satisfactory at low frequencies where junction effects at the points where the stubs are connected to the main line are negligible and where the separation between stubs could be kept very small in terms of wavelength.^t

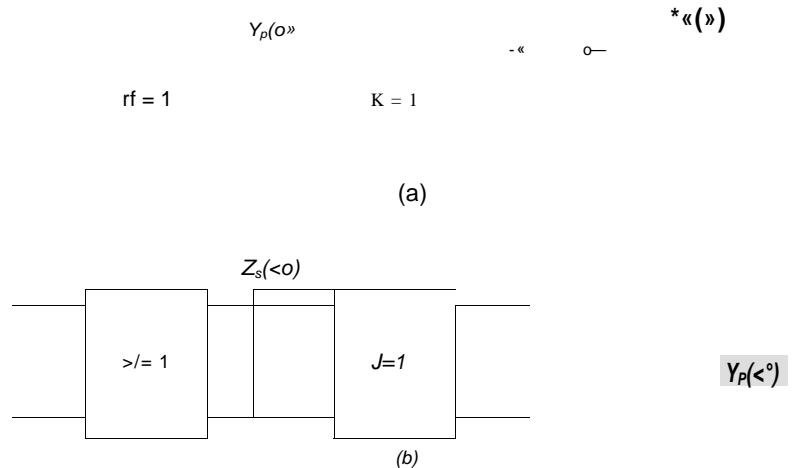
At microwave frequencies the alternative occurrence of series and parallel tuned circuits in the filter derived from the low-pass prototype filter by a frequency transformation is an undesirable feature. It is difficult to construct a filter of this type using microwave elements. However, it is possible to convert the filter with series tuned elements into an equivalent filter containing only parallel tuned circuits in a cascade connection, or to convert the circuit into one containing only series tuned circuits. The desired transformation may be obtained by the use of impedance and admittance inverters or quarter-wave transformers, as discussed in the next section.

IMPEDANCE AND ADMITTANCE INVERTERS

An impedance inverter is an ideal quarter-wave transformer. A load impedance connected at one end is seen as an impedance that has been inverted with respect to the characteristic impedance squared at the input. Impedance inverters may be used to convert a bandpass-filter network of the type shown in Fig. 8.25 into a network containing only series tuned circuits. By using admittance inverters the bandpass filter can be converted into a network using only parallel tuned circuits. Furthermore, by choosing the inverters correctly all of the inductors and capacitors can be chosen to have the same values. Thus impedance and admittance inverters enable us

^tThis transformation is known as Richard's theorem.

If a filter is made up of transmission lines of commensurate length and resistors, the only frequency variable occurring is $\tan(w/c)$. In this case the frequency transformation $iw = \tan(\omega/c)$ permits the use of conventional low-frequency-network synthesis techniques to be applied.

**FIGURE 8.25**

(a) Impedance inverter used to convert a parallel admittance into an equivalent series impedance: (b) admittance inverter used to convert a series impedance into an equivalent parallel admittance.

to use identical resonators, either series or parallel tuned, throughout the network.

In our initial discussion of inverters, we will assume that they are ideal frequency-independent elements. Later on we will consider the practical realization of inverters and how to compensate for their frequency-dependent characteristics.

Consider the parallel admittance element $Y_p(a>)$ with an ideal impedance inverter with characteristic impedance K connected on D sides as shown in Fig. 8.25a. A short circuit at the output will be transformed to an open circuit in parallel with Y_p . The input impedance given by

$$Z_{in} \sim li \quad K \quad Y_p = Y_p - Z_s$$

Thus the shunt element with two impedance inverters converts the admittance into an equivalent series impedance $Z_s(co) = Y_p U_o >^2 & Ig$ parallel tuned resonator with $Y_p = > C - j/coL = j\omega C(l) \sim \frac{a}{f} \%$ with converted into a series tuned circuit with $Z_s = j\omega L(l) \sim \frac{a}{f} \%$ the inductance L in henries having the same numerical value capacitance C in farads. If we want to convert an admittance $j\omega C(jd - wl/\omega)$ into a particular series tuned circuit with arbitrary

tance L , then we must choose K so that

$$K'VOJCAI - \frac{ju}{GJ} = \frac{wQ}{1 - \frac{wQ}{\omega_0}}$$

Thus we can transform a shunt element into a series element having the same dependence on frequency and an arbitrary impedance level.

Consider next a series element $Z_s(o)$ with an admittance inverter with characteristic admittance J connected on both sides as shown in Fig. 8.256. An open circuit at the output is transformed into a short circuit in series with $Z_{jo})$ so the input admittance will be

$$Y_m = J = Y,$$

Thus the series element with the two admittance inverters will appear from the terminals to be an equivalent shunt admittance. If Z_s is the impedance $juL(1 - u^2/o/w^2)$ of a series tuned circuit, it is converted into a parallel tuned circuit with admittance $Y = jwLil - wj/w^2) = jcoCil - to'i/w^2$, where C is numerically equal to L . For a series tuned circuit $7'a)L(l - (o^2Q/U)^2)$ to be transformed into an arbitrary parallel tuned circuit with $Y_p = ;'wC(l - ia/m^2)$, we must choose $J'L_x = C$ or

From another point of view we see that a series element can be replaced by a shunt element with an impedance inverter at the input and output. Similarly, a shunt element can be replaced by a series element with an admittance inverter connected at each port. The admittance and impedance functions for parallel tuned and series tuned resonators can be expressed in the form

The factor $J(T/L)$ is the admittance level of the resonator and $y(L/C)$ is the impedance level. The remaining factor is the frequency variable.

c—TRRP-

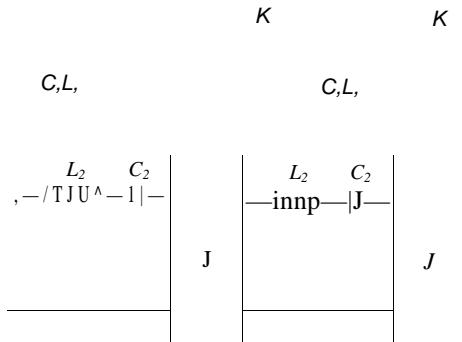


FIGURE 8.26

Use of inverters to convert a section of a filter into a network using only identical parallel series tuned resonators.

We will illustrate the use of inverters to convert the circuit shown in Fig. 8.26 into one with two identical parallel tuned resonators or one with two identical series tuned resonators. For the first case we choose K so that $K^2 J \rightarrow C^l - WQ/OJ^2 = ./aiL_2(1 - a>aoA^2)$ which requires that $K = yjL_2/C_l$. For the second case we choose J so that $J^2 jwL_2(l - CO\backslash<D^l) = jmC_l(l - co^2)/to\sim$ which requires $J = yC_l/L_2$. Impedance and admittance level changing can be accomplished by using different input and output inverters as shown in Fig. 8.27. For example, in Fig. 8.27a the parallel admittance appears as a series element $K^2 Y_p$ at the left-side port and as a series element $K^2 Y_p$ at the right-side port. Similarly, in Fig. 8.27b the series element Z_s is made to appear as a parallel admittance JfZ_s at the left-side port and as an admittance JfZ_s at the right-side port. We can modify the impedance and admittance inverters to accommodate an arbitrary change in the impedance and admittance levels of a resonator. In Fig. 8.27c the impedance level $y/L_x/C_x$ of the resonator is changed to \hat{L}_u/C_a by changing the impedance K' of the inverters to K , where K is chosen so that

K'^2	K^2	K^-	K''
$\bullet JLJC-L$	$/Lo/C_o$	$aioLy$	CO_QL_o

or

$$K = K V -$$

From the terminals the new circuit is equivalent to the old one. of the 8.270" we show a similar transformation of the admittance $yj^2i/ \backslash ^e$ parallel tuned resonator to a value $y/C_o/L_o$ obtained by rep

mra-

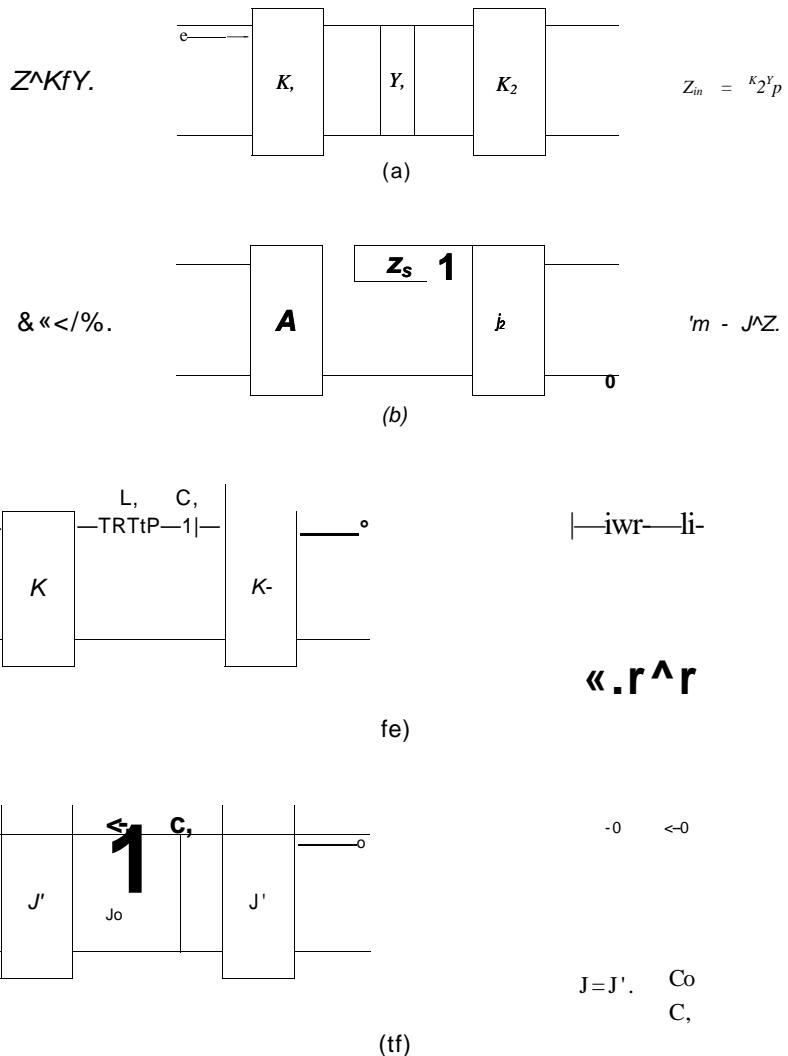


FIGURE 8.27

(o) Impedance level changing property of two impedance inverters; (6) admittance level changing properties of two admittance inverters; (c) equivalent basic sections with a series tuned resonator; *id*) equivalent basic section with a parallel tuned resonator.

inverters with characteristic admittance J' by new inverters with characteristic admittance $J = J' y J C_Q / C_I$.

Impedance and admittance inverters are ideal quarter-wave transformers. There is no basic difference in their inverting properties. The only distinction that we make is to use the symbol K to denote the characteristic impedance of the impedance inverter and we use J to denote the character-

istic admittance of an admittance inverter. An impedance in characteristic impedance K is equivalent to an admittance $\frac{1}{K}$. When $K = J = 1$ we have $\frac{1}{J} = \frac{1}{K}$. The inverter and it does not make any difference whether we identify the impedance inverter with $K \sim 1$ or as an admittance inverter with $J \sim 1$.

In Fig. 8.28a we show a bandpass filter terminated in a load resistance R_L and a source resistance R^* . We will use impedance inverters (this circuit into one using identical series tuned resonators with L_0 and C_n and terminated at both ends in a resistance R . The transformation process is easiest to understand by carrying it out as a successive several intermediate transformations. We first introduce unity inverters to convert all of the parallel tuned resonators into series tuned resonators. Thus, with all $J_i = K_i$ equal to unity, we obtain the network shown in Fig. 8.28b, where $L_1 = C_1$, and $Z_{13} = C_3$. We now change the impedance level of resonators 1 and 3 to $-JL_0/C_0$ by changing the first and second inverters to inverters with $K_1 = K'_1 = -JL_0/L_1 = -yJL_0/C_1$ and the third and fourth inverters to inverters with $K'_3 = K'_4 = -L_0/L'_3 = -yJL_0/C_3$. The new network is shown in Fig. 8.28c. This impedance shifting is the same as that illustrated in Fig. 8.27c.

As a next step we change the impedance level of resonator 2 from $y/L_2/C_2 = o>_0 L_2$ to $u>_0 L_2$ by changing inverters 2 and 3 so that

$$\frac{\mathbf{Kl}}{o>_0 L_0} \quad \frac{\mathbf{K?}}{u>_0 L_2} \quad \begin{matrix} \#3 \\ o>_0 L_0 \end{matrix} \quad \begin{matrix} \wedge 3 \\ ft) \end{matrix}$$

and hence

$$\frac{\mathbf{Li}}{\sqrt{\frac{L_0}{L_2}}} = \sqrt{\frac{L_0^2}{L_2 C_3}}$$

The resultant network is shown in Fig. 8.28d.

In order to change the load termination from R_L to R , we modify first inverter so that

$$\frac{K_1}{R} \quad \frac{K_f}{R}$$

and thus choose

The last modification needed to complete the transformation is to change the fourth inverter to one with an impedance

$$K_A = K'_4 \cdot \frac{1/X_0}{L_4} \quad \frac{L_1}{y L_4 C_s}$$

(a)

$$L_4 \quad C_{\infty} \quad k^{\frac{L_3}{H} \circ \beta} \longrightarrow L^{\alpha} r S h$$

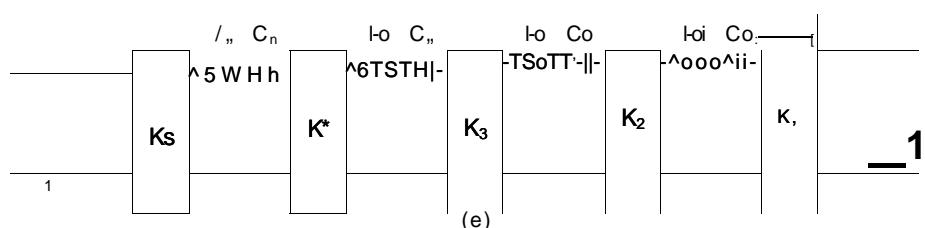
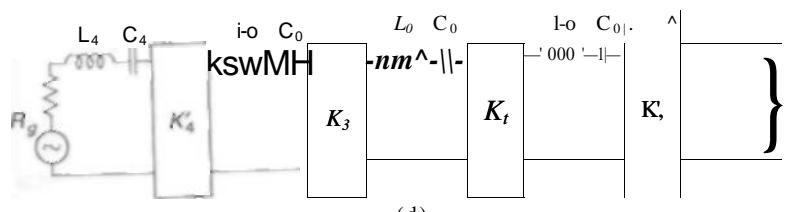
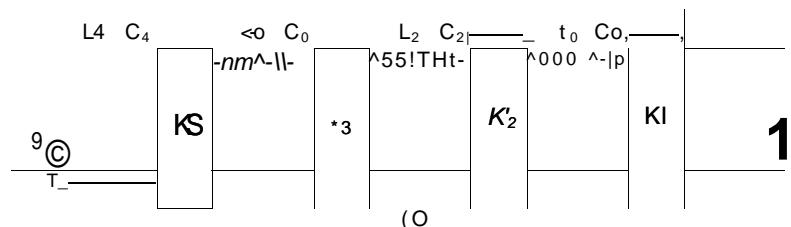
K=1

K=1

K=1

K=1

(6)



S S S L⁸? successive transformations used to convert the band^{as} filter •" (^ ^ network with identical series tuned resonators and equal .nput and output term.na resistance R as shown in (e).

and to introduce a fifth inverter between R_g and the fourth resonator place an inverter with impedance K'_5 chosen so that

If \mathbf{W}_e

* •
-F-*-

we can replace R_g by the new source resistance JR . After introducing the fifth inverter to accommodate the change in source resistance we "and the change it again when L_4 is replaced by L_0 . The new impedance K_5 must be chosen so that

and hence

$$K_5 = K'_5 \sqrt{\frac{L_0}{L_4}} = \sqrt{\frac{L_0 R R_g}{L_4}}$$

The final transformed network is the desired one and is shown in Fig. 8.28e.

The required inverters that will transform the original filter network to the final form shown in Fig. 8.28e have impedances given by

$$\begin{aligned} & / L_Q R \\ & \mathbf{A} \mathbf{I} \cdot \mathbf{C}, \end{aligned}$$

$$K_2 = J j_l \frac{1}{L_2 C}, \quad (8.1W^*)$$

$$K_3 = \quad (8.114c)$$

$$K_4 = \frac{\mathbf{r}^2}{L_4 C_3} \quad (8.11*0)$$

$$K_5 = \frac{J L_0 R R_g}{L_4} \quad V \frac{I L_0 R_\infty}{L G_*} \quad (8.H^{4e})$$

If instead of specifying identical resonators we choose $L_{0V} \wedge R_{g5}$ and L_{04} for the new inductances, f_{L00} for the new load resistance, $\wedge R_{g5}$ for the new source resistance (this is viewed as a parallel element)

series of transformations would give

$$K_1 = \frac{1}{R_L} = \frac{^01}{C_L R_L} \quad (8.115a)$$

$$K_{-2} = \frac{1}{L_2 C_1} = \frac{^01}{L_2 C_1} \quad (8.115o)$$

$$\#3 - \frac{1}{L_2 C_2} = \frac{^03}{C_3 L_2} \quad (8.115c)$$

$$\kappa_4 \sim \#13 = \frac{^04}{G_e i-4} \quad (8.115d)$$

$$\#5 \sim \%U\sim = \frac{^05}{G_e i-4} \quad (8.115e)$$

These expressions exhibit a regular pattern and can, by induction, be extended to a network with a larger number of resonators or to one with fewer resonators. For example, if we have three resonators then in the expression for K_{43} we interpret L_{04} as the new source resistance R_M and interpret L_4 as the old source resistance. For a network with five resonators, we would obtain

$$* M = \frac{^05}{\sim C7L7} \quad ^04$$

and

$$* 65 = \frac{^06}{R_e c_5} \quad ^05$$

The inverters shown in Fig. 8.28 can also be viewed as admittance inverters with $J_i = 1/K_i$.

By using admittance inverters the bandpass filter shown in Fig. 8.28a can be converted to a network that uses only parallel tuned resonators as shown in Fig. 8.29. The analysis can be carried out in a manner similar to that used for impedance inverters as a series of elementary transformations of the type illustrated in Figs. 8.266 and 8.27d. The required expressions

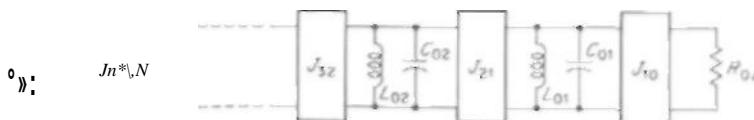


FIGURE 8.29

Equivalent filter network obtained by use of admittance inverters.

for the characteristic admittances of the inverters can be obt l
expected, by interchanging the role of inductance and resistance win!
of capacitance and conductance. It is readily found that

$$\overline{J_1 = \frac{1}{L_1 C_1}}^{\text{/ } \wedge 01 \wedge 00} \quad (8.116a)$$

$$\overline{J_2 = \frac{1}{L_2 C_t}}^{\text{/ } \wedge 02 \wedge 01} \quad (8.116b)$$

$$\overline{J_3 = \frac{1}{j C_3 L_2}}^{\text{/ } \wedge 03 \wedge -02} \quad (8.116c)$$

$$\overline{J_4 = \frac{1}{j C_4 R_g}}^{\text{/ } \wedge 04 \wedge 03} \quad (8.116d)$$

$$\overline{J_s = \frac{1}{j C_5 R_g}}^{\text{/ } \wedge 05 \wedge 04} \quad (8.116e)$$

If we denote the load *and source terminations* by R_L and R_g in the unmodified bandpass filter and by R_{0L} and R_{0g} for the transformed network, then the general expressions for $K_{kk,x}$ and J_{kik} for a filter with N resonators are

$$\#io = \overline{01 \text{ **0L}} \quad (8.117a)$$

$$K_{\star \star -!} = \overline{0t \text{ - } 0 \text{ * } -1} \quad fodd \quad (8.117b)$$

$$K_{A,A-1} = \overline{0 \text{ * } 0 \text{ * } -1} \quad A: \text{even} \quad (8.H7^C)$$

$$ft_{N+LN} = \overline{RogL'ON} \quad Nodd \quad (8.H7^D)$$

$$ft_{N+LN} = \overline{0g \text{ " } \wedge \text{ } 0A} \quad jV \text{ even} \quad (8.n^{7e})$$

and

$$' " - V C K S ; \quad (8118f)$$

$$t f > i * s y t t * . ! f e o d d \quad (8118g)$$

$$* * - ' = V \sim L \sim C \frac{en}{\dots} \quad (8.118c)$$

$$' ' - V m Z f c \quad N o d d \quad (8 U M)$$

$$J w - U H - V \frac{p}{V} \frac{r}{K} \frac{-}{og^A v} \quad N e v e n \quad (8.118e)$$

The impedance and admittance inverters used in the preceding analysis were assumed to be ideal frequency-independent quarter-wave transformers of electrical length $jr/2$ independent of frequency. Such ideal inverters do not, of course, exist. Nevertheless, the theory can be applied in practice. For very narrowband bandpass filters, say bandwidths of 1 or 2 percent, a quarter-wave length of transmission line or waveguide does not depart appreciably from the ideal inverter having a $ir/2$ electrical length. A filter designed on the basis of ideal inverters would have a response very close to the theoretical response in this case. For greater bandwidths the departure of a quarter-wave transformer from the ideal can be incorporated into the design by splitting the transformer into an ideal one with twt additional short lengths of transmission line on either side to account for the excess or deficit in phase shift from the ideal phase shift of 90° . For example, a transformer of length $0.3A_0$ at a wavelength of A_0 can be treated as a transformer of length $0.25A_0$ plus sections of line of length $0.025A_0$ on both sides. The excess length of line may be incorporated into and made part of the resonant circuit on either side of the inverter.

Lumped-element circuits that act as impedance inverters and admittance inverters are illustrated in Fig. 8.30 along with their equivalent characteristic impedance or admittance. These circuits involve negative elements and are frequency-dependent. However, the negative elements may be absorbed into the elements of the adjacent resonant circuits if they eliminate them from the overall network. The resultant filter then consists of tuned circuits coupled by single reactive elements, as illustrated in Fig. 8.31. Application of these techniques to microwave filter design is discussed in the following sections.

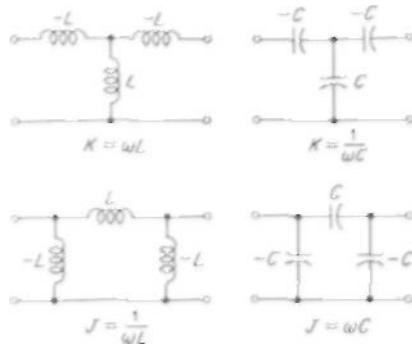
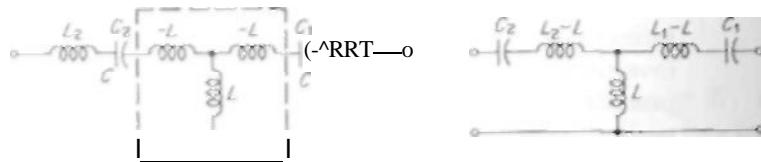


FIGURE 8.30
Lumped-element inverters.

Example 8.1. In this example we will design a two-resonator filter final form as shown in Fig. 8.32a. The filter is to be used in an intermediate^ frequency amplifier with center frequency of 10 MHz, a bandwidth of 0.5 MHz and having a load and source termination of 1,000 ft. The passband ripple is 2 dB.



$\text{o-omHh-lf} \quad | \quad \text{l(jHwo^X)}$

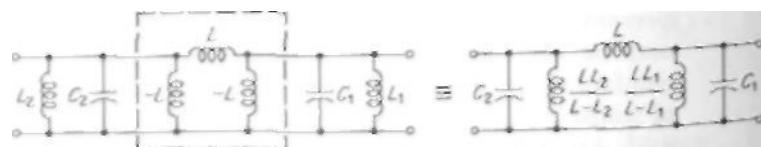
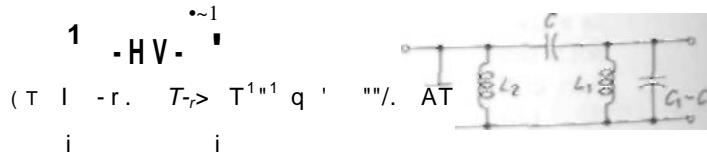
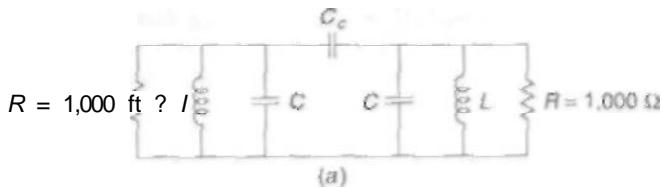


FIGURE 8.31 Reactance-coupled resonator-type circuits obtained by use of lumped-elements

inverted



92

 $L_2 \quad C_2$

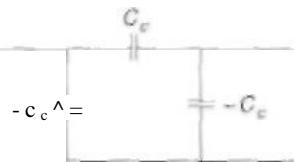
$$R = r \cdot 9, \$1 \quad *; \quad = rC, k - |$$

(*)

 $L, \quad C, \quad C, \quad L,$

Ji

1



(O)

(e)

FIGURE 8.32

(a) Bandpass filter; (b) low-pass prototype circuit; (c) bandpass prototype circuit; (d) transformed bandpass prototype circuit; (e) admittance inverter used to obtain the circuit shown in (a).

We begin the design by first designing a low-pass prototype circuit with a cutoff radian frequency $\omega_c = 1$ as shown in Fig. 8.326. For a 2-dB ripple the maximum power loss ratio in the passband is $10^{0.3} = 1.5849 = 1 + f^2$. The parameter p is

$$P = \ln \frac{\sqrt{T} T k^i + 1}{T \wedge i} = 2.1661$$

From (8.101) we obtain $R = 2fc^2 + 1 - 2j\% / 14 k^2 = 0.244175$ for the low-pass filter input termination. We also obtain from (8.101) $a_1 = \sin \theta / r / 4 = 0.7071$, $a_2 = \sin 3\pi / 4 = 0.7071$, $b_1 = 1.32306$, $b_2 = 0.32305$, and $g_x = 2.4881$ and $g_2 = 0.6075$. Hence, for the low-pass prototype circuit in Fig. 8.326, we have $C_1 = g_1 = 2.4881 \text{ F}$ and $Z_{12} = g_2 = 0.6075 \text{ H}$. The values of $a >$ at the band edges are ω_1 and ω_2 . Since $\omega_1 \omega_2 = u > \%$ and $u >_2 - \theta_2 = 2\pi \times 0.5 \times 10^3$, we get $\omega_1 = OJ_X + TT \times 10^6 = ai\% / oY$. By solving this equation we obtain $\omega_1 = 6.1281 \times 10^7$, $\omega_2 = 6.4423 \times 10^7$. The corresponding frequencies are 9.7532 and 10.2532 MHz. For the bandpass-h Iter prototype circuit, we use (8.1116) and (8.1126) to obtain $C_1 = 2.4881 / TT \times 10^6 = 0.79199 \text{ MF}$ and $L_2 = 0.6075 / CT \times 10^6 = 0.19337 \mu\text{H}$. The elements L_1 and C_2 are found using

$$\text{and } L_1 C_1 = a j L_2 C_2 = I \text{ and are } L_x = 3.198 \times 10^{10} \text{ F and } C_2 \approx 1.3099 \times 10^{-9} \text{ F.}$$

By introducing two unity inverters, we can convert the resonator to a parallel tuned resonator with $L'_2 = C_2 = 1.3099 \times 10^{-9} \text{ H}$ and $C_V = L_2 = 0.19337 \text{ M}\Omega$. In order to shift the admittance level of the resonator from $o_d C_2$ to OJQC, we change the admittance inverters to ones with $J_2 = 2 \times 10^{-9} \text{ H}$.

$$J_2 = \frac{7.9199 \times 10}{1.9337 \times 10} = 2.0238$$

For the admittance inverters we use the capacitive circuit shown in Fig. 8.70 for which $J_x = J_2 = OJ_0 C_C$; thus $C_c = (2.0238/2\pi) \times 10^7 = 0.0322 \text{ M}\Omega$. Since we want to use the same input and output terminations, we change J_1 so that $R = 1$ appears the same to the network looking through the original inverter; thus $J'_2 \times 1 = |J| \times R$ so $J_2 = 2.0238 \times 0.244175 = 1$. This result shows that a 141 source resistance appears the same to the network as the 0.244175 resistor seen through the inverter with $J = J_2$. Thus we can discard the input inverter.

We now scale the impedance level to 1,000 Ω by multiplying L by 1,000 and dividing C_x by 1,000. We also absorb the $-C_c$ in the remaining inverter as part of C . Thus we obtain $L = 0.3198 \text{ nH}$, $C = (C_1 - C_c)/1,000 = 760 \text{ pF}$, and $C_c = 32.2 \text{ pF}$, for the element values in the final filter shown in Fig. 8.32a.

The double-tuned coupled circuit that we have designed is a classic one that has been widely used in low-frequency radio IF amplifiers. The power transmission coefficient is equal to $1/P_{LR}$ and is shown in Fig. 8.33. If the circuit is designed for a maximally flat response, it is said to be critically coupled. When the response has a dip in the center, the circuit is overcoupled. The coupling coefficient for the circuit equals $C_C/(C + C_C)$. For critical coupling the coupling coefficient equals $1/Q = w_0 L/R = 0.02$. For the circuit that we designed, the coupling coefficient equals 0.0406 so clearly the circuit is overcoupled. Two other well-known properties of this classic circuit are that

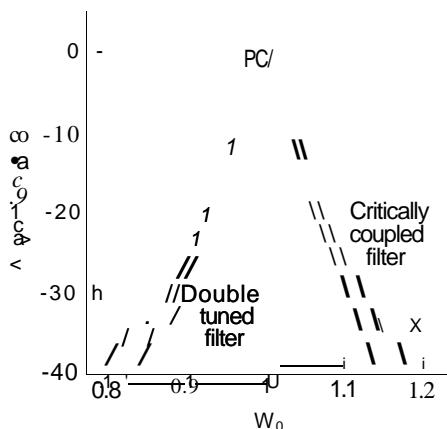


FIGURE 8.33 Power transmission coefficient for the double-tuned filter shown in Fig. 8.32a. The curve shows the response for a circuit with a 3-dB bandwidth equal to Δf .

the 3-dB fractional bandwidth equals $\sqrt{2}/Q$ for critical coupling, and for an overcoupled circuit having a 3-dB dip at ω_o , the coupling coefficient equals $(1 + \sqrt{2}J)$ times critical coupling and the 3-dB fractional bandwidth equals $2i/2114/Q$.

19 A MICROSTRIP HALF-WAVE FILTER

In Chap. 7 we showed that open-circuited or short-circuited transmission-line sections one-quarter or one-half wavelength long were equivalent to either series or parallel tuned lumped-parameter LC circuits. Thus it is not surprising to find that sections of transmission lines are widely used as resonators in filters. Such structures are particularly appropriate for microstrip filters because of *easy fabrication and low cost*. In this section we will describe a microstrip filter that uses one-half wavelength open-circuited microstrip lines for the resonators. The resonators are coupled by means of the capacitance of the gap between the resonator sections. A typical half-wave filter consisting of three resonator sections is shown in Fig. 8.34.

For an open-circuited transmission line the input admittance is given by

$$F_{jn} = y \cdot y; tan^{\wedge}$$

We can expand Y_{in} in a Taylor series about the frequency ω_0 where $\partial^3(Y_{in})/\partial(\omega)^3 = 0$; thus

$$\frac{dp}{dco} = \frac{[\sec^2/3(w_0)Z](w - w_0)}{r < a_o P'_o \quad ft' - ft''}$$

where β_0 equals $d/3/dio$ evaluated at circuit,

For a parallel tuned LC

$$= J \ll_0 C \frac{1}{w c \ll_0} (w - w_{\infty})(w + w_0) = 2J O J_0 C \frac{(D - (O_c f_c > 0))}{f_c > 0}$$

Thus, for frequencies within around ± 10 percent of a_0 , the open-circuited transmission line is equivalent to a parallel tuned LC circuit if we choose

$$C \quad a >_n 7r/3'_{n-1}$$

s₁ **s₂** **s₃**

FIGURE 8.34

A microstrip half-wave filter with three resonators.

In the half-wave filter the open-circuited transmission lines can be used equivalent parallel tuned resonators. We can choose all of the resonators to be identical provided we insert appropriate admittance inverters between each resonator.

The network shown in Fig. 8.35 functions as an admittance inverter. It consists of two transmission lines with *negative* electrical length δ and a series capacitor C_g in the center. For this network a load admittance Y_L connected at one end is transformed into

$$Y_L = \frac{Y_c - jY_{Lt}}{Y_c + jY_{Lt}}$$

at the location of $B = \omega C_g$ and where $t = \tan \delta$. Just to the left of C_g , we have

$$K > \frac{jBY_{in}}{JB + Y_{in}}$$

and at the input

$$Y_{in} = \frac{y_{rK} - m}{m} = \frac{Y_c^2 Y_{Li} - 2BYt + jY^2 L_B (t - t^2)}{m}$$

We now equate the coefficient of Y_L in the numerator and the constant term in the denominator to zero. Both conditions give

$$\frac{t}{I - t^2} = \frac{1}{2} \tan 2\theta \quad (8.120a)$$

-e -0

FIGURE 8.35

A simple admittance inverter used in the filter.

crostr-P * * —

With this condition imposed

$$Y_+ = \frac{Y^2 t(2B - Yj)}{(Y_c + 2Bt)Y_L} = \frac{Y^2 \tan \theta}{Y_-} \quad (8.1206)$$

upon substituting for B . When the relations specified by (8.120) hold, we obtain an admittance inverter with

$$J = y_c \tan \theta \quad (8.121)$$

The fact that the network in (8.35) involves transmission-line sections with negative electrical lengths does not cause any difficulty since we can add a length θ at each end of the transmission line and then add lengths of electric length $\sim \theta$ to serve as part of the inverter network. Since both $B = <i>C$, and θ are proportional to w the inverter is not an ideal one. However, for narrowband filters J does not change very much over the passband that the filter must operate.

In the next example we carry out the design of a three-resonator filter of the type we just described.

Example 8.2 Design of a three-section half-wave bandpass filter. We will design a filter with a center frequency of 10 GHz, a bandwidth of 1 GHz, and having a passband tolerance $k^d = 0.1$ which gives a maximum VSWR equal to 1.863.

We first design the low-pass prototype circuit shown in Fig. 8.36a. From (8.101) we readily find that $g_s = C_s = g_{s,i} = C_{s,i} = 1.5062$ and $g_{\%} = L_2 = 1.1151$. For the bandpass-filter prototype network shown in Fig. 8.366, we require $C' = C_s / (y_3 - w_s) = C\ell = (1.5062/2-) \times 10^{-9} = 239.72 \text{ pF}$. $L'_s = L_2 / (a_{s,i} - \theta_s) = (1.1151/2-) \times 10^{-9} = 177.47 \text{ pH}$. By introducing unity-inverters we can convert the series tuned resonator to a parallel tuned resonator with $C_2 = L'_s = 177.47 \text{ pF}$. We now change these inverters into inverters with J chosen so that $J^2/C_2 = 1/C_2$ or $J = |jC\ell| = ijC\ell/L'_s = 1.16221$. The network in Fig. 8.36a* is the final one after we scale all impedances by a factor of 50 so that the input and output transmission lines will have a characteristic impedance of 50 ohms instead of 1 SI. Thus we require $C = (239.72/50) \text{ pF} = 4.7944 \text{ pF}$, and $J = 1.16221/50 = 0.023244$. The ratio of the resonator admittance $|jC\ell| = w_0 C$ to V' , of the input and output lines is $2 \times 10^{10} \times 4.7944 \times W^{-2}/0.023244 = 15.062$.

If we use 50-fl transmission-line sections for our resonators, then the ratio of the resonator admittance given by (8.119) to the characteristic admittance of the input and output lines is $a_{s,i} Tr p_s'' / 2l_{in}$. If we neglect dispersion, then $0_0 = \sqrt{\mu_0/c}$ and $p_0/p_n = \sqrt{\mu_0/\mu_n}$, where $c = 3 \times 10^8 \text{ m/s}$ is the free-space velocity of light. In this case the ratio equals $-1/2$ which is different from the required ratio of 15.062. This problem is overcome by inserting an admittance inverter between the input line and the filter and also one between the output line and the filter. These inverters are chosen so that the line admittance Y_c will appear to be 15.062 times smaller than the

La

$$f_i = 1 \quad \mathbf{c}, \mathbf{4} = \quad T = C, \quad |f| = 1$$

(a)

$$\begin{array}{ccccccc} C_j & L_J & j \xrightarrow{TJTO} & L_V & C; \\ 1^* & \cancel{\mathbf{4}} & & C_j & & \cancel{\mathbf{4}} & *1 \\ & & " \underline{\underline{II}} " & & & & \end{array}$$

(b)

$$Co = U, \quad C'i = Lg, \quad 14 = c,$$

$$1 \ a = f = j' \quad J'$$

(c)

5012-

50 O

(O

FIGURE 8.36

(o) Low-pass prototype circuit; (6) bandpass filter; (c) transformed bandpass niter, filter network.

final

resonator admittance v_{jlt} . Thus we require

$$\frac{4}{Y_c} = \frac{TIV_s}{2 \times 15.062} \quad \text{or} \quad \frac{4}{Y_c} = \frac{f}{V_s} \frac{1}{2 \times 15.062} \quad y = 0.3229V,$$

If dispersion is important then we note that

$$\frac{dp}{d\theta_0} = \frac{d^2 \cdot o_j/c}{du^2} = \frac{y/ej}{c} = \frac{a}{2y/T_d c} \quad do$$

which now gives

(8)

$$2 \backslash \epsilon do$$

in place of (8.119) for the resonator admittance. This expression can be evaluated once the substrate material and microstrip width W have been chosen to give a 50-JI transmission line. The frequency dependence of e_{r} can be evaluated using Kobayashi's formula (3.176). For simplicity, we will neglect dispersion effects.

Our last step is to design the required inverters. For the two middle inverters we require the ratio of J to the resonator admittance to be $0.023244/C = 0.023244/(2 \times 0.047944) = 0.07716$. Hence we must scale J from the value 0.023244 to $0.023244 \times 7rF_r/2w_0C$ since we are using resonators with admittances equal to $-Y_e/2$. Our new and final value for J is $0.07716 \cdot Y_e/2 = 0.1212Y_r$. From (8.121) we find that we require $0 = \tan^{-1} 0.1212 = 0.1206$. From (8.120a) we obtain $B = 0.00245$ and hence a gap capacitance C_g of $0.00245/w_0 = 0.039 \text{ pF}$.

For the input and output inverters, $J_0 = 0.323Y_r = Y_c \tan l_0$ so $l_0 = 0.3124$ and $B_0 = 0.01 \tan 2l_0 = 0.00721$ which gives a gap capacitance C_{g0} of 0.1147 pF .

In this example we determined the inverter parameters from basic principles in order to illustrate again the theory involved. We could have made the evaluations more directly by using the general formulas (8.118). Thus we would have

$$\gamma c^{\wedge} Z = J_{\text{TM}} \sim J_{\text{H}}$$

$$Coi^m$$

where $C_{01} = C_{02} \sim TrY_r/2t \gg_0$ is the chosen value for the resonator capacitances and $C_1 = C_2$ and $L_1 = L_2$ are the values used in the prototype bandpass-filter circuit. The reader can readily verify that the above expressions give the same results that we found after a series of steps.

The equivalent circuit of a symmetrical microstrip gap (equal width input and output lines) is a symmetrical capacitive n network shown in Fig. 8.37a. Therefore, when we introduce a gap capacitance C_g for our inverters, we will also introduce a shunt capacitance C_s at the ends of the two coupled fines. These capacitances are equivalent to a small effective lengthening of each line. For the two circuits in Fig. 8.376, we have

$$Y_{ss} = > C_s = jY_c \tan /3 M * > jY_e 0 M$$

and hence the line will appear to be

$$, I = g \quad mm$$

longer due to the capacitive loading. Thus we need to reduce the physical length of each resonator by an amount proportional to the capacitive loading at each end. For the two middle gaps we will let $A/$ be the length reduction required and for the two outer lengths we will let the length reduction be denoted by $A/,,$.

The designed filter is symmetric about the midsection. In Fig. 8.38e we show one-half of the filter with the various parameters that we have

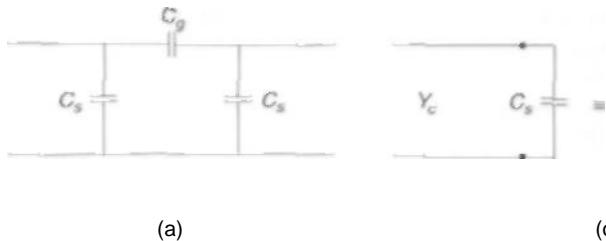


FIGURE 8.37

(a) Equivalent circuit of a gap in a microstrip line; (b) equivalent transmission-line used to represent shunt capacitive loading.

determined. At each end of the transmission-line sections, we add and subtract lines of length \$l_1\$ and \$-l_0\$ or \$-\theta_0\$ and \$-\theta_{0e}\$. The lines with negative electrical lengths are then part of the required inverter circuits. The additional lengths \$\theta_0\$ and \$\theta_{0e}\$ are absorbed as part of the resonators and require that the physical length of each resonator be shorter by the corresponding amounts. We must also account for the line lengthening due to capacitive end loading. The length of the first resonator is chosen so that \$\theta_0 l_1 + p_0(l_0 + \Delta\theta_0 + \theta_0 + 8) = \pi r\$; thus

$$r \rightarrow t - \$ - M_0 - M$$

Similarly, we find that the required physical length of the second resonator

$$h = \frac{IT - 26}{-2M}$$

$$\bullet \bullet \parallel \bullet \bullet \text{---}' - \text{li}^* - p_0(l_2 + 2 Al) \rightarrow g - B .. - O \quad 6$$

(a)

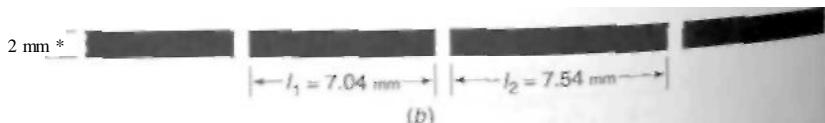


FIGURE 8.38

(a) Half-wave filter equivalent circuit; (b) metalization pattern for narrow microstrip construction on a substrate with thickness \$0.5\$ mm and dielectric constant \$2.65\$.

The third resonator has the same length as the first one. We can ignore the effects of A_0 and θ_0 at the input and output lines since the length of these lines are arbitrary.

We will complete the design by assuming that the filter will be built using microstrip construction and a substrate 1 mm thick and with a dielectric constant $\epsilon_r = 4.2$. This thickness of substrate at 10 GHz results in negligible dispersion of the characteristic impedance. The required strip width is 2 mm for a 50- Ω line. The low-frequency effective dielectric constant is 3.2. By using Kobayashi's formula we find that $\epsilon_c = 3.37$ at 10 GHz. The dielectric constant is about 5 percent larger at 10 GHz and since $f_i = \sqrt{\epsilon_r} k_n$ this will result in about a 2.5 percent shorter resonator length than what would have been specified if we used the low-frequency value of ϵ . If we did not change the resonator lengths, the filter center frequency would have been about 2.5 percent below the design value. Since the filter bandwidth is 10 percent, this amount of detuning is significant so a satisfactory design would not have resulted if we had ignored the dispersion in t_{ir} .

For the two outer gaps we need a capacitance of $0.1147/2 = 0.0557$ pF/mm. For the two inner gaps we need a capacitance of $0.039/2 = 0.0195$ pF/mm. From the data given in Fig. 7.10c, we see that the required gap spacings will be very small, to the extent that the shunt capacitance C_s is negligible; so we do not need to adjust the resonator lengths to compensate for C_s . When C_t in Fig. 7.10c can be neglected, the gap capacitance $C_A = C_0/2$. For the inner gaps we thus require $C_0 = 0.39$ pF/cm, which is obtained using a gap spacing of about $0.22 W = 0.44$ mm. The data in Fig. 7.10c do not give values for the capacitance for S/W less than 0.05 for which $C_s = 0.6$ pF/cm. For the outer gaps we require $C_0 = 1.147$ pF/cm. The data in Fig. 7.10 apply to infinitely thin strips. For very small gaps the strip thickness will be comparable to the spacing, so that the parallel plate capacitance of the end faces must be accounted for. If we use 1-oz copper-clad board, the strip thickness will be approximately 0.036 mm. If the gap spacing were 0.05 mm, the parallel plate capacitance between the two end faces would be

$$C_{PP} = \frac{0.036 \times 2 \times 10}{77 \times 0.05} = 0.0127 \text{ pF}$$

for a strip 2 mm wide. This amount of capacitance is about 10 percent of the required amount. Since the data for determining the outer gap spacing are not available, they need to be found either experimentally or numerically before the filter design can be completed.

The resonator lengths can be specified since we can neglect A_0 and A_{10} . Thus for the center resonator we require

$$l_2 = \frac{-20}{3} = \frac{77 - 2 \times 0.1206}{1/3 \times 37 - 2/3} = 0.754 \text{ cm}$$

The required length for the two outer resonators is

$$l_1 = \frac{v - e - e_0}{\gamma} = \frac{77 - 0.1206 - 0.3124}{0.754} = 0.704 \text{ cm}$$

The metalization pattern for the filter is shown in Fig. 8.386.

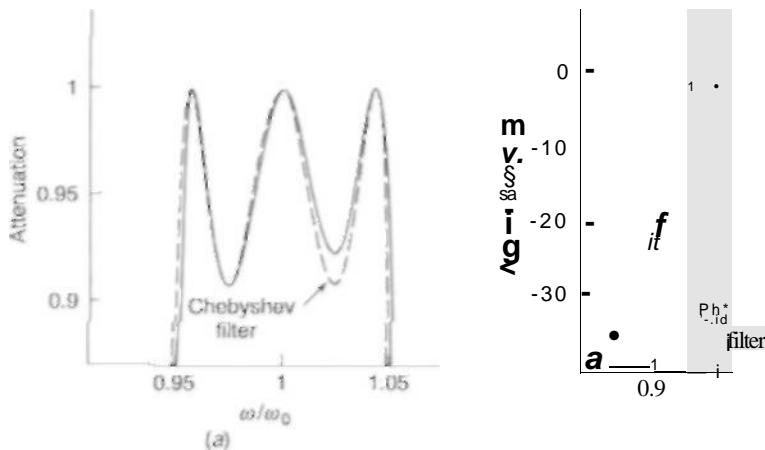


FIGURE 8.39

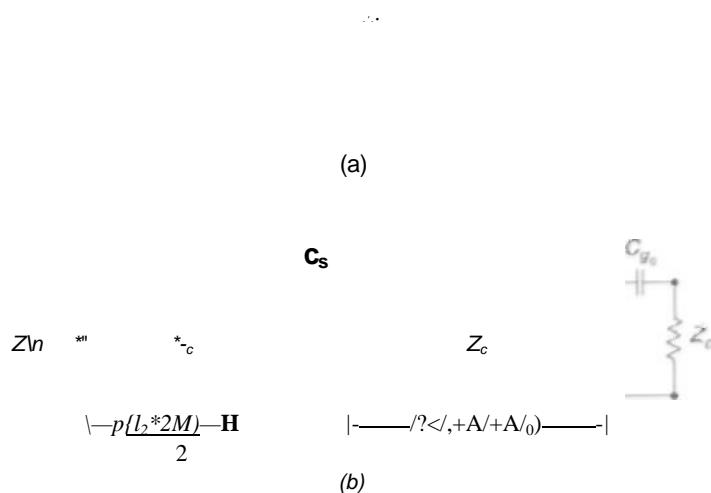
(a) Comparison of attenuation of half-wave filter and ideal Chebyshev filter in the passband
 (b) comparison of half-wave and ideal Chebyshev filter attenuations as a function of normalized frequency over an extended frequency band. The dotted curve shows the deterioration in the filter response when the two outer gap capacitances are reduced in value by 10 percent-

From the attenuation of the microstrip line, which is approximately 0.0029 Np/cm, we obtain the intrinsic resonator $Q = p/2a = 665$. The external Q of the first resonator is $\langle o_0 C/G \rangle = \langle o_0 CY/Jjj \rangle = (fLTr^2/2)/(0.323y_c)^2 = 15.06$. Since this external Q is much smaller than the unloaded resonator Q , the losses in the resonators will have a negligible effect on the performance of the filter.

The calculated insertion loss or attenuation for the filter is shown in Fig. 8.39 as a function of the normalized frequency ω/ω_0 . Also shown is the response or attenuation of the ideal Chebyshev filter which has the power loss ratio

$$PI_{\text{dB}} = 1 + fe^2 r_3^2 (\omega/\omega_0) = 1 + *^2 (4\omega^8 - 3a/\omega^2)$$

where $\omega_0 = (a/\omega_0 K_a)/u_a - w_Q/fo = U^2 - w^2/a > 0$. Since $P_{LR} = 1A^2 \sim P_{\text{in}}$ where $f >$ is the magnitude of the input reflection coefficient, the $p <$ transmission coefficient $1 - p^2$ equals $1/P_{LR}$. The filter insertion loss Δ attenuation in decibel units is given by $10 \log P_{LR}$. In Fig. 8.39a we see that the filter response is very close to that of the ideal Chebyshev filter with the main difference being a somewhat smaller Δ over $O/I/OIQ \Rightarrow 1.025$. In Fig. 8.39b the two responses are shown in decibel units over the not-normalized frequency range 0.8 to 1.25. The response of the filter is remarkably close to that of the Chebyshev filter with the same Δ over $O/I/OIQ \Rightarrow 1.025$. The above the center frequency (33.8 dB versus 40.9 dB at $\omega/\omega_0 = 1$) function half-wave filter was designed on the basis that the admittance $Y_e \tan \pi l$ for an open-circuited transmission line could be approximated as $P_{\text{in}} = Y_e \tan \pi l$.

**FIGURE 8.40**

(a) Illustration of point at which the impedance mismatch is evaluated; (b) equivalent circuit used to evaluate Z'' .

$jY_c v(a) - w_0)/io_0$ in the vicinity of the frequency at which the resonator was one-half wavelength long. Thus it is not surprising that there should be some difference in the attenuation at frequencies that are far from the center frequency. The results shown in Fig. 8.39 do, however, show that the approximations made are acceptable and lead to useful filter designs. If we use $w' = (a - w_0)/u > 0$ in place of the expression (8.109) in the Chebyshev polynomial, we will obtain a symmetric response about the point α , and a closer agreement with the response of the half-wave filter. The latter has a symmetrical response curve because $\tan \theta$ is symmetrical about $H = 0$.

The impedance mismatch M at the input to the filter is equal to $1 - p^l$. In Sec. 5.7 we showed that in a lossless network the impedance mismatch is invariant throughout the network. Hence we can evaluate it at the center of the filter. If we let Z_J and Z_{in} be the impedance at the center of the filter when looking toward the right and left as shown in Fig. 8.40a, then the impedance mismatch is given by

$$A/ = \frac{4^{\wedge}; / ? - (R, :,)^2}{iz + z, j^2" |z:i^a}$$

since the filter is symmetrical so $Z_{in} = Z_r$. The evaluation of Z'' is easily carried out using the circuit model shown in Fig. 8.40b. This method was used to evaluate the filter attenuation shown in Fig. 8.39 and is much simpler than evaluating the input impedance for the whole filter network.

The construction of a filter requires good control on the electrical parameters if the theoretical performance is to be met. In Fig. 8.396 we show the response of the filter when the two outer gap capacitances are reduced in

H

• b.

JT^t

S A Parallel

FIGURE 8.41

coupled microstrip filter

value by 10 percent from the design specification of 0.1147 to 0.1032 DF F the figure it can be seen that the passband attenuation at $u/a_1 = n^o$ increased to 1.43 dB which exceeds the passband tolerance of 0.414 dR I about 1 dB. Nevertheless, the overall filter response is still quite good

8.20 MICROSTRIP PARALLEL COUPLED FILTER

The parallel coupled filter shown in Fig. 8.41 is more compact than the half-wave filter described in the previous section. Since the coupling between resonators occurs over a quarter-wave-long side of each resonator the slot width is larger and the tolerance on the slot width is not as critical. The ends of each resonator section may be open-circuited or short-circuited. The design of parallel coupled filters is readily carried out by using an equivalent circuit of the filter which is easily designed. This equivalent circuit for each pair of coupled resonators is derived below.

Consider the pair of coupled microstrip lines shown in Fig. 8.42a. The strips will be considered to have unequal widths. When the voltage applied to each strip is the same (even mode) as shown in Fig. 8.42b, the currents on the two strips will not be the same because of the different widths. We will let the currents be $I_1 = Y_v^a V$ and $I_2 = Y_v^b V$, where Y_v^a and Y_v^b represent the characteristic admittance of strip a and strip b relative to the ground plane. For the odd mode with a voltage V applied to strip a and $-V$ applied to strip b , the currents are given by $I_1 = Y_o^a V$ and $I_2 = -Y_o^b V$ where Y_o^a and Y_o^b are the characteristic admittances of the two strips for the odd mode. We will show that the coupled-line circuit shown in Fig. 8.42a is equivalent to the circuit shown in Fig. 8.42c, where

$$Y_x = y(Y_o^a - Y_e^a)(Y_o^b - Y_j^b) \quad (8-124a)$$

$$\mathbf{r} \cdot \mathbf{W} + \mathbf{m} \cdot \mathbf{r}, \quad (8-124b)$$

The equivalence will be established by showing that the input admittance port 1 with port 2 terminated in an open circuit or a short circuit is the same for both circuits.

The voltage and current waves on the circuit shown in Fig. 8.42a can be expressed in terms of a superposition of odd and even modes. Thus on strip a we express the voltage V_a as

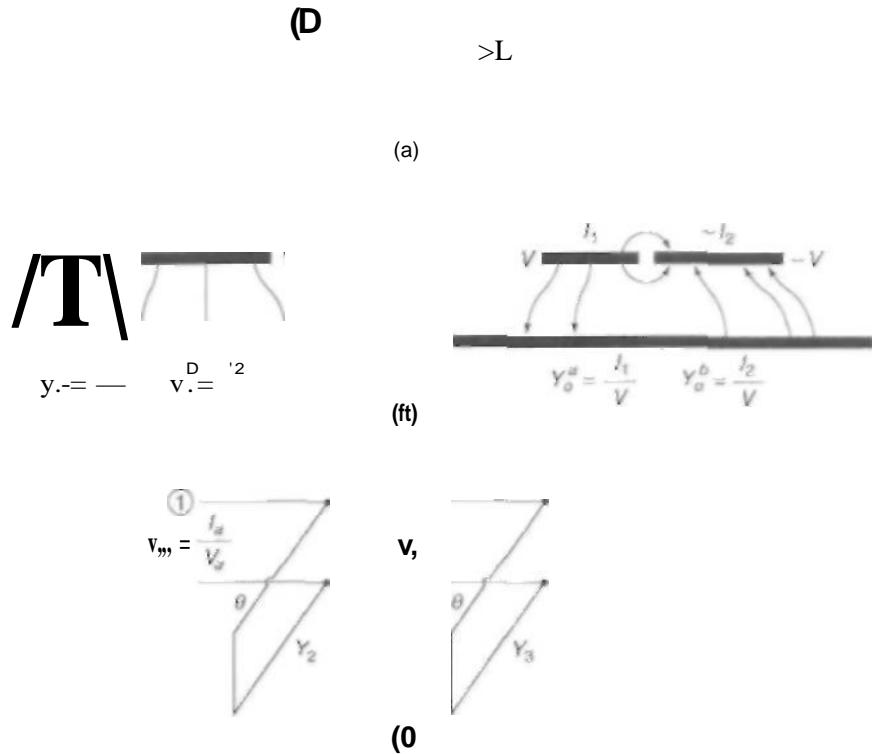


FIGURE 8.42

(a) A pair of coupled microstrip lines; (b) illustration of even and odd modes; (c) equivalent circuit of coupled strips shown in (a),

waves in the form

$$V_a(z) = V; e^{-Jlii} + V \sim e^m + V^* e^{-**} + V,, e^{i5''}$$

$$IA^*) = Y?V?* \sim^{i0*} - Y?V \sim e^{if^*} + Y?V^* e & * - Y^*V,, e^{JlU}$$

It is assumed that both modes have the same propagation constants. On strip b the voltage and current waves are

$$\mathbf{V}^*(z) = \mathbf{V} e^{-1**} + \mathbf{I} T^{**''} - \mathbf{T}^{**'*} - \mathbf{K} \ll^{**}$$

$$/_6(z) = I^*V + e^{-**} - Y_e^b V; e^\wedge - Y_0^h V?e^{-J'<>} + y,,^*V_n T e^{-i2}$$

When $2 = 0$ we must have $V_A(0) = 0$; thus

$$(K^+ + V;) - (V; + V-) = 0 \quad (8.125a)$$

since the end at $2 = 0$ on strip 6 is short-circuited. If we place a short

circuit at port 2, we will require that at $z = l$, where $\beta z = 0$, which gives $V_a(l) = 0$.

$$(V_r^+ - V_s^-)e^{-j0} + \{V_s^+ - V_r^-\}e^{jB} \quad (8.125b)$$

On strip a we must have $V_a(l) = 0$ since this is also a short-circuited thus

$$(V_s^+ + V_r^+)e^{-j0} + (V_s^+ + V_r^-)e^{jB} = 0 \quad (8.125c)$$

The last terminal condition is that $V_a(0)$ must equal the applied voltage at port 1 which is V_x . Hence

$$\hat{V}_a(0) = K_s^+ + K_s^- + V_{o+} + V_{o-} \quad (8.125d)$$

From the above four equations we can readily solve for V^* , V_s , V^* and V_s . The input admittance at port 1 is given by

$$\begin{aligned} \frac{Y_s}{V_s} &= \frac{h - Y_s \{V_s^+ - V_r^-\} + Y_s \{V_s^+ - V_{o+}\}}{V_t + V_s^+ + V_s^- + V_{o-}} \\ &= -j \cot \theta \end{aligned} \quad (8.126)$$

When we place an open circuit at port 2, the terminal condition $V_b(l) = 0$ given by (8.1256) is replaced by $Y_b(l) = 0$ or

$$Y^* \{V_s^+ e^{j\theta} - V_{o+} e^{j\theta}\} - Y^* \{V_s^+ e^{-j\theta} - V_{o+} e^{-j\theta}\} = 0$$

The other terminal conditions remain the same. After solving for voltage amplitudes for this case and using (8.126), we find that

$$Y^*_{in,oc} = \frac{J}{2(Y_J + Y_0^h)}$$

$$\frac{Y^*_{in,oc}}{Y_b} = \frac{C_T X_t}{(Y_J + Y_0^h)} \ll 1$$

For the network shown in Fig. 8.42c, a straightforward evaluation I that when port 2 is short-circuited

$$Y_{in,oc} = -J(Y_i + Y_2) \ll 1 \quad (8.128)$$

and when port 2 is open-circuited

$$Y_{in,oc} = \frac{Y_i \tan \theta}{Y_i + Y_l + Y_s} \ll 1 \quad (8.129)$$

From a comparison of (8.126) and (8.128a), we see that

$$Y_1 + Y_2 = y_{\infty}'' + Y_f$$

If we excite the networks at port 2 and place a short circuit or an open circuit at port 1, the expressions for Y_{in}^a and $Y_{m\text{dB}}$ are the same as given by the above equations but with the superscripts a and b interchanged and the subscripts 1 and 3 interchanged. Thus, for the short-circuit condition at port 1, we will find that

$$\mathbf{F}_i + \mathbf{y}_{\infty} - \frac{Y^h + Y^*}{Y_n^h + Y^*}$$

Hence we have

$$\begin{aligned} * & i - \bullet \frac{\mathbf{N} + Y_{\infty}''}{\mathbf{Z}} \bullet - * i \\ \mathbf{y}_3 & = - \frac{Y_n^h + Y^*}{Y_n^h + Y^*} \mathbf{T} \mathbf{A} - \mathbf{y}. \end{aligned}$$

When we use the expression for $V_{in} + Y_{in}$ in the coefficient for $\tan ti$ in (8.1286), we easily find that

Thus we have shown that the circuit in Fig. 8.42a is equivalent to that in Fig. 8.42c. For the special case when the coupled strips have the same width, $Y_e^b = Y^o = Y_c$ and $Y^* = Y^o = Y_0$ and we obtain

$$Y_1 = \frac{Y - Y}{g - \pm} \quad (8.129a)$$

$$Y_z = Y_3 = Y_e \quad (8.129b)$$

Consider now the filter structure shown in Fig. 8.43a. By replacing each coupled strip pair by its equivalent circuit in the form shown in Fig. 8.42c, we arrive at the circuit shown in Fig. 8.43b. This circuit is readily reduced to the one shown in Fig. 8.43c by combining the admittances of the stubs that are connected in parallel. A Chebyshev filter based on the circuit shown in Fig. 8.43c is readily designed.

For parallel coupled filters using microstrip construction, it is desirable to use open-circuited coupled microstrip sections in place of short-circuited coupled sections. A parallel coupled filter using open-circuited coupled transmission-line sections is shown in Fig. 8.44a. The equivalent circuit for this filter is shown in Fig. 8.44b. A basic coupled section is shown in Fig. 8.44c and its equivalent circuit is shown in Fig. 8.44d.

The equivalent circuit for the basic section is readily obtj set of equations (8.124) through (8.129). If we regard all voltage v representing currents, current variables as representing voltages, voltage_s^{**} , admit.^{**} tances as representing impedances, and open circuits/short representing short circuits/open circuits, then all the equations and "1" final conditions remain the same. Thus we obtain immediately termi

$$Z^a = \sqrt{(z^a_0)(Z_e^h Z)} \quad (8.130a)$$

$$Z_i = \sqrt{Zl + Zl} \sim Z, \quad (8.130b)$$

$$Z_z = |W_+ + Z_-| \sim Z_l \quad (8.130c)$$

and when $Z_{\infty} = 2^* Z_{f_t} = Z$,

$$Z_s = Z_{\infty} - Z, \quad (8.131a)$$

$$Z_s = Z_{\infty} = Z, \quad (8.131b)$$

The equivalent circuit shown in Fig. 8.44b for the filter is obtained by replacing each coupled section by its equivalent circuit shown in Fig. 8.44a" and combining the adjacent series-connected transmission-line stubs. This filter and its equivalent network is the dual of the one shown in Figs. 8.42 and 8.43. In extracting the square root of (8.130a) for the case of symmetrical strips, we choose the positive root which is the one given by (8.131a) since $Z_s > Z$.

For both of the parallel coupled filters described, each section is $A/4$ long at the center frequency and hence each transmission line is $A/2$ long. In the equivalent circuits of these filters, we have transmission-line stubs separated by $A/4$ transformers. These quarter-wave transformers function as nonideal admittance and impedance transformers. In the circuit shown in Fig. 8.43c, each short-circuited stub is the approximate equivalent of a parallel tuned LC resonator, while in Fig. 8.44b each open-circuited stub is the approximate equivalent of a series tuned LC resonator.

Equations for designing Chebyshev bandpass filters, with bandwidth up to one octave, using parallel coupled strips have been derived by Matthaei.^t The filter design equations were obtained by setting the i impedance and phase of each section of the network shown in rig- equal to that for a bandpass-filter prototype circuit of the form \backslash_e . Fig. 8.29 at the center frequency ω_0 and at the passband edge $w^m = w_+$. The design equations for the filter circuit shown in Fig. 8.44b are the dual of those for the filter circuit in Fig. 8.43c. We will give the equations below without derivation.

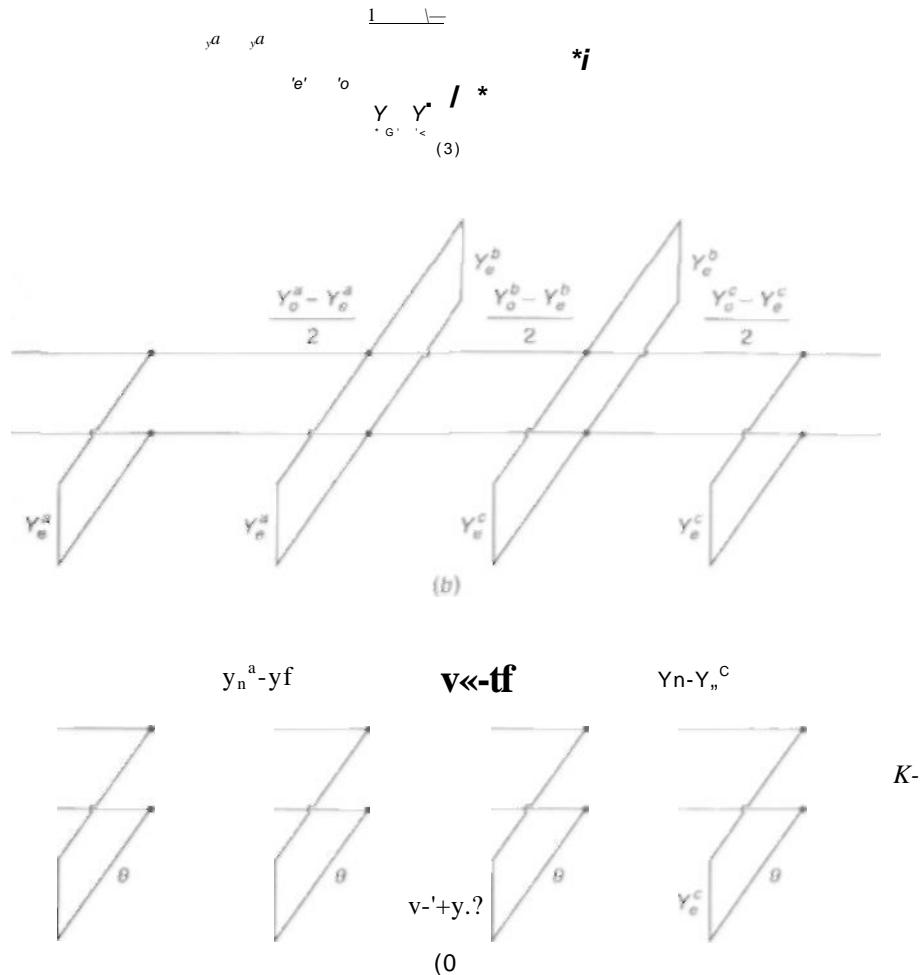


FIGURE 8.43

(a) A parallel coupled filter; (b) equivalent circuit; (c) reduced equivalent circuit.

For a Chebyshev filter with N sections, there are $N + 1$ impedance inverters and $N + 1$ even- and odd-mode line impedances to specify. The filter is assumed to be terminated in input and output lines with characteristic impedance Z_c . Each resonator has an electrical length of IT at the center frequency co_0 . The frequency at the lower edge of the passband is w , and $fil_e = 0j$ at this frequency, where l_c is the effective length of each resonator after correcting for the capacitive end loading at each open-circuited end. The impedances of the input and output impedance inverters are

given by

$$K_{in} = K_{N-1, N} Z_c \quad (8.132a)$$

The following parameters are also required:

$$\begin{aligned} & TT_{to}, \\ & \gg 1 - 2w_0 \end{aligned} \quad (8.132b)$$

$$P \sin \theta_I = \frac{K_{lo} Z_c}{[U_{an} O_{-h} (K_{lo} Z_c)^{f_1 V^2}]} \quad (8.132c)$$

$$S = I_{tan ft.} + CX^2/Z_c^2 \quad (8.132rf)$$

$$Z^* = Z f^{+1} \wedge (1 + P \sin f_1 J) \quad (8.132e)$$

$$Z_{l1} = Z^* \wedge f_c (1 - P \sin 0) \quad (8.132f)$$

The remaining impedance inverters and even- and odd-mode impedances are given by

$$Z_c = k_{A-l,i} \quad k = 1, 2, \dots, N-1 \quad (8.133o)$$

$$N_{t+i,ft} = \frac{K}{-2} + 5 \tan^2 \theta_x \quad (8.1336)$$

$$(8.133c)$$

$$(8.133rf)$$

In the above equations the g_k are the element values for a loss-pass prototype filter with cutoff frequency $c_0 = 1$, and are given by (8.101).

For the dual network all $K_k + u_k$ are replaced by $\langle * + i * \rangle d \wedge$ replaced by Y^* , and all 2^* are replaced by Y_c^* .

In the next example we will illustrate the application of the t equations to the design of a two-section Chebyshev filter.

Example 8.3 Parallel coupled Chebyshev bandpass filter. The filter will be designed for a 30 percent bandwidth, a passband tolerance of 1%, and a center frequency of 6 GHz. The input and output lines have a characteristic impedance of 50 ohms. The filter is illustrated in Fig. 8.4-5a.

The element values for the low-pass prototype circuit are using (8.101) and are $g_0 = 1$, $g_1 = 2.33754$, $g_2 = 0.62634$, and $g_3 = 3.7321$. The frequency at the lower band edge is $f_0 = 0.15 f_c = 0.15 \times 6 \times 10^9 = 9 \times 10^8$ Hz. The value of ω_0 is $0.85n/2$ and $\sin \theta_0 = 0.97237$, $\tan \theta_0 = 4.165$. By using (8.133) we obtain $K_{lo} = 0.65406 Z_c$, $K_{2l} = 0.826447 Z_c$, $P \sin \theta_0 = 0.03$.

-26-

fa)

$$7^{a*} \quad 7^{\circ} \quad z^c + z''$$

$$Z-z; \quad z^b - z'' \quad z^c - z^c \quad zi-4$$

m

$$|A \quad Id)$$

FIGURE 8.44

(a) Parallel coupled filter using open-circuited coupled-line sections; (b) equivalent circuit of filter; (c) basic coupled-line section; (d) equivalent circuit of basic section.

$s = 19.9165$, and $JV_{12} = 2.8865$. The following even- and odd-mode impedances are found

$$Z_j = 70.64 \quad ZJ = 29.36 \quad Z^2 = 73.95 \quad Z^2 = 41.03$$

In the equivalent transmission-line circuit shown in Fig. 8.456, the line impedances are found to be $Z_1 = Z_a$, $Z_2 = (Z_f - Z_{a'})/2$, $Z_3 = Z_{a'} + Z_f$, $Z_4 < (Z^2 - Z^2)/2$, $Z_5 = Z_b$, $Z_6 = Z_2$, and $Z_7 = Z_c$. For this filter it is a straightforward transmission-line circuit analysis problem to evaluate the impedances $Z_m = Z_{a'}$ at the center of the filter and evaluate the impedance mismatch $(R_{in})^2/ZJ$. In Fig. 8.46 we show the computed filter response with that of the Chebyshev filter with power loss ratio $1 + k^2 T\% (o>')$ and using $to' = (u) - u>_0/io_0$ as the frequency variable.

The design equations are approximate ones only. Since we are specifying a 30 percent bandwidth for a filter using only two sections, it is not surprising

(a)

$$Z_r = Z_s \quad Z_7 = Z_1$$

$$Z_e = Z_s \quad (b)$$

FIGURE 8.45

(a) A two-section parallel coupled filter; (b) equivalent transmission-line circuit.

that the actual filter performance does not quite meet the specifications. The response curve shown in Fig. 8.46 indicates that the actual filter has a 25 percent bandwidth but a more rapid increase in attenuation than the ideal Chebyshev filter has.

We have also designed a three-section filter with the same passband tolerance and bandwidth. The calculated parameters are $g_1 = g_3 = 2.5547$, $\kappa_w = \kappa_M = 0.62565$, $K_{2l} = K_{32} = 0.672$, $P \sin 0_\theta = 0.39776$, $s = 20.21$, $Z_1 = 69.89$, $Z_2 = 30.11$, $Z_e^2 = Z_f = 57.807$, and $Z_f = Z_0^3 = 30.645$. In Fig. 8.47 we show the computed response of this filter. The agreement is now much closer to that of the corresponding three-section Chebyshev filter. The difference in the attenuation does not exceed 0.3 dB over the frequency range 0.7/ f_n to 1.3/ f_n . At the band edges where $f = 0.85f_n$ and $1.15f_n$, the parallel coupled filter has an attenuation of 2 dB instead of the theoretical design value of 1.761 dB, which is a difference of only 0.24 dB.

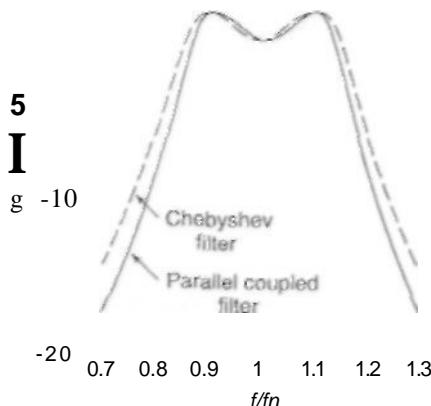
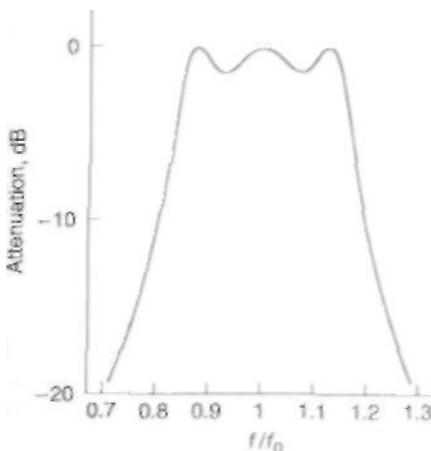
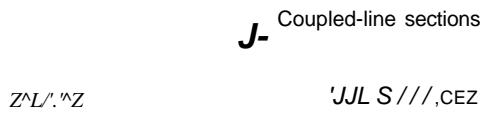


FIGURE 8.46
A comparison of the response of the two-section parallel coupled filter with that of the ideal Chebyshev filter. The $P_{\text{attenuation}}$ is 0.3 dB and the fractional bandwidth

**FIGURE 8.47**

Response of a three-section parallel coupled filter. The passband tolerance is 1.76 dB and the fractional bandwidth is 0.3.

**FIGURE 8.48**

Shielded suspended microstrip line used for constructing parallel coupled filters.

The design formulas for parallel coupled filters are based on the assumption that the even- and odd-mode phase velocities are equal. Thus, if conventional microstrip-line construction is used, it would be necessary to use a dielectric overlay so as to equalize the two phase velocities. The parallel coupled filter can also be built using strip-line or shielded suspended microstrip line as shown in Fig. 8.48. The suspended microstrip line employs a thin low dielectric constant sheet on which the transmission-line sections are printed. Since the dielectric sheet is thin, it has about the same effect on the phase velocity of both the even and odd modes.

For parallel coupled strips a ratio Z/Z_n less than 3 is needed in order to avoid very close spacing between the strips. For the filters discussed in Example 8.3, the largest ratio was 2.4, so that these filters can be built using edge-coupled strips. When tighter coupling is required, the adjacent strips can be overlapped by printing every other transmission-line section on the opposite sides of the supporting dielectric sheet.

QUARTER-WAVE-COUPLED CAVITY FILTERS

Quarter-wave-coupled cavity filters are similar to the filter discussed in the preceding section except that the transmission-line stubs are replaced by cavities. The filter is realized in practice by placing diaphragms in a waveguide. To understand the basis for design in the narrowband case, we must first consider the equivalent circuit of a section of waveguide loaded with two identical diaphragms a distance l_k apart, as shown in Fig. 8.49.

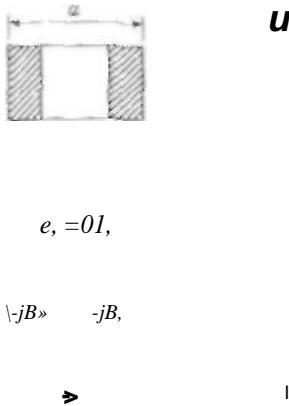


FIGURE 8.49

Ia) Rectangular waveguide loaded with twr, inductive diaphragms to form a cavity, (6) exact equivalent circuit. (\leftrightarrow) approximate equivalen! circuit.

For a waveguide, the important frequency variable is not cu but rather $(\beta/k_0)u = fie$, since waveguide diaphragms have susceptances that vary very nearly as β or jl' and the electrical length of a section of guide is proportional to fi . The normalized frequency variable $to/w_0 = A_0/A$ is therefore replaced by $A_{g0}/A_g = \beta/\beta_0$, where λ_{g0} is the guide wavelength at $w = \omega_0$ and A_g is the corresponding value at any w . Consequently, in all design formulas, we replace w by βc , where c is the velocity of light.

The exact equivalent circuit for a waveguide loaded with two identical inductive diaphragms with normalized *susceptance* $-jB_k$ is shown in Fig. 8.496. For filter design according to the methods developed in preceding sections, we must replace the exact equivalent circuit by an approximate shunt circuit. Mumford has shown that an equivalent circuit of the form illustrated in Fig. 8.49 c has the same frequency characteristics as the exact equivalent circuit has over a narrow band of frequencies around $\omega_0 T^*$. The results obtained by Mumford are presented here without derivation. The derivation is straightforward, and may be found in Mumford's paper. The shunt susceptance B is expressed in the form

$$B =$$

where $A/3 = 0 - \beta_0$ is small. When a resonant circuit of this OT* ^ connected across a transmission line, it is loaded by a shunt condu

27. PP

normalized value unity on each side. The loaded Q of the circuit is thus

$$Q_L = \frac{1}{\gamma} (\gamma_0 C) C = \frac{1}{\gamma} fC \quad (8.134)$$

since $\gamma_0 C = (LC)^{-1/2}$. Hence we may express B in terms of the loaded Q : thus

$$B = 4Q_k - \frac{Po}{\gamma} \quad (8.135)$$

The value obtained for Q_k by Mumford is

$$\begin{aligned} Q^* &= \frac{j\pi \tan^{-1}(2/B_t)}{2 \sin(\gamma + 4lf)^{1/2}} \\ &= \frac{(B_t! + 4B_t^2)!}{4} \sim \frac{\tan^{-1}\gamma}{B_k} \quad (8.136) \end{aligned}$$

since B_t is large compared with unity for a narrowband (high-Q) filter. The required diaphragm spacing γ to give perfect transmission through the cavity at $\omega = \omega_0$ is given by

$$tmf_i ol_i = -w \quad (8.137)$$

The two sections of line with electrical length θ_{lk} in the circuit of Fig. 8.49c are to be chosen so that

$$/M_A + 2e_u = \xi_A + 20_{lf} = TT \quad (8.138)$$

at the frequency ω . These additional lengths of line in the equivalent circuit of a single cavity are absorbed into and made part of the quarter-wave coupling lines in the filter.

The design of maximally flat and Chebyshev filters with N odd is straightforward. If the prototype circuit of Fig. 8.29 is used, it is only necessary to make

$$Q_k = \frac{1}{\gamma} \frac{C_{kk}}{\gamma - \gamma^*} \quad <8.139>$$

and to choose C_{ok}, L_{ok} so that $C_{ok}L_{ok} = (0_o c)^{-2}$ and all J_{kk} , equal unity. The section of waveguide between cavity k and $k + 1$ has an electrical length equal to $\pi/2$. Since this includes tt_{UL} , and θ_{lk} from the adjacent cavities, the physical length of the quarter-wave coupling line between

-JS,
 $r^{\Lambda}rr$

FIGURE 8.50
 Quarter-wave-coupled waveguide-cavity filter.

cavities k and $k + 1$ will be

$$\frac{1 - iir}{\frac{*_g o I *k + h}{2} \frac{1}{2T} \frac{1}{2}} = \frac{0 \gg H ij}{T_4^{A_4 0}} \quad (8.140)$$

upon using (8.138). A schematic illustration of the filter is given in Fig. 8.50. Formulas for the required diaphragm dimensions to yield the specified value of $B_{c,}$ are given in Sec. 5.8. The power loss ratio for the filter is obtained by replacing $i\omega/(a_c)$ by $/3//V$ For a Chebyshev filter it is given by [see (8.104) and (8.109)]

$$i_u = 1 + k^2 T^2 \frac{P_o}{A} \frac{P_-}{-ft} = 00 \quad (8.141)$$

where f_2 and $/3$, are the values of $/3$ at the edges of the passband. If $/3$, and 13_c , are specified, then

It should also be noted that for the waveguide filter, w_0 , $W]$, and $a)$ design formulas (8.111) and (8.112) must be replaced by $V - Pi^c < a^n d & f'$ where c is the velocity of light; i.e., replace $k_0 = co/c$ by $/3$. To illustrate the procedure, we shall evaluate the required susceptance $-jB_j$ for the cavity in the filter shown in Fig. 8.50.

Assume that a five-element filter will be needed. The response is of the Chebyshev type with a passband tolerance $k^2 = 0.0233$. The waveguide to be used has a width of $a = 0.9$ in. The passband is to extend from $\lambda = 10,000$ MHz to $f_2 = 10,400$ MHz. The corresponding values of co/c are 2.1 and 2.18 rad/cm. The values of 0 , and $/3_2$ are thus

$$fix = (2.1) - \frac{| - |}{a} = (4.4 - 1.89)^{1/2} = 1 - 5.9$$

$$0a = [(2.18)^2 - 1.89]^{1/2} = 1.7$$

The center of the band occurs at $p_0 = (j6/3_2)^{1/2} = 1.64$, which gives $\gamma'' = 10,200$ MHz. From Table 8.2 we find $g_x = 1.1468$. Using (8.1126) gives

$$' \quad 0_2 - \text{ft } /V \quad 0_2 - 0, f_0C$$

Since J_{00} is to equal unity and $R_L = i\gamma_{00} = 1$, we have $C_{00} = CJ$ from (8.118a). Using (8.139), we obtain

For the feth resonator we should obtain

$$Q - JK - 4L \quad (8.143)$$

For Q, we obtain 8.56. and from this result we can determine B, by using (8.136). For S_A large, we can replace $\tan^{-1}(2/B_k)$ by $2/B_k$, and we then find that

$$B_s = 2(Qf - 1)''' \quad (8.144)$$

Thus we find that $B_s = 17$. The required diaphragm dimensions can now be determined, and also the cavity length from (8.137). The above procedure has to be repeated for each cavity in the filter chain.

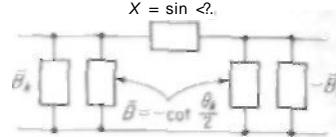
DIRECT-COUPLED CAVITY FILTERS

Direct-coupled cavity filters have the advantage that the physical structure is more compact than the corresponding quarter-wave-coupled cavity filter. A design procedure for direct-coupled cavity filters that is accurate for bandwidths up to 20 percent has been developed by Cohn.^t Cohn's design method is based on the use of the network in Fig. 8.28 as a prototype. The design formulas will be presented here without the detailed derivation.

The waveguide cavity and its equivalent circuit shown in Figs. 8.49a and b may also be represented by a II network shunted with inductive susceptances at each end, as in Fig. 8.51. The two shunt susceptances $B = -\cot(0_s/2)$ may be neglected^t compared with B_y since B_k will be large, and 0^s is nearly equal to $-r$; so B is small compared with unity. The series arm X is thus used as the series resonant circuit in the prototype filter.

For impedance inverters Cohn uses the shunt inductive reactance plus two short sections of waveguide (equivalent transmission lines), as shown in

^tS. B. Cohn, Direct Coupled Resonator Filters, *Proc. IRE.* vol. 45, pp. 187-196. February, 1957.



1a)

FIGURE 8.51
A waveguide cavity and its equivalent circuit.

Fig. 8.52. For this circuit the impedance inverting properties are obtained "r

$$k = \frac{1}{2} \tan^{-1} \frac{i}{B}, \quad (8.145a)$$

$$B, \quad K \quad (8.1456)$$

where K is the characteristic impedance of the quarter-wave impedance inverter. With θ_0 and B_k determined at a frequency w_0 , it is found that the inverter does not depart appreciably from its ideal characteristics over a 20 percent band.

In the vicinity of $t_0 = t_0$, where $\theta_0 = v$, the series reactance X behaves as

$$X = \sin \theta_k = \sin(\theta_0 + tr + TT)$$

$$= -(0, -77) = -(p - 0_0) / = \text{ft, -?} \quad rr(0) \quad \text{aIA}$$

where $/3_0/ = IT$. This frequency behavior is similar (apart from the sign, which is immaterial) to that for a series resonant circuit for which $X = jL/C(a>/u>_0 - \ll, / \ll) = 2/L/C(a> - t_0)/t_0$ if $\ll, / \ll_0$ is replaced by H new frequency variable $/3//3_0$.

When the negative line lengths of the impedance inverters are absorbed as part of the cavity length, the physical length of the k th inverter becomes

$$l_i = A * o + r \ll + e \quad ^8-146)$$

FIGURE 8.52
An impedance inverter.

$f \wedge H$ $-SM$

$h \ll -M -0i,$



FIGURE 8.53
A direct-coupled waveguide-cavity filter.

In the prototype circuit of Fig. 8.28, we must choose all \hat{L}_{ok}/C_{0l} equal to $7r/2$ to obtain a correspondence with the type of series resonant circuit employed here. In addition, we choose

$$\hat{O}k\hat{O}k \sim (\$0^c)^{-2} \quad (8.147)$$

The impedance-inverter parameters as given by (8.115) thus become known in terms of the C_k and L_k which are related to the g_{lt} in the low-pass prototype. From the known values of the $Z_t^{\wedge, l}$, the shunt susceptances B_k may be found. The filter is illustrated schematically in Fig. 8.53. The design formulas obtained as outlined above are

$$\frac{1}{yWtfi} = \frac{1 - w/g_l}{w} \quad (8.148a)$$

$$B_2 = \frac{1}{w} \quad (8.148b)$$

$$g^* = \frac{yJSkSk-i}{gkSk} \quad (8.147c)$$

$$N = \frac{I}{R} = \frac{wR/g_s}{U'R/gN-l} \quad R = 1 \text{ for } N \text{ even} \quad (8.148rf)$$

where

$$w = \frac{P-i-Px}{2 da}$$

and the g^* are the element values from the low-pass prototype filter. Note that $R = 1$ for N even, and also for N odd in the case of maximally flat filters; otherwise R is given by (8.98). The length of the A th cavity at $/3 = 0o$ is

$$* = \frac{-go}{2} \quad \frac{"go_i}{4\pi I} \quad \rightarrow B_{*+i} + \tan^{-1} \frac{2}{B_{*+i}} \quad (8.149)$$

The power loss ratio is given by substituting

$$\mathbf{P}_0$$

$$P_s - P_i \quad P_o \quad P$$

for w' in the low-pass prototype filter response.! Note also that $8^a -$

$\text{oifl}_2.$

8.23 OTHER TYPES OF FILTERS

In the preceding sections we have discussed only a few specific filters in order to illustrate the insertion loss method of filter design. In addition to the maximally flat and Chebyshev-type filters, there exist two other types that have the feature that, in addition to producing equal-ripple response in the passband, they produce a number of attenuation poles outside the passband. These filters are the elliptic filters and the generalized Chebyshev response filters. By producing attenuation poles, i.e., frequencies at which the attenuation is infinite, the skirts of the filter response curve are much steeper so the attenuation will increase much faster outside the passband. It is much more difficult to find microwave circuit components that can be used to implement elliptic-type filters so these filters are not very commonly used.

Another important filter topic that we have not addressed in the phase response of a filter. In order that the filter should not produce a distortion of the signal transmitted through it, the amplitude response should be independent of frequency and the phase response should be a linear function of u so that the group delay will be a constant. If we express the transmission coefficient of the filter in the form $A(w)e^{j\phi(w)}$, then the time delay experienced by a signal propagating through the filter is given by $d = b/dw$ as shown in Sec. 3.19. Thus $\phi(w)$ should be a linear function of w so as to give the same time delay for each frequency component of the signal. Real filters only satisfy this criterion in an approximate way. For critical applications, especially when broadband signals are involved, it may be necessary to insert a phase equalization circuit in series with the filter.

The reader is referred to the references at the end of this chapter for a discussion of a number of filter topics that we could not include because of space limitations.

PROBLEMS

- 8.1. Find the $\mathbf{Rf}^* \mathbf{S}^*$ matrix for the following networks: (a) a shunt, S_y , series jB . (b) a series reactance jX , (c) a shunt reactance jX_x followed by a reactance jX_2 .

Cohn uses a somewhat different frequency variable, which, however, is considered, is essentially the same as we have used.

(or the $b_i^* d_i^*$)

- 8.2. Derive the relations (8.14) and (8.15).
- 8.3. Consider a T network terminated in a load Z . Evaluate the input impedance Z_m and show that the condition that Z transforms into $Z_{m\bar{m}}$, that is, $Z_m = Z_{m\bar{m}}$, leads to the characteristic values Z_S for the periodic structure.
- 8.4. Show that the eigenvalue equation for the propagation constant of a Bloch wave on a transmission line loaded at intervals d with a series reactance jX is $\cosh yd = \cos k_d d - (X/2)\sin k_d d$.
- 8.5. Show that (8.21) may be expressed in the form

$$\cosh yd = \frac{\cos(\beta d)}{1 + 2j}$$

where β is the phase angle of S_{12} , and S_n is the scattering-matrix off-diagonal element for the unit cell (Sec. 4.8).

- 8.6. Express λ_B in terms of the A_u by noting the similarity between (8.20) and (8.6) and that between I''_e and Z_B .
- 8.7. Show that the wave-amplitude transmission matrix for a shunt susceptance jB is

$$2+jB \quad B$$

$$\begin{matrix} [A] & B & 4 + B^2 \\ j2 & & 2(2 + jB) \end{matrix}$$

- 8.8. A load Z_L on an ordinary transmission line gives a reflection coefficient $R_L = (Z_L - V)/(Z_L + V)$. Show that (8.30), giving V_L fw a Bloch wave, may be expressed as

$$\mathbf{r}_B = \frac{\mathbf{i}^* \mathbf{r} \mathbf{j}}{\mathbf{i} + \mathbf{J}^* \mathbf{i} \mathbf{l} - \mathbf{r}}$$

where r_B are the characteristic reflection coefficients of the component waves making up the Bloch wave.

- 8.9. For Bloch waves in the capacitively loaded coaxial line, show that the TEM voltage waves between any two consecutive diaphragms are given by

$$V = y_{-e^-} /*(iU - e^{jk_z z}) + y_{+e^+} /*(iU + e^{jk_z z}) - y_{-t-g} /*(n - e^{jk_z z}) + p_i y_{-g} /*(1 - e^{jk_z z})$$

for the Bloch wave propagating in the $+z$ direction, and

$$V = f_y /*(J^* f) + V_{LI} V; e^{jk_z n z}$$

for the Bloch wave propagating in the $-z$ direction. The zeroth terminal plane has been chosen as $z = 0$ and $V_B = V*(l + V')$, $V_B = V_0(l + T_f)$.

- 8.10. Consider an infinite transmission line loaded with shunt capacitive susceptances jB at $z = nd$, $n = -\infty$ to ∞ . Show that the current and voltage waves that make up a Bloch wave are given by

$$V_B = V^* e^{-jk_z n z} + V^- e^{jk_z n z}$$

$$I_B = I^* e^{-jk_z n z} + I^- e^{jk_z n z} = V^* e^{-jk_z n z} - V^- e^{jk_z n z}$$

where $V = -V(1 - e^{-Jz})/(1 - e^{2\pi f_0 d})$ and $0 = k_0 d$. Let $V_p(z) = e^{-Jz}$ and expand $V_p(z)$ into an infinite series of spatial harmonics. Then the relation between V^* and K_{∞} may be obtained by using the convolution integral $\int_0^\infty V_p(z) e^{-Jkz} dz = \frac{1}{J} V^*$ and that $V_{\infty}(z = d) = e^{-Jk_0 d} V^*$ and that V^* is given by (8.10a).

- 8.11.** For the sheath helix show that the eigenvalue equation for the n th mode is

$$\frac{I' i V + n f i a \cot \theta}{(k, ka^2 \cot \theta / l!)^n} = \frac{K_n(ha)}{l_n(ha)}$$

- 8.12.** Consider an N -section filter made up of a capacitively loaded coaxial line unit cells. The filter is terminated in a resistive load R equal to the characteristic impedance at zero frequency, i.e., equal to 1. The generator at the input has an internal resistance equal to R . Show that the power delivered to the load is given by

$$P = \frac{V^2}{I^* Z_J + Z_{in}}$$

where V is the generator voltage and Z_{in} is given by

$$Z_{in} = \frac{R + Z_J \tanh y N d}{R + Z_J + R \tanh y N d}$$

and Z_J , $y d$ are the image parameters at any frequency. In the passband where Z_J is real and $\tanh y d = j \tan Npd$, verify that

$$\frac{V^2 Z_J R / (1 + t^2)}{4 R^2 + (R^2 + Z_J)^2 t^2} = \frac{V^2 (l + t^2)}{l R^2 - Z^2} \quad t = \tan Npd$$

Thus show that the power loss ratio becomes

$$\frac{P_{un-}}{P_{in-}} = 1 - \frac{R^2 - Z_J^2}{2 R Z_J} \sin^2 Npd$$

Plot P_{un-}/P_{in-} as a function of N in the passband for the case where $R = 1$, and $N = 4$. See (8.10a) and (8.16) for expressions giving Z_J .

$$R = \frac{\sinh^2 Npd}{\sinh^2 Npd + 1}$$

- 8.13.** (a) For the two circuits shown in Fig. PS.13a show that the mismatch is the same.
 (b) Similarly, show that the two circuits shown in Fig. P8.13b have the same insertion loss when an $K = Z_{in}$ or $J = Y_t$ is inserted at either end of the filter.

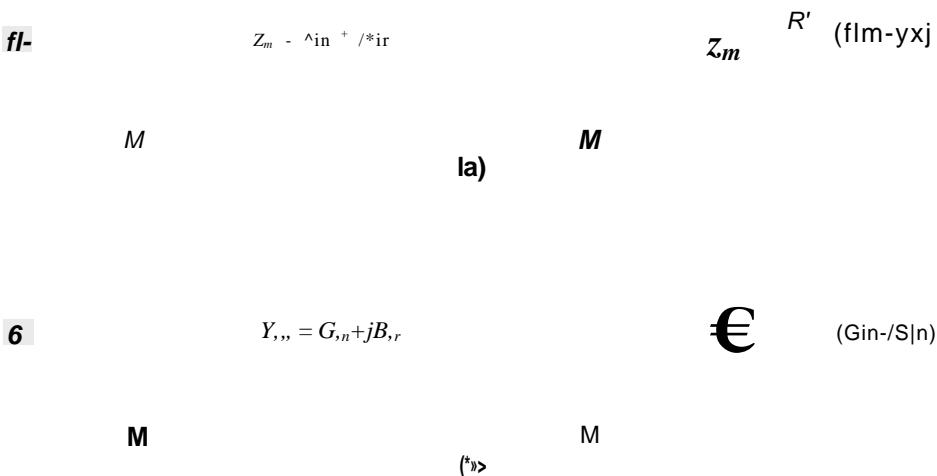


FIGURE P8.13

- 8.14.** Design a two-section lumped-element bandpass filter similar to that in Example 8.1. Assume that $k^2 = 1$ and that the bandwidth equals $0.1 f_{\text{c}}$. The terminating resistances equal 1,000 Ω and $f_0 = 1$ MHz. Verify that the coupling coefficient defined in Example 8.1 equals $(1 + \sqrt{2})/Q$, where Q is the loaded resonator Q .
- 8.15.** Design a three-section lumped-element filter of the form shown in Fig. P8.15. Assume a passband tolerance $k' \approx 0.5$, a bandwidth equal to $0.1 f_{\text{c}}$, terminating resistances of 1,000 Ω , and a center frequency of 455 kHz. Plot the filter response using $P_{\text{r}} = 1 + k^2 T(u)$, where u is given by (8.109).

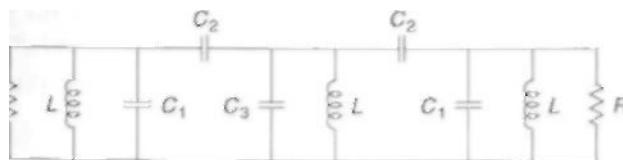


FIGURE P8.15

- 8.16.** Design a two-section half-wave filter with a 5 percent bandwidth and a passband tolerance $k^2 = 0.2$. The frequency of operation is 5 GHz. The filter uses 50-ft transmission lines. Specify the three required gap capacitances.
- 8.17.** Design a three-section half-wave filter with the same specifications as in Prob. 8.16. Specify the required gap capacitances.
- 8.18.** Design a two-section parallel coupled filter with a bandwidth of $0.1 f_0$ and a passband tolerance $k^2 = 0.2$. Specify the required even- and odd-mode characteristic impedances for each section. Assume that $Z_r = 50$ Ω .

- 8.19. Verify the parameter values given for the three-section parallel ^{coupled} filter discussed in Example 8.3.
- 8.20. Design a three-section parallel coupled filter with a bandwidth $\Delta f = 10\text{ MHz}$ and a passband tolerance $k^2 = 0.1$. Specify all even- and odd-mode ^{harmonic} impedances. Assume $Z_c = 50 \Omega$.
- 8.21. Design a three-cavity quarter-wave-coupled filter with the following specifications: waveguide width $a = 0.9$ in, band edges at $f_1 = 10,000 \text{ MHz}$, $f_2 = 10,400 \text{ MHz}$, passband tolerance $It^2 = 0.0233$. Chebyshev response. Inductive diaphragms with circular holes are to be used. Determine the hole radii and diaphragm spacings.
- 8.22. Design a four-cavity direct-coupled cavity filter having Chebyshev response. The passband tolerance is $k^2 = 0.0233$, band edges occur at $f_1 = 9500 \text{ MHz}$, $f_2 = 10,500 \text{ MHz}$, and the guide width is 0.9 in. Specify the diaphragm dimensions and spacing. Use any convenient inductive diaphragm.
- 8.23. Design a four-cavity direct-coupled maximally flat waveguide filter with the specifications given in Prob. 8.22. Note that the maximally flat filter designed from the low-pass prototype has a passband tolerance of 1. To obtain a passband tolerance of It^2 between 0_1 and 0_2 , the design must be carried out for a wider bandwidth, say $0'_1$ to $0'_2$. Thus we should have

$$\begin{matrix} \text{flo} [& 0 & 0 \\ & 0 & 0 \\ 02 - 0'i & \backslash & 0o \end{matrix}] = 1 \quad \text{for } 0 = 0'_1, 0'_2$$

Also $0'_1, 0'_2 = It^2$. Determine 0_1 and 0_2 so that

$$\begin{matrix} 00 \\ & 0 \\ 02 - 0', 10 \end{matrix} = k^2 \quad \text{when } 0 = 0, \text{ and } 0_2$$

- If the design is carried out using these values of $0'$, and $0'_2$, the required passband tolerance of k^2 will be maintained in the passband between 0_1 and 0_2 . Show that, in general, $0'_1, 0'_2 = 0, 0_2 = 0$ and $0'_2 - 0' = (0a - 0i)^{1/2}$
- 8.24. For the circuits in Figs. 8.496 and c show that the normalized input impedances are given by

$$(l-B^2 \sim 2B, \cot \theta_k) \quad \text{and} \quad \frac{2sm28_u + \cos 2Bu^2 B}{2\sin 20j^* + \cos 20u + BI} \quad - \sqrt{l/2}$$

When $w = i\omega_Q$, $B = 0$, show that B_k must be related to θ_k by $\tan^{-1} \frac{B_k}{Z}$, equal to unity. Show that the image phase constants ϕ for the are given by

$$\cos \phi = \cos Q_k \quad B_k \sin fl_k \quad \text{and} \quad \cos \phi = \cos 2 \pi n^2 \cdot Y \sin 2 \pi n^2 \cdot Y$$

and will be equal at w_0 if $\theta_A + 2\theta_u = IT$.

;FEBENCS

Periodic structures

1. Brillouin, L.: "Wave Propagation in Periodic Structures," 2d ed.. Dover Publications, Inc., New York, 1953.
2. Slater, J. C: "Microwave Electronics," D. Van Nostrand Company, Inc., Princeton, N.J., 1950.
3. Watkins, I). A.: "Topics in Electromagnetic Theory," John Wiley & Sons, Inc., New York, 1958.
4. Bevensec, R. M.: "Electromagnetic Slow Wave Systems," John Wiley & Sons, Inc., New York, 1964.

Microwave filters

5. Rhodes, J. D.: "Theory of Electrical Filters," John Wiley & Sons, Inc., New York, 1976.
6. Malherbe, J. A. G.: "Microwave Transmission Line Filters," Atech House Books, Dedham, Mass., 1979.
7. Matthews, EL (ed.): "Surface Wave Filters," John Wiley & Sons, Inc., New York, 1977.
8. Alseyab, S. A.: A Novel Class of Generalized Chebyshev Low-Pass Prototype for Suspended stripline Filters," *IEEE Trans.*, vol. MTT-30, pp. 1341-1347, 1982.
9. Mobbs, C. I., and J. D. Rhodes,: A Generalized Chebyshev Suspended Substrate stripline Bandpass Filter, *IEEE Trans.*, vol. MTT-31. pp. 397-402. 1983.
10. Williams. A. E.: A Four-Cavity Elliptic Waveguide Filter," *IEEE Trans.*, vol. MTT-18. pp. 1109-1114, 1970.
11. Makimoto, M., and S. Yamashita: Bandpass Filters Using Parallel Coupled stripline Stepped Impedance Resonators, *IEEE Trans.*, vol. MTT-28, pp. 1413-1417, 1980.
12. Wenzel. R. J.: Exact Theory of Inter-Digital Bandpass Filters and Related Coupled Structures. *IEEE Trans.*, vol. MTT-13. pp. 558-575. 1965.
13. Cohn, S. B.: Microwave Bandpass Filter Containing High-Q Dielectric Resonators. *IEEE Trans.*, vol. MTT-16. pp. 218-227. 1968.

CHAPTER 9

MICROWAVE TUBES

9.1 INTRODUCTION

Microwave tubes are the prime signal sources in high-power radar systems. The magnetron is the tube most frequently used and *can* provide *many* kilowatts of continuous-wave (CW) output power and a megawatt or more of peak power with pulsed operation. Magnetrons are also used for industrial heating applications and in microwave ovens for consumer use. The traveling-wave-tube amplifier with power outputs up to 10 W or more is the workhorse in satellite communications. The klystron tube can function as an oscillator or as an amplifier. It can be designed for either low or high output power applications. In low-power applications the klystron was once widely used as the local oscillator in microwave receivers but has now been replaced by solid-state oscillators. Solid-state oscillators are replacing microwave tubes in many low-power transmitter applications also. Even though many of the applications for microwave tubes have been taken over by solid-state devices, the requirements for high power can only be met by microwave tubes, so they are an essential device for many systems.

Conventional low-frequency tubes, such as triodes, fail to P^e/I_w at microwave frequencies because the electron transit time from the cathode to the grid becomes an appreciable *fraction* of the period of the sinusoidal signal to be amplified. In other words, propagation times become significant, and the same limitations that are inherent in low-frequency tubes are present in low-frequency tubes also. Microwave tubes must be utilized to utilize the wave-propagation phenomena to best advantage.

Broadly speaking, there are two basic types of microwave tubes, klystrons and some magnetrons.

those that employ slow-wave circuits (traveling-wave tubes). Both types of tubes utilize an electron beam on which space-charge waves and cyclotron waves can be excited. The space-charge waves are primarily longitudinal oscillations of the electrons and interact with the electromagnetic fields in cavities and slow-wave circuits to give amplification. The properties of cavities and slow-wave circuits have already been discussed. What remains to be done is to examine the propagation of space-charge waves on electron beams and then to consider the interactions that take place between electron beams and the fields in cavities and slow-wave circuits.

The purpose of this chapter is to examine the nature of electron beams and the space-charge waves that they can support. In addition, the interaction of the beam with a microwave cavity or slow-wave circuit is to be studied in order to explain the operating principles of a number of different microwave tubes. Space does not permit a detailed treatment of the many different varieties of microwave tubes in existence. We shall concentrate on fundamentals that form, more or less, the basic operating principles of all microwave tubes.

Two approaches may be used in analyzing the dynamic behavior of the electron beam. The earliest approach used was the ballistic, or lagrangian, approach. In this method the motion of an individual electron is studied in detail, and it is assumed that all other electrons behave in a similar way. The ballistic approach has the advantage of permitting certain nonlinear, or large-signal, effects to be treated fairly easily.

The other approach is the field approach, sometimes called the eulerian, or hydrodynamical, approach. In this method the electron beam is essentially treated as a charged fluid. Field variables that describe the velocity, charge density, ac current, etc., at an arbitrary point as a function of time are introduced. However, no attempt is made to follow the motion of a single electron. The field approach, which leads to the space-charge waves, is more unifying and lends itself to the treatment of all different types of microwave tubes within the same general mathematical framework. Therefore only the field approach is used in this text.

An exact analysis of a microwave tube would be very difficult and laborious to carry out. As in any other physical problem, it is necessary to introduce a number of simplifying assumptions in order to arrive at a mathematical model that can be analyzed without too many complications. The success of a simplified theory must then be judged by the extent to which it predicts and agrees with experimental results.

The first few sections of this chapter discuss a number of models used for the electron beam and the propagation of space-charge waves on these beams. The governing equations are Maxwell's equations and Newton's laws, together with the Lorentz force equation. The equation of motion for a charge element is a nonlinear equation, but may be linearized by assuming small-signal conditions; i.e., all ac quantities are small compared with dc quantities. We shall consider only the small-signal situation since this will

suffice to develop the operating principles of microwave tubes. The analysis is a great deal more difficult, and the theory, in general, is not fully developed.

After treating the dynamics of the electron beam, the klystron traveling-wave tube are examined in detail. A number of other types are also discussed, but in a more qualitative way.

9.2 ELECTRON BEAMS WITH DC CONDITIONS

By means of a suitable electron gun consisting of a cathode, accelerating electrodes, and focusing electrodes, a beam of electrons with essentially uniform velocity v_0 can be produced.¹ Figure 9.1 is a schematic illustration² of a cylindrical electron beam with a radius a . If the potential difference through which the electron is accelerated is V , the velocity u_0 is given by

$$\frac{2Ve}{m} = 5.93 \times 10^5 V^{1/2} \text{ m/s} \quad (9.1)$$

where $-e$ is the electron charge and m is the mass of the electron. For $V = 1,000$ V, $v_0 = 1.87 \times 10^7$ m/s = 0.0625c. The beam perveance is defined by the quantity TV^{-2} , where T is the total beam current.

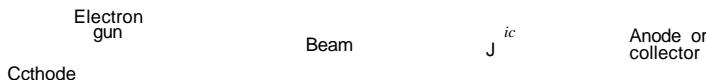
The coulomb repulsive force, or dc space-charge force, will tend to cause the electron beam to disperse, i.e., cause outward radial motion of the electrons. The space-charge force will be proportional to the density of the beam, i.e., to the number of electrons per unit volume. For the usual density of beams employed in microwave tubes (10^{12} to 10^{16} electrons per cubic meter), the dispersion of the beam due to space-charge forces is negligible if the drift space is short (d is small in Fig. 9.1). This condition exists in many klystrons, but in traveling-wave tubes the beam must travel over distances which are so long that considerable dispersion may take place unless some means of keeping the beam together or focused is employed. The means by which the dc space-charge forces are counteracted leads to three common used beam models. These models are discussed below.

Ion-Neutralized Beam

Even with the high vacuum employed in a microwave tube, a few neutral gas particles are still present. Many of these gas molecules are ionized by means of collisions with the relatively high energy electro-

¹The design of electron guns is not treated in this text. For a discussion of these, see J. R. Pierce, "Theory and Design of Electron Beams," D. Van Nostrand Co., Princeton, N.J., 1950.

²K. R. Spangenberg, "Vacuum Tubes," McGraw-Hill Book Company, New York, 1948.



4 H H —

FIGURE 9.1
A cylindrical electron beam.

presence of positive ions will tend to neutralize the negative space charge of the electron beam. The positive ions, however, need not be considered in the interaction of a high-frequency electromagnetic field with the beam because their mass is at least 1,800 times greater than the electron's mass, and hence the ac motion of the ions is negligible by comparison with that of the electrons.

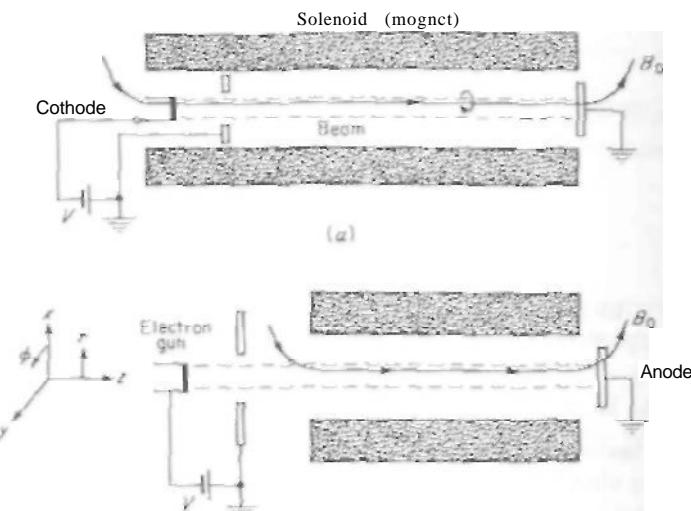
Although all electron beams are ion-neutralized to some extent, complete electron space-charge neutralization is rarely achieved. However, for the purpose of mathematical analysis, a completely ion-neutralized electron beam is sometimes postulated as a model. Beam spreading due to space-charge forces is discussed in Spangenberg's book.

with Axially Confined Flow

If a very large static magnetic field B_0 in the direction of the beam velocity is applied, the effect is to constrain the electrons from moving in the radial direction. The space-charge forces tend to impart a radial velocity to the electron. The magnetic field B_{\parallel} produces a force $-ev_x B_0$, which causes the electrons to execute circular motion about the magnetic field lines and thus prevents the beam from dispersing in the radial direction.

In the magnetically focused beam the field B_{\parallel} has its flux lines threading through the cathode surface, as in Fig. 9.2a. Some electron diffusion across the magnetic field lines will occur, but if B_{\parallel} is made large enough, the amount of beam dispersion can be kept small.

For the purpose of mathematical analysis, it is convenient to assume that B_0 is made infinite since in this case no electron motion in a transverse-direction can take place. The analysis of the behavior of the beam under ac conditions is thereby greatly simplified since electron motion can now occur only in the axial direction (one-dimensional motion). The axially confined flow model is commonly used in the treatment of traveling-wave tubes.

**FIGURE 9.2**

(a) Magnetic focusing for axially confined flow; (b) magnetic focusing for Brillouin

Brillouin Flow

In Brillouin flow (to be described), the axial magnetic field B_0 is not permitted to thread through the cathode surface. Since the field lines are continuous, they must move away from the beam region in the radial direction near the cathode, as shown in Fig. 9.26. When the beam enters the magnetic field region, it is given a uniform rotation at the Larmor frequency $w_l = eB_0/2m$ by the magnetic field. In cylindrical coordinates, r, θ, z , equation of motion for an electron,

$$\frac{dv}{dt} = -e(E + v \times B)$$

may be written in component form as

$$\frac{d^2r}{dt^2} = \frac{I}{e} \frac{dd\theta}{dt} \quad (9.2a)$$

$$\frac{d^2\theta}{dt^2} = \frac{dr}{dt} \frac{dtf}{dt} - \frac{e}{m} \frac{dr}{dt} \quad (9.2b)$$

$$\frac{d^2z}{dt^2} = \frac{dz}{dt} = U_n \quad (9.2c)$$

in the region where $B = B_0 a$, and is uniform. It is assumed that $E = 0$. The radial electric field may be found by using Gauss' law-

charge density of the beam is $-p_0$, then $2-rD_r = -7rr^2p_0$, or $E_r = -rp_0/2e_0$. The radial space-charge force on an electron is thus $-eE_r = -ep_0/2e_0$. If $d^2d/dr^2 = 0$, we find from (9.26) that

$$\frac{dr}{dt} \sim \frac{e}{m} \mathbf{v} \times \mathbf{B} \quad (9.3)$$

If this solution is to satisfy (9.20) and also make d^2r/dt^2 vanish, we require

$$rwf = \frac{e}{m} [\frac{rPo}{m} + \mathbf{r} \cdot \mathbf{S}_a]$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{v}} = \frac{2m e_0}{2} \mathbf{v}$$

where $\omega_p = (ep_0/ne_0)^{1/2}$ is called the plasma frequency. Typical values of ω_p for beams used in microwave tubes range from 10^7 to 10^9 . If the focusing field B_0 is chosen to satisfy (9.4), there will be no radial acceleration of the electrons. The equilibrium condition in the radial direction is actually a balance of outward radial forces $-eE_r$ due to space charge and $m\omega^2 r$ due to centrifugal acceleration against the inward magnetic radial force $e\omega B_r$. Electron-beam flow under these conditions is referred to as Brillouin flow.

Although we have given the conditions for steady-state Brillouin flow within the uniform B_r field region, we did not show that a beam leaving a cathode with a velocity v_{0a} will assume Brillouin-flow characteristics as it enters into the uniform B_r field region through the nonuniform field region in front of the cathode. To show this requires demonstrating that the change in angular momentum of the beam from an initial value of zero to its final value for Brillouin flow is equal to the time integral of the torque $ev_0 B_{ar}$ produced by the radial magnetic field component in the nonuniform region. The reader is referred to Brillouin's original paper for the derivation.

The conditions required for Brillouin flow can be achieved in practice. Even if the beam is partially ion-neutralized, as long as p_0 is not zero, a value for B_u such that (9.4) holds can be found. However, the behavior of a beam with Brillouin flow under ac conditions is more difficult to treat since transverse motion of the electrons is permitted. For this reason the ideal axially confined flow model is more commonly used.

ff.. Brillouin, A Theorem of Larmor and Its Importance for Electrons in Magnetic Fields. *Phys. Rev.*, vol. 67, p. 260. 1945.

W. G. Dow, Nonuniform D.C. Electron Flow in Magnetically Focused Cylindrical Beams, *Advan. Electron. Electron Phys.*, vol. 10. 1958.

Pierce, *op. cit.*

In the magnetron-type (M-type) traveling-wave tube, a r beam is used. For this type of beam an analogous flow, referred to as Snel Brillouin flow, can take place. The properties of sheet beams are H^S-planar in Sec. 9.11, dealing with M-type traveling-wave tubes, and hence discussed are not covered in this section.

9.3 SPACE-CHARGE WAVES ON BEAMS WITH CONFINED FLOW

This section is devoted to an analysis of space-charge waves on a confined electron beam inside a cylindrical waveguide of radius h . The radius of the beam is a , as in Fig. 9.3. Small-signal conditions are assumed.⁸

The beam is considered to be uniform in density in a cross-sectional plane. The dc charge density is $-p_0$, and the axial velocity is u_0 . The dc current density in the z direction is $J_{dc} = -p_0 v_0$. The dc parameters p_v and J_a are independent of space and time coordinates. Under ac conditions with time dependence $e^{j\omega t}$, there will be ac components of charge density, velocity, and current that vary with time and the spatial coordinates. These ac components are denoted by p , v , and J . The ac fluctuation in electron density from the dc value N will be denoted by n .

The electromagnetic field satisfies the equations

$$\nabla \times E = -jwff_u H \quad (9.5a)$$

$$\nabla \times H = \mu_0 E + J \quad (9.5b)$$

$$\nabla \cdot E = 0 \quad (9.5c)$$

$$\nabla \cdot B = 0 \quad (9.5c)$$

$$\nabla \cdot J = -j(t)p \quad (9.5c)$$

A unit volume of the beam with charge density $-p_0 + P$ and charge-mass ratio n equal to that for electrons, that is, $TJ = e/m$, has a motion governed by the equation

$$(N + n)m \ddot{v} = (p - p_0)(E + v_x B_x + v_y B_y) \quad (if)$$

where $N + n$ is the number of electrons per unit volume and $v = \sqrt{v_x^2 + v_y^2}$ the total velocity. For small-signal conditions, ($B_0 \ll |B|$; so the i is

$$\overline{B_0}$$

$v \times B$ can be neglected compared with $v \times B_{dc}$. In addition, since $B = x_0 H = j_0 Y f E = E/c$, we see that $|v \times B|$ is smaller than $|E|$ by a factor v/c . Hence ac magnetic forces are negligible. The total velocity v is a function of x , y , z , and t . In addition, the position x, y, z of a charge element is a function of time. Consequently,

$$\frac{d\mathbf{v}}{dt} + \mathbf{v} \cdot \nabla \mathbf{v} + \mathbf{d} \cdot \mathbf{r} \frac{d}{dt} + \nabla \mathbf{V} \cdot \mathbf{v} = \frac{d}{dt} (\mathbf{v} \cdot \mathbf{V}) \mathbf{v} \quad (9-6)$$

since $\mathbf{v}_t = a_x \frac{dx}{dt} + a_y \frac{dy}{dt} + a_z \frac{dz}{dt}$. Thus expansion of $d\mathbf{v}/dt$ leads to a nonlinear term $\mathbf{v} \cdot \nabla \mathbf{v}$, depending on of . However, for an ac velocity v that is small compared with the dc velocity v_0 , we have

$$[(\mathbf{v} + \mathbf{v}_0) \cdot \nabla](\mathbf{v} + \mathbf{v}_0) = (\mathbf{v} + \mathbf{v}_0) \cdot \nabla \mathbf{v} = (\mathbf{v}_0 \cdot \nabla) \mathbf{v} \quad (9.7)$$

since \mathbf{v}_0 is constant and the second-order term $(\mathbf{v} \cdot \nabla) \mathbf{v}$ is negligible and may be dropped. Thus we obtain

$$\frac{(AT + n)m}{pc} \nabla \cdot \mathbf{v} + (\mathbf{v}_0 \cdot \nabla) \mathbf{v} = (\mathbf{p} - >_0)(\mathbf{E} + \mathbf{v} \times \mathbf{B}_0)$$

But the terms involving n and p are products of two ac quantities and may be dropped for small-signal conditions. Hence the first-order linearized equation of motion becomes

$$\frac{d\mathbf{v}}{dt} + (\mathbf{v}_0 \cdot \nabla) \mathbf{v} \underset{\ll}{=} -i \mathbf{j} (\mathbf{E} + \mathbf{v} \times \mathbf{B}_{dc}) \quad (9.8)$$

since $Ne = p_0$ and $e/m = \pi$.

For the cylindrical beam under consideration, we also have $\mathbf{v}_0 = a, v_0$ and a time dependence $e^{i\omega t}$. If we let B_n approach infinity, the transverse components of v must vanish, so that the term $v \times B_{dc}$ in (9.8) will vanish. Thus v has a component in the z direction only, and (9.8) gives

$$\frac{Hu}{jwv + u_0} = -7E, \quad (9.9)$$

When v has only a z component, the ac current density J has only a z component since the total current is

$$\begin{aligned} J_0 + J &= (f_0 + p)(v_0 + v) = -p_0 V_Q + (pV_{dc} - >_0 v) + pv \\ &= -p_0 v_{dc} + (p v_0 - p_0 v) \end{aligned} \quad (9.10)$$

after dropping the second-order term pv , which is the product of two small ac quantities. The dc and ac currents are thus

$$J_0 = -p_0 V_Q \quad (9-Ha)$$

$$J = p v_0 - p_0 v \quad (9.116)$$

From the continuity equation (9.5e) we obtain

$$\frac{dj}{Hz} = -Jwp$$

Equations (9.9), (9.116), and (9.12) permit us to express J as a function of E_z . Maxwell's equations (9.5a) and (9.56) may then be solved in the manner to obtain wave solutions.

Since we are looking for wave solutions, we may assume that the quantities have a z dependence $e^{j\beta z}$. In this case (9.9) and (9.12) give

$$(j \ll -JPv_0)v = -TJE, \quad (9.13a)$$

$$jpj = jwp \quad (9.13b)$$

For convenience, ω/v_0 will be denoted by β_0 , which may be interpreted as the dc propagation constant for the beam. Using (9.116) and (9.13), we find that

$$J = \sim_j \frac{Pl^A E_s}{(\mathbf{P}_0 - \mathbf{P})} \quad (9.14)$$

where $\omega_p = p_0 v^2 / o$ is the plasma frequency squared, and $-p_0$ is the electron charge density of the beam.

To solve Maxwell's equations for the beam inside a cylindrical guide, it will be convenient to introduce the vector potential. For a mode having azimuthal symmetry (no θ dependence), all boundary conditions can be satisfied by a vector potential having only a z component $A_z(r, z) = i/(r)e^{-j\beta z}$. The equation satisfied by A_z is

$$\nabla^2 A_e + k^2 A_z = -n_0 J$$

From A_z , we obtain

$$E_z = -j\omega A_z + \frac{\nabla V \cdot -A}{jV(t_0 \epsilon_0)} = \frac{i}{r} + \frac{d^2 A_z / dz^2}{a^2 \epsilon_0} = \frac{k^2 \omega^2 p^2}{Ji \gg H_0^2} \quad (9.15)$$

Using (9.14) to express E_z in terms of J gives

$$\gg *' -i -i \otimes TtW * \quad (9.16)$$

The Helmholtz equation for A_z now becomes

$$\nabla^2 A_t + k^2 A_z = 0 \quad 0 < r < a$$

$$\nabla^2 A_t - k^2 A_z = 0 \quad a < r < b$$

(9.16)^{fl}

(9.16)^{fr}

where we have replaced V^2 by $V^2 - p^2$ and put

$$\mathbf{p} \cdot \mathbf{H} = \frac{p^2 - k_0^2}{\mu_0} \mathbf{E}$$

$$- - (\star \star - \star \ll 1 - \mathbf{0} \ll - \mathbf{1} \mathbf{8}) \quad (9.17a)$$

$$/i^2 = /3^2 - fe \quad (9.176)$$

after using (9.15) to express J in terms of A .

The analysis, when completed, will show that the space-charge waves are slow waves, with $0 \sim p_0 \sim k_0$, and hence p and h will be real. With no θ variation, (9.16) reduces to

$$\frac{dH}{dr'} = \frac{1}{r} \frac{d \langle l \rangle}{dr} - h' = 0 \quad (9.18)$$

where $A_z = ili(r)e^{-j\beta r}$. The equation for $\langle l \rangle$ is Bessel's equation of order zero, and the solutions are $J_0(pr)$, $Y_0(pr)$, $J_0(jhr)$, and $Y_0(jhr)$. Instead of using the Bessel functions with imaginary argument, we use the modified Bessel functions $I_0(hr)$, $K_0(hr)$. In the region $0 < r < a$ we cannot use Y_0 since it becomes infinite. Therefore we let

$$i''(r) = C_1 J_0(pr) \quad 0 < r < a$$

$$i''(r) = C_2 I_0(hr) + C_3 K_0(hr) \quad a \leq r < 6$$

where C_1 , C_2 , C_3 are arbitrary constants. The axial electric field must vanish at $r = 6$ and must be continuous at $r = a$. These conditions hold for A_z and hence for i'' also. Thus

$$C_1 J_0(pa) = C_2 I_0(ha) + C_3 K_0(ha) \quad (9.19a)$$

$$0 = C_2 I_0(hb) + C_3 K_0(hb) \quad (9.196)$$

Besides E_z , the only other field components present are E_r and H_{ll} . These are given in terms of $\langle l \rangle$ by

$$\begin{aligned} \mathbf{p} \cdot \mathbf{M}_a &= \mathbf{Pc} \cdot \mathbf{M}, \\ wii_0 e_0 \langle l \rangle &\quad k_0 dr \\ H_l &= \frac{1}{\mu_0} \frac{dA}{dr} \quad \bullet Y_0 E_r \end{aligned}$$

Continuity of H^\wedge at $r = a$ requires

$$C_1 p J_n(pa) = C_2 h l'_a(ha) + C_3 h K'_0(ha) \quad (9.20)$$

where the prime denotes differentiation with respect to the argument pa or ha . In order for (9.19) and (9.20) to have a nontrivial solution for C_1 , C_2 ,

and C_3 , the determinant of the coefficients must vanish. Thus we find that

$$\frac{J''(pa)}{J'(pa)} = \frac{K_0(l_0(ha)) - K'_0(l_0(ha))I_0(kb)}{J_0(pa) + K_0(ha)I_0(kb)} \quad (9.21)$$

This transcendental equation, together with the relations (9.17) and (9.21), determines the propagation constant β . This gives

$$p^2 = -k^2 + \frac{a^2 h^2}{(p_0 - /1^2 + ktf)} \quad (9.22)$$

determines the propagation constant β .

Two special cases are now examined. First consider the case where a and a are made very large. Then, since

$$\beta \approx (\frac{p}{V}) - \frac{1}{2\pi c} \sim \frac{p}{V}$$

for large x , we find that (9.21) gives $p \tan(pa - \pi/4) = h$. But since we are letting a go to infinity, the only possible solution independent of a is $p = h = 0$. From (9.17) we then obtain the following solutions for (i) :

$$P = \pm k_0 \quad (9.23a)$$

$$0 = /3_0 (1 \pm \sqrt{\omega_p^2/c^2}) \quad (9.23b)$$

But with $p = h = 0$, $(i = \pm k_0)$, all field components vanish as reference to the equations given earlier for E_r , H_{lb} , and E_r shows. Thus this is a trivial solution. The other solutions $/3 = /3_0(1 + \omega_p^2/c^2)$ correspond to the space-charge waves. The wave velocities are

$$1 \pm \frac{\omega_p}{c} \quad \text{and} \quad v_r = \frac{c}{\sqrt{1 \pm \frac{\omega_p^2}{c^2}}}$$

since $\omega_p < K$ for conditions that are typical in microwave tubes. The wave velocities are slightly greater and slightly smaller than the dc beam velocity v_0 . The two waves are called the fast and slow space-charge waves. For $p = 0$, both E_r and H_r vanish but E_r remains finite. The space-charge waves may thus be viewed as a longitudinal oscillation of the electron beam. When $\omega_p = \omega_0$, one solution corresponds to $/3 = 0$, that is, no propagation. Thus it is seen that the plasma frequency is a natural frequency of oscillation for electrons in an infinite beam. Since the trap fields are zero, the space-charge waves are not changed even if a is finite as long as the beam has infinite radius.

As a second special case consider the situation $a = \infty$ so that the beam fills the waveguide. From (9.21) we now see that the denominator becomes infinite since the denominator vanishes. Thus we require $\omega_p = 0$. Hence pa takes on values typical of those for TM_{0,1} modes in a rectangular guide. The lowest-order solution is (Sec. 3.18) $pa = 2.405$, or $m = 1$.

$p_a = p_{0m}$, where the p_{0m} are given in Table 3.5. Using (9.22), we now find that

$$\begin{aligned} P_o &= -K' \quad 1 \quad - \quad \frac{Pi}{(ft, \sim yfh^2 + kl)} \\ &= (*2''/3^2) \quad 1 \quad \frac{Pi}{w \quad I \quad (Po-P)} \end{aligned} \quad (9.24)$$

For the field waves we expect j_3 to be approximately equal to k_0 . Then, since $w_p \ll w$ and $P_o \gg \&_0$, an approximate solution of (9.24) is

$$F = k\% - \frac{Po}{P}, \quad (9.25)$$

This is the unperturbed propagation constant for TM_{,,,} modes in a cylindrical guide. A correction to $/3$ may be obtained by using the solution given by (9.25) in the term multiplied by w^2 in (9.24).

Of greater interest are the space-charge waves for which $ji \sim po$. For these $k'l \ll /3^2$; so (9.24) may be approximated by

$$P_o, \quad = -\&' \quad 1 \quad - \quad \frac{A}{<\bullet> \quad (Po-P)}$$

which is a quadratic equation in $/3^2$. To obtain an approximate solution, let $P = po(l + 8)$, where <5 will be small. Then we obtain

$$P_o, \quad = -Pl \quad co$$

which gives

$$5 = \frac{\pm(w - /M)P_c}{/3_0^2 + (Po_m/<*)T}$$

and hence

$$P=Po \quad \frac{\bullet i_1 + rf}{PW}^{1/2} \quad (9.26)$$

Note that 5 is small, which justifies the approximations made. We may express (9.26) in the same form as (9.236) by introducing an effective plasma frequency $u>q$ given by

$$\gg, = \ll, \quad 1 \quad P_l^2 \quad = Fto_r \quad (9.27)$$

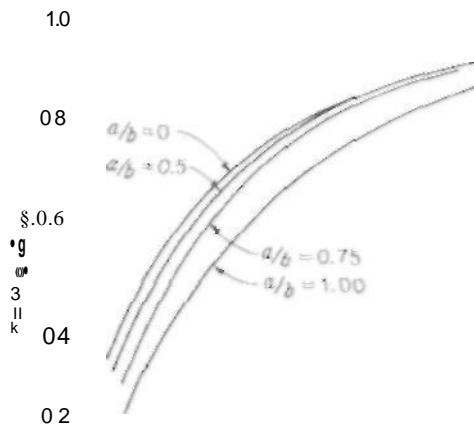


FIGURE 9.4

Space-charge reduction factor for a cylindrical electron beam of radius a , velocity v_∞ , inside a circular guide of radius b . The data apply to the dominant TE₀₁ space-charge mode.

where F is called the space-charge reduction factor; for example, $w^* = F^2 p_0 r / e_a$ so the effective space charge is $F^2 p_0$. Hence we can write

$$\beta = \beta_0 \frac{1}{\text{ft}} \quad (9.28)$$

For the beam completely filling the guide, there are again a slow and a fast space-charge wave, with velocities slightly greater and slightly smaller than the beam velocity v_0 . However, the effective plasma frequency is reduced because of transverse variations in the field. Nevertheless, the space-charge waves have very small transverse field components E_r, H_y . Only the axis electric field E_z is large.

In the general case, when $a \neq b$, the solution for β is tedious. The results may be expressed in the form (9.28) for the space-charge $w^* j^*$, introducing the effective plasma frequency or the space-charge reduction factor. Some typical results computed from curves given by Branch and Mihran are shown in Fig. 9.4.t

The space-charge-wave theory was first developed by Hahn and in 1939. Since that time space-charge waves under a variety of conditions have been studied. The references cited at the end of this chapter

^tG. M. Branch and T. G. Mihran, Plasma Frequency Reduction Factors in Electrons, *IHE Trans.*, vol. ED-2, pp. 3-11, 1955.

provide an introduction to the literature on this topic. For the analysis of the ordinary, or O-type, traveling-wave tube, the model discussed above is accurate enough to describe the main operating characteristics.

4 SPACE-CHARGE WAVES ON UNFOCUSSED BEAMS

Many low-power klystrons employ electron beams without magnetic field focusing when the distance (drift-space length) the beam must travel is short. The propagation of space-charge waves on this type of beam is therefore of interest. We shall consider a beam of radius a and with dc parameters $-p_0, v_0, a_z$. For space-charge waves with axial symmetry, the only field components present are E_r, E_θ , and H_A . The governing equations are Maxwell's equations (9.5) and the force equation (9.8), with B_0 equated to zero.^t

For space-charge waves we may assume a z dependence e^{-iz} . In component form (9.5a) to (9.5c) and (9.8) are

$$JPE_r + \sim = jwp \cdot Q H_A \quad (9.29a)$$

$$JtH_{il} = jlo\epsilon_n E_r + J_r \quad (9.29b)$$

$$\begin{matrix} \mathbf{i} & \mathbf{a} \\ \mathbf{r} & \text{or} \end{matrix} \quad (9.29c)$$

$$\frac{Id}{r} - \frac{r}{\gg r} E_r - m \wedge \frac{p}{e_0} \quad (9.29d)$$

$$\frac{JvEr}{w - pv_0} \quad (9.29e)$$

$$v = \frac{JvE_z}{v} \quad (9.29f)$$

In addition, we have the relation

$$J = pv_0 - p_0v$$

which gives

$$<f^* - -Pn>r''' Br \quad (9.30a)$$

$$J_z = \sim PoV_z + P^yO = <^r E_z + PV_0 \quad (9.30b)$$

where (9.29e) and (9.29f) have been used, and the effective conductivity $\ll r$

^tStatic space-charge forces are assumed to be negligible, which is a valid assumption for a low-density beam and a short drift space. Alternatively, the beam may be assumed to be ion-neutralized.

has been introduced as follows:

$$\alpha = \frac{-JVPo}{u > - / 3u_0} \quad w \sim fv_Q \quad (9.3i)$$

When we make use of (9.30), the continuity equation $\nabla \cdot J =$ found to give

$$^a I \sim j; ^r E r \sim JPEt | - J&HP = -Je>P$$

If (9.29d) is used to replace the term in parentheses, we obtain

$$-jp v_0 + Jo\Psi P = 0 \quad (9.32)$$

If the ac charge density p does not vanish, we must have

$$-JPv_a + ju = 0$$

This requires that ji be given by

$$\begin{aligned} & \langle i \rangle \pm \langle i \rangle \\ & V, \quad \backslash \quad \text{ft} \end{aligned}$$

The corresponding wave solutions are the space-charge waves in an infinite beam for which $E_r = H^\wedge = 0$. For a beam with finite radius the boundary conditions at $r = a$ cannot be satisfied with these values of 0. Consequently, the space-charge waves that we are looking for must have different value of p and, in addition, must have a zero ac space-charge density p , so that (9.32) will hold. For these waves, (9.30) now gives $J = crE$, or

$$J = J^{we} \frac{e}{\epsilon_0} \frac{-E}{c}$$

Maxwell's curl equation for H becomes

$$\nabla \times H = jioe_0 E + J = jioe_0 I + \frac{a}{jwe_0} E = . / WE$$

where the effective permittivity e of the beam is given by

By using the effective permittivity ϵ , the beam may be treated as a dielectric cylinder. The equation satisfied by E , is thus

$$\nabla^2 E + k^2 E_t = V?E_t + (k^2 - p)E_z - 0 \quad (9.34)$$

where

$$k^* = i\omega p_0 \epsilon = kl \quad \text{t-S-1} \\ <\omega(w - pu_0)$$

For space-charge waves we anticipate *slow* waves, for which $/S \gg k_0$. Hence a suitable solution for E_z is

$$E_r = CJ_n(pr) \quad 0 < r < a \quad (9.35a)$$

$$E_s = C_2 K_0(hr) \quad r > a \quad (9.35b)$$

where $p = (3^2 - *^2)^{1/2}$, $ft = (3^2 - A^2)^{1/2}$, and C, C_2 are amplitude constants to be determined.

The boundary conditions at $r = a$ are different from those for a beam with confined flow, for the following reasons.^t The electrons have a radial ac velocity, and hence the boundary of the beam does not remain at its dc position $r = a$ since electrons will oscillate back and forth about the dc boundary. The resultant boundary thus becomes rippled, as shown in Fig. 9.5a. The positive charge shown in this figure represents a deficit of negative charge. The effect of the rippled boundary on the radial electric field under small-signal conditions may be accounted for by replacing the rippled boundary by a layer of surface charge. This is, of course, exactly what is done in the case of a dielectric boundary in which polarization charge oscillates back and forth about a mean boundary surface. The surface charge is given in terms of the dielectric polarization P by $P \cdot n = (D - \epsilon_0 E) \cdot n$. For the electron beam the corresponding surface-charge density p_s arises from two causes, namely, charge flowing toward the boundary because of the radial current J_r and charge carried to a given point z owing to surface charge moving with the beam. That is, the rate of increase of surface charge is (Fig. 9.5b)

$$— \equiv j(p_s) = J_r \quad t' — = J_r + J_p v_{LJ} p_s$$

The term $-1/\epsilon_0 Hpx/dz$ arises as follows: Let the surface charge density at a point z be p_s . Then at the point $z - dz$ the charge density is approximately $p_s - (\epsilon_0 Hpx)(dz)$. The rate at which charge is carried away from the point z on the boundary is $v_{LJ} p_s$, and the rate at which charge from the adjacent

^tSee also W. C. Hahn, Small Signal Theory of Velocity Modulated Electron Beams, *Gen. Elec. Rev.*, vol. 42, pp. 258-270. 1939.

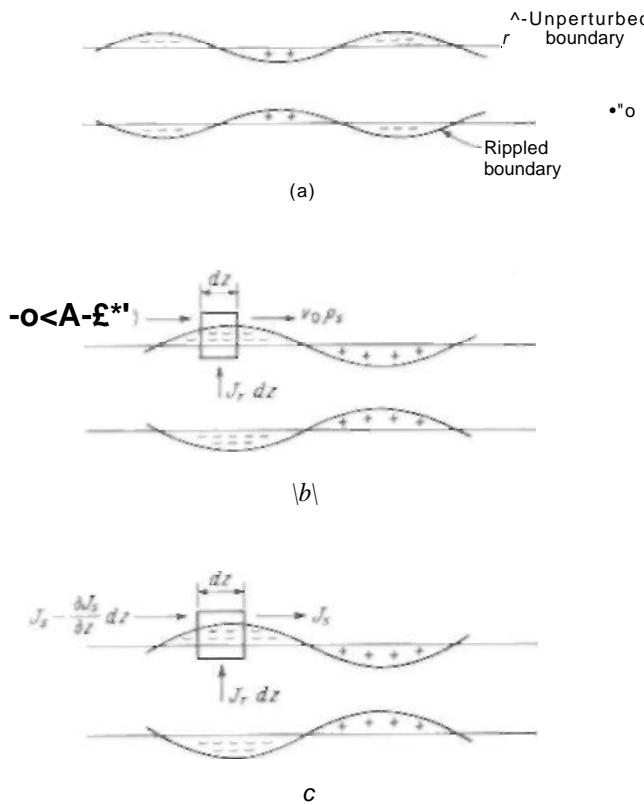


FIGURE 9.5
Boundary conditions at the surface of a rippled beam.

point $z - dz$ flows toward z is $v_0 [P_s - O_p j dz] dz$. The net rate of accumulation of charge in an interval dz , due to the finite beam velocity u_0 , is $-v_0 V_p / dz dz = J I 3 v_0 P_s dz$, since the z dependence is $e^{(o-pv_0)z}$. The density is obtained by dividing by dz . Our final expression for the equilibrium surface charge density is

$$\rho_s = \frac{crE}{j(t_o - pv_0)} \sim \frac{-VP\theta^E r}{j(o-pv_0)Y} \quad (9.36)$$

The amount of charge which has crossed the unperturbed boundary ρ_{0r} , where r is the ac displacement of a unit volume of charge, is

$$\rho = \rho_r = \frac{-A}{at} + \frac{(v_0 \cdot V)r}{at} = j(a) \sim Pv_0 r$$

we see that

$$= \sim \mathbf{P} \mathbf{o}^r = j(c\omega - l3v_0) \quad j(\omega - Pv_a)$$

which is an alternative derivation of (9.36).

The *boundary condition* to be applied to the radial electric field is

$$E_{2r} \sim E_{lr} = \frac{3}{\epsilon_0 \epsilon_r (w - 3U_0)} = \frac{p}{(to - fv_0)} r_A \quad (9.37)$$

where the subscripts 1 and 2 refer to the fields in the regions $r < a$ and $r > a$, respectively. The above boundary condition may also be expressed in the form

$$\hat{\mathbf{E}}_{2r} = \circ \quad \begin{matrix} a \\ \mathbf{j} \mathbf{l} \mathbf{v} - \mathbf{f} \mathbf{a} \end{matrix} \quad \mathbf{K}$$

This latter result is similar to that which holds at the boundary of a dielectric cylinder except that ω is replaced by $w - fv_0$ because of the uniform motion of the cylinder, in the z direction, with velocity v_0 .

Associated with the equivalent surface charge is *an* equivalent surface current of density $J_s = J_{sa}$. To obtain an expression for J_s , consider Fig. 9.5c. The total rate at which charge flows into a small region of length dz on the boundary is

$$J_s dz = \frac{W}{dz}$$

and must equal the rate of increase $J a p_s dz$ of surface charge density in an interval dz . Consequently,

$$- \frac{dJ_s}{dz} = J a p_s = j f p, \quad - J_r = J f i o D p$$

$$\text{or} \quad J_s = f t B, \quad (9.38)$$

after replacing J_r by $j(i\omega - fv_0)p_s$ from (9.36). For $w \gg w_p$, the total surface current is usually much larger than the total volume current, and it is therefore very important to include it.

The boundary condition to be applied to H^\wedge is

$$\#2^* - \#u = \wedge = P_s \#o \quad (9.39)$$

When H_{lr} satisfies this discontinuity relation, the boundary condition on the radial electric field is also satisfied. With the above boundary conditions we are now able to complete the solution to our beam problem.

When we combine (9.29a) and (9.296), we obtain

$$H_{\text{ext}} = \frac{j(ae_0 + a)}{p^2 - k_0^2 + ja} e_Z = \frac{jwe}{f^2 - k^2} e_Z$$

Referring to (9.35), we thus find that

$$H_{\text{ext}} = H_h = \frac{j(oe)}{\gamma y^C i^r} = \frac{jioe}{C J_l(p)} \quad r < a \quad (9.40)$$

$$H_{\text{ext}} = H_{2il} = \frac{/we}{C K^h r} = \frac{/we}{C K^h r} \quad r > a, \quad (9.40)$$

We require E_z to be continuous at $r = a$ and $H_{z,4}$ to satisfy the condition (9.39). Therefore

$$\begin{aligned} CM Pa) &= C K_0(ha) \\ /we, & - /we \\ -CMka) & - C J_l(p) = p_{s,0} = f L \% \end{aligned} \quad (9.41a)$$

From (9.296) we find $E_{lr}(y^w e_0 + a) = j p H_{VU}$, and hence the second boundary condition becomes

$$\frac{C K_0(ha)}{h} - \frac{C J_l(p)}{p} = \frac{w_r}{(to - p v_0 y)} \quad (9.41b)$$

Dividing (9.41b) by (9.41a) gives

$$\frac{7, (pa)}{/we, (pa)} = \frac{1}{(w - p v_0 y)} \frac{K_x(ha)}{h K_0(ha)} \quad (9.42)$$

The propagation constant p is determined by a solution of this equation, together with the relation

$$02 = p^2 + k^2 = p^2 + k^2 - \frac{1}{w(w - p v_0)} * S - t f + A j \quad (9.43)$$

For space-charge waves, $p \gg k_0$, so that h and p can be replaced by P' approximation that may be used to simplify (9.42) and (9.43).

The results obtained by solving (9.42) and (9.43) may be expressed in the form (9.28):

At microwave frequencies, w/w is usually in the range 0.01 to 0.1.

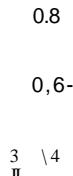


FIGURE 9.6

Plasma-frequency reduction factor F for an unfocused cylindrical electron beam as a function of fia .

differs from p_{co} by only a few percent or less. In Fig. 9.6 the plasma-frequency reduction factor $F = w_q/u>p$ is plotted as a function of fia . If p and h are replaced by $j3$ in (9.42), that equation may be put into the form

$$\beta_0 a = \beta_n a \left(1 + \frac{\frac{KJ_0}{co}}{K_0 I_x} \right)^{-1/2}$$

from which it is seen that

$$\frac{\beta_0 a}{\beta_n a} = \frac{fia}{K_0(130)I_x(130)}^{-1/2} \quad (9.44)$$

This expression was evaluated to obtain F , as given in Fig. 9.6. The data in Fig. 9.6 may be used to solve for the corresponding values of $\beta_n a$; that is,

$$\beta_0 a = \frac{fia}{l \pm F(co_p/co)} \quad (9.45)$$

In a beam with confined flow only axial ac convection currents are permitted. For this reason an ac charge density p exists since the current J_2 is a function of z and must have associated with it space-charge-density fluctuations. For the unfocused beam, ac radial convection currents are also present, and this makes it possible for the total current $J + jwe_n E$ to be solenoidal, i.e., to form continuous closed flow lines without terminating in ac space charge.

ac POWER RELATIONS

The ac power associated with the space-charge waves on an electron beam is of importance for understanding the gain mechanism of traveling-wave

Lubes and, of course, for power calculations. The small-signal DO rem for beams with confined flow was first derived by Chu f p" ¹
extensions have been made by Haus and Bobroff? and Kluver. § Th ^{?^rt.^tr}
tion is straightforward, but the interpretation of the various te [^]
enter in is not always clear, depending on the type of beam ^{irmri, ' \thal}
considered.

Maxwell's equations for the fields associated with the beam are

$$\mathbf{V} \times \mathbf{E} = -\mu_0 \mathbf{H} \quad \mathbf{V} \times \mathbf{H} = -\omega_0 \mathbf{E} + \mathbf{J}$$

where the small-signal ac current density is given by (the surface current" is included for generality)

$$\mathbf{J} = -n_0 \mathbf{v} + p_{v,} + p_s \delta_Q$$

In addition, we have the equation of motion

$$j\mathbf{c}\mathbf{v} + (\mathbf{v}_0 \cdot \nabla)\mathbf{v} = -\mathbf{T}\mathbf{J}(\mathbf{E} + \mathbf{v} \times \mathbf{B}_0)$$

Expanding the following expression and using the above equations, we obtain

$$\begin{aligned} \mathbf{V} \cdot (\mathbf{E} \times \mathbf{H}^*) &= \mathbf{H}^* \cdot \mathbf{V} \times \mathbf{E} - \mathbf{E} \cdot \mathbf{V} \times \mathbf{H}^* \\ &= -jioi\mu_0 H \cdot \mathbf{H}^* + j<\omega_0 E \cdot \mathbf{E}^* - \mathbf{E} \cdot \mathbf{J}^* \end{aligned} \quad (9.46)$$

The continuity equation in the interior of the beam may be written as

$$\begin{aligned} \mathbf{v}, \mathbf{V} \cdot \mathbf{J} &= -jwpv_0 = -jai(p\mathbf{v}_0 - p_0\mathbf{v}) - jwpv_0 \\ &= -ja)J - jojpov \end{aligned}$$

If we multiply the complex conjugate of this equation by v/ij , we obtain

$$\frac{\mathbf{V} - \mathbf{V}_0}{v} \quad \mathbf{J}^* = \frac{-\mathbf{V}}{J} \quad J^* + \frac{\mathbf{V} \cdot \mathbf{V}^*}{n} \quad \wedge \wedge \wedge \wedge \wedge \wedge$$

Multiplying the equation of motion by J^*/ij gives

$$J^* \cdot \mathbf{E} + J^* \cdot \mathbf{v} \times \mathbf{B}_0 = \frac{>}{V} \mathbf{v} - J^* \frac{J^*}{V} (\mathbf{v}_0 \cdot \mathbf{V}) \mathbf{v}$$

Since the transverse components of \mathbf{J} and \mathbf{v} are in the same direction, the term

$$J^* \cdot \mathbf{v} \times \mathbf{B}_0 = J^* \times \mathbf{v} \cdot \mathbf{B}_0 = 0$$

T.L. J. Chu, A Kinetic Power Theorem, paper presented at the IRE-PGED Electr Research Conference, Durham, N.H., June, 1951.

tH. A. Haus and D. Bobroff, Small Signal Power Theorem for Electron Beams, *J. Appl. Phys.*, vol. 28, pp. 694-703, June, 1957.

SJ. W. Kliiver, Small Signal Power Conservation Theorem for Irrotational Electro *J. Appl. Phys.*, vol. 29, pp. 618-622, April, 1958.

The addition of the above two equations thus gives

$$\frac{J^* \cdot E}{7} - \frac{\nabla \cdot V^*}{V} = - \frac{\nabla \cdot r^*}{\nabla J^*} - \frac{I^*}{I}$$

If we introduce the term $V \cdot (v - v_0 J^*) = v \cdot v, V \cdot J^* + J^* \cdot V(v \cdot v_0)$, we obtain for the right-hand side

$$\frac{-V}{1} - \frac{\nabla \cdot V}{V} - \frac{r[J^* - I^*]}{V} = \frac{7V - (r^* - Vv \cdot v_0)}{V}$$

By expanding the bracketed term in rectangular coordinates, it is found to reduce to

$$-a \cdot X J^* \cdot V X v$$

when $v_0 = v_0 a_z$. Hence we have

$$\frac{J^* \cdot E}{V} - \frac{\nabla \cdot V^*}{V} = - \frac{V \cdot V^*}{V} = a \cdot X J^* \cdot V X v \quad (9.47)$$

The ac kinetic-power theorem is obtained by adding (9.47) to (9.46) to obtain

$$\begin{aligned} T - E \cdot X H^* &= \frac{V \cdot V_n}{V} J^* / = \frac{PO}{TO} H \cdot H^* - \frac{PO}{TO} v \cdot v^i \\ &\quad + jwe_0 E \cdot E^* + \frac{a \cdot X J^* \cdot V X v}{V} \end{aligned} \quad (9.48)$$

The term $(p/2n)v \cdot V^*$ is the ac kinetic-energy density in the beam since $PO/TO = ((>\rho/e)m$ is the mass density per unit volume.

Let us now specialize to the case of a beam with confined flow for which $J = a, J$, and the last term on the right-hand side of (9.48) vanishes. The real part of the above equation then gives

$$ReV \cdot E \cdot X H^* - \frac{V \cdot V^*}{V} = 0 \quad (9.49)$$

The volume integral over a volume enclosing the beam between cross sections S_1 and S_2 as in Fig. 9.7 may be converted into a surface integral

5c

-rfs

S_1

S_2

FIGURE 9.7

A section of an electron beam.

over the surface $S_1 + S_2 + S_c$. Thus the following power-current is obtained:

$$\operatorname{Re} \oint \mathbf{E} \times \mathbf{H}^* \cdot d\mathbf{S} = -\operatorname{Re} J''_{s, 27}$$

$$+ \operatorname{Re} f_{a, dS} '9.50$$

The term $(-\mathbf{v} \cdot \mathbf{V}_h)/TJ = -im/e)v \cdot \mathbf{V}_h = V_h$ has the dimensions of a voltage, and is called the kinetic voltage. It is the term that gives the conversion of kinetic energy of the beam into electromagnetic energy. In order that the net amount of electromagnetic energy may flow out of the surface S quantity

must be negative.

A fuller appreciation of the above relations may be obtained by considering the application to space-charge waves in an axially confined infinite cross-section electron beam. For the slow and fast space-charge waves, the ac kinetic power may be obtained by using (9.13a), (9.14), and (9.49). It is found that

$$Be[-V_{kf}Jf] = Re \frac{-v_Q v_f J_f}{2n} \frac{\omega_0 p e_0}{2v_l(f) - p_0} m_s$$

$$\operatorname{Re} \frac{1}{1} = I_* = - \frac{tow^2 p e_c}{2uS(0, -/30)} E^1$$

Since $p_s > p_0 > f_i$, the slow space-charge wave has a negative ac kinetic power, whereas the fast space-charge wave has a positive ac kinetic power. If the beam is excited, with only the slow space-charge wave, the significance of the negative ac kinetic power is that some of the dc flow energy, the beam has been extracted and converted into negative ac energy during the excitation process. The negative ac kinetic energy must in turn be converted into electromagnetic energy flow in order to maintain power consumption. In the discussion of the traveling-wave tube, it will be seen that the space-charge wave is the one that produces amplification.

VELOCITY MODULATION

The preceding sections have established the existence of space-charge waves on electron beams. We must now examine the problem of exciting these waves, i.e., producing an ac velocity modulation on the beam. In the tubes velocity modulation is commonly produced by passing the

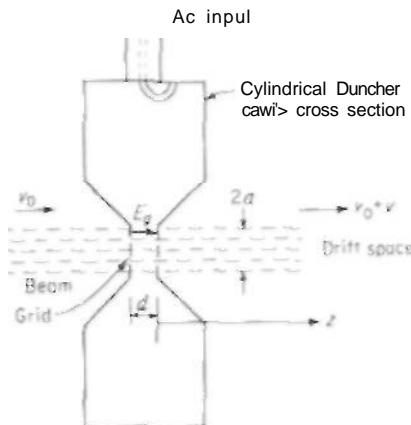


FIGURE 9.8
Velocity modulation of an electron beam.

through two closely spaced grids located at the center of a cylindrical reentrant cavity, as in Fig. 9.8. The particular form of cavity used is chosen in order to satisfy the requirement of high ac electric field strength across the grids (which requires small grid spacing) and yet maintain a high cavity Q . The latter requires a large volume-surface area ratio. If we let $\text{Re}(E_g e^{i\omega t}) = E_g \cos \omega t$ be the cavity electric field across the gap (often referred to as the buncher gap), those electrons entering the gap when $E_g \cos \omega t$ is directed in the negative z direction will be accelerated and will leave with a velocity greater than v_0 . Electrons entering the gap region when $E_g \cos \omega t$ is directed in the positive z direction are slowed down and will leave with a velocity less than v_0 . It is apparent, then, that an applied ac electric field between two parallel grids will velocity-modulate an electron beam. The analytical details of the modulation process are discussed below.¹

We shall consider an unfocused electron beam of the type discussed in Sec. 9.4. Let the cavity field in the gap region be $E_f \cos \omega t$. Electrons will traverse the gap with essentially the entrance velocity v_0 . If the time at which a particular electron passes the midplane $z = -d/2$ is t_0 , then the field in the cavity at time t when this electron is at a position $z = -d/2 + v_0(t - t_0)$ is

$$E \cos \omega t = E_f \cos \omega t + f_0(t - t_0)$$

The work done by the cavity field on the electron during its transit through

¹This analysis is based on a ballistic formulation, and not on a field approach, since the former is more straightforward.

the gap is

$$W = -C \sim eE_g \sin(\theta_0) P_o |Z + \frac{e}{M} e^{j\theta_0} \rangle dz$$

(9.1)

where $\theta_0 = ai/v_0$ is the dc propagation constant for the beam. The coupling parameter M is defined to be

$$\frac{\sin(fL/2)}{L/2} \quad (9.52)$$

For an electron passing the midplane at time t , the work done on the electron is clearly given by

$$-eE_g dM \cos \theta_0$$

The work done on an electron results in an increase in its kinetic energy. If the exit velocity from the buncher cavity is $v_0 + v_2$, we have

$$jm(v_0 + v_2 f - \theta_0) = m(2u_0 i_2 + vf) = mv_0 u_t = -eE_g dM \cos \theta_0 \quad (9.53)$$

since for small-signal conditions, $v_2 \ll v_0$. In complex form (9.53) may be written as

$$v_2 e^{J'''} = E_g dM e^{J''''}$$

Thus the axial ac beam velocity at the exit grid has a value

$$v_2 = -M(E_g d) \quad (9.54)$$

The foregoing first-order analysis predicts that there will be zero average work done in bunching the electron beam since the average of (9.51) over one period from f to C , $+ L/f$ is zero. This result is not *correct*, and in actual fact, a net amount of average work is required to velocity-modulate the beam. To determine the average work done, a second-order analysis must be performed.¹ The principal effect of requiring a finite amount of work to velocity-modulate the beam can be represented by an equivalent shunt conductance loading the buncher cavity (beam loading of the buncher cavity). The magnitude of this shunt conductance is typically such a

¹K. R. Spangenberg, "Vacuum Tubes," chap. 17, McGraw-Hill Book Company, New York, 1948.

M. Chodorow and C. Suskind, "Fundamentals of Microwave Electronics," McGraw-Hill Book Company, New York, 1964.

reduce the unloaded Q of the buncher cavity by a factor of 2 or so. However, even though the first-order analysis given above is not sufficiently accurate to give the beam-loading equivalent conductance, it does give a satisfactory answer for the velocity modulation of the beam, which was the information we were interested in obtaining from the analysis.

Since we now know the velocity modulation of the beam at the exit grid of the buncher cavity, we are in a position to evaluate the amplitudes of the space-charge waves that will be excited on the beam in the drift space beyond the buncher cavity. We shall treat the case of an unfocused beam in detail. The case of a beam with confined flow is somewhat easier to analyze and is the model usually assumed in the analysis of the klystron, even though it is not the type of beam used in most klystrons. However, it turns out that the results for the unfocused beam and the beam with confined flow are essentially equivalent, the main difference being that, for the unfocused beam, the major contribution to the ac current comes from the equivalent surface current on the beam, whereas for the beam with confined flow, the ac current is a volume current distributed over the cross section of the beam. During the course of the analysis, the results for the case of the beam with confined flow will be given for comparative purposes.

In the drift space $z > 0$, space-charge waves will be launched because of the ac velocity modulation of the entering beam. At the plane of the exit grid, the radial electric field of the *space-charge waves* is short-circuited and must be zero. This condition can be met by a suitable combination of the fast and slow space-charge waves. If we let \mathbf{E}_r , and E_{rf} be the radial electric field of the slow and fast space-charge waves, we require $E_{rs} = -E_{rf}$ at $z = 0$. But E_{rs} and E_{rf} depend on r according to the first-order modified Bessel function $I_x(pr)$, where p is different for the fast and slow waves. However, for typical beams, $p \sim (5 = ?)$; so the radial dependence can be taken as $I_x(flor)$ with negligible error, in which case the boundary conditions at $z = 0$ can be satisfied without introducing higher-order space-charge modes.

The required boundary conditions on E_r can be met if we choose the two space-charge waves to have the same amplitude and to combine in phase for the ac velocity v_x at $z = 0$. Thus let o_{ls} and v_{st} be the amplitudes of the slow and fast space-charge-wave axial velocities, so that we may write

$$v_g = v_{zs} e^{-ip_s z} + v_{zf} e^{-ip_f z} = u_z (e^{-J'' r} + \text{c.c.})$$

where $p_s = \beta_0 o_l + o^2/c_0$ and $p_f = \beta_0(1 - o_{sf}/o_f)$, and io_e is the effective plasma frequency equal to Fw_p . Introducing these expressions gives

$$v_z = 2v_{zf} \cos p_f z e^{-J'' r} \quad (9.55)$$

where $f\beta_q = (\beta_0 w_g/o_i) = (o_q/v_n)$. Note that v_a is a function of r according to $I_0(P, r)$. However, $\beta_0 a$ is small, so that v_z is almost constant. Thus we may

equate $2v_{zf}$ at $r = 0$ to v_z as given by (9.54) to obtain

$$v_2 = -MV_g \cos p_q z l_0(p_0 r) e^{-J^*} \quad (9.56)$$

for the ac axial velocity at any point in the drift space $z > 0$. In (9.56) we have put V_R for the exciting gap voltage E_{gd} .

For the unfocused beam the ac space charge density p_i is zero from (9.306) we find $J_s = -p_0 u_2$. Consequently, the ac axial current density in the drift space is

$$I_s = V P_o M V_g \cos p_q z I_0(p_0 r) e^{-J^*} \quad (9.57)$$

The surface current J_s will be evaluated later, and will turn out to be more important than the volume current.

To find E_p we use (9.29/ii) to obtain

$$E_p = \frac{\frac{d}{dr} \text{ftPe}}{in} \quad <0, \\ E_{pf} = \frac{U_j - P_f U_0}{D_{fr}} = \frac{d}{dr} \quad >0$$

Hence the axial electric field is given by

$$E_z = -A A f (e^{-\frac{2W}{V}}) = -v_{zf} e^{-JP_o Z} s m p_q z \quad (9.58)$$

Introducing the earlier expression for v_{zf} gives

$$E_z = -p_q M V_g I_0(p_0 r) \sin p_q z e^{-J^*}$$

Note that E_z vanishes at the exit grid, where $Z = 0$.

From (9.29tO) we have

$$-rB_n - j f_l B_{n\perp} = -(P_0 + P_q) j \cdot f'' \quad ' \wedge \wedge$$

$$-rE_{rr} = JP_f E_{rf} = (f_t - f_l q) - \pm v_r \sim -W \cdot f$$

$$1 \text{ a} \quad w_n$$

since $\beta_0 \gg \beta_1$. The boundary conditions on f_r are consequently $\text{sat}, S_{\text{Qf}}$ of the very good degree of approximation. If desired, a small adjustment $\wedge \wedge j$, amplitudes of the two space-charge waves could make $E_{rs} + r f$

exactly at a particular value of r . However, since we have already approximated $I_t(pr)$ by $J_f(\epsilon_0/0)$ for both waves, the present approximation of dropping p_q relative to 0_0 is consistent with our earlier assumptions.

If we examine the expressions for J_z and v_z , we see that, because of the beating or interference between the two space-charge waves, the ac current and velocity vary according to $\cos f_i q z$ in the drift space. Maximum volume current density occurs when

$$2 = "Y \quad (9.59)$$

where $X_q = 2v/p_q$ is the space-charge wavelength and n is an integer.

If we had considered a beam with confined flow as discussed in Sect. 9.3, we should have

$$J_z = \frac{-\bullet}{w - Pv_0}$$

by using (9.116) and (9.12). In this case

If v_z varies according to $\cos(3/2)$, then J_z will vary as $\sin(i_q z)$ for a beam with confined flow. The different behavior of the two beam models arises because of the zero ac volume space charge density in the unfocused beam, a condition that can exist because radial ac motion of the electrons is permitted. The axial current for the two beam models is given by

$$J_z = -N^t t + P^u 0 \quad J_z = -Pffiz$$

The space charge p changes the relationship between J_z and v_z from $J_z = -p_Q v_z$ to $J_z = -p_0 v_0 t / ((i - p v_0))$, which in turn produces the difference in z variation of the current amplitude. In some klystrons where a high power and long drift spaces are used, magnetic focusing is employed. In this case the confined-flow-beam model would be the appropriate one to use. For the beam with confined flow, the axial current density at $r = 0$ would be found to be

$$J_z = \frac{IVPn}{v_0 \% t} \text{ MV, } \sin \text{ " } / J_a z e^{-\bullet} \text{ " } \quad (9.60)$$

The peak current density is a factor co/co greater than the maximum current density for the unfocused beam as given by (9.57). This suggests that the confined-flow beam is superior. This is usually not the case because the surface current J_s for the unfocused beam may contribute in a very substantial way to the total ac axial current, as the following analysis will show.

The surface current density is given by (9.36) and (9.38) p_n
space[^]-charge waves, we obtain - — or the two

$$W > 0 \quad JL - \wedge St^* L$$

When we combine (9.29a) and (9.296), we find

$$7/3 \quad 3JS. \quad 7/3$$

where $C_s I_0(pr) = E_z$. Consequently,

$$J_s = - /, (j8_0 a)(C_1/c - >^* + C_1 j e^{-\lambda} \cdot *)$$

Now $E_{zs} = -(oi_q/jr)v_{zf}$ and $E^{\wedge} = (w_q/jr)v_{zf}$ so we have $C_{1s} = -Uo_q/j_r v_{1s}$, $C_{lf} = (<o_q/J7l)v_{2s}$, where $2u_{z/} = -(ij/i_0)Jlf^{\wedge}$. The final expression for J_a becomes

$$J = \text{—————}^{\wedge} / . (\text{So o j s i n } \wedge \text{e e} - \wedge) \quad (9.61)$$

We shall now compare the relative contributions to the total axial current. The total surface current is given by

$$/s = 2ira - MV_g I_l (l3_0 a) Bmp_q z e^{-J**} \quad (9.62a)$$

and the total volume current flowing in the axial direction is

$$7, = ira - MV_g I^{\wedge} 0 a) \cos P_q z e^{-\lambda} \quad (9.62b)$$

where we have used the result

$$\int_{-n}^{+a} I_0(flo r) 27rr dr = \frac{2ira}{\pi} - I^{\wedge} a)$$

The ratio of the peak amplitudes is

$$(4). \quad 0 \\ (h), \quad (9.63)$$

This gives the very interesting result that the total surface current^{curre^&}_{v0} is [ne case of an unfocused beam is a factor $oj/a > q$ larger than the total \wedge_e current. In fact, the total surface current given by (9.62a) is equal[^] to the total volume current in a small-radius confined-flow beam, that is, ira

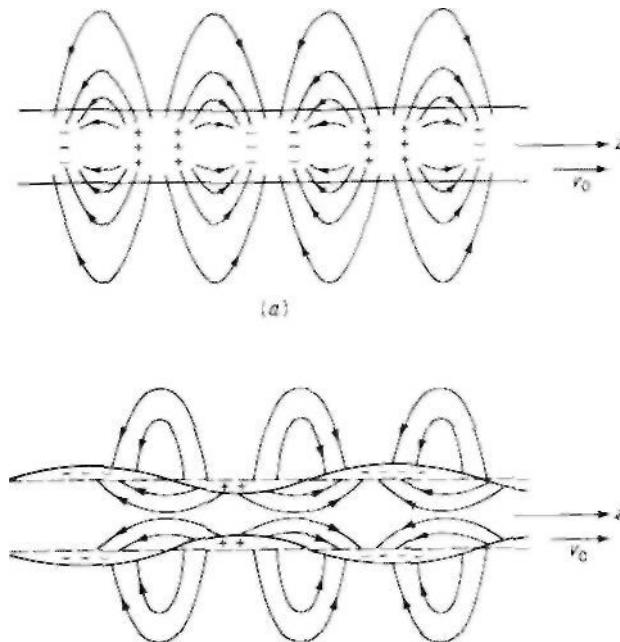


FIGURE 9.9

Electric field lines associated with ac space-charge bunching in (a) a beam with confined (low, (b) an unfocused beam.

where J , as given by (9.60) for confined flow has a peak amplitude equal to that of J_S if we replace $7(\ell, a)$ by $13_n a/2$. We therefore conclude that both types of beams are about equally efficient, at least for short drift spaces, where beam spreading would not be important. It is apparent that ac space-charge bunching is an important mechanism in the production of high-density ac currents in the velocity-modulated beam. For confined flow the ac space-charge bunches are formed within the beam, and for the unfocused beam the ac space-charge bunches appear on the beam surface in the form of a rippled boundary. A sketch of the electric field associated with the two beam models is given in Fig. 9.9. The positive charge shown is only an equivalent charge that accounts for a net migration of electrons out of the region, leaving a net negative charge density less than p_0 , which can be viewed as a superposition of a small positive charge density on the constant dc background density $\sim p_0$.

In judging the relative amplitudes of the volume current and surface current in an unfocused beam, the ratio $to/u_p>$ must be known. Usually (o_q) does not differ by more than a factor of ℓ or so from the plasma radian frequency w_p . In Fig. 9.10, $f_p = < n_p / 2\pi r >$ is plotted as a function of the beam

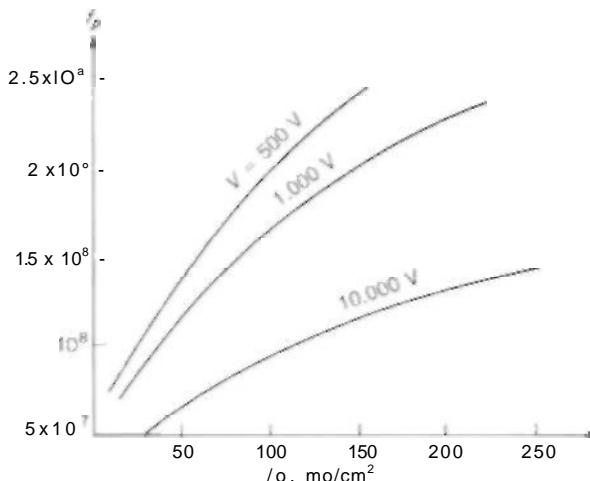


FIGURE 9.10
Plasma frequency as a function of beam current density.

current density J_0 in milliamperes per square centimeter for several values of beam-accelerating voltage V . Note that f_p is proportional to $JQ^2V^{1/4}$.

7 TWO-CAVITY KLYSTRON

A schematic illustration of a typical two-cavity klystron amplifier is shown in Fig. 9.11. The first cavity is excited by the input signal, which can be coupled to the cavity by a coaxial-line loop or a waveguide aperture. The first cavity acts as the buncher and velocity-modulates the beam. The

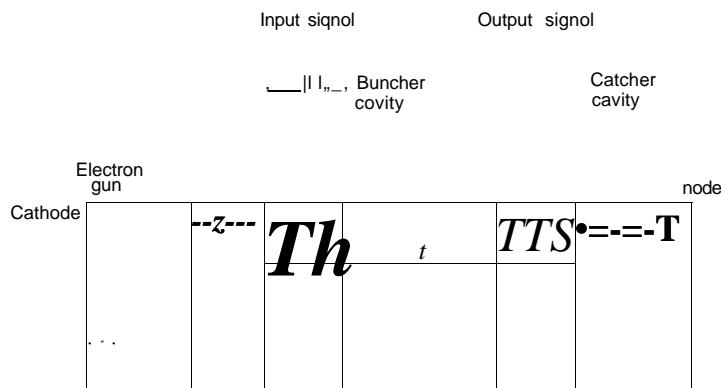


FIGURE 9.11
The two-cavity klystron amplifier.

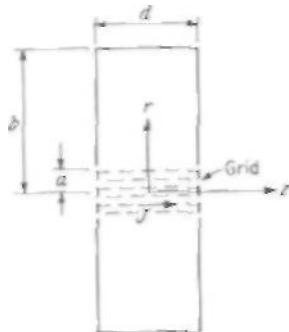
second cavity is separated from the buncher by a drift space of length l , which should ideally be chosen so that the ac current at the second (sometimes called the catcher) cavity is a maximum. The second cavity is thus excited by the ac signal impressed on the beam in the form of a velocity modulation with a resultant production of an ac current. The ac current on the beam is such that the level of excitation of the second cavity is much greater than that in the buncher cavity, and hence amplification takes place. The output signal is taken from the second cavity. If desired, a portion of the amplified output can be fed back to the buncher cavity in a regenerative manner to obtain self-sustained oscillations.

One form of klystron analysis begins with an assumed lumped-parameter equivalent circuit for the output cavity and evaluates the current flowing in this equivalent circuit from an electron beam passing through the cavity by calculating the rate of change with time of the charge induced on the grids at the center of the cavity. This analysis gives a correct picture of klystron behavior, but it fails to illustrate the mechanism of the electromagnetic field interaction with the beam as it actually takes place within the cavity. A more satisfactory approach is to begin with a field analysis that will eventually lead to an equivalent circuit and the basis for a circuit-type analysis of the problem. This is the approach presented below. We first evaluate the field set up in a cavity by the passage of an ac current through the cavity. This leads to an equivalent circuit for the cavity in the vicinity of one of the resonant frequencies. The next step is to evaluate the response of the cavity (or its equivalent circuit) to the passage of a velocity-modulated electron beam on which the ac current is in the form of a propagating current wave. This leads to a definition of the beam coupling coefficient, which is a measure of how effective the modulated electron beam is in exciting a response in the cavity. The third step, which we have presented in Sec. 9.6, is the evaluation of the ac current produced on an initially unmodulated electron beam when it passes through a cavity in which an oscillating electric field exists. These three phases of the analysis substantially provide the complete picture of the operation of a klystron.

of a Cylindrical Cavity

In a klystron cavity it is desirable to have a small grid spacing in order to make the beam coupling parameter $M = [\sin(l_i d/2)]/(p_0 d/2)$ close to unity. The transit angle $f_0 d = wd/v_0$ should be kept small. In addition, a high cavity Q is desired, and this leads to the use of a reentrant cavity. The analysis of the modes in a cavity of this configuration is difficult, and therefore we shall consider instead an ordinary cylindrical cavity. The principal features involved in the excitation of the latter type of cavity by a modulated electron beam are the same as for the reentrant-type cavity.

The cylindrical cavity to be studied is illustrated in Fig. 9.12. The cavity radius is b , and the cavity length is d . Two small cylindrical holes are



8FIGURE 9.12
Cylindrical cavity excited by an axial current.

cut in the center and replaced by grids to allow an electron beam to pass through. The beam radius is a , and is considered very small compared with the cavity radius r . We shall first study the excitation of this cavity by an axial ac current of the form

$$J_z = J e^{Jml} \quad 0 < r < a \quad (9.64)$$

Later on we shall consider a traveling-wave current $J e^{JU_{\text{mod}}(z)}$ of the type existing on a velocity-modulated beam and shall find that, for the latter, the cavity response is modified by the beam coupling parameter M .

In view of the uniformity of the current in the z direction and the axial symmetry, only TM_{11m0} modes are excited. These have E_r , E^θ , H_r , and H^θ equal to zero. It is convenient to introduce the vector potential A_z , which is a solution of

$$\nabla^2 \mathbf{A}_z + k_z^2 \mathbf{A}_z = \begin{cases} -H_0 J & 0 < r < a \\ 0 & r > a \end{cases} \quad (9.65)$$

At $r = b$ we have $A_z = 0$, so that E_z will vanish on the boundary. From (9.65) we obtain

$$E_t = -ju) A_z \quad (9.66)$$

$$\frac{1}{r} \frac{\partial A_z}{\partial r} = \frac{i Q}{dr} \quad (9.66)$$

since there is no z or θ variation.

The natural modes of the cavity are solutions of the equation

(9.67)
of the

where $k_{Qm0} = a_0 \omega_0 \sqrt{M_0 \epsilon_0}$ and ω_0 is the resonant frequency of the TM_{01m0} mode. With no z or θ variation, V^z becomes

and the solutions to (9.67) are Bessel functions. That is,

$$A_{2,0m} = C_m J_0(\gamma j f) \quad (9.68)$$

where p_{0m} is chosen so that $\omega_0(p_{0m}) = 0$, C_m is an arbitrary constant, and for no axial variation $k_{0m0} = p_{0m}/b$ (Sec. 7.4).

For a solution to (9.65) we may choose]

$$A = \frac{v}{L} \frac{\partial}{\partial r} \left(r J_0 \left(\frac{p_{0m} r}{b} \right) \right) \quad (9.69)$$

since the Bessel functions are analogous to the sine and cosine functions and may be used as such in a Fourier series expansion of the vector potential. When we substitute (9.69) into (9.65), we obtain

$$\mathbf{E} \cdot \mathbf{C}_n (k_{0m0}^2 - \omega^2 \mathbf{g}) \mathbf{j}_0 (\gamma j f) = \omega_0 \mathbf{J} \quad 0 < r < a \quad (9.70)$$

since (9.69) is a solution of (9.67).

The following orthogonality property holds:

$$\int_0^b J_0 \left(\frac{p_{0m} r}{b} \right) J_0 \left(\frac{p_{0n} r}{b} \right) r dr = \begin{cases} 0 & n \neq m \\ \frac{b^2}{2} J_1^2(p_{0m}) & n = m \end{cases} \quad (9.71)$$

Thus, if we multiply both sides of (9.70) by $r J_0(p_{0m}r/b)$ and integrate, we obtain

$$\frac{2 \omega_0 J_0(P_{0m}a/b) c^2}{P_{0m}} \quad (9.72)$$

after replacing k_0 by w/c , and similarly for k_{Qm0} . This equation holds for all values of m . We see immediately from this expression that only for those modes for which $w = w_{0m0}$ will the excitation amplitude be large. In addition, we see that, if $co = w_{0m0}$ for a particular value of m , C_n becomes infinite. An infinite response occurs because the cavity is ideal and is driven at one of its natural resonant frequencies. A practical cavity has a finite Q and will not respond with an infinite amplitude. As shown in Chap. 7, the effect of a finite Q is to replace the resonant frequency by

$$1 + \frac{1}{2Q_{0m0}}$$

where Q_{0m0} is the Q of the TM_{0m0} mode. The unloaded cavity Q is given

by (7.57) as

$$Q = \frac{P_{out}C}{(1 + b/d)w_{0m0}8}, \quad (9.73)$$

Thus the excitation amplitudes for a cavity with finite Q are given by

$$C_{zz} = -\frac{2Mo<fc/cV>(p_{0m}q/6)}{Pa_mM^2(p_{0m})(<o_{0m0} + w)(w - w_{0m0} - >_c^2/2Q^2)}, \quad (9.74)$$

If we choose $t_0 = t_{010}$, then C_{zz} will be large and all the other C 's will be small. In this case

$$\dot{A}_z = \frac{n_i l^{P+n_r}}{C, J_0} \cdot \frac{2>oO'^3J^r_1/(J_{01}a/6)Q}{6v f(p_{01})w} \cdot \frac{1}{p_{01}rx} \quad (9.75)$$

where w and Q now refer to the TM_{010} mode. Note that A_z is proportional to the Q , and hence a high Q is desirable. If the cavity is coupled to an external load, we must replace Q by the loaded Q , Q_L .

The total ac current is $I = wa^2J$. Also, since $a \ll 6$, we can replace $J(Pa^2/b)$ by $p_{01}a/2b$. In the vicinity of the resonant frequency w_{010} , that is, for $t_0 = w_{010} + A$ to, we find that the electric field E_z is given by

$$E_z = -J^* = \frac{1}{27r^26_0J^2(p_{01})(At_0-yt_0/2Q)} \quad (9.76)$$

If we introduce an equivalent voltage V_0 as the line integral of E_z across the cavity gap at $r = 0$, that is, $V_n = -E_z d$, we may define an admittance Y_e for the cavity as follows:

$$y_{zz} = -V_n / I = Aa / (27r^26_0J^2(p_{01})Q) \quad (9.77)$$

For a lumped-parameter LCG_0 circuit with resonant frequency $w_{010} = (LC)^{-1/2}$ as in Fig. 9.13, we have

$$y_{zz} = G_0 + 2jC\omega_0 = G_0 / (1 + j\omega_0 Q)$$

where $Q = \omega_{010}C/G_0$. Comparison with (9.77) shows that the equivalent cavity conductance G_0 is given by (note that $\epsilon_0 = Y_0/c$)

$$G_0 = Trb^2Y_0J/(p_{01})^2 \quad (9.78)$$

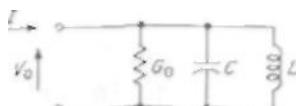


FIGURE 9.13
Equivalent circuit of excited cavity with no $t_0 < 0$.

The above equivalent circuit thus seems a possible circuit to represent the cavity in the vicinity of the resonant frequency w_{01} , of the TM_{01} , mode. However, we must show that it correctly accounts for the properties of the cavity even in the absence of the beam current J .

For the TM_{01} , mode, the energy- stored in the electric field is given by [we assume $E_r = J_u(p_0 r/b)$]

$$\frac{4 \pi \epsilon_0 A_0}{2 \pi} dz = -7 b^2 dJ'f(p_0) dz$$

The corresponding voltage across the cavity at $r = 0$ is $V = E_r d = d$. Since $Q = 2aiW_b/P_b$ the power Joss in the cavity is given by

$$\frac{P_i}{Q} = \frac{2 \pi}{V' Q}$$

If we define a conductance G so that $P_i = |V' G|$, we find that

$$G = \frac{2P_i}{V'^2} = \frac{4 \pi}{V' Q} = \frac{T r e_0 b^2 o >_{\text{m}} J' f(p_{as})}{dQ} = G,$$

since $<_o = w_{010}$. Thus the two definitions for the cavity conductance lead to consistent results.

If required, C and L are given by $t_{0nw}C = GG$, and $t_{0fU}LC = 1$. Thus L and C can be found from the known values of Q and G .

Cavity Excitation by a Velocity-Modulated Beam

In a velocity-modulated unfocused beam the ac current is predominantly the ac beam surface current, $I_s = 2t r a J_s$, where J_a is given by (9.61) since $co \gg io_q$ for typical beams. The z dependence of the current is $\sin(i^2 z e^{-\alpha z})$ when z is measured from the output grid of the buncher cavity. The output cavity should be located a distance $l = (n + j X A_c / 2)$ from the input cavity so that $\sin f_l q z$ will equal unity and I_s will have a maximum. Because $f_l q \ll l$, we have $f_l d \ll 1$ and the ac current variation with z can be taken as $e^{i f_l z} e^{-\alpha z}$ in the output cavity, where we have replaced z by $l + z$, so that the new origin is at the center of the output cavity. Thus the cavity is excited by a traveling-wave current.

When the ac current has a z dependence, all the TM_{lm} , modes are excited. The solutions for A_t are thus of the form

$$G_{lm}, t(l) = \frac{1}{2} \int_{-l}^{l} J_0 \left(\frac{r i \pi z}{d} \right) \cos \left(\frac{r i \pi z}{d} \right) dz$$

We then have

$$E_r = E_0 \cos \left(\frac{2 \pi n z}{\lambda} \right) = f_l J_2(r, z) \quad (9.79)$$

in place of (9.70). However, if $OJ = <^0_{010}$, only the $m = 1, 7, -Q$ mode is excited with a large amplitude. To find C_{U_0} , we multiply (Q 7 $rJ_0(p_0m r/b) \cos(n\pi r z/d)$ with $m = 2, n = 0$ and integrate over r and obtain

$$\text{£-10} \quad \int_{-d/2}^{2M0} (ki_{10} - ki) by?_{(Poi)} J_0 / J_t(r) r J_0 dr \sim e^{-jB_z} dz$$

This latter result shows that the excitation amplitude is modified by factor

$$\int_{-d/2}^{d/2} d/za \sim e^{-jB_z} \sin(p_0 d/2) M$$

which is the beam coupling factor. The integral over r may be replaced by $a J_s$ without the factor e^{-jB_z} since $J_2 = 0$ except at $r = a$. $J_s(p_0 a/b) \sim 1$ for $a \ll b$. Hence, for $w = <^0_{D_{OU}} |a\rangle$, we obtain for E , $r = 0$,

$$E = 2 \cdot 6 \% j;^{-i}(p_{01})(A_w - A_{w_{010}}/2Q) \quad (9.80)$$

where $7,$ is obtained from (9.62a) and is

$$7, = 27ta - M/\lambda/V^* \sim TT-a^e, 'doMV_e \quad (9.81)$$

since for $\beta_0 a \ll 1$ we have $\beta_0/\beta_0 a = \beta_0 a/2$ and are assuming that an optimum length l of drift space is employed so that $\sin P_q l$ equals unity. If we compare (9.80) with (9.76), we find that the only change in E , is replacing the current $/$ by $7,$ and multiplying by the beam coupling coefficient M and an irrelevant phase factor $e^{-jB_z}.$

The voltage developed across the cavity is modified by the factor M . Consequently, the effective current that flows in the equivalent circuit I product the voltage V is

$$I, MI, \quad \langle M \rangle J$$

from which it is seen that the term beam coupling coefficient is $\langle M \rangle J$ appropriate. For good coupling between the beam and the cavity « transit time t_{0d} is required.

When an external load is coupled to the output cavity, I can be represented by an additional conductance G_L in shunt $\wedge t$ where equivalent circuit of the output cavity is shown in Fig. 9-7. I is the equivalent current that flows in the circuit supplied to the external load G_L is

$$Po = \frac{1}{2} (G_0 + G_y) \langle XG, \quad (9.83)$$

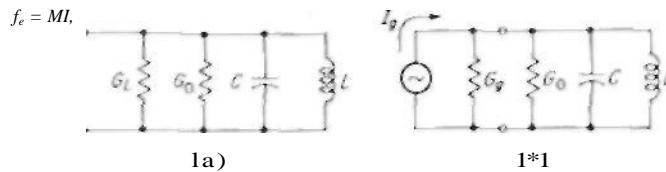


FIGURE 9.14

(a) Equivalent circuit of output cavity; (b) equivalent circuit of input cavity.

A similar equivalent circuit may be assumed for the input cavity, as in Fig. 9.14b. For identical cavities the cavity conductance is G_0 for both. If we assume that the beam produces negligible loading on the input cavity, the total conductance in the equivalent circuit of the input cavity is $G_0 + G^*$ where G^* is the equivalent conductance of the signal source.!

The input power to the buncher cavity at resonance is

$$P_{in} = \frac{1}{2} \frac{H/G_0}{(G_g + G^*)^2} \quad (9-84)$$

which results in a voltage

$$L = i I P^{1/2}$$

developed across G_0 . This is also the velocity-modulating voltage. We can now evaluate the power gain, or amplification, of the klystron. Combining (9.81), (9.83), and (9.85), we obtain

$$\begin{aligned} A &= \frac{E_O}{P_{in}} = \frac{V/G_L(G_0 + G^*)}{I^2 G_0 / (G_0 + G_L)} \\ &= \frac{Tr a^2 Y_0 p_0}{C} \frac{V^2}{G_Q^i (G_0 + G_L)} \quad (9.86) \end{aligned}$$

^tThe conductance representing the beam loading on the input cavity is given by $G^* = 1/2(I_0/V)M[M - \cos(fS_0 d/2)]$ and is easily taken into account by adding it to $G_0 + G_M$ when it is not negligible. Note that I_0 is the dc beam current and V the dc accelerating voltage. See Spangenberg, *op. cit.*, chap. 17.

As an example consider $G_0 = G_{\text{cav}}$, and a beam radius $a = ($ also assume a beam current density of 100 mA/cm^2 and an voltage of $1,000 \text{ V}$, we find, from Fig. 9.10, that $w = i 02?wvP^{tift}8 / \bullet = 10''$, $a = 6.28 \times 10''$ and $\beta_0 = w/c_0 = 33.6 \text{ rad/cm}$. Hence $* \text{ For } 6.72$. Since $u > p < e w$, we have $jia - /i_a a$, and Fig. 9.6 then shows $aj_q * 0.10J_p$. Using these data, the power amplification A is found to be $0.094(y_i/G_{\text{cav}})^2 M^4$. To evaluate G_0 , we make use of (9.78) and have $p_{01} = 2.405$, and since $p_{ni} = 0.106$, we get

$$d = \frac{2.405c}{6.28 \times 10''} = 115 \text{ cm}$$

To keep p_{od} small, we must choose d very small. If we take $c/ = 0.05$, $f_n d = 1.68$ and $M^4 = 0.62$. It would be desirable to make d even smaller but then the Q , and hence G_0 , become small for the type of cavity we are considering. For a copper cavity we find from (9.73) that $Q = 785$, which is not very large. If a reentrant-type cavity were used, a Q about 10 times larger could be obtained. Using (9.78) gives $G_0 = 0.06V_0$. The power amplification is thus 16.2, or 12 dB. Considerably higher gain would be obtained by using a reentrant-type cavity since d could then be made smaller and still a high unloaded Q maintained. However, even with the nonoptimum cavity that we have considered, the gain is quite good. For the particular example we have evaluated, the beam loading conductance G_h is very small compared with the cavity conductance G_0 . For a more efficient cavity with a much higher unloaded Q , the cavity conductance G_0 would be much smaller and the beam loading conductance G_h might not then be negligible.

In order to obtain greater power gain than can be obtained from a two-cavity klystron, multicavity tubes are used. The gain increases exponentially with the number of cavities employed. In the multicavity klystron the first cavity is again used to provide the initial velocity modulation of the beam. The last cavity in the chain is used as the output cavity. Intermediate cavities are kept unloaded by any external circuits and used to increase the modulation and hence the ac current on the beam. Power gains of 50 to 60 dB can be achieved with multicavity klystrons.

9.8 REFLEX KLYSTRON

The reflex klystron is an oscillator tube with a built-in feedback mechanism. It uses the same cavity for bunching and for the output cavity. A sketch of the reflex klystron is given in Fig. 9.15. The operation is as follows: When an initial ac field is applied to the cavity, the beam will be decelerated and reversed (reflected) by the large dc field set up by the reflector electrode at potential $-V_r$. Thus the beam is able to pass through the cavity again, but in the opposite direction. The choice of the reflector voltage V_r , the beam can be made to pass through the cavity again, but in the opposite direction.

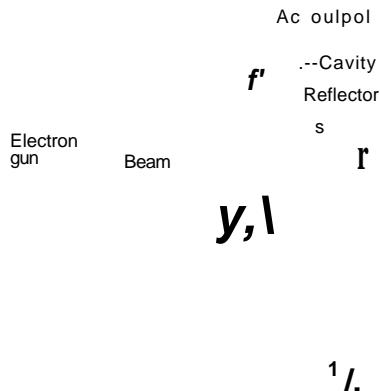


FIGURE 9.15
The reflex klystron.

cavity on its return flight when the ac current phase angle is such that the field excited in the cavity by the returning beam adds in phase with the initial modulating field. The feedback is then positive, and oscillations will build up in amplitude until the system losses and nonlinear effects prevent further buildup.

When the velocity-modulated beam enters the drift space, it is subjected to a constant decelerating field V_r/s , where s is the cavity output grid-reflector spacing. As a result, the beam propagation constant $\beta_0 = \sqrt{\omega_0^2 - V_r^2/s^2}$ is gradually reduced to zero, and then increased back up to $-\beta_0$. The total phase change undergone by the ac current on the beam will be given by

$$B = 2f'' p_0(z) dz$$

where z_m is the maximum distance an electron can penetrate into the drift space. We can evaluate I in terms of the transit time T for an electron to return to the cavity. We have, in the drift space,

$$\frac{dv_a(z)}{dt} = -\frac{q}{s}$$

which integrates to $v_a(z) = v_0 - T(V_r/s)$. Hence

$$v_Q(z_m) = v_0 - \frac{V_r T}{s} = 0$$

which gives $T/2 = v_0 s / V_r$. The return time is equal to $T/2$ also, so that

$$\theta = wT = \frac{2v_0 s w}{V_r} \quad (9.87)$$

If we let V_g be the accelerating-gap ac voltage, the ac beam current reflected back through the cavity is given by (9.62a) when f_{0z} is replaced

by 0 and $/3,2$ by $P_q \theta / P_o = t a_H d / u$. Thus

$$I_s = \frac{J T r a^{\wedge} P o P o M V}{m_i} \sin \frac{w O}{t o} \frac{J L - e^{\sim J n}}{}$$

where we have approximated $/,(/3,0)$ by $(S_o a / 2)$. In a reflex klystron usually quite small, so that $\sin w_0 / W$ may be replaced by $u > 0^{\circ}$. The effective current for excitation of the cavity is $I_e = 7, A_f$, and is given by

$$B_0 f > e^{\wedge} = -f M^{\wedge} \sin \theta + j \cos \theta \quad (9.88)$$

where I_{dc} is the total dc beam current $T r a^{\wedge} p_0 u_0$ and V_a is the accelerating voltage from which $v\% = 2nV_0$.

The ac electronic admittance of the beam is denoted by

$$Y_e = - = -6(s \sin \theta + j \cos \theta) \quad (9.89)$$

The equivalent circuit of the reflex klystron consists of the electronic admittance Y_e in shunt with the equivalent circuit of the loaded cavity, as in Fig. 9.16. Oscillations can take place when the net conductance is less than zero, or more specifically when

$$y_e + (G_L + G_0) \left| \frac{2 \text{ Aw}}{1 + 7 \frac{w_0}{10} Q_L} \right| = 0 \quad (9.90)$$

where Q_L is the loaded Q of the cavity. Since 0 is a function of the reflector voltage, as given by (9.87), oscillations depend on an appropriate choice of

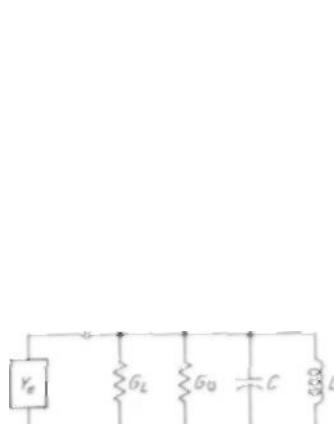


FIGURE 9.16
Equivalent circuit for a reflex klystron.

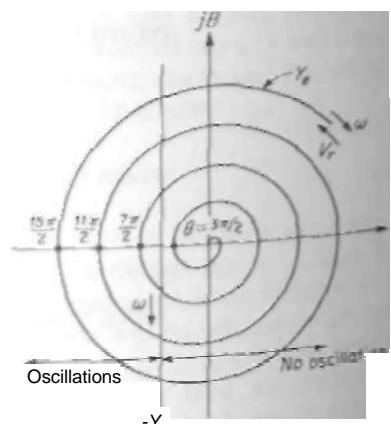


FIGURE 9.17
Admittance diagram for a re

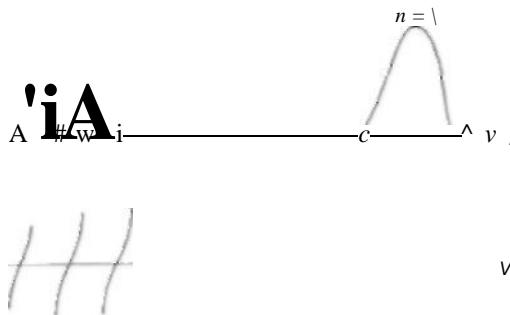


FIGURE 9.18
Reflex-klystron tuning curves.

V_r . In Fig. 9.17 we have plotted the admittance Y in polar form as a function of θ . Note that $|Y_e|$ increases with θ . On the same plane we have plotted the negative cavity admittance

$$-F = -(G_0 + G_L)l + j \frac{2 Aw}{Q_L}$$

which is a straight line parallel to the jB axis at $G = -(G_0 + G_L)$, provided we assume G_a, G_L, Q_L independent of w in the vicinity of the resonant frequency w_{010} . The construction shows that oscillations are possible for θ in the vicinity of $3\pi/2, 7\pi/2$, etc., since in this region $G_c + G_0 + G_L < 0$. In addition, oscillations will take place for a range of values of V_r about the points that make $\theta = 3\pi/2 + 2n\pi$. Each value of n gives a mode of oscillation. In typical klystrons as many as seven or more modes of oscillation can be obtained. For stable oscillations $Y_c + Y = 0$, and consequently the frequency of oscillation varies as V_r is varied to tune across a given mode. Typical tuning curves giving power output and frequency as a function of reflector voltage are shown in Fig. 9.18. Physically, the various modes arise because of the increased transmit time for electrons into the drift space when V_r is reduced. Oscillations occur when the transit time T equals $(l + n)f'$ or $\lambda + n$ ac periods since the ac current has the proper phase under these conditions.

Commercially available reflex klystrons range from small-size units producing 100 mW of power up to units capable of delivering several watts of power under continuous operation. Klystron amplifiers employing two or more cavities are available in a size range from units capable of handling a few hundred milliwatts up to several hundred kilowatts of amplified output power.

9.9 MAGNETRON

This section is devoted to a qualitative description of the magnetron. The basic structure of a magnetron is a number of identical resonators arranged in a cylindrical pattern around a cylindrical cathode, as shown, ^{**} Fig. 9.19. A permanent magnet is used to produce a strong magnetic field B_0 normal to the cross section. The anode is kept at a high positive voltage V_a relative to the cathode. Electrons emitted from the cathode are accelerated toward the anode block, but the presence of the magnetic field B_0 produces a force $\sim evB_0$ in the azimuthal direction which causes the electron trajectory to be deflected in the same direction. If the cathode radius is a and the anode radius is b , the potential at any radius r is $V(r) = V_0 \ln(r/a)/\ln(b/a)$. The velocity of an electron at this radius is given by

$$v(r) = |2_e V(r)|^{1/2}$$

The electron can execute circular motion, at the radius r , about the cathode if the outward centrifugal force mv^2/r and the radial electric field force $-eE_r = eV_a/[r \ln(6/a)]$ are exactly balanced by the inward magnetic force $ev(r)B_0$. For circular motion at radius r , we therefore have

$$\frac{mv^2}{r} + \frac{eV_a}{rln(b/a)} = evB_0 \quad (9.91a)$$

or since $v = (o_e r)$, where o_e is the electron's angular velocity,

$$we \sim v^2 Q U_e + \frac{eV_a}{r^2 ln(b/a)} = 0 \quad (9.91b)$$

For later reference, we solve (9.91) for the cathode-anode acceleration

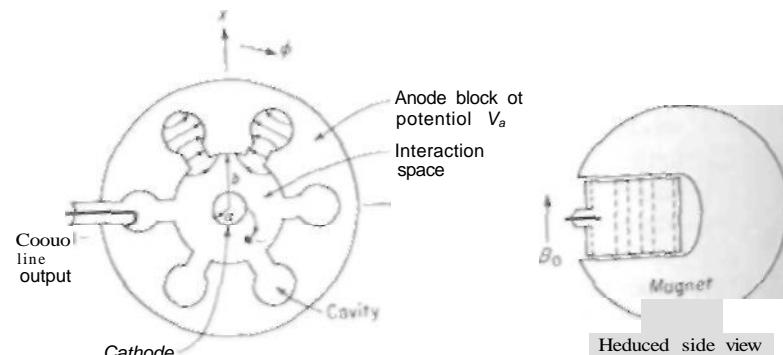


FIGURE 9.19
A multicavity magnetron.

voltage V_Q :

$$\mathbf{v}_a = \frac{1}{kr^*} \mathbf{f} \mathbf{a} \frac{bM}{a/\sqrt{\epsilon_0}} \frac{\mathbf{w}_r S}{\mathbf{J}} \quad (9.92)$$

This value of V_a^* will permit an electron to execute circular motion at a radius r and with an angular frequency ω_e . If now there is present an ac electromagnetic field that propagates in the azimuthal direction with a phase velocity equal to the electron velocity w_r , strong interaction between the field and the circulating electron cloud can take place. The possibility of this type of electromagnetic field being present is discussed below.

The multicavity magnetron is a periodic structure in the azimuthal, or $\langle l \rangle$, direction. If there are N cavities, the period in ϕ is $2\pi/N$. According to Floquet's theorem, each field component can be expanded in the form

$$n = -x$$

$$= \sum_{n=-x}^L (\$, Lr) e^{\sim JM' - J^{nm}} \quad (9.93)$$

where the period $p \sim 2\pi/N$. But since the structure closes on itself, $(Mr, 2TT) = i\langle r, 0 \rangle$. The only possible values of i that will make $|i|^2$ equal a multiple of $2\pi r$ are

$$fr = m \quad m = 0, \pm 1, \pm 2, \dots \quad (9.94)$$

With the value of $/3$ specified, a corresponding frequency w becomes specified, say io_m , which is the resonant frequency for the m th mode. In other words, when $w = m\omega$, we obtain a value m for p''_m . Thus a typical field component will have the form

$$\langle l \rangle_{in}(r, \langle t \rangle) e^{jn - nm} = \sum_{n=-x}^x L^{\wedge} (r) \langle \cdot, \cdot \rangle + \langle \cdot, \cdot \rangle + \cdot \cdot \cdot$$

The phase velocity in the azimuthal direction θ for the n th spatial harmonic of the m th resonant mode is

$$\theta_{in} = \frac{\omega_m r}{m + nN} \quad (9.95)$$

at the radius r ; that is, angular phase velocity is $w_m/3, \dots$

The usual mode employed in a magnetron oscillator is the *TT* mode, where the phase change between adjacent cavities is *TT* rad, or 180° . Each cavity with its input gap acts as a short-circuited transmission line a quarter wavelength long, and hence has a *maximum electric* field across the gap. For the TT mode the field is oppositely directed at adjacent cavities. A sketch of the electric field lines in two cavities is given in Fig. 9.19. For the TT mode,

$Pm4 > = mtf >$ must equal $-n$ for a change in $\langle \rangle$ equal to one 1
Hence $m = N/2$, and the phase velocity for the r th spatial harmonic becomes $v_p \sim N$

$$N(l + 2n) \quad (9.96)$$

In order to obtain interaction between the electron cloud and the spatial harmonics at a particular radius r , we must choose V_v . So that $w_{,r} = v(r) = v_{p,nN/2}$ or

$$AT(1 + 2n)^{1/2} \quad (9.97)$$

The required voltage V_a to obtain synchronism between the electron cloud and the ac field may be found from (9.92). If we choose a value of r midway between the cathode and anode., that is, $r = (b + a)/2$, and note that in typical magnetrons b/a is small enough so that $\ln(b/a) = 2(b - a)/(b + a)$, we obtain

$$V = \frac{6^2 - o^2}{m n N} B_0 \quad (9.98a)$$

in general, and

$$K = \frac{2w_{N/2}}{N(l + 2n)} \frac{6^2 - a^2}{B_0} \frac{2(\theta_{N/2})}{nN(l + 2ra)} \quad (9.98ft)$$

for the $-$ mode, where $m = N/2$.

From a physical viewpoint the synchronism between the electron cloud and the r th spatial harmonic and the ac field implies that those electrons located in the field where $f_{,r}$ acts to slow down the electrons will give up energy to the field. As the electrons slow down they move radially outward [see (9.91)], and eventually are intercepted by the anode. Electrons that are accelerated by the ac field move in toward the cathode until they get into a proper phase relationship such as to give up energy to the field. When latter happens, they begin to slow down and spiral out toward the $^m \wedge ^o \wedge$ Thus the only electrons that are lost from the interaction space are too; that have given up a net amount of energy to the ac field.

The ac power may be coupled out from one of the cavities coaxial-line loop as shown in Fig. 9.19 or by means of a waveguide.

9.10 O-TYPE TRAVELING-WAVE TUBE

The ordinary, or O-type, traveling-wave-tube employs a $^m \wedge ^a \wedge$ cused electron beam and a slow-wave structure such as a helix, disc $^s \wedge ^d$ Chap. 8. The electron-beam velocity is adjusted to be approximately

the phase velocity for an electromagnetic wave propagating along the helix. Under these conditions a strong interaction between the beam and the field can take place. From another viewpoint we can consider the presence of the slow-wave circuit to modify the space-charge wave-propagation constant in such a manner that it becomes complex and represents a growing wave. We shall present a more satisfactory picture of the gain mechanism after we have analyzed a particular tube configuration in detail. A full appreciation of the physical principles involved is somewhat difficult to obtain without a detailed study.

For simplicity we shall use the sheath-helix model discussed in Sec. 8.10 and an axially confined flow beam model (S_{\parallel} , infinite) of the type treated in Sec. 9.3. In addition, we shall assume that the beam completely fills the region interior to the helix. This assumption is not true in practice, but we make it, nevertheless, in order to simplify the analysis. The basic principle of operation of the tube is not changed by this assumption. The traveling-wave tube is operated in an axially symmetric mode; so all field quantities will be independent of the angle $\langle f \rangle$.

Figure 9.20 illustrates the construction of a typical traveling-wave tube. The main components are an electron gun, a helix, a solenoid to produce the focusing field B_{\parallel} , and suitable input and output ac coupling to the helix. The helix is taken to have a radius a and a pitch angle $\langle l \rangle$. It is approximated by a cylindrical sheath with infinite conductivity along the direction of the winding and zero conductivity in the perpendicular direction.

In Sec. 8.10 it was shown that both TM and TE modes were required in order to satisfy the boundary conditions at $r = a$. However, for a beam with axially confined flow, where only a z component of ac velocity is permitted, the TE modes are not affected by the beam since these have $E_z = 0$. Hence the field components for the TE mode for $n = 0$ are those given by (8.72) in Sec. 8.10. Similarly, for $r > a$, that is, outside the helix, the field components for the TM mode are those given by (8.72) in Sec. 8.10.

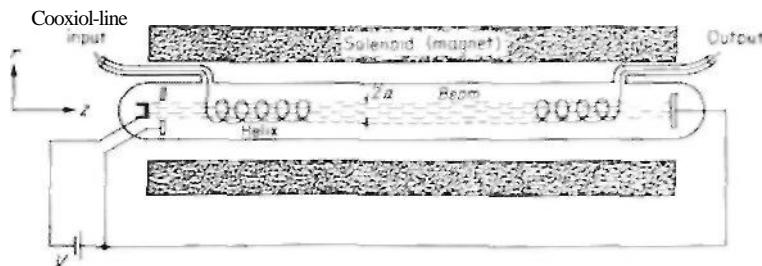


FIGURE 9.20
O-type traveling-wave tube.

Inside the helix region the TM field in the presence of the beam is that a_v in Sec. 9.3. However, the pertinent equations are repeated here for convenience. The vector potential A , is a solution of (9.16a),

$$\nabla_t^2 A_r + p^2 A_r = 0$$

where

$$* \rightarrow - (\leftarrow - * \bullet) \quad 1 - \frac{0}{\infty} \text{ to } I \quad \underline{\text{Po-P}} \quad (9.99)$$

For the present problem p^2 will turn out to be negative; so we shall replace p^2 by $-g^2$. The solution for A_z is then proportional to $I_0(gr)$. Since E_z is proportional to A_z , we can choose

$$K_z = a_0 I_0(gr) e^{-J_i z}$$

where a_0 is an amplitude constant. The field components E_r and H_{ib} are readily found from Maxwell's equations; i.e.,

$$E_r = \frac{j/i}{p^2 - K_z^2} \hat{U} \sim \frac{!> E_r}{\bullet \bullet} \quad H_i = \frac{k_0}{\mu_0 \epsilon_0} \hat{P} \sim \hat{Y} \hat{O} E_r$$

Thus we can write the following expressions for the fields in the two regions for the $n = 0$, or axially symmetric, case:

For TE modes,

$$H_2 = c_0 J_0(O) e^{-j\theta z}; *$$

$$M_r = -c_0 U h r e^{-j\theta z} \quad r < a$$

$$E_{rr} = -c_0 l_0 h r e^{-j\theta z}$$

$$B_z = d_0 K_a(kr) e^{-j\theta z}$$

$$H_r = -d_0 K_x(hr) e^{-j\theta z} \quad r > a$$

$$E_{rz} = -d_n K_l(hr) e^{-j\theta z}$$

For the TM mode,

$$E_z = a_0 I_0(gr) e^{-\lambda}$$

$$E_r = \frac{J P g}{h} q_0 I_0(gr) e^{-\lambda} \quad r < a$$

$$E_{\perp} = b_n K_0(hr) e^{-\lambda}$$

$$E_r = -\lambda b_0 K_i(hr) e^{-\lambda} \quad r > a$$

$$H_{\perp} = -\frac{J}{\lambda} \pm b_{li} K_i(hr) e^{\lambda}$$

where $h^2 = p^2 - k^2$.

The boundary conditions at $r = a$ for the sheath helix are given by (8.69). For the present problem they yield

$$\frac{-juv-u}{h} c_0 / (ia) \cos tli + a_0 I_a(ga) \sin ip = 0$$

$$d_0 K_l(ha) \cos \lambda i + b_0 K_Q(ha) \sin tli = 0$$

$$a_0 h(\$^a) \cos I/I + \frac{\text{Mo}}{\text{Mo}} c_u / (ha) \sin 0$$

$$= b_n K_0(ha) \cos \lambda i - d_0 f [(ia) \sin \lambda i]$$

and

$$c_0 I_0(ha) \sin ib + \frac{J W f_0 g}{2} a_n I_l(ga) \cos \lambda i$$

$$= d_0 K_0(Aa) \sin \lambda A - \&,,/?,(/m) \cos i/r$$

If we solve for c_0 and c_0 from the first two equations and substitute into the latter two equations, we obtain two homogeneous equations for a_0 and b_0 . For a nontrivial solution the determinant must vanish. Equating the determinant to zero gives

$$\frac{8}{I_0(ga)} \frac{I_l(ga)}{k_0} / i^a \tan^2 \lambda i - \frac{I_p(ha)}{I_l(ha)} + \frac{K_0(ha)}{K_l(ha)}$$

For most traveling-wave tubes the parameters are such that ha are large. In this case the ratio of the Bessel functions $\frac{J_1}{J_0}$ approaches unity, and we obtain

$$g = 2 \frac{h^3}{J_0} \tan^2 \frac{h}{\lambda} \quad (9.101Q)$$

From (9.99) we have

$$g^2 = h^2 \left(1 - \frac{\frac{Po}{Po-P}}{\sin \frac{h}{\lambda}} \right)^2$$

and hence

$$1 - \frac{\frac{Po}{Po-P}}{I} = \frac{h^2}{2 \tan^2 \frac{h}{\lambda} - 1} \quad (9.1016)$$

where $h^2 = f^2 - k^2$. The above is a sixth-degree equation in p , and cannot be solved exactly.

Since we are dealing with a slow-wave system, β^2 will be large compared with k^2 , and $h^2 = f^2$. In addition, we can equate $k_0 c \omega_0$ to 0, since the phase velocity of the helix in the absence of a beam is chosen equal to the beam velocity v_b . That is, $k_0 \sin 4i$ is the propagation constant for the helix, and for β small, $\sin iji$ can be replaced by $\tan iji$. We thus obtain

$$\frac{Po}{Po-P} \approx \frac{2p^2}{PJ} - 1$$

We now assume that $\beta = y_3(1 + \delta)$, where δ is small. With this substitution we get

$$\begin{aligned} &= \delta^2(1 + 4\delta + 2\delta^2) \\ &= 4S^6 + 16S^4 + 20S^2 + 8S^3 + 8S \end{aligned} \quad (9.102)$$

For δ small, we can drop all but the term involving the lowest power of δ . This is the cubic term, and thus

The three solutions for δ are $(w_c/\omega_0)^{2/3}$ multiplied by the cube

of

- 1, which are - 1 and $(1 \pm \sqrt{v}/3)/2$. Hence

$$* " I^{ \% >^2 } \quad (9.103a)$$

$$1 / 01. \sqrt{v}/3 \quad (9.1036)$$

$$(9.103c)$$

Since t_0/c_0 is small, the assumption that δ was small is justified. The corresponding propagation constants are

$$JPt = JPo \left| 1 - \frac{1}{2} \frac{">,"}{w},^{2/3} \right| \quad (9.104a)$$

$$JP\pm = JPu \left| 1 + \frac{1}{4} \frac{\lambda_n}{w},^{2/3} \right| \quad (W73) \quad [9.1046]$$

$$JP3=.)P0 \left| 1 + \frac{1}{4} \frac{w_p}{co},^{2/3} \right| \quad i + fVf \quad (9.104c)$$

The first solution corresponds to a wave with a phase velocity slightly greater than the beam velocity. The other solutions have phase velocities slightly less than the beam velocity, and in addition j/S_2 corresponds to a decaying wave, whereas $\sqrt{3}_3$ corresponds to a growing wave. The growth constant a_g is

$$\ll, - * - T - \frac{V3}{w_r},^{2/3} \quad (9.105)$$

If all three waves are present at the input, only the latter wave will predominate at the output.

There are additional solutions to the eigenvalue equation (9.1016). We should expect a wave propagating in the $-z$ direction, with $(i \sim -k_Q \csc \theta \sim -0_0)$, which is not significantly perturbed by the beam. We therefore assume that $f_i = -jG, (1 + \delta)$ and consider δ small. Substituting into (9.1016) and retaining the smallest power term in d give

$$\delta = - \frac{1}{32} \frac{1}{co}$$

Hence a fourth solution is

$$JPA = -JPo \left| 1 - \frac{1}{32} \right| = -JPo \quad (9.106)$$

The remaining two solutions of (9.1016) give values for $\sqrt{3}$ approximately

equal to $\pm k_0$. However, the eigenvalue equation (9.1016) is a solution to the true eigenvalue equation (9.100), obtained by assuming a_{oxini} and ha are large and that f_i is large compared with k_0 . Therefore h^{Sa} solutions $p = \pm k_0$ to the sixth-degree equation (9.1016) are not $^{two}_{olutions}$ of (9.100) and do not correspond to physical waves.

The ac current and velocity are given by (9.14) and (9.13Q) as?

$$v = u_o(Po - P) \quad f9n^2 07)$$

$$J = -J_- \rightarrow (fio - P)^* e^{-s_p E} \quad (9108)$$

These equations show that v and J are negligible for the three waves for which p is significantly different from p_n . Thus v and J arise from the first three slow waves discussed. The fourth wave can be excited by reflection at the output end of the tube. If it is reflected at the input end also, it will be amplified and, with continued reflection and amplification, will result in oscillations. To avoid this undesirable feature, an attenuating resistive vane or an integral ferrite isolator is built into the traveling-wave tube.

At the input end we must have the total ac current and velocity associated with the three forward slow waves vanish. Thus the initial conditions at the input $z = 0$ are

$$t/[+ J_2 + J_3 = 0 \quad v_l + v_{.2} + v_3 = 0$$

When we assume that

$$E_z = I_0(p_0 r) \{ C_{ie} e^{-\lambda z} + C_{tf} e^{-j \omega t} + C_3 e^{-\alpha z} \}$$

and make use of (9.108) and the initial conditions, we find that

$$^2 = g/Zir/3 \quad -^* = e^{-j 2 \pi / 3}$$

Consequently, all three waves at the input have equal magnitudes; i.e., we find that $G_j = C_2 = C_3$. The growing wave will have an amplitude equal λ one-third that of the input signal. Therefore the amplitude gain c traveling-wave tube is

$$\frac{E_0}{E_i} = \frac{1}{3} e^{-\alpha l}$$

where a_e is given by (9.105) and l is the tube length. The power gain g decibels is

$$A = 20 \log 0.333 + 20 a_e / \log e$$

$$= -9.54 + 3.75/V \quad \omega = 2/3 \quad 9109)$$

With the aid of the preceding results we can now describe the physical mechanism of the gain. We note that the growing wave has a phase velocity slightly less than the beam velocity. This growing wave is the perturbed slow space-charge wave. The ac kinetic-power density of the fast and slow space-charge waves are [see (9.49), (9.13a), and (9.14)]

The slow space-charge wave has $f\tilde{v}_s > 0_0$ and hence has a negative ac kinetic-power density, whereas the fast space-charge wave has a positive ac kinetic-power density. Since the slow wave grows, it therefore loses energy, and the conservation theorem (9.49) then requires that the electromagnetic power increase. The ac current of the slow wave will have a phase angle relative to E , such that $\text{Re}(E_{..}, J^*)$ is negative and the current continually gives up energy to the field. This may be verified by substituting $/3_3$ for $/3$ in (9.14) to obtain

$$-f) \quad f\epsilon\mathbf{E}, /$$

As a further aid to the understanding of the traveling-wave tube, it may be noted that it can be viewed as a large number of closely spaced cavity gaps operating as a multicavity klystron. The adjacent turns of the helix are then considered as constituting a gap.

The main advantage of the traveling-wave tube over the klystron is its relatively broad frequency band of operation. Typical units provide gains of 30 to 50 dB over an octave or more in frequency. Power-handling capability ranges from milliwatts to megawatts.

M-TYPE TRAVELING-WAVE TUBE

The magnetron-type (M-type) traveling-wave tube is a linear version of the cylindrical magnetron. Figure 9.21 is a schematic illustration of an M-type tube using a corrugated, or comblike, slow-wave circuit. The electron beam is much wider than it is thick and approximates a sheet beam. A potential V_a is applied between the sole and the anode block. A large static magnetic field is applied in a direction perpendicular to the beam velocity $v_n a_v$, and the static electric field $-f_{..} a_v$ arises from the anode to sole potential V_a . The electrons moving upward from the cathode at potential V_e are deflected by the magnetic field into a beam moving in the positive z direction. The desired type of flow is the one where there is only a z -directed velocity

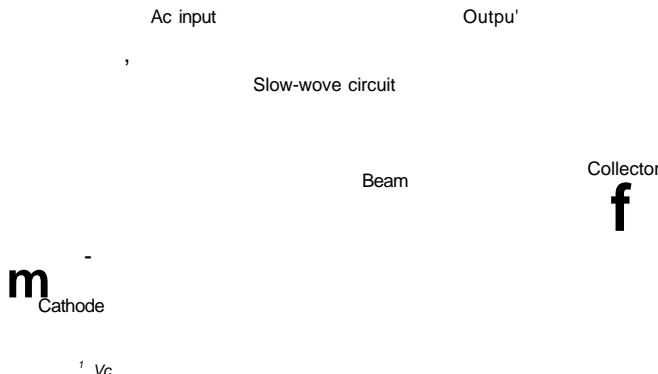


FIGURE 9.21
M-type traveling-wave tube.

$v_a(y)$, which in general is a function of y . Electron flow takes place in a crossed E and B field, which is typical of magnetron-type tubes.

For stable flow, $v_0(y)a^{\wedge}$ does not vary with y . If we denote by $V(y)$ the potential at an arbitrary value of y between the sole and anode block, we must have a balance between the magnetic force

$$-ev_0(y)a_x \times a_z B_0 = -eB_0 u_0(y)a_y$$

and the electric field force $a_y e < V/dy$. Thus

$$\frac{mv}{ay} = v_0(y)B_c \quad (9.110)$$

The velocity $u_0(y)$ may be found from the energy equation

$$imv^2_0(y) = e(V - V_r)$$

The derivative with respect to y gives

(9.11D)

The potential $V(y)$ arises from the applied potential V_{∞} and from space charge within the beam. Under equilibrium conditions of this $-e(E + v \times B_0)$ acting on an electron is zero. The divergence equation thus gives

$$V \cdot E = V \times B_0 = -\frac{P_0 s}{B_0} + B_0 \frac{dv}{dy} \quad (9.112)$$

since $v = v_0 a$, and $B_0 = B_0 a$. In this equation $-P_0$ & the space charge density and s is a factor giving the fraction of space charge which is not neutralized by positive ions. For no y

- negative negative negative j_{ns}

present, $S = 1$. If we assume that $S = 1$, the set of relations (9.110) to (9.112) can hold only if

$$\frac{2}{Or} = \frac{n S B_2}{T} \approx \frac{I^{m11}}{\longrightarrow} = \frac{2}{Up}$$

or

$$m_e = w; \quad (9.113)$$

as can be determined by eliminating $8V/9y$ and $9v_e/9y$. When this condition holds, the flow is referred to as planar Brillouin flow.

With the above model for the beam, it is possible to solve for space-charge waves that can propagate on the beam. In the presence of a slow-wave structure, the propagation constants become perturbed and a growing wave is produced similar to that in the O-type tube. For a detailed analysis the reader is referred to the citations given at the end of this chapter. The principles involved are not sufficiently different from those already discussed to warrant inclusion in this text.

12 GYROTRONS

Magnetrons and klystrons require resonant cavities to support the electromagnetic field that interacts with the electron beam. The traveling-wave tube requires a slow-wave structure. These structures have dimensions linearly proportional to the operating wavelength and become very small at millimeter wavelengths. The consequence of having to reduce the dimensions as the frequency increases is that the available area for the electron beam decreases and the power output that can be achieved decreases rapidly, approximately proportional to $1/\sim^2$. Thus, at frequencies of 100 GHz and above, conventional microwave tubes are not capable of producing power outputs in the kilowatt range. A relatively new tube, the gyrotron, has been developed in more recent years and does not rely on the use of resonant cavities or slow-wave structures. In the gyrotron the electromagnetic field interacts with the cyclotron motion of the electrons in a strong static magnetic field. When an electron is acted upon by the force of a steady magnetic field, its motion in the plane perpendicular to the magnetic field is a circular-one. By using a sufficiently strong magnetic field, the frequency of rotation, called the cyclotron frequency, can be in the frequency range corresponding to millimeter waves. The interaction of a microwave field having the same frequency as the cyclotron frequency of the electrons results in growing waves. Thus the waveguide through which the electron beam passes and which supports the electromagnetic field is not restricted in diameter by the need to provide either a resonant structure or a slow-wave structure. As a result the fundamental size restrictions of conventional microwave tubes are not present in the gyrotron.

There are three common forms of gyrotron tubes. These are the gyromonotron oscillator; the gyro-TWT, a traveling-wave amplifier tube;

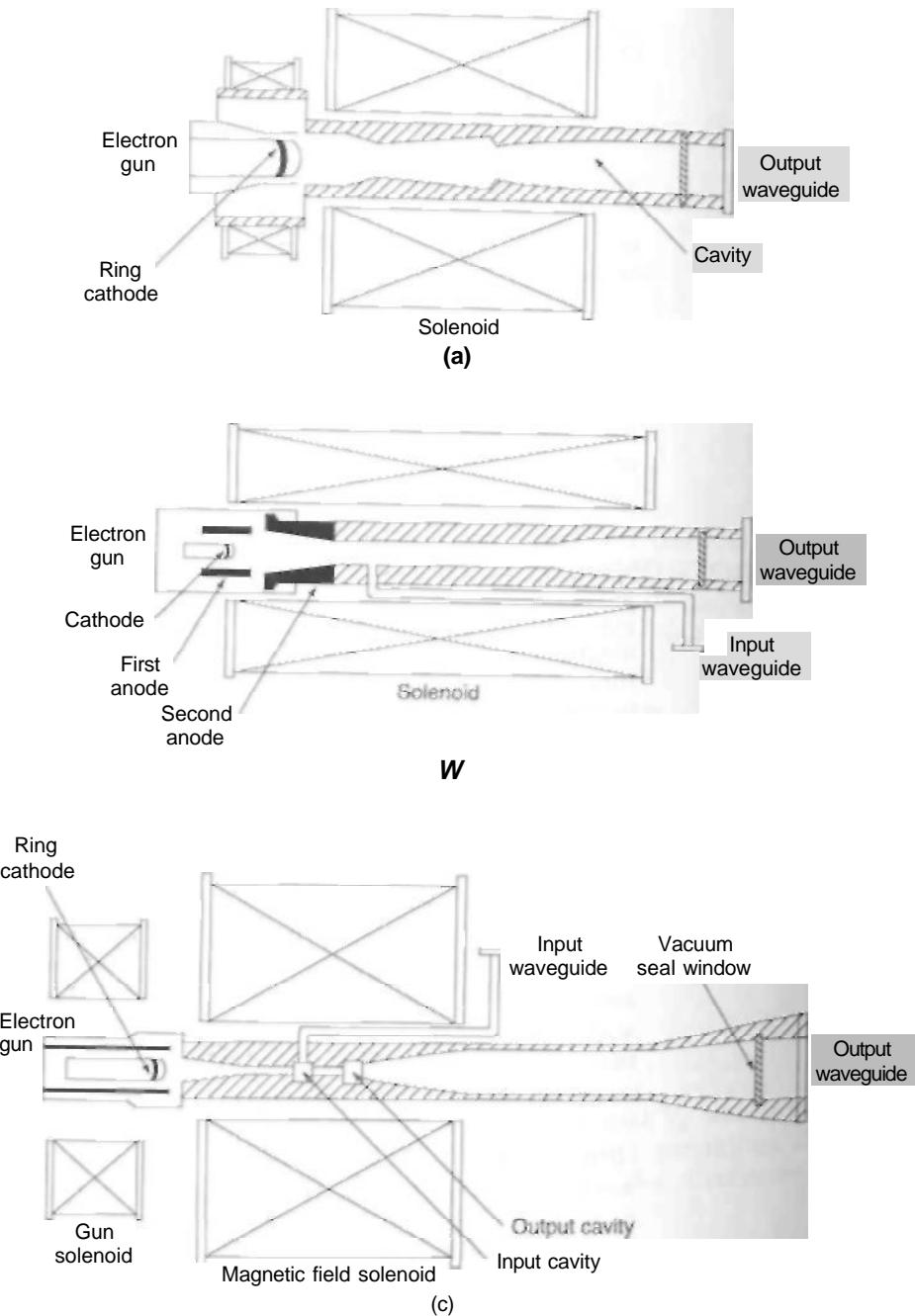


FIGURE 9.22

(a) The gyromonotron oscillator; (b) the gyro-TWT amplifier; (c) the gyrokylystron am Jifier

and the gyrokylystron, another amplifier tube. Simplified drawings of the three gyrotron tubes are shown in Fig. 9.22. Each tube has a magnetron-type electron gun which imparts a high radial velocity to the electrons before they enter the high magnetic field region. The large static magnetic field is provided by either a liquid cooled solenoid or a superconducting solenoid. In the gyromonotron shown in Fig. 9.22a. the interaction region is an enlarged circular waveguide that can support many different propagating modes. The output is taken from an output waveguide through a transparent window which is also used as a vacuum seal for the tube. In the gyro-TWT amplifier the input signal is coupled into the input of the interaction region through a waveguide as shown in Fig. 9.22b). The input microwave signal provides the initial bunching of the electrons in the beam. The input signal is amplified in the circular waveguide whose dimensions are large enough to support many possible propagating modes. The electron beam is in the form of a hollow beam with a radius such that it interacts strongly with only one, or at most only a few, of the circular waveguide modes. The gyrokylystron uses an input and output cavity as shown in Fig. 9.22c. The signal to be amplified is coupled into the input cavity. The output signal is taken from the output cavity.

Gyrotron amplifiers that provide power gains of as much as 24 dB and output powers as high as 50 kW at 5 GHz have been built. Ferguson, Valier, and Symons describe a 5.2-GHz gyrotron tube producing 128 kW of output power.^t This tube uses an 8-A-65 kV electron beam. Pulsed power outputs from gyrotron oscillators have been produced at levels of several hundred megawatts. It has been reported that 28-GHz gyrotrons with 200 kW of continuous-wave output power are in operation at Oak Ridge National Laboratory.[?] The high power capability of gyrotrons has been amply demonstrated; so these tubes will become more important for millimeter-wave systems in the future.

Particle Interaction in a Gyrotron

In a gyrotron electron beams having a very large azimuthal velocity are used, so that the relativistic increase in the mass of the electron must be taken into account. The electron gun injects the beam into the high magnetic field region with the electrons initially having a large radial velocity component. The $v_{\text{d}} \times B_{\text{r}}$ force then causes the electrons to follow a helical path with a velocity $v_{\text{r}} = v_{\text{ih}}a_{\text{r}} + v_{0j}a_{\text{r}}$ with v_{ih} being several times larger than v_{0j} . If we treat the electrons as a charged fluid, then under the action

^tP. E. Ferguson, G. Valier, and R. S. Symons, Gyrotron-TWT Operating Characteristics, *IEEE Trans.*, vol. MTT-29, pp. 794-799. 1981.

[?]J. T. Coleman, "Microwave Devices," Reston Publishing Company, Inc., Reston, Va. 1982.

of a microwave field the velocity of a differential volume element charged fluid will have an ac component v in addition to the dc component v_0 ; so the total velocity field will be $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}$. The effective mass $m^* = m/c^2 r^2$, where r is the relativistic factor. The momentum is my . The particle density is $N + n$, where the ac variation in the number density from the average value n is n particles per unit volume. The charge density is given by $-e(AT + p_0) - p$. If \mathbf{E} and \mathbf{B} represent the microwave field and \mathbf{B}_0 is the static magnetic field, then the equation of motion for a volume element of the electron fluid is [see (9.5'')]

$$\frac{N + n}{m} \frac{d\mathbf{v}}{dt} = (p - p_0)(\mathbf{E} + v_{\perp} \mathbf{XB} + v_z \mathbf{XB}_0) \quad (9.114)$$

When there is no ac field present, the steady-state motion is a drift with a constant velocity v_{0z} along the direction of \mathbf{B}_0 , which we take as the z axis, and rotation about \mathbf{B}_0 at the cyclotron frequency ω_c given by

$$\mathbf{n} = \frac{p \mathbf{B}_0}{N m y} e \mathbf{B}, \quad (9.115)$$

where $y_0 = (1 - v_0^2/c^2)^{1/2}$. By using a very large magnetic field, ω_c can be in the microwave or millimeter-wave range of frequencies. For example, $v_0/c = 0.8$ and $B_0 = 3 \text{ W/m}^2$ (30 kG), we get $\omega_c = 8.8 \times 10^{11} \text{ rad/s}$, which corresponds to a frequency of 140 GHz. A magnetic field as large as 3 requires the use of superconducting solenoids. A gyrotron can operate harmonics of the cyclotron frequency. The advantage of operating at harmonic of the cyclotron frequency is that a smaller static magnetic field is required, but this is accompanied by a lower efficiency for power generation.

The possibility of field interaction with the beam at a harmonic of the cyclotron frequency is readily demonstrated by considering the current associated with a circulating electron. Consider an electron with azimuthal velocity $v_{0\phi}$ and located at $\theta = \phi$, $r = r_0$, at time $t = 0$. The current is in the form of an impulse $J_\phi = -ev^2 S \delta(r - r_0) S(z) \delta(\phi) - \delta(t)$, where the delta functions localize the current element at the position r_0 , $z = 0$, $\phi = d\phi$. We can make a Fourier series expansion of J_ϕ in terms of the angle ϕ :

$$J_\phi = S(r - r_0) S(z) I_\phi e^{i\phi} \quad (9.116)$$

where the I_ϕ are given by

$$\begin{aligned} & \mathbf{L} \cdot \mathbf{i} e^{i\phi} = (-e \omega_0 \#)^j (* - * i)^k \\ & -\mathbf{fit}, -** \\ & 2\nu \end{aligned}$$

[9.H]

Hence we have, at $t = 0$,

$$\mathbf{J}^* = -\nabla \mathbf{H} \cdot \mathbf{r} - \mathbf{r} \cdot \nabla S(z) \quad \text{Eq. } 9.118$$

At a later time the position of the electron will be at $\langle \mathbf{r} \rangle = \langle \mathbf{r}_0 \rangle + \mathbf{v}_0 t$. A Fourier series expansion of the current at time t may also be carried out and gives

$$\mathbf{J}^* = -\nabla \mathbf{H} \cdot \mathbf{r} - \mathbf{r} \cdot \nabla S(z) \quad \text{Eq. } 9.119$$

upon replacing \mathbf{r} , by $\mathbf{r}_0 + \mathbf{v}_0 t$. This expansion shows that the current associated with a single rotating electron is composed of an infinite number of equal-amplitude harmonics of the cyclotron frequency, a result due to the impulsive nature of the current.

Consider now a very large number N of electrons spaced at random around the orbit. Each electron contributes a current given by (9.119) but with $\langle \mathbf{r}_0 \rangle$ replaced by \mathbf{r}_n for the n th electron. The total current is obtained by averaging over all electrons and will involve the average of the following quantity over all phase angles θ :

$$\langle \mathbf{J}^* \rangle = \frac{1}{N} \sum_{n=1}^N \mathbf{J}^*(\mathbf{r}_n) = \frac{1}{N} \sum_{n=1}^N \left(-\nabla \mathbf{H}(\mathbf{r}_n) \cdot \mathbf{r}_n - \mathbf{r}_n \cdot \nabla S(\mathbf{r}_n) \right)$$

For large N the average will be zero except for the $n = 0$ term which gives a factor N . Thus the current becomes

$$\mathbf{J} = \frac{1}{N} \sum_{n=1}^N \mathbf{J}^*(\mathbf{r}_n) = \frac{1}{N} \sum_{n=1}^N \left(-\nabla \mathbf{H}(\mathbf{r}_n) \cdot \mathbf{r}_n - \mathbf{r}_n \cdot \nabla S(\mathbf{r}_n) \right) \quad (9.120)$$

which is a dc current. These results show that, in order to obtain an ac current at the cyclotron frequency or its harmonics, the electron distribution around the orbit must be nonuniform. We require bunching of the electrons around the orbit. If electron bunching occurs, then we will obtain an ac current with which the microwave field can interact. We will show next that the dependence of the electron mass on the velocity provides a mechanism that will cause electron bunching to occur.

If we take a scalar product of the equation of motion with \mathbf{v} , we obtain

$$(N+m) \mathbf{v} \cdot \frac{d}{dt} (\mathbf{v} \cdot \mathbf{v}) = (\mathbf{p} - \mathbf{p}_0) \cdot \mathbf{v} - \mathbf{E} \cdot \mathbf{v} \quad (9.121)$$

We now note that

$$\begin{aligned} \frac{d}{dt} (\mathbf{v} \cdot \mathbf{v}) &= \mathbf{v} \cdot \frac{d}{dt} \mathbf{v} + \mathbf{v} \cdot \frac{d}{dt} \mathbf{v} = 2 \mathbf{v} \cdot \mathbf{a} \\ &= 2 \mathbf{v} \cdot \mathbf{F} \end{aligned}$$

since $d(v, \cdot \backslash)/dt = 2v, \cdot dv/dt$. We also have

$$\frac{dy}{dt} = \frac{d}{dt} \left[\frac{2e}{c^2} \sin i \right] = \frac{2e}{c^2} \frac{dv}{dt} \quad \text{in } \mathbf{A}$$

By using these results we find that

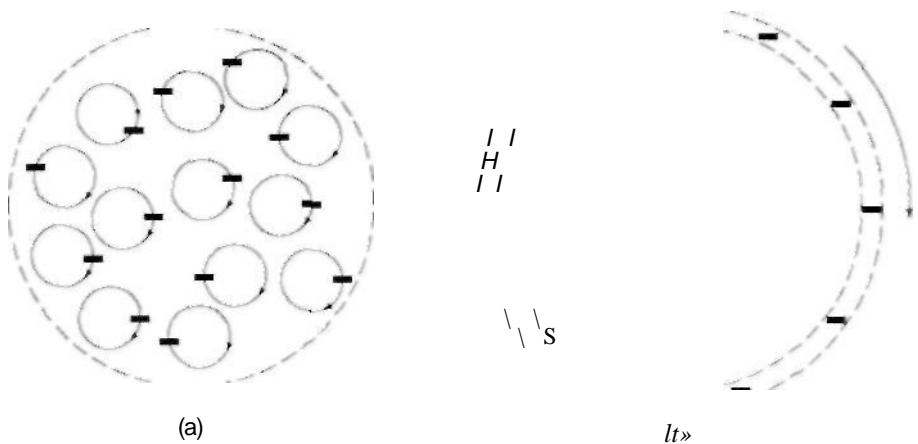
$$\frac{dy}{dt} = \frac{y}{m} \frac{d\mathbf{v}}{dt} - \frac{e}{mc^2} \mathbf{r} \cdot \mathbf{K} + \frac{e^2 \mathbf{d}}{mc^2} = \frac{e^2}{mc^2} \frac{\mathbf{B} \times \mathbf{E}}{dt} \quad \text{so} \quad \frac{dy}{dt} = \frac{e}{m(N+n)} c^2 \frac{\mathbf{B} \times \mathbf{E}}{dt} \quad (9.122)$$

This equation tells us that when $v, \cdot E$ is positive y will decrease and when $v, \cdot E$ is negative y will increase. From (9.115) we see that when y decreases the cyclotron frequency will increase and when y increases the cyclotron frequency will decrease. Consequently, those electrons that have a phase angle greater than $+90^\circ$ relative to the electric field will have their cyclotron frequency reduced, whereas those with phase angles less than $\pm 90^\circ$ relative to that of the electric field will have an increased cyclotron frequency. This process results in bunching of the electrons in the azimuthal direction in a manner similar to the longitudinal bunching that occurs in a klystron. When the electric field adds energy to the electron, its azimuthal velocity increases. Paradoxically, this increases the value of y but reduces the cyclotron frequency. What happens is that the radius of the orbit increases so that even though the azimuthal velocity has increased it takes longer for an electron to execute one circuit around the orbit so the cyclotron frequency is lower.

Gyrotrons generally use hollow cylindrical beams as shown in Fig. 9.23. In Fig. 9.23a we show a conventional beam in which the electrons revolve around individual *magnetic field lines* and do not have a common center of rotation (guiding center). This type of beam is best suited for field interaction at the cyclotron frequency. The beam shown in Fig. 9.23b is used in large-orbit gyrotrons. All of the electrons in this beam have common center of rotation. As a typical example a beam with a cyclotron frequency of 10 GHz and an azimuthal velocity of 0.8c will have an orbital radius equal to $r_M = 2.4 \times 10^{11} / (2\pi \times 10^{10}) = 3.82$ mm. This beam is large enough to provide good field interaction with microwave fields of various harmonics of the cyclotron frequency.

The small-signal analysis of a gyrotron can be based on the linearized equation of motion. If only terms linear in the ac quantities are retained (9.114), the equation of motion, we obtain

$$yoNm \sim v + 4 \cdot -p_Q E - p_0(v_0 \cdot xB + v \cdot XB_0) + / > v_o x B, \quad (9.123)$$

**FIGURE 9.23**

(a) Cross section of a solid electron beam with each electron having its own guiding center. (b) Cross section of a cylindrical sheath beam in which all electrons have a common guiding center.

In addition, the current J is given by

$$Po^v + />_0^v \quad (9.124)$$

and the continuity equation

$$\nabla \cdot J = - \frac{dp}{m} \quad (9.125)$$

must hold. The above *equations can be solved For J and p in terms of the ac fields E and B*. Maxwell's equations must then be solved in the circular (or other) waveguide, both within the beam region and outside, including the source terms. The results will show that growing waves are produced. All ac quantities can be assumed to have the form

•x

where $C_{..}$ is an expansion coefficient for the quantity of interest. Since the equations are linear, the solution can be carried out for the n th term by itself. The electromagnetic field in the circular waveguide can be described in terms of left and right circularly polarized waves. Only those fields that rotate in synchronism with the electrons will interact with the beam ac currents.

For large power applications nonlinear effects must be included. The commonly used approach is to solve for the perturbed orbits of the electrons

and then evaluate the field interaction that take place. It is necessary to use numerical methods in order to solve the nonlinear equations. Typical results obtained by this method are given in the papers cited at the end of this chapter.

9.13 OTHER TYPES OF MICROWAVE TUBES

In addition to the main types of microwave tubes already discussed there are a variety of others as well. In one form of traveling-wave tube the resistance-wall amplifier, the helix is replaced by a circular guide lined with a resistive material. The resistive lining enables a slow wave to propagate in the guide, a wave that is highly attenuated in the absence of a beam. If an electron beam is present, amplification takes place with a growth constant a_B large enough to offset the attenuation due to the resistive lining. Thus a net overall amplification is obtained,

In another form of traveling-wave tube, the double-stream amplifier, two parallel electron beams are used. In this tube one of the beams provides the slow-wave structure, or circuit, for the other beam.

It is also possible to amplify the space-charge waves directly by passing the beam through a succession of accelerating and decelerating regions. This type of tube is called a velocity-jump amplifier because the beam velocity v_{fl} is periodically changed, or jumped, to new values.

For both the O-type and M-type traveling-wave tubes, it is possible to adjust the beam velocity so that it is equal to the phase velocity of any one of the spatial harmonics making up the Bloch wave that can propagate along the periodic structure used for the slow-wave circuit. In particular, interaction between the beam and one of the backward-propagating spatial harmonics is possible. Consider a Bloch wave propagating in the $-z$ direction. For this wave, E , has the expansion

where p is the period of the periodic structure in the z direction. If we want interaction between the beam and the $n = -1$ spatial harmonic, it is necessary to choose

$$\omega = v_p = -\frac{2\pi}{P} - 2\pi r/P \quad 2\pi r/p \sim P$$

If the period p is small enough, the $n = -1$ spatial harmonic has a phase velocity directed in the $+z$ direction and its group velocity is in the $+z$ direction. If the $n = -1$ spatial harmonic is amplified, all the other spatial harmonics are attenuated.

harmonics are also amplified, since they must all be present with very definite amplitudes in order that the boundary conditions may be satisfied. The amplification of the noninteracting spatial harmonics comes about because of the increasing surface current and charge induced on the metallic boundaries by the amplified spatial harmonic that interacts with the beam. Tubes employing interaction with a backward spatial harmonic are usually used as oscillators and are called backward-wave oscillators, or carcinotrons. They have their output coupling at the cathode end.

There are still other forms of microwave tubes, and no doubt more will be developed. For more extensive discussion the cited references at the end of this chapter should be consulted.

- 9.1. Consider an electron beam of radius a , velocity u_{∞} , and space charge density $f_i n$. The dc current density is then $J_Q = p \ll O_0$. Show that a magnetic field

$$\frac{rp_0v_0}{0 < r < a}$$

$$\begin{array}{ll} H_{\infty}, & a^2 p_0 v_0 \\ & \frac{2r}{r > a} \end{array}$$

is produced. Verify that the compression force $-ev_{\infty} X B_{\infty}$ is much smaller than the radial outward force due to the space-charge electric field and may therefore be neglected.

- 9.2. Show that an electron with velocity v perpendicular to B_{∞} executes circular motion at the cyclotron frequency $\omega_{c\infty} = eB_0/m = TJZ_{\infty}$, by equating the centrifugal force to the $-ev X B_u$ force.
- 9.3. An electron beam has a radius of 0.2 cm. The accelerating voltage is 1,000 V. The total beam current is 0.03 A. Calculate the beam perveance, the space charge density p^* and velocity v_{∞} , the number of electrons per cubic meter, and the radial electric field due to space charge. Estimate the radial displacement of an electron located at the beam boundary during the time it takes the beam to move a distance $d = 5$ cm. Use the equation $m d^2 r/dt^2 = -eE_r$, and assume E_r to be constant and equal to its value at the beam boundary. Is the beam dispersion significant in this case if d is kept less than 5 cm?
- 9.4. Consider an electron beam with dc parameters λ_{∞} , $v_0 = v_{\infty} a$, immersed in a field $B_0 = a D_n$. Assume a time dependence $e^{j\omega t}$ and a z dependence $e^{-j\omega z}$ and solve the linearized equation of motion (9.8) for $v_z = v_x a_x + L_v a_y + t' a_z$, to obtain

$$\begin{array}{lllll} j(\omega - pu_0)/\lambda & -\mathbf{a} \cdot \mathbf{v} / \mathbf{A} & 0 & E_x \\ co/A & j(<\omega - pv_0)/\pm & 0 & E_y \\ -n & 0 & 0 & A' \\ & & l/j(a, -liv_{lt}) & \end{array}$$

where $A = co^2 - (co - fivo)^2$

From the continuity equation (9.5e) and (9.116), show that

$$\mathbf{J} = \frac{\mathbf{v}_0 \mathbf{V} \cdot \mathbf{J}}{-j\omega_0} - \mathbf{P}_{\text{out}}$$

$$\mathbf{J} = \frac{\mathbf{u}_0 \mathbf{V} \cdot \mathbf{J}}{-j\omega_0} - \mathbf{P}_{\text{out}}$$

$$\mathbf{J} = -\mathbf{P}_{\text{out}} \mathbf{V}$$

$$\mathbf{V} \cdot \mathbf{J} = -j\omega_0 \mathbf{P}_{\text{out}} \cdot \mathbf{V}$$

$$\therefore \mathbf{J} = \frac{j\omega_0 \mathbf{P}_{\text{out}} \cdot \mathbf{V} - W_{\text{out}} V_t}{\omega_0 - f_i u_n}$$

- 9.5. Using the results of Prob. 9.4, obtain solutions for p for waves in an infinite electron beam when all ac quantities are independent of x and y . Note that for space-charge waves, $E_x = E_y = 0$ but E_z is finite. For the field waves' $E_z = 0$.

Hint: Note that $\mathbf{V} \times \mathbf{J} = -y/\omega_0 \mathbf{a}_z$, $\mathbf{X} \cdot \mathbf{H} = -y/\omega_0 \mathbf{a}_z$, $\mathbf{T} \times \mathbf{E} = -y/\omega_0 \mathbf{a}_z$, $\mathbf{X} \cdot \mathbf{E} = 0$, which leads to the equation $j\omega_0^2 - k^2 E_z = w/i_s J_z = -c_0/\omega_0^2 > 0$.

Answer: For field waves, p is a solution of

$$(\omega - \omega_c)^2 - \mathbf{r} \cdot (\mathbf{0}^* - \mathbf{A}) = \pm \frac{\omega_c^2}{\omega - \omega_c}$$

where \mathbf{A} is given in Prob. 9.4. Note that two solutions are given by $w - f_i v_a = \pm 1$. These are the cyclotron waves. For $p = k_0$, so that $w \gg \omega_0$, four other approximate solutions are

$$P = \pm \omega_0^2 \left(\frac{1}{\omega - \omega_c} \right)$$

- 9.6. Compute the gain of a klystron amplifier of the type considered in the text where the following data apply: Beam radius = 0.3 cm, beam current density = 100 nA/cm². Accelerating voltage = 1,000 V, frequency = 3,000 MHz. $G_L = G_0$, cavity width $d = 0.2$ cm, cavity conductivity = 5.8×10^7 S/m. Compute the gain for $d = 0.3$ cm also, and compare with the earlier calculation.
- 9.7. Consider a reflex klystron employing a cylindrical cavity of the form shown in Fig. 9.12. The data of Prob. 9.6 apply, with $d = 0.2$ cm. The external loading $G_L = G_0$. The cavity grid-reflector spacing s is equal to 1 cm. Calculate an plot the electronic admittance spiral as a function of reflector voltage V_r at a frequency of 3,000 MHz. Plot also the negative cavity admittance $-Y_{\text{on}}$ same susceptance plane. Determine the reflector-voltage variation ΔV_r across the $n = 2$ and $n = 3$ modes. Evaluate the change in oscillation frequency as the modes are tuned across.
- 9.8. Consider a cylindrical waveguide of radius a lined by a resistance sheet so the boundary conditions at $r = a$ are $E_z = -Z_{\text{ave}} H_A$, where $Z_{\text{ave}} = \sqrt{\mu_r \epsilon_r}$ is the surface impedance of the wall. Analyze this structure as a transverse wave.

tube when an electron beam (axially confined flow) with velocity i_{a} , a. completely fills the guide. Determine an appropriate value of Z_m in order to obtain amplification. Find the optimum value of Z_{a} to give a maximum gain.

Answer: f_i is a solution of

$$g = I^a ga \quad (P-1\$)$$

For ga large, so that $i_0/7 = 1, ji = (1 - S)\ell$, with S small. S^a is given by

$$\frac{g^2}{J} = \frac{.2kUR_m/Z_{\text{a}} - (o_p/w)}{lil + j2kl(R_m/Z_{\text{a}})y}$$

where $Z_{\text{a}} = (1 + .)/?_{\text{a}}$.

- 9.9. Consider a cylindrical waveguide of radius a uniformly filled with a stationary plasma (an ionized gas with an equal number of electrons and ions per unit volume). At high frequencies the motion of the ions may be neglected because their mass is much greater than that of the electrons. Thus the guide will have the same properties as one filled with an electron beam with zero axial dc velocity. Use (9.33) to show that the guide may be considered as filled with a dielectric medium with permittivity $\epsilon = e_{\text{a}}(1 - atg/\omega_p)$, where ω_p is the plasma frequency for the plasma. Find a solution for the lowest-order circularly symmetric E mode and show that for $w < w_0$ such that ϵ is negative the wave impedance is inductive.
- 9.10. The results of Prob. 9.9 may be used to analyze the beam-plasma amplifier. Consider an electron beam passing through the plasma-filled guide. Use a confined-flow model to describe the beam and show that for a beam completely filling the guide the equations of Sec. 9.3 are valid provided e_{a} is replaced by $\epsilon = e_{\text{a}}(1 - WQ/IU^2)$ throughout, where ω_p is the plasma frequency for the plasma. In particular, (9.24) to (9.26) hold. Thus in (9.26), if $u_{\text{a}} = (ep_0/m_{\text{e}})^{1/2}$ is replaced by $(ep_{\text{a}}/m_{\text{e}})^{1/2} = (ep_{\text{a}}/m_{\text{e}})^{1/2} - (\omega_p/\omega)^{1/2}h$ it is seen that β becomes complex for $\omega < w_0$ and a growing and decaying pair of waves are obtained. Show that the gain constant is given by

$$\text{Re } 0_0 \quad \text{epe} \setminus \frac{ph}{a-fl} \quad (\mathbf{V-4})$$

and is very large when w is close to ω_0 . Note also that for a finite-radius beam with unconfined flow passing through an unbounded plasma medium the equations of Sec. 9.4 apply with e_0 again replaced by e . For this model of the beam-plasma amplifier (9.45) may be used in place of (9.26) and will predict a gain constant of the same order of magnitude as does the confined-flow model.

- 9.11. Show that when both the plasma electrons and the beam electrons are subjected to the confined-flow condition the only change which occurs in the result given in Prob. 9.11 is the replacement of w_0 by Fw_0 , where $F = (1 + Pom/Po^{1/2})^{1/2}$ is the plasma-frequency reduction factor.

REFERENCES

1. Slater, J. C.: "Microwave Electronics." D. Van Nostrand Company Inc. Printrot
1950.
2. Pierce, J. R.: "Traveling Wave Tubes," D. Van Nostrand Company, Inc., Prin ceton, N.J..
2950.
3. Kleen, W. J.: "Electronics of Microwave Tubes," Academic Press, Inc., New York, lae
4. Beck, A. H. W.: "Space Charge Waves," Pergamon Press, New York, 1958,
5. Mutter, R. G.: "Beam and Wave Electronics in Microwave Tubes," D. Van Niwt
Company, In&S, Princeton, N.J.. 1960. oatrand
6. Spangenberg, K. R.: "Vacuum Tubes," McGraw-Hill Book Company, New York 194R
7. Hamilton, D. R., J. K. Knipp, and J. B. Horner Kuper: "Klystrons and Microwav
Triodes," McGraw-Hill Book Company, New York, 1948.
8. Collins, G. B.: "Microwave Magnetrons," McGraw-Hill Book Company, New York 194R
9. Chodorow, M., and C. Susskind: "Fundamentals of Microwave Electronics," McGraw-Hill
Book Company, New York, 1964.
10. Reich, H. J., P. F. Ordung, H. L. Krauss, and J. K. Skalnik: "Microwave Theory and
Techniques." D. Van Nostrand Company, Inc., Princeton, N.J., 1953.
11. Reich, H. J., J. K. Skalnik, P. F. Ordung, and H. L. Krauss: "Microwave Principles,"
D. Van Nosrland Company, Inc., Princeton. N.J., 1957.

Space-charge wave theory

12. Ramo, S.: The Electronic Wave Theory of Velocity Modulated Tubes, *Proc. IRE*, vol. 27,
p. 757, 1939.
13. Ramo, S.: Space-Charge and Field Waves in an Electron Beam, *Phys. Rev.*, vol. 56,
p. 276, 1939.
14. Hahn, W. C: Small Signal Theory of Velocity Modulated Electron Beams, *Gen. Elec.
Rev.*, vol. 42, p. 258, 1939.
15. Chodorow, M., and L. Zitelli: The Radio Frequency Current Distribution in Brillouin
Flow. *IRE Trans.*, vol. ED-6, p. 352, 1959.
16. Rigrod, W., and J. Lewis: Wave Propagation along a Magnetically Focused Cylindrical
Electron Beam, *Bell System Tech. J.*, vol. 33, p. 399, 1954.
17. Brewer, G. R.: Some Effects of Magnetic Field Strength on Space-Charge Wave Propaga
tion. *Proc. IRE*, vol. 44, p. 896. 1956.

Gyrotrons

18. Sprangle, P. and A. T. Drobot: The Linear and Self-Consistent Nonlinear Theory of the
Electron Cyclotron Maser Instability, *IEEE Trans.*, vol. MTT-25, pp. 528-544, 1977.
19. Li, Q. F., S. Y. Park, and J. L. Hirshfield: Theory of Gyrotron Traveling-Wave Amplifi
ers, *IEEE Trans.*, vol. MTT-34. pp. 1044-1058, 1986.
20. Silverstein, J. D., M. E. Read, K. R. Chu, and A. T. Drobot: Practical ^{Consi^r^o^g}
in the Design of a High-Power 1 mm Gyromonotron, *IEEE Trans.*, vol. MI
pp. 962-966. 1980. Compact
21. Vitello, P., W. Miner, and A. Drobot: Theory and Numerical Modeling of a ^A ^ J ^
Low-Field High-Frequency Gyrotron. *IEEE Trans.*, vol. MTT-32, pp. 373-38

CHAPTER 10

SOLID-STATE AMPLIFIERS

The first solid-state amplifiers for microwave applications were negative-resistance diodes such as the tunnel diode. This was followed by the development of parametric amplifiers that used a variable-capacitance diode (varactor) and an oscillator (pump source) to vary the junction capacitance at the pump frequency. An outstanding feature of parametric amplifiers was the low noise that could be achieved by cooling the diode to liquid-nitrogen temperatures. The theory and design of parametric amplifiers is described in Chap. 11.

Parametric amplifiers became the prominent and most widely used solid-state amplifiers during the period 1958 to about 1970. By 1970, improvements in materials preparation and processing technology had resulted in development of *n-p-n* silicon bipolar transistors with a maximum frequency of oscillation greater than 10 GHz. During the next two decades further progress in the design and manufacture of high-frequency microwave bipolar transistors and field-effect transistors was dramatic. The key to successful microwave transistor design is miniaturization which is a necessity in order to reduce device and package parasitic capacitances and lead inductances and to overcome the finite transit time of the charge carriers. An appreciation for the need to reduce parasitic capacitance and inductance can be obtained by referring to Table 10.1 where representative values of reactances are given at several frequencies. For example, an inductance of 0.1 nH at 10 GHz represents a reactance of 6.28 Ω which is not a negligible value in a 50- Ω system. A capacitance of 0.1 pF at 10 GHz has a reactance of 159 f Ω and would be a significant shunt reactance across a 50- Ω transmission line.

Transit times are dependent on the electron mobility and saturation velocity in the semiconductor material. In this regard gallium arsenide

TABLE 10.1
Reactance as a function of frequency

Frequency K;Hz	1	10	100
Reactance			
/ - 0.1 nH	0.628	6.28	62.8
L = 1 nH	6.28	62.8	628
C = 0.1 pF	1592	159	15.9
C = 1 pF	159	15.9	1.6

(GaAs) is significantly better than silicon for high-frequency devices. By 1980, the design and fabrication of metal-semiconductor field-effect transistors (MESFETs) were well established. In the frequency range above 5 GHz, MF.SFET devices are widely used.

In order to achieve the high-frequency performance in transistors it was necessary to develop the technology that would enable key device dimensions to be less than 1 μ m, e.g., gate lengths with submicron dimensions. By means of molecular beam epitaxy (MBE), it has been possible to grow high-quality epitaxial layers and controlled doping profiles in highly localized regions. MBE techniques also led to the development of heterostructures which, in turn, led to the development of the high-electron-mobility transistor (HEMT) which can operate at frequencies as high as 100 GHz.

Microwave amplifiers are usually constructed either as hybrid microwave integrated circuits (MIC) or as monolithic microwave integrated circuits (MMIC). In hybrid construction the transmission lines and matching networks are usually realized as microstrip circuit elements on a suitable substrate material and then the discrete components such as chip capacitors, resistors, and transistors are connected in place by soldering or using wire-bonding techniques. Discrete devices are available with beam leads for easy insertion into the hybrid circuit.

The word monolithic is derived from the two Greek words *monos* meaning single and *lithos* meaning stone. Thus a monolithic microwave integrated circuit is a circuit where all active devices, e.g., transistors, and passive circuit elements such as transmission lines, capacitors, resistors, and spiral inductors are fabricated on a single semiconductor crystal. The substrate material used has typically been gallium arsenide because of its high resistivity in the undoped state and because of its superiority in high-frequency field-effect device construction. A number of processes such as ion implantation for active devices, metal deposition and evaporation, forming ohmic contacts, electrode pads, and transmission lines, via hole etching and plating, dielectric deposition, etc., is involved in monolithic construction. The overall design and mask making is facilitated by t

of computer-aided design (CAD) programs. Electron-beam lithography and plasma-enhanced etching and deposition techniques are used for fabrication of submicron device elements.

The cost of a monolithic microwave integrated circuit is related directly to how many circuits can be built on a single wafer since the processing of a single wafer is generally a fixed-cost operation. Consequently, in the frequency range below 10 GHz, where distributed circuit elements are relatively large, the hybrid form of construction is often less costly than monolithic construction. However, in the frequency range of 0.1 to 10 GHz, the ability to produce miniature inductors and capacitors has led to the development and production of many MMIC systems using lumped circuit elements instead of distributed circuit elements. In the millimeter-wavelength band monolithic microwave integrated circuit construction promises to be more cost effective and to yield circuits with greater reliability and uniformity.

In this chapter we are primarily concerned with the design of microwave amplifiers from the engineer's point of view. That is, starting from the measured or manufacturer's given two-port parameters of the device, we want to design an amplifier that meets a set of given system requirements such as gain, noise figure, bandwidth, and input and output VSWR. For this reason we only give a short discussion of the main characteristics of bipolar and field-effect transistors. Bias requirements and some typical bias circuits are also described.

The design methodology⁷ that we will develop is based on the use of the scattering-matrix parameters for the device. At high frequencies a transistor will have some intrinsic feedback from the output to the input, usually caused by a finite capacitance from collector to base and emitter lead inductance or in an FET the capacitance from drain to gate. Thus the device may be potentially unstable, and unless the amplifier is properly designed, it may oscillate, in which case it would not be useful as an amplifier. Thus, after the basic equations for amplifier gain have been derived, we examine the conditions under which the device is unconditionally stable or potentially unstable. If it is potentially unstable we will find that only for a certain range of values for the source and load impedances will the amplifier be stable. The available source and load impedances are easily displayed by constructing the source and load stability circles on a Smith chart. The equations for these stability circles will be derived and their interpretation and use will be examined.

The Smith chart is an indispensable aid in the visualization of the different constraints that the engineer must take into account in the design of a microwave amplifier. In addition to the input and output port stability circles already mentioned, there are a number of other useful circles that aid the design process and can be plotted on the Smith chart. The most important of these are circles of constant gain, circles of constant noise

figure, and circles of constant input and output mismatch. For these other circles will be derived and their interpretation in amplifier design will be discussed.

The chapter will conclude with the description of a design strategy for single- and two-stage amplifiers that are subjected to various system requirements. A semi interactive computer program that implements the design strategy will also be described. This computer program reduces the drudgery of carrying out all the computations that are required and gives the user valuable experience in a design process where a number of constraints are imposed and tradeoffs among conflicting requirements must be made.

Most of the relationships involved in amplifier design are of the form of a bilinear transformation from one complex variable to another complex variable. If Z and W are two complex variables, then an equation of the form

$$W = \frac{AZ + B}{CZ + D}$$

where A , B , C , and D are complex constants, is a bilinear transformation of Z into W . This transformation has the property that circles in the Z plane will map into circles in the W plane, with straight lines as limiting forms of circles, with infinite radii, and some points as circles with zero radius. The relationship between impedance and reflection coefficient, namely,

$$_ - 1 + T$$

is a bilinear transformation. For this transformation the straight $R = \text{constant}$ and $X = \text{constant}$ in the Z plane map into circles in the W plane. These mapped circles make up the Smith chart. Since the bilinear transformation occurs over and over again in amplifier design, we discuss its circle-mapping properties before we take up the design theory. This will provide results that enable us to readily identify circle mappings, in particular, the center and radii of mapped circles, from the particular bilinear transformation involved.

10.1 BIPOLAR TRANSISTORS

The basic principle of operation of a bipolar transistor designed for microwave applications is the same as that of low-frequency transistors. The device must be biased to set the operating point. In a common-emitter amplifier circuit, an input network must be designed so that a signal can be applied to the base. A suitable load impedance must be connected across the collector and the output signal is developed across this $\text{m}^{\text{i}}\text{P}^{\text{n}}$. The main difference in the analysis of the operation of a transistor amp

microwave frequencies relative to low-frequency operation is due to the fact that the intrinsic device, along with package parasitic capacitances, resistances, and lead inductances, requires a much more complex equivalent-circuit model. In addition, there is often sufficient capacitive feedback from the collector to the base so that the device is potentially unstable and will be prone to oscillate unless the input and output circuits are designed to prevent oscillations from occurring.

In the frequency range below 5 GHz, silicon bipolar transistors are generally preferred over GaAs FETs except for very low noise amplifier designs. Silicon bipolar technology is more mature and manufacturing costs are less. The gain obtained from a bipolar transistor is inherently greater than that from a field-effect transistor because of a much higher transconductance g_m . Bipolar transistors are suitable for oscillator and power amplifier applications in addition to small-signal amplifiers. Power gains of 15 to 20 dB can be obtained at 2 GHz with noise figures of around 2 dB. At 10 GHz the power gain for many presently available bipolar transistors is around 5 dB and the GaAs FET is then a better alternative. It is expected that, by perfecting the technology for making the critical dimensions of the emitter structure smaller, silicon bipolar transistors with a maximum frequency of oscillation approaching 100 GHz can be achieved. In recent years the heterojunction technology, originally applied to the construction of field-effect transistors, has also been used to improve the high-frequency performance of bipolar transistors. These transistors are called heterojunction bipolar transistors (HBTs) and exhibit very low base resistance, high current gain, and a speed increase by a factor of 2 to 3. An AlGaAs/GaAs HBT with a cutoff frequency of 105 GHz and maximum frequency of oscillation of 175 GHz has been reported.¹ Thus the future application of bipolar transistors can be expected to extend well into the millimeter-wave-wavelength region.

The basic construction used in a microwave bipolar transistor involves a multifinger interdigitated emitter-base construction. A simplified drawing of the cross section and top view of a bipolar transistor is shown in Fig. 10.1. The use of this particular design is for the purpose of overcoming transit-time limitations and yet maintain a sufficient emitter area. An equivalent-circuit model of the intrinsic bipolar transistor and the additional parasitic elements added by the package is shown in Fig. 10.2. A drawing of a packaged microwave transistor with beam leads is shown in Fig. 10.3.

¹N. H. Sheng, *et al.*, High Power GaAlAs/GaAs HBT's for Microwave Applications, 1987 IEEE Int. Electron Devices Meeting Digest, pp. 619-622, 1987.

C. 11. Liechti, High Speed Transistors: Directions for the 1990's Microwave pp. 165-177, September, 1989.

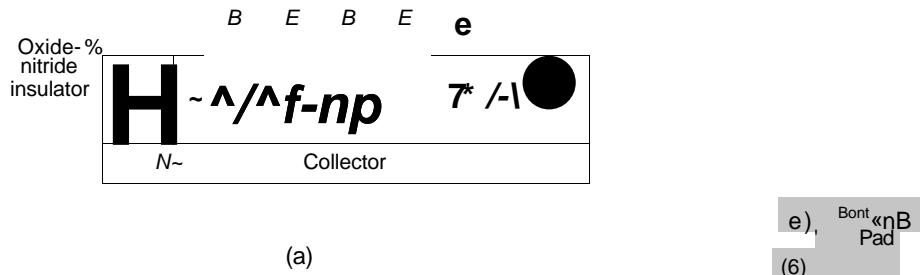


FIGURE 10.1

(a) *Cross section of a microwave silicon bipolar transistor; (b) top view showing interdigitated emitter-base construction.*

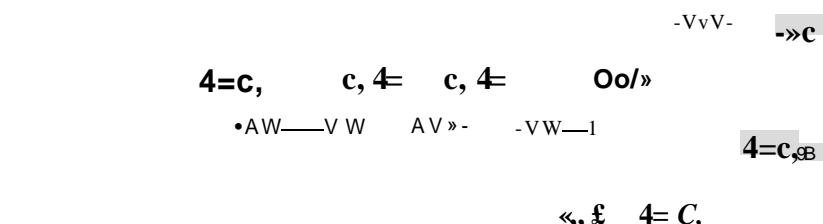
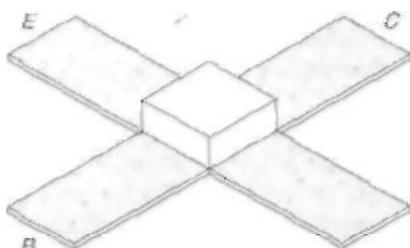


FIGURE 10.2

Equivalent-circuit model of a silicon bipolar transistor.

FIGURE 10.3
A packaged transistor with beam leads.

The equivalent-circuit model shown in Fig. 10.2 is based on that given by Vendelin, Pavio, and Rohde.¹ In this circuit model the various circuit elements are identified as follows:

- C_{bp} —base bond pad capacitance
- C_{ep} —emitter bond pad capacitance
- $R_{b\mu}$ —base contact resistance
- R_m —emitter contact resistance
- $/?, R_2, R_3$ —base distributed resistance
- C_v, C_2, C_{bj} —collector-base distributed capacitance
- R_e —dynamic emitter-base diode resistance
- C_e —emitter-base diode junction capacitance
- R_c —collector resistance
- L_b, L_t —base and emitter bond wire inductances

For an Avantek AT-60500 silicon bipolar transistor operated with a collector current of 2 mA and a collector-emitter voltage of 8 V, typical values for these parameters are:

$$C_{bp} + C_3 = 0.055 \text{ pF} \quad C_v = 0.01 \text{ pF} \quad C_2 = 0.039 \text{ pF}$$

$$C^A \ll 0.026 \text{ pF} \quad C_e = 0.75 \text{ pF}$$

$$R^A + Ra = 4.2(2) \quad R_{bb} = 0.66 \quad /?, = 7.5 \text{ } 12$$

$$R_2 = 10.3 \text{ } 12 \quad R_e = 12.9 \text{ n} \quad R_r = 5 \text{ } 12$$

$$L_b = 0.5 \text{ nH} \quad L_t = 0.2 \text{ nH}$$

This transistor has a base cutoff frequency of 22.7 GHz. The common base current gain is given by

$$1 + jf/f_h$$

where f_h is the base cutoff frequency and τ_d is the collector depletion region delay time (6.9 ps).

The only device parameters that can be easily measured at microwave frequencies are the scattering-matrix parameters S_{11} , S_{12} , S_{21} , and S_{22} . These can be measured by embedding the transistor in a 50-12 microstrip transmission line and using a network analyzer for the measurements. It is also possible to measure the S_{11} parameters with on-wafer probes. By measuring the S_{11} parameters over a range of frequencies, the equivalent-

¹G. D. Vendelin, A. M. Pavio, and U. L. Rohde, "Microwave Circuit Design Using Linear and Nonlinear Techniques," chap. 3, John Wiley and Sons, Inc., New York, 1990.

circuit parameters can be adjusted, by using a computer program they produce a circuit model with calculated scattering parameter that correlate with the measured ones. The equivalent-circuit model provides better physical understanding of the various circuit parameters that affect the operation of the transistor. However, for amplifier design it is easier to use the scattering-matrix parameters instead of the equivalent-circuit model.

Transistor Biasing

There are two main considerations involved in the design of a bias circuit: (1) The biasing circuit must provide a stable operating point that is insensitive to variations in the device parameters and temperature changes and (2) the biasing circuit must be isolated from the high-frequency circuit so that high-frequency signal currents do not flow in the dc biasing circuit. The first requirement is met by incorporating dc feedback in the biasing circuit. The second requirement is met by inserting high-impedance high-frequency circuit elements in series with the dc components and by using low-impedance capacitive bypass circuits to shunt high-frequency currents around the dc circuit elements. The overall bias circuit and RF matching circuits must provide stable terminations for each active device outside the frequency band of interest in order to ensure that oscillations do not occur at any frequency.

The bias circuit shown in Fig. 10.4a provides a stable operating point. It is commonly used in low-frequency electronic circuits and can also be used in microwave amplifier circuits but with more difficulty because of parasitic inductance associated with the capacitor leads. The bias circuit is shown isolated from the transistor by incorporating series inductors (RF chokes) between the device terminals and the bias-circuit resistors. The

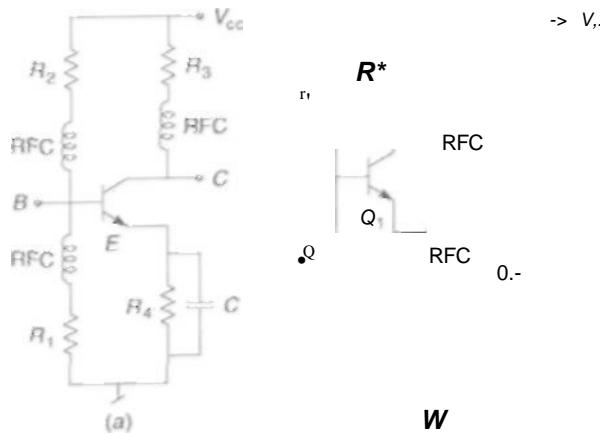


FIGURE 10.4
(a) A passive bias circuit,
an active bias circuit.

emitter terminal is maintained at RF ground by means of the bypass capacitor C.

An active bias circuit is shown in Fig. 10.46. In this circuit the collector current in transistor Q_1 is established by means of the resistors R_y , r_{T2} , and i_{T3} . The base current in the microwave transistor Q_u is the collector current of Q_1 . Since the bias circuit for Q_1 is a stable one, the collector current of Q_u and hence the base current of Q_2 , is maintained at a value that is essentially independent of the transistor parameters. This circuit has the advantages that it consumes less dc power and requires only two RF chokes for isolation as compared with the circuit shown in Fig. 10.4a which is shown with three RF chokes and a bypass capacitor for isolation purposes.

In a microwave amplifier the RF chokes are often replaced by a quarter-wave high-impedance transmission line or a combination of transmission-line sections.

FIELD-EFFECT TRANSISTORS

There are two main characteristics of field-effect transistors that make them superior to bipolar transistors in microwave amplifiers. These are the lower noise characteristics and the higher frequency of operation. The higher operating frequency is due to the higher electron mobility in gallium arsenide, which is the material used in field-effect transistors, compared to that in silicon, the standard bipolar transistor material. The higher electron mobility and the absence of shot noise are important features that result in low noise. The first, microwave field-effect transistors were metal-semiconductor field-effect transistors (MESFETs). High-frequency operation required a very short gate length, typically less than 1 μm . Thus it was only after the processing technology had advanced to the stage that submicron device features could be reliably made that microwave solid-state device development advanced rapidly. Currently produced MESFETs have gate lengths of the order 0.3 to 0.5 μm . The frequency at which the short-circuit current gain becomes equal to unity is given approximately by the relationship $f_T = v_s/2\pi rL_g$, where v_s is the electron saturation velocity and L_g is the gate length. In gallium arsenide (GaAs) the maximum drift velocity is about $2 \times 10^7 \text{ cm/s}$, so that for a gate length of 0.5 μm , $f_T = 60 \text{ GHz}$. Clearly, very short gate lengths are essential for high-frequency operation.

Beginning around 1980, a new technology involving heterojunctions began to find applications in device construction. A heterojunction is a junction formed at the interface of say an aluminum-gallium-arsenide (AlGaAs) doped alloy and an undoped GaAs layer. The use of a heterojunction enables a channel with a very high electron mobility to be obtained. The field-effect transistor which is made using a heterojunction is called a high-electron-mobility transistor (HEMT). Since high mobility is achieved by doping only the large bandgap material, the name modulation-doped

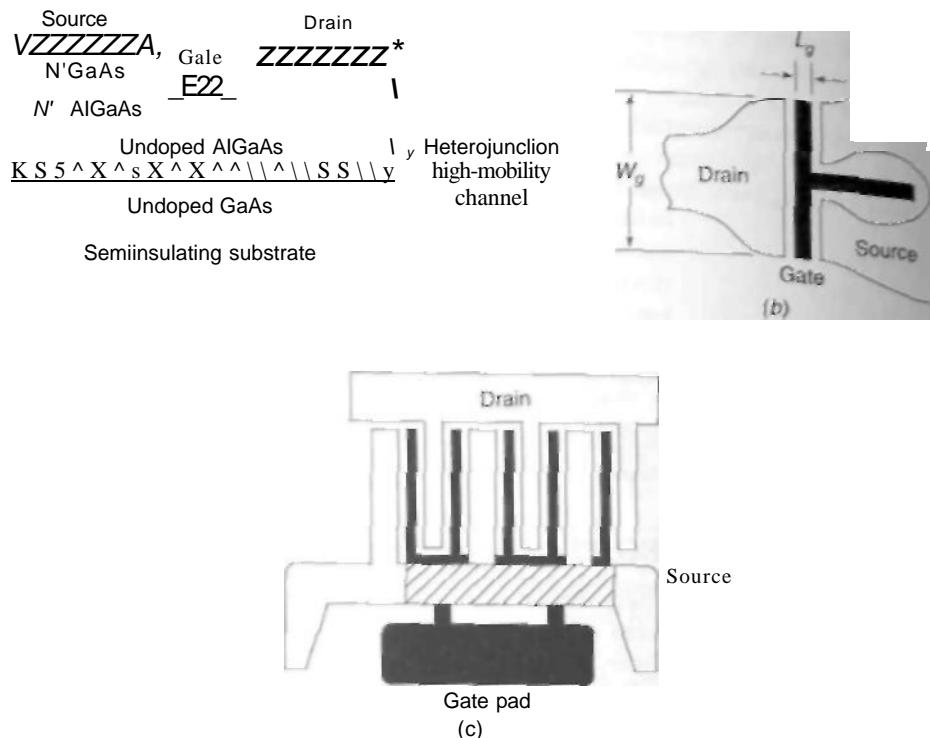
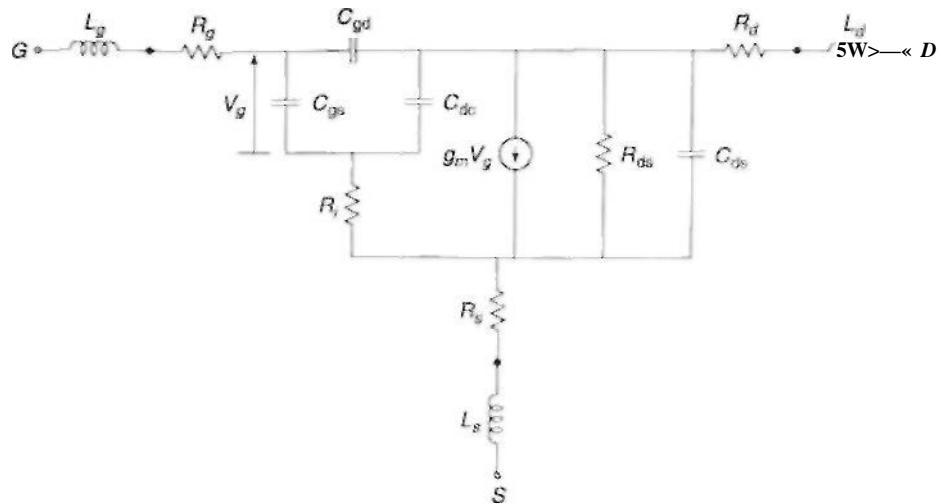


FIGURE 10.5

(a) Cross section of a high-electron-mobility transistor (HEMT); (b) source, drain, and gate structure for a low-power FET; (c) interdigitized construction used for a power FET. The source metalization passes above the gate structure and is insulated from it.

field-effect transistor (MODFET) is also used. The HEMT device has a higher frequency of operation and a lower noise figure than the standard MESFET device. A good discussion of the technology and fabrication of HEMT devices and the theory of operation can be found in the books by Chang and Pengelly.

Figure 10.5a shows the cross section of a typical HEMT device, while Fig. 10.5b shows a top view of the source, drain, and gate structure. The gate width W_g is much greater than the gate length L_g . In a power FET many as 10 source and drain fingers adjacent to gate fingers are used to obtain large drain currents and hence large output powers. The interc

**FIGURE 10.6**

Simplified small-signal equivalent-circuit model of a microwave GaAs FET.

tated construction is shown in Fig. 10.5c. The transconductance g_m of the FET is increased by using a large gate effective width and this is needed for large power output. However, with a larger gate width the input capacitance is increased.

A simplified small-signal equivalent-circuit model of a microwave GaAs FET is shown in Fig. 10.6. Typical values for the circuit parameters for a device with a 1-*i.m* gate length and a 300-*fj.m* gate width are listed below:

$$\text{Gate-to-source capacitance } C_{ss} = 0.4 \text{ pF}$$

$$\text{Gate-to-drain capacitance } C_{Rd} = 0.01 \text{ pF}$$

$$\text{Channel resistance } R_{ds} = 500 \text{ ohms}$$

$$\text{Transconductance } g_m = 30 \text{ mS}$$

$$R_s = 3il$$

$$C_{dl} = 0.015 \text{ pF}$$

The circuit parameters R_g , R_d , R_s , and C^{\wedge} are extrinsic elements. The gate, drain, and source ohmic contact resistances R_g , R_{db} , and R_s are typically a few ohms. The drain-to-substrate capacitance $C_{ds} = 0.07 \text{ pF}$. The inductors L_g , L_d , and L_s have inductances in the range 0.05 to 0.3 nH. The equivalent current source is $V_g g_m$, where V_g is the signal voltage across C_{ss} . The microwave FET can also be described in terms of measured scattering-matrix parameters.

Gallium-arsenide MESFET devices can give a single-stage gain of 8 to 15 dB at 2 GHz with noise figures below 1 dB. For HEMT devices a

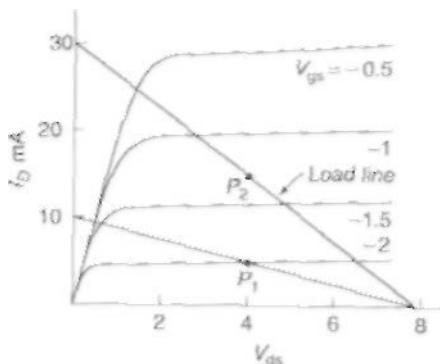


FIGURE 10.7

Output characteristics for a GaAs MESFET showing suitable operating points \$P_1\$ and \$P_2\$ for small- and large-signal inputs.

single-stage gain of 15 dB at 8 GHz and 6 dB at 50 GHz can be achieved with corresponding noise figures of 0.4 and 1.8 dB, respectively.¹ For power applications output powers of several watts from a single device can be obtained. Several devices may be operated in parallel using power-combining techniques to achieve higher output powers.

FET Biasing

The output drain current versus drain-to-source voltage \$V_{ds}\$ is shown in Fig. 10.7 as a function of gate-to-source voltage for a typical microwave GaAs MESFET. For small-signal application an operating point in the vicinity of the point \$P_1\$ would be suitable. For maximum dynamic range the operating point should be placed in the central region of the output characteristics which is depicted by point \$P_2\$. In either case the dc voltage of the gate must be negative with respect to that of the source for a depletion mode device.

This bias condition can be achieved by grounding the gate through an RF choke or a high-impedance quarter-wavelength transmission line and obtaining the desired bias voltage from the voltage drop across the source resistance \$R_s\$ as shown in Fig. 10.8. The required value of \$R_s\$ is given by \$R_s = -V^{\wedge}/I^{\wedge}\$. The source resistance \$R_s\$ should be bypassed to ground for RF signals by means of a capacitor \$C_s\$. In small-signal low-noise applications, the best noise figure is obtained for a dc drain current equal to about 20 percent of the drain saturation current at zero gate-to-source voltage. At low values of drain current, the transconductance is reduced, so that gain must be sacrificed to achieve a low noise figure.

¹K. Chang, *loc. cit.*, p. 465.

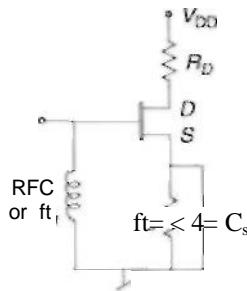


FIGURE 10.8
FET bias circuit-

10.3 CIRCLE-MAPPING PROPERTIES OF BILINEAR TRANSFORMATIONS

A circle with center at x_0, y_0 and with radius R is described by $(x - x_0)^2 + (y - y_0)^2 - R^2 = 0$

$$x^2 + y^2 - 2x_0x - 2y_0y + (x^2 + y^2 - R^2) = 0$$

Now let $Z = x + jy$, $Z_0 = x_0 + jy_0$. The circle equation can be written as $|Z - Z_0|^2 - R^2 = 0$ or $(Z - Z_0)(Z^* - Z_0^*) - R^2 = 0$ which gives

$$ZZ^* - Z\bar{Z}_0 - \bar{Z}Z_0 + (Z_0Z_0^* - R^2) = 0 \quad ;(10.1)$$

Consider now the bilinear transformation from the complex Z plane to the complex W plane given by

$$W = \frac{AZ + B}{CZ + D}$$

This transformation will map circles in the Z plane into circles in the W plane (straight lines are limiting cases). Consider the circle $|W| = p$ or $WW^* - p^2 = 0$. Using the transformation, we get

$$\frac{AZ + B}{CZ + D} \cdot \frac{A^*Z^* + B^*}{C^*Z^* + D^*} - p^2 = 0$$

By expanding this equation we get

$$\begin{aligned} & ZZ^*(AA^* - p^2CC^*) - Z(p^2CD^* - AB^*) - Z^*(p^2C^*D - A^*B) \\ & + BB^* - p^2DD^* = 0 \end{aligned} \quad (10.2)$$

By comparison with (10.1) we see that this is a circle with center at (coefficient of $-Z^*$)

$$Z_c = \frac{p^2C^*D - A^*B}{AA^* - p^2CC^*} \quad ; \quad |A|^2 - p^2|C|^2 \quad (10.3)$$

The constant term equals $|Z_0|^2 - R^2$, so that we can identify the as given by

$$R = Z_0 Z Q = \frac{|\mathbf{B}|^2 - p^2 \mathbf{W}}{IA_1^2 - p^2 |\mathbf{C}|^2}$$

from which we get

$$R = \frac{|AD - BC|}{|(A - p^2 |\mathbf{C}|^2)|} <^{10-4}$$

If the circle in the W plane is $|W - W_a| = p$, then we note that

$$\begin{array}{lll} W - W_n & AZ + B & (A - CW_0)Z + (B - DW_0) \\ & CZ - D & CZ + D \end{array}$$

If we define $A' = A - CW_0$, $B' = B - DW_0$, then the above formulas (10.3) and (10.4) apply with A, B replaced by A', B' .

Many of the relationships that occur in amplifier design involve bilinear transformations and their circle-mapping properties can easily be identified by comparison with the above equations.

10.4 MICROWAVE AMPLIFIER DESIGN USING S_u PARAMETERS

At microwave frequencies impedance and admittance parameters of a transistor cannot be directly measured. The scattering-matrix parameters can be measured and therefore a design methodology based on using the S_u parameters is widely used. The S parameters are measured by inserting the transistor into a test circuit with 50- Ω input and output lines, applying appropriate bias voltages and currents, and measuring the S_{ij} parameters. In any design using S_{ij} parameters, it should be kept in mind that these vary with bias conditions, temperature, and from transistor to transistor even if it is the same device number. Thus the design should leave some margin for S_{ij} variations.

The following are the usual microwave amplifier design goals:

1. Maximum power gain.
2. Minimum noise figure for the first stage. This requires a specific $s^{\circ u r \wedge}$ impedance Z_s for the input stage. The optimum Z_s giving the low noise figure is generally given by the manufacturer of the transistor.
3. Stable gain, i.e., no oscillations.
4. Input and output VSWR as close to unity as possible.
5. Adequate gain and uniformity of gain over a specified frequency "

6. Phase response that is a linear function of ω (no distortion, orly group delay).
7. Insensitivity to nominal changes or variations in the device S_{ij} , parameters.

These objectives cannot all be realized at the same time; so the design procedure must trade off one objective against another one, e.g., gain must be sacrificed for stability. Input VSWR must be sacrificed for a low noise figure since the normalized source impedance Z_s is fixed for minimum noise.

Designing a microwave amplifier using potentially unstable transistors is like designing a bridge using interconnected beams whose lengths can be telescoped and with joints that are free to rotate with insufficient external constraints to define a rigid structure. It is a "loose-jointed" problem without a unique solution. The many design specifications are all interrelated which makes the problem almost unmanageable without some computer optimization strategy. In the next several sections we will examine in detail the constraints that are imposed by stability requirements, by the need for large power gain, low noise, and low input and output VSWRs. We can obtain a good physical understanding of these often conflicting requirements from a study of the stability circles, the constant power-gain circles, the constant impedance-mismatch circles, and the constant noise-figure circles plotted on a Smith chart. From the insight obtained from such a study, we will be able to formulate a design strategy using some computer optimization that will lead to satisfactory designs. The reader will need to bear with us as we work our way through the maze of details but will in the end be rewarded with the satisfaction of having obtained the necessary insight and understanding to be able to implement the theory for practical amplifier design.

All the various low-frequency amplifier circuits such as balanced push-pull amplifiers, cascode amplifiers, and traveling-wave amplifiers can also be used at microwave frequencies. A device that is potentially unstable can be stabilized by resistive loading at the input, output, or both input and output, with a resultant reduction in power gain and increase in noise figure. Stabilization can also be achieved by using negative feedback. Space limitations do not allow a detailed discussion of the variety of amplifier circuits and configurations that are used in practice. Many of these circuits are described in the book by Pengelly and the one by Vendelin, Pavio, and Rohde already cited. The reader is referred to these texts for a detailed discussion. In this text we will limit our attention to amplifier design based on the use of scattering-matrix parameters and linear two-port design.

For high-power amplifier design, it is necessary to consider the nonlinear characteristics of transistors and to pay more attention to device power

dissipation and the design of adequate heat sinks. Small-signal »
design can often provide a first approximation to the design of lar? • r
amplifiers, particularly in connection with stability. Thus the linear t ^{SI***ra}
design methodology developed in this chapter will also provide useful' ^
into the large-signal design problem even though we do not cover the 1 t
topic.

10.5 AMPLIFIER POWER GAIN

We will begin our discussion by deriving expressions for the power gain of an amplifier. There are several definitions used for the gain of an amplifier and they are given below.

$$\text{Power gain } G_p = \frac{\text{power delivered to load}}{\text{input power to amplifier}} \quad (10.5a)$$

$$\text{Transducer gain } G = \frac{\text{power delivered to load}}{\text{available input power from source}}$$

$$\text{Available power gain } G_a = \frac{\text{available load power}}{\text{available input power from source}} \quad (10.5c)$$

If the device is unconditionally stable, then conjugate impedance matching can be used at both the input and output. If this is done then $G_p = G$ $G_a = G_{max}$ - maximum gain. For a device that is only conditionally stable, conjugate impedance matching at both the input and output cannot be used when the stability parameter $K < 1$. The power gain achieved in this case is G . Power gain is the most useful definition in practice since it applies to any actual amplifier independent of whether conjugate impedance matching is used.

For the two-stage amplifier shown in Fig. 10.9, the incident power u given by

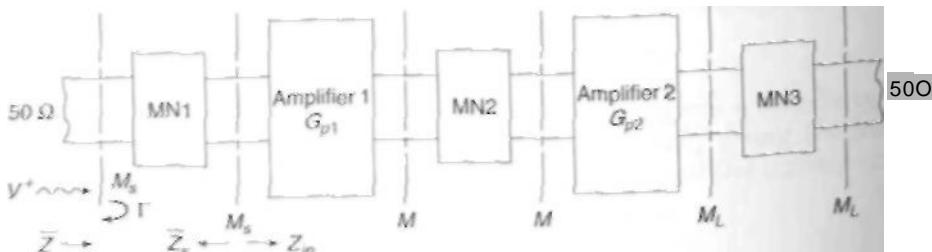


FIGURE 10.9
A two-stage amplifier with matching networks.

and the input power is

$$P_{in} = (i - i_n V_{ta})$$

where Y_c is the characteristic admittance of the input line and Y is the input reflection coefficient. The power P_L delivered to the output 50-J1 line is given by $G_p/G_2 P_{in}$. Hence the two-stage power gain is

$$G_p = \frac{P_L}{P_{in}} = G_p/G_2 \quad (10.6)$$

and the corresponding transducer gain is

$$G_t = \frac{P_L}{P_{in}} = (1 - |1f|) G_{in} G_{out} \quad (10.7)$$

For lossless matching networks the impedance mismatch is the same on the input side as on the output side as shown in Sec. 5.7. If M_x is the impedance mismatch between the first amplifier input and its source impedance Z_s , as seen looking into the output side of the first matching network, then

$$M_x = \frac{4fL_5 fL_{12}}{iz_s + z_j^2}$$

where $Z_{in} = R_m + jX_{in}$ is the amplifier input impedance. On the input line

$$\frac{4Z_{in}}{(Z_C + Z)^2} \frac{4R}{|1 + Z|^2}$$

where $Z = (1 - D)/(1 + D)$. When we use the relations

$$\frac{4R}{(Z_C + Z)^2} \sim \frac{1 + T}{1 - r} \frac{1 + V^*}{1 - r}$$

and

$$\frac{4R}{|1 + Z|^2} \sim \frac{1 + r}{1 - r}$$

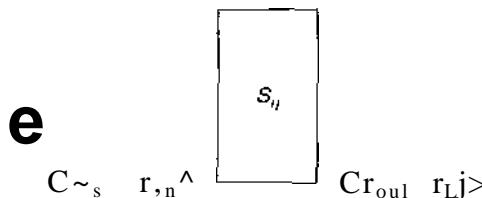
we find that

$$M_x = i - rr^* = I - n^2 = M_x \quad (10.8)$$

upon substituting for R and $1 + Z$ in terms of 1^* and simplifying the resultant expression. The input voltage standing-wave ratio is given by

$$VSWR_i = \frac{i + i_n}{1 - HI} = \frac{i + i/i - M}{1 - \sqrt{1 - M}} = \frac{11 i / i - j^2 t}{1 - j^2 t / M_s} \quad (10.9)$$

Thus the degree of mismatch between Z_s and the input to stage 1 determines the input VSWR. If we have a constraint on $VSWR_i$ in our design and we want to use an optimum source impedance Z_s for minimum noise, then we must terminate stage 1 in a load that will produce an impedance mismatch M_s that will keep the input VSWR, at the specified value (oi-

FIGURE 10.10
A basic amplifier circuit.

lower). The required load termination must, of course, not lead to unstable (oscillating) amplifier and must also yield good power gain. A power gain of 10 or more is desirable, since this will make the effect of following stages on the noise figure small. Sometimes the input VSWR constraint has to be relaxed because it is in conflict with other requirements.

•Derivation of Expressions for Gain

Consider the basic amplifier circuit shown in Fig. 10.10. The source and load are viewed as connected to the amplifier by means of transmission lines with characteristic impedance Z_c and having negligible lengths. For this circuit the source and load reflection coefficients are given by

$$\frac{Z_s - 1}{Z_s + 1} \quad \text{and} \quad \frac{r_L}{Z_c + i} = \sim$$

For the amplifier we have

$$v_f = s_{11} v_r + s_{12} v_2$$

$$V_2 = S_{21} V_f + S_{22} V_i$$

But $v_2 = r_L y$, so $V_2 = S_{21} v_f + S_{22} r_L v_2$ or $v_2 = S_{21} y / a - S_{22} r_L i$ we use this in the first equation to get

$$v_f = v_r s_{11} + \frac{s_{12} s_{21} r_L}{i - s_{22} r_L}$$

and hence

$$= r^* \quad 1 - S_{22} T_L \quad (10.10)$$

where $A = S_{11} S_{22} - S_{12} S_{21}$

In a similar way we find that

$$r_{in} = \frac{S_{22} - AI}{1 - s_{11} r_L} \quad (10.11)$$

The input power to the amplifier is given by

$$\text{W/R,}$$

$$\wedge \quad 8^* \cdot \quad * \quad 2IZ_+ + ZJ^*$$

where $|V_S| / 8R_S$ is the available power from the source. First, we will express M_s in terms of I_s and I_{jn} . To do this, we use

$$\text{and } 2R_{ss} = Z_s + Zf \quad \begin{matrix} i + r_s & i + r^* & 2(1 - |rf|) \\ i - i; & i - ir & ii - i; \end{matrix}$$

since $(1 - r_s X l - r^*) = |l - lf|$. Similarly, we get

$$li - r_s J^2$$

Hence we obtain

$$\begin{aligned} M_s &= \frac{4R_m R_s}{Z_s + 2J} \quad \begin{matrix} 4(i - irj^2)(i - |rf|) \\ i - \cancel{|rf|}i - if \end{matrix} \quad \begin{matrix} i + r_s & 1 + r_s \\ i - r_s & + I - r_s \end{matrix}, \\ &\quad \frac{4(i - ir, j^2)(i - ir)}{li - r_{in}^2 |i - rf|} \quad ii - r / li - r_{in} \\ &\quad \frac{(i - ir, l^2)(i - wf)}{li - r, r, j^2} \end{aligned} \quad (10.12a)$$

Similarly, at the output

$$M_o = \frac{li - i'J'ouJ^2}{(10.12o)}$$

The load voltage $V_L = V_2 + Vg = Kfd + F_L$. The load current is $I_L = Y_c(V_2 - Vf)$, so that the power delivered to the load is given by

$$p_L = |Rev,/* = iy_2 i v_2^2 (i + rj(i - vt))$$

Now

$$(i + r_t)(i - Tt) - 1 - rjT + r_E - rr = 1 - |r, \quad 2Imr,$$

$\wedge \bullet = W(i - irj^2) y_r$

so

as expected. We now use the expression for V_2 given above (in ini

\circ btain

The input power is $i|V_2|^2(1 - |f_{in}|^2) Y_c$. Hence

$$\stackrel{G}{P_m} = \frac{(i - irj^2)is_{21}i}{(i - ir_m i^2)u - s_{22}r_L} \quad (10.13)$$

The transducer gain is given by

$$G = \frac{PL}{P_{ava}} \quad W//8R,$$

The available power P_{ava} is related to the input power by $P_{ava} = MP = M_s P_{ava}$ and thus the expression for transducer gain becomes

$$\stackrel{G=EL_M}{P_m} = \frac{C^l - irj^2)(i - ir_l i^2)is_{21}i^2}{u - s_{22}r_L i^2 |i - r_{in}r|} \quad (10.14)$$

We now eliminate r_{in} from the gain expressions by using (10.10) for r_{in} , that is,

$$i - irj^2 = i - \frac{i(i - s_{22}rj)}{ii - s_{22}rj} \quad \frac{li - s_{22}rj^2 - is_u - Arj}{ii - s_Biy}$$

The elimination of i_{in} gives us

$$G = \frac{(i - irj^2)is_{21}i^2}{n - s_{22}rj^2 - |s_H - Ar_L i|^2}$$

Next we use $|1 - S_{22}V_L|^2 = (1 - S^H X 1 - S & I?)$ and similarly $|S_n - & T_L|^2$ to get for the denominator

$$i + s * s & r_t r f - s_{22}i - s_{22}n - s_{22}sf + s_{22}A^*T_L^*$$

$$= 1 - |S_{22}|^2 + irj^2(|S_{22}|^2 - |A|^2) - 2Re(r_t(S_{22} - ASf))$$

Hence

$$G = \frac{(i - ir, r)is_{21}i}{i - is_{22}i^2 + irj^2(is_{22}i^2 - |A|^2) - 2Re(r_t s_{22}ASf)} \quad \text{io.i}^5$$

The same substitution for $|V_s|$ reduces the expression for G to

$$G = \frac{(i - irj^2)(i - ir)/is_{21}i^2}{DL - \ll r_i - s_u r_f + Ar_r r_i r} \cdot \frac{(i - ir_f i^2)(i - ir, i^a)is_2 i^a}{|(i - s_{21}r_j(i - s_n r_j - s_{12}s_{21}r_s r_j)|} \quad (10.16)$$

These are the final forms for the power gain and transducer gain.

When the device (transistor) is absolutely stable, we can use conjugate impedance matching, i.e., choose $Z_s = Z_j^*$, $Z_L = Z_{ul}^*$. In this case $V_s = T^A$, $V_L = r_{ul}^*$, and $M_s = M_r = 1$. Clearly, when $M_s = I$, the power gain Q equals the transducer gain G . For conjugate impedance matching we require

$$i^* = r_o = \quad \text{and} \quad r_i = r_{in}^* = \frac{s_{22}^* - I^* V^*}{1 - S_f I^*} \quad (10.16a)$$

We can substitute the second equation into the first one and solve for Y_s which gives

$$r_o = \frac{sM}{2B_2} A_r + (\lambda - 4IB_r I^2)^{1/2} \quad (10.17a)$$

Similarly,

$$r_L \sim r_{LM} \sim \frac{2B_2}{A_r} \pm (A_r^2 - 4|B_r|^2)^{1/2} \quad (10.17b)$$

where

$$A_r = 1 + IS_r I^2 - |S_{22}|^2 - |A_r|^2$$

$$A_{r2} = 1 + |S_{22}|^2 - |S_n|^2 - |A_r|^2$$

$$B_r = S_r - AS_{22}$$

$$B_{r2} = S_r^2 - AS_{22}^*$$

The minus sign is used when $A_r > 0$ and the plus sign is used when $A_r < 0$ in order to get $|r_{LM}| < 1$ and $|V_{LM}| < 1$. We will show in the next section that for an absolutely stable device $A_r > 0$ and $A_{r2} > 0$, so the minus sign is the appropriate one to use in (10.17). When $|A_r| < 4|B_{r2}|^2$ the solution for V_{LM} can be expressed as $[A_{r2} \pm \sqrt{4|B_{r2}|^2 - A_r^2}]^{1/2}/2B_{r2}$. For this case the magnitude of Y_{LM} is equal to unity which corresponds to a pure reactive load termination and zero power gain. Thus, for an absolutely stable transistor, we must have $|A_{r2}| > 2|B_{r2}|$. It is this condition that leads to the choice of sign to use in (10.17).

In the next section we will show that a device is absolutely stable if the stability parameter K is greater than one, i.e.,

$$K = \frac{|A|^2 - |S_{12}|^2 - |S_{22}|^2 + |A|^2}{2|S_{12}S_{21}|} \quad (10.18)$$

In terms of the parameter K , the power gain for an absolutely stable using conjugate impedance matching, is given by

$$G_p = G_{p,\max} = G = G_{\max} = \frac{|A|^2}{|S_{12}|} (\# - V / r^{\wedge} T) \quad (10.19)$$

The parameter $|S_{21}/S_{12}|$ is called the "Figure of Merit" for the transistor. When $K = 1$ it gives the maximum stable gain $G_{MSG} = |S_{21}/S_{12}|$. The expression for G_{\max} for an amplifier with conjugate impedance matching can be derived following the steps outlined below. By direct expansion we can show that

$$A = 4 \cdot 5 / = 4|S_{12}S_{21}|^2 / (ir^y - 1)$$

From the equations for A_2 and K , we can show that

$$1 - |S_{12}|^2 = A_2 + |A|^2 - |S_{22}|^2 = 2K|S_{12}S_{21}| \sim (|A|^2 - |S_{22}|^2)$$

so by addition and subtraction we get

$$1 - |S_{12}|^2 = K|S_{12}S_{21}| + \dots$$

$$|A|^2 - |S_{12}|^2 = K|S_{12}S_{21}| - \dots$$

We also take note of the relationship

$$2W S_{22} - S_{22}^* A = A_2 \pm (A^2 - 4|5|)^{1/2}$$

which is a real quantity. By using these expressions the denominator (10.15), which is the equation for G_p , can be written as

$$\begin{aligned} & K|S_{12}S_{21}| + A_2 \pm |H| H^* - K|S_{12}S_{21}| - A_2 + 2|S_{12}S_n| JK \\ & = (i - ir, /) + 2iS_{12}S_{21}iV^{\wedge}r^{\wedge}T \quad \frac{i}{r} \frac{A}{f} \frac{1}{K} \frac{S_{12}}{S_{21}} \frac{S_{21}}{S_{22}} \end{aligned}$$

The fast step is to use the solution F_{LM} for T_L to get

$$\frac{i - irJ - 8 \cdot a \gg h f^r I}{B, r} (4 - is \cdot sa \cdot l^{\wedge} 1)$$

With this substitution and a few more algebraic steps, we get

$$\frac{G_p}{|S_{V2}S_2|/(K+i/K^2-l)} = \frac{iS_{,,i'}}{S_{,,}} \frac{(ff \pm V^2 - 1)}{S_{,,}}$$

There are two possible solutions for G_p which correspond to the two possible solutions for V_{LM} . Since G_p should not keep increasing for $K \gg 1$, only the solution given by (10.19) corresponds to the gain obtained with a passive load. We will have a further comment on these two possible solutions at a later point. For now we note that the chosen solution corresponds to using the minus sign in (10.17/); which implies that $A_2 > 0$.

6 AMPLIFIER STABILITY CRITERIA

When the transistor is potentially unstable, which occurs when the stability factor K given by (10.18) is less than one, a stable amplifier can still be designed but only for restricted values of source and load impedances. Furthermore, it will not be possible to use conjugate impedance matching at both the input and output ports. In this section we will derive the expressions for the allowed terminating impedances in terms of input and output reflection-coefficient stability circles. These stability circles can be plotted on a Smith chart and will show what values of source and load impedances can be used in order to achieve a stable (nonoscillating) amplifier.

The conditions for amplifier stability are established by requiring that the reflected power from the amplifier ports be smaller than the incident power. This means that the reflection coefficients looking into the amplifier ports must have a magnitude less than one for all passive source and load impedances. If a reflection coefficient has a magnitude greater than unity, the amplifier input or output impedance would have a negative real part, e.g., if $Z_{,,} = -R + jX$ then

$$|U| = \frac{-R + jX - Z_c}{-R + jX + Z_c} \frac{(R + Z_c f)^4}{(Z_c - R)^2 + X^2} < 1$$

If $Z_m = -R_{,,} + jX_w$ then the input current is

$$I = \frac{V}{R_s - R_{,,} + j(X_m + X_j)}$$

If $R_x = R_m$ and $X_m + X_s = 0$ which can occur at some frequency, then I becomes infinite. We can set $V_s = 0$ and thermal noise in the input can produce self-sustained oscillations at the frequency where the total loop impedance in the input equals zero. Oscillation at any frequency generally makes the amplifier unusable.

The conditions for stability are

$$|r_{\text{L}} - AJ| < i \quad \text{for all } |r_L| < 1 \quad (10.20a)$$

$$\frac{|S_{22} - AI|}{|i - S_u i|} < 1 \quad \text{for all } |r_j| < 1 \quad (10.20b)$$

When these conditions hold conjugate impedance matching can be used. In general, we find that $|\Gamma_{\text{in}}| < 1$ for only a restricted set of values for Γ_L (i.e., Z_L) and $|\Gamma_{\text{out}}| < 1$ for only a restricted set of values for the load impedance Z_s . In this circumstance the device is said to be condition T stable. If $Z_L = Z_s$, then $|r_j| = 0$ and $|r_j| < 1$ only if $|S_{11}| < 1$. Similarly if $Z_s = Z_C$, $|r_{\text{out}}| < 1$ only if $|S_{22}| < 1$. Hence two necessary conditions for absolute stability are

$$|S_{11}| < 1 \quad \text{and} \quad |S_{22}| < 1$$

Values of T , that result in $|r_m| < 1$ are called stable ones. The corresponding region of the Smith chart is the stable region. The boundary between stable values of T_L and unstable ones is the circle in the V_L plane that corresponds to the mapping of the circle $jF^J = 1$ in the T_{in} plane. From the bilinear transformation

$$r_{\text{m}} = \frac{S_u - A r_L}{i - S_{22} r_L} = \frac{A r_L - S_u}{S_{22}^2 r_L + 1} \quad (10.21a)$$

we find that the center of the mapped circle is at ($A = A$, $-S_u = B$, $S_{22} = C$, $-1 = D$)

$$r^* = \frac{S_u A^* - S_{22}^2}{|A|^2 - |S_{22}|^2} \quad (10.21a)$$

and the radius of the circle is

$$R_{LC} \sim \frac{|r_m|}{\sqrt{(|A|^2 - |S_{22}|^2)|}} \quad (10.216)$$

This load stability circle may or may not include the origin $r_L = 0$ as shown in Figs. 10.11a and b. When $Y_L = 0$, $T_{\text{in}} = S_{11}$, and $|S_{11}| < 1$ for a $\ll J$ circuit. Hence, if the stability circle encloses the origin, then all values of r_L inside the circle will give values for Y_m such that $|T_L| < 1$. These r_L are stable ones. If $r_L = 0$ lies outside the circle, then all values of r_L outside the circle are stable ones. The origin is included only if R_{LC} is greater than the distance $|r_{\text{in}}|$ to the center. If this is the case, then we require $R_{LC} > 1 + |T_{LC}|$, so that all values of r_L lie within the stability circle and represent stable values for all possible passive load networks. If the origin is not included, then we require that $|T_{LC}| > 1 + R_{LC}$.

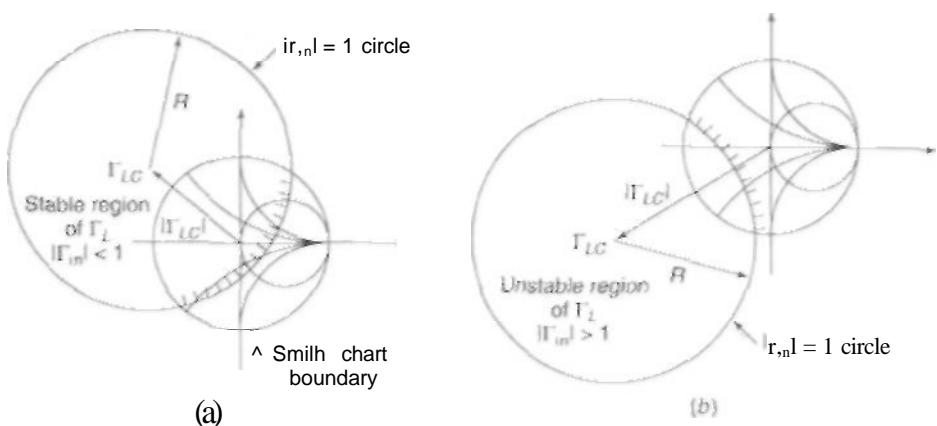


FIGURE 10.11

Load stability circles plotted on the Smith chart, (a) Origin is included within stable region; (b) origin is excluded so load impedances outside stability circle are stable values.

the entire stability circle $|F_{in}| = 1$ maps into a circle outside the Smith chart boundary $|T_L| = 1$. Then again all values of $|V_L| < 1$ will be stable ones.

We need to find an expression that will enable us to state, in terms of the scattering-matrix parameters S_w , S_{12} , S_{21} , and S_{22} of the device, whether or not it is an absolutely stable device. If the amplifier will be stable for all passive source and load impedances, then the device is absolutely stable. If it is stable only for a limited set of source and load terminations, then the device is only conditionally stable. There are two cases to consider and they both lead to the same statement or criterion for stability.

Case 1. Origin ($I^{*A} = 0$) lies outside the circles of V_L values that make $|r_{n,l}| = 1$. As noted earlier this case corresponds to all values of I' , outside the load stability circle $|I''_m| = 1$ being stable values. All passive values of Z' will be acceptable if the $|T_m| \sim 1$ circle lies outside the Smith chart boundary $|I''J| = 1$. This requires that the distance $|V_{LC}|$ to the center of the circle be greater than the radius plus one, that is, $|f_{tc-}| > 1 + R_{LC}$. In order to have $r_{n,l} > 1 + R_{lb}$, we must have $|f_{tc-}| > f_{fc}$. From (10.21) we see that this requires $|f_{tc-}|^2 - S^2 T^2 > S^2 S^2 J^2$. By direct expansion we can show that

$$O_{u-i} \quad & 22I \sim I^{*12} I^{*21} (1 - |S_n|^2) (IS_{22} - |A|^2) \quad (10.22)$$

Hence our stated condition is equivalent to

$$(1 - |S_{22}|^2) (IS_{22} - |A|^2) > 0$$

which is possible only if $|S_{22}| > |A|$ since $|S_n| < 1$. Thus Case 1 can occur only when this condition is true. When the condition is true, then $|(S_{22}I - |A|^2)| = |S_{22}|^2 - |A|^2$. The stability condition $1 + R_{n,l} < |V_{LC}|$ can be stated in the form

$$(Su^* - Sf) > d + R_{LC}f(M^2 - |S_{22}|^2)$$

upon using (10.21a). From (10.216) we obtain

$$\begin{aligned} & \text{is}_{22}\text{r} - |\text{A}|^2 + \text{is}_{12}\text{s}_{21}\text{i} \\ & \ll \text{IS}_{22}\text{I}^2 \sim |\text{A}|^2 \end{aligned}$$

and hence

$$|\text{S}_{12}\text{A}^* - \text{S}|_2^2 > (|\text{S}_{22}|^2 - |\text{A}|^2 + |\text{S}_{12}\text{S}_{21}|)^f$$

Upon using the expansion (10.22), we can restate the condition for stability

$$\begin{aligned} & \text{IS}_{12}\text{s}_{21}\text{r}^2 + (1 - |\text{S}_{12}|^2)(|\text{S}_{22}|^2 - |\text{A}|^2) \\ & > (\text{IS}_{12} - |\text{A}|^2 + |\text{S}_{12}\text{S}_{21}|)^f \end{aligned}$$

After multiplying out the term on the right, canceling the common factor $|\text{S}_{12}\text{S}_{21}\text{r}|$ on both sides, and also canceling a common factor $\text{IS}_{12} - |\text{A}|^2$, we get

$$(|\text{S}_{22}|^2 - |\text{A}|^2) + (2|\text{S}_{12}\text{S}_{21}|) < 1 - |\text{S}_{12}|^2$$

We now introduce the stability parameter K and express the final result in the form

$$\frac{1 - \text{IS}_{12} - |\text{S}_{22}|^2 + |M|^2}{2\text{IS}_{12}\text{S}_{21}\text{I}} \geq 1 \quad (10.23a)$$

$$|\text{S}_{12}| < 1 \quad (10.236)$$

Case 2. It may happen that the circle of T_L values giving $|F_{LC}| = 1$ encloses the origin ($F_L = 0$, $T_{in} = S_u$). In this case all Y_L values inside the circle are stable ones. For absolutely stable devices the circle must then be large enough to enclose the entire Smith chart $|Y_L| < 1$, so that any passive load impedance Z_L can be used. This requires that the radius R_{LC} be greater than one. With reference to Fig. 10.12 it is seen that we now require $d > 1$ or $R_{LC} \sim |LC|^{-1}$ or equivalently $|F_{LC}| < R_{LC} - 1$. This case can occur only if

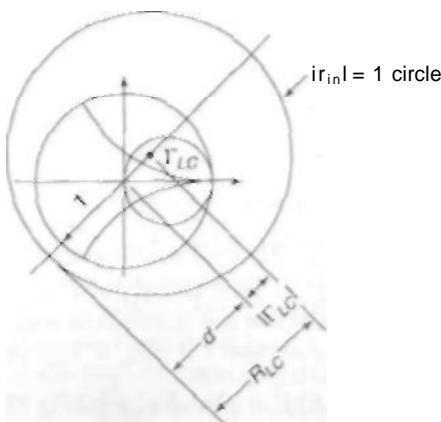


FIGURE 10.12 Illustration for deriving stability condition when the stability circle encloses the origin

using (10.21) we find that we require $|S_{22}I| < |A|$. We now use (10.21a) and (10.21A) along with the expansion for $|S_n A^* - S_2 J|^2$ given by (10.22) to state the stability criterion in the form

$$\begin{aligned} & [|(|A|^2 - |S_{22}|^2)| - |S_{12}S_{21}f| > |S_{12}S_{21}i|^2 + (1 - |S_n f|)(|S_n f| - |A|^2)] \\ \text{or} \quad & (|AI^2 - IS_{12}S_{21}|^2 + IS_{12}S_{21})^2 - 2|S_n S_{21}((|Ar| - [S_{21}]) \\ & > |S_{12}S_{21}|^2 - (1 - |S_n|^2)(|A|^2 - |S_{22}|^2)) \end{aligned}$$

When we cancel the term $IS_{12}S_{21}i^2$ on both sides and a common factor $|A| - IS_{12}S_{21}$ in the remaining terms, the final result is the same as obtained earlier for Case 1.

The requirement that R_{LC} be greater than one places a further restriction on the condition for absolute stability. From (10.2.16) we readily see that, when $|AI| > |S_{21}I|$ in order for $R_{LL} > 1$, we must have

$$|A|^2 - |S_{22}|^2 < |S_{12}S_{21}|$$

From the expression for K we get

$$\frac{1 - |S_n|^2}{|S_{12}S_{21}|} \frac{|A|^2 - |S_{22}|^2}{|S_{12}S_{21}|} = 2K$$

Let $(|A|^2 - IS_{22})^2 / |S_{12}S_{21}| = 1 - .5$, where $.5$ is a positive quantity since this term is less than one. Thus we get

$$\frac{1 - |S_n|^2}{IS_{12}S_{21}I} = 2K - 1 + .5 > 1$$

since $K > 1$. Hence we require that

$$IS_{12}S_{21}I < 1 - IS_{12}S_{21}$$

in order that we can have $R_{lc} > 1$. The condition $K > 1$ is a necessary condition for absolute stability but may not always be sufficient since the condition that $R_{lc} > 1$ when $|A| > |S_{22}J|$ may be a more stringent condition as we will shortly see.

The source stability circle is the circle of source reflection-coefficient values that make $|V'_{out}| = 1$. By direct analogy with the derivation of (10.21), we find that the center for the source stability circle, in the r_s plane, is given by

$$r_s = -\frac{1}{|A|^2 - IS_{12}S_{21}} \quad (10.24a)$$

and its radius is given by

$$R_{sc} = \frac{1}{|A|^2 - IS_{12}S_{21}} \quad (10.246)$$

These equations are the same as (10.21) with S_n and S_{22} interchanged.

Since K is symmetrical in the variables S_{11} and S_{22} , it can be inferred that the same condition $K > 1$ for absolute stability must be met from the requirement $|r_{out}| < 1$, along with the two conditions $|S_{12}S_{21}| < 1 - |S_{22}|^2$ and $|S_{12}S_{21}| < 1 - |S_{11}|^2$.

The necessary and sufficient conditions for absolute stability are thus

$$\frac{1 - |S_{11}|^2 - |S_{12}f + A|^2}{2|S_{12}S_{21}|} > 1 \quad (10.25a)$$

$$|S_{11}I| < 1 \quad (10.25b)$$

$$|S_{22}| < 1 \quad (10.25c)$$

$$|S_{12}S_{21}| < 1 - |S_{11}|^2 \quad (10.25d)$$

$$|S_{12}S_{21}| < 1 - |S_{22}| \quad (10.25e)$$

By adding (10.25a) and (10.25e) and using (10.25a), we get

$$\begin{aligned} 2|S_{12}S_{21}| &< 1 - |S_{11}|^2 - |S_{22}|^2 + |A|^2 + (1 - |A|^2) \\ &< 2|S_{12}S_{21}|K + (1 - |A|^2) \end{aligned}$$

Thus

$$* > 1 + \frac{|A|^2 - 1}{2|S_{12}S_{21}|K} \quad <^{UJ} V >$$

When $|A| < 1$ then clearly this condition holds whenever $K > 1$. However, when $|A| > 1$ the condition (10.25f) is more stringent than (10.25a). For most devices $|A| < 1$, so that (10.25a) to (10.25c) are sufficient to guarantee absolute stability.

Conditionally Stable Devices

For GaAs MESFETs in the common source connection and bipolar transistors in the common emitter connection, one generally finds that $|S_{11}|$ and $|S_{22}|$ are both less than unity. For bipolar transistors in the common base connection, $|S_{11}|$ and $|S_{22}|$ are usually greater than one. Like these, and MESFETs in the common gate connection, one often finds that $|S_{11}|$ will be greater than one. The same is true for the common connection.

For an unstable device there are four possible load configurations and a similar number for the source stability circles. These four cases are described below, along with the necessary conditions to occur.

Case 1. The load stability circle may lie entirely outside the Smith chart since $|S_{11}| > 1$. In this case all values of f_L on the Smith chart are unstable, since the origin is an unstable point. The device in this case

unstable and would be of no interest or use for an amplifier. The necessary conditions for this case to occur are

$$K < -1$$

$$|S_{22}| < |A|$$

$$W_n f > l$$

Case 2. The load stability circle may enclose the Smith chart, and when $|S_{12}I| > 1$ all values of l'_L on the Smith chart will be unstable values. Again, the device will be an absolutely unstable one. The necessary conditions for this case to occur are

$$K < -1$$

$$IS_{22}I > |A|$$

$$IS_{12}S_{21}I < 1 - (2K + 1)$$

Case 3. The load stability circle may lie entirely inside the Smith chart. In this case the device is conditionally stable since a region of the Smith chart will represent stable values of V_L . In order for the stability circle to lie inside the Smith chart, we require that $|r_{f,n}| + R_{LC} < 1$ and $R_{LC} < 1$. When $|S_{22}I| < |A|$ these two conditions lead to the following necessary conditions for this case to occur:

$$K > 1$$

$$|S_{22}| < |A|$$

$$|S_{12}S_{21}| < 1 - (2K - 1)$$

The last condition is equivalent to $R_{LC} < 1$. When $|S_{22}| > |A|$ the corresponding necessary conditions for the load stability circle to lie entirely inside the Smith chart are

$$K < -1$$

$$|S_{22}I| > 141$$

$$|S_{12}S_{21}|^2 < 1 - (2K + 1)|S_{12}S_{21}I|$$

These conditions are obtained by requiring that $|Y_{LC}|^Z < (1 - R_{LV})^Z$ and using (10.21) and (10.22).

When $K > 1$, $|S_{22}| < |A|$, and $|S_{12}S_{21}|^2 < 1 - (2K - 1)|S_{12}S_{21}I|$, the device is absolutely stable. Devices for which $K > 1$ are almost always stable, and since devices with $K < -1$ are not likely to occur, Case 3 is not likely to be encountered in practice.

Case 4. The usual situation that occurs in practice when a device is only conditionally stable is the one for which the load stability circle intersects the boundary of the Smith chart. This case is the one on which we will focus our attention. The load stability circle $|r_{jn}| = 1$ then maps into a circle of $|l'_L|$ values that intersects the Smith chart boundary if $J = 1$ at two points. The load stability circle may or may not enclose the origin. In order to establish

whether values of $|I''_{ln}| = 1$ circle are stable
following rules apply:

- | | | |
|-----|---------------------------|---------------------------------------|
| (a) | $\underline{ir}_{nl} = 1$ | circle encloses the origin |
| | $i < 1 < i$ | all interior values are stable values |
| | $i > 1$ | all exterior values are stable values |
| (6) | $r_c = 1$ | circle does not enclose the origin |
| | $i < i < r_c$ | all exterior values are stable values |
| | $i > r_c$ | all interior values are stable values |

These rules are established by simply noting that the origin $I''_n = 0$ where $r_c = S_n$ is a stable point when $|S_n| < 1$ and is an unstable point when $|S_n| > 1$.

A similar set of rules applies to the source stability circle which is a plot of all values of the source reflection coefficient J_s that gives $|J_s| = 1$. For example, when $|S_{22}| < 1$ the origin $J_s = 0$ giving $r_{sm} = S_{22}$ is a stable point. Thus, if the circle $|J''_{nl}| = 1$ encloses the origin, all values of r_c inside the circle are stable values. They are unstable values when $|S_{22}| > 1$.

By plotting the load stability circle whose center and radius is given by (10.21), it is easy to determine what values of Y_L and hence load impedances Z_L can be used to ensure that $|f_{jn}| < 1$. A similar plot of the source stability circle $|J''_{nl}| = 1$ shows what values of source reflection coefficient Y_s can be used and will ensure a stable circuit.

It is useful to have an analytical test of when the load or source stability circle will enclose the origin. Such a test is easily established by examining the ratio $WLC \wedge RIC - F_{ROM}$ (10.21) and using the expansion (10.22). we obtain

$$\frac{|J''_{nl}|}{|J''_{nl}| - 1} = 1 + \frac{(1 - |S_U I|^2)(|S_{22}| - |A| I^2)}{|S_{22} S_{2V}|}$$

In order for the origin to be enclosed, we must have $|S_U I| < 1$ and $|A| > |S_{22}|$ requires that the last term on the right-hand side of the above equation negative, since the ratio must be less than one. Hence we can state following rule:

The load stability circle encloses the origin when $|S_U I| < 1$ and $|A| > |S_{22}|$ or when $|S_U| > 1$ and $|A| < |S_{22}|$.

When these conditions do not hold, the origin is not enclosed. A similar rule applies to the source stability circle. The origin is enclosed if $|S_{22} I| < 1$ and $|A| > |S_n|$ or when $|S_{22} I| > 1$ and $|A| < |S_n|$.

After the derivation of the equations for the load and source reflection coefficients needed for conjugate impedance matching at both ends of the device, we promised to show that for an absolutely stable device i

are positive. The expression for A_2 is

$$A_2 = (1 - |S_n|^2) - (|A|^2 - |S_{22}|^2)$$

Since $|S_{11}| < 1$ for an absolutely stable device, the first term is positive. For $|S_{22}| > |A|$ the second term is also positive and A_2 is positive. When $|S_{22}| < |A|$ then we must have $|A|^2 - |S_{22}|^2 < |S_{12}S_{21}|$ in order to get $R_{LC} > 1$. This condition comes directly from (10.216) which gives R_{LC} . In order for A_2 to be positive, we require

$$1 - |S_{11}|^2 > |A|^2 - |S_{22}|^2$$

But our stability requirements specify that $1 - |S_{11}|^2 > |S_{12}S_{21}| > |A|^2 - |S_{22}|^2$ and hence A_2 is positive. A similar proof shows that A_1 is positive.

Example 10.1 Stability circles. A MESFET has the following scattering-matrix parameters at 5 GHz:

$$\begin{aligned} S_{11} &= 0.75 + j120^\circ & S_{21} &= 0.08 + j50^\circ \\ S_{21} &= 3.9 + j90^\circ & S_{22} &= 0.4Z - 25^\circ \end{aligned}$$

and at 10 GHz:

$$\begin{aligned} S_{11} &= 0.72/170^\circ & S_{12} &= 0.1Z40^\circ \\ S_{21} &= 2.3/j45^\circ & S_{22} &= QA/L - 55^\circ \end{aligned}$$

We wish to find the stability parameter K , the maximum stable gain, and the load and source stability circles. At 5 GHz we find from (10.23a) that $K = 0.64234$: so the transistor is not absolutely stable at 5 GHz. By using (10.21) we obtain the load stability circle parameters which are

$$Y_{LC} = 3.7678 + j78.3712 \quad K_{LC} = 8.5057$$

From (10.24) we obtain the source stability circle parameters which are

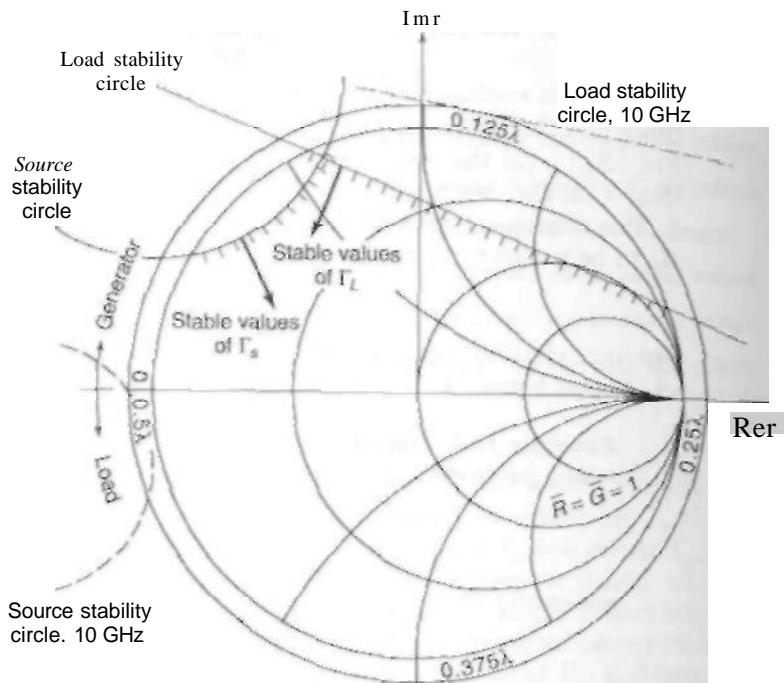
$$I'_{sc} = -0.9971 + jL.19296 \quad R_{sc} = 0.7104$$

These circles are shown plotted in Fig. 10.13. Neither circle encloses the origin, and since $|S_{11}| < 1$ and $|S_{22}| < 1$, all values of I' , and r , outside their respective stability circles are stable values and may be used in an amplifier design.

At 10 GHz we find that for the given parameters $K = 1.0891$; so at this frequency the transistor is absolutely stable. The stability circle parameters are

$$\begin{aligned} I'_{sc} &= -2.1677 - jL0.5364 & R_{sc} &= 11.8549 \\ r_{sc} &= -1.47156 - jL0.28336 & R^* &= 0.47035 \end{aligned}$$

For the load stability circle $|I'|_{sc} = 10.757$, and since $R_{LC} > |r_{LC}| + 1$, the circle encloses the Smith chart. Since $|S_{11}| < 1$ all values of F_L inside this circle are stable values. The circle is a very large one, so that only a small portion of it is shown in Fig. 10.13 (dashed line). The source stability circle is also shown in Fig. 10.13. This circle lies entirely outside the Smith chart, and since $|S_{22}| < 1$ all values of H , outside this circle are stable values.

**FIGURE 10.13**

Load and source stability circles for a MESFET at 5 GHz (solid circles) and 10 GHz (circles). The device is conditionally stable at 5 GHz and absolutely stable at 10 GHz.

10.7 CONSTANT POWER-GAIN CIRCLES

It is convenient to introduce the normalized power gain g_p given by

$$g_p = \frac{G_{31}}{|S_{31}|^2} \quad (10.26)$$

From (10.15) we obtain

$$g_p = 1 - |S_{11}|^2 + V_r J_t (|S_{22}|^2 - |A|^2) - T_L (8_n - AS^*) - I? (S & - * * * -) \quad (10.27a)$$

upon writing $|T_L|^2 = Y_h X_l$ and using the relationship $2 \operatorname{Re} Z = \lambda$ can rewrite the expression for normalized power gain in the form

$$g_p = \frac{g_{31} (S_{22} - AS^*) r_t + g_s (S_{22}^* - A^* S_{11}) r_L^* - (i - \lambda u Q i r j J)}{(|S_{22}|^2 - |A|^2) g_p + 1} \quad (10.27*)$$

by multiplying g_p by the denominator on the right-hand side and bringing the term $1 - |Y_L|^2$ over to the left-hand side and then dividing by the coefficient of V^*V . When we compare the above expression with (10.1), we see that it describes a circle in the T_L plane. The center of the circle is at $\backslash_L = V_{Lg}$, where V_{Lg} is the coefficient multiplying T_L . The radius R_{Lg} of the circle is obtained by noting that the constant term is equal to $|Ij_L| - Rf$. Thus we readily find that the constant normalized power-gain circles have a center at

$$\mathbf{r}_{\text{center}} = \frac{\mathbf{r}_i + \frac{\mathbf{r}_o}{|A|^2} \mathbf{f}^*}{IS_{22}I^2 + |A|^2 g_p + 1}$$

and a radius given by

$$Ri = \frac{1 - 2K_{gp}|S_{V2}S_{21}| + g^2|S_{12}S_{21}|^2}{|(\mathbf{WJ}^* - IA\mathbf{I}^2)^2 + 1|} \quad (\text{io.28ft})$$

where we have used the expansion (10.22) to replace $|S_{22} - A^*S_i|^2$ and expressed $1 - |S_n|^2$ as $|S_{22}|^2 - |A|^2 + 2K|S_{12}S_{21}|$ in order to simplify the expression for R_{Lg} . When g_p approaches infinity the equations above become those for the load stability circle. Thus, on the load stability circle, the power gain is infinite. When $g_p = 0$ we obtain $\Gamma_{Lg} = 0$, $R_{Lg} = 1$, which is the boundary of the Smith chart. On the boundary of the Smith chart, $|Y_L| = 1$, which corresponds to a pure reactive load impedance and consequently no power is delivered to the load and $g_p = 0$, as the above equations verify.

It is common practice to plot the normalized constant power-gain circles that correspond to gains 1 dB, 2 dB, 3 dB, etc., less than the maximum normalized power gain $(K - iJK^2 - 1)/|S_{12}S_{21}|$ [see (10.19)] for an absolutely stable device. For a potentially unstable device, constant power-gain circles for normalized gains 1 dB, 2 dB, etc., less than the normalized "Figure of Merit" gain $(|S_{21}I/|S_{12}|)/|S_{21}|^2 = 1/|S_{12}S_{21}|$ are usually plotted.

The equation for the center of a constant power-gain circle lies on the same ray from the origin as the center of the load stability circle does, since both r_{ic} and \backslash_{Lg} have the same complex numerator except for the factor

By plotting the constant power-gain circles and the load stability circle, we can easily determine those values of load reflection coefficients T_L that will give the largest power gain and yet result in a stable amplifier design. Those values of T_L in a stable region and lying on a given $g_p = \text{constant}$ circle will give a power gain $G_p = |S_{21}|^2 g_p$. Later on we will find that there will be additional constraints imposed on T_L because of low-noise and low input and output VSWR requirements.

Properties of the Constant Gain Circles

By examining the geometric properties of the constant gain circles obtain considerable insight into the operational characteristics of a microwave amplifier. The characteristics of the constant power-gain circles are quite different for absolutely stable devices and potentially unstable devices, so we examine each case separately.

Stable Devices

For a stable device one set of constant gain circles lies entirely inside the Smith chart. The $g_p = 0$ circle coincides with the boundary of the Smith chart. As g_p increases from zero, the radius R_{Lg} decreases and becomes zero when

$$K \pm \sqrt{K^2 - 1}$$

$$\text{or } |S_{12}S_{21}|$$

For circles inside the Smith chart, $R_{Lg} = 0$ when

$$K - \sqrt{K^2 - 1}$$

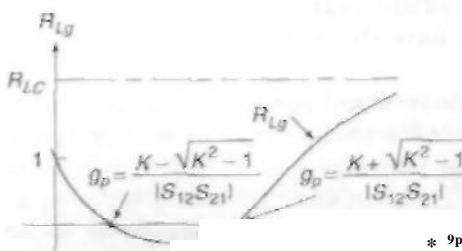
$$|S_{12}S_{21}|$$

When $K = 1$, $R_{Lg} = 0$ when $g_p = 1/|S_{12}S_{21}|$, which gives the maximum gain. A second set of circles exists outside the Smith chart and gives values of g_p for values of $|V_L| > 1$, which implies the use of an active load. For passive loads these circles are not relevant. In the region

$$S_{12} = \frac{K - JK^2 - 1}{|S_{12}S_{21}|} \quad \text{to} \quad \frac{K + fK^2 - l}{|S_{12}S_{21}|}$$

there are no real solution for R_{Lg} when $g_p > 0$ as illustrated in Fig. 1014. Typical constant power-gain circles for a stable device are shown in Fig. 10.15 for the case when the load stability circle lies outside the Smith chart.

Inside the Smith chart boundary the maximum normalized power gain



* No real solution for R_{Lg} in this region for $g_p > 0$

FIGURE J0.14
Illustration of region where there is no real solution for the radius of the constant gain circles.

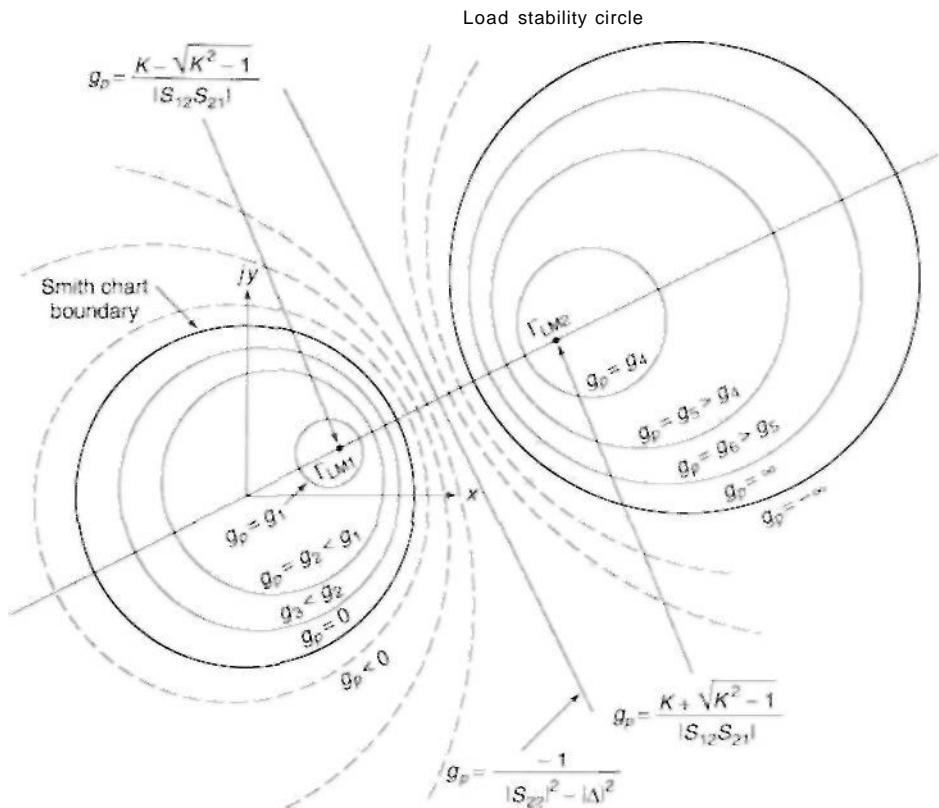


FIGURE 10.15

Circles of constant normalized power gain for an absolutely stable transistor when the load stability circle lies outside the Smith chart. The circles of constant negative power gain are shown by the dashed circles.

$(K - yK^2 - 1)/|S_{12}S_{21}|$ is obtained at the single point where $V_L = V_{LM}$ as given by (10.176). This is the gain for the conjugate-impedance-matching condition. Away from this point the constant gain circles have increasing radii and decreasing gain values. The limiting circle coincides with the Smith chart boundary on which $g_p = 0$.

A second set of similar constant power-gain circles exists inside the load stability circle. In this region the minimum normalized gain is $(if + JK^2 - 1)/|S_{12}S_{21}|$ and occurs for $Y_L = V_{UMZ}$, where now V_{UMZ} is the other solution given by (10.176) for conjugate impedance matching, i.e., using the positive sign in front of the square-root term. As we move away from this point, the constant gain circles have increasing radii and increasing gain values. The limiting circle is the load stability circle on which the gain is infinite. Inside the load stability circle, which is the unstable region

of F_L values, each F_L results in an input reflection coefficient I' magnitude greater than unity. Since $|V_J| > 1$ also in this region it is true that both Z_{in} and Z_L have negative resistive parts. This type of impedance could arise by connecting an unstable microwave amplifier output of the amplifier under discussion. When $|T_{in}| > 1$ and IF $I' > 1$, power is reflected from the amplifier input and load termination incident on either one. The input power and load power are thus negative but the ratio is positive, so that the power gain is positive. A boundary of the load stability circle is approached, which is the circle of values that give $|F_{in}| = 1$, the gain becomes infinite since, when IF $I' = 1$ there is zero input power to the amplifier.

In the region between the Smith chart, and the load stability circle $|F_L| > 1$, but since this is a stable region each F_L produces a I' with $|F_{in}| < 1$. We now have a situation where there is a finite positive input power to the amplifier but a negative load power since, with $|F_L| > 1$, the load termination reflects more power than is incident upon it. Thus the power gain is negative. Indeed, if we assume a negative power gain, we will get another system of constant negative power-gain circles that fills the region between the Smith chart boundary and the load stability circle. As the load stability circle is approached from this region, the gain approaches minus infinity. These negative power-gain circles are shown as dashed circles in Fig. 10.15. The circle with infinite radius occurs when

$$SP = |A|^2 - |S_{21}|^2 / \quad 10.29$$

i.e., when the denominator in the expression (10.286) for R_{Lg} vanishes. On this straight line the gain is constant and is given by (10.29). In order to obtain finite values for F_L from (10.276) when g_p is given by (10.29), the numerator of the second term in (10.276) must vanish. If we let $V_L = x + jy$, we find that setting the numerator equal to zero gives the following equation for the straight line which is the constant power-gain circle having an infinite radius

$$y = * \frac{\operatorname{Re}(S_{22} - AS_{12})}{\operatorname{Im}(S_{22} - AS)} \quad |A| - |S_{21}| + |S_{11}| - 1 \quad 10.30$$

In Fig. 10.16o we show the gain profile along the line ^{connected}_{from} the origin of the Smith chart to the center of the load stability circle. In regions in which the power gain is negative.

When $|A| > |S_{22}|$ the load stability circle encloses the Smith chart. In this case the constant power-gain circles have the configuration shown in Fig. 10.17. For this case the circles again begin at the point $I = 1$ inside the Smith chart and increase in radii and decrease in gain until the circle is IF $J = 1$ on which $g_p = 0$. Beyond the Smith chart boundary the circles are inside the load stability circle, IF $I > 1$, $|F_{in}| < 1$, and the power gain

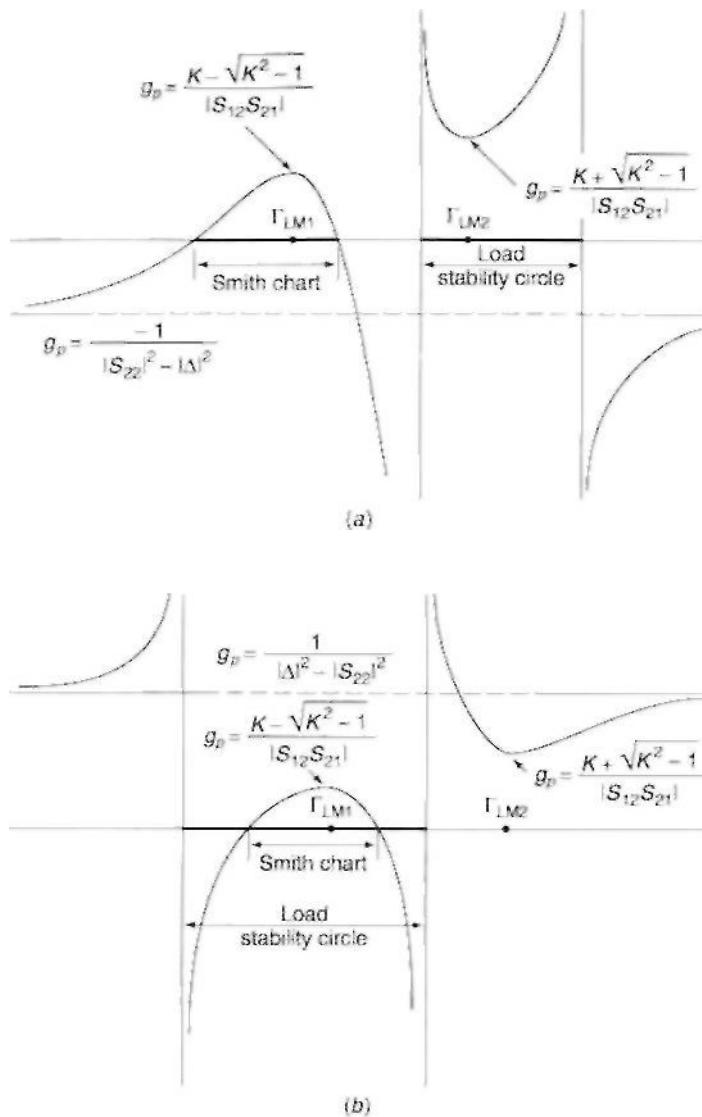


FIGURE 10.16

Gain profile along the line joining the Smith chart center and the load stability circle center.
 (a) Gain profile when the load stability circle lies outside the Smith chart; (b) gain profile when the load stability circle encloses the Smith chart.

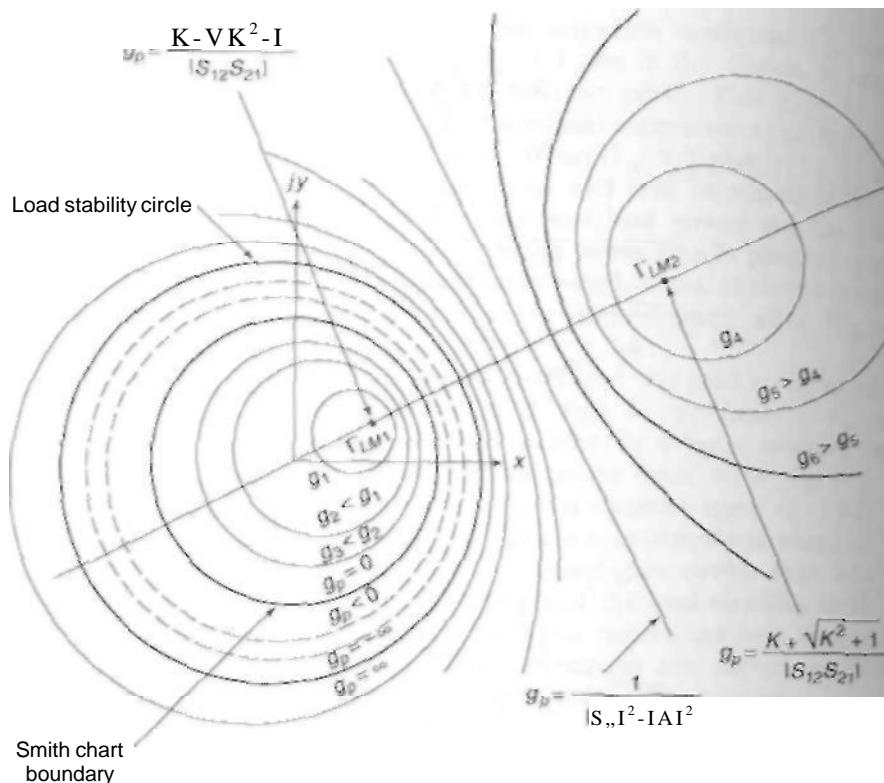


FIGURE 10.17

Circles of constant normalized power gain for an absolutely stable transistor when the load stability circle encloses the Smith chart. The dashed circles are circles of constant negative power gain.

creases in a negative direction and approaches minus infinity on the *load stability circle*. Beyond this circle the radii keep increasing until the straight line given by (10.30) is reached. On this line g_p is given by (10.29) but is now positive since $|A| > |S_{22}I|$. The radii then begin to decrease until the limit point at $r_L = Y_{LM2}$, the second solution given by (10.176), is reached and at which the normalized gain is $(K + J\sqrt{K^2 - 1})/S_{12}S_{21}l$. A gain profile along the ray from the origin to the center of the load stability circle for this case is shown in Fig. 10.16*. When WJ approaches infinity the g_p approaches the value given by (10.29).

Unstable Devices

For an unstable device the load stability circle will usually ^{interf} Smith chart and may enclose the origin. For unstable devices with I_e no value of g_p that is real will make $R_{Lg} = 0$. Again the $g_p = \infty$

boundary of the Smith chart. For most unstable devices the load stability circle will cut the Smith chart boundary at two points. We can show that these two points are invariant points that all constant gain circles pass through. The constant gain circle on which the gain is given by (10.29) is the circle with infinite radius. This circle is the straight line given by (10.30).

We can prove that all constant gain circles intersect the Smith chart boundary at the same two points that the load stability circle does as follows: The load stability circle is given by

$$\text{fit} - hd^2 = R'tc = irj^2 - r ir_{LC}l^2 - r, j, *, - iyr_u.$$

We now let $\mathcal{V}_L = x + jy$. The stability circle intersects the Smith chart when $|T_L| = 1$. By setting $|V_L| = 1$ we then find that the above equation is that of a straight line given by

$$y = \frac{\text{Rer}_{LC}}{\text{Im } l'_{LC}} - \left(\frac{|r_u|^2}{2 \text{Im } r_{iC}} \right) \sim \frac{\text{Rer}_{LC}}{\text{Im } l'_{LC}} - \left(\frac{|r_u|^2}{2 \text{Im } r_{iC}} \right)$$

This line intersects the circle $|T_i| = 1$ at two points. In a similar derivation we find that the constant gain circles will intersect the circle $ll'J = 1$ at the two points where the line

$$\frac{\text{Rer}_u}{\text{Im } T_i} \quad \frac{\underline{\text{ir},/-\text{fl}}, + \text{i}}{21mr_{i,f}}$$

intersects the $|W_L| = 1$ circle. By substituting for $r_{i,r}$, R_{LC} from (10.21) into the first equation and for \mathcal{V}_L , R_{Lg} from (10.28) into the second equation, we find that both lines have the same slope and intercept. We omit the detailed algebra involved but it is straightforward. As a consequence of the above property, the constant normalized power-gain circles for an unstable device have the configuration shown in Fig. 10.18. When Y_L approaches either one of the two invariant points, both the numerator and denominator in the expression (10.27a) for the power gain approach zero. The ratio is thus indeterminate and must be evaluated as a limit with the result that the limit depends on the direction along which the invariant points are approached. Thus we can have many constant gain circles with different gains at the two invariant points, as is clearly evident from an examination of Fig. 10.18.

We will now examine the constant gain circles for the four special cases that can occur in practice. The usual situation is the one where $|S_n| < 1$. When this is true the origin, that is, $F_i = 0$, is a stable point. The load stability circle may or may not enclose the origin. When the origin is not enclosed, the constant gain circles within the Smith chart but outside the load stability circle are in the stable region and \mathcal{V}_L may be chosen to lie on one of these gain circles as shown in Fig. 10.19a. In the region common to both the Smith chart and the load stability circle, the power gain is

$$g_o = |A|^2 - |S^*|$$

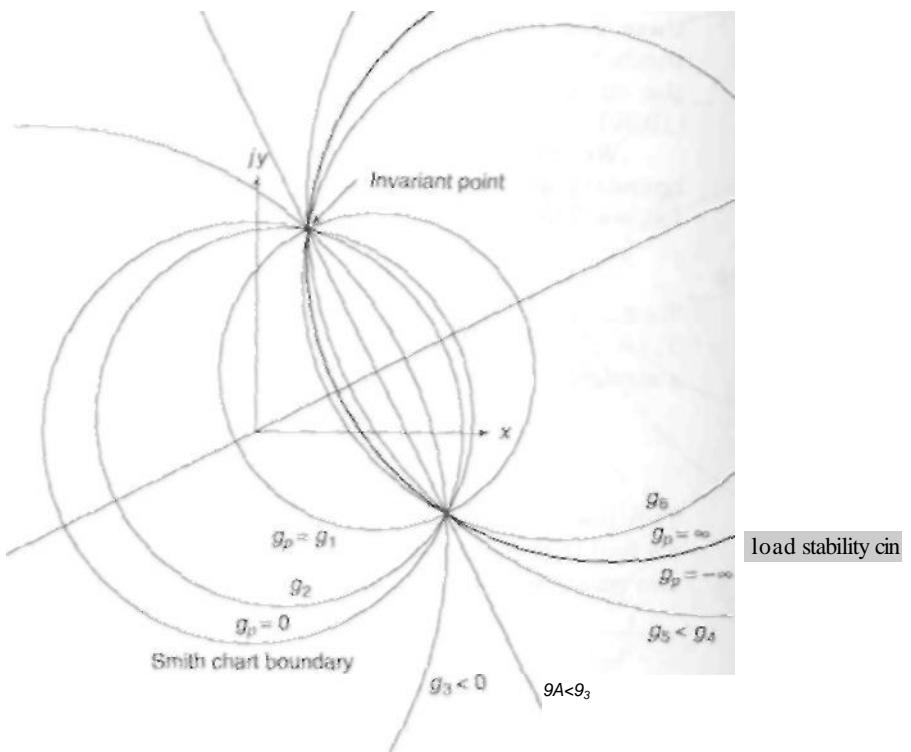


FIGURE 10.18
Constant normalized power-gain circles for an unstable device.

negative. In the region outside both circles, the power gain is also negative; in the region inside the stability circle but outside the Smith chart, the gain is again positive. The gain always changes sign when the circle $|fJ| = \sqrt{-\rho}$ which $g_p = 0$ is crossed and also when the load stability circle on which the gain is infinite is crossed. On one side of the load stability circle boundary, $g_p = \infty$, and on the other side of the load stability circle boundary, $g_p = -\infty$.

For this case the centers of the constant gain circles move from zero ^ Y_{LC} as g_p increases from zero to infinity, as shown in Fig. 10.19a. for o locations of the centers on the ray joining the origin to the stability ^ center, the constant gain circles have negative gain values.

When the stability circle encloses the origin, then the stable region ^ the region inside the load stability circle as shown in Fig. 10.19b ^ stable amplifier design Y_L must be chosen to lie on a constant gain circ this region. For this case the centers of the positive constant gain cir move from zero to infinity in the direction away from Y_{LC} and return

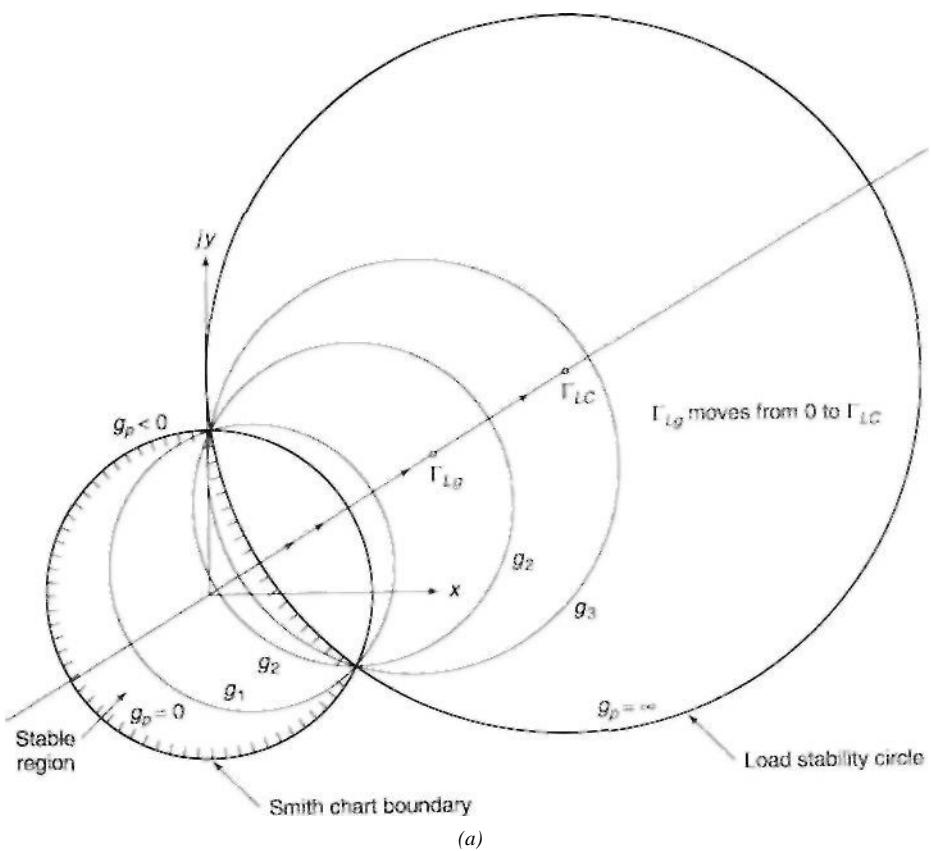


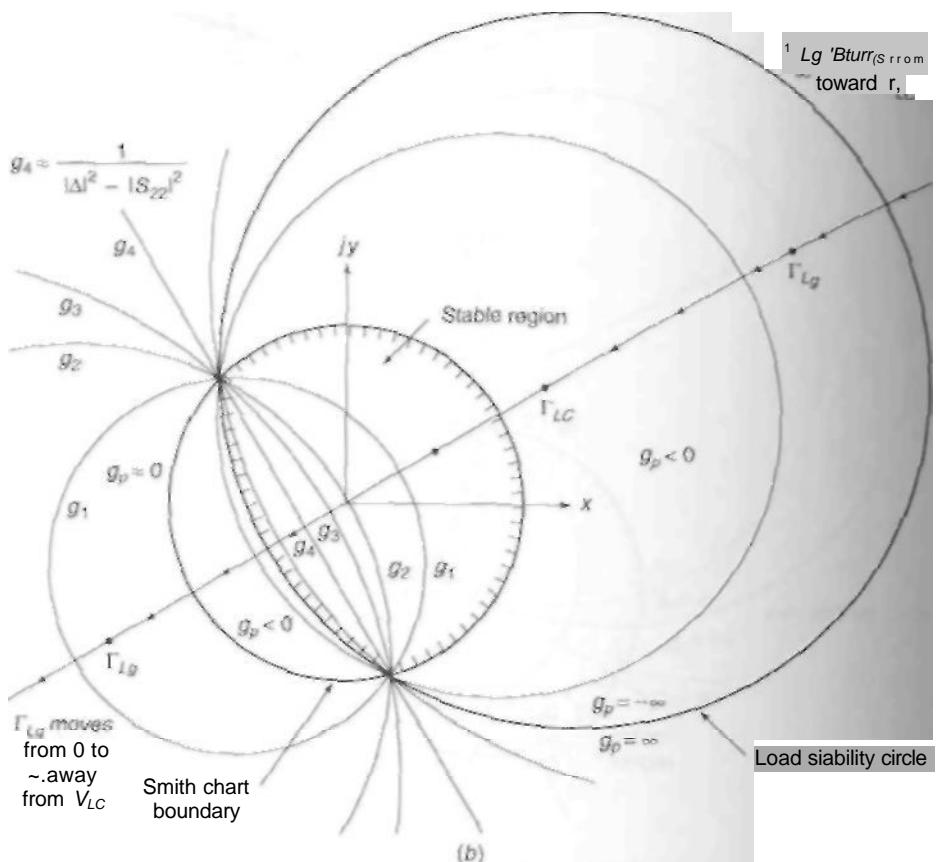
FIGURE 10.19

Constant normalized power-gain circles when $|S_U| < 1$ so that the origin is a stable point, (a) Constant gain circles and stable region when the origin is not enclosed; (b) constant gain circles and stable region when the origin is enclosed.

infinity toward Y_{LC} as the gain increases from zero to infinity, as shown in Fig. 10.196.

The other two cases occur when $|S_U| > 1$ for which the origin is an unstable point. The constant gain circles for these two cases are shown in Fig. 10.20. The stable and unstable regions are interchanged from those shown in Fig. 10.19. The values of Y_L must be chosen to lie on a constant gain circle in that part of the Smith chart in which the origin is not included. Hence, when the stability circle encloses the origin, Y_L must be outside the stability circle for a stable design. When the stability circle does not enclose the origin, then Y_L should be chosen as a point inside the stability circle.

For an absolutely unstable device, $K < -1$, and there are two values of g_p that are negative and correspond to circles of constant gain that have

FIGURE 10.19 *Continued.*

zero radii. These gain values are given by $g_p = \sim(K \pm vK'' - 1)/IS_{12}S_{21}j$. The constant gain circles are similar to those shown in Figs. 10.15 and 10.17 but g_p is negative inside the Smith chart. For the case when the load stability circle lies inside the Smith chart, K may be greater than 1 or less than -1. In both cases there are two points corresponding to constant gain circles with zero radii. The constant gain circles are again similar to those shown in Figs. 10.15 and 10.17. On the load stability circle g_p equals plus or minus infinity and on the Smith chart boundary $g_p = 0$. If the interior of the load stability circle is a stable region, g_p will be positive in this region. Between the load stability circle and the boundary of the Smith chart it will be negative, and outside the Smith chart it will be positive. The interior of the load stability circle is an unstable region, the sign of g_p will be opposite to that described above.

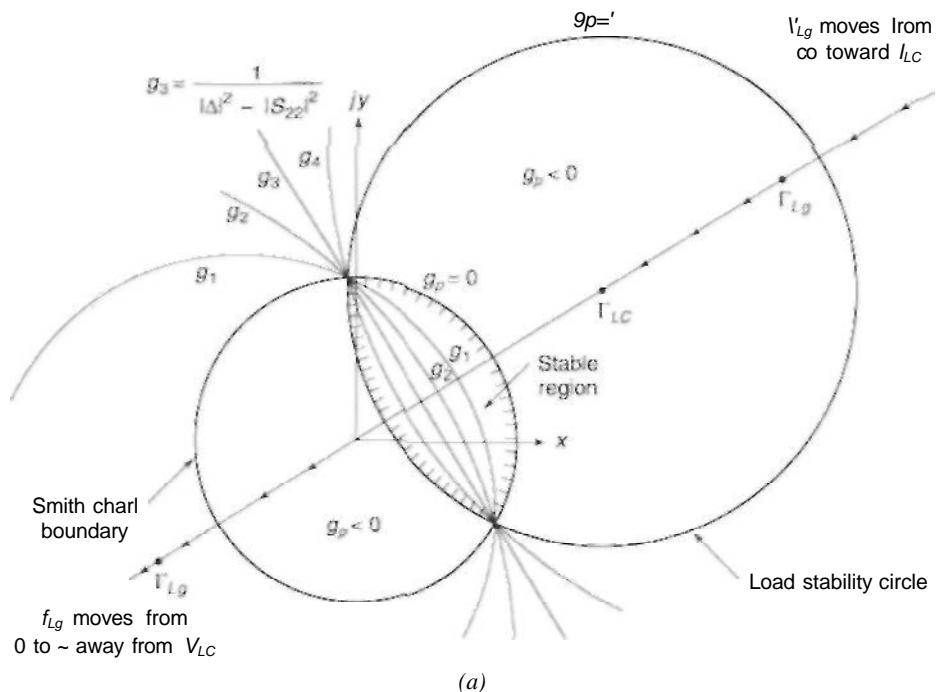


FIGURE 10.20

Constant normalized power-gain circles when $|S'U| > 1$ so that the origin is an unstable point.
(a) Constant gain circles and stable region when the origin is not enclosed; (b) constant gain circles and stable region when the origin is enclosed.

Example 10.2 Preliminary amplifier design. A given microwave transistor has the following scattering-matrix parameters: $S_n = 0.9Z.60"$. $S_{12} = 0.06/160^\circ$, $S_{21} = 3^120^\circ$, $S_{11} = 0.82^1 - 30^\circ$. We will use this device in the design of a single-stage amplifier having 50-12 input and output transmission lines.

For the given transistor we find that

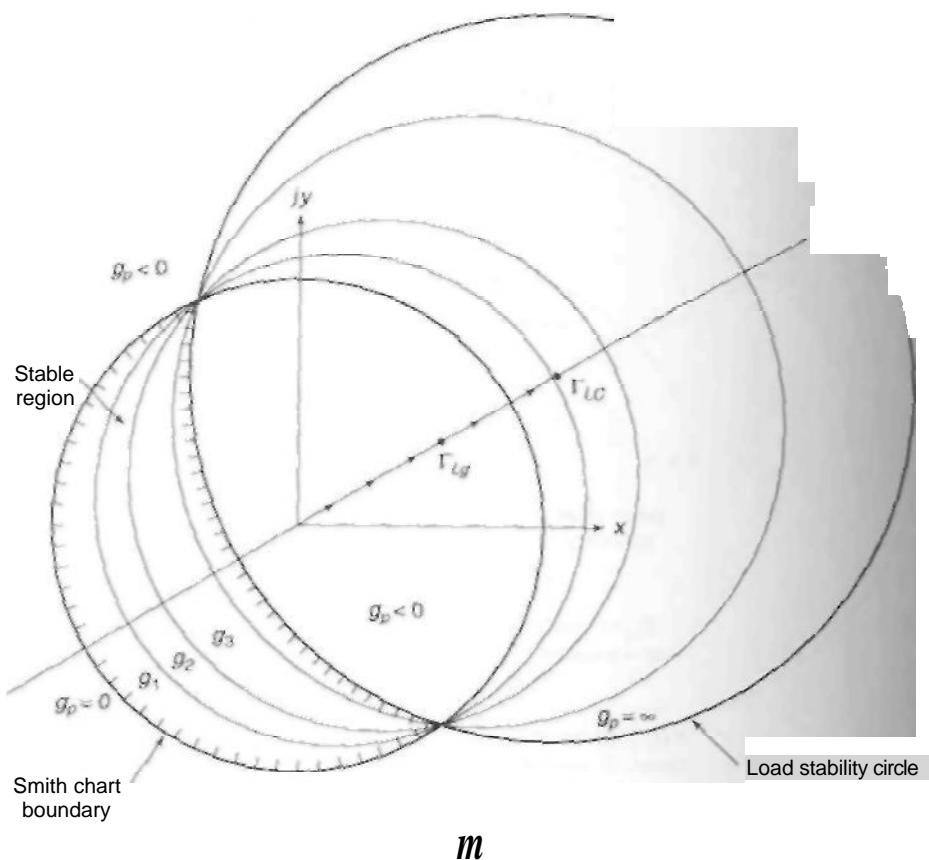
$$\text{Stability parameter } K = \frac{1 - 0.9 - 0.82^* + |A|^2}{2 \times 3 \times 0.06} = 0.902026$$

where $|M| = |S_u S_{2-2}| \sim S_{12} S_{21} = 0.898404$. The transistor is potentially unstable since $K < 1$. The "Figure of Merit" gain is $|S_{21}/S_{12}| = 50$ or 17 dBi. The normalized value is $50/|S_{12}|^2 = 5.555$. The parameters for the stability circles are obtained using (10.521) and (10.24) and are

$$r_{tc} = -0.40026 + j0.46312 \quad K_{lc} = 1.336$$

$$r_{sc} = 6.8262 - j63.243 \quad R_{gc} = 62.7068$$

The centers and radii of the constant normalized power-gain circles for $g_p = 5$,

FIGURE 10.20 *Continued.*

4, and 3 are

$$\begin{array}{ll}
 8., & r_{Lg} = 0.8262 -j0.95596 \quad R_{Lg} = 1.3227 \\
 g_p = 4 & T_{Lg} = 0.4678 \dots -0.5413 \quad R_{Lg} = 1.01606 \\
 g_p = 3 & V_{Lg} = 0.2715 -j0.31417 \quad R_L = 0.9456 \wedge \wedge \wedge \wedge
 \end{array}$$

The stability circles and constant gain circles are shown in Fig- ' . The load stability circle encloses the origin. The source stability circle ^{interse*} only a small portion of the Smith chart, so that almost all values of source reflection coefficients T_s will be stable values. Since $(S_u l < 1)$ represents a stable point, so that all values of f_L inside the load stability are stable values. The source stability circle appears as a straight line in Fig. 10.21 since its radius is very large, that is, $R_{sc} = 62.7$.

From the figure we see that $I_t^* = 0$ is a stable point and would give a normalized gain somewhat greater than 5. From (10.27a) we see that the normalized gain, when $T_s = 0$, is $1/(1 - |S_u l|^2)$ which equals 5.2b. The advantage gained by using $Y_L = 0$ is that no output matching net*

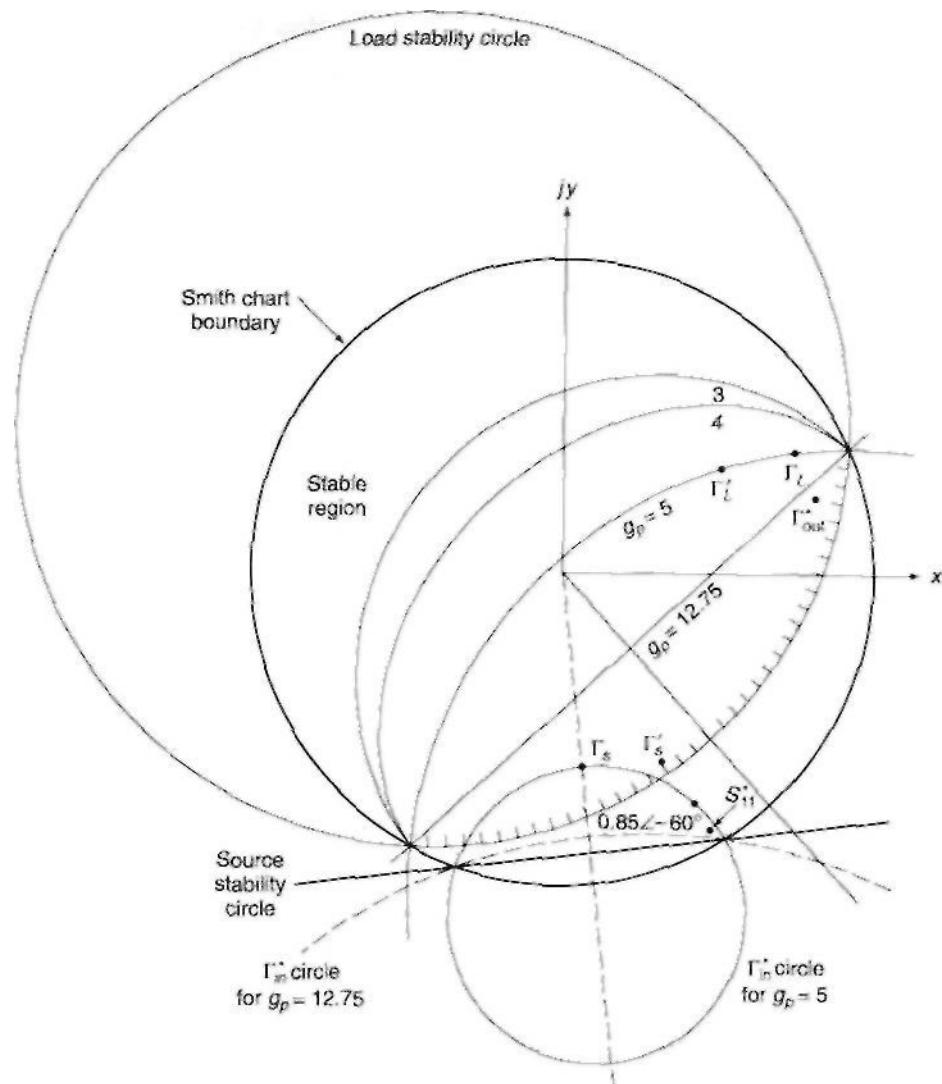


FIGURE 10.21

Load and source stability circles and constant normalized power-gain circles for the amplifier designed in Example 10.2.

required. On the straight line joining the two invariant points, $g_p = 1/(|AI - IS_{22}|) = 12.75$ which is considerably larger than the "Figure of Merit" gain. We could choose I_{\sim} to be on this constant gain circle and still have $|r_{in}| < 1$ with a sufficient stability margin. However, as we will see, the choice of too large a gain will result in poor stability at the output of the amplifier or a very high input and output voltage standing-wave ratio.

If we choose $I_s = 0$, then $r_{ss} = S_{11}$. For conjugate impedance matching at the input, $r_s = T^* = S_{21}$. If we choose this for the source Z_f coefficient, then the amplifier output reflection coefficient becomes

$$r_{out} = \frac{S_{22} - Ai}{S_{22} + AS_1} = \frac{5 - 0.2838 + 70.3384}{5 + 0.2838 - 70.3384} = 0.4416$$

The magnitude of the output reflection coefficient is 0.4416. Since $r_s = S_{21}$, the output VSWR is $(1.4416)/(1 - 0.4416) = 2.58$.

The point $r_s = S_{21}$ is shown in Fig. 10.21 and is seen to lie quite close to the source stability boundary. A relatively small increase in T_L could lead to oscillations in the output circuit of the amplifier since $|r_s| > 1$ when V_L lies inside the source stability circle. A more conservative choice for r_s would be $0.85Z_0 = 60^\circ$. This point is also shown in Fig. 10.21. For this choice the input is not conjugate matched. The mismatch at the input is, from (10.12B)

$$M = \frac{(1 - |r_{in}|^2)(1 - |r_s|^2)}{|n - r_s|^2} = \frac{0.19 \times 0.2775}{0.05523} = 0.9547$$

The input VSWR is now

$$\text{VSWR}_i = \frac{1 + \sqrt{1 + M}}{1 - \sqrt{1 - M}} = 1.54$$

With the above choice for I_s , we find that $r_{ss} = 0.3846 + j0.1615$ and the output VSWR is 2.43. For this choice we improved the stability margin but at the expense of having to accept an input VSWR of 1.54.

Before we leave this example, we will introduce a useful method that helps one to choose a value for I_s that gives a good stability margin, a low input VSWR, and a given normalized power gain. The input reflection coefficient is given by

$$\frac{\text{Sn-Ary,}}{i - S_{21}r_L}$$

The complex conjugate of this equation is

$$\frac{A^*I_s - Sf_1}{sf_2r_L^* - i}$$

This is a bilinear transformation. Thus all values of V_L that lie on a constant gain circle will map into a circle of T^* values. If we choose r_s to lie on the T^* circle, then the input will be matched. We will discuss the FJ circles in more detail in Sec. 10.9. For now we note that the centers of these circles lie on a ray from the origin to the center of the source stability circle. The r_{in} circles also have the interesting property that all circles cut the Smith chart at the same two points that the source stability circle does. This is similar to the corresponding property of constant gain circles for u devices. In Fig. 10.21 we show the F^* circle for all values of T_L on the F^* circle. Some of the T^* values will correspond to Y_L values outside the U^* chart. The best stability margin is obtained by choosing F^* as the point as T_s in Fig. 10.21, since this is farthest away from the source stability circle.

For this example the T^* circle has a center at $0.11833 - j1.09625$ and a radius of 0.46922 . The value of $\|$ at the illustrated point is $0.06797 - j0.6297$. With this choice for $T_s = I^*$, the *input is matched and we get a normalized power gain of 5 or a real power gain of $5 \times |S_{21}| = 45$ or 16.53 dB*. However, we must check if the corresponding value of I^* is satisfactory. With the above choice of T_s , we find that $r_{s,u} = 0.7819 - j0.2077$. For the chosen value of r^* , we can solve for V , using the equation

$$r_A = \frac{S - I}{A^* o_r} = 0.7097 + j70.36165$$

The value of $\|_L$ for the chosen F_{ss} is well within the load stability circle. In Fig. 10.21 we show the points T_L and $r_{s,u}^*$. The output VSWR is easily calculated and is 2.52. Thus this design gives a matched input, a good power gain of 45, good stability, and an output VSWR equal to 2.52. However, we do need an *output matching network that will transform the line impedance of 50 Ω into a load impedance given by*

$$Z_L = \frac{1 + \|_L}{1 - \|_L} Z_o = 85 * j168.2$$

A better output VSWR could be obtained by relaxing the input VSWR. By trial and error we find that, by using $T_s = 0.35 - j0.6062$ and $\| = 0.47237 - j0.31859$, we can achieve a normalized power gain of 5, an input VSWR of 1.366, and an output VSWR of 1.8515. These final values of T_s and $\|_L$ are shown in Fig. 10.21 as the points I^* and $\|_L$. This last design results in good gain, low input and output VSWR, and good stability margins.

The last item that we will examine is what happens if we try to design the amplifier for a gain significantly larger than the "Figure of Merit" gain. A very large gain means that the amplifier is very close to oscillation. Hence it can be anticipated that either T_s , $\|_L$, or both will be close to their respective stability circle boundaries. The I^* circle for all values of I^* , on the $g_p = 12.75$ gain circle is shown as the dashed circle in Fig. 10.21. This r^* circle lies very close to the source stability circle. For a matched input, I^* , would have to be chosen to lie on this I^* circle and clearly this would not give an adequate stability margin. If we choose T_s sufficiently far away from this circle to obtain a good stability margin, then we will end up with a very poor impedance match at the amplifier input. The consequence of a poor impedance match at the input is that only a small fraction of the incident power on the input transmission line will be delivered to the amplifier input. Thus, even though the power gain will be high, the transducer gain will be low. In practice, microwave amplifiers are generally designed for power gains no greater than the "Figure of Merit" gain because of the poor stability and high input and output VSWR that results for larger gain designs.

As is apparent there is no direct way to control all design objectives individually, since improvements in some areas generally are accompanied by a deterioration in some other characteristic.

In Sec. 10.9 we will discuss low-noise amplifier design and will find that this will result in an additional constraint on V_s since the best noise figure is obtained only for a particular value of I_s .

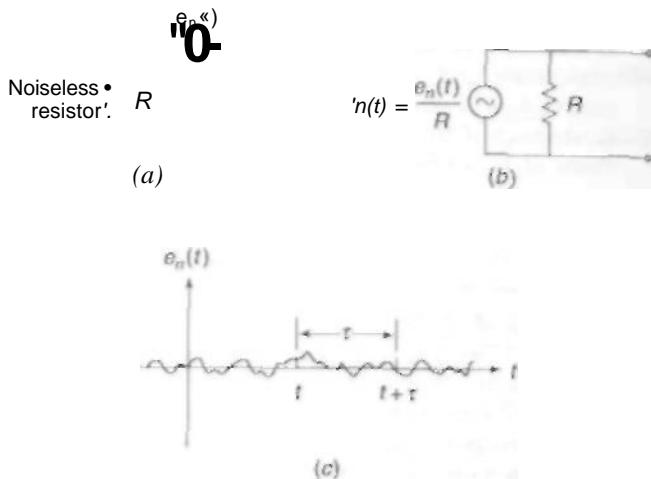


FIGURE 10.22

(a) Thevenin equivalent circuit which uses a noise voltage generator; (b) equivalent circuit for a noisy resistor in which a noise current source is used; (c) typical noise voltage waveform.

10.8 BASIC NOISE THEORY

As a result of thermal agitation, the electrons in a resistor have an inherent random motion which results in a random voltage appearing across the resistor terminals. This random voltage is referred to as noise. There is no analytical way to describe the exact voltage waveform; so we must be content with a description of certain average characteristics of the noise. We can model a noisy resistor as a noise-free resistor in series with a noise voltage generator $e_n(t)$ or in shunt with a noise current source $i_n(t)$ as shown in Pigs. 10.22a and b. In Fig. 10.22c we show a typical noise voltage waveform that might be produced.

Noise is a random process and its effects in a linear system are analyzed using statistical methods. For this purpose we construct an ensemble of macroscopically identical systems, e.g., we consider an infinite number of resistors with each one producing its own noise voltage. Averages $\langle \dots \rangle$ various products of the noise voltages at different times, such $\langle e_{n1}(t_1) e_{n2}(t_2) \dots \rangle$, etc., are obtained by averaging over the ensemble of noise waveforms.

Thermal noise is generally regarded as a stationary ergodic noise process which is a random process for which ensemble averages $\langle \dots \rangle$ are replaced by time averages. Thus, in our brief summary of basic noise theory we will use time averages. Since our objectives are only to obtain the results needed to derive the equations required for low-noise amplifier design, our discussion will be brief. Much more complete treatments of noise are widely available in texts on statistical communication theory.

the reader should consult some of these for the details that are missing in our treatment.

The time-average value of the noise voltage, given by

$$\langle e_{av}(t) \rangle = Hm \overline{\langle e_{av}(t) \rangle} dt = 0 \quad (10.31)$$

is zero. The correlation function for the noise voltage is the average value of the product of the noise voltage at time t and that at a later time $t + r$, thus

$$\begin{aligned} C(r) &= Hm \overline{\langle e_{av}(t) e_{av}(t+r) \rangle} \\ &= \langle e_n(t) e_n(t+r) \rangle \quad (10.32) \end{aligned}$$

where $C(r)$ is the correlation function. If $r = 0$ we obtain the average power $\langle e^2 \rangle$ associated with the noise (the noise voltage is thought of as being applied to a 1-ohm resistor, so that the dimensions are those of power). The average power in noise is distributed over a broad band of frequencies because noise voltage waveforms contain a broad spectrum of frequencies. The power spectral density $S_{av}(w)$ of noise is given by the Fourier transform of the correlation function; thus

$$S_{av}(w) = \int C(r) e^{-jwr} dr \quad (10.33a)$$

The inverse transform relationship is

$$C(r) = \frac{1}{2\pi} \int S_{av}(w) e^{jwr} dw \quad (10.33b)$$

The power spectral density represents the noise power in the spectral domain; so $S_{av}(w) A f$ is the noise power in a frequency increment Δf .

At room temperatures the power spectral density of thermal noise is constant up to frequencies of the order of 1,000 GHz and decreases at higher frequencies. Thus, at microwave frequencies and below, we can assume that the spectral density is a constant or is flat. This is equivalent to having a correlation function that is a constant multiplying the delta function $\delta(r)$, that is,

$$C(r) = C_0 \delta(r)$$

since the Fourier transform of $\delta(r)$ is a constant equal to unity. Noise with a constant power spectral density is called white noise and is uncorrected noise.

The power spectral density is an even function of fr ; so we can choose the spectral density such that only positive values of w need to be considered. For thermal noise in a resistor, the power spectral density for the noise voltage is given by Nyquist's formula

$$S_c(t_0) = 4kT R \quad t_0 > 0 \quad (10.34)$$

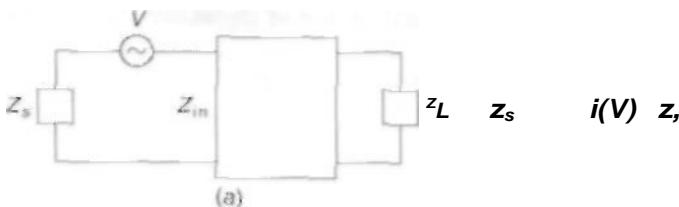


FIGURE 10.23

A two-port network connected to (a) a voltage source; (b) a current source.

where $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant and T is the absolute temperature of the resistor R . Thus the amount of noise power P_{noise} in a frequency interval Δf is given by

$$P_{\text{noise}} = UTRAf \quad (10.35)$$

Thermal noise in a resistor is also called Johnson noise, after one of the early investigators of noise.

If we use the equivalent current source model shown in Fig. 10.226, then the average power, if the current $i_n(t)$ flows in a $1\text{-}k\Omega$ resistor, is given by $(if,0)$ and has a power spectral density given by

$$S_f(\text{ft}^2) = \frac{4kT}{R} \quad w > 0 \quad (10.36)$$

for thermal noise.

Filtered Noise

Consider a sinusoidal voltage generator with a complex rms voltage V that is connected in series with a source impedance Z_x and a two-port network as shown in Fig. 10.23a. The input current produced by V is $V/(Z_{in} + Z_m)$, where Z_m is the input impedance to the network. The input power produced by V is given by

$$P_{\text{in},1} = \frac{V^2}{Z_x + Z_{in}} \cdot \frac{47?}{IZ_S + Z_{ia}} \cdot \frac{i_{fe}^2}{M} \quad (10.37)$$

where $|V|^2/4R_S$ is the available power and M is the impedance-mismatch factor. The power transfer function is $M/4i?_s$ and is a function of ω . Z_x and Z_{in} are functions of ω . If the voltage generator is replaced by a voltage source $e(j\omega)$ with a power spectral density $S_e(j\omega)$, then the noise power in a frequency band Δf centered on ω is given by P_{noise} the source power spectral density by the power transfer function

factor A /'; thus

$$M(w)$$

(10.38)

The total input noise power is obtained by integrating over all frequencies; thus

$$,* M(ui) dai$$

.10.39)

Consider next the circuit shown in Fig. 10.236. The input current from the current source is $Z_s I / (Z_s + Z_{in})$ and the input power is

$$P_{in,2} = \left| \frac{iz_s}{Z_s + Z_{in}} \right|^2 fl.,, = \frac{|I|^2 |Z|}{4R_s} \frac{4fl,*_{in}}{Z_s}$$

$$\frac{|I|^2}{4\&^M} \quad (10.40)$$

where $G_s = \text{Re}(1/Z_s) = R_s / Z_s$. If the sinusoidal current source is replaced by a noise current source $i_{in}(t)$ with power spectral density $S_{in}(w)$, then the input noise power in a frequency band Af is given by

$$P_{in,2}(w) = S_{in}(w) A f - \frac{M(w)}{4G_s} \quad (10.41)$$

The power spectral density of the noise power delivered to Z_s would, by analogy, be given by the product of the power spectral density of the source and the power transfer function from the source to the output load impedance. Since we can describe the power delivered to Z_L by the product of P_m with the power gain $G_p(co)$, we see that the noise current source will produce an output power in Z_L with a spectral density given by $G_{out}(w)Af(w)S_{in}(w)/4G_s$. A similar expression holds for the output power spectral density produced by the noise voltage source $e_n(t)$ acting alone.

When both sources V and I are acting, the input current will be

$$V + /Z.$$

and the input power will be given by

$$(V + iZ_s)(V^* + i^*Z^*) \quad R_{out}$$

$$\frac{iVr}{4fl<} \frac{i/r}{M + T^4 G_s} \frac{Vi^* Z^* R_{in}}{M + 2Re - iZ_s} \quad (10.42)$$

Because of interaction between the two sources, the input is simply the sum of that from each source acting independently.

When we have two noise sources $e_{rl}(t)$ and $i_n(t)$ acting simultaneously there is no input power caused by the interaction between them if $e_{rl}(t)$ and $i_n(t)$ are uncorrelated or statistically independent sources. For this case the input noise power in a frequency band A_f is the sum of that given by (10.38) and (10.41). When there is a dependence correlation between $e_{rl}(t)$ and $i_n(t)$, there will be some input noise due to the source interaction.

The cross-correlation between the current source $i_n(t)$ and voltage source $e_{rl}(t)$ is given by

$$C_x(T) = \text{hm} \int_{-\infty}^{\infty} wJ (TJMKit + r) dt \quad (10.43)$$

The Fourier transform of CJT gives the cross-power spectral density $S_x(o)$, that is,

$$S_x(w) = S_{rl}(o) + jS_{xl}(u) = \int_{-\infty}^{\infty} C_x(r)e^{-jmr} dr \quad (10.436)$$

If we replace to by $-to$, we see that $S_x(-u) = S_x(to)$ since $C_x(r)$ is real. From this result we find that $S_x(to)$ is an even function of to and $S_r(o)$ is an odd function of to . For input noise power calculations, we replace $|V|$ by the noise voltage source power spectral density $S_e(to)$, replace $|I|$ by the power spectral density $S_i(to)$ of $i_n(t)$, and replace VI^* by the cross-power spectra density $SJ(to)$ in (10.42). Thus, for partially correlated noise sources, the total input noise in a frequency band A_f is given by

$$P_n = A_f(S_e(\frac{M}{4a}) + S_r(to)\frac{M}{4a} + \frac{4[S_{xr}(to)R_s + S_{xi}(to)X_s]}{|z_r + z_i|} - K) \quad (10.44)$$

The extra factor of 2 in the last term is due to the fact that we have combined the contribution from negative values of to with those non-positive values of to using the fact that $S_{xr}R_s$ and $S_{xi}X_s$ are even junctions of to . Also, we have defined the spectral densities $S_e(to)$ and $S_i(to)$ such that only positive values of to are to be integrated over to get the total input noise power.

The power spectral density of the noise produced in a network is from that of the noise source because the network response depends on the frequency. The source noise spectrum is filtered by the network. The source noise spectrum is flat (white noise), the noise spectrum at some point in the network is not flat. Noise with a nonconstant power spectral density is called colored noise.

From (10.44) we find that the power spectral density of the noise power delivered to Z , is

$$\frac{M}{4R} \cdot \frac{M}{4G} \cdot " = M G_b(\text{co}) \quad (10.45)$$

after multiplying by the power gain G_p of the two-port network. The total output noise power delivered to Z_L is

$$P_{\text{out},n} = f \int S(\theta) d\theta \quad (10.46)$$

Noise in Active Devices

In an active device such as a transistor, there are three main contributions to the noise produced by the device: (1) thermal noise in the resistive elements that are present, (2) shot noise due to the discrete nature of the charge carriers that constitute the current flow across $p-n$ junctions, and (3) flicker noise that has a power spectral density proportional to $1/\theta^2$ in the frequency domain. The current that flows across a $p-n$ junction is not a smooth stream of charge. It consists of discrete charges that cross the junction in a more or less random manner like raindrops falling on a roof. The noise associated with this current is called shot noise and is directly proportional to the dc bias currents. Shot noise, like thermal noise, has a flat power spectral density. For low noise an active device is operated with low dc bias currents. However, since the transconductance, and hence the gain of the device, decreases with a reduction of the bias current, there is an optimum lower value of bias current that gives the best compromise between low noise and gain and produces the optimum noise figure for the device. In Fig. 10.24 we show the typical behavior of noise figure versus drain current for a MESFET.

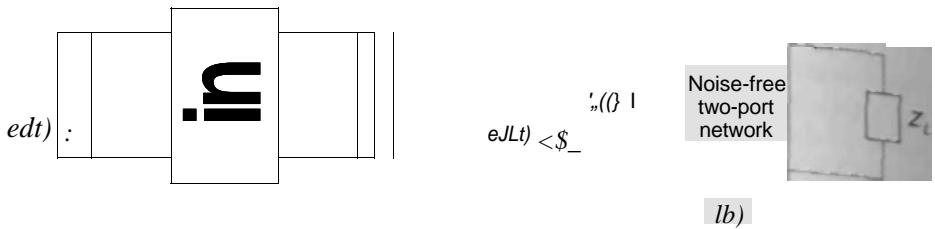
2 -

1.5

10	20	30	40
mA			

FIGURE 10.24

Variation of the noise figure F of a MESFET as a function of drain current.

**FIGURE 10.25**

Equivalent input noise sources for a noisy linear two-port network.

Flicker noise is usually negligible relative to thermal noise and shot noise at frequencies greater than a few kilohertz for bipolar transistors but may be important in MESFETs up to frequencies as high as 100 MHz.

Noisy Two-Port Networks

In analyzing the noise produced at the output of a linear two-port network due to the internal noise sources, we can replace all of the internal noise sources by a series noise voltage generator $e_{n0}(t)$ and a shunt noise current generator $i_{n0}(t)$ at the input as shown in Fig. 10.25. The total noise power at the output can be found by evaluating the noise output produced by $e_{n0}(t)$, $i_{n0}(t)$ and the thermal noise in the resistive component R_s of the source impedance. Two equivalent noise sources are needed at the input because if the input is short-circuited, that is, $Z_s = 0$, the source $i_{n0}(t)$ does not produce any output noise, yet the noisy two-port does have a noise output under short-circuit conditions at the input so a noise voltage source $e_{n0}(t)$ is required. Similarly, under open-circuit conditions $e_{n0}(t)$ does not produce any output noise; so a noise source $i_{n0}(t)$ is needed to represent the equivalent input noise source under open-circuit conditions. The two noise sources $e_{n0}(t)$ and $i_{n0}(t)$ are not completely independent since a part of $e_{n0}(t)$ and $i_{n0}(t)$ may arise from the same basic noise-producing mechanism within the two-port network. Thus, in general, there is some cross-correlation between $e_{n0}(t)$ and $i_{n0}(t)$ with a resultant nonzero cross-power spectral density. j

It is common practice to express the power spectral density as S_{xx}^n with the two noise sources $e_{n0}(f)$ and $i_{n0}(t)$ in a form similar to that in (10.34) and (10.36) for thermal noise. When this is done, nicker noise, which is low-frequency noise with a X/f spectrum, is not included, since microwave amplifiers do not produce an output for low-frequency signals, the neglect of flicker noise can be justified. We will thus specify

spectral densities as follows:

$$\text{For } e_{in}(t) \quad S_{in}(w) = 4A77?.$$

$$\text{For } i_{in}(t) \quad S_{in}(w) = AkTG,$$

and

$$2[S_{xr}(w) + JS_{xv}(w)] = 4kT(y_r + jy_i)$$

where R_c is an equivalent noise resistance, G_h is an equivalent noise conductance, and $y_r - jy_i$ is a complex equivalent noise impedance. A total of four parameters, R_c , G_h , y_r , y_i , are needed to describe the noise properties of a noisy two-port network.

In terms of the above spectral densities, we can express the total noise input to the noise-free two-port network, in a frequency band A' , as follows by using (10.44):

$$\begin{aligned} P_{nin} &= hTlfM + kTAf^{\Delta}M + kTLf^{\Delta}M \\ &\quad + 2kTbf^{\Delta}M \end{aligned} \quad (10.47)$$

In this equation the first term on the right, $hTlfM$, is the input thermal noise from the source resistance r_s . The output noise in Z_o in a frequency band A' is obtained by multiplying by the power gain $G_o(ui)$ of the network. The noise produced at the output termination by the equivalent sources e_{nit} and $i_{it}(t)$ placed at the input of the network is fully equivalent to that produced by the internal noise mechanisms in the real noisy two-port network.

10.9 LOW-NOISE AMPLIFIER DESIGN

In a typical microwave communication system, the information to be communicated is modulated onto a microwave carrier and radiated into space by means of an antenna. At the receiving site an antenna is also used to intercept a small portion of the radiated signal energy. The receiving antenna will also pick up a certain amount of noiselike radiation from atmospheric disturbances, radio stars, the sun, and other celestial bodies. The received signal, along with some noise, is very weak and must be amplified to a level where it can be used to produce the desired video, audio, or digital output information that was transmitted. The function of the first amplifier stage is to amplify the signal with the addition of a minimum amount of extra noise. Thus the first amplifier stage should be designed for minimum noise. If the power gain of the first stage is around 10 or more, the signal will be sufficiently large at the output of the first stage, so that additional noise contributed by the following amplifier stages will have a

negligible degrading effect on the overall signal-to-noise power ratio provided that the noise contribution of the second stage is moderate. In the design of the first stage, the minimum noise requirement is more important than maximum power gain or output VSWR, provided a power gain of $\frac{1}{F}$ more can be achieved. Hence we can relax the gain and output voltage requirements in order to achieve the objective of a minimum noise contribution from the first stage.

In this section we will discuss noise figure and the design of amplifier for minimum noise. We will show that there is an optimum source impedance Z_s (or source reflection coefficient V_s) that will result in the lowest noise figure. We will also introduce constant noise-figure circles that can be plotted on the r_s plane and which will show in a pictorial way the increase in noise figure that occurs when the optimum source reflection coefficient cannot be used. For transistors that are not absolutely stable, the use of the optimum source reflection coefficient for minimum noise could result in an unstable amplifier, in which case a noise figure somewhat larger than the minimum one will have to be accepted.

Noise Figure

With reference to the circuit shown in Fig. 10.25a, the definition of noise figure F (also called noise factor) is

$$F = \frac{\text{signal-to-noise ratio at input}}{\text{signal-to-noise ratio at output}} \quad (10.48)$$

The output noise is the amplified thermal noise from the source resistance plus the noise produced by the amplifier. The standard definition of F requires the source to be conjugate impedance matched to the network, that is, $Z_s = Z^*$, and the source resistance R_s to be at the standard temperature $T_0 = 290$ K. Very often in practice the source resistance is at a different temperature and the source is not matched to the network. In this case the definition (10.48) gives the operating noise figure. If F is given at a single frequency and is based on the noise power in a small frequency band Δf then the noise figure is called the spot noise figure. When all of the noise sources are referred to the input as equivalent noise sources, then the spot noise figure can be defined as follows:

$$\frac{\text{total input noise power to network}}{\text{thermal noise input power from source resistance}}, \quad (10.49)$$

where the noise powers are those in a narrow frequency band Δf . In this latter definition, the spot noise figure for the system shown in Fig. 10.25a is

10.25 may be obtained from (10.47) by dividing by $kT A f M$. Thus we get

$$\frac{F}{G_s} = \frac{R_s y_r + X_a n}{R_s^2 + G_s^2} \quad (10.50)$$

This noise figure is seen to depend on the source impedance as well as on the noise parameters R_s , G_s , y_r , and X_a . In (10.50), $G_s = R_s j R_s' i + X_a^*$. The noise figure does not depend on the frequency bandwidth Af . However, the input signal-to-noise ratio will deteriorate if the amplifier bandwidth is greater than that required to accommodate the signal.

The optimum source impedance that will minimize the noise figure is obtained by setting $dF/dR_s = 0$, $dF/dX_a = 0$, and solving for R_K and X_a . It is readily found that

$$X_a = \frac{R_s}{G_s} \quad (10.51a)$$

and $R_s^* + X_a^* = R_s/G_s$, which gives

When these values for R_s , and X_a are used in (10.50), we obtain the minimum noise figure F_m .

We now replace y_r by $-X_a/G_s$ in (10.50) and consider $F - F_m$ which is given by

$$\begin{aligned} F - F_m &= \frac{[R_s + (8\pi f X_a^*) G_s - 2X_m X_a G_s]}{R_s^2 + G_s^2} \\ &\quad - [R_s + (R_s^2 + R_m^2) G_s - 2X_m X_a G_s] \\ &\quad - j[R_s + (R_s^2 + R_m^2) G_s + (X_m^2 + X_a^2) G_s^2] \\ &\quad - i[R_s + (R_s^2 + R_m^2) G_s + (X_m^2 + X_a^2) G_s^2] \\ &\quad - [R_s + (R_s^2 + R_m^2) G_s + (X_m^2 + X_a^2) G_s^2] \\ &\quad + [(R_s^2 + R_m^2) G_s + (X_m^2 + X_a^2) G_s^2] \\ &\quad + [(R_s^2 + R_m^2) G_s + (X_m^2 + X_a^2) G_s^2] \\ &\quad + [(R_s^2 + R_m^2) G_s + (X_m^2 + X_a^2) G_s^2] \end{aligned}$$

upon using $2R_s R_m G_s / R_s = 2R_m^2 G_s / R_m$ and putting this term with the

factor $-I/R_m$. The factor $R_e \sim (R_m^2 + X^2)G$, is zero as shown in (10.51). Thus we obtain

$$G = -rm + YI <^R> - * . .) * + (* . - x_m)^2] \quad (10.52)$$

This is a very useful relationship in practice, since it determines the noise figure in terms of the minimum value F_m obtained when the optimum source impedance $R_m + jX_m$ is used along with only one additional parameter G . By adjusting the source impedance, both the optimum source impedance and the minimum noise figure can be determined experimentally. Transistor manufacturers will often give the minimum noise figure F_m , the optimum source impedance $R_m + jX_m$, or source reflection coefficient $r_{ss} = (R_m + jX_m - Z_C)/(R_m + jX_m + Z_C)$, and the noise conductance G , or noise resistance R_e . If R_e is given the equation $i?^2 + Xf = R_e/G$ which gave (10.516), can be used to find G ; thus

$$G = \frac{R_e}{Rl + Xi} \quad (10.53)$$

The given data can be used in (10.52) to find the noise figure F .

Noise Figure for Cascaded Stages

Figure 10.26 shows a two-stage microwave amplifier. The power gains of the two stages are G_{P1} and G_{P2} . The input impedance mismatch is M_1 for the first stage and M_2 for the second stage. The equivalent noise sources for stage 2 are designated as $e_n(t)$ and $i_n(t)$. The source impedance for stage 2 is the output impedance of stage 1. The noise figure for stage 1 will be called F_1 , and that for stage 2 will be called F_2 . The noise figure for the system will be called F .

From the definition (10.49) for noise figure, we see that the total input noise to the first stage is $F_1 kTA_f M_1$. At the output of stage 2, this noise

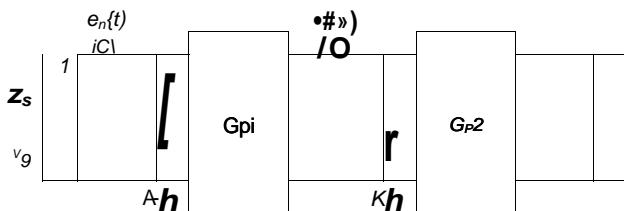


FIGURE 10.26

A two-stage microwave amplifier showing equivalent noise sources.

has been amplified by the factor $G_{P1}G_{P2}$ to give an output noise contribution

$$P_m = G_{P1}f_i^{\wedge}F.kT_l fM,$$

A similar analysis shows that the noise sources $e'_{\text{in}},(t)$ and $i'_{\text{in}},(t)$ produce an input noise power equal to $kT_s/M_2(F_2 - 1)$ to stage 2. Note that we do not include the thermal noise from the source resistance for stage 2, since this comes from the resistive part of the output impedance of stage 1. The latter is an internal noise source for stage 1 and its effects have already been included in the equivalent noise sources for stage 1. The total output noise power in a frequency band Af is readily seen to be

$$P_n^{\wedge} = G_{P1}G_{P2}F_i kTAfM_1 + G_{P2}(F_2 - Y)kTAfM_2$$

where the last term is the amplified noise produced by $e'_{\text{in}},(t)$ and $i'_{\text{in}},(t)$. The output noise can be considered to be the amplified thermal noise in $/?$, multiplied by the two-stage amplifier noise factor F ; thus

$$P_{n,m}^* = G_{P1}f_i P_2 F kT_l fM,$$

By equating these two expressions for output noise power, we get

$$F = \frac{P_{n,m}^*}{P_n^{\wedge}} = \frac{G_{P1}f_i P_2 F kT_l fM}{G_{P1}G_{P2}F_i kTAfM_1 + G_{P2}(F_2 - Y)kTAfM_2} \quad <10,54>$$

This result shows that the second-stage noise does not produce a large change in the noise figure above that of stage 1 alone, since the contribution from the second stage is divided by G_{P2} , provided G_{P1} is of order 10 or more, and the second-stage noise figure is on the order of that of the first stage.

For a cascade of three or more amplifier stages, a similar analysis shows that the overall noise figure is given by

Clearly, successive stages do not degrade the overall noise figure significantly.

The noise in an amplifier can be accounted for by imagining that it comes from the thermal noise in the source resistance by assigning an equivalent noise temperature T_e to the source resistance. Thus we can write

$$\begin{aligned} P_{n,m} &= G_{P1}G_{P2} \cdots M_s kT_s f \\ &= G_{P1}G_{P2} \cdots FM^{\wedge} kTX_f \end{aligned}$$

and hence $T_e = FT$. The excess temperature $T_e - T = (F - l)T$ is called the noise temperature of the amplifier. A low-noise amplifier with a noise figure of 1.4 would have a noise temperature of $0.4 \times 290 = 116$ K.

If all sources of noise in a system, which includes amplifier noise, thermal noise, and radiation noise picked up by the antenna, are regarded

as thermal noise in the source resistance R_s , then the temperature must be assigned to R_s so as to give the same amount of total noise. This is called the system noise temperature.

Constant Noise-Figure Circles

For the purpose of low-noise amplifier design, it is useful to plot constant noise-figure circles on the source reflection-coefficient plane. Since

$$\begin{aligned} Z_s &= Z_c \\ Z_{s+Z_c} &\bullet - Z_m + Z_r \end{aligned}$$

the expression (10.52) for noise figure can be written as

$$F = \frac{G}{R} = \frac{i + r_s}{i - r_s} \frac{i + r}{i - r}$$

$$\begin{aligned} G &= 4ir_s - r_s \\ R &= i(i - r) / i(i - r) \end{aligned}$$

Next we use

$$\begin{aligned} 2ft &= z_s + z_r = 2 \operatorname{Re} \frac{i + r}{i - r} \\ &= 2 \operatorname{Re} \frac{(i - r_s)(i - r_s^*)}{ii - r_s} = 2(1 - irj^2) \\ &= 2 \operatorname{Re} \frac{ii - r_s}{ii - rj} \end{aligned}$$

to get

$$F \sim F_m = AG \frac{K - U}{li - rj^2(i - irj)} \quad (10.56a)$$

where $G = G_s Z_t$ is the normalized value of G . A similar expression can be derived involving $R_N = R_e/Z_c$, namely,

$$F = F_m = \frac{N}{U + rj^2(i - irj)} \quad (10.56*)$$

We now introduce a parameter N , defined as

$$AT = \frac{(fl - irju + rj^2)}{4ft_N} \quad (F_m) ii - r, \quad (10.57)$$

where W corresponds to a chosen value F' for F . From (10.56)

now

obtain

$$N = \frac{r - r_i}{i - ir_j^2} \quad \frac{rr^* - rr^* - rr + rr^*}{i - r, r};$$

which can be written in a form such as (10.1) and identified as describing a circle in the r, r^* plane. The center of the circle is located at

$$r_s = J^{\wedge} - , 0.58_a,$$

and the radius of the circle is given by

$$R_s = \frac{yV + N, (i - ir_m i^2)}{(10.586)}$$

We can plot the circles for various values of N_n determined by (10.57) for chosen values of F_s . Each circle shows the values of Y_s that can be used in order to get a noise figure equal to F_L . For $N_s = 0$ the circle degenerates to a single point at Y_m giving the minimum noise figure F_m . If $F_s = Y_m$ is a stable point, then the amplifier can be designed to give a minimum noise figure F_m . Usually, we can choose the load Z_L so that $F_s = Y_m$ gives a stable design. However, sometimes the input VSWR is too high if we use $F_s = F_m$. In such a case a choice for T_s on a constant noise-figure circle with $F_s > F_m$ would be used in order to obtain a better input VSWR.

Example 10.3 Low-noise amplifier design. A silicon bipolar transistor has the following parameters at 4 GHz:

$$S_u = 0.36X148^\circ \quad S_{11} = 0.11X42^\circ \quad S_{21} = 1.57X27^\circ$$

$$S_{22} = 0.67Z_r - 64^\circ \quad T_m = 0.38^\wedge - 153^\circ \quad R_N = 0.4$$

$$F_m = 1.905 \text{ (2.8 dB)}$$

By using (10.25a) we find $K = 1.2421 > 1$; so the transistor is an absolutely stable device. From (10.19) we find that the maximum power gain is 7.2123 for which the maximum normalized gain is $7.2123/IS_{21}I = 2.926$. Since the device is absolutely stable, we can use conjugate impedance matching. From (10.17) we find that this requires $F_s = -0.53287 j0.40911$ and $T_L = 0.34159 + y 0.74723$. With conjugate impedance matching, the input and output VSWR equals unity. If we use the above value of T in (10.566), we find that the noise figure is 2.49 or 3.96 dB. This is 1.16 dB greater than the minimum value. It is desirable to have a lower noise figure and yet not sacrifice any gain which already is on the low side. We can accept an increase in VSWR at both the input and the output of the amplifier. A design using $g_p = 2.9$, which is very close to the maximum value, will be attempted. For an aid in the design process, the two constant normalized power-gain circles $g_p = 2.9$ and $g_p = 2.3$

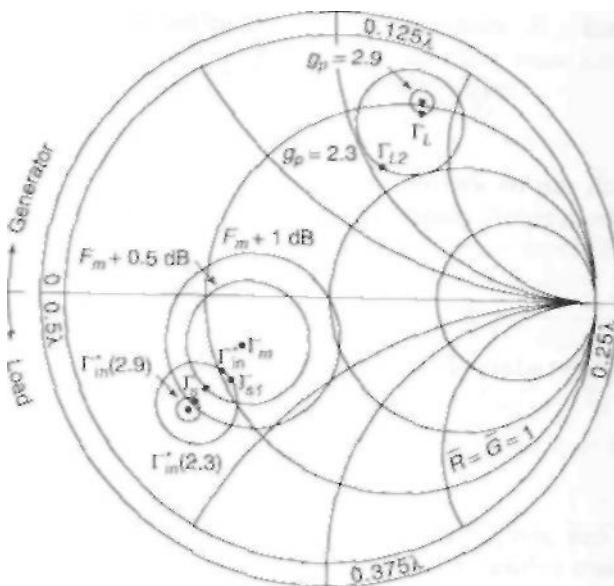


FIGURE 10.27

Constant normalized power-gain circles, constant noise-figure circles, and J^* the low-noise amplifier design example.

m

have been constructed and are shown in Fig. 10.27. The centers and radii of these circles were found using (10.28) and are

$$r_{Lg} = 0.34026 + j0.7443 \quad R_{Lg} = 0.03574 \quad \text{for } g_p = 2.9$$

$$r'_{Lg} = 0.30532 + j0.6679 \quad R_{Lg} = 0.2052 \quad \text{for } g_p = 2.3$$

As explained in Example 10.2 the values of V_L on a $g_p \sim$ constant circle generate a set of values for r_{in} and hence a circle of 1^*_{in} values. The 1^*_{in} circles for $g_p = 2.9$ and 2.3 are also shown in Fig. 10.27. The centers and radii of these two circles [the equations for these circles are given in Sec. 10.10 a (10.59a) and (10.596)] are

$$1^*_{in} = -0.53198 - j0.40843 \quad R_{in} = 0.03029 \quad \text{for } g_p = 2.9$$

$$1^*_{in} = -0.51225 - j0.39328 \quad R_{in} = 0.148 \quad \text{for } g_p = 2.3$$

Also shown in Fig. 10.27 are the constant noise-figure circles for $F = 2$ (3.3 dB) and 2.4 (3.8 dB), which correspond to 0.5 dB and 1 dB greater t^* than F_m . The centers and radii for these circles are found using (10.57) and are

$$r_{nf} = -0.31706 - j0.16155 \quad R_r = 0.23444 \quad \text{for } F_{in} + 0.5 \text{ dB}$$

$$r_{nf} = -0.2959 - j0.1508 \quad R_f = 0.3317 \quad \text{for } F_m + 1 \text{ dB}$$

If we want an input VSWR of unity, then we must choose f_s to J^* circle. The dot shown inside the smallest f^* circle is the value of J in

conjugate impedance matching is used. The figure clearly shows that if we choose T_s to lie on this point, the noise figure will be more than 1 dB greater than F_m . Our earlier calculation gave a value of 1.16 dB greater. If we are willing to relax the gain requirement to $g_p = 2.3$, then the figure shows that we can obtain a unity input VSWR and a noise figure somewhat better than $F_m + 0.5$ dB by choosing F , as the point V_{si} shown in Fig. 10.27. If we choose $r_t = i'_m$ and $g_p = 2.3$, then the best input VSWR is obtained by choosing a l'_L that will produce a Tf_n that lies as close as possible to $\parallel = Y_m$. This is the point $rj = -0.3386 - j'0.1725$ shown in Fig. 10.27. The corresponding value of V_L is $l'_L 2 = 0.1818 + ./0.504$ and is also shown in the figure. These choices result in a minimum noise figure, an input VSWR equal to 1.392, and an output VSWR equal to 2.069. The noise figure is a minimum value and the VSWRs are acceptable but the gain is only $2.3 \times |S_{21}|f = 5.67$.

If we insist on having $g_p = 2.9$ which gives a power gain of 7.15 and we also want a good input VSWR, then we have to accept some increase in noise figure. If we accept a noise figure 0.5 dB greater than F_m , then for the best input VSWR we choose r_s to lie on the $F = F_m + 0.5$ dB constant noise-figure circle and as close as possible to the I^* circle for $g_p = 2.9$. This point is $T_s = -0.471 - ./0.338$ and is shown in Fig. 10.27 as the point I_{ss} . The required value of T_L that will place I^* as close as possible to \parallel , is $I_s = 0.3217 I - yO.7137$ and is shown in Fig. 10.27. These choices result in an input VSWR equal to 1.22 and an output VSWR equal to 1.065. This last design can be considered to be acceptable.

The input and output impedance-matching networks using parallel open-circuited transmission-line stubs can be designed using the method described in Sec. 5-6 (i.e., (5.29a) and (5.296>i). Since we use parallel stubs each stub is required to produce only one-half of the susceptance given by (5.29a). From the chosen values of l' , and Y_L given above, we readily find that the required normalized source and load admittances are

$$Y_s = 1.683 + j1.716 \quad Y_L = 0.1715 - ./0.6327$$

The matching networks along with the stub positions and lengths are shown in Fig. 10.28. This figure does not show the dc bias circuit which is also required.

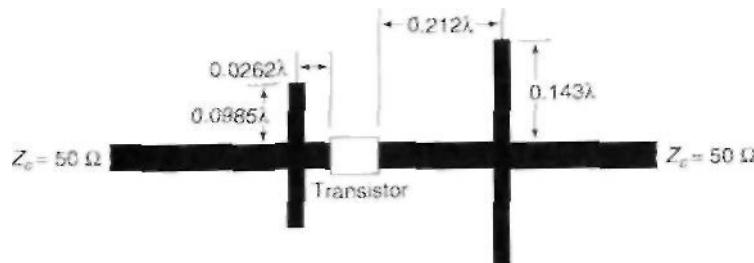


FIGURE 10.28
Matching networks for the low-noise amplifier

After all of the calculations for the above amplifier design determination of the values of T_L and η that give the best input and output VSWR for the chosen g_p and F were carried out using the computer program MICROAMP. The matching networks were also designed using this PTO^

10.10 CONSTANT MISMATCH CIRCLES

In Examples 10.2 and 10.3 we used the f_f circle as an aid in the design of microwave amplifier with a low input VSWR. The parameters that describe the T_{j*} circle will be derived here. For an amplifier terminated in a load having a load reflection coefficient r_L , the input reflection coefficient r_{in} is given by

$$r_{in} = \frac{a^* r_L^* - s^*}{1 - S_{22} V_L}$$

The complex conjugate of this equation is

$$r^*_{in} = \frac{a^* r_L^* - s^*}{1 + S_{22} V_L}$$

This is a bilinear transformation, so that all values of T_t that lie on a constant g_p circle will map into a circle of r_{in} values. The center r_{Lj} and radius R_{Lg} for a constant normalized power-gain circle are given by (10.28). By using these circle parameters, we can obtain the corresponding parameters for the circle of T_{jn} and circle of P_j^* values as described in Sec. 10.7. The center f_f for the circle of f_f values is given by

$$f_f = \sqrt{\frac{1 - r_{in}}{1 + r_{in}}} = \sqrt{\frac{1 - r_{Lj}}{1 + r_{Lj}}} = \sqrt{\frac{1 - R_{Lg}}{1 + R_{Lg}}} = \sqrt{\frac{1 - g_p}{1 + g_p}}$$

s is given by

$$Rir = \frac{1 \ll a SJR_g^*}{K^* |S_{22}| - |S_{22} r_{Lg} - s|^2} \quad (10.596)$$

In the design of a microwave amplifier, the choice for the reflection coefficient is constrained by the requirements that are to obtain a stable amplifier with a low noise figure. Unity input VSWR can be obtained if T_s can be chosen to lie on the f_f circle for the P_j^* circle. If the constraints do not allow this choice or, as sometimes the f_f values inside the Smith chart are unstable values for I_s . . should be located as close as possible to the f_f circle to obtain the input VSWR. When the choice for f , has been made, then, if f_s lies

T_s^* circle, we will require $F_{in} = F^*$ and we can then find the required value of F_L using

If T_s cannot be placed on the r_n^* circle, then the best value for T_n^* is the one that is closest to r_s but on the specified r_L circle. The complex conjugate value of this should be used for r_n in (10.60) to find T_L .

For an unstable device a part of the $g_p = \text{constant}$ circle will lie outside the Smith chart boundary. The values of λ_L on the $g_p = \text{constant}$ circle outside the Smith chart produce values of λ_n^* that lie outside the Smith chart. As explained in Example 10.2 the F^* circles have two invariant points for an unstable device. These invariant points coincide with the two points at which the source stability circle intersects the Smith chart boundary as illustrated in Fig. 10.21. The proof that these points are invariant points is as follows: For an unstable device we have shown that the $g_p = \text{constant}$ circles have two invariant points that coincide with the two points at which the load stability circle intersects the Smith chart boundary (see Sec. 10.7). Since F^* is given by

$$\begin{aligned} & A^* F_f - S_f, \\ & \text{in } s_2^* r^* - 1 \end{aligned}$$

it is clear that these two particular values of F , which we will call V_L and r_{12} , that are invariant points for the $g_p = \text{constant}$ circles will map into two fixed values for F_f that are common to all λ_n^* circles. Thus the r_n^* circles have two invariant points. Since F_{in} and r_{12} also lie on the load stability circle that makes $|r_{in}| = 1$, the two points, which we will label as r_{in} and r_{12} , must lie on the Smith chart boundary. Thus the two invariant points for the F^* circles lie on the Smith chart boundary. It remains to be shown that these points coincide with the two points at which the source stability circle intersects the Smith chart boundary.

The source stability circle is the circle of F_s values that make $|F_{011}| = 1$, where

$$T = \frac{A r_s - S_{22}}{s_s - 1} \quad (10.61)$$

We can rewrite (10.60) in the following form:

$$1 - \frac{1}{A \frac{r_s}{r_{in}} - S_{22}} \quad (10.62)$$

which is a bilinear transformation of the same form as in (10.61). Let us choose $r_{s1} = F^*$. Since $|V_{12}| = 1$ we have $r_{ri} = 1/F_{in}$. From (10.62),

$1/r_{in}$, maps into the point $L \wedge'_{L1}$ for which $|r_{t1}| = 1$. The bilinear transformation in (10.61) is the same, so it follows that $F_{SI} = 1/f$. The point r_{out} , for which $|r_{t1}| = 1$, is a point on the source stability circle. For the same reasons r_{s2} is a point on the source stability circle. Consequently, $F_{SI} = V_{si}$ and $T_{s2} = T_{s2}$ must be on the source stability circle. Hence the invariant points for the $|f|$ circles coincide with the points at which the source stability circle intersects the Smith chart boundary.

Constant Input Mismatch Circle

If we want to design an amplifier with a specified input VSWR, then if the load reflection coefficient V_L has been chosen, there will be a circle of f values that can be used which will produce the specified input VSWR. Let the required input VSWR be $VSWR_r$. The reflection-coefficient magnitude is given by

$$\frac{VSWR_r - 1}{P} = \frac{VSWR_j + 1}{VSWR_j - 1}$$

and the input impedance mismatch M_x is given by $M_x = 1 - p^2$. From (10.12a) the input mismatch is given by

$$4R_s R_{ID} \left(\frac{1 - i r_{in} |f|^2}{1 + i r_{in} |f|^2} \right) \quad (10.63)$$

Let f_r be the chosen value of V_L . From (10.60) we can calculate the corresponding value of f_r , say r_{in} . By using this particular value of r_{in} in (10.63), we can express (10.63) in the form of an equation for a circle in the $|f|$ plane. The center of this circle is located at

$$sM \quad | - o - iiWJ^2 \quad (10.64a)$$

and the radius of the circle is given by

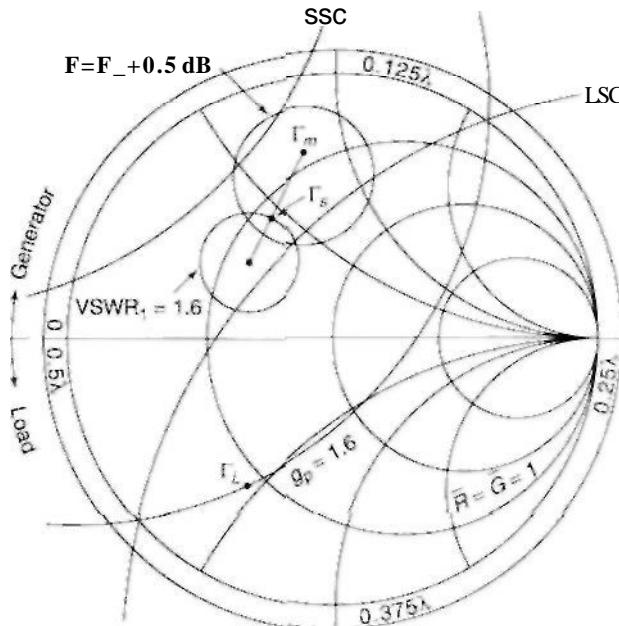
$$R_M = \frac{|r - nr(i - iy^*)|}{i - (i - M_r)ir;_n r} \quad (10.64b)$$

The following example will illustrate the application of the constant input mismatch circle in amplifier design.

Example 10.4 Application of constant input mismatch circle in amplifier design. An FET has the following parameters:

$$\begin{aligned} S_M &= 0.8Z - 140^\circ & S_{22} &= 0.2Z30^\circ & S_{21} &= 2.8Z60^\circ \\ S_{22} &= 0.2Z1150^\circ & T_{ss} &= 0.7Z100'' & R_N &= 0.4 \\ F_{ss} &= 1.5 \text{ (1.76 dB)} \end{aligned}$$

We want to design an amplifier with low noise and an input VSWR

**FIGURE 10.29**

Constant gain, constant noise-figure, and constant input impedance-mismatch circles used for the amplifier design in Example 10.4.

1.6. The required value of the input impedance mismatch is found to be $M_{in} = 0.94674$.

In Fig. 10.29 we have plotted the $\text{G}_{\text{ref}} = 1.6$ constant normalized power-gain circle, the load stability circle, the source stability circle, and the optimum value Γ_{opt} for Γ_s to obtain a noise figure equal to the minimum value F_m . The $F = 1.683$ constant noise-figure circle is also plotted. This noise figure is 0.5 dB greater than F_m . For this example we have chosen $\Gamma_L = 0.307 \angle 0.55^\circ$, which is a point on the $\text{G}_{\text{ref}} = 1.6$ constant gain circle and not too close to the load stability circle, for this choice of $\Gamma_L = \Gamma_{in}$ the input reflection coefficient $r_{in} = -0.3302 \angle 0.3062^\circ$. The input mismatch circle that will give $\text{VSWR}_{in} = 1.6$ has a center and radius given by

$$[\bullet, \#] = 0.316 (\sqrt{0.293}) \quad \text{ft.}, \quad = 0.186$$

and is also plotted in Fig. 10.29. All values of V_H on this circle will give an input VSWR equal to 1.6. In order to obtain the best possible noise figure, we should choose Γ_L to lie on this circle and on the line that joins the center of this circle to the optimum point Γ_{opt} . The best choice for Γ_L is shown in Fig. 10.29 and lies inside the $F = F_m + 0.5$ dB constant noise-figure circle, so that the input VSWR requirement can be met with a noise figure somewhat less than 1.683. The power gain obtained is $1.61 S_{21}^2 = 12.54$, which is quite close to the "Figure of Merit" gain of 14.

Output Impedance-Mismatch Circle

If it is required to design an amplifier with a specified output VSWR for each chosen value of $I_s = \sqrt{s}$, we can plot a circle of r_i values that ensure that the specified output VSWR is obtained. The equation of the output impedance-mismatch circle are of the same form as (10.641) if the center of the circle is located at

$$\mathbf{0}^* = \frac{MAKucf}{i - (i - A f_2) i r^2} \quad (10.65a)$$

and the radius of the circle is given by

$$CM = \frac{-/i - 2(i - ir;_u, j^2)}{1 - (i - M_2)ir;_u i^2} \quad (10.65b)$$

where

$$M_2 = 1 - \frac{(VSWR_2 - 1)^2}{(VSWR_2 + 1)}$$

$$= \frac{AS \cdot S_{i2}}{S_n r_i - I}$$

In the design of a two-stage amplifier, the design of stage 1 leads to a specified output impedance mismatch for stage 1. Consequently, the design of stage 2 is constrained by the requirement that the input mismatch to stage 2 be equal to the output mismatch of stage 1, since the impedance mismatch is conserved in the lossless matching network that is used to couple the first and second amplifier stages. For this reason the constant impedance-mismatch circles described above are useful aids in the design of a two-stage amplifier. The application of the constant impedance-mismatch circles in two-stage amplifier design is described more completely in the next section.

10.11 MICROWAVE AMPLIFIER DESIGN

In this section we will present a design strategy for designing narrowband one- and two-stage amplifiers. The first stage can be designed for a low noise figure. The method to be described can be used with both stable and potentially unstable transistors. It is assumed that the scattering-matrix parameters S_{ij} , the optimum source reflection coefficient Y_m for minimum noise, the minimum noise figure F_{min} , and the normalized noise resistance R_N or noise conductance G , are all known. The design specifications assumed to be a power gain greater than some minimum value, a noise figure no greater than a specified maximum value, and input and output VSWRs that do not exceed specified maximum values.

There is no unique design for an amplifier that meets the design specifications. Also, there is no unique method for carrying out the design. The method described in this section works quite well for achieving a satisfactory design, but many other systematic approaches could also be developed. In general, we have to examine a range of possible load and source reflection coefficients in order to obtain an optimum design. It would be very tedious to carry out the required optimization procedures using hand calculations. Consequently, in practice a suitable computer program is used. The design strategy that is described in this section has been implemented as the computer program MICROAMP.

Stage Amplifier Design

The first stage of a multistage amplifier or a single-stage amplifier is normally designed for a minimum noise figure, maximum power gain, and a chosen maximum input and output VSWR. In a multistage amplifier the output VSWR of the first stage is usually not a critical parameter. When the constraint on the output VSWR is relaxed, there is a greater degree of freedom available that makes it easier to achieve the other design requirements. The design of the second stage is also easier to carry out when the first-stage output VSWR is relatively large. We will describe the design of the first stage as a series of steps or procedures to be carried out.

1. The first step is to evaluate the stability parameter K given by (10.25a), i.e.,

$$\%|S_u S_n|$$

and also to check if (10.256) to (10.257') are satisfied. When these conditions hold and $K > 1$, the transistor is absolutely stable and steps 2 to 4 should be followed. If $K < 1$ the device is potentially unstable and steps 5 and 6 should be followed.

2. For $K > 1$ conjugate impedance matching can be used. The required values for the source and load reflection coefficients are given by (10.17). Let r_s^* be the solution for Γ_s given by (10.17a). From this value of Γ_s , the noise figure F can be calculated by using (10.56). The power gain with conjugate impedance matching is given by

$$G_p = G_{l>,max} = (K^{-\frac{1}{2}} - l)^{-1}$$

If the noise figure is acceptable, then the design is finished except for the design of the input and output matching networks. In practice, it usually turns out that the noise figure obtained using conjugate impedance matching is not satisfactory. In order to obtain a better noise figure, it will be

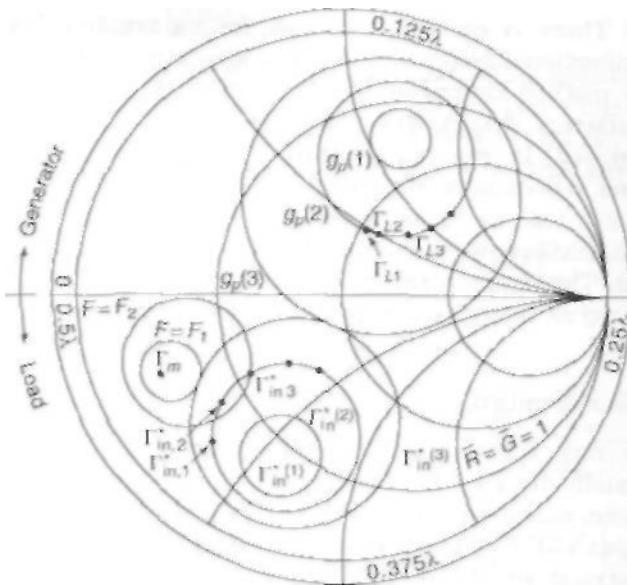


FIGURE 10.30
Constant gain, Γ^* , and constant noise-figure circles used for amplifier design.

necessary to design for a lower gain and *some mismatch at the input and output ports*, as described in the following steps.

3. In order to visualize the design procedure, we have plotted three constant normalized power-gain circles, the corresponding three V_n^* circles, and two constant noise-figure circles for a hypothetical device in Fig. 10.30.

We first try to obtain a satisfactory design using $\Gamma_s^* = T_m$ for minimum noise and a chosen normalized power gain. For example, we can choose $\xi_P = \xi_i/2$ in Fig. 10.30. We now construct an objective function

$$OF = W_1 M_1 + W_2 M_2$$

where M_1 is the input mismatch and M_2 is the output mismatch and W_1 and W_2 are weights that can be chosen to place different levels of importance on the input and output VSWRs. A value of V_L , say r_t , is chosen on the $g(2)$ constant gain circle. From this value of f_L we can calculate \dots , using

$$n_{11} = \frac{M_1 - S_u}{S_{22} S_{11}}$$

We can also calculate f_{out} using $T_s = F_{11}$; thus

$$r_{11} = \frac{AT_{11} - S_{22}}{S_{11} R_m - t}$$

From these we calculate M_1 and M_2 using

$$M_1 = \frac{(i - ir_{in}i^2)(i - ir_j i^2)}{li - r_{in}r_{out}i^2}$$

$$M_2 = \frac{(i - ir_{out}i^2)(i - ir_L i^2)}{li - i; u^2 i_A}$$

The objective function for this value of P_i is now evaluated. The best input and output VSWRs are obtained by maximizing the objective function. Thus we search the $g_p(2)$ circle for the value of Y_L that maximizes OF. This search is carried out by incrementing λ in the direction that increases OF, that is, r_L is set equal to $r_{1/2}, P_{1/2}, \dots$, as shown in Fig. 10.30. If satisfactory values for the input VSWRj given by

$$\text{VSWR}_j = \frac{1 + fl - M_1}{l - fl - M_1}$$

and the output VSWR₂ given by

$$\text{VSWR}_2 = \frac{1 + / l - M_2}{l - / l - M_2}$$

are obtained, then the design process is terminated.

If we obtain a good value for the output VSWR₂ but an unacceptable value for the input VSWR_j, then we can set the weight $W_i = 0$ and $W_l = 1$ so as to place all of the emphasis on achieving a good input VSWR_j. With all of the emphasis placed on M_l , the optimum value of T_L will be that value which produces a F''_n that is as close as possible to r_m , since this produces the best input impedance match. However, M_l and $M_{>l}$ have the same dependence on r_L so the largest values of M_l and $M_{>l}$ occur for the same value of T_L , so changing the weights will not improve the results. It will then be necessary to search for an optimum value of T_i on a lower constant gain curve, say $g_i(i)$. The corresponding r^* circle is larger: so clearly HJ, can be brought closer to r_m , thereby improving the input VSWR_j.

The above process is repeated until the lowest acceptable constant gain circle has been searched. If this does not result in acceptable input and output VSWRs, the only alternative left is to accept a noise figure greater than F_m . In this case it is helpful to compile a table of best possible input VSWR_j values for a given noise figure and normalized power gain, as described in step 4.

4. Figure 10.31 shows the $F = F$, constant noise-figure circle and the $[^2(2)$ circle for $g_p = g_p(2)$. If we are designing the amplifier for this noise figure and power gain, then the optimum choice for Y^* and T^* that will result in the best input VSWR_j is the value of T_s on the $F = F_x$ circle and the value of F_f on the $V_{ad}^*(2)$ circle that are as close together as possible.

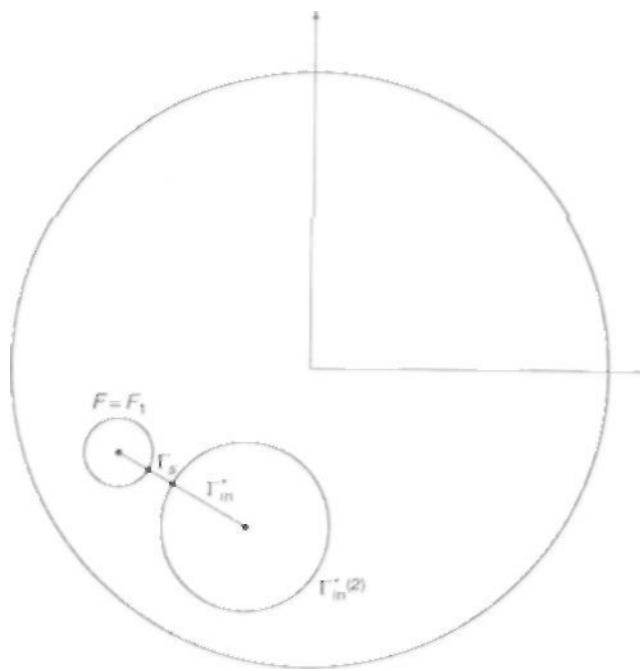


FIGURE 10.31

A constant noise-figure circle and a F^* circle and the optimum choice for V_x and T^* , that will give the lowest input VSWR.

These points lie on the line that joins the center of the $F = F_j$ circle to the center of the $F^*(2)$ circle as shown in Fig. 10.31.

The center for the noise-figure circle is \vec{r}_f given by (10.58a) and the radius R_f of the circle is given by (10.586). The center $V_{n,0}^*$ and radius r_m^* for the F_j^* circle are given by (10.59). The vector from the center of the $F = F_j$ circle to the center of the $F_{n,0}^*(2)$ circle is given by

$$\mathbf{r} = \mathbf{r}_{n,0}^* - i\mathbf{v}$$

A unit vector pointing from T_s to $f_{n,0}^*$ is $\mathbf{r}/|\mathbf{r}|$. The point T_s lies a distance R_f from $V_{n,0}^*$ and in the direction of \mathbf{r} ; hence

$$\mathbf{r}^* = \mathbf{r}_{n,0}^* + R_f \frac{\mathbf{r}}{|\mathbf{r}|}$$

The optimum value of f^* is similarly given by

$$T_{in}^* = -E_{in} = f_i$$

Thus we can calculate the optimum choice of T_s and T^* and from these

evaluate the input mismatch M_i and VSWR,. The value of V_L that produces r_{jjj} is given by < 10.60 . The above calculation can be repeated for various choices of g_p and F and allows us to compile a table of best VSWR, values as a function of g and noise figure F . By consulting such a table we can easily see the tradeoffs between power gain, noise figure, and input VSWR,. We now choose f_e for the lowest noise figure consistent with the lowest value of acceptable power gain and the largest acceptable value of input VSWR,. The next step is to repeat the optimization procedure described in step 3 by searching the identified $g_p = \text{constant gain circle}$ so as to optimize the objective function. The maximum value of the objective function might not correspond to the optimum choice for V^* that maximizes $M|$. Also, the output VSWR₂ might be higher than specified. If this is the case, then either the design requirements have to be relaxed or a different transistor must be used.

5. If the transistor is potentially unstable, then we cannot carry out a design with conjugate impedance matching. For this case a design for minimum noise using $\text{l}' = \text{l}_{\text{m}}$, should be explored first. This requires that we determine the source stability circle and check that l_{m} , lies in a stable region of the Smith chart and not too close to the boundary of the source stability circle. When T_m is a stable value, the design procedure is the same as described in step 3, i.e., a chosen $g_p = \text{constant gain circle}$ is searched for the best value of V , that will maximize the objective function OF given in step 3. The "Figure of Merit" gain is $|S_{21}/S_{12}|$ with a normalized value $|S_{12}S_{21}|$. It is good practice to design an amplifier for a normalized power gain that does not exceed this value. Thus constant power-gain circles with $S_p = |S_{a1}S_{12}|$ and smaller are searched. If acceptable values of g_p , VSWR,, and VSWR,, are obtained, then the design process is terminated. The resultant amplifier will have a minimum noise figure F_m . If a satisfactory design cannot be obtained using $T_m = f'_{\text{m}}$, or if l_{m} , is an unstable value, then the procedure outlined in step 6 should be followed. If f_{m} , is an unstable value, it would be advisable to use another transistor.

6. In order to minimize the amount of searching for the optimum value of V , subject to the constraints on g_p , F , VSWR,, and VSWR₂, it is again helpful to compile a table of best VSWR, values as a function of g_p and F . By consulting such a table we can determine if the design objectives can be met and/or the best power gain, noise figure, and input VSWR, that can be obtained using the chosen transistor. When the I'^* values inside the Smith chart represent stable values of I' ,, then the input VSWR, for the optimum choice of T^*_n and f_s on a chosen $F = \text{constant circle}$ are calculated the same way as for a stable device. The procedure is described in step 4. The optimization of the input VSWR, and output VSWR,, for the chosen normalized power gain and noise figure can be carried out in the same way as described in step 4.

A bipolar transistor used in a common base configuration or an FET used in a common gate configuration often have $|S_{11}| > 1$ and $|S_{22}I| > 1$.

When this latter set of conditions holds true, it is not possible to design an amplifier with a low input or output VSWR. The reason is that stable if*? ^ of T_s produce values of V_n* that lie in the unstable region of the r_L plane so we cannot choose T_s equal to 1*. Similarly, stable values of r produce r* values that lie in the *unstable region of the r_L plane* so the output noise cannot be matched. Since the common base and common gate connections usually have poorer noise performance as well, the common base or common gate circuits are not used in low noise amplifiers. In Example 10.5 we illustrate the impossibility of matching the input port of a common base amplifier.

Example 10.5 Common base amplifier. A bipolar transistor in the common base connection has the following parameters at 5 GHz:

$$\begin{array}{lll} S_u = 1.3^{\wedge}140^\circ & S_{12} = 0.2^{\wedge}130^\circ & S_{21} = 2^{\wedge}-85^\circ \\ S_{22} = 1.1f_iZ - 50^\circ & r_{ss} = 0.7^{\wedge}135^\circ & F_m = 2.5 \quad R_N = 0.4 \end{array}$$

For this transistor K = -0.579 so it is potentially unstable. In Fig. 10.32 we have plotted the load stability circle, the source stability circle, the g_p = 1.5

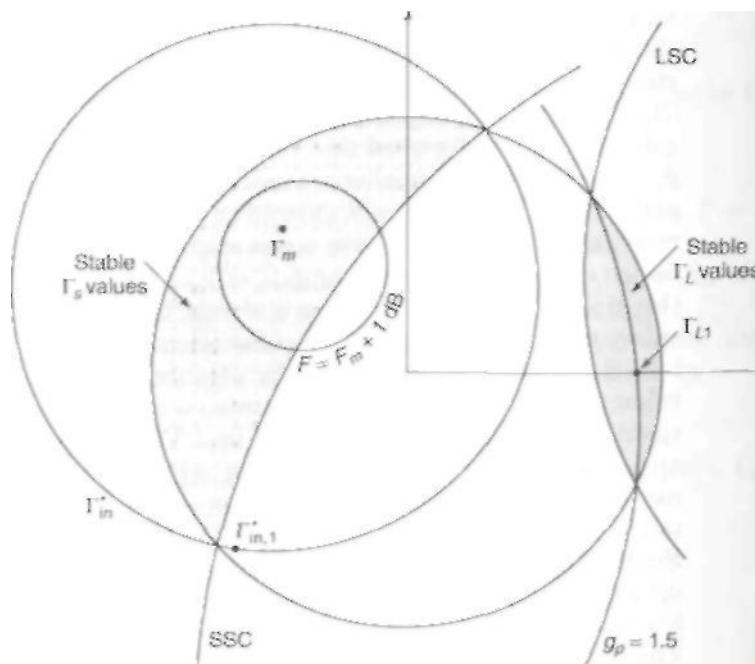


FIGURE 10.32

Load stability circle, source stability circle, $g_p = 1.5$ circle and corresponding $F = F_{ss} + 1$ dB constant noise-figure circle for the bipolar transistor, in a common base configuration as used in Example 10.5.

circle and corresponding $!^*$ circle, and the $F = F_{m,} + 1 \text{ dB} = 3.147$ circle. Since $|S_{11}| > 1$ and $|S_{22}| > 1$, the origin is an unstable point for both $\|\cdot\|$ and $!/\cdot$. The regions of the Smith chart where stable values of T_v and V_s are located are shown cross-hatched. The point r_L on the $g_p = 1.5$ circle maps into the point l^* , which lies in the unstable part of the T_v plane. All values of l'_L in the stable region map into T^* values that lie in the unstable part of the Y_s plane. Hence we cannot choose $l''_v = l^*$ and consequently the input port cannot be matched,

If we try to design a low-noise amplifier by choosing $I_v = [^*]$, and with a normalized power gain of 2, the resultant input VSWR₁ = 15.42 and the output VSWR₂ = 19.03. If we reduce the gain requirement G_p to 5, for which $g_p = 1.25$, we obtain VSWR₁ = 12.76 and VSWR₂ = 26.56 for a design with minimum noise. There is a small improvement in the input VSWR, but the output VSWR₂ is increased.

From Fig. 10.32 it is quite clear that any value of r_{in} on the $F = F_m + 1 \text{ dB}$ noise-figure circle and in the stable region will still be far away from all I^* values inside the Smith chart. Thus a low-noise **amplifier** with acceptable gain and input and output VSWRs cannot be designed using the above transistor in a common base circuit.

Example 10.6 Low-noise amplifier design. A GaAs FET has the following parameters at 10 GHz:

$$S_u = 0.73^{+42^\circ} \quad S_{12} = 0.2A - 58^\circ \quad S_{21} = 1.52i - 66''$$

$$S_{22} = 0.5^{+34^\circ} \quad r_{in} = 0.52Z_0 - 70^\circ \quad F_{m,} \Rightarrow 1.25 \quad A_{IV} = 0.75$$

We want to use this device in a low-noise amplifier design which meets the following specifications:

Noise figure $/^\bullet < 1.5$

Input VSWR₁ < 1.5

Output VSWR₂ < 1.5

Power gain G_p , as large as possible

For this device $K = 1.071$ so that the FET is absolutely stable. For a conjugate-impedance-matched design, we get $G_p = G_{pmm} = 5.222$. VSWR₁ = VSWR₂ = 1, and $F^* = 1.7$. The noise figure is too large; so we must consider a design that is not matched at the input and output. The power gain which can be obtained is no greater than 5.222; so we do not want to sacrifice much power gain for an improved noise figure. The best VSWR, that can be achieved using $g_p = 2.25$, which gives $G_p = 5.1984$, for $F = F_{m,} - i 0.5 \text{ dB}$ is 1.3644 and for $F = F_{m,} + 1 \text{ dB}$ it is 1.026. This shows that the design objectives can be met by allowing a noise figure equal to $F_m + 0.5 \text{ dB} = 1.4$. By searching the $g_p = 2.25$ circle for the optimum value of I_v , and using the optimum value of I''_v , which was determined so as to lie on the $F = F_m + 0.5 \text{ dB}$ noise circle and give VSWR₁ = 1.3644, we obtain $G_p = 5.1984$, VSWR₁ = 1.3643, VSWR₂ = 1.1179, and $F = 1.4$. The required values of T_v and V_s are $I'_v = 0.389 - /'0.534$ and $IV = -0.0028 - /0.32S$. This design meets all of the specifications.

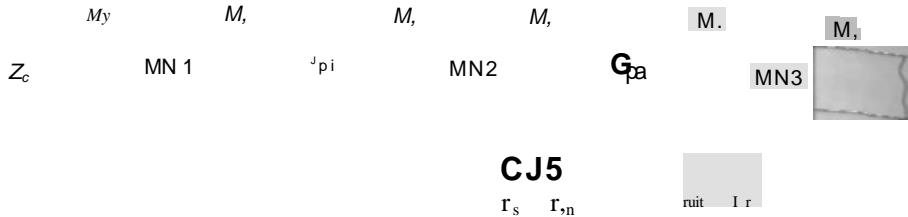


FIGURE 10.33
A block diagram of a two-stage amplifier.

Design of Second Stage for a Two-Stage Amplifier

The design specifications for the second stage of a two-stage amplifier emphasizes power gain and output VSWR. The noise figure of stage 2 is not very critical since the noise contribution of the second stage is reduced by the power gain of the first stage as shown by (10.54). However, at the higher microwave frequencies, the power gain of the first stage is often not very large, so that some consideration of the noise figure of stage 2 is necessary. If we assume that the matching network used between the output of stage 1 and the input to stage 2 is a lossless network, then, since the impedance mismatch is constant throughout a chain of lossless networks, the impedance mismatch at the input to stage 2 is the same as that at the output of stage 1. This places a constraint on the design of stage 2, namely, that the input impedance mismatch must equal M_2 , where M_2 is the output mismatch for stage 1 and was determined in the design of stage 1. In order that the interstage matching network be physically realizable, this constraint cannot be violated.

In Fig. 10.33 we show a block diagram for a two-stage amplifier. The source reflection coefficient, input reflection coefficient, output reflection coefficient, and load reflection coefficient for stage 2 are identified by a superscript prime. The output impedance mismatch of stage 2 is M_3 and the corresponding output VSWR will be called $VSWR_3$. It is *assumed* that the same type of transistor that was used in stage 1 is also used in stage 2.

The design of stage 2 will be based on the optimization of the noise figure and output VSWR or mismatch M_3 for a chosen power gain $g_p \cdot T^{hf}$. Optimization is carried out subject to the constraint that the input impedance mismatch equals M_2 . The following objective function is used for the design of stage 2:

$$OF = W_1 M_3 + W_2 - \frac{F}{F_m}$$

where W_1 and W_2 are suitable weights. We choose F_m/F for the second term, since this quantity is of the same order of magnitude as M_3 and becomes larger for smaller values of F . Thus our goal is to find the

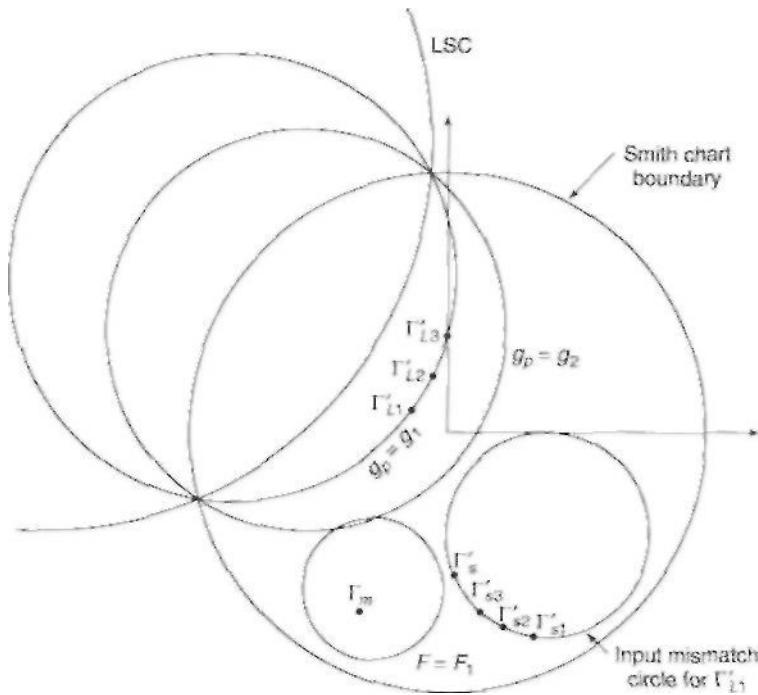


FIGURE 10.34

A constant power-gain circle and a constant input impedance-mismatch circle used in the design of the second stage of a two-stage amplifier.

value of r_L on a chosen $g_p = \text{constant}$ gain circle so as to maximize OF. The design procedure is the same for both stable and unstable devices.

A visualization of the optimization process can be obtained by referring to Fig. 10.34. In this figure we show the load stability circle, two constant power-gain circles, and a constant noise-figure circle for a hypothetical transistor. Let us assume that we will choose $g_p = g_s$. We then pick an initial value of V_s , say T_{s1} , on the g_s circle. From this value of V_s we can calculate I_{T_s} using

$$\frac{nr_{II} - S_u}{1 + \frac{Q}{r_s} - \frac{1}{r_s}}$$

The input mismatch M_2 is given by

$$\frac{(i - irj')(i - ir)/}{|i - r_s|^2}$$

Since M_2 is fixed by the stage 1 design, this equation determines a circle of

V_s values that can be used. The center and radius of this input circle are given by (10.64) and are

$$KM \sim M_2 d 7 \text{nf}$$

$$R'SM = i - (i - M_2)ir; j^2$$

Each value of T_s such as 1^\wedge , r_{12} , etc., shown in Fig. 10.34 enables us to calculate a corresponding noise figure using (10.56). Also for each IT we can calculate an output reflection coefficient I''_L , .. From r_{ul} , and the chosen value Y'_L for V_L , we can evaluate the output mismatch Af_3 . Thus we can evaluate the objective function. Our procedure is now to search the input mismatch circle for the value of l''_s that maximizes the objective function. The maximum value is recorded. We now increment Y''_L to a new value $Y',..$. This results in a new input mismatch circle which is searched for the value of l''_a that maximizes OF. This value of OF is compared with the previous one, and if it is greater than l''_L is incremented to a new value $T\ell_3$ in the same direction on the $g_p = g_x$ circle. If the second value of OF is smaller than the first one, then l''_L is incremented in the opposite direction. The search process is continued until the optimum values of V_L and T_a are found. The optimization routine would be *very time consuming to carry out* with hand calculations but can be done very quickly on a computer. The optimization carried out on the input mismatch circle will result in a choice for H that lies as close as possible to the optimum point l_m if the weight W , is set equal to zero (see Fig. 10.34). When W_L is not zero, TJ will generally deviate from this point in order to get a lower output VSWR₃.

If satisfactory values of noise figure and output VSWR₃ are not obtained from a value of Y'_L on the chosen $g_p = \text{constant}$ circle, then the procedure outlined above must be repeated on a constant gain circle having a lower gain. If, after searching the lowest acceptable power gain circle, an acceptable design is not obtained, then the design specifications will have be relaxed, a different transistor used, or a third stage added. A third stag* can be designed using the same approach as used for the second stage-

Some of the V_s values on an input mismatch circle may he "/* *?" unstable region of the Smith chart. If Y^\wedge is an unstable point, the "te" r produces an output reflection coefficient r_{ul} , with a magnitude 8^\wedge than unity. When $ir_{ul} > 1$ the output mismatch M_3 and output VbW $\wedge j$ will be negative. Consequently, M_3 will contribute a negative quanti / the objective function. Since the objective function is being maxim' ^ unstable values of f^\wedge are not selected since they tend to mininuz objective function.

When the design of the second stage has been completed, the two-stage amplifier will have a power gain

$$G_p = G_{pfp2}$$

and a noise figure

$$F = F_X + (F_a - 1) \frac{M_2}{M_{fp}^2}$$

The interstage matching network must transform the output admittance Y_{IM1} of stage 1 into the required source admittance F_J for stage 2. At the same time it must transform the input admittance V_n' of stage 2 into the required load admittance Y_L for stage 1. This matching network is physically realizable because of the constraint that was placed on the stage 2 input impedance mismatch when it was designed. The design of an interstage matching network is described in Sec. 5.7.

Example 10.7 Two-stage amplifier design. At 6 GHz an PET has the following parameters:

$$\begin{array}{lll} S_n = 0.8^\circ - 130^\circ & Sts = 0.2,630'' & S_{21} = 3z60^\circ \\ S_{21} = 0.3Z.140'' & r_{in} = 0.6/L16(T) & PL = 1.6 \quad R_x = 0.6 \end{array}$$

The design specifications are:

Two-stage power gain $O > 120$

Total noise figure $F_s < 1.8$

Input VSWR, < 1.5

Output VSWR, < 1.5

For this device $K = 0.513$, so that the transistor is not absolutely stable. The "Figure of Merit" gain equals $[f_{jj}/f_{c1}] = 15$. The normalized value is $g_p = 1.666$.

For the design of the first stage, the following table of minimum VSWR, values as a function of normalized power gain and noise figure was compiled:

gp	F	VSWR,	r
1.6	$*_m$	2.39	-0.564 +70.205
1.4	F_a	2.084	-0.564 +70.205
1.2	F_m	1.809	-0.564 +70.205
1.6	1.68	1.71	-0.581 -70.304
1.4	1.68	1.49	-0.576 70.306
1.2	1.68	1.294	-0.569 t-70.307

From the above table we see that we can obtain a noise figure of 1.68, a normalized power gain of 1.4, and an input VSWR, = 1.49 or, with a reduced normalized gain of 1.2, an input VSWR_t of 1.294. We will choose a power gain $g_p = 1.2$ and a value of $I_{\%}$ midway between those giving VSWR, equal to 1.294

and 1.809, so as to obtain a noise figure somewhat smaller than 1.68. Thus choose $r_s = -0.568 + j'0.27$. We now use this value and search the $g = i$ gain circle for the value of Γ_L' that will give the smallest values for VSWR₁ VSWR₂. The result is $T_L = 0.458 + .014$ and an amplifier having a nT figure 1.572 and an input VSWR₁ = 1.484 and a power gain $G_1 = 10.8$. The output VSWR₂ = 6.5015 which gives $M_2 = 0.4621$.

The design of stage 2 is carried out using the procedure described earlier. The objective function $W^AT_1 + W_2F_{nT}/F$ is optimized by searching for the value of l'' , on a chosen $g_p = \text{constant}$ gain circle and the value of r_j on the associated input impedance-mismatch circle using M_2 determined above. The search on the $g_p = 1.6$ circle resulted in a design with VSWR₂ = 1.65 and $F_2 = 3.693$ using $W_1 = 1$, $W_2 = 0$. For this example a search on lower gain circles gave higher values of VSWR₂. If we relax the design specifications to allow the somewhat larger output VSWR₃ value, then for the two-stage amplifier the following performance is obtained:

$$G_p = 10.8 \times 14.4 = 155.5 \quad \text{VSWR}_1 = 1.484 \quad \text{VSWR}_3 = 1.65$$

$$F = F_1 + M_2 = 1.572 + 0.4621 \times \frac{2.693}{0.962 \times 10.8} = 1.692$$

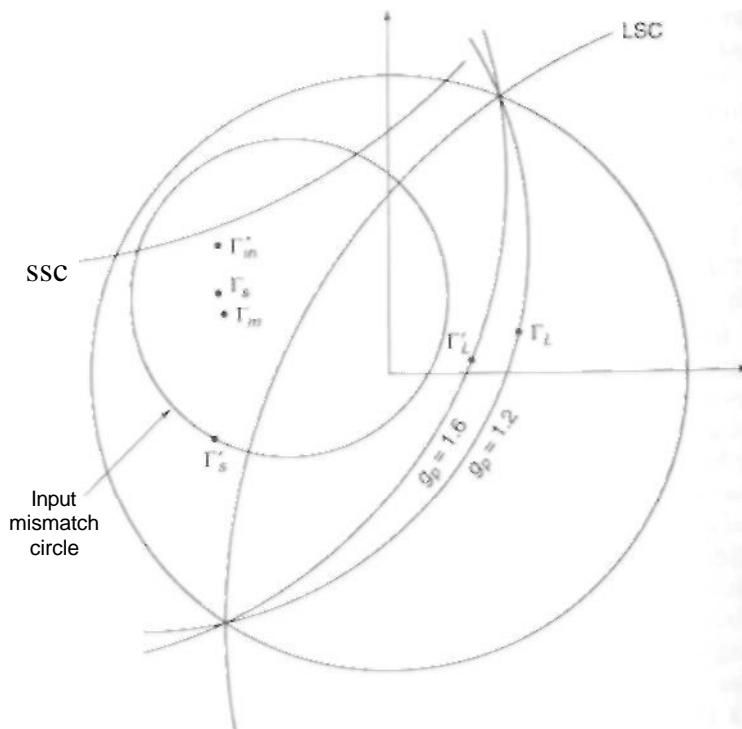


FIGURE 10.35

Illustration for the two-stage amplifier designed in Example 10.7.

For all specifications except the output $VSWR_3$, this design meets the stated criteria. For practical applications an output $VSWR_a$ equal to 1.65 instead of 1.5 is acceptable.

In Fig. 10.35 we show the load and source stability circles, the $g = 1.6$ and 1.2 constant gain circles, the values of r , and f' , for stage 1, the values of i''_s and \backslash''_L for stage 2, and the point $f_s = !''_{\text{m}}$. Also shown is the point I''_{m} for I''_N and the input mismatch circle for V'_L . Note that V'_s lies on this circle but is not as close as possible to the optimum point I_m . The reason for this is that V was chosen to obtain the best output $VSWR_3$, not the best noise figure. The figure clearly shows that t_v , \backslash_L , P , and \backslash''_L are sufficiently far away from the stability circle boundaries so the design has an adequate stability margin.

A second design was carried out for which the input stage was designed for a power gain of 7.2 and a minimum noise figure and a resultant input $VSWR| = 1.358$ was obtained. For the second stage we were then able to obtain $G_{ps} = 14.4$, $VSWR_3 = 1.517$, and $F = 4.28$. For the two-stage amplifier we obtained $G_p = 103.7$, $VSWR, = 1.358$, $VSWR, = 1.517$, and $F = 1.685$. This design has an input $VSWR$, lower than required, a noise figure essentially the same as for the first design, an output $VSWR,$ nearly equal to 1.5, but a significantly lower gain. The small increase in output $VSWR,$ in the first design is only a small price to pay for the much larger power gain that was obtained, so the first design is a better one.

10.12 OTHER ASPECTS OF MICROWAVE AMPLIFIER DESIGN

Microwave amplifier design as described in the preceding section is only a small part of the overall design problem. Once it has been verified that the specified performance can be obtained, a decision has to be made as regards to whether hybrid construction or monolithic integration will be used in the fabrication of the physical amplifier. A decision of whether to use microstrip circuits or eoplanar-waveguide circuits must also be made, as well as a decision of whether to use lumped elements or distributed elements for the impedance-matching networks. The layout of the circuit must be designed and the circuit must incorporate both the bias circuit and the RF elements with suitable decoupling of the RF circuit from the dc bias circuit. The physical dimensions of all transmission lines and other printed-circuit elements must be calculated and will depend on the substrate material used. Suitable masks must be prepared for use in the fabrication of the amplifier. After the amplifier has been built, it must be tested to determine if it meets the design specifications.

After the RF circuit and dc bias circuit have been designed, a theoretical evaluation or computer simulation of the amplifier should be carried out to verify that it is stable at all frequencies. This check should be performed before the amplifier construction is undertaken.

For a broadband amplifier the design of the matching networks is considerably more complex than that for narrowband amplifiers. The

matching networks must be designed so as to provide adequate stability, gain throughout the passband, and constant group delay. The latter requires that the phase function $\angle M(\omega)$ in the overall amplifier transfer function $H(\omega) = |H(\omega)|e^{-j\angle M(\omega)}$ be a linear function of ω (see Sec. 3.19).

For a power amplifier, dynamic range, nonlinear distortion, and intermodulation characteristics must be taken into account. Also, suitable provision must be made to remove the heat produced in the active devices.

Companies that manufacture microwave amplifiers use a number of computer-aided design (CAD) software packages to facilitate the overall design. The effort expended in carrying out a thorough design before construction is undertaken pays large dividends since there is very little that can be changed in either the circuit component values or the circuit layout once the amplifier has been built, particularly so for monolithic microwave integrated circuits (MMICs).

Hewlett-Packard manufactures modular microcircuit packages that are very useful in the early stages of microwave amplifier development. These packages provide a ready-made miniature box with input and output miniature 3-mm coaxial-line connectors and dc bias terminals that allow for easy mounting of a prototype circuit so that it can be tested. A photograph of these package modules is shown in Fig. 10.36.

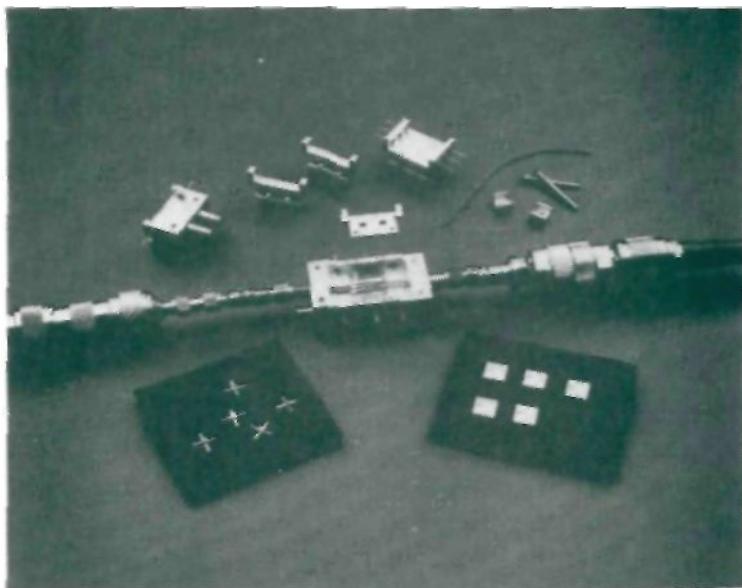


FIGURE 10.36
Modular microcircuit package for prototype circuit design and testing. (Photograph by Ray Moskaluk, Hewlett-Packard Company.)

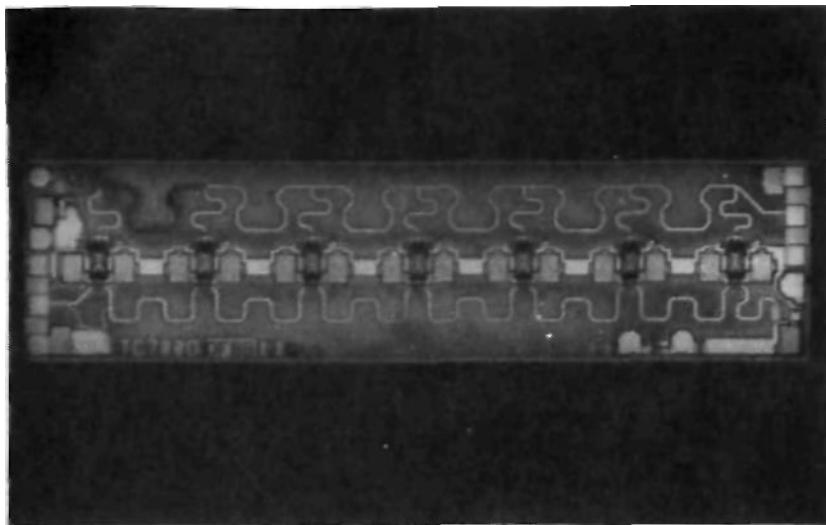


FIGURE 10.37

A broadband traveling-wave amplifier (MMIC) circuit. All components including matching networks are built on a single chip. (Photograph courtesy of Kay Moskaluk. Hewlett-Packard Company.)

A typical MMIC circuit is shown in Fig. 10.37. This is a broadband general-purpose traveling-wave amplifier. It is a GaAs MMIC chip using seven MESFET gain stages and has a flat gain of 8.5 ± 1 dB over the frequency range 2 to 26.5 GHz. A large number of gain stages is required since for any amplifier the gain-bandwidth product tends to remain constant, so that the gain per stage is necessarily low in a very broadband amplifier. The noise figure ranges from around 5 dB at the low-frequency end to 7 to 8 dB at the high-frequency end. The input and output VSWRs are less than 1.5.

PROBLEMS

- 10.1.** A bipolar transistor has the following scattering-matrix parameters at 2 GHz:

$$S_{11} = 0.56Z \angle 170^\circ \quad S_{12} = QM \angle -75^\circ$$

$$S_{21} = 4.04 Z \angle 76^\circ \quad S_{22} = 0.41Z \angle -23^\circ$$

Find the stability parameter K , and if $K > 1$ also find the maximum stable gain. Find the load and source stability circle parameters and plot these on a Smith chart.

- 10.2.** A silicon bipolar transistor is used in a common base amplifier circuit at 5 GHz. Its scattering-matrix parameters are:

$$\begin{aligned} S_{11} &= 1.3^{\wedge}140^\circ & S_{12} &= 0.2Z130'' \\ S_{21} &= 2Z. - 85^\circ & S_{22} &= 1.15/. - 55^\circ \end{aligned}$$

Find the stability parameter K , the maximum stable gain, and the load source stability circle parameters. Plot the stability circles on a Smith chart.

- 10.3.** A GaAs FET has the following scattering-matrix parameters at 2 GHz?

$$\begin{aligned} S_{11} &= 0.91Z - 42^\circ & S_{12} &= 0.05^{\wedge}33^\circ \\ S_{21} &= 6Z.105^\circ & S_{22} &= 0.62Z-95^\circ \end{aligned}$$

Find the stability parameter K and the load and source stability circles. Plot the stability circles and show the regions of the Smith chart where stable values of the load and source impedances lie.

- 10.4.** A bipolar transistor is used as a common collector (*source follower*) amplifier at 5 GHz. Its scattering-matrix parameters are:

$$\begin{aligned} S_{11} &= 0.63A - 96^\circ & S_{12} &= 0.8^{\wedge}15^\circ \\ S_{21} &= 2.3Z. - 53^\circ & S_{22} &= 0.62Z98'' \end{aligned}$$

Evaluate the stability parameter K . Find and plot the stability circles and show the regions of the Smith chart where stable values of Y_L and i_s occur.

Answer: $K = 0.3993$, $r_{Lc} = -1.298 + j4.0718$, $R_{Lc} = 4.5735$, $r_{s1} = 2.347 - j3.8792$, $R_{sc} = 4.8395$. Stable values lie inside the stability circles.

- 10.5.** An engineer adds an external network to a bipolar transistor and finds that in a common base connection its scattering parameters at 5 GHz are

$$\begin{aligned} S_{11} &= 1.3^{\wedge}140^\circ & S_{12} &= 0.2^{\wedge}130^\circ \\ S_{21} &= 2zL85'' & S_{22} &= 1.1SZ - 55^\circ \end{aligned}$$

Find the stability parameter K and the stability circles. Show in what regions of the Smith chart stable values of Y_L and T_s occur.

Answer: $K = 1.439$ but the device is only conditionally stable since $|S_{11}| > 1$ and $|S_{12}| > 1$; $V_{ub} = 0.174 + j0.6258$, $R_{ub} = 0.2172$, $T_{sc} = -0.2179 - j0.495$, $f_{sc} = 0.2714$. Origin is unstable.

- 10.6.** For the amplifier discussed in Example 10.2 and using the *design* that requires $Z_L = 85 + j168.2$, design a matching network consisting of an open-circuited stub located a distance d from the output that will transform the 50-11 line impedance into the required load impedance (see Sec. 5.6).

- 10.7.** Redesign the amplifier discussed in Example 10.3 so as to get a noise figure of 2 and a normalized power gain of 2.9. Find the required values of I_{sc} , w and the resultant input and output VSWRs.

Hint: Construct the $F = 2$ noise-figure circle. The best choice for I_s is a point on this circle lying on the line joining the center of the F - & P^* circles with the center of the P^* circle for $g_p = 2.9$. This I_s can be found since centers and radii of the circles are known. The corresponding E can also be found. From P , calculate P_{sc} , and from F^* calculate T_{sc} . Input and output impedance mismatches can now be found using (10.126). From these the input and output VSWRs can be found.

- 10.8.** A MESFET has the following parameters at 8 GHz:

$$\begin{array}{ll} S_{11} = 0.65Z - 150^\circ & S_{1a} = 0.12^{\wedge} 32^\circ \\ S_{21} = 2.2Z.6T & S_{22} = 0.1/150 \\ \Gamma_{ss} = 0.45^{\wedge} 130^\circ & R_K = 0.32 \quad F_{ss} = 1.3 \end{array}$$

Use the computer program MICROAMP to design a single-stage low-noise amplifier using the MESFET described above. The design requirements are: power gain of 9 or more, noise figure equal to 146 or less, input VSWR no greater than 1.5. What noise figure is obtained if conjugate impedance matching is used?

Answer; $G_s = 9.196$, $F = 1.4586$, input VSWR = 1.346, output VSWR = 1.166, $i_s = -0.5043 -j'0.3715$. $\backslash\backslash = -0.2188+.0.1157$. For conjugate impedance matching $F = 1.87$.

- 10.9.** Show that for a microwave amplifier the available power gain G_a can be expressed as $G_a = Mf_i/M_{ss}$, where G_p is the power gain and A_s , and M_{ss} are the input and output impedance-mismatch factors. By using relations of this type, show that the noise figure for a cascade connection of amplifier stages can be expressed as

$$\gamma_s = 1 + F_{ss} - 1$$

- 10.10.** Verify that the matching networks shown in Fig. 10.28 will transform the line characteristic admittance $1' \ll 0.02$ S into the required source and load admittances needed for the amplifier discussed in Example 10.3.

- 10.11.** Design a single-stage low-noise amplifier using an FET having the following parameters:

$$\begin{array}{lll} S_{11} = 0.74Z - 115^\circ & S_{12} = 0.14Z40^\circ & S_{21} = 2.7z87^\circ \\ S_{22} = 0.13Z - 60^\circ & r_{ss} = 0.5^{\wedge} 100' & \gamma_{ss} = 1.3 \quad R_x = 0.24 \end{array}$$

The design specifications are: $G_s > 15$, VSWR < 2 , $VSWR_2 < 2$, and $F < 1.5$. If you cannot meet the design specifications, relax one or more of the requirements.

Answer: A design with $G_s = 18.95$, $F = 1.33$, $VSWR_j = 2.11$, and $VSWR_2 = 2.265$ can be achieved using $\backslash\backslash = -0.17+.0.61$.

- 10.12.** The transistor whose parameters are given in Prob. 10.11 is to be used in a two-stage amplifier. The design calls for $G_p > 300$, $F < 1.5$, $VSWR_s < 1.5$, $VSWR_{s,i} < 1.5$. Plot the load and source stability circles, the $g_p = 2.6$ and 2.4 constant gain circles, and the $F = F_{ss} + 0.5$ dB constant noise-figure circle. On this figure show your final design values for V_s and $\backslash L$ for the first stage and I''_s and V'_s for the second stage. Does your final design have a good stability margin?

- 10.13.** A transistor has the following parameters:

$$\begin{array}{lll} S_{11} = 0.5Z 160^\circ & S_{12} = 0.06^{\wedge} 50^\circ & S_{21} = 3.6/.60^\circ \\ S_{22} = 0.5Z -45^\circ & l_{ss} = 0.4/.145^\circ & R_s = 0.4 \quad F_m < * > 1.6 \end{array}$$

Design a single-stage amplifier with the best possible noise figure subject to the constraints $G_s > 10$, $VSWR_s < 2$, $VSWR_2 < 2$.

10.14. A transistor with the following parameters is to be used in a two-stage amplifier:

$$\begin{array}{lll} S_u = 0.85L159^\circ & S_{12} = 0.06^{\wedge} - 74^\circ & S_{21} = 1.8z - 34^\circ \\ S_{2,,} = 0.66, 1 - 162^\circ & I_m = 0.8A - 160^\circ & F_{,,} = 1.74 \quad R_N = Q_3 \end{array}$$

The design objectives are: $G_s = 250$, $VSWR_j < 2$, $VSWR_r < 2$, $F < 22$ F your final design construct the load and source stability circles the $6, 7, 8, 9$ gain circles, and the $F = F_{,,} + 0.5$ dB and $F = F_{,,} + 1$ dB conS'am noise-figure circles. On this figure show the design values of T_s , r_s , and I^* at which you arrived.

REFERENCES

1. Vendelin, G. D., A. M. Pavio, and U. L. Rohde: "Microwave Circuit Design Using Linear and Nonlinear Techniques," John Wiley & Sons. Inc.. New York, 1990. This is a very comprehensive text covering small-signal and low noise amplifier design, power amplifier design as well as oscillator and mixer design.
2. Pengelly, R. S.: "Microwave Field-Effect Transistors—Theory, Design, and Applications," Research Studies Press, Hertfordshire. England. 1986, distributed by John Wiley & Sons. Inc., New York.
3. Gonzalez, G.: "Microwave Transistor Amplifiers. Analysis and Design," Prentice-Hall, Inc.. Englewood Cliffs, N.J.. 1984.
4. Vendelin, G. D.: "Design of Amplifiers and Oscillators by the S-Parameter Method," John Wiley & Sons, Inc., New York, 1982.
5. Ha, T. T.: "Solid-State Microwave Amplifier Design." John Wiley & Sons, Inc., New York, 1981.
6. Gentile, C: "Microwave Amplifiers and Oscillators," McGraw-Hill Book Company, New York, 1987.
7. Carson, R.: "High Frequency Amplifiers," 2nd ed., John Wiley & Sons, Inc., New York, 1982.
8. Chang, K. (ed.): "Handbook of Microwave and Optical Components. Microwave and Solid-State Components," vol. 2, John Wiley & Sons, Inc., New York, 1990. This is a very good reference source on semiconductor theory, device design and fabrication, device modeling, and applications, for microwave transistors and diodes.

CHAPTER 11

PARAMETRIC AMPLIFIERS

A parametric amplifier is an amplifier utilizing a nonlinear reactance, or a reactance that can be varied as a function of time by applying a suitable pump signal. The time variation of a reactive parameter can be used to produce amplification. This is the origin of the term parametric amplifier. The possibility of parametric amplification of signals was shown theoretically, as long ago as 1831, by Lord Rayleigh. The first analysis of the nonlinear capacitance was given by van der Ziel in 1948.^f He pointed out that this device could also be useful as a low-noise amplifier since it was essentially a reactive device in which no thermal noise is generated. The first realization of a microwave parametric amplifier was made by Weiss, following the earlier proposal (1957) by Suhl, suggesting the use of the nonlinear effect in ferrites (Sec. 6.7). In the following few years, the semiconductor-diode (sometimes called a varactor, for variable reactance) parametric amplifier was developed through the combined efforts of many workers. At the present time the semiconductor junction diode is the most widely used parametric amplifier. For this reason we limit the discussion in this text to this particular type of parametric amplifier. The *p-n* junction diode has a nonlinear capacitance. If a pumping signal at frequency w_p , and a small-amplitude signal at frequency ω , are applied simultaneously, the device behaves like a time-varying linear capacitance at the signal frequency ω . As we show in later sections, a time-varying capacitance or a nonlinear

^fA. van der Ziel, On the Mixing Properties of Nonlinear Capacitances. *J. Appl. Phys.*, vol. 19, pp. 999-1006, November, 1948.

capacitance can be incorporated into a circuit to give linear amplification of a small-amplitude signal. Before presenting this analysis a brief description of some of the properties of junction diodes is given, followed by a presentation of the Manley-Rowe relations. The latter are a set of power-conversion relations, of considerable value in determining the maximum gain and other performance features of parametric amplifiers.

11.1 *p-n* JUNCTION DIODES

The diodes used for parametric amplifiers consist of a junction of *n*-type and *p*-type semiconductor material. An *n*-type semiconductor has an excess supply of electrons, which is why it is called *n*-, or negative-, type. An example of an *n*-type material is a pure semiconductor such as germanium with a small amount (about 1 part in 10⁵) of impurity doping with an element such as arsenic or antimony. Germanium has a valence of 4 whereas arsenic or antimony has a valence of 5. Thus, at each site in the germanium or host crystal where an arsenic or antimony atom replaces a germanium atom, four of the valence electrons are used up to form the bond, and this leaves one excess valence electron. The valence electrons left over are relatively free to move around in the crystal under the influence of applied electric fields and make the material a donor of electrons, or *n*-type.

In *p*-, or positive-, type material, the impurity atoms are chosen to have a valence less than that of the host atoms. For example, gallium, with a valence of 3, may be used in a germanium crystal. When a gallium atom replaces a germanium atom, there are only three available valence electrons to form the required bond. A stable bond requires four valence electrons, and consequently, at each site where a gallium atom is located, a hole is created which can be filled by an electron that may be passing by. If an electron from some other bond moves over to fill the hole, the result is the creation of a new hole at some other point. The overall effect is as though the holes were positive carriers of electricity, i.e., equivalent positive electrons, that can move through the crystal. The holes do, in fact, behave as equivalent positive carriers, and thus *p*-type material can be considered essentially the same as *n*-type material except that the signs of the charge carriers are opposite.

Consider now a linearly graded junction of *n*-type and *p*-type material, as shown in Fig. 11.1a. In the linearly graded junction, the *n*-type material changes gradually and in a linear fashion over to *p*-type material in distance *d*. This variation is obtained by gradually reducing the doping concentration of donor atoms, down to zero in the region $x = d/2$ and then linearly increasing the concentration of acceptor atoms in the region $x = -d/2$ to $x = 0$. If the number density of acceptor atoms is N_0 and the density of donor atoms is N_d , the difference will vary as

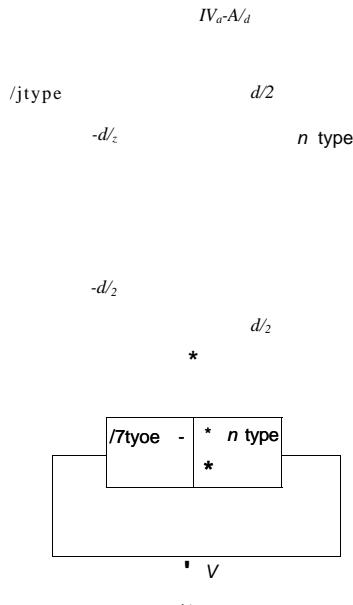


FIGURE 11.1
The linearly graded junction.

with x across the junction; thus

$$N-N_0 = kx \quad (11.1)$$

where k is a suitable constant.

When there is a gradient in the impurity-concentration densities, electrons will diffuse from a region of high concentration to one of low concentration. Holes will diffuse in a similar manner. Thus the electrons will diffuse into the p-type side of the junction and holes will diffuse into the n-type side. This diffusion continues until a space-charge distribution, together with a resultant electric field, is set up of sufficient strength to produce a force that is equal and opposite to the effective force created by the concentration gradients. When equilibrium has been reached, a small region, called the depletion region, substantially free of charge carriers, is produced at the junction. The space charge built up on either side of the depletion region, together with the electric field existing across the depletion region, constitutes an equivalent capacitor. If a reversed-bias voltage is applied across the junction, the electron distribution and hole distribution will be forced farther apart. This widening of the depletion layer results in a decrease in the junction capacitance. It is now apparent that if an ac pumping voltage is superimposed on the bias voltage, the equivalent junction capacitance can be varied as a function of time.

In the graded junction the space-charge density will vary linearly across the junction so that a depletion layer completely free of carriers is not

produced. The effect of having the space charge vary linearly across the junction instead of being concentrated at $x = \pm d$ is, however, the same. If we consider a linearly varying space-charge density (per cross-sectional area) $p = qx$, where q is a suitable constant, Poiss $^{(11-1)}$ equation gives

$$\frac{d^2\phi}{dx^2} = \frac{p}{\epsilon} = \frac{q}{\epsilon}$$

for the potential ϕ . Integrating and using the boundary conditions that the electric field, and hence $d\phi/dx$, is zero for $|x| > d/2$ and that $\phi = 0$ at $x = 0$ from symmetry considerations, we get

$$\phi = \frac{qT^*}{2\epsilon} \left(\frac{x}{3} - \frac{d}{4} \right)^2 \quad (11-2)$$

The potential difference across the junction is

$$V = \frac{qd^3}{127} \quad (11-3)$$

Under equilibrium conditions this potential difference must equal the contact potential T_c plus the negative applied bias voltage $-V$; thus

$$\phi - V = \frac{qd^3}{12e} \quad (11-3)$$

The total stored charge per unit area is given by

$$Q = \int qxdx = \frac{qd^2}{8}$$

Eliminating d by means of (11.3) gives

$$Q = \frac{8}{8} \cdot \frac{(c p, - V)^{2/3}}{12e} \quad (11.4)$$

Since the capacitance is a function of the voltage, it must be defined as the ratio of an incremental change in Q to incremental change in V . The capacitance per unit cross-sectional area is

$$C = \frac{dQ}{d(V)} = \frac{dQ}{d(\%V)} = \frac{\epsilon}{12e} \cdot \frac{8}{(c p, - V)^{1/3}} \quad (11.5)$$

As seen from this equation, the junction capacitance C is nonlinear since it depends on the voltage V . If C is a linear element, $Q = CV$. In an abrupt junction diode C is proportional to $(\phi_c - V)^{-1/2}$.

If we denote $\Delta - V$ by V_0 and apply in addition a pumping voltage $v_p = V_p \cos u_p t$, the capacitance becomes a function of time:

$$C_w = e I_u^* \frac{1}{2} \left(1 + \frac{v_p}{V_0} \cos u_p t \right) \quad (11.6)$$

We now have a nonlinear capacitance that is also a function of time. The capacitance is a periodic function of time, and can be represented by a Fourier series expansion of the form

$$C(t) = \sum_{n=0}^{\infty} c_n m s n a_p t \quad (11.7)$$

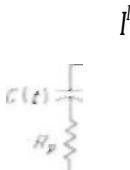
The coefficients are given by

$$c_n = \frac{2\pi \cdot 12fV_0}{(2eV_i)} \int_{-\pi/2}^{\pi/2} \left(1 + \frac{v_p}{V_0} \cos \theta \right)^{-1/3} \cos n \theta d\theta$$

$$c_n = \frac{V}{V_0} \int_{-\pi/2}^{\pi/2} \left(1 + \frac{v_p}{V_0} \cos \theta \right)^{-1/3} \cos n \theta d\theta$$

where $0 = \omega_{0,0}$. To evaluate the coefficients would require a numerical procedure. However, we do not need to know the values of the C_n in order to analyze the general properties of a parametric amplifier. The important feature brought out in the foregoing analysis is that $C(t)$ is a function of time that can be represented by a Fourier series involving all harmonics of the pumping frequency f_p . It is important to note that the coefficients are not, in general, linear functions of the ac voltage V or the voltage V_0 . Thus, since the junction capacitance $C(l)$ is a nonlinear capacitance, the principle of superposition *does not* hold for arbitrary ac signal amplitudes. Under small-signal conditions, a Taylor series expansion of $C(l)$ about an operating point may be used and only the linear term in the signal amplitude retained. In this case superposition *does* hold. The situation here is no different from that in any other amplifying device since all are nonlinear for sufficiently large applied signals.

In addition to the capacitance associated with the diode junction, there is a shunt conductance arising from the bulk resistance of the material in the depletion layer. This shunt conductance is proportional to the area, and since C is also proportional to the area, the ratio is independent of the cross-sectional area of the diode. The shunt conductance of the depletion layer is small, and can often be neglected. Of more importance is the series resistance of the n - and p -type semiconductor material outside the depletion layer. When the $p-n$ junction is encapsulated and connecting leads are put on, an additional shunt capacitance C_p due to the cartridge and a series inductance arising from the leads are also present. The overall equivalent circuit is thus of the form shown in Fig. 11.2. Typical values of C_p and L_s are somewhat less than 1 pF and 1 nH, respectively, at microwave frequencies.



Equivalent circuit Encapsulated
diode

FIGURE 11.2
Equivalent circuit of a parametric diode.

cies. The junction capacitance C is also about 1 pP, and typical values of R are a few ohms.

11.2 MANLEY-ROWE RELATIONS

Manley and Rowe have derived a set of power-conservation relations that are extremely useful in evaluating the performance which can be achieved from a parametric device consisting of a nonlinear reactance.^t These relations are also derived below.

The circuit considered by Manley and Rowe is shown in Fig. 11.3. It consists of resistive loads in series with ideal bandpass filters connected in shunt with a lossless nonlinear capacitance. Two sinusoidal signals at frequencies f_1 , and f_2 are applied. The *nonlinear capacitance causes* (re) frequencies at the harmonics of f_1 , and f_2 to be generated. Each bandpass filter is considered to pass only one harmonic component $n f_1 + m f_2$. The overall circuit thus isolates all the harmonics and dissipates their power in separate resistive loads. The Manley-Rowe relations establish two constraints governing the conversion of input power at the frequencies f_1 and f_2 into power at other frequencies.

Let the charge Q on C be a single-valued function of the ac voltage $v = L\dot{I} + v_2 = V_1 \cos \omega_1 t + V_2 \cos \omega_2 t$ applied across it. Thus $Q = Q(v)$. We may expand Q in a Taylor series in v to obtain

$$Q = Q(0) + \frac{dQ}{dv} v + \frac{1}{2} \frac{d^2Q}{dv^2} v^2 + \dots \quad (11.8)$$

where all derivatives are evaluated at $v = 0$. Since all powers of v occur, it is clear that, because $v = <V_1/2X_e^{f_1+1} + e^{-j\omega_1 t}> + (V_2/2X_e^{-f_2})e^{j\omega_2 t}$, the charge Q will have frequencies at all harmonics of f_1 , and f_2 . If currents •

^tJ. M. Manley and H. E. Rowe, Some General Properties of Nonlinear Elements, P ^ General Energy Relations, Proc. IRE, vol. 44, pp. 904-913, July, 1956. See also Proc. IRE*, 47, pp. 2115-2116. December, 1959.

$$\begin{array}{cccccc} f_i \cdot i & | & \cdot & f_i & | & K^* \cdot 2 \cdot ? \\ \sim f_y & | & \sim & h & | & \cdot n_f + m_f \end{array}$$

$$C(I)$$

FIGURE 11.3
Circuit for illustration of the Manley-Rowe relations.

the various harmonics are permitted to pass through C , the voltage developed across C will also contain all possible harmonics. In this case Q is a function of all voltages present. However, the expansion (11.8) will still be valid, except that now the coefficients will have different values. Consequently, the general expansion of Q has the form

$$Q = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} Q_{nm} e^{j(nu_1 t - mu_2 t)} \quad (11.9)$$

The charge is a real function of time; so we must have $Q_{nm} = Q_{-n,-m}^*$ in order that the n,m and $-n,-m$ terms will combine to form a real function of time with frequency n to, $+m$ w₂.

The total voltage V can be expressed as a function $V(Q)$ of the charge. A similar Taylor series expansion of $V(Q)$ then shows that V can be expressed in a form similar to (11.9); thus

$$V = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} V_{nm} e^{j(nu_1 t - mu_2 t)} \quad (11.10)$$

For V to be real, we must have $V_{nm} = V_{-n,-m}^*$.

The current through C is the total rate of change of Q with time, and is given by

$$I = \frac{dQ}{dt} = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} V_{nm} (u_1 + mu_2) Q_{nm} e^{j(nu_1 t - mu_2 t)} \quad (11.11)$$

$$I = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} Y_{nm} j(nu_1 t - mu_2 t) \quad (11.11)$$

where $I_{nm} = j(nw_1 + mw_2) Q_{nm}$.

Since C is a pure reactive element, there can be no net power into or out of C . If we assume that w_1 and io_2 are incommensurable, there will be no time-average power due to interacting harmonics. The average power at

the frequencies $\pm|nuj_l + mw_1|$ is

$$\begin{aligned} \mathbf{P}_{nm} &= (\mathbf{V}^* \mathbf{I}^* + \mathbf{V}^* \mathbf{I}) \\ &= (\mathbf{V}^* \mathbf{I}^* + \mathbf{V}^* \mathbf{T}^*) = \mathbf{P} \end{aligned} \quad (11.12)$$

since the time average of

$$\mathbf{I} \rightarrow \mathbf{V}^* \mathbf{I}^* = \mathbf{V}^* \mathbf{T}^* = \mathbf{V}^* \mathbf{I}$$

$$\frac{VI}{nm'nm} + \frac{V}{-n'm'nm} I = \frac{V}{nm'nm} T^* + \frac{V^*}{nm'nm} I = \frac{V}{nm'nm} / + \frac{T}{-n'm'nm} / = \frac{7^*}{-n'm'nm}$$

Conservation of power can therefore be expressed as

$$\sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} P_{nm} = 0 \quad (ii.i3)$$

since $\sum_{n,m} = \sum_{-n,-m}$ from (11.12). To obtain the Manley-Rowe relations, we multiply each term by $(nco_1 + mco_2)/(nio_1 + muo_2)$ and split the sum into two parts; thus

$$\sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} nP = co_2 \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} z \cdot p = 0 \quad (ii.i4)$$

We can now show that each double sum must vanish separately. We can replace each $I_{nm}/(nw_1 + mw_2)$ by jQ_{nm} and then $P_{nm}/(ra + ma)$ becomes $-JV_{nm}Q_m^* - JV_{-n,m}Q_{-n,m}^*$ and does not depend explicitly on OJ_1 or OJ_2 . For any choice of w_1 and w_2 , we can always adjust the network external to C so that the currents which result keep all the voltage amplitudes V_{nm} unchanged. The Q_{nm} are then also unchanged since they depend only on the V_{nm} . When this is done we see from (11.14) that it is possible to change w_1 and w_2 arbitrarily but keep the two double sums involving $P_{nm} = -JV_{nm}Q_m^* - JV_{-n,m}Q_{-n,m}^*$ unchanged. Consequently, (11.14) can hold for arbitrary w_1 and w_2 only if

$$\begin{aligned} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} nP &= 0 \\ \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} mP &= 0 \end{aligned}$$

That is, the coefficients of w_1 and w_2 must vanish separately. The above two relations are the Manley-Rowe relations. They are usually written in somewhat different form, however. We may write the first sum as

$$\sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} nP = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} nw_1 - \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} mo_2$$

where n and m have been replaced by E_n and $-m$ in the second term.

Since $P_{-n-m} = P_{nm}$, the two parts are equal; so we obtain

$$\sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \frac{n P_{nm}}{nco + into^*} = 0 \quad (11.15Q)$$

Similarly, we can obtain

$$\sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} \frac{m P_{nm}}{nco + into^*} = 0 \quad (11.156)$$

These are the standard forms for the Manley-Rowe relations. The Manley-Rowe relations are general power-conservation relations, and do not depend on any specific circuit such as that in Fig. 11.3. This is apparent since no reference to an external circuit was made in the derivation.

For an example of the application of the Manley-Rowe relations, consider a circuit similar to that in Fig. 11.3 with generators at frequencies f_x and f_2 . Let all harmonics be open-circuited except f_x , $+f_2$. Thus currents at the three frequencies f_x , f_2 , and $f_x + f_2$ are the only ones that can exist. The $n = +1$, $m = 0$ and $n = 0$, $m = +1$ and $n = m = \pm 1$ terms in (11.15) are the only ones present. Thus we get

$$P_{10} + P_{01} = 0 \quad (11.16a)$$

$$P_{10} = 0 \quad (11.1661)$$

Since power is supplied at the frequencies w_1 and w_2 we must have P_{10} and P_{01} positive. Therefore P_{10} is negative, and power is delivered from the nonlinear capacitor C at the frequency $w_1 + w_2$. If a_i is the input-signal frequency and m_2 is the pump frequency, then $\omega_3 = co + OI$, is the output frequency. The maximum signal gain is

$$\frac{-P_{10}}{P_{01}} = \frac{M_1 + M_2}{W_1} = 1 \quad (11.17)$$

as obtained from (11.16a). A parametric amplifier of this type is called an up-converter. Because of losses that are always present in a practical amplifier, the gain will be less than ω_3/a . The Manley-Rowe relations give the maximum gain possible and hence provide a criterion by which a practical up-converter can be judged.

P-3 LINEARIZED EQUATIONS FOR PARAMETRIC AMPLIFIERS

Consider a linear capacitance C for which the charge-voltage relationship is $Q = Cv$. The current flowing through C is given by

If C is made a function of time, for example, a parallel-plate capacitor •
a plate separation that is varied with time, the current will be given by

$$i = \frac{d}{dt} (Cv) = v \frac{dC}{dt} + C \frac{dv}{dt} \quad (U.18)$$

If instead of a time-varying linear capacitance we have a nonlinear capacitance, where the charge Q is a nonlinear function $Q(v)$ of the volta
the current is given by

$$i = \frac{dQ}{dt} \frac{dv}{dt}$$

If the voltage v is the sum of a pump voltage v_p at frequency ω_p and
signal voltage v_s at frequency ω_s and $|v_s| \ll v_p$, we can expand

$$Q(v) = Q(v_p + v_s)$$

in a Taylor series about the point $v_s = 0$. Thus we obtain

$$Q(v_p + v_s) = Q(v_p) + \left. \frac{dQ}{dv} \right|_{v=0} v_s + \frac{1}{2} \left. \frac{d^2Q}{dv^2} \right|_{v=0} v_s^2 + \dots$$

For v_s sufficiently small, we can obtain satisfactory accuracy by retaining
the first two terms only. The current is then given by

$$i = \frac{dQ(v_p)}{dt} + \frac{d}{dt} \left[\left. \frac{dQ}{dv} \right|_{v=0} v_s \right] \quad (11.19a)$$

Let the quantity dQ/dv for $v_s = 0$ be denoted by $C(t)$, in which case we can
write

$$i = \frac{dQ(v_s)}{dt} + \frac{d}{dt} [C(t)v_s] \quad (11.19b)$$

If we compare this result with (11.18), we see that the nonlinear capacitance
behaves like a *time-varying* linear capacitance for signals with amplitudes
that are small compared with the pump signal amplitude. The first term,
 $dQ(v_p)/dt$, in (11.19) gives a current at the pumping frequency, and is not
related to the signal current. If the pumping voltage is also small compared
with the dc bias voltage in a junction diode, we can assume $C(t)$ to have tri
form [see (11.6)]

$$C(t) = C_0(1 + 2M\cos\omega_p t) \quad (11.20)$$

since $V_0 + V_p \cos\omega_p t \approx V_0 + V_p \cos\omega_p t$ for $V_p \ll V_0$. The
linearized equations (11.19) and (11.20) are the ones we shall use in the
analysis of parametric amplifiers.

The equivalent circuit of a *p-n* junction diode is illustrated in Fig.
11.2. For the purpose of analysis it is more convenient to use an equivalent
series circuit of the form shown in Fig. 11.4. The two circuits are equivalent

$\bullet T^{Cs}$ FIGURE 11.4
 I Equivalent series circuit for a junction diode.

if we choose

$$R. \quad \frac{R_p C^2}{(u > CC_p R_p f + (C + C_p f)} \quad \mathbf{C} = \frac{(<\omega C C_p R_p f + (C + C_p f)}{(a > CR_p f c_p + C + C_{ss})}$$

which makes the input impedance the same for both circuits at the frequency w . In most diodes the resistance R_p is small compared with the cartridge reactance $1/i\omega C$; so $c\omega C.R \ll 1$. In this case we find that

$$R. \quad \frac{\mathbf{C}}{\mathbf{C} + \mathbf{C}_{ss}} \quad R. \quad \mathbf{C}. \quad \mathbf{C}_{ss}$$

When these approximations are valid, the two circuits are equivalent, independently of frequency. This is a necessary requirement if the series circuit is to be useful for analysis purposes, since in a parametric amplifier currents and voltages at several different frequencies are simultaneously present. When C is a function of the voltage, R_s will also be voltage-dependent. The effect of a voltage-dependent resistance R_s will, for simplicity, be neglected, since it is not too important. In other words, we shall consider R_K to be a constant resistance.

PARAMETRIC UP-CONVERTER

In the up-converter, a pump voltage at frequency $<u$ and a signal at frequency w_s are applied to the diode, and the output signal is taken at the higher frequency $>u + co_p$. Mixing effects take place that give rise to all possible harmonics of w_p and $>u$. However, in the up-converter, the circuit external to the diode is chosen so as to permit currents only at the signal frequency w_s , the pump frequency $>u_p$, and the output frequency $<\omega_0$, which is chosen as the sum of the pump and signal frequencies, that is, at $w_0 = <u + w_s$. There will, consequently, be a voltage across the diode at the

three possible frequencies. If we let the diode voltage be represented

$$v_s = \operatorname{Re}(V >^{\text{TM}}) = i(V_5 e^{-\omega_s t} + V_0 e^{j\omega_a t})$$

$$v_p = \operatorname{Re}(V_p e^{j\omega_p t}) = |(V_p e^{j\omega_p t} + V_0 e^{j\omega_a t})|$$

$$u_a = \operatorname{Re}(V_0 e^{j\omega_a t}) = |(V_0 e^{j\omega_a t} + V_0 e^{j\omega_p t})|$$

we may generalize (11.19) and (11.20) to give

$$i = \frac{dQ(v_p)}{dT} = \frac{1}{2d} \frac{d}{d\omega} C_{\text{eff}} (1 + 2M \cos \omega)$$

$$x(V, e^{-\omega_s t} + V_f e^{-\omega_p t} + V_0 e^{j\omega_a t} + V_0 e^{j\omega_p t}) \quad (\text{U.21})$$

Let the current at the frequencies ω_s and ω_0 be expressed as

$$i_s \sim k(l e^{j\omega_s t} + l^* e^{-j\omega_s t}) \quad (\text{11.22})$$

$$i_o \sim k(o_0 e^{j\omega_0 t} + o_0^* e^{-j\omega_0 t}) \quad (\text{11.23})$$

When the time derivative of (11.21) is taken and the terms at frequencies ω_0 and $\pm io_0$ only are retained, we obtain (a knowledge of the pump current at frequency ω_0 , will not be required; so we do not need to evaluate it)

$$\begin{aligned} & * + *** - fU''?_* - Ju_e V^* e^{-j\omega_0 t} + jo_0 V_0 e^{j\omega_0 t} - j<o_0 V^* e^{j\omega_0 t} \\ & + jw_0 M V_s e^{j\omega_0 t} - ja>_s M V f e^{-j\omega_0 t} - ju_0 M V^* e^{-j\omega_0 t} + j^* M V^* e^{j\omega_0 t} \end{aligned}$$

Using (11.22) and (11.23) gives

$$h = -j^* \& y_s + \mathbf{M} \mathbf{A}^* \quad (\text{11.24a})$$

$$/o = J^* o C_0 V_0 + o C_0 M y_s \quad (\text{11.24b})$$

These two equations show that, for the input signal current i_s and output signal current i_o , the junction capacitance may be represented by an admittance matrix such that

$$\begin{array}{ccc} I_s & jw_s C_0 & j(o_s C_0 M) \\ I_o & ju_a C_0 M & jo_0 C_0 V_0 \end{array} \quad (\text{11.24c})$$

The parameter M is proportional to the pump voltage and gives the coupling between the voltages at the two frequencies ω_a and ω_0 .

Figure 11.5 illustrates an equivalent-circuit model of an up-converter. The series tuned circuits are chosen so that the three circuit loops have resonant frequencies of ω_s , ω_0 , and ω_p and only currents with these respective frequencies can exist in each loop. Thus, in the input circuit 1 P only I_s is present. The three circuit loops are coupled together through time-varying part C of C_s only. Therefore, for the two frequencies ω_s and ω_0 , the equivalent circuit can be reduced to that illustrated in Fig. 11-6. The box labeled $C(t)$ in this circuit is an equivalent impedance network which maintains the relationship given by (11.24) between the terminal currents.

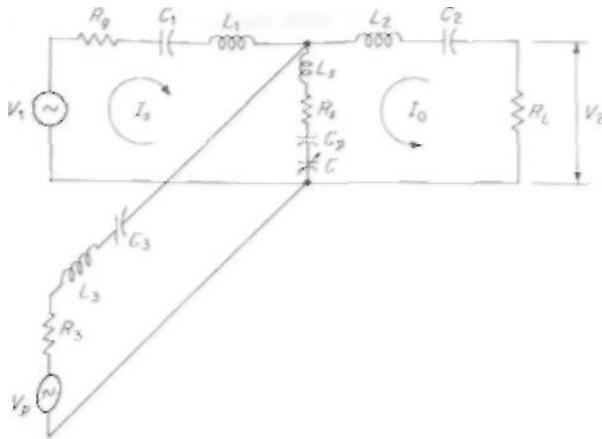


FIGURE 11.5
Equivalent-circuit model of an up-converter.

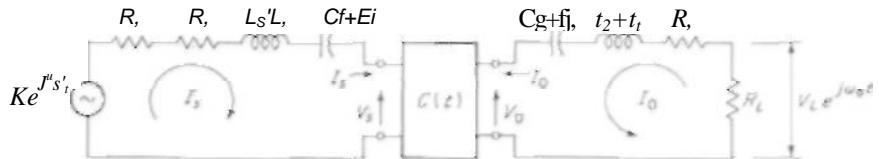


FIGURE 11.6
Reduced equivalent circuit for an up-converter.

and voltages. The analysis of the parametric amplifier is a conventional network-analysis problem since the diode has been replaced by an equivalent linear two-port network with terminal relations described by (11.24). Each loop in the circuit is assumed to provide a very high impedance to currents at all frequencies present except the resonant frequency for that loop. The resonant circuits in Fig. 11.5 have been assumed to have zero loss. Circuit losses can be considered included in R_p and R_e . At the end of the analysis, R_g and R_L can then be split into two parts so as to separate the circuit losses from the generator and load impedances. In practice, circuit losses are small compared with the loss arising from the diode resistance R_p and the external loading represented by R_e and R_L .

We may solve (11.24c) for V_v and V_0 in terms of I_a and I_0 to obtain

$$\frac{V_v}{1 - M^2} = \frac{-M/jco_0C_0}{1 - M^2} \quad 11.25)$$

^aCare must be exercised in the analysis since currents and voltages at several different frequencies are simultaneously present.

For the input circuit we may now write

$$\Delta = I_s f_t + /?_s + j * J_i L_a + L_s) + \mathbf{y}_{ws}(\mathbf{c}_s + \mathbf{c}_o) + v_o$$

$$= & f_t + /?_s + >_s (L_s + L_J + \frac{i}{y_w C_p + C_j}) + \frac{1}{y_w s (i - M^2) C_0}$$

$$y_s (1 - M^2) C_0$$

whereas for the output circuit

$$\mathbf{0} = I_s * * + *_s + \mathbf{M}(\mathbf{f}_s + \mathbf{f}_o) + \frac{1}{j m_0 (\epsilon_p + C_2)} + V_o$$

$$= I_s, \frac{1}{>_o (C_s + C_2)} + \frac{1}{j w_0 (1 - M^2) C_o}$$

M7.

$$j o_s (1 - M^2) C_u$$

Let us now assume that the circuits are tuned so that the following conditions hold:

$$\theta > i_s = C_p + C_o + (1 - A^2) C_0$$

$$oi_b = \frac{i}{L_s + U} \frac{1}{C_p + C_2} \frac{1}{(1 - M^2) C_0}$$

We then obtain

$$\begin{aligned} & y_s / o \\ & w_0 (1 - M^2) C_0 \\ \mathbf{0} = & 7_0 (f_t, + f_o) \frac{j M 7_s}{\ll, (1 - A f^2) C_o} \end{aligned}$$

We may solve for 7_0 to obtain

$$\mathbf{h} = \frac{-J V M w o C^2 l - M^2}{M^2 + (R_g + f_t j (f_t, + f_o)) \gg, (1 - A^2) C_o} \quad (11.26)$$

The maximum available input power from the generator is and the output power developed in f_t, o is $|7_0|^2 f_t, o$. The midband transdu

power gain is thus given by

$$\begin{aligned}
 G_0 = & \frac{4|J_0|^2 f_i t / ?}{Vf} \\
 & AR_L R_g M^2 \\
 & a \cdot ^2 (1 - M T C^2) (R_a + R_t) / (R_h + R_s) + \frac{M''}{<o_0 c o_s (l - M^2 V C^*)} \\
 & (11.27)
 \end{aligned}$$

when (11.26) is used to express I_0 in terms of V . If desired, circuit losses can be included at this point by replacing R_g and R_L in the denominator in (11.27) by $R_g + R_u$ and $R_s + R_{2t}$, where R_u and R_{2t} represent the loss resistance in the input and output circuits. For simplicity we shall take

To achieve maximum gain requires adjustment of R and R_L . Since R_K and R_s occur symmetrically in the expression for G_m , the optimum values of R_L and R_a are equal. Hence we need to maximize

$$\begin{aligned}
 G_0 = & \frac{ARIM'}{M^2} \\
 & <o^*, (I-M^*) \sim C^* (R_L + R_s) +
 \end{aligned}
 \quad |2$$

Equating dG_0/dR_L to zero and solving for R_L give

$$R_L = R^* I + \frac{M^2}{jw, J?^2 (I-M^2) C_u^2} \quad |2 \quad (11.28)$$

The effective Q of the diode may be defined as

$$\frac{1}{i \cdot , \wedge (1 - M^2) C_a} \quad |1 \quad (11.29)$$

We then find that

$$R_L = R_s - \frac{1}{(M Q)^2} \quad |2 \quad (11.30)$$

and the maximum gain is

$$G_0 = \frac{1}{m * (i + v T T T)} \quad |1 \quad (11.31)$$

where $S = (w_s / o >_0 X M Q)$. According to the Manley-Rowe relations discussed in Sec. 11.2, the maximum gain of an up-converter is $w_0 / o >_0$. The quantity $8/(1 + fl + <5)^2$ may therefore be regarded as a gain-degradation

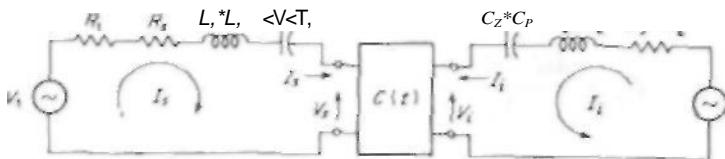


FIGURE 11.7
Equivalent circuit for a negative-resistance parametric amplifier.

factor. As the diode Q approaches infinity, that is, as R_s goes to zero, η approaches infinity, and the gain-degradation factor becomes equal to unity. Hence, for a lossless diode, the gain becomes equal to $\langle \alpha_0/a \rangle_s$, as predicted by the Manley-Rowe relations. In a typical microwave diode, MQ could be equal to 10. If $\alpha_0/w_s = 10$ also, the maximum gain as given by (11.31) is 7.3 dB.

To achieve high gain with an up-converter requires a large ratio ω_w/ω_0 of output-to-input frequency. At the higher microwave frequencies this is not a very practical requirement, and for this reason up-converters are usually restricted to operation at signal frequencies f_s below 1,000 MHz. Higher gain can be obtained from the negative-resistance parametric amplifier, which is discussed in the next section.

11.5 NEGATIVE-RESISTANCE PARAMETRIC AMPLIFIER

In the negative-resistance parametric amplifier currents are permitted to exist at the signal frequency ω_0 , the pump frequency ω_p , and the idler frequency $\omega_i = \omega_p - \omega_0$. The equivalent-circuit model that will be analyzed is shown in Fig. 11.7.

When we replace the voltage ϕ_0 in (11.21) by $v_s = \{ (V^A' + V_t * e^{j\omega t}) \}$ and introduce the idler current $i_s = \{ (I_s * e^{j\omega t} + I_r * e^{-j\omega t}) \}$, we may solve for I_s and I_r in terms of V_s and V_t , in the same manner that was used to treat the up-converter. It is readily found that, for $w_s = w_p - w_s$,

$$ju.^{\wedge}o \quad jcOiCoM \quad V^* \quad (11.32a)$$

and

$$\begin{array}{cccccc} & & M & & & \\ V, & \begin{matrix} Jo > iC_Q \\ -M \end{matrix} & & h & & (11.32^*) \\ V^* & 1 - M^2 & -1 & 7^* & & \end{array}$$

For the circuit of Fig. 11.7, we can write the following equations:

$$\begin{aligned} \mathbf{W}_1 &= R_s + jw_s(L_l + L_s) \quad \rightarrow x(C_0 + C_2) \quad \frac{1}{ju_s(lM^2)C_0} \\ &\quad + \frac{M}{(1 - M^2)C_0} \quad ! 11.33a) \end{aligned}$$

$$\begin{aligned} V_2 &= 7, R_2 + R_s + j\omega_o(L_2 + L_s) \quad J_o > (C_0 + C_2) \quad > (1 - M^2)C_0 \\ &\quad + \frac{M}{(1 - M^2)C_0} \quad ! 11.336 \end{aligned}$$

If we impose the tuning conditions

$$\begin{aligned} \ll ! &= C_0 + C_2 + (1 - A f^2)C_0 \\ &\quad + \frac{1}{L_0 + L_p C_p + C_2} \quad (1 - M^2)C_0 \end{aligned}$$

we obtain

$$V_2 = (*, + \ll,) /, - \quad ! 11.34a)$$

$$V_2 = (R_2 + R_s)I_s - \frac{1}{a} \quad ! 11.346$$

Let us now assume that $V_2 = 0$. We can then determine $/$, from (11.34) and evaluate the gain $G_0 = 4i?_1/?_2|/|V|^2$. We readily find that

$$4Z?_1f_1M^2$$

$$\begin{aligned} \text{On } &= R_s + R_l - \frac{M^2}{(R_2 + R_s) - Q(lM^2)} \\ &\quad (\mathbf{K}_2 + I^* \ll, \mathbf{K} \ll, (\mathbf{i} - \mathbf{J} \mathbf{n} \mathbf{cs}) \quad ! 11.35) \end{aligned}$$

The term $-M^2/(R_2 + R_s) - M^2C_0^2$ may be interpreted as an equivalent negative resistance $-/$. Introducing $i?$, we may express G_0 as

$$G_0 = \frac{4i?_1rt_1(t, (1 - M^2)C_0/f)^V}{M^2(f_1 + R_s - R_s') \quad ! 11.36}$$

It is clear that a very large gain can be obtained if R is made almost equal to $i? + R_s$. However, care must be taken not to make $7?$ too close to $R_l + R_s$ because a small change in parameters will then cause large changes

in the gain and will cause oscillations to occur if R becomes $\frac{1}{R + R_x}$ equal to

The parametric amplifier discussed above is called a negative resistance converter. It is possible to take the output at the same frequency as the input. If we split R_j into a generator internal resistance R_{di} to be load resistance R_L , the power delivered to R , is $5 \cdot R_{\text{JJ}}^2$. The transd power gain will be

$$G_0 =$$

We may evaluate I from (11.34) to obtain

$$G_0 = \frac{4RR_L}{(R_g + R^* + R_x - Ry)} \quad (11.37)$$

where, as before,

$$R = (R_z + R_s) < o_s < o_s(l \sim M^2)^2 C^* \quad (11.38)$$

The effective negative resistance $-R$ arises in the following manner: The application of signal plus pump power to the nonlinear capacitance causes frequency mixing to occur. When current is permitted to exist at the idler frequency $o_p - w_s$, further frequency mixing of power at the pump and idler frequencies occurs. This latter mixing causes harmonics of o_p and $(o_p - OJ_s)$ to be generated; in particular, power at the frequency w_s is generated. When the power generated through frequency mixing exceeds that being supplied at the signal frequency w_s , the diode appears to have a negative resistance. If idler current is not permitted to exist, the negative resistance vanishes, as reference to (11.38) shows when R is made infinite (open circuit for the idler signal).

The negative-resistance parametric amplifier with input and output at the same frequency is not very stable. The reason is that in a microwave system R_g and R_L are the impedances seen looking into the input and output transmission-line ports. If the loads connected to these transmission lines are not matched, reflected waves occur. Reflected waves in the output fine return to the amplifier and are amplified and fed into both the input and output lines. The result is that the gain becomes a sensitive function of the external generator and load impedances. The stability of the amplifier greatly improved by the use of a circulator, as illustrated in Fig. H-^o. The use of a circulator makes the load termination R_L for the amplifier equal to the characteristic impedance of the transmission line independently of the external generator or load impedances Z_g and Z_L . The available power

the generator is still given by $\frac{V^2}{4R_g}$. However, the amplifier power

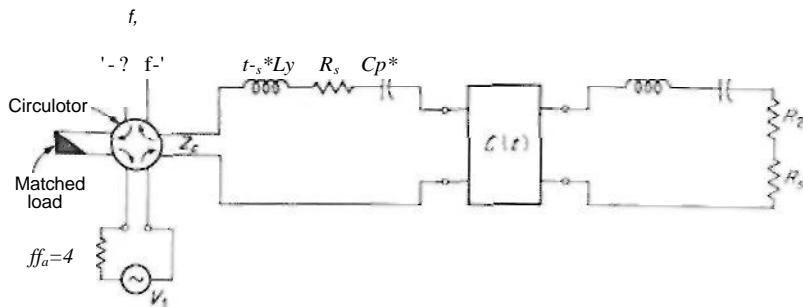


FIGURE 11.8
A negative-resistance parametric amplifier using a circulator.

now all delivered to the load $R_L = Z_c$, and none of it is dissipated in the internal generator impedance R_g . Consequently, the power gain is nearly four times greater since I_s is nearly twice as large, because the series resistance in the input circuit is now $R_L + R_{ss}$ instead of $R_L + R_g + R_s - 2R_L$, when $R_L = R_g \gg R_s$. The power gain is the square of the voltage reflection coefficient, and is given by

$$G_0 = \frac{Z_n - Z}{Z_m + Z_r} \quad (R_s + R_t \sim R)$$

since $Z_m = R_s - R$ at resonance and we are taking $R_L = Z_c$. For high gain, $R \sim R_L + R_s$ and is large compared with R_a . Consequently, the gain can be expressed as

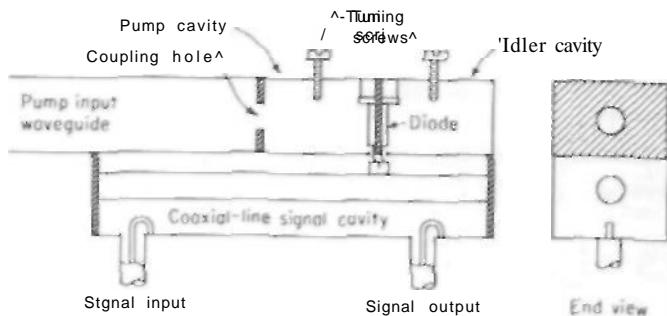
$$G_0 = \frac{4*1}{(R_L + R_s - R)} \quad (11.39)$$

The maximum value of negative resistance that can be obtained is determined by the diode that is used. If we make $R_2 = 0$ in (11.38), we see that maximum R is R_{ss} , where

$$R_m = -(MQVR) \quad (11.40)$$

(11.29) having been used to introduce the diode Q . If we regard R as fixed, we see that, for the amplifier without a circulator, we must make $R_g + R_L + R_s = R$, as (11.37) shows. But with $R_g = R_L = Z_c$ large compared with R_s , we get $R_g = R_L \sim R/2$, and (11.37) gives

$$G_0 = \frac{R'}{(2R_L + R, - R)} \quad (11.41a)$$

**FIGURE 11.9**

A microwave negative-resistance parametric amplifier.

whereas for the amplifier with a circulator, $R_s = R$ and

$$4R^2$$

$$G_{dB} = \frac{1}{4} \ln \left(\frac{R_L + R_s}{R_L - R_s} \right) \quad (11-41)$$

Note that when a circulator is used, the optimum value of R_L is twice what it is when a circulator is not used. Thus the denominators in (11.41a) and (11.41b) are the same. These relations then show that the use of a circulator gives 6 dB more gain for the same amount of diode loading, and hence will have a gain-bandwidth product twice as great.

If the pump frequency ω_p is chosen equal to twice the signal frequency ω_s , the idler frequency $\omega_i = \omega_p - \omega_s = \omega_s$ is equal to the signal frequency. In this case the amplifier is called a degenerate negative-resistance amplifier. For the degenerate amplifier the signal and idler circuits would be a single resonant circuit. The analysis of the degenerate amplifier is similar to that already carried out, and so will not be presented here (some results on noise properties are given in Sec. 11.6).

There are many different ways of building microwave parametric amplifiers. Transmission lines, waveguides, or a combination of the two may be used to construct suitable cavities to use as resonant circuits. A typical microwave negative-resistance amplifier is illustrated in Fig. 11.9. The pump and idler cavities are formed in an X-band rectangular waveguide. The signal cavity is a coaxial-transmission-line cavity. The varactor diode is mounted in the center of an inductive diaphragm located between the pump and idler cavities. Coupling to the signal cavity is achieved by having the diode terminate in the center conductor of the coaxial-line signal cavity. The pump is coupled to its cavity by an aperture. The pump frequency is around 9,200 MHz, and the idler frequency is 7,900 MHz. The mpu

output signal frequencies are the same and equal to 1,300 MHz. This amplifier gives a gain of 25 dB or more with a bandwidth of about 5 MHz.

The bandwidth over which high gain can be obtained in a negative-resistance amplifier of the type discussed above is small. The negative resistance has the effect of increasing the loaded Q , which results in a *high-Q* resonant circuit with a narrow bandwidth. To analyze the bandwidth properties, we shall assume that the circuit model given in Fig. 11.7 is valid. For high-Q circuits the impedance of the signal and idler circuits may be expressed in the form

$$Z_s = (R_i + R.M + 2j\hat{Q}_s) \quad (11.42a)$$

$$Z_i = (R_2 + R_{jll} + 2j\hat{Q}_i) \quad (11.42b)$$

where

$$\frac{Li + L_t}{R_s + R_i} \quad \text{and} \quad \frac{L_2 + L_i}{R_s + R_i}$$

The derivation of these expressions is as follows: From (11.33a) the impedance of the signal circuit at a frequency $w_s + Aw_s$, is

$$Z_s = R_s + j(\omega_s + A\omega_s) \quad (* - i^+ f.)$$

$$(A\omega_s + Aw_s)^2 C_s + C_s \quad C_0 M^2 C_0$$

$$= \frac{1}{R_s + j(A\omega_s + Aw_s)} \quad \text{and} \quad \frac{Ml}{(A\omega_s + Aw_s)^2 C_s + C_s}$$

$$= fl_s + R_s + j \frac{(2w_s Aw_s + Aw_s^2)}{co_c + AOJ_K}$$

$$= (\#_s + R_s) \mid l + j2Q_s Aw_s$$

A similar derivation holds for Z_i .

From (11.33) we obtain

$$V_s = IJR_s + R_s(i_s + 2jQ_s) \quad \text{and} \quad \frac{Mlf}{(I-M')j(w_s + Aw_s)C_0} \quad (11.43a)$$

$$V_i = I(R_2 + R_s) \mid l + 2jQ_2 - L_i + \frac{MI!}{(1-M^2)7(w_s + Aw_s)C_Q} \quad (11.43b)$$

If $i?j = R_g + R_L$ and the output is taken at the frequency ω_0 , the e_i gain is given by

$$G = \frac{4ft_g f_l J_s}{V^2}$$

When we solve (11.43) for J_s , we find that (note that $V_2 = 0$ in this case)

$$\begin{aligned} G &= \frac{2}{Z_0 - \frac{(A w_s + A w_s)(1 - M^2)}{C_0^2 Z_0^2}} \quad (H.44) \\ &\text{AT} \end{aligned}$$

Since the pump frequency is fixed, $Aw_s = -Aw_s$. The midband gain G_0 is given by (11.37). To determine the bandwidth, we equate G given by (11.44) to $G_0/2$ and solve for Aw_s . When the gain G_0 is high, we find that, to a good approximation,

$$\frac{Aw_s}{(J_s^2 + f_l^2(Q_1 + Q_2)^2)} = \frac{Rt + R_s - RY}{(Rt + R_s - RY)^2 + (f_l^2(Q_1 + Q_2))^2} \quad (11.45)$$

Thus the gain-bandwidth product becomes

$$\frac{Aw_s}{G_0} = \frac{Rt + R_s - RY}{(Rt + R_s - RY)^2 + (f_l^2(Q_1 + Q_2)^2)} \quad (11.46)$$

If we assume as a typical case $R_g + R_L = ft_s$ and $R_g = R_L$ and note that, for high gain, $ft_s + R_s = R$, we get

$$\frac{2/G_0}{w_s} = \frac{Aw_s}{Q_1 + (w_s/w_s)Q_2}$$

The smallest possible value of Q , occurs if $C_1 + C_2$ equals zero. In this case

$$\frac{2/G_0}{w_s} = \frac{\omega_i + \omega_s}{w_s(1 - M^2)C_0(ft_s + ft_p)} = \frac{\omega_i + \omega_s}{(1 - M^2)C_0ft_s}$$

Similarly, the smallest possible value of Q_2 is obtained if $C_2 + \omega_p$ and is

$$\omega_a = w_s(1 - M^2)C_0(ft_p + ft_s)$$

If we refer to (11.38) for ft_p , we now find that $M^2 Q_1 Q_2 = 1$. Thus

$$\frac{2/G_0}{w_s} = \frac{Aw_s}{Q_1 + w_s/M^2 w_s Q_2}$$

This expression has a maximum value for

$$1 \sqrt{\frac{\omega_s}{\omega_i}}$$

Hence the maximum gain-bandwidth product is

$$V_s / \text{max} = \sqrt{\frac{\omega_i}{\omega_s}} \quad (\text{U.47})$$

For 20-dB gain we obtain a bandwidth of

$$2 Aa \ll s \quad Af \quad rWj$$

$$W_s = 2oy \ll T$$

Usually M is no greater than about 0.2, and consequently the fractional bandwidth in percent is approximately equal to Jia/w_s . In a practical amplifier the gain-bandwidth product will be less since the capacitances $C_l + C_p$ and $C_2 + C_p$ cannot be reduced to zero.

The parametric amplifier may, of course, be broadbanded by using broadband circuits at the signal and idler frequencies. An alternative scheme for obtaining broadband operation is the traveling-wave parametric amplifier, where *resonant circuits* are avoided *entirely*. In the traveling-wave amplifier a waveguiding system is loaded periodically with varactor diodes. With the application of pump power and signal power, mixing occurs, with resultant power generation at the signal frequency. For a detailed analysis the references cited at the end of this chapter may be consulted.

16 NOISE PROPERTIES OF PARAMETRIC AMPLIFIERS

The noise produced by parametric amplifiers is the thermal noise in the resistances in the equivalent circuit. For the up-converter illustrated in Fig. 11.6, the input thermal noise at frequency w , is that arising from the generator resistance R_g . In analyzing the noise power in a circuit, it is useful to consider noise as consisting of a spectrum of noise waveforms with an effective root-mean-square (rms) voltage $e(o>)$ equal to the square root of the noise power spectral density. Thus, for the generator resistance R_g , the mean-square noise voltage across R_g is $e^2 = 4kT R_g A f$. If we neglect circuit losses compared with the diode resistance R_s (this is a reasonably good approximation), the only other noise originating in the signal circuit is that generated in R_s across which a mean-square noise voltage $e^2 = 4kT_s R_s A f$ at frequency ω_a exists. We denote the equivalent noise temperature of the diode at frequency ω_s by T_s . The noise at frequency ω_s in the signal circuit is amplified and converted to noise at the output frequency ω_o .

The amplified output noise at midband is given by (replace V_j by $\frac{J}{V} \sim$ in Fig. 11.6)

$$P_i = U_{i1}X = \frac{R_f(e^2 + e^2)MW_0C^*(l - M^2)'}{[M^2 + (R_g + R_s)(R_L + R_s)a\cos_l(l - M^2)^2 C^*]}$$

where the noise current I_n is obtained from (11.26), with V_n replaced by $(e^2 + e^2)^{1/2}$. If we introduce the midband gain G_0 , we obtain [see (11.27)]

$$\frac{e^2 + e^2}{G_0} = \frac{kA_f(R_J + R_sT_s)}{R_s} G_0 \quad (11.48)$$

There is also noise generated in the output circuit at frequency $u >$ in the resistance R_s . If the diode noise temperature at the frequency ω_0 is T an equivalent noise voltage $e_j = 4kT_0R_sA_f$ appears across R_s in the output circuit. When we solve for the noise current I_2 that results and compute $P_2 = |I_2|^2R_s$, we obtain

$$\begin{aligned} P_2 &= \frac{w^*(l-M^2)C^*(R_s + R_s)^2G_0el}{4R_sM^2} \\ &= \frac{\cos^2(l - M^2)C^2/R_s + R_jkT_0R_s A_f G_0}{EM^2} \end{aligned} \quad (11.49)$$

The total output noise power at frequency w_0 is $P_n = P_i + P_2$. The available input noise power from R_n is $kT_A f$. Hence the noise figure is given by ($T = 290^\circ$ in the definition for F)

$$\frac{P_i + P_2}{G_0kT_A f} = \frac{R_sT_s}{R_gT} \frac{\llbracket l-M^2 \rrbracket C^2(R_s + R_g)R_sT_0}{R_gM^2 T} \quad (11.5)$$

Note that the gain G_0 was defined as output power divided by the available power from the input signal source, so that F as evaluated above conforms to the accepted definition of noise figure. If we take the diode noise temperature at the two frequencies w_s and w_0 equal to T_d , we obtain

$$F = \frac{1}{R_g} + \frac{R_s}{R_g} \frac{T_d}{T} \left(1 + \frac{(R_s + R_g)}{M^2 Q^2 R_s T_0} \right) \quad (11.5D)$$

after introducing the diode Q from (11.29). For maximum gain,

$$R_s = R_L = R_s \left(1 + \frac{M^2 Q^2}{R_s} \right)^{1/2}$$

from (11.30). Thus the noise figure under maximum-gain conditions is given by

$$F = l + \frac{T_d'}{T_x} \left(1 + \frac{\alpha_s}{M^2 Q^2} \right) \frac{V_{\text{ff}}}{V_{\text{in}}} + \frac{(\mathbf{i})}{M^2 Q^2} + \frac{[\mathbf{i} \mathbf{M} \mathbf{V}^{\text{ff}}]}{M^2 Q^2}$$

$$= \mathbf{i} + Y(l + \frac{\alpha_s}{M^2 Q^2}) \quad (11-52)$$

As a typical example consider $T_d = T$, $\alpha_s = 10W_{\text{in}}$, and $MQ = 10$. We then find that $F = 1.36$, or 1.3 dB. This example clearly demonstrates the low-noise property of the parametric amplifier.

The diode noise temperature T_d is somewhat greater than the ambient temperature T because of shot noise that arises from random motion of the carriers across the junction. In addition, the resistance R_s is usually not exactly the same at the two frequencies. However, this difference can be taken into account by choosing appropriate values of T_o and T_{id} . The thermal noise arising from circuit losses can be included in (11.50) very simply by replacing $R_s T_s$ by $R_u T + R_s T_s$ and $R_s T_o$ by $R_u T + R_s T_o$, where R_u and R_{s2} represent the equivalent-circuit resistances. However, since the effective Q of most diodes, that is, $[(1 - M^2)\alpha_s C_s R_s]^{-1}$ would rarely exceed 50 whereas circuit Q 's would normally be 1,000 or more, the resistances R_u and R_{s2} are negligible compared with the diode resistance R_s .

The noise properties of negative-resistance parametric amplifiers have been analyzed and measured by Uenoharat and others. The noise theory for a negative-resistance parametric amplifier employing a circulator, and with the output signal taken at the same frequency α_s as the input, is presented below.

With reference to Fig. 11.10, the following sources of noise are the main ones that need to be considered: (1) input noise from $R_g = Z_c$ at temperature T_x and frequency ω_1 ; (2) input noise at the idler frequency $\omega_2 = \omega_{p2} - \omega_1$ that arises in i_2 ; (3) noise arising in the diode resistance R_s at frequency ω_2 (equivalent noise temperature T_s); and (4) noise arising in R_s at the idler frequency ω_2 (equivalent noise temperature T_i). It is important to consider noise sources at the idler frequency because these noise signals are converted into noise at the frequency ω_1 by frequency mixing that takes place in the diode. The equivalent circuit and noise voltage sources are shown in Fig. 11.10.

¹M. Uenohara, Noise Considerations of the Variable Capacitance Parametric Amplifier, *Proc. IRE*, vol. 48, pp. 169-179, February, 1960.

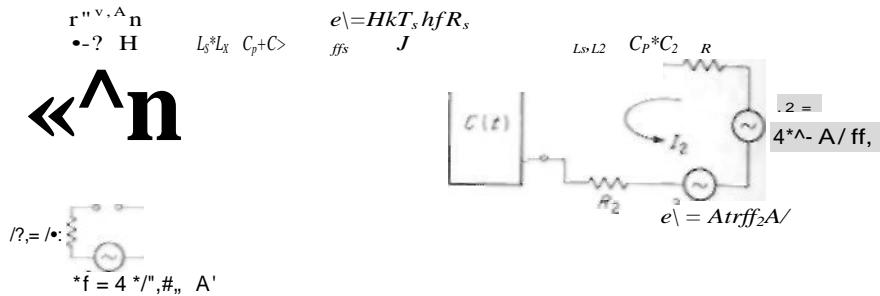


FIGURE 11.10
Noise circuit for a negative-resistance parametric amplifier.

The noise-power input from R_g is $kT_x A f$. This will appear at the output load R_L as amplified noise of amount

$$P_x = G_d kT_x t S f = \frac{4R_j kT_x A f}{(R_L + R_s) R} \quad (11.53)$$

where G_0 was obtained from (11.39) for the high-gain case.

The noise contributed by the amplifier is represented by the voltage sources e_2 , e_3 , and e_4 . Since the three noise sources are uncorrelated, the noise powers add. Hence the effective noise voltage in the idler circuit is $e_i = e_2^2 + e_3^2 + e_4^2$. The equations describing the noise currents I_1 and I_2 at midband are obtained from (11.34). Thus

$$e_2 = (R_L + R_s) I_{1-} - \frac{j M I^*}{0,; (1 - M^2) C_0} \quad (11.54a)$$

$$e_3 = (R_2 + R_s) I_{2-} - \frac{m t}{w_s (1 - M^2) C_0} \quad (11.54b)$$

since the loading R_x of the signal circuit as seen by the source e_2 is $Z_c = R_L$.

When we solve (11.54) for I_v we can evaluate the noise power

$$P_2 = |h?Ri|$$

delivered to the external load R_L . Giving the final results, we find that

$$\frac{-e_2}{R_L + R_s - R} + \frac{J M e^*}{0,; (1 - M^2) C_0 (R_L + R_s - R) (R_2 + R) * \bullet}$$

We must find the noise power contributed by e_2 and e_3 separately,

noise voltages do not add. From e_2 we obtain

$$P(e_2) = \frac{e^2 R_s M^2}{(R_L + R_s)^2}$$

and from e_1 we obtain

$$\begin{aligned} P(e_1) &= \frac{e^2 R_s M^2}{[<, (1 - M^2) C_0 (R_L + R_s - fL)(H_2 + 8,)]} \\ &\quad efM_L Mm_a \\ &\quad * > / ?^* + fL \cdot A H K_2 + \ll J \end{aligned}$$

Hence

$$P_2 = P(e_2) + P(e_1) = \frac{G_0}{4R_L} \left(e_2^2 + \frac{\omega_s}{\omega_i} \frac{e_i^2 R}{R_2 + R_s} \right) \\ \frac{G_0 k \cdot f}{R_L} \left[R.T. + \frac{\omega_s}{\omega_i} \frac{R}{R_2 + R_s} (R_s T_i + R_i T_s) \right] \quad (11.55)$$

If we assume that $T_a = T_s = T = T_d$ we obtain the simplified expression

$$P_2 = G_0 k \cdot f T_d \cdot \frac{R_K}{l} + \frac{co_s}{10} \frac{R}{R_L} \quad (11.56)$$

The corresponding noise figure F is given by

$$F = \frac{R}{kT_d f G_0} = 1 + \frac{R}{T_d R_L} \quad (11.57)$$

where T_d must be taken to be 290 K for the standard definition of F . In (11.57) R may be replaced by $R_s + R_i$ since this is the requirement for high gain. We then obtain

$$F = 1 + \frac{T_d (w_p R_L)}{d} + \frac{R_s}{R_L} \quad (11.58)$$

As an example, assume $T_d = T_s$, choose $w_p = 1.5\omega_t$, and let ω_t be much greater than jR_s . We then obtain a noise figure of 1.5, or 1.76 dB.

The foregoing analysis is valid only when the two frequencies w_s and w_i are spaced by an amount greater than the passband of the signal and idler circuits. When this is not the case, $\omega_s \sim \omega_i$, and the amplifier is a degenerate negative-resistance amplifier. Noise current from the input source resistance R_e at the frequency ω_i will now exist in the signal circuit. Similarly, noise currents at the frequencies w_s and w_i arising from R_s will be present in both the signal and idler circuits. In the degenerate amplifier the signal and idler cavities are identical and the equivalent circuit is that shown in Fig. 11.11. If the amplifier passband is symmetrical about $0^\circ/2$,

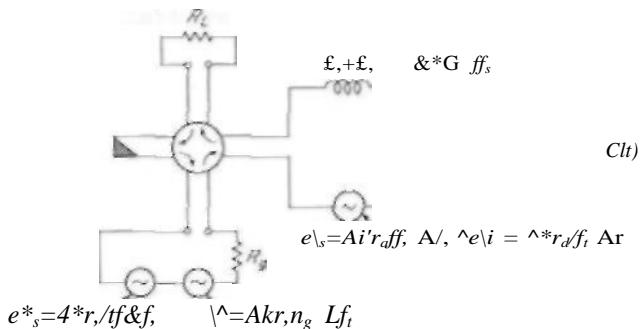


FIGURE 11.11
Noise circuit for a degenerate parametric amplifier.

noise at the frequency ω , $\omega_p/2$ is amplified to give noise output at the frequency ω_{s_1} , and also converted and amplified to give noise-power output at the frequency $\omega_s = \omega_p - \omega_{s_1}$. The two frequencies ω_{s_1} and ω_{s_2} are symmetrically located about $\omega_p/2$ and fall within the amplifier passband. Therefore we must consider the noise from R_g and R_s at the two frequencies ω_{s_1} and ω_{s_2} as separate uncorrelated noise in evaluating the total output noise power. The present situation is illustrated schematically in Fig. 11.12. The passband is split into two equal parts A/ω , and $A\omega/2$, located symmetrically on either side of $\omega_p/2$. Noise in the band A/ω , is amplified to give noise in the same band $A\omega/2$ at the output. In addition, noise in the band $A\omega/2$ is amplified and converted into noise in the band A/ω at the output. The lower sideband A/ω , will be regarded as the signal sideband.

Since $\omega_{s_1} = \omega_{s_2}$, the noise power in the band A/ω , arising from R_g and R_s produces the same amount of noise power in the output bands A/ω and $A\omega/2$ as does the noise power arising from R_g and R_s in the band A/ω . Hence we need to consider only the noise in the lower band A/ω in detail.

The noise from R_g in the band A/ω is represented by an equivalent voltage source e_u in Fig. 11.11. To evaluate the noise power delivered to R_i from e_{s_1} , we shall make the approximation of taking the series impedance of the circuit as $R_s + R_L$ throughout the passband. We denote the noise currents at frequencies ω_{s_1} and ω_{s_2} by I_1 and I_2 .

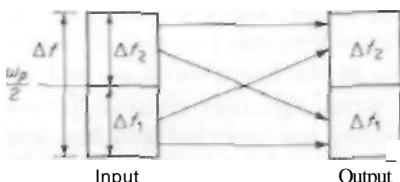


FIGURE 11.12
Illustration of noise conversion.

At the signal frequency ω_s the available gain is $G_0 = |\Gamma|^2$. Under high-gain conditions $R_L + R_s = R$ so (11.39) may be used for G_0 . The available input noise power from e_u is $e_u^2/R_g = e_u^2/4R_L$, since we assume that $R = R_L = Z_c$. The noise power P_u delivered to R , in the band A/\prime , is thus

When there is no impressed voltage at the frequency ω_{f_x} the currents I_s and I_a are related by [see (11.34) with $R_2 = R_L$]

$$(R_s + R)I_s = \frac{jMI}{j - j}$$

The power P_h delivered to $i\prime$, in the band A/\prime_2 is given by

$$P_h \ll I_s/R_L = \frac{fpP_u}{U_{il}} = \frac{R}{w + H_s} P_u \quad (11.60)$$

since $fL = M^2/[w^2(1 - M^2)^2 C_0^2 (ft. + H_g)]$ when $w = 0$.

The source e_u delivers an amount of power equal to P_{ls} to $i\prime_L$ in the band A/\prime_2 and an amount equal to P_u into the band $A f_x$. Hence the total noise power delivered to R_L from R is the same in the two bands A/\prime , and $A f_x$, and is given by

$$Pi = Pu + P_u = G_0 A A 11 + R_L^R + R) = 2G_0 A f_x = G_0 A f_x \quad (11.61)$$

since $R = R_L + R_s$ and $2 A f_x = A f_x$.

To evaluate the noise power delivered to R_L from R_s , we consider the noise in the band A/\prime , first. The equivalent voltage source is e_{2s} , and the circuit equations are (we can put $w_i = w_s$)

$$fci - (Ji_A + *)/ - , - \frac{t^1}{w, (1 - M^2)C_0} \quad (11.62o)$$

$$O = (* A + *)/ - \frac{jfft}{2} \quad (11.626)$$

$$\ll J, (1 - M)C_0$$

The output noise in the two bands $A f_x$ and $A f_2$ is given by

$$p_{2s} = \sqrt{fR_L} \quad p_{2l} = u_j \gg j r_x$$

When we solve for J , and I_2 , we obtain

$$p_{2s} = < * _t + * . - * > * - \wedge - \ll ; \quad (11.63)$$

The output noise from e_{2s} in the band $A f_2$ is given by

$$\frac{R}{2i} \sim P g^* =$$

The noise source e_2 , contributes a noise power P_{2s} in the band $\setminus r$ and P_{2l} in the band $A /"$. Hence the total noise power delivered to $7? f_r$ the internal amplifier noise sources is (in each band $A f_i$ and $A f$)

$$P_z = P^* + P_{2l} = \frac{2kT_d A f R_s G_0}{RL} + \frac{kT_r R f i t i f}{RL}; \quad (n.65)$$

The total output noise from both R_g and R_s is

$$P_{ss} = P_i + P_2 = kA f G_0 T_r + \frac{J_o}{RL}; \quad (11.66)$$

If the degenerate amplifier is used as a single-sideband amplifier (signal input ix the lower band $A /"$, only), the single-sideband noise figure F_{SSB} is defined by the ratio of the total output noise power in the band $A f = 2 A f_i$ divided by the available input noise power in the signal band $A /"$. Thus

$$\hat{F}_{SSB} = \frac{kA f_i G_a T_l}{PL} = 2^{-1} T_y R_L^5; \quad (H-67)$$

It is seen that, for single-sideband operation, the noise figure cannot be less than 2, or 3 dB. The signal-to-noise power ratio in the signal band $A f_Y$ at the input is $P_j k A /"$, $7 \setminus$). The signal-to-noise power ratio in the band $A /"$, at the output, given by $G_0 P_s / P_n$, is worse by a factor equal to the single-sideband noise figure F_{SSB} . ^{noise} degradation is due to noise entering in the idler band $A f_2$, in which no signal is present.

For double-sideband operation input signal power is present *in hot* bands $A f_x$ and $A f_2$. In this case the available input noise power is taken to be that in the band $A f$. Hence the double-sideband noise figure F_{DSB} is

$$F_{DSB} = \frac{kA f G_o T_r}{kA f G_J} = 1 + \frac{T_d}{T_r} \frac{R_s}{R_g} + \frac{T_d}{T_r} \frac{R_s}{R_g}; \quad (11.68)$$

and is a factor of 2 (3 dB) better than for the single-sideband case.

The double-sideband noise figure has been measured by Uenohara^{ab o g II} a number of different diodes. It is found that the theory given ^{ab o g II} reasonably accurate. Typical noise figures that were measured ranged 0.9 to 4.5 dB. For diodes with $w_s C_0 R_s$ less than 0.1, the noise fig^{ur*} $J^{w \Delta g e}$ dB or better. There was a strong correlation between the measured i^* figure and the diode quality factor Q . This is predicted by the theory as »

Since $R_L + R_s = R$ for high gain and R can be expressed by

from (11.29) and (11.38), we see that the optimum value of $R_L + R_s$ is MQR_X . The factor R^2/R_L in the expression for noise figure may now be replaced by $(MQ - 1)^{1/2}$ to give

$$\text{^DSB} = \frac{1}{T} + \frac{1}{MQ - 1}$$

showing that the noise figure improves with diode quality factor Q .

Parametric amplifier noise is primarily thermal noise in the diode resistance R_s . The equivalent amplifier noise temperature is given by $T_A = (F - 1)T_h$ where $T_h = 290$ K [for the degenerate amplifier with single-sideband operation $T_A = (F - 2)T_h$]. By cooling the amplifier to liquid-nitrogen temperature, noise temperatures below 100 K have been obtained.

- 11.1. Consider a square-law mixer for which the output current $U(t) = k[v(t)]^2$, where $v(t)$ is the applied voltage and k is a constant. Let a local-oscillator signal $V_p \cos i\omega_p t$ and a signal $V_s \cos coj$ be applied, with $V_s \ll V_p$. Show that the output current at the sum or difference frequencies to $\pm m_x$ is a linear function of V_s when $V_s < K$ V. Thus the square-law mixer is a linear converter for small-signal amplitudes.
- 11.2. Consider a parallel-plate capacitor with capacitance C_0 . Let a voltage $V = V_s \cos \omega t$ be applied. At time $t = 0$ the plate separation is suddenly increased to change the capacitance from C_0 to $C = C_0 / (1 + AC)$. Since the charge cannot change instantaneously, the voltage must increase. At time $t = (4f_s)^{-1}$, when $V = 0$ and $Q = 0$, let the plate separation be brought back to its original value. There is no change in V produced since $V = 0$ at this time. When $t = (2f_s)^{-1}$, let the capacitance be suddenly decreased to a value C again. When this process is continued, the resultant voltage across C is amplified and will have the waveform illustrated. This is an example of a linear capacitance varied at a rate twice that of the signal frequency. Evaluate the incremental change in voltage that occurs every half cycle and the power supplied by the pump. To evaluate the latter, determine the change in stored energy that occurs every time C is suddenly decreased.

FIGURE PI 1.2

- 11.3. A down-converter is a parametric amplifier with an input signal frequency $w_{in} = a\cos\omega_p + b\sin\omega_p$, and the output signal taken at frequency $\omega_{out} = r\cos\omega_p - s\sin\omega_p$. The circuit of Fig. 11.6. show that the down-converier gain (actually ≈ 1) is given by (11.27), with a, b , and r, s , interchanged.
- 11.4. Derive (11.32).
- 11.5. Derive the expression (11.35) for the gain of the negative-resistance parametric amplifier.
- 11.6. Derive (11.49) For the noise power P_2 .
- 11.7. Obtain an expression for the equivalent noise temperature of a parametric up-converter.
- 11.8. Derive an expression for the gain of the negative-resistance degenerate parametric amplifier illustrated in Fig. 11.11.
- 11.9. Consider the degenerate parametric amplifier with circulator shown in Fig. 11.11. Assume an input generator voltage V , at frequency w , in place of e_{in} and e_{out} . The generator sends a wave with current I_{in} into the amplifier where I_{in} must be equal to $V/(4R_g + Z_c) = V_H/2Z_{in}$, when $R_g = Z_{in}$, since V_H sees a matched load. A reflected wave is set up with a current $I_{refl} = -I_{in}$. The load current in R_L is I_{out} , apart from a phase angle. The total amplifier current at frequency w , is $I_{out} = I_{in} + I_{refl}$. With this information determine the appropriate circuit equations, analogous to (11.34), for I_{in} and I_{out} .
- 11.10. A parametric diode has the following parameter values: $C_0 = 2 \text{ pF}$, $R = 1 \text{ ft}$. The modulation index $M = 0.25$. The frequency $f_s = 5,000 \text{ MHz}$ and $f = 12,000 \text{ MHz}$. Evaluate the diode effective Q . Determine the load resistance R_L to give a gain of 20 dB for a negative-resistance amplifier of the form shown in Fig. 11.8. Assume $R_g = R_L = R_2$. Calculate R for 20-dB gain.

REFERENCES

- Blackwell, L. A., and K. L. Kotzebue: "Semiconductor-Diode Parametric Amplifiers," Prentice-Hall Inc., Englewood Cliffs, N.J., 1961.
- Penfield, P., and R. P. Rafuse: "Varactor Applications," The M.I.T. Press, Cambridge, Mass., 1962.
- Chang, K. K. N.: "Parametric and Tunnel Diodes." Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.
- Cullen, A. L.: A Traveling Wave Parametric Amplifier, *Nature*, vol. 181, February, 1968.
- Honey, R. C., and E. M. T. Jones: A Wide-Band UHF Traveling Wave Variable Reactance Amplifier, *IRE Trans.*, vol. MTT-8, pp. 351-361, May, 1960.
- Heilmeier, G. H.: An Analysis of Parametric Amplification in Periodically Loaded Transmission Lines. *RCA Rev.*, vol. 20, pp. 442-454, September, 1959.

Broadbanding techniques

- Matthaei, G. L.: A Study of the Optimum Design of Wideband Parametric Amplifiers Up-Converters, *IRE Trans.*, vol. MTT-9, pp. 23-28, January, 1961.
- Gilden, M., and G. L. Matthaei: Practical Design and Performance of Nearly Optimum Band Degenerate Parametric Amplifiers, *IRE Trans.*, vol. MTT-9, pp. 484-490, No. 1961. Wjlie
bgl

CHAPTER 12

OSCILLATORS AND MIXERS

In this chapter we will examine the operating characteristics of two types of negative-resistance solid-state devices, namely, Gunn devices and IMPATT diodes. These devices are widely used for low-power oscillators in microwave and millimeter-wave transmitters. The Gunn device is useful as a local oscillator in receiver front ends. We will also discuss the use of bipolar and FET transistors in oscillators.

The last part of the chapter provides an introductory treatment of mixers. A mixer is a nonlinear device, very often a diode or several diodes in a bridge arrangement, that will cause the microwave signal and local-oscillator (LO) signal to mix to produce a translation of the signal spectrum to a lower frequency called the intermediate frequency (IF). All superheterodyne receivers use a mixer for this purpose. The advantage gained is that the amplification of the signal, before demodulation, is more easily accomplished at the lower fixed IF frequency. The IF amplifier also establishes the bandwidth of the system since selectivity in the high-frequency RF amplifier is usually low. The main purpose of the RF amplifier is to increase the signal amplitude to a level such that the mixer noise will not produce a significant degradation of the signal-to-noise ratio.

Apart from economic factors the requirements of an oscillator include:

1. Excellent frequency stability, i.e., negligible variation in the frequency of oscillations due to variations in temperature, power-supply voltage, and oscillator loading (load pulling).
2. Adequate power output for the intended use.
3. Low amplitude-, phase-, and frequency-modulation noise.

4. Variable tuning, including mechanical tuning and voltage control
5. Capability to be modulated in amplitude (AM), frequency (FM) or phase (PM).
6. Simple circuit requirements.

A variable-frequency oscillator (VFO) whose frequency is varied by means of an applied control voltage is called a voltage-controlled oscillator (VCO).

Good frequency stability and low noise are obtained by employing *high-Q* resonator in the oscillator circuit. The resonator should have a resonant frequency that is insensitive to variations in the ambient temperature, and this usually translates into low thermal expansion of the resonator. For critical applications the oscillator may be placed in a temperature-controlled oven.

The frequency of an oscillator will vary with the dc bias voltages that are applied. This effect is called oscillator pushing and can sometimes be used to advantage to fine-tune an oscillator over a narrow band of frequencies.

The frequency of oscillation is determined by the resonant frequency of the input and output networks. Consequently, any change in the impedance of the load connected to the oscillator will result in a change in the oscillator frequency. This effect, which is referred to as oscillator load pulling, is usually undesirable. The pulling effect can be minimized by using loose coupling between the oscillator and the load, i.e., the external Q should be large. A disadvantage of using a loosely coupled load impedance is that the output power will be smaller and the oscillator efficiency will be reduced. Load pulling can also be minimized by using a very high Q resonator as the main frequency-determining element in the oscillator circuit.

Oscillator pulling can be determined by measuring the change in oscillator frequency as a function of the phase angle of the load reflection coefficient.

12.1 GUNN OSCILLATORS

Some bulk semiconductor materials, such as gallium arsenide (GaAs), dium phosphide (InP), and cadmium telluride (CdTe), have two closely spaced energy bands in the conduction band. A typical energy versus momentum band structure is shown in Fig. 12.1. At low electric field strengths in the material, most of the electrons will be located in the lower-energy band. At high electric field strengths, most of the electrons be transferred into the high-energy band. In the high-energy band effective electron mass is larger and hence the electron mobility is than what it is in the low-energy band. Since the conductivity is proportional to the mobility, there is an intermediate range of electric

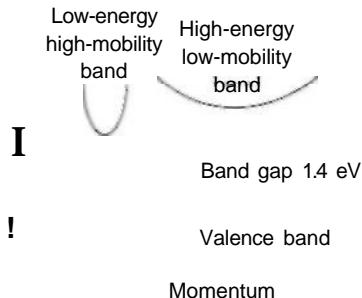


FIGURE 12.1

Typical double-energy conduction band for a Gunn material such as GaAs.

strengths for which the fraction of electrons that are transferred into the high-energy low-mobility conduction band is such that the average mobility, and hence conductivity, decreases with an increase in electric field strength. Thus there is a range of applied voltages over which the current decreases with increasing voltage and a negative incremental resistance is displayed by the device. A typical current-voltage characteristic for a Gunn device is shown in Fig. 12.2. A Gunn device is also called a transferred-electron device since the negative resistance arises from the transfer of electrons from the low- to the high-energy band.

The oscillations that occur in materials with the energy band structure noted above was discovered by J. B. Gunn. The possibility of obtaining negative differential resistance had been predicted earlier by Ridley and Watkins.

There are two principal modes of operation that result in oscillations for a Gunn device. When the applied voltage exceeds the threshold value, a dipole domain (a region of electron concentration and depletion) forms near the cathode end with most of the voltage drop appearing across the high-resistance part of the device. A short section of the input region is in the low-energy high-mobility state and electrons leave the cathode with a large

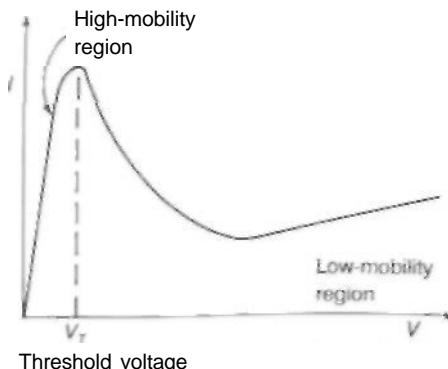


FIGURE 12.2

Current-voltage characteristics for a Gunn device. Note the negative-resistance region-

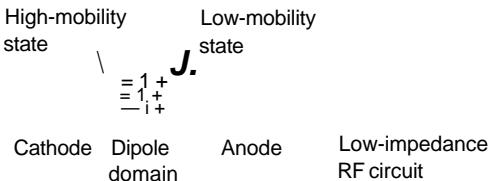


FIGURE 12.3

A simple Gunn Oscillator circuit in the transit-time mode of oscillation.

velocity. At the point in the material that separates the high-mobility and low-mobility states, electrons accumulate on the left side and are depleted on the right side by virtue of the different mobilities. This dipole arrangement of charge is shown pictorially in Fig. 12.3. This dipole domain sweeps across the device, and when it arrives at the anode, the device is in a high-mobility state and a new dipole domain forms at the cathode end and moves toward the anode. This mechanism is self-repeating and represents an oscillation with a period equal to the transit time. This mode of oscillation has a low efficiency (a few percent) of power generation and a frequency that cannot be controlled by the external circuit. This mode of oscillation is called the transit-time mode or Gunn mode.

The second mode of oscillation is the limited-space-charge (LSA) mode. Operation of a Gunn oscillator in the LSA mode can produce several watts of power with efficiencies of around 20 percent or more. The power outputs that have been obtained decrease with frequency and are below 1 W at frequencies greater than 10 GHz. Output power of several milliwatts can be obtained at 100 GHz.

In the LSA mode the Gunn device is incorporated as part of a resonant circuit as shown in Fig. 12.4a. The frequency of the resonant circuit is adjusted so that it is several times greater than that of the transit-time mode. As a consequence, dipole domains do not have sufficient time to form and the device operates essentially as a negative-resistance device. The dc bias is adjusted to a value somewhat greater than the threshold voltage. The RF voltage of the oscillations will build up to a peak-peak value approximately equal to the voltage increment over which the device resistance is negative as shown in Fig. 12.4b. The resonator loading, represented by a resistor R , is adjusted to a value about 20 percent greater than the maximum negative resistance of the device. This will ensure that oscillations will start. The amplitude of the oscillations will build up until the average negative resistance of the Gunn device becomes equal to the resonator resistance R .

If the resonator frequency is adjusted to a value slightly ^{above} _{the} the transit-time mode, the Gunn device will operate *very much like* the basic Gunn mode, but the dipole domain will be quenched before it ^{arrives} at the anode by the negative-going oscillation voltage. This type of operation is called the *quenching mode*.

mode is called a quenched-domain mode. Oscillations can also occur by adjusting the resonator frequency, so that it is lower than the frequency of the Gunn mode. In this case the dipole domains have sufficient time to sweep across the device and arrive at the anode. However, the initiation of a new dipole domain is delayed until the oscillation voltage rises above the threshold value. This mode of operation is called the inhibited or delayed mode.

Circuits

The equivalent circuit of a Gunn device operating in the LSA mode is a negative resistance $-R_d$ in parallel with a capacitance C_d as shown in Fig. 12.5. The negative resistance has a value that typically lies in the range -5 to -20 ft. The required resistive loading from the cavity and the external load should be around 20 percent higher than the Gunn device resistance so that the parallel combination $-RR_d/(R - R_d)$ will be negative. The cavity used for the resonator must generally have an impedance-transforming property in order to reduce the high impedance of the output waveguide to

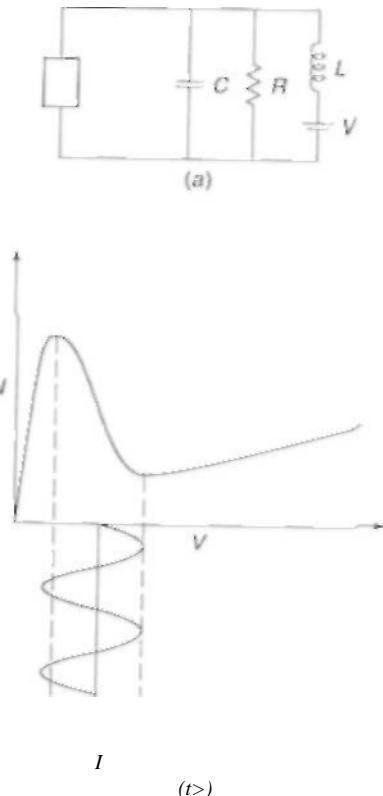


FIGURE 12.4

(a) A basic Gunn oscillator operating in the LSA mode; (b) RF oscillating voltage across a Gunn device operating in the LSA mode.

C,, ϵ - R ,

FIGURE 12.5

Equivalent circuit of a Gunn device operating in the LSA mode

M

the appropriate Jow value required by the *Gunn device*. A simple cavity structure is shown in Fig. 12.6. The Gunn device is located under a post in a rectangular waveguide. The cavity is resonated at the desired frequency by adjusting the short-circuit position. The degree of coupling to the external waveguide is adjusted by changing the window opening in the inductive diaphragm located at the front of the cavity. The top of the post is insulated from the waveguide. The dc bias voltage (typically around 12 V) is applied to the post. There is sufficient capacity between the post and the surrounding waveguide to provide an adequate low-impedance RF bypass capacitance and thus RF currents do not flow through the bias voltage supply. Fine tuning of the cavity can be obtained by means of a tuning screw.

Another simple cavity arrangement for a Gunn oscillator is shown in Fig. 12.7. In this cavity the high impedance of the waveguide is transformed into a low impedance at the location of the Gunn device by means of quarter-wave transformers. The cavity resonant frequency can be adjusted by changing the location of the short circuit. A tuning screw can be used for fine tuning of the cavity.

The cavity shown in Fig. 12.6 is easily modified to have two posts, one for mounting the Gunn device and a second one for mounting a varactor diode. The capacitance of the varactor diode is a function of the control voltage V_c . By varying V_c the resonant frequency of the cavity can be varied.

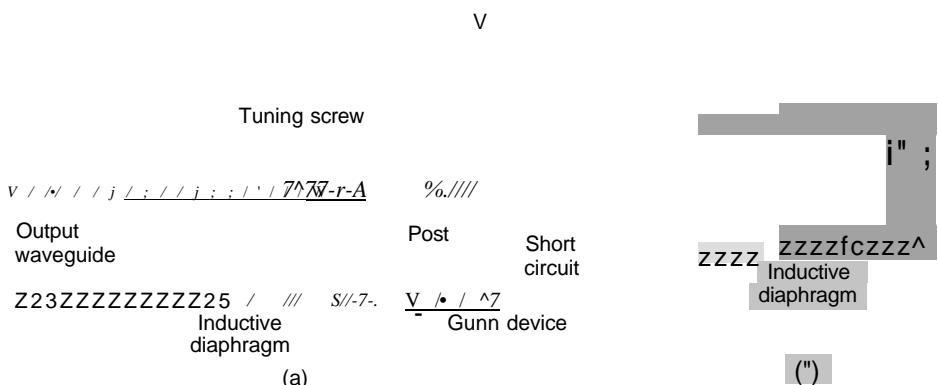


FIGURE 12.6

(a) A simple waveguide cavity for a Gunn oscillator; (b) inductive diaphragm used to coupling between the output waveguide and cavity.

j - ^ t the

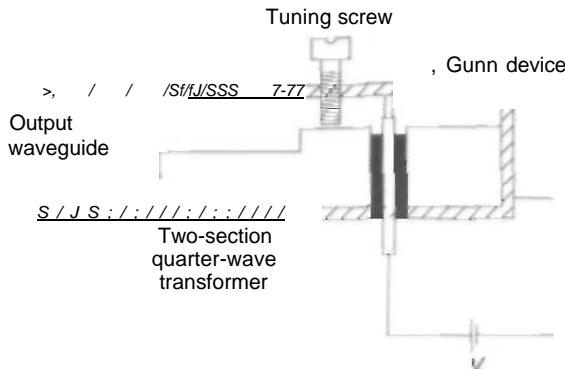


FIGURE 12.7

A Gunn oscillator cavity which uses a two-section quarter-wave transformer to transform the high impedance of the waveguide to a low impedance at the Gunn device.

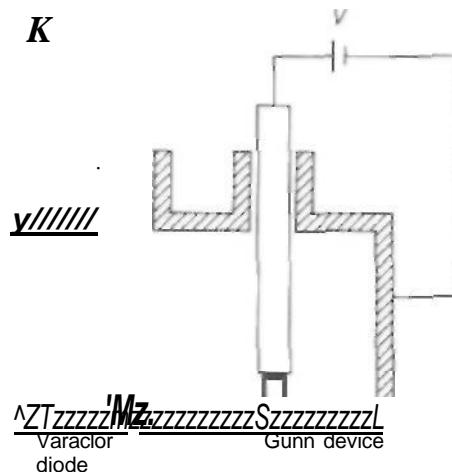


FIGURE 12.8

A Gunn oscillator cavity which has a post-mounted varactor diode for cavity tuning. A sawtooth sweep voltage applied to the varactor diode will produce frequency modulation of the oscillator.

If a sawtooth sweep voltage is applied to the varactor diode, the Gunn oscillator will be frequency-modulated. The cross section of the cavity is shown in Fig. 12.8.

The Gunn device can be operated as a pulsed oscillator by applying the dc bias voltage in the form of a pulse train of short rectangular pulses. If the duty cycle is low enough and very short bias pulses are used, the peak power output will be limited only by the peak current, since thermal heating of the device will be negligible during the short on time. For pulsed oscillator applications, the IMPATT diode, or variations of it, is preferred because of higher output power. The Gunn device can also be, and has been, used as a negative-resistance amplifier.

DIODES

The acronym IMPATT stands for *IMPact Ionization Avalanche Transit Time* and describes the phenomenon associated with reverse voltage breakdown in a *p-n* junction diode and the transport or transit of charge carriers



FIGURE 12.9

Structure of a Read (IMPATT) diode and the electric field profile across the diode.

through a drift region. W. T. Read had proposed in 1958 that there should be a phase delay of more than 90° between an applied RF voltage and the avalanching current if the RF voltage caused the total voltage to exceed the reverse breakdown voltage in a diode. In 1965, R. L. Johnson verified Read's prediction. When the current lags the RF voltage by more than 90° , the diode will exhibit a negative resistance and can be used as a source of microwave power in an oscillator circuit.

The Read diode, designated as a $p^*n\bar{n}^*$ type, consists of a heavily doped p^* region, a normally doped n region, an undoped or intrinsic semiconductor section, and an n^+ region as shown in Fig. 12.9. The $p\sim n$ diode junction will break down when the applied reverse bias voltage exceeds a threshold value. The current-voltage characteristic is shown in Fig. 12.10 and is similar to that in the familiar zener diode. The rapid increase in current at the breakdown voltage is caused by avalanche multiplication of the density of the holes and electrons. If a Read diode is placed in a cavity and a reverse bias somewhat smaller than the breakdown voltage is applied, along with a small RF voltage, then breakdown will occur when the RF voltage becomes positive. When breakdown is initiated a large number of holes and electrons are created at the p^*n junction. The electrons are swept across the n region into the intrinsic semiconductor drift region. After a transit-time delay, the electrons are collected at the n terminal. The current pulse moves through the diode from right to left. When the time for avalanche charge buildup plus that for charge transit through the drift region exceeds one-half RF period, the output current will lag the RF voltage by more than 90° . With these conditions the diode will exhibit a negative resistance for RF currents. In an oscillator circuit the initial RF voltage comes from the cavity resonant-frequency component. Once the oscillations start they grow in amplitude until the average negative resistance of the diode becomes equal to the total equivalent resistance of the cavity and the external load.

Reverse voltage
breakdown

FIGURE 12.10
Current-voltage characteristic of a diode.
Note the reverse voltage breakdown re-

Since the introduction of the Read diode as a generator of microwave power, a number of other diode structures have been developed that will also produce microwave oscillations. The two most common variations of the Read diode are the BARRITT (BARrier /njection Transit-Time) diode and the TRAPATT (Ti?A pped Plasma Avalanche Triggered Transit) diode. The BARRITT diode is a p^*tip' or back-to-back diode. The charge carriers that traverse the drift region in a BARRITT device consist of minority carriers that are injected from the forward-biased p^+n junctions. Since the BARRITT diode does not involve an avalanche breakdown, it produces less noise than an IMPATT diode does. However, the power output and efficiency is less.

The TRAP ATT diode is a $p'nn^*$ diode and is driven by a large repetitive pulse of current. Breakdown *will* occur at one of the $p'n$ diode junctions, and since the current drive is very large, a large collection of electrons and holes (a plasma) is generated. The violent breakdown creates a high electric field shock front that moves across the n-type drift region. After passage of the shock front, the plasma is located in a low-field region and is said to be trapped because it takes a long time to clear the drift region of charge carriers. When the plasma has been cleared from the drift region, the cycle will repeat. In a TRAPATT diode, oscillations start with the diode operating as an IMPATT device. When the amplitude of oscillation becomes large enough, the TRAPATT mode of oscillation is established. The TRAPATT diode will not operate at as high a frequency as the IMPATT diode does. It is also noisier, so its use is decreasing.

IMPATT diode oscillators are used with the same cavity structures that are used with Gunn devices. IMPATT diodes are much noisier than

Gunn devices and so are generally not used for local oscillators E crowave and millimeter-wave receivers. For power generation, IMPATT diodes are superior to Gunn devices. Output CW powers of as much as 10 W at a few gigahertz and around 1 W at 100 GHz can be obtained from a sin 1 device. At frequencies above 100 GHz, the output power from currentJ produced IMPATT diodes decreases approximately as $1/f^3$.

12.3 TRANSISTOR OSCILLATORS

Silicon bipolar transistors are a good choice for oscillators at frequencies UD to 5 GHz. From 5 GHz up to about 40-GHz, MESFET devices can be used in oscillator circuits. In the frequency range 30 GHz up to around 100 GHz the high-electron-mobility transistor (HEMT) would be used because of it's higher frequency of oscillation.

In Chap. 10 we were concerned with the general problem of designing amplifiers that would not oscillate. We noted that many microwave transistors were only conditionally stable and only a restricted range of load and source impedances would ensure amplifier stability. The stable and unstable regions were shown graphically by plotting the input and output stability circles on the Smith chart. In the design of an oscillator, we choose the input and output port terminations in the unstable regions. In the unstable regions both the input and output impedances of the transistor circuit will have a negative resistance and oscillations will occur at a frequency at which the total reactance in the input and output circuits vanish (resonance condition). Transistor oscillators can thus be viewed as negative-resistance oscillators. If a transistor is absolutely stable, it can be made unstable by using feedback from the output to the input of the device. Common base or common gate circuit configurations tend to have the greatest amount of instability as Example 10.5 showed. One suitable feedback arrangement is the use of a reactance in series with the emitter or source lead. A series reactance can also be used in the common base or common gate circuits.

The criteria for a transistor terminated in an impedance Z_s at the input port and Z_L at the output port to oscillate are readily established. Initially, assume that a voltage source V_g acts in series with $Z_s + Z_{in}$. The input current to the transistor will be $V_g/(iZ_s + Z_{in}) = I_{in}$. If we now reduce V_g to zero but at the same time make $Z_s + Z_{in}$ tend to zero, we can maintain the current I_{in} and a finite output power. Thus, for oscillations occur when $V_g = 0$, we must have

$$Z_s = R_{in} + jX_{in} \quad (12.1a)$$

or

$$\frac{I_{in}}{V_g} = -\frac{Z_s}{Z_{in}} \quad ;$$

Thus the transistor must have a negative input resistance and the i P must be tuned to resonance. The frequency of oscillation is determined by (12.16). From the relations $f_s = (Z_s + D)/(Z_s + 1)$ and $T_m = (\#in) \backslash$

$(Z_{in} + 1)$, we readily find that the condition for oscillations can also be stated as

$$r_s r_m = i \quad (12.2)$$

Since R_m is negative $|r_{in}| > 1$, which is in accordance with our requirement for instability as given in Chap. 10.

We will now show that when (12.1) or (12.2) hold, at the output $Z_L = -Z_{<m}$, or

$$\hat{r}_{out} = -\%ot \quad (12.3a)$$

$$Mmt'' \sim JX_L \quad (12.36)$$

$$I'J'out = 1 \quad (12.3c)$$

Thus the conditions for oscillations are satisfied at both ports if they are satisfied at one port. The proof of this property is as follows: The following relations hold

$$= \underline{S_{22}} - \underline{M_i} - \underline{S_{22}} - \underline{Vr_{in}} = \underline{S_{22}r_{in}} - \underline{A} \\ \text{"* } \underline{i} - ' ; \underline{s_n} " \underline{i} - \underline{s_u} / \underline{r_{,n}} " \underline{r_{,n}} - \underline{s_u}$$

upon using (12.2). We also have

$$r_s = s_{,} - \&r_b$$

which can be solved for T , to give

$$h = \begin{matrix} \text{'in} \sim s_u \\ \hat{r}_{22'm} - A \quad r_{out} \end{matrix}$$

which is the relationship we wanted to prove.

Since the impedance looking into both ports has a negative real part, the transistor will deliver power to the external circuit at both ports. Since it is operating in the unstable region, the power gain is negative as explained in Sec. 10.7. An absolutely stable transistor with a suitable external feedback network can be viewed as a new modified potentially unstable device for which the above relations will hold true. We normally view an oscillator as an amplifier with a feedback loop that feeds a fraction of the output power to the transistor input and the input circuit. Part of the power fed back is absorbed in the resistance associated with the input circuit. When an unstable transistor is used as an oscillator, the feedback takes place internally and is described in terms of the reverse transmission coefficient $S_{,2}$. Thus the power dissipated by the input circuit as well as the input power that drives the oscillator is provided by internal feedback of power from the output to the input of the oscillator. The only significant difference in the two oscillator types is the feedback path which can be either external or internal,

In order for oscillations to start, it is necessary to choose $R_s < |R_m|$ using the small-signal scattering-matrix parameters to evaluate R_m . Oscil-

lations will then build up until the nonlinear characteristics of the traⁱtor cause the power gain to saturate. Thus steady-state oscillations wll cause the transistor to operate under large-signal conditions. For la signals the nonlinear behavior means that the large-signal scattering-mat ^ parameters will be different from those that apply for small-signal cond" tions. Furthermore, the nonlinear behavior will cause harmonics of th fundamental frequency to be generated. Usually the presence of harmonic in the output is undesirable unless the oscillator is specifically designed for an output at one of the harmonics. The basic problem in oscillator design is choosing the port terminations so that the desired output power is obtained the harmonics are adequately suppressed, and the desired frequency of oscillation with good stability against variations due to load, temperature and bias conditions is obtained.

Some general guidelines that should be followed in order to achieve the above objectives are:

1. The loaded Q of the output circuit should be at least 10 to give good harmonic suppression.
2. The input reactance jX_a should have a large-frequency derivative or slope dX_s/dw . This will mean that a small change in jX_{in} can then be matched by a small change in to to bring jX_s back to $-jX_m$. Usually this requirement is met by incorporating a high-Q temperature-stable resonator in the input circuit or in the feedback path.
3. For minimum harmonic generation the oscillator should operate as a class A oscillator.
4. For best efficiency and largest amount of output power, the oscillator should be designed for class B or class C operation. This will require a self-bias circuit, so that initially the circuit operates as a class A oscillator in order for oscillations to start.

In class A operation the collector (drain) current flows continuously over a complete RF cycle. In a class B oscillator the current flows for one-half of the RF cycle, while in class C operation the current flows for less than half of one period in each RF cycle. In class C operation the maximum collector (drain) current flows when the RF voltage at the collector (drain) is negative. Thus the power dissipation in the device is minimized and efficiency (RF output power/dc input power) can be high. At lower frequencies class C efficiencies as high as 80 percent or more can be achieved, tu microwave frequencies the efficiency is much lower because of circuit 1< relatively low power gain, and limited supply voltages to avoid damaging transistor. The efficiency of microwave oscillators and amplifiers is described in terms of the power-added efficiency. The power-added efficie is defined as follows:

$$\text{Power-added efficiency} = \frac{\text{RF output power} - \text{RF input power}}{\text{dc input power}} \quad , \quad , .4)$$

4 THREE-PORT DESCRIPTION OF A TRANSISTOR

In order to facilitate the analysis of an oscillator circuit, when an impedance is connected between one of the transistor terminals and the ground plane, it is useful to have a scattering-matrix description of a transistor viewed as a three-port circuit. In Fig. 12.11 we show a transistor with a microstrip line connected to each terminal. For clarity we have labeled the base terminal as port 1, the collector terminal as port 2, and the emitter terminal as port 3. Any other labeling could be used equally well. The normalized amplitudes of the incident and reflected voltage waves will be denoted by a_1, a_2, a_3 and b_1, b_2, b_3 , respectively. The three-port circuit is described by the scattering-matrix relation

$$\begin{vmatrix} \mathbf{6}, \\ b, \\ b, \end{vmatrix} = \begin{vmatrix} \mathbf{S}_{11}, & \mathbf{S}_{12}, & \mathbf{S}_{13}, \\ \mathbf{S}_{21}, & \mathbf{S}_{22}, & \mathbf{S}_{23}, \\ \mathbf{S}_{31}, & \mathbf{4.}, & \mathbf{SS} \end{vmatrix} \begin{matrix} a_1, \\ a_2, \\ a_3 \end{matrix} \quad (12.5)$$

where S_{ij} are the three-port scattering-matrix parameters. We will show that the three-port parameters have the property that the sum of all elements in any row or in any column equals unity; thus

$$i = 1, 2, 3 \quad (12.6a)$$

$$\sum_{j=1}^3 S_{ij} = 1 \quad i = 1, 2, 3 \quad (12.6b)$$

Consequently, the three-port scattering-matrix elements are not all independent. In fact, they can be expressed in terms of the two-port scattering-matrix parameters that describe the transistor when one of the terminals is grounded, e.g., common emitter parameters when the emitter is grounded.

In order to prove (12.6a) we note that, provided there is negligible stray capacitance between each transistor terminal and the ground plane, the sum of all currents entering the three-transistor terminals must be zero. Thus we have

$$\sum_{i=1}^3 (0, -\mathbf{6},) = 0 = \sum_{i=1}^3 L_i - \sum_{j=1}^3 S_{ij} a_j = 0$$

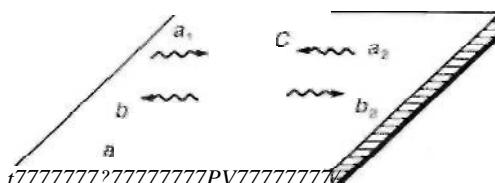


FIGURE 12.11
A transistor viewed as a three-port network.

We can choose the incident-wave amplitudes independently; so if we choose $a_2 = a_3 = 0$, we obtain

$$\langle\langle \cdot - E \rangle\rangle_{,3} = 0$$

or

$$\sum_i a_i S_{ii} = 1$$

which is one of the relations in (12.6a). By choosing in turn a_1 and a_3 as the nonzero amplitude, the other two relations are obtained.

The relations given by (12.66) are obtained by noting that, when all port (terminal) voltages are equal, the input current at each port will be zero, provided again that there is negligible capacitance between each terminal and the ground plane. Thus we must have $b_1 = a_1$, when $a_1 = a_2 = a_3 = 0$; consequently, $b_2 = b_3 = a$ and

$$\sum_i a_i S_{ij} = \sum_j a_j V_j = 0$$

which gives the desired result.

The total normalized voltage at port 3 is $V_3 = a_3 + b_3$. Let us define new voltage-wave amplitudes at ports 1 and 2 as follows:

$$V_r = \frac{V_1 - V_2}{2} \quad (12.7a)$$

$$V_f = V_1 - \frac{V_3}{2} = a_1 - \frac{a_3 + b_3}{2} \quad (12.7b)$$

$$V_s = V_2 - \frac{V_3}{2} = a_2 - \frac{a_3 + b_3}{2} \quad (12.7c)$$

$$V_r = 6, \quad V_f = 6, \quad V_s = 6 \quad (12.7d)$$

The total new voltages for ports 1 and 2 are now $V_x = V_f + V_s = a_1 + a_2 + b_3$, and $V_r = a_3 + b_3 - V_3$ and are thus referenced to the port A voltage. The above definitions for the new incident- and scattered-wave amplitudes were chosen so as to leave the port currents unchanged, since $V_f - V_s = a_1 - a_2$ and $V_2^+ - V_2^- = b_1 - b_2$. The common emitter (port 3 terminal) two-port scattering matrix is the scattering matrix relates V_r and V_s to V_f and V_s ; thus

$$V_f = S_{11} V_r + S_{12} V_s \quad (12.8a)$$

$$V_s = S_{21} V_r + S_{22} V_s \quad <^{12-86}$$

Upon expressing the V_r and V_s in terms of the a_i and b_i , using 1(12.7) we

obtain

$$6_2 - \mathbf{H} \mathbf{I}_1 \wedge = S_{21} a_1 + \frac{\mathbf{L}}{2} \wedge + S_{22} \mathbf{j}_a \wedge$$

Since the sum of all currents flowing into the three transistor terminals is zero (Kirchhoff's law), we must also have

$$0_1 - 6_1 + Q_2 - 6_2 + a_3 - 6_3 = 0$$

The above three equations are easily rearranged to give three equations expressing the b_i in terms of the a_i . The first two equations can be written in the form

$$*_i = -S_{nO} + 5.23:2 + f_f \mid V^*$$

$$b_i = S_{21}a_1 + S_{22}a_2 + "T> \wedge$$

where $r_u = 1 - S_{12}$ and $o''_{22} = 1 - S_{22} - S_{12}$. The third equation can be solved for b_3 to give $b_3 \sim a_1 + G_2 + a_3 - 6_1 - b_2$. When we substitute for 6_1 and 6_2 , we readily find that

$$b_3 = -\frac{2a_9}{4 - r_J} - \frac{2r_2}{4 - (T_z - cr)} + \frac{a}{4 - cr} - a_3 \quad (12.9a)$$

where $o_{12} = 1 - S_{12} - S_{21}$, $o'_{21} = 1 - S_{21} - S_{12}$, $ir = S_n + S_{12} + S_{21} + S_{22} = 2 - cr_{12} - (T_z = 2 - ir_u - IT_{...})$. By using this expression in the equations for 6_1 and 6_2 , we obtain

$$"T> = (a_1 + T \wedge) "T> + I_s \frac{IT_{...}; < T_{...}}{4} + Trf h + 5 \wedge 0 a \quad (12.9c)$$

Equations (12.9a) to (12.9c) provide the three-port scattering-matrix description of the transistor in terms of the two-port scattering-matrix parameters. The S_{ij} are given by

$$\begin{aligned} c_{21} - \frac{Q}{321} + \frac{IT_{12}(Ta)}{4 - a} & \quad c_{22} - \frac{o}{322}, \quad \frac{r_{TM}(r)*i}{4 - ff} \quad \frac{6}{a23} \quad \frac{2fr_{TM}}{4 - ff} \\ * & \quad \frac{2o''_{12}}{32} \quad \frac{3}{32} \quad \frac{2''_{21}}{4^7} \quad \frac{5}{a3} \quad \frac{o''}{4 - a} \\ ** & \quad \frac{-O}{4} \end{aligned}$$

where

$$\begin{aligned} a &= S_n + S_{l2} + S_{2i} + S_{22} & \langle r_u &= l - S_u - S_{l2} \\ tr_{l2} &= 1 - S_n - S_{2i} & \langle r_{22} &= 1 - S_{22} - S_{l2} \\ a_{2i} &= 1 - S_{22} - S_{l2} \end{aligned}$$

The reader can readily verify that the above scattering-matrix parameter¹⁷ satisfy the relations (12.6).

If the three-port scattering-matrix parameters have been measured it is easy to obtain expressions for the two-port scattering-matrix parameters. Let us assume that we want to find the common emitter two-port scattering parameters. This requires that we make $V_a = 0$ or $b_a = -a_3$. By setting $b_3 = -a_3$ in the last equation in the set given by (12.5), we can solve for a , in terms of a , and a_2 . By using this solution in the first two equations, we obtain the solutions for b_1 and b_2 in terms of a , and a_2 from which the two-port scattering-matrix parameters are readily identified. It is found that

$$\begin{aligned} \frac{\overset{\circ}{A}}{1 + \overset{\circ}{A} \overset{\circ}{A}^H} &= S_{11} - S_{12} - \frac{\overset{\circ}{S}_1 \overset{\circ}{S}_2}{\overset{\circ}{A} \overset{\circ}{A}^H \overset{\circ}{A} \overset{\circ}{A}^H} & (12.11) \\ S_{21} &= -S_{12} & \frac{\overset{\circ}{A} \overset{\circ}{A}^H}{1 + \overset{\circ}{A} \overset{\circ}{A}^H} &= S_{22} - S_{12} \\ \overset{\circ}{A} \overset{\circ}{A}^H &= 1 + S_{33} \end{aligned}$$

Consider now the case when a series normalized impedance Z_s is connected between the emitter and the ground_plane. For this situation, as shown in Fig. 12.12, $a_3 = T b_3$ where $T = (Z_s - 1)/(Z_s + 1)$. The same procedure used to obtain (12.11) can be used to find the two-port scattering matrix for the transistor with a series feedback impedance Z_s connected in series with the common emitter lead. It is readily found that the new scattering-matrix parameters are given by (12.11) upon replacing $S_{33} +$ by $S_{33} - r^{-1}$. As a final note we point out that, if the relations (12.10) are used to express the $\overset{\circ}{S}_u$ in terms of the S_{ij} (12.11) is satisfied identically. i.e., it gives $S_0 \cdot S_{ij}$

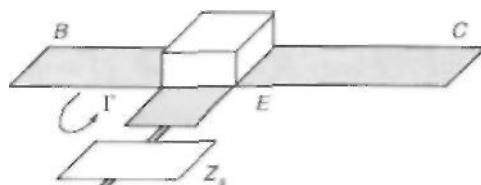


FIGURE 12.12
A transistor with a series feedback impedance in the common emitter lead.

The scattering parameters S_u and $S_{.,z}$ are given by

$$S_{II} = \frac{A_1}{A_2} T \quad (12.12fl)$$

$$A_2 r - 3_{22}$$

where $A_1 = S_{11}S_{33} - S_{13}S_{31}$ and $A_2 = S_{22}S_{33} - S_{23}S_{32}$. The above equations can be solved for f to give

$$r = \frac{c''}{\circ 33^\circ H \sim^a} \quad d2.13a)$$

$$r = * \frac{I}{\circ .S3 - 22 \sim^a S} \quad < 12-136)$$

If we restrict the series impedance to be a pure reactive element, then $|l''| = 1$. The circle of $|T| = 1$ values maps into circles in the S_u and S_{22} planes in accordance with the bilinear transformations given by (12.13). The center and radius of the S_u circle are

$$\text{Center} = \frac{S_{11} - Mf_3}{i - iW} \quad (12.14a)$$

$$\text{Radius} = \frac{|S_{13}S_{31}|}{II - IS33I} \quad (12.146)$$

while those for the $S_{.,z}$ circle are

$$\text{Center} = \frac{\& \sim A J?}{1 - I \& \sim A J?} \quad (12.15a)$$

$$\text{Radius} = \frac{\& \& I}{n - 1 \& \& 1} \quad (12.156)$$

By plotting these circles it is possible to visually see the range of values that can be obtained for S_{11} and S_{22} by varying the series feedback reactance. For each value of jX the stability parameter K for the equivalent two-port network can also be evaluated [see (10.18)]. This parameter is a useful measure of the degree of instability a series feedback impedance can produce. Some impedances, particularly resistive ones, will actually improve the stability of the device. Such impedances would be useful in amplifier design where instability is undesirable. Some caution should be exercised in using the above two-port to three-port relations since they are based on the assumption that the stray capacitance from each transistor terminal to the ground plane is negligible.

The above results show that a knowledge of the three-port scattering matrix parameters is very useful in determining the equivalent two-port scattering-matrix parameters with a series feedback impedance connected 'one common lead. The following example illustrates an application of the above results.

Example 12.1 Transistor with common source feedback impedance
A MESFET device has the following common source two-port scattering-matrix parameters at 10 GHz:

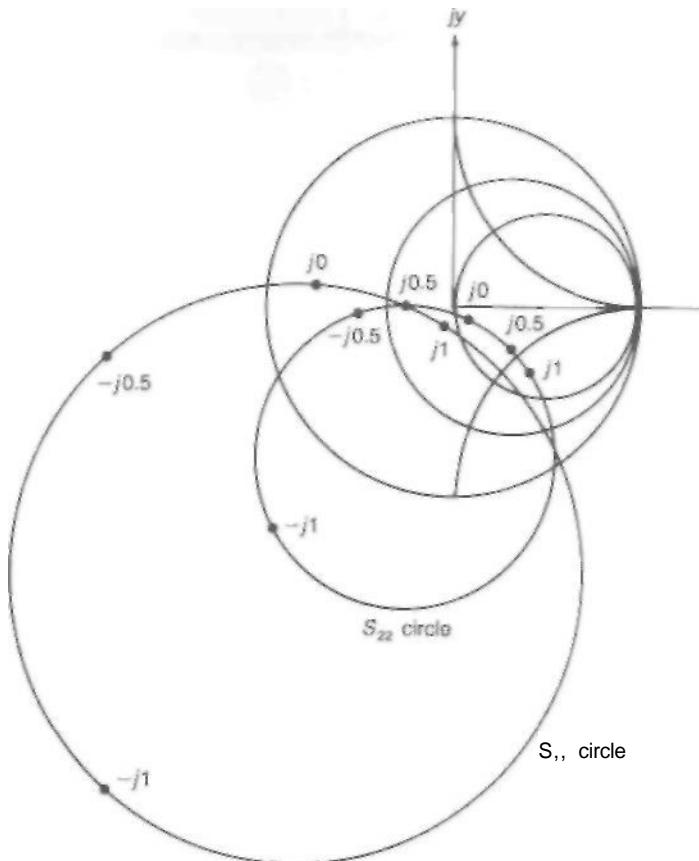
$$\begin{aligned} S_n &= 0.73/1172^\circ & S_{12} &= 0.093,139^\circ \\ S_{21} &= 2.31/144^\circ & S_n &= 0.09Z -55^\circ \end{aligned}$$

This device has a stability parameter $K = 1.13 > 1$ and is absolutely stable. By using (12.10) the following three-port scattering-matrix parameters were computed.

$$\begin{aligned} S_{,,} &\gg 0.735/1241.5^\circ & S_{l2} &= 0.52^Z27.8^\circ \\ S_{,:1} &= 0.98^Z24.4^\circ & S_{21} &= 1.71/161.4^\circ \\ S_{22} &= 0.517Z -79.2'' & S_{23} &= \sqrt{L} -85^\circ \\ S_{31} &= U -58^\circ & S_{32} &= 0.517/130.9^\circ \\ S_w &= 0.59,187.8^\circ \end{aligned}$$

If a normalized series reactance jX is inserted into the common source lead, a potentially unstable equivalent transistor or two-port network can be obtained. The possible range of S_n and S_{22} values that can be obtained are shown by the circles plotted in Fig. 12.13. Also shown on these circles are the values of S_n and S_{21} obtained using a normalized inductive reactance of $j0.5$ and jl and the values obtained using normalized series capacitive reactance of $-j0.5$ and $-jl$. The computed values of the equivalent two-port network and the stability parameter K for the four reactances considered are tabulated below.

$$\begin{aligned} jX &= j0.5 & K &= 0.921 \\ S_u &= 0.244/1181.7^\circ & S_{l2} &= 0.33Z66.3^\circ \\ S_{21} &= 1.83/140.3^\circ & S_{22} &= 0.342/1 -38.9^\circ \\ jX = jl & \text{ AT } = 0.915 \\ S_u &= 0.126Z263.3^\circ & S_{l2} &= 0.463/-.65-1^\circ \\ S_{21} &= 1.57/139.7^\circ & S_{22} &= 0.475^Z-41.37^\circ \\ jX = -j0J & \text{ K } = -0.49 \\ S_n &= 1.818/1188.2^\circ & S_{,2} &= 0.556^Z -63.5^\circ \\ S_{21} &= 3.23/158.7^\circ & S_M &= 0.545/1185.7^\circ \\ jX = -jl & \text{ K } = -0.453 \\ S_u &= 3.12/1235.2^\circ & S_{12} &= L52Z -19.7^\circ \\ S_{21} &= 3.57/198.7^\circ & S_{22} &= 1.52Z.231.1'' \end{aligned}$$

**FIGURE 12.13**

Circles showing the values of the scattering-matrix parameters S_{22} and S_{11} for a MESFET with a series reactance in the common source lead.

For the above reactance values $K < 1$, so the MESFET with common source series feedback is unstable (will oscillate). Note that a series capacitive reactance produces a highly unstable device, in particular, for $jX = -j\lambda$ the reverse transmission coefficient S_{12} is very large relative to normal values and both $|S_{11}|$ and $|S_{22}|$ are greater than unity.

The results given above were obtained using the computer program TRIPORT.

^j2.5 OSCILLATOR CIRCUITS

A microwave oscillator can be designed using any of the standard low-frequency oscillator circuits such as the Hartley, Colpitts, or Clapp circuits. Various variations of these circuits can also be used. The frequency stability of the oscillator is generally achieved by incorporating a resonator in either

r~i

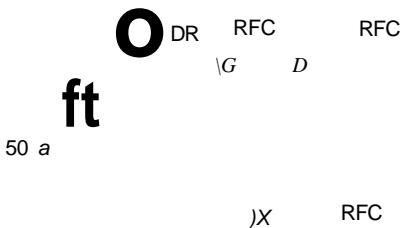


FIGURE 12.14
A 5-GHz FET oscillator using a dielectric resonator DR in the input circuit for frequency stabilization. The feedback is obtained using a series capacitive reactance in the common source lead.

the input or output circuits or as part of the feedback loop. A disk resonator can be used but its Q is relatively low; so the resultant frequency stability will not be very high. A high- Q metallic cavity can be used, but because of its large size the compact high- Q , temperature-stable, dielectric resonator is often the preferred choice.

In Fig. 12.14 we show a 5-GHz FET oscillator that is stabilized by using a dielectric resonator in the input circuit. The magnitude of the source reflection coefficient is controlled by the coupling to the resonator which can be varied by changing the spacing d between the resonator and the microstrip line. The phase angle of the source reflection coefficient is controlled by the length l of the input line. The output circuit is a standard stub-matched circuit that transforms the 50- Ω load impedance to the required value for the oscillator. The FET is made to oscillate by using a series capacitive reactance in the common source lead. This feedback arrangement makes the equivalent transistor two-port circuit unstable as was shown in Example 12.1. The dc bias currents are applied through K_r chokes. The output load is isolated from the oscillator dc voltages by the low-impedance dc blocking capacitor C_t .

In Fig. 12.15 we show an FET oscillator using a dielectric resonator in the feedback path from the drain to the gate. The amount of feedback can be adjusted by the coupling to the dielectric resonator. The resonator is located a distance $A/4$ from the open-circuited ends of the coupling line since the standing wave of current and magnetic field is greatest at this location from the open ends and thus provides the strongest coupling to the resonator. The correct phase of the feedback voltage is controlled by the

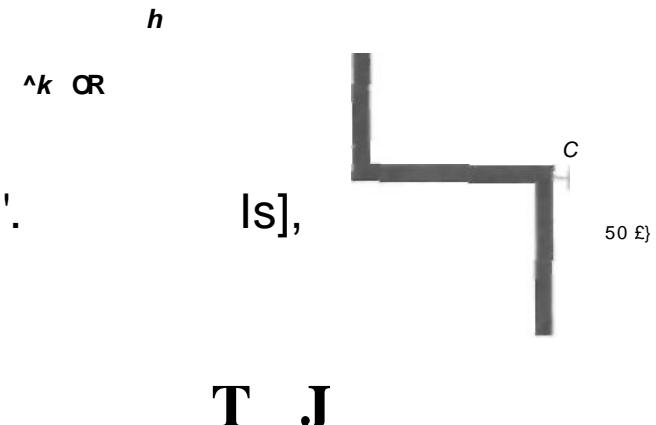


FIGURE 12.15

An FET oscillator using a dielectric resonator in the feedback path from the drain to the gate.

lengths l_1 , and l_2 , as is the phase of the source reflection coefficient *seen at the gate terminal*.

Electronic control of the oscillator frequency can be obtained using a varactor diode as part of the input circuit. Variable-frequency oscillators are also built using a yttrium garnet (YIG) ferrite sphere whose resonant frequency is controlled by the dc magnetic biasing field.

OSCILLATOR DESIGN

When maximum power output and efficiency are not of prime importance, a satisfactory oscillator design can be achieved using the small-signal scattering-matrix parameters. The amount of power generated can be varied by adjusting the dc bias voltages. The major shortcoming of small-signal oscillator design is that it does not provide any way of predicting the steady-state oscillating signal level. Oscillator design based on large-signal scattering-matrix parameters is much more difficult because of the difficulty of obtaining large-signal parameters. Two approaches are possible, namely, measuring the scattering-matrix parameters under large-signal conditions or obtaining these from computer simulations using a theoretical nonlinear model of the transistor. Neither method is easy to carry out so as to achieve a high accuracy. Space limitations will not allow us to discuss the large-signal approach. However, the references at the end of this chapter provide information on methods that have been developed and used successfully. We will only consider the small-signal design approach, and even this in a limited way, by means of two examples.

Example 12.2 Oscillator design using an unstable transistor. A "ail" bipolar transistor has the following scattering-matrix parameters at 6 GHz:

$$S_{11} = 0.65^{\angle 130^\circ} \quad S_{12} = 0.2^{\angle 80^\circ}$$

$$S_{21} = 2^{\angle 142^\circ} \quad S^{TM} = 0.4Z_0 - 60^\circ$$

The stability parameter $K = 0.646$, which is less than 1, so the transistor is potentially unstable. In Fig. 12.16 we have plotted the input and output stability circles. One of the primary effects of large-signal operation is reduction of gain because of gain saturation. For simplicity, we will assume that under large-signal conditions S_{21} changes to a value $1.5^{\angle 42^\circ}$ and all other scattering-matrix parameters stay the same. Thus, under large-signal conditions, we find that the stability parameter K has increased to 0.798. As a result the stability circles move. The large-signal stability circles are shown by the dashed circles in Fig. 12.16.

Let us choose a source reflection coefficient $\Gamma_s = 1^{\angle 210^\circ}$ which corresponds to a pure capacitive reactance load at the base. For this value of Γ_s

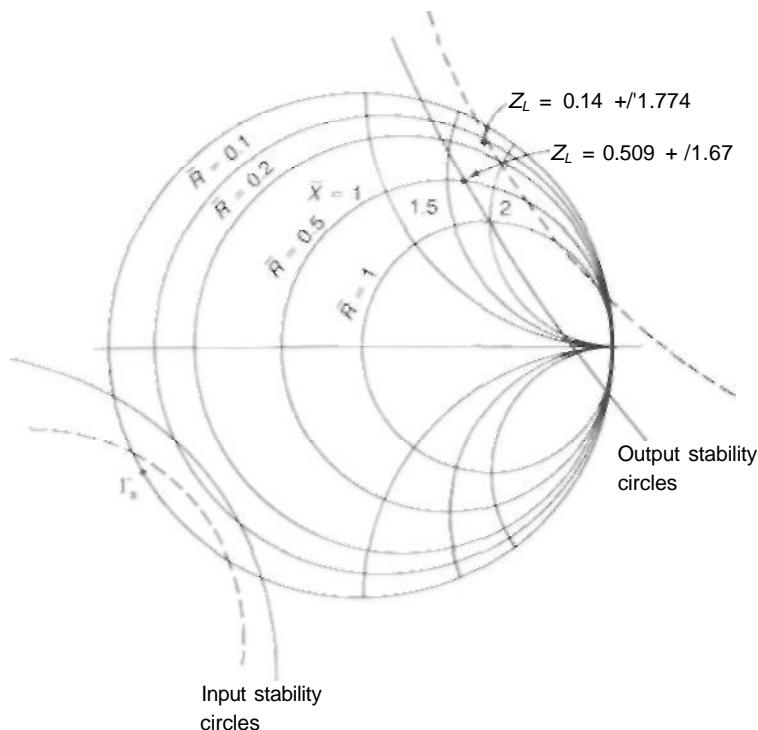


FIGURE 12.16

Stability circles for small- and large-signal conditions for the oscillator discussed in Example 12.2. The source reflection coefficient and two values of the load impedance are also shown.

we find that the output normalized impedance of the transistor is

$$Z_{out} = -0.509 - j1.67 \quad \text{for small signal conditions}$$

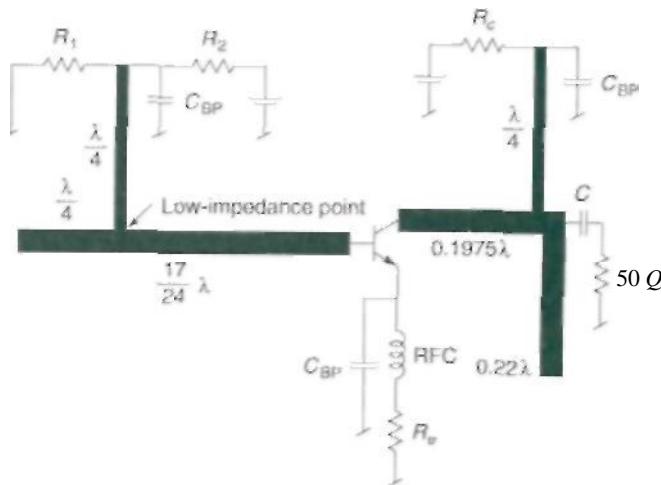
$$\text{and} \quad Z_{out} = -0.139 - jl.774 \quad \text{for large signal conditions}$$

In order to satisfy the conditions (12.3) for oscillations, we must choose a load termination Z_L at the collector where $Z_L = 0.509 + j1.67$ for low-level oscillations and $Z_L = 0.139 + jl.774$ for large-signal oscillations. The reactive parts are almost of the same value, but the resistive part for large-signal conditions is only 0.27 of the required value for small-signal conditions. In the small-signal design approach, it is usual practice to choose the load resistance around a factor of 3 smaller than what is required by the condition (12.3a) for steady-state oscillations. This will allow the oscillations to build up in amplitude until gain saturation makes $-R_{out}$ equal to R_L . There will be some shift in the resonant frequency as the oscillations build up since X_{out} changes and thus the frequency must change until $X_{out} = -X_L$. When the input network produces a very rapid change in the phase of F_s with frequency, only a small frequency shift will occur. Some means of tuning the oscillator is normally used so as to establish the desired frequency of oscillation.

In the example we are discussing, we will choose $Z_L = 0.14 + j1.774$. This impedance point is shown in Fig. 12.16, as is the point $Z_L = 0.509 + j1.67$. The former point lies on the stable side of the assumed large-signal stability circle and the oscillator would not oscillate under these conditions. Our choice for Z_L will, in actual practice, limit the oscillation amplitude at a value for which $|IS_{21}|$ is somewhat greater than 1.5 so as to keep the point Z_L in the unstable region. Since we have used a pure reactive termination at the input, there is no power delivered to the input circuit. Thus the power gain of the circuit is negative and infinite, so that the stable point of oscillation occurs when the stability circle moves outwards so as to make Z_L lie on the circle since the output stability circle coincides with the infinite gain circle. When the input termination has a resistive part, the power gain must be negative and finite, so that Z_L must lie on the appropriate negative gain circle and inside the unstable region (see Chap. 10 for a discussion of negative gain circles in the unstable region).

The circuit for the oscillator is shown in Fig. 12.17. An open-circuited transmission line is used to produce the input reactance. Since $r_s = e^{j2\pi f L}$ the minimum length L is equal to $5A/24$ in order to make the phase angle of F_s equal to -150° , which is equivalent to 210° . The frequency sensitivity of the phase angle can be increased by using a transmission line $A/2$ longer, that is $L = 17A/24$. A 1 percent change in frequency or $\Delta f/f$ will then change the phase angle of F_s by $(-150 - 1801/100)^\circ = -3.3^\circ$. The computed values of Z_{out} for $\Delta f/f = \pm 1\%$ change in the phase angle of T_s are $Z_{out} = -0.083 - j1.697$ for $F_s = 1/1218^\circ$ and $Z_{out} = -0.202 - jl.863$ for $F_s = 1/222^\circ$ under large-signal conditions. It is apparent that Z_{out} is quite sensitive to small changes in the phase angle of T_s .

The output matching network consists of a 50-ft transmission line of length $0.1975A$ and an open-circuit 50- Ω stub of length $0.22A$. This circuit transforms the 50- Ω load to a normalized impedance of $0.14 + j1.774$ at the collector. The bias voltages are applied through 150-ft quarter-wave

**FIGURE 12.17**

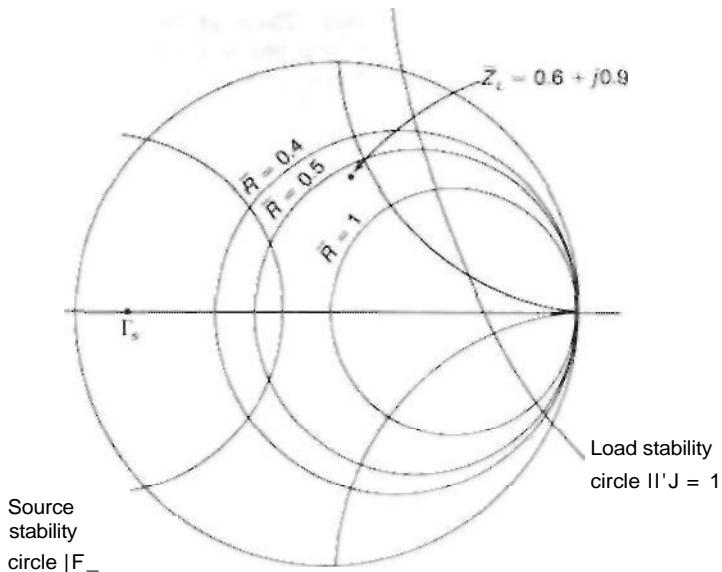
The oscillator circuit designed in Example 12.2.

transmission lines that are bypassed to the ground plane by capacitors C_{BP} . These lines are connected at low-impedance points on the input and output circuits. The 50- Ω output is isolated from the transistor by the dc blocking capacitor C .

The design carried out above is a hypothetical one since we do not know how the scattering-matrix parameters change with signal level. The example does show the basic physical principles involved in determining the steady-state operation of the oscillator. The small-signal design approach that uses $Z_L = -(R_{out}/3 + jX_{out})$ usually leads to a reasonably satisfactory design.

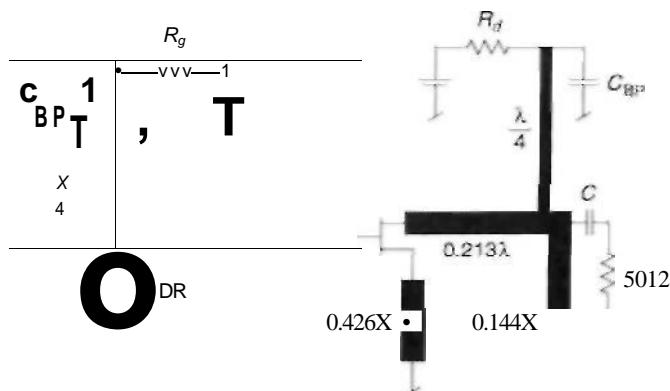
Example 12.3 Oscillator design using a dielectric resonator. In this example we will use the MESFET described in Example 12.1 with a series reactance $-j0.5$ in the common source lead. The equivalent two-port scattering-matrix parameters are $S_u = 1.818Z.188.2^0$, $S_{12} = 0.556A -63.5^\circ$, $S_{21} = 3.23^58.7^\circ$, $S_{22} = 0.545^185.7^\circ$, and the stability parameter $K = -0.49$. The input and output stability circles for the equivalent two-port network are shown in Fig. 12.18. Since $|S_{22}| < 1$ the origin is a stable point for r_{ss} , so that values of Y_s inside the source or input stability circle are unstable ones. The values of Z_L outside the load stability circle are unstable ones since $|S_{11}| > 1$.

We will choose $T_s = -0.8$ for the initial design of the oscillator. This value of T_s is shown in Fig. 12.18 and is in the unstable region. For this value of T , we find that $Z_{om} = -1.6 - j0.906$. For Z , we will choose $Z_L = 0.6 + j0.906$ for which $R_L = -0.375R_{out}$. The oscillator circuit used is shown in Fig. 12.19. The resonator equivalent circuit is a parallel combination of R and C which is series-coupled to the microstrip line by an ideal transistor with turns ratio $n : 1$ as shown in Fig. 7.23. We can choose R equal to unity

**FIGURE 12.18**

The input and output stability circles for the oscillator in Example 12.3. The design values of T , and Z_j are also shown.

choosing an appropriate value for the turns ratio n . The resonator Q is given by R/w_0L , which determines the inductance in terms of the resonator Q . The capacitance in the equivalent circuit is given by the resonance condition $IOQLC = 1$. At $w = i\omega_0$ the impedance coupled into the microstrip line is a series resistance equal to $n^2R = n^2$, so n^2 represents the coupling coefficient. The transmission line of length $A/4$ can be replaced by a short circuit at the

**FIGURE 12.19**

The oscillator circuit designed in Example 12.3.

location of the resonator. Thus, at the resonator location, the reflect coefficient on the microstrip line is $I'' = (n^2 - 1) / (\epsilon^2 + 1)$. In order to make T_n less than 10^{-4} , we must choose $\epsilon^2 = (1 + D) / (1 - D) = 0.2 / 1.8 = 0.111$; so the resonator is undercoupled. By using a transmission line $A/2$ long between H resonator and the gate, we obtain $T_s = V = -0.8$.

The output circuit consists of a $50\text{-}il$ transmission line 0.213A long and an open-circuited 50-ft stub 0.144A long. This network transforms the $50\text{-}\Omega$ load impedance to the required normalized value $0.6 + j0.906$ at the drain terminal. The capacitive feedback normalized reactance of $-j0.5$ is obtained by using a short-circuited 50-fl stub of length equal to 0.4262A in the common source lead.

In general, the resonator impedance coupled into the input microstrip line is given by

$$\epsilon = \frac{n^*}{1 + 2jQ \frac{\epsilon_{\text{FI}} - *}{\epsilon_0}}$$

in the vicinity of the resonant frequency ϵ_0 . If we assume that the resonator Q equals 500, then a 0.1 percent change in frequency will change T from -0.8 to

$$r = \frac{n^2 Z - 1}{n^2 Z + l} = \frac{n^2 - 1 - / - 8 \text{ fft}}{n^2 + l + j} = 0.895 L 186.38^\circ$$

The new value of Z_{ou} becomes $-1.37 - j1.41$, where we have assumed, for a 0.1 percent change in frequency, that there is a negligible change in the S_y parameters and the transmission-line length. If we used stronger coupling to the resonator, we would obtain a larger change in I'' and this would improve the frequency stability of the oscillator, since the resonator would have a stronger control of the oscillation frequency. A better choice for I''_s would be -0.4 which gives $n^2 = 3/7$ and $Z_{oul} = -0.494 - j0.089$. A 0.1 percent change in frequency would now give

$$T = \frac{3 - 7 - 77}{3 + 7 + j7} = \frac{-4 - / 7}{10 + jl} = 0.6605^{\circ} 205.3^\circ$$

and $Z_{oul} = -0.523 - j0.905$. The change in the phase angle of I'' has been increased by a factor of about 4 by using the larger coupling. The frequency stability is also increased by about the same amount. The design of the oscillator circuit using $f_s = -0.4$ is left as a problem to be solved (Prob. 12.5).

12.7 MIXERS

In Fig. 12.20 we show a block diagram of a microwave superheterodyne receiver. The signal from the antenna is first amplified by a low-noise amplifier. After amplification, the signal is mixed with a local-oscillator signal to obtain the original signal translated to a much lower frequency called the intermediate (IF) frequency. The mixer is a nonlinear device such as a diode or dual-gate FET. If the microwave carrier frequency is ω_c

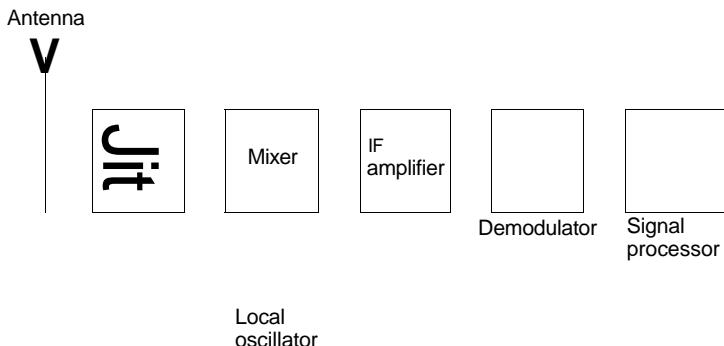


FIGURE 12.20
A block diagram of a microwave receiver.

the oscillator frequency is ω_0 . The nonlinear mixer device will produce signals at the IF frequency $\omega_{IF} = \omega_c - \omega_0$, at the frequency $\omega_0 + \omega_c$, and, in general, at many harmonic frequencies $n\omega_0 \pm m\omega_c$, where n and m are integers. The signal at the IF frequency is further amplified, then demodulated, and finally processed for the intended output application.

A single-tone AM signal with modulation frequency ω_m has the form

$$(1 + M \cos(\omega_m t) \cos \omega_c t) = \cos \omega_c t - jM[\cos(\omega_c + \omega_m)t + \cos(\omega_c - \omega_m)t]$$

After mixing, the sideband spectrum becomes

$$\cos[(\omega_r - \omega_0)t + \omega_m t] + \cos[(\omega_r - \omega_0)t - \omega_m t]$$

when the carrier frequency ω_r is greater than the local-oscillator frequency, ω_0 . When the carrier frequency is less than the local-oscillator frequency, the sideband spectrum, after mixing, remains unchanged because an AM signal has symmetrical sidebands. Thus a local-oscillator frequency greater or smaller than the carrier frequency can be used.

For an FM or phase-modulated signal of the form $\cos[\omega_0 t + \phi(t)]$ the spectrum, after mixing, is of the form $\cos[(\omega_r - \omega_0)t + \phi(t)]$ when $\omega_r > \omega_0$ and $\cos[(\omega_0 - \omega_r)t - \phi(t)]$ when $\omega_r < \omega_0$. For the case when ω_r is greater than the carrier frequency ω_0 , the sideband spectrum is reversed with the high frequencies becoming low frequencies, and vice versa. This phenomenon occurs because FM and phase-modulated signals do not have symmetrical sidebands. The subtraction of the spectrum of $\phi(t)$ from the IF frequency reverses the high- and low-frequency components. In order to avoid this inversion of the signal spectrum, the local-oscillator frequency must be less than the carrier frequency in FM and phase-modulated systems.

In this section we will discuss those mixer characteristics that are important from a systems point of view. For the purpose of this discussion,

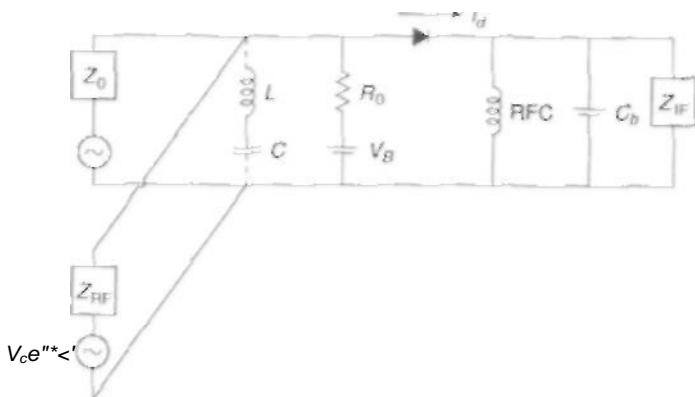


FIGURE 12.21
A single diode mixer circuit.

we will use a simple single-diode-mixer circuit to help clarify some of the operational characteristics of diode mixers. In the following section we will examine some of the more complex mixer circuits that enhance the overall mixer performance.

In Fig. 12.21 we show a simplified circuit for a single diode mixer. The RF signal with carrier frequency w_r is connected to the diode through a filter network with impedance Z_{RF} . Similarly, the local-oscillator signal at frequency w_0 is applied to the diode through the impedance Z_0 . The IF signal appears across Z_{IF} which represents the input to the IF amplifier. If a point contact diode is used, a dc biasing circuit is not used but the diode must have a dc current return path to ground which is through the RF choke RFC. When a Schottky barrier diode is used, a small forward bias is normally applied to the diode. This serves to overcome the barrier potential and increases the sensitivity of the diode. In a conventional $p-n$ diode, the junction capacitance is quite large and will shunt the RF and LO signals across the junction, thus making these diodes ineffective as mixer diodes at frequencies greater than 1 GHz. The minority carriers also limit the diode recovery time. The Schottky diode, consisting of a metal-semiconductor junction does not have a depletion layer and also has very little stored charge at the junction. Hence it has a very small junction capacitance. It is usable as a mixer diode for frequencies in the microwave and millimeter-wave range and beyond.

The IF frequency is much lower than the LO and RF frequencies, for example, if the RF frequency $f_r = 10$ GHz and a typical IF frequency of 1 MHz is assumed, then the local-oscillator frequency must be (10 ± 1) GHz. In this case the IF frequency is a factor of 200 smaller than the RF and LO frequencies. Hence the capacitor C_b in the circuit shown in Fig. 12.21 can be chosen large enough to short-circuit the high-frequency components at the IF amplifier input. Consequently, the IF port is isolated.

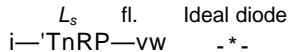
**C,**

FIGURE 12.22
Equivalent circuit for a microwave diode.

the RF and LO ports. On the other hand, the RF and LO frequencies are almost equal, so it is not practical to use filter networks Z_0 and Z_{RF} that provide good isolation between the RF and LO ports. We can, however, assume that Z_n and Z_{RF} are essentially zero-impedance elements at the IF frequency. We could, for example, insert a series resonant circuit across the RF and LO ports (shown by dashed lines in Fig. 12.21) with L and C chosen so that $LCU_{IF} = 1$, where ω_{IF} is the IF frequency. This circuit provides a zero-impedance path across the RF and LO ports at the IF frequency. At the RF and LO frequencies, $j\omega L$ represents a very high reactance that shunts the RF and LO ports and can be neglected.

The equivalent circuit of the diode is shown in Fig. 12.22. In this circuit we have an ideal diode shunted by the diode junction capacitance C , along with a series resistance R_s , a series inductance L_s , and a package shunting capacitance C_p . The ideal diode is described by the equation

$$i_d = I_s (e^{v_d / (kT)} - 1) \quad (12.16)$$

where i_d is the diode current, I_s is the reverse saturation current, v_d is the voltage across the diode; $y = e/kTn$, where e is the electron charge, k is Boltzmann's constant, T is the absolute temperature, and n is a diode-dependent parameter having a value between 1 and 1.5. At room temperatures y has the approximate value of 40.

For the purpose of the discussion in this section, we are going to neglect all of the parasitic elements in the equivalent circuit of the diode. This can be justified only if we assume that the RF and LO frequencies are quite low. We will carry out a more careful analysis of the diode mixer in a later section. With the assumptions we have made, the mixer equivalent circuit for RF, LO, and IF signals reduce to those shown in Fig. 12.23. The network that the diode is embedded in has been assumed to have a zero impedance at all frequencies except those in the vicinity of the RF, LO, and IF frequencies.^t Thus the only voltages that can exist across the diode are those at the RF, LO, and IF frequencies since all other frequency components are short-circuited by the embedding network.

^tWhen diode biasing is used, the RF and LO circuits must have nonzero dc resistance so as not to short-circuit the dc voltage applied to the diode.

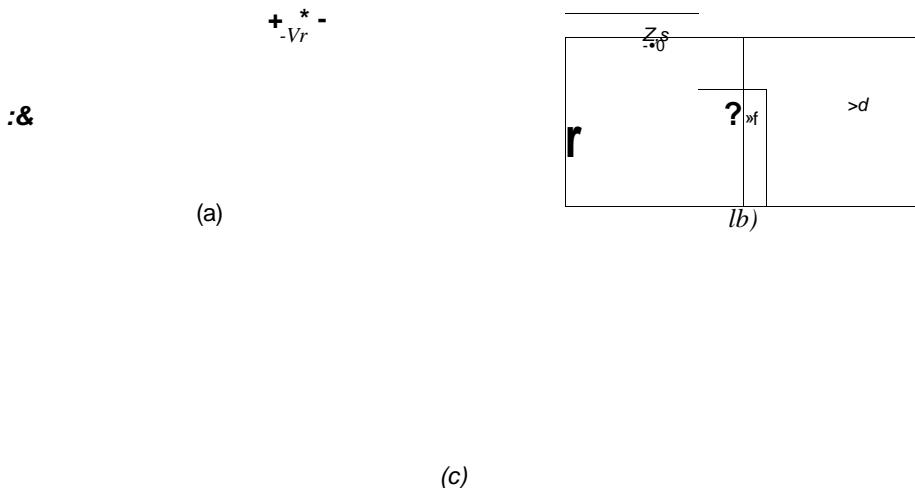


FIGURE 12.23

(a) Equivalent circuit for RF signals in the mixer; (b) equivalent circuit for LO signals;
(c) equivalent circuit for IF signals.

The current i_d through the diode can be expanded in a Taylor series about the dc operating point. This is equivalent to assuming a power-series expansion of the form

$$h = J_0 - a_1 v + a_2 v^2 + a_3 v^3 + \dots \quad (12.17)$$

where J_0 is the dc biasing current, v is the voltage excursion about the operating point, and a_1, a_2, a_3, \dots are suitable coefficients. We will assume that the voltage v consists of a local-oscillator signal $v_0 = V_0 \cos \omega_0 t$, an RF signal $U_j = V_r \cos \omega_r t$, and an IF signal $-v_{IF}$ at the IF frequency ω_{IF} . The IF signal has the form $w_{IF} = |V_{IF}| \cos(\omega_{IF}t + \phi)$, where d is the phase angle. The complex-phasor IF voltage is $V_{if} = |V_{IF}| e^{-j\phi}$. When we expand (12.17) we obtain

$$\begin{aligned} i_d = J_0 &+ a_1(v_0 + v_I - u_{RF}) + a_2(vl + v_I + v_{IF} + 2v_0v_I - 2v_0v_{IF} - 2W_{IF}) \\ &+ a_3(vl + u_I - u_{RF} + Sv'l v^2 - Svl u_w + 3y_{0u} - 3f_{off}F - S i^2 p - B a_0 v_I v_w) \end{aligned} \quad (12.18)$$

In a number of published analyses of mixers using a power series such as that in (12.17), it is assumed that the voltage across the diode consists only of the RF and LO signals. This assumption would imply that the embedding network that the diode is connected to has a zero impedance at the IF frequency, which is an unrealistic assumption. If power is to be delivered to one port of the embedding network at the IF frequency, then the network must have a nonzero impedance at the IF frequency at

port. The impedance of the embedding network at the various harmonics of the applied signals determines the harmonic voltages across the diode.

Operation

In practice, the local-oscillator signal has an amplitude much larger than the RF signal and the IF signal. When the RF signal amplitude is small, we can neglect terms that are proportional to higher-order powers of the RF and IF signals. Thus, when we retain only those terms that are linear in u , and v_{IF} , the diode current is given by

$$\begin{aligned} i_d = & I_0 + c_1(j'_0 + \Delta - v_w) + a_2(4 + 2v_0v_l - 2f_0d/I_F) \\ & + a_3(u\% + 3v\$v_t - 3u^2i_{IF}) \end{aligned} \quad (12.19)$$

A term such as u^2 equals $V_0^2 \cos^2 a_2 t = iV^2 + |v_*| \cos 2aj_2t$. The product term $2U_0L_1$ equals $V_0V_s[\cos(w_0 - \omega)/ + \cos(w_0 + co)t]$ and has an IF frequency component. The product term $-2v_{IF}v_{IF}$ equals $-V_0|V_s| \{ \cos[(a_0 - w_{IF})t - \phi] - \cos[(w_0 + w_{IF})t + \phi] \}$ and does not contain frequency components falling within the IF amplifier passband. The term $-3a_3i_{IF}$ contributes an IF current component equal to $-fcig^2v_{IF}$. The total IF current is given by

$$j_{IF} = -a_1i_{IF} + a_2y_0V_l \cos w_{IF} - fa_3V_{IF}i_{IF}$$

We now let $i_{IF} = \text{Re } /_{IF}e^{j\omega_{IF}t}$ and use phasor analysis to obtain

$$ht = -\langle\langle !*\rangle\rangle + \rangle\langle W - hs^V\$W$$

At the IF frequency the circuit equation or constraint imposed by the embedding network shown in Fig. 12.23c gives $V_{IF} = /_{IF}2_{IF}$. Thus we find that

$$V_{IF} = \frac{a_0 V_s Z_{IF}}{1 + (o_0 + 1.5a_3V_0^2)Z_{IF}} \quad (12.20)$$

The power-series expansion up to terms in a_3 is valid only for small values of V_0 . Consequently, the term $1.5a_3V_0^2$ is usually small relative to a_0 . The above equation shows that under the assumptions made, the voltage at the IF amplifier input is linearly proportional to the RF voltage amplitude V_0 . In this *operating range* the mixer functions as a linear mixer.

The conversion loss, in decibels, for the mixer is given by

$$L = \text{conversion loss} = 10 \log \frac{\text{available RF power}}{\text{IF input power}} \quad (12.21)$$

Typical values for conversion loss for a single-diode mixer are 6 to 10 dB.

Nonlinear Mixer Operation

For larger values of the RF voltage, we must take additional terms in account. From the term multiplied by a_3 , we have a term

$$-3a_3w^2f_{1P} = -1.5a_3V^2U_K - 1.5a_3V_1^2u_{IF}\cos 2w_1 i$$

and a term

$$\begin{aligned} -a_3u_F &= -a_3|V_{IF}|^3 \cos^3(o_{IF}t + \angle j) = -0.75a_3|V_{IF}|^3 \cos(w_{IF}t + \angle \phi) \\ &\quad - 0.25a_3|V_{IF}|^3 \cos(3w_{IF}t + 3\phi) \end{aligned}$$

The IF frequency component of this latter term can be expressed as $-0.75a_3|V_{IF}|^2t_{IF}$. We can use (12.20) to obtain an approximate solution for $|V_{IF}|^2$ and then find that

$$\frac{V_y}{V_{IF} + a_t + 1.5a_3V_0^2} = \frac{\frac{a_2V_nZ_{IF}V}{V^2}}{\frac{V_0^2}{V_0^2 - 2(1 + (a_t + 1.5a_3V_0^2)Z_{IF})^2}}$$

$$(12.22)$$

where $Y_{IF} = Z_{IF}$. This equation shows that the effect of the nonlinear terms $-a_3i_F^3$ and $-3a_3u^2L_{IF}$ is to reduce the IF voltage. Thus a mixer will exhibit nonlinear saturation and the range of allowed input RF voltage amplitudes (dynamic range) must be limited in order to avoid nonlinear distortion of the down-converted RF signal. The dynamic range of a mixer is specified by giving the RF power level at which a compression of 3 dB occurs in the IF power. In Fig. 12.24 we show a plot of IF power versus RF input power and the 3-dB compression point. The units are dBm or decibels relative to 1 mW.

The nonlinear behavior of a mixer also results in intermodulation distortion. Consider an input signal consisting of two closely spaced sinusoidal signals at the frequencies w_1 and w_2 , where $|w_1 - w_2| \ll \omega_{IF}$. The presence of two closely spaced RF signals results in the generation of a larger number of frequency components. For linear operation we obtain IF signals $|V_{IF1}|\cos[(a_{IF1} - u_{IF1})t + \angle \phi_1]$ and $|V_{IF2}|\cos[(a_{IF2} - u_{IF2})t + \angle \phi_2]$. The cubic term $-a_3i_F^3$ will result in new additional frequency components come from the products $V_{IF1}u_{IF2}$ and $u_{IF1}V_{IF2}$. The new frequencies generated, which fall within the IF amplifier passband, are $a_{IF1} - 2a_{IF2} + \angle \phi_2$

^tSometimes the 1-dB compression point is used.

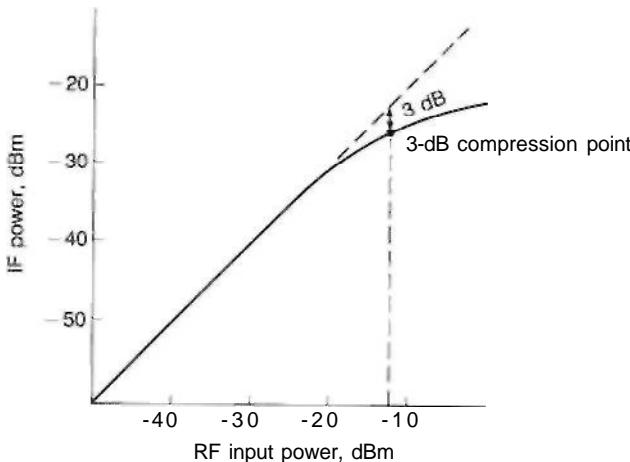
**FIGURE 12.24**

Illustration of nonlinear mixer response and the 3-dB compression point.

$w_1 - 2w_2 + w_3$. We will let $\omega_{IF,i} = (w_n - w_i) + (w_2 - w_1)^{an}$ and $\omega_{IF,i} = (u_{IF,i} - u_{IF,1}) + (o_i) - d\omega_2$. For a self-consistent solution for (12.18), we see that we must assume the presence of at least four IF signals, namely,

$$\begin{aligned} v_{IF} = & v_{IF1} \cos(\omega_{IF,1}t + \phi_1) + V_{IF,2} \cos(\omega_{IF,2}t + \phi_2) \\ & + V_{IF,3} \cos(\omega_{IF,3}t + \phi_3) + V_{IF,4} \cos(\omega_{IF,4}t + \phi_4) \end{aligned} \quad (12.23)$$

When v_{IF3} and u_{IF4} are included, then the term multiplied by ω_3 will also have product terms of the form $v_{IF1}v_{IF3}$, $v_{IF1}v_{IF4}$, etc. These product terms result in additional new frequency components that fall within the IF passband. Since v_{IF1} and L_{IF2} are proportional to V_1 , and V_2 , we see that I_{IF3} and u_{IF4} will be proportional to $V_1^2V_2^*$ and $V_2^2V_1$, respectively. A term such as i^{fip} is thus proportional to $V_1^*V_2$ and can be neglected since its amplitude will be much smaller.

The production of new frequency terms that arise from the mixing of v_{SF1} and u_{IF2} is called two-tone intermodulation distortion. If $V_1 = V_2 = V$ these terms are proportional to the third power of the RF signal amplitude. The IF power associated with the intermodulation terms will increase proportional to the third power of the input RF power. The intermodulation distortion of a mixer is specified by giving the RF power level at which the intermodulation IF power becomes equal to the IF power associated with the desired IF signals v_{IF1} and v_{IF2} . This is a theoretical value that is obtained by extrapolating the linear mixer response or IF power output at the desired frequencies ω_{IF1} and ω_{IF2} and that of the IF power output for the intermodulation terms until they meet. On a decibel scale the slope of the curve giving the intermodulation IF power is three times that of the

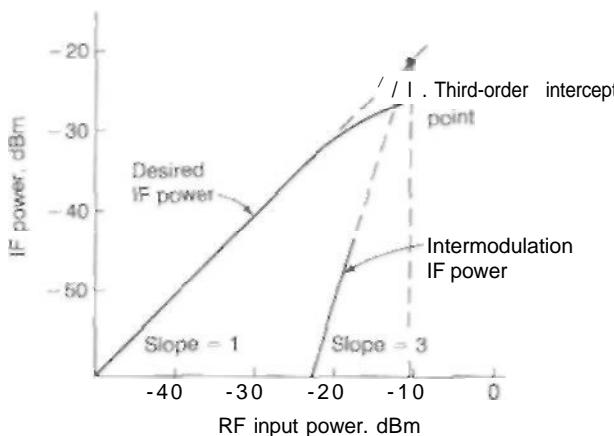


FIGURE 12.25

Illustration of diode mixer response for the desired IF power and the IF power at the intermodulation frequencies. The intersection of the two extrapolated response curves gives the third-order intercept point.

curve giving the desired IF signals. The intercept point is called the third-order intercept and is illustrated in Fig. 12.25. In practice, the RF input power to a mixer must be kept below the 3-dB saturation point and also below the third-order intercept point to avoid signal distortion.

When the amplitude V_2 of the RF signal at the frequency co_2 is held constant, but V_j is increased, then the slope of the line giving the intermodulation power in the IF passband will be 2.

12.8 MIXER NOISE FIGURE

The noise figure of a mixer is another important parameter that specifies the performance of a mixer. Thermal input noise at frequencies around $<o_0 - w_{IF}$ will be down-converted into noise in the IF passband. Thermal noise at the image frequencies around $w_0 + \omega_{IF}$ when mixed with the LO signal will also fall within the IF passband. The mixer will also introduce additional noise with frequency components in the IF passband. Additional noise arises primarily from thermal noise in the resistive elements of the mixer and from shot noise.

The noise figure of a mixture is defined by the relation

$$F = S/N_o \quad (12.24)$$

where S/N_i is the input signal-to-noise power ratio and S_o/N_o is the output signal-to-noise ratio. If the local-oscillator frequency is $<o_0$, the signals at $<o_0 + <o_{IF}$ and $<o_0 - w_{IF}$ are converted to IF signals. Frequency $u > o_0 + co_{ie}$ is called the upper sideband and the frequency $w_0 \sim$

the lower sideband. Noise at both sidebands is converted into noise in the IF passband. In most systems the signal is present in one sideband only. The unused sideband is called the image frequency. When the signal is present in one sideband only, the noise figure is designated as the single-sideband (SSB) noise figure F_{SSB} . If the signal is present in both sidebands, the corresponding noise figure is called the double-sideband (DSB) noise figure and denoted by F_{DSB} . A double-sideband signal would be of the form $A[l + m(t)][\cos(\omega_0 + (O_p)t + \cos(\omega_0 - w_{IF})t)]$, where $m(t)$ is the modulation. When down-converted the IF signal voltage will be proportional to $2A[l + rn(t)]\cos(\omega_0 - w_{IF})t$, and consequently, the IF power is increased by a factor of 4. Thus for double-sideband signals, the ratio S_v/S_u is one-half of that for a single-sideband signal since the output IF power is increased by a factor of 4 but the input RF power is increased by a factor of 2 only. This assumes that the mixer conversion loss is the same for both sidebands. As a consequence of the above, the single-sideband noise figure is twice as large as the double-sideband noise figure, that is $F_{SSB} = 2^F_{DSB}$. The double-sideband noise figure is easiest to measure since noise sources are usually very broadband and would provide noise signal input in both sidebands.

Schottky diode mixers have typical single-sideband noise figures in the range of 4 to 8 dB. The IF amplifier noise figure will also increase the noise figure of the mixer-IF amplifier combination because of the mixer conversion loss. The noise figure of the mixer-IF amplifier combination is given by

$$F = F_m + L_c(F_{IF} - 1) \quad (12.25)$$

where F_m is the mixer noise figure, F_w is the IF amplifier noise figure, and L_c is the numerical value of the mixer conversion loss. As an example assumes that the conversion loss is 4 (6 dB), the IF amplifier noise figure is 2 (3 dB), and the mixer noise figure is 4. The noise figure of the combination is thus 8 which is a 3-dB increase over that of the mixer by itself. Since the signal level at the output of a mixer needs additional amplification, the mixer conversion loss has a significant effect on the overall noise figure. For this reason it is important to have sufficient signal amplification ahead of the mixer. The noise figure of the preamplifier will then govern the system noise figure.

9 BALANCED MIXERS

There are several disadvantages associated with a single-diode mixer. These are: no isolation between the RF and LO ports, poor isolation between the IF port and the RF and LO ports, a high level of oscillator noise input to the IF amplifier, and the generation of many spurious signals. Balanced mixers can improve the characteristics of a mixer and alleviate many of the listed shortcomings.

The basic circuit for a single balanced mixer is shown in Fig. 12.26. The two diodes need to be well matched in their electrical characteristics in

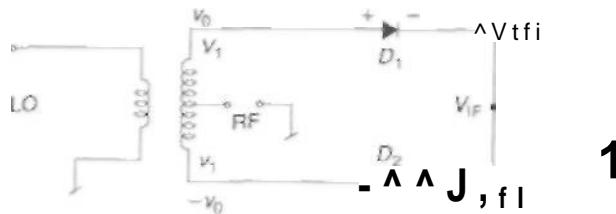


FIGURE 12.26
Basic circuit for a single balanced diode mixer.

order to achieve a well-balanced mixer. From the symmetry properties of the circuit, we can see that the voltage across diode D_x is $u_0 + v - v$ while that acting across diode D_2 is $u_0 - u_x + u_{IF}$. If we retain terms up to a_2 in the expression (12.18) for the diode current, we find that for diode D

$$\begin{aligned} i_d = & I_0 + o^Uo + V_t - i_{IP}) + a_2(u_g + v_f + v^ \\ & + 21, , - 2u_0u_{IF} - 2u,u_{IF}) \end{aligned}$$

while for diode D_2 ,

$$\begin{aligned} i_{d2} = & h + O(l^o - y_I + u_{IF}) + Q_2(t^o + u_1 + 1IF \\ & - 2^0, , + 2y_0y_{IF} - 2v_1v_{IF}) \end{aligned}$$

The input current to the IF low-pass filter is

$$i_{dl} = 2a_1(u, - o,p) + 4o_8(u_0o_1 - u_0o_{IF})$$

The IF frequency component is given by

$$I_{IF} = -2a_1V_{IF} + 2a_2y_0V_1 \quad (12.26)$$

From the above equations we see that there is no local-oscillator voltage at the IF port so the LO and IF ports are isolated. Amplitude-modulation noise from the oscillator, as well as up-converted (or down-converted) thermal noise in the oscillator circuit, will consequently not be present at the I^* input. From an examination of the circuit in Fig. 12.26, it is also apparent that the LO and RF ports are isolated.

For broadband mixers operating up to 1 GHz or so, the single balanced mixer can be built using transformers consisting of windings on a toroidal core. At microwave frequencies the single balanced mixer can be constructed using a 180° 3-dB hybrid junction such as the magic T or hybrid ring described in Chap. 6. A single balanced mixer using a magic T is shown in Fig. 12.27. A 90° hybrid junction can also be used, but the mixer will then not be fully balanced with respect to LO and IF port interactions. In a hybrid junction reflected waves at the output port appear at both mP ports with a phase that makes the reflected waves add out of phase.

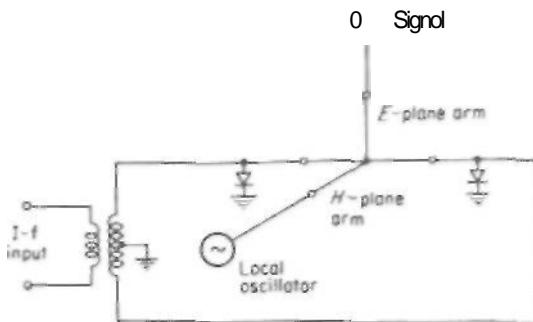


FIGURE 12.27
Single balanced mixer using a wave-guide magic T.

the LO and RF port input VSWRs tend to be small. For a 180° hybrid junction, the reflected waves add in phase at each input port so the resultant input VSWRs are higher for the same degree of output impedance matching. A VSWR of 1.5 using a 90° hybrid junction can be readily achieved, whereas the VSWR for a 180° hybrid junction is around 2. It is relatively difficult to provide a good impedance match between a diode and a transmission line over a wide band of frequencies.

By using the four diodes in a bridge network, a double balanced mixer can be built. The basic circuit for a double balanced mixer is shown in Fig. 12.28 and provides isolation between the RF and LO ports, as well as between the IF and the RF and LO ports. In addition, many more spurious signals cancel at the IF port. The double balanced mixer suppresses the even harmonics of both the RF and LO signals and hence has a lower level of intermodulation distortion than that of the single balanced mixer. At the lower frequencies double balanced mixers are readily constructed using transformer hybrids. At microwave frequencies transmission-line baluns are used and the overall mixer configuration can become quite complex.

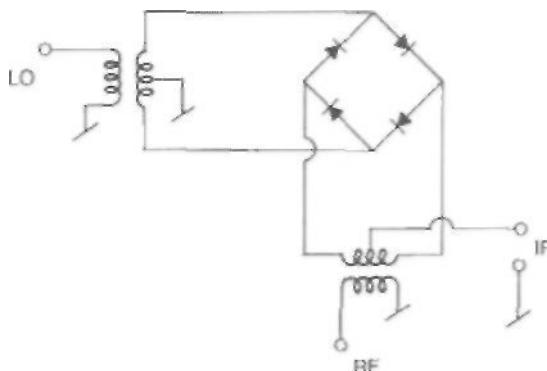


FIGURE 12.28
Basic circuit for a double balanced mixer.

OTHER TYPES OF MIXERS

When the local-oscillator frequency is higher than the RF signal frequencies, the IF spectrum contains those signal components in the frequency band $\omega_0 - \omega_{IF}$. The undesired frequencies in the image band $GJ_0 + w$ are suppressed by means of a filter ahead of the mixer. It is possible to arrange two mixers in such a manner that IF signals corresponding to signals in the upper and lower sidebands are down-converted and appear at separate output ports. The unwanted sideband signals can be dissipated in a resistive termination. A mixer of this type is called an image-rejection mixer and is particularly useful in wideband signals where it is difficult to eliminate signals in the undesired sideband by filtering.

The up-converted signal proportional to $V_0 V_1 \cos(\omega_0 + w)$ represents a transfer of RF power to an undesired frequency. By reactively terminating this frequency component, it can be reflected back to the mixer and mixed with the second harmonic of the LO signal to produce a desired IF signal. With a proper adjustment of the circuit parameters, the conversion loss can be reduced by 1 to 2 dB. This type of mixer is called an image-recovery or image-enhanced mixer.

At millimeter wavelengths it is not always possible to build a suitable oscillator at the very high required frequency. In such circumstances a lower-frequency oscillator can be used and the desired IF frequency signals can be obtained by mixing the RF signal with one of the harmonics of the oscillator. These mixers are referred to as subharmonic mixers. By using two well-matched diodes connected in antiparallel as shown in Fig. 12.29, a second harmonic mixer is obtained. This particular diode arrangement results in an absence of all mixing products that involve odd harmonics of the local-oscillator signal.

FET mixers can also be configured to operate as balanced mixers, although current practice is to use an FET in an unbalanced circuit configuration. Discussions of FET mixers can be found in the references given at the end of this chapter.

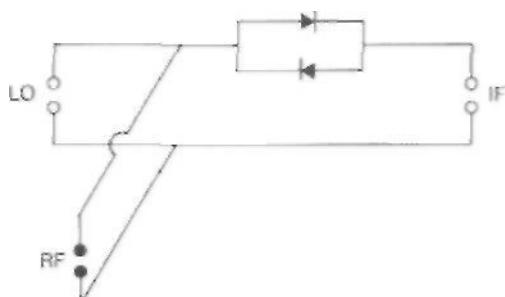


FIGURE 12.29
Basic circuit of a subharmonic mixer where the RF signal is mixed with the second harmonic of the oscillator signal.

j2.11 MIXER ANALYSIS USING HARMONIC BALANCING

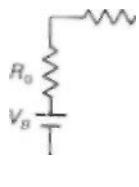
A power-series expansion of the diode current, such as that given by (12.17), is valid only when the voltage across the diode is of the order of a few millivolts. In practice the local-oscillator-signal voltage across the diode is much larger in order to obtain a relatively low conversion loss. The harmonic-balance method of mixer analysis is based on first analyzing the nonlinear mixer circuit with only a local-oscillator signal applied. This leads to a description of the diode in terms of a time-varying conductance and capacitance. In this section we will develop the basic concepts associated with the harmonic-balancing method, which are similar to those used in describing the operation of parametric amplifiers in Chap. 11.

The equivalent circuit for the diode is shown in Fig. 12.22. The junction capacitance C_j is a function of the voltage across the ideal diode junction. The other circuit elements, namely R_s , L_s and C_p , can be considered to be part of the embedding network. We can replace the network as seen from the terminals of the ideal diode by a Thevenin equivalent circuit which includes an equivalent local-oscillator source as shown in Fig. 12.30a. The impedance $Z(w)$ of this network has a different

H

Zlu.)

LO
(a)



lb)

FIGURE 12.30

(a) Thevenin equivalent circuit of a diode mixer as seen from the terminals of the ideal diode; (b) Thevenin equivalent circuit of a diode mixer at $m = 0$.

value at each harmonic n a_{in} of the local-oscillator signal. The total current i_T flowing through the diode and junction capacitor is $i_T = i_d + j$. The current through the capacitor is given by

$$, \quad \frac{dQ(vj)}{dt} = \frac{dQ^{\wedge} dv^{\wedge}}{dVj dt}$$

where $Q(vj)$ is the charge stored at the junction. The incremental capacitance $dQ/dVj = CjiVj$ is a voltage-dependent capacitance since, in general Q is not a linear function of Vj . The current i_d through the ideal diode is given by

$$i_d = I_s(e^{\wedge} - 1)$$

When only the local-oscillator signal is applied to the circuit, the voltage Vj and the currents i_T, i_d, i_c will contain all harmonics naj_0 of the fundamental frequency $a > 0$. Hence we can write

$$\begin{aligned} jj =& x \\ & \text{x} \\ & \ll \& - E v^{\wedge \prime \prime \prime} \\ & n = - \infty \end{aligned}$$

$$h - E i_c, e > \ll \ll$$

Since i_d and v_t are real, $/_{\perp}, = 7^*$ and $V_{-n} = V^*$. The voltage \triangleright , is periodic, so $CjiVj$ is a periodic function and has a Fourier series representation of the form

$$C_{\perp} = \sum_{n=-\infty}^{\infty} C_n e^{-jnaj_0 t}$$

where the C_n will be dependent on Vj . C_n is a time-varying circuit element. The presence of a dc bias voltage simply changes the values of J_0, V_0 of the $n = 0$ terms.

Consider now the presence of an additional small-signal voltage across the junction, with $|v| \ll |Vj|$. The current through the diode is now given by

This equation can be expanded in a Taylor series about Vj , the dyn^{31}

operating point; thus

$$i_r + i - IAt^*j-l) + \frac{dQ(vj)}{dt} = \frac{d}{dt} IdQ(Vj)) +$$

$$\frac{d}{dt}$$

The current i is given by

$$i = |yl|, e^{\lambda} + \frac{dC_s}{dv} + CJ(VJ) \sim \frac{dv}{dv} \quad (12.28)$$

This is a *linear* differential equation of first order with time-dependent coefficients. Under small RF signal conditions the diode behaves as a time-varying linear circuit element and superposition holds for the RF signal.

If an RF signal at frequency w , is applied to the circuit, the time-varying diode-operating characteristics will cause currents and voltages at all the frequencies $co_1 + nio_0$, $n = 0, \pm 1, \pm 2, \dots$, to be generated. Hence i_x and y , will have the expansions

$$i_x = l_s e^{jw t} + l_s e^{-jw t} \quad , \quad y = y_d + y_n e^{jw t} - y_n e^{-jw t}$$

There are complex-conjugate terms at $-w$. The term $yl_s e^{yw}$ = $y(i_d + I_s)$ = yi_d since I_s is very small. Thus $yl_s e^{yw}$ can be replaced by yi_d . Each term y_n , has the dimensions of a conductance and will be called g_n , that is, $yl_n \sim S_n$. From the basic equation describing the diode small-signal behavior

$$I = Y \frac{x}{n} i e^{jw t} + Y \frac{x}{n} i e^{-jw t} \quad (12.29)$$

$$+ \frac{L}{s} \frac{C_s}{x} \frac{f}{n} vJ(o_{jl} + n<o_0>) e^{jw t} - e^{-jw t} \quad (12.29)$$

The harmonic terms on each side of this equation must be equal (principle of harmonic balancing); so upon putting w , $(s + n)io_0 = w$, $+ mw_p$, where $m = s + n$, we can write

$q = -$ or

$$= f [* \sim, + / (\sim i + T M \sim o) C_m J_{yn} \quad (12.30)$$

In matrix form we get

$$[\gg] = [Y_{nm}][v_m] \quad (12.31)$$

where $Y_{nm} = g_{n,m} + j(\omega_0 + nw_0)C_{n,m}$. The $[Y_{nm}]$ matrix describes the mixing action of the diode. It relates the harmonic amplitudes of the Rp signal current through the diode to the harmonic amplitudes of the Rp signal voltage across the diode junction. Each harmonic current $i_j e^{j\omega t}$ has a constraint imposed on it by the external circuit (embedding network) in that the voltage around any closed loop, including the diode, must sum to zero(Kirchhoff's law) at each frequency $\omega_0 + nw_0$. These circuit constraints will be given later.

In order to determine the elements in the mixer conversion matrix $[Y_{nm}]$, it is necessary to know the function that describes the junction capacitance $C_J(V_J)$ in terms of the junction voltage. The solutions for i_T , j , i_c , and V_J are constrained by circuit relations imposed by the embedding network. The Thevenin equivalent circuit of the mixer for dc currents is shown in Fig. 12.30b. At $w = 0$ we must have

$$V_B = I_0(R_L + R_S) + V_0 \quad (12.32a)$$

At the local-oscillator frequency, we have

$$j = I_{Tz}Z(w_0) + V, \quad (12.32b)$$

$$\wedge = I_{Tx}Z(-\omega_0) + V_x = IfcZ^*(\omega_0) + v^* \quad (12.32c)$$

At all other harmonics of the local-oscillator frequency

$$0 = I_{Tn}Z(n\omega_0) + V_n \quad (12.32d)$$

These equations along with the known function $C(y)$ and the ideal diode equation must be solved in order to find the current amplitudes $i_{,,}$, the voltage amplitudes $V_{,,}$, and the capacitance elements $C_{,,}$. Numerical methods are required because of the nonlinear behavior of the diode.

After the elements of the mixer conversion matrix have been found, we can solve the small RF signal problem by solving a linear network problem. The embedding network can be replaced by a Norton equivalent network at each frequency $\omega_l + m\omega_0$ as shown in Fig. 12.31. The RF signal source can be chosen as $I_{eo}e^{j\omega l}$. The circuit equation at the frequency w , is

$$\frac{v_n}{Z(f\omega_l)} = \frac{\wedge}{Z(i\omega_x)} + \frac{y}{m} \quad (12.33a)$$

At the frequency $\omega_0 + mo_0$ the circuit equation is

$$\frac{v_n}{Z(i\omega_v + nw_0)} - L Y_{nm} v_m = 0 \quad (12.33b)$$

At the frequency $w + nw_0$ the current through the impedance $Z(o\omega_l + no\omega_0)$

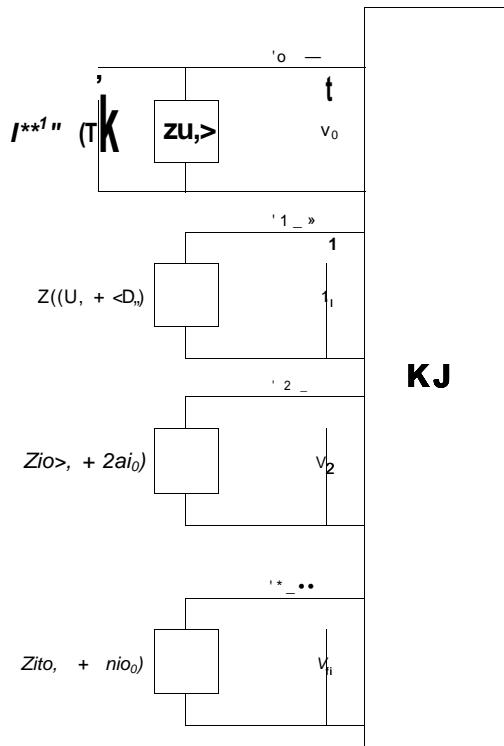


FIGURE 12.31
Equivalent circuit for a mixer for linear RF signal response.

is given by the first term on the left-hand side in (12.336). The current i_n is given by the second term in (12.336). The equation states that the sum of these two currents must be zero.

In principle, the solution of the mixer problem using harmonic balancing is straightforward. In practice, it is a complex problem that requires numerical evaluation using a computer program. Space limitations do not allow us to develop the method in greater detail. However, some of the references given at the end of this chapter provide more details.

- 12.1. The three-port scattering-matrix parameters are given by (12.10). Verify that the sum of the elements in any row or column equals one.
- 12.2. In a transistor a series impedance producing a reflection coefficient F is connected in the common lead. Show that the equivalent two-port network scattering-matrix parameters are given by (12.11) upon replacing $S_{33} + 1$ by $\bullet S_{33} - T$.
- 12.3. For the unstable bipolar transistor in Example 12.2, find the smallest series resistance to be added in the common emitter lead to obtain an absolutely stable two-port network. Use the computer program TRIPORT.

- 12.4.** Design an oscillator following the procedure described in Example f_j assume that $\beta_c = 0.9/1225^\circ$. For this design it will be necessary to design a matching network that will transform a 50- Ω input termination into a source impedance at the base that corresponds to the chosen value of r .
- 12.5.** Redesign the oscillator in Example 12.3 using a resonator coupling with $n^2 = 2/7$. Specify the lengths of all transmission lines used in the oscillator circuit.
- 12.6.** Design a 10-GHz common gate oscillator using a GaAs MESFET having the following scattering-matrix parameters:

$$S_u = 1.1/1160 \quad S_{11} = 0.16^{\wedge}130^\circ$$

$$S_{21} = 2.7/1 - 62^\circ \quad S_{22} = 1.4/1 - 57^\circ$$

Use a pure reactive termination for the input (source terminal) and a circuit topology similar to that in Fig. 12.17. The parameters of the stability circles can be found by using the MICROAMP program (arbitrary values for T_c and noise resistance can be entered. Arbitrary gain circles can also be chosen since this information is not used). The output impedance for the chosen value of r , can be found using the TWOPORT program but the scattering-matrix parameters must be entered in rectangular coordinate form (real and imaginary parts).

- 12.7.** Design an 8-GHz MESFET oscillator similar to that in Example 12.3 using a device with the following common source scattering-matrix parameters:

$$S_{11} = 0.25^{\wedge}8^\circ \quad S_{12} = 0.11^{\wedge}145^\circ$$

$$S_{21} = 3.5/1168^\circ \quad S_{22} = 0.43/1 - 60^\circ$$

This device is absolutely stable so a series reactance must be connected into the common source lead. Choose this reactance to obtain a value for the stability parameter K that is less than 0.6.

- 12.8.** For the oscillator in Fig. 12.19 the dc gate voltage is applied through a high-impedance line $A/4$ long. The feed line is connected at a point where $Z_m = R_m$. Find the distance from the load end at which the feed line should be connected.
- 12.9.** Show that a term such as $V^4 \cos^3 w_c t$ has signal components at the two frequencies ω_o and $3\omega_o$ while $V^4 \cos^4 u_i t$ has components at the frequencies 0 , $2\omega_o$, and $4\omega_o$.
- 12.10.** Show that a term such as $V_o^2 V_2^2 \cos^2 w_o t \cos^2 u_i t$ has signal components at $\omega_o = 0, 2\omega_o, 2\omega_o + 2(\omega_o \pm \omega_2)$.
- 12.11.** Find all of the frequencies generated from the term multiplied by $a_A H$ (12.18).
- 12.12.** When $a_3|V_{IF}|^3 \ll a_2|V_{IF}|$, show that, by using (12.20) to evaluate the term $-fa_3|V_{IF}|^2$ in the expression for the diode current, in place of (12.20) the input voltage is given by (12.22). 2
- 12.13.** When (12.19) with terms up to a_2 is used to evaluate i_{IV} and v_{IV} and V_{ffv} show that the term $-3a_2((\frac{1}{2}V_{IF}^2)^2 - \frac{1}{2}V_{IF}^2)$ results in the following contribution to the intermodulation component of the diode current in the If

amplifier passband:

$$\frac{\sim^{3q3} \times V_{IF}}{2|1+f_1, Z_{IF}|^2}, [V_2^2 V_2 \text{ consta} - 2w, + wz)i + V, V_2^2 \text{ consta} - 2\ll, + \text{tol})r] \frac{\sim^{V \circ |Z|} J}{2(1+n, Z_{IF})}$$

Assume that Z_K is constant in the passband.

- 12.14.** A mixer has a noise figure of 8 dB and a conversion loss of 10 dB. The If amplifier noise figure is 3 dB. Find the noise figure of the combination.
- 12.15.** Find the noise figure of a mixer and IF amplifier combination if the conversion loss is 5 dB. the mixer noise figure is 6 dB, and the IF amplifier noise figure is 4 dB.

REFERENCES

1. Chang, K.: "Handbook of Microwave and Optical Components. Microwave Solid State Components." vol. 2, John Wiley & Sons. Inc., New York, 1990,
2. Vendein, G. D., A. M. Iavio, and U. L. Rohde: "Microwave Circuit Design Using Linear and Non-Linear Techniques," John Wiley & Sons. Inc., New York, 1990.
3. Bahl I., and P. Bhartia: "Microwave Solid-State Circuit Design," John Wiley & Sons, Inc.. New York, 1988.
4. Maas, S. A.: "Microwave Mixers," Artech House Books, Norwood. Mass., 1986.
5. Maas. S. A.: Two-Tone Intermodulation in Diode Mixers. *IEEE Trans.*, vol MTT-35. pp. 307-314. 1987.
6. Koilberg, E. L. ted.>: "Microwave and Millimeter Wave Mixers," IEEE Press, Piscataway. N.J., 1984.
7. Kerr, A. R.: A Technique for Determining the Local Oscillator Waveforms in a Microwave Mixer, *IEEE Trans.*, vol. MTT-23, pp. 828-831. 1975.
8. Held, D N.. and A. R. Kerr: Conversion Loss and Noise of Microwave and Millimeter-Wave Mixers. I. Theory. *IEEE Trans-*, vol. MTT-26. pp. 49-55, 1978.

APPENDIX **I**

USEFUL RELATIONS FROM VECTOR ANALYSIS

1.1 VECTOR ALGEBRA

Let vectors A and B be expressed as components along unit vectors a_1, a_2, a_3 in a right-hand orthogonal coordinate system. Then

$$A \pm B = (A_1 \pm B_1)a_1 + (A_2 \pm B_2)a_2 + (A_3 \pm B_3)a_3 \quad (1.1)$$

$$A \cdot B = |A| |B| \cos \theta = A_1 B_1 + A_2 B_2 + A_3 B_3 \quad (1.2)$$

where θ is the angle between A and B.

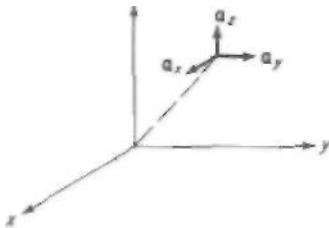
$$\begin{aligned} A \times B &= a_1(A_2 B_3 - A_3 B_2) + a_2(A_3 B_1 - A_1 B_3) \\ &\quad + a_3(A_1 B_2 - A_2 B_1) \end{aligned} \quad (1.30)$$

$$|A \times B| = |A||B|\sin\theta \quad (1.36)$$

$$A \cdot B \times C = A \times B \cdot C = C \times A \cdot B \quad (1.4)$$

$$A \times B = -B \times A \quad (1.5)$$

$$A \times (B \times C) = (A \cdot C)B - (A \cdot B)C \quad (1.6)$$

FIGURE 1.1
Rectangular coordinates.

1.2 VECTOR OPERATIONS IN COMMON COORDINATE SYSTEMS

Rectangular Coordinates

$$\nabla \cdot \mathbf{A} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \quad (1.7)$$

$$\operatorname{div} \mathbf{A} = \nabla \cdot \mathbf{A} = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z} \quad (1.8)$$

$$\operatorname{curl} \mathbf{A} = \nabla \times \mathbf{A} = a_x \hat{i} + a_y \hat{j} + a_z \hat{k} \quad (1.9)$$

$$\nabla^2 \mathbf{A} = \frac{\partial^2 a_x}{\partial x^2} + \frac{\partial^2 a_y}{\partial y^2} + \frac{\partial^2 a_z}{\partial z^2} \quad (1.10)$$

$$\nabla^2 \mathbf{A} = a_x \nabla^2 \mathbf{A}_x + a_y \nabla^2 \mathbf{A}_y + a_z \nabla^2 \mathbf{A}_z \quad (M1)$$

Cylindrical Coordinates

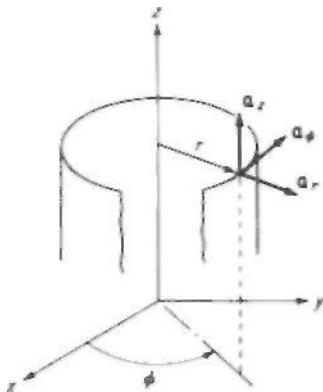
$$\nabla \cdot \mathbf{A} = \frac{1}{r} \frac{\partial a_r}{\partial r} + \frac{1}{r} a_\theta^* \frac{\partial a_\theta}{\partial \theta} + a_z^* \frac{\partial a_z}{\partial z} \quad (1.12)$$

$$\mathbf{V} \cdot \mathbf{A} = \frac{1}{r} a_r \frac{\partial}{\partial r} (r A_r) + \frac{1}{r} a_\theta \frac{\partial}{\partial \theta} a_\theta + a_z \frac{\partial}{\partial z} a_z \quad (1.13)$$

$$\nabla \cdot \mathbf{A} = \frac{1}{r} a_r \frac{\partial a_r}{\partial r} + \frac{1}{r} a_\theta \frac{\partial a_\theta}{\partial \theta} + a_z \frac{\partial a_z}{\partial z} \quad (1.14)$$

$$\nabla^2 \mathbf{A} = r r \cdot \mathbf{A} - \mathbf{v} \times \mathbf{r} \times \mathbf{A}$$

Note that $\nabla^2 \mathbf{A} \neq a_r \cdot \nabla^2 \mathbf{A}_r + a_\theta \cdot \nabla^2 \mathbf{A}_\theta + a_z \cdot \nabla^2 \mathbf{A}_z$ since $\nabla^2 a_r A_r \neq a_r \nabla^2 A_r$, etc., because the orientation of the unit vectors a_r, a_θ, a_z varies with the coordinates r, θ, z .

FIGURE 1.2
Cylindrical coordinates.**Spherical Coordinates**

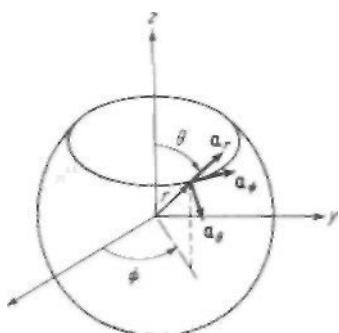
$$\nabla^* = \mathbf{a}_r \frac{dQ}{dr} + \mathbf{a}_\theta \frac{1}{r} \frac{dP}{d\theta} \quad (1.17)$$

$$\frac{1}{r^2} \frac{d}{dr}, \quad \frac{1}{r \sin \theta} \frac{d}{d\theta} (\sin \theta A_\theta), \quad r \sin \theta \frac{d}{d\phi} \quad (1.18)$$

$$\nabla \times \mathbf{A} = \frac{1}{r \sin \theta} \left(\mathbf{A}_\theta \sin \theta - \frac{dA_r}{dr} \right) \hat{\mathbf{a}}_r + \frac{1}{r \sin \theta} \frac{dA_\theta}{dr} \hat{\mathbf{a}}_\theta + \frac{1}{r} \frac{dA_\phi}{dr} \hat{\mathbf{a}}_\phi \quad (1.19)$$

$$\nabla^2 \mathbf{A} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \mathbf{A}}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\frac{1}{\sin \theta} \frac{\partial \mathbf{A}}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \mathbf{A}}{\partial \phi^2} \quad (1.20)$$

$$\nabla^2 \mathbf{A} = \nabla \cdot \mathbf{V} \mathbf{A} - \mathbf{V} \times \nabla \times \mathbf{A} \quad (1.21)$$

FIGURE 1.3
Spherical coordinates.

3 VECTOR IDENTITIES

$$\nabla(\$A) = \nabla A + \nabla \cdot A \quad (1.22)$$

$$\nabla \cdot (\wedge A) = A \cdot \nabla + \operatorname{tfr} V \cdot A \quad (1.23)$$

$$\nabla \cdot (A \times B) = (V \times A) \cdot B - (V \times B) \cdot A \quad (1.24)$$

$$V \times (e \wedge A) = (V \times A) \times A + (A \times V) A \quad (1.25)$$

$$V \times (A \times B) = AV \cdot B - BV \cdot A + (B \cdot V)A - (A \cdot V)B \quad (1.26)$$

$$\begin{aligned} V(A \cdot B) &= (A \cdot V)B + (B \cdot V)A \\ &\quad + A \times (V \times B) + B \times (V \times A) \end{aligned} \quad (1.27)$$

$$V \cdot YS = V^2 * \quad (1.28)$$

$$V \cdot V \times A = 0 \quad (1.29)$$

$$V \times V^* = 0 \quad (1.30)$$

$$V \times V \times A = VV \cdot A - V^2 A \quad (1.31)$$

If A and ∇A are continuous functions with at least piecewise continuous first derivatives in V and on S (or on S and the contour C bounding S),

$$\int_V V \cdot P dV = \int_S \nabla \cdot P dS \quad (1.32)$$

$$\int_V V \cdot A dV = \int_S A \cdot dS \quad (\text{divergence theorem}) \quad (1.33)$$

$$\int_V V \times A dV = \int_S A \times N dS \quad dS = n dS \quad (1.34)$$

$$\int_S V \cdot Q dS = \int_C Q \cdot d\tau \quad (1.35)$$

$$\int_C V \times A \cdot d\tau = \int_C A \cdot d\tau \quad (\text{Stokes' theorem}) \quad (1.36)$$

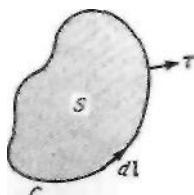


FIGURE 1.4
Surface S bounded by contour C .

GREEN'S IDENTITIES

If \mathbf{A} , \mathbf{B} , \mathbf{P} , and \mathbf{i} are continuous with piecewise continuous first derivatives

$$\int_V (\nabla \cdot \mathbf{D} + \nabla V \cdot \mathbf{D}) dV = \int_S \mathbf{D} \cdot \mathbf{n} dS \quad (1.37)$$

which is Green's first identity. Green's second identity is

$$\int_V (\nabla^2 V - \nabla \cdot \mathbf{D}) dV = \int_S (\mathbf{D} \cdot \mathbf{n} - \nabla V \cdot \mathbf{n}) dS \quad (1.38)$$

In two dimensions (1.37) becomes

$$\int_S (V_{,x}^2 + V_{,y}^2 + V_{,z}^2) dS = \int_C V_{,z} T_z ds \quad (1.39)$$

where $V_{,z}$ is the del operator in two dimensions and T is a unit vector normal to C and in the plane of S . The two-dimensional form of (1.38) is similar.

The vector forms of Green's identities are

$$\begin{aligned} \int_V (\mathbf{f} \cdot \mathbf{v} - (\mathbf{A} \times \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v}) dV &= \int_V [\mathbf{f} \cdot (\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \mathbf{B}) \cdot \mathbf{A} - \mathbf{v} \cdot (\mathbf{v} \times \mathbf{B})] dV \\ &= \oint_S \mathbf{A} \cdot (\mathbf{v} \times \mathbf{B}) dS \end{aligned} \quad (1.40)$$

$$\begin{aligned} \int_V (\mathbf{f} \cdot \mathbf{v} - (\mathbf{A} \times \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v}) dV &= \int_V [\mathbf{f} \cdot (\mathbf{A} \times \mathbf{v} \times \mathbf{B}) - \mathbf{B} \cdot (\mathbf{v} \times \mathbf{A})] dV \\ &= \oint_S (\mathbf{A} \times (\mathbf{v} \times \mathbf{B}) - \mathbf{B} \times (\mathbf{v} \times \mathbf{A})) \cdot \mathbf{n} dS \end{aligned} \quad (1.41)$$

APPENDIX IT

BESSEL FUNCTIONS

1 ORDINARY BESSEL FUNCTIONS

The wave equation and Helmholtz's and Laplace's equations are separable in cylindrical coordinates. The differential equation describing the radial dependence of the solution is *BesseVs* differential equation. *BesseVs* equation is

$$\frac{1}{r} \frac{d}{dr} \left(r \frac{df}{dr} \right) + n^2 f = 0, \quad r \sim J$$

When k^2 is real and positive, the two independent solutions of Bessel's equation are called Bessei functions of the first and second kind, denoted by $J_n(kr)$ and $Y_n(kr)$, respectively. These solutions may be expressed as power series as follows:

$$\begin{aligned}
 J_n(kr) &= \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(n+m)!} \langle \frac{d^m}{dr^m} J_n(kr) \rangle_{r=0} \\
 Y_n(kr) &= \frac{2}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \left[\frac{J_{n+1}(kr)}{J_n(kr)} - \frac{1}{2} \sum_{k=0}^{m-1} \frac{(n-m-k)!}{k!} \right] T_m(kr) \\
 &\quad + \frac{2}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{(kr/2)^{2m}}{m!(n+m)!} \\
 &\quad \times i \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{1}{(3k+1)!} + \frac{1}{m!} i \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{1}{(3k+2)!} + \cdots + \frac{1}{(n+m)!} i \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{1}{(3k+n+1)!}
 \end{aligned} \tag{II 3}$$

where $y = 0.5772$ is Euler's constant. The subscript n denotes the order of the function and is usually an integer in physical problems. The Y_n functions become infinite at $r = 0$. For large values of kr , the Bessei functions approach damped sinusoids:

$$\lim_{r \rightarrow \infty} J_n(kr) = \frac{z}{irkr} \cos kr - \frac{-rr}{4} \frac{nrr}{2} \quad (II-4a)$$

$$\lim_{r \rightarrow \infty} Y_n(kr) = \frac{I}{irkr} \sin kr - \frac{IT}{4} \frac{nn}{2} \quad (11.46)$$

A few of the lowest-order Bessei functions are plotted in Fig. II. 1.

To represent radially propagating waves, linear combinations of the J and T , are formed, called Hankel functions of the first and second kind. Thus the Hankel function of the first kind is

$$W_n(kr) = J_n(kr) + jY_n(kr) \quad (11.5a)$$

and the Hankel function of the second kind is given by

$$H^{\wedge}kr = J_n(kr) - jY_n(kr) \quad (11.56)$$

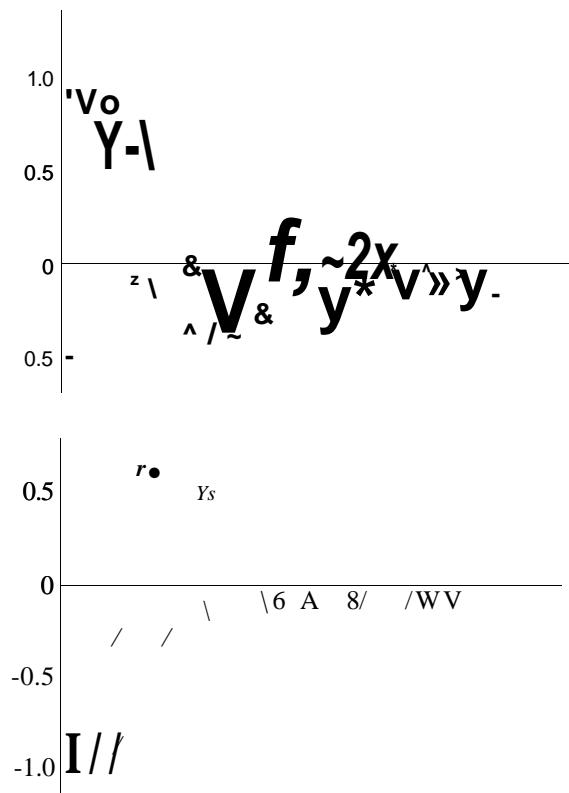


FIGURE H.1
Ordinary Bessei functions.

For large values of kr , the Hankel functions are given by the following expressions:

(11.6a)

$$-j(Ar - 7r/4 - mr/2) \quad (11.66)$$

Some useful relations that hold for any of the Bessel functions J_n , Y_n , or H_n are given below, where $Z_{\cdot\cdot}$ denotes $J_{\cdot\cdot}$, $Y_{\cdot\cdot}$, or $H_{\cdot\cdot}$.

$$xZ'_{\cdot\cdot}(x) = nZ_n(x) \quad -xZ_{nl}(x) = -nZ_n(x) + xZ_{\cdot\cdot\cdot\cdot}(x) \quad (II.7)$$

where the prime denotes differentiation with respect to x .

$$f \quad Z_n(kx)Z_n(lx)xdx = ^{\wedge-\wedge}[kZ_a(lx)Z_{n+1}(kx)] - lZ_n(kx)Z_{\cdot\cdot\cdot\cdot}(lx) \quad (II.8)$$

$$f \quad Z'_n(kx)xdx = ^{\star-2}Z'_n(kx) + \langle I - \frac{1}{k^2x^2} \rangle ZZ(kx) \quad (II.9)$$

2 MODIFIED BESSEL FUNCTIONS

When A^2 is negative, k is pure imaginary. If we let $k = jh$, the solutions are given by $J(jhr)$ and $Y_n(jhr)$. However, for convenience, new modified Bessel functions are introduced and denoted by $I_{\cdot\cdot}(hr)$ and $K_{\cdot\cdot}(hr)$. The modified Bessel function of the first kind is $I_{\cdot\cdot}(hr)$, and is given by

$$I_n(hr) = j''J_n(jhr) = j''J_n(-jhr) \quad (11.10)$$

and the modified Bessel function of the second kind is given by

$$\underline{K_j(hr)} = \langle j^{nil}[J_n(jhr) + jY_n(jhr)] \rangle = ^{\wedge}J^{\wedge}B\&kr \quad (H.11)$$

For large values of hr we have

$$Uhr) \sim \bullet J2irhr \quad (II.12a)$$

$$K_n(hr) \sim \frac{77}{2hr} e^{-hr} \quad (11.126)$$

The first few modified Bessel functions are plotted in Fig. II.2.

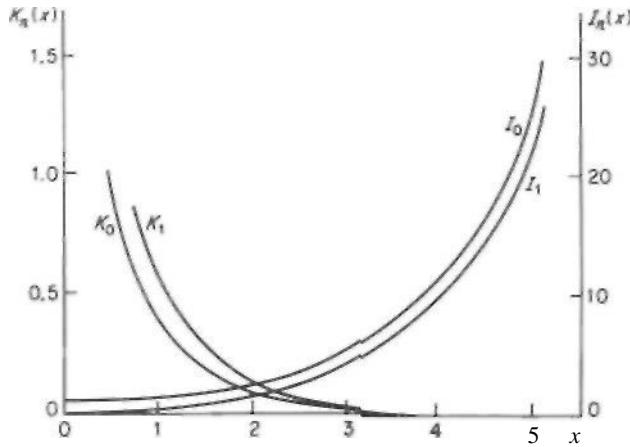


FIGURE II.2
Modified Bessel functions.

A number of useful relations that hold for the modified Bessel functions are given:

$$xI'_n(x) = nl_n(x) + xl_{n+l}(x) = -nl_n(x) + *_{l,n-l}(* \quad \text{(II.13a)}$$

$$Io(x) = I_l(x) \quad \text{(II.136)}$$

$$fx''I_n + I_l(x)dx = x''I_n(x) \quad \text{(II.140)}$$

$$fx''I_{n-l}(x)dx = x''I_n(x) \quad \text{(II.146)}$$

When $n > -1$, we have

$$\int (kx)I_n(lx)xdx = -^{\wedge}[kl_n(lx)l_n(kx)] - ll_n(kx)l_n + l(lx) \quad \text{(II.15)}$$

$$\int fZ(kx)xdx = -^X_r n|kx) - | + k^2 x^2 n(kx) \quad \text{(II.16)}$$

$$xK'_n(x) = nR_n(x) - xK_{n+l}(x) = -nK_{n-l}(x) - xK_{n-l}(* \quad \text{(II. 17a)}$$

$$K'_0(x) = -K^0 x \quad \text{(11.176)}$$

$$\int fx''K_r^{\wedge}l(x)dx = -x''K_n(x) \quad \text{(II.18a)}$$

$$\int fx''K_{n-l}(x)dx = -x''K_n(x) \quad \text{(11.186)}$$

When $\operatorname{Re}(k + 1) > 0$, we have

$$iy, Akx)K_{n+1}(lx)dx = -\frac{1}{2} \int_0^{\infty} [kK_n(lx)K_n(kx) - lK_{n+1}(kx)K_n(lx)] dx \quad (11.19)$$

For $\operatorname{Re} k > 0$

$$\int_0^{\infty} x^n K_{n+1}(kx) dx = \frac{x^{n+1}}{k} K'_{n+1}(x) - \frac{1}{k} \int_0^{\infty} x^{n+1} K_{n+1}'(x) dx \quad (11.20)$$

REFERENCES

1. Watson, G. N.: "Theory of Bessel Functions." Cambridge University Press, New York, 1922.
2. McLachlan, N. W.: "Bessel Functions for Engineers," 2d ed., Oxford University Press, Fair Lawn, N.J., 1948.
3. Bowman, F.: "Introduction to Bessel Functions," Dover Publications, Inc., New York, 1958.

APPENDIX in

CONFORMAL MAPPING TECHNIQUES

The determination of the distributed capacitance and inductance of a transmission line requires a solution of Laplace's equation in two dimensions. For transmission-line structures like those used in many planar transmission lines, it is difficult to construct solutions for Laplace's equation if one used the system of coordinates that is the natural one for describing the transmission-line configuration. A powerful method for solving two-dimensional potential problems is to use conformal mapping to map the boundaries into a simpler configuration for which solutions to Laplace's equation are easily found. This conformal mapping technique is equivalent to a coordinate transformation and its application to planar transmission lines is described in this appendix.

ni.1 CONFORMAL MAPPING

Let $\xi = cr + jri$ be a complex variable and consider the function

$$\sin \xi = \sin(c r + j r) = \sin c r \cosh r j + / \cos c r \sinh r j$$

This function is periodic along the real a axis with a period of $2ir$. The function takes on all its possible values in a strip extending from $-ir/2$ to $ir/2$ along cr as shown in Fig. III.1. Along the contour labeled $A-B-C-D$ shown in the figure, the $\sin \xi$ function goes from -1 to 1 and then to -1 and finally to 1 . If we let a new complex variable W be defined by

$$W = u + jv = \sin \xi \quad (\text{If1.1})$$

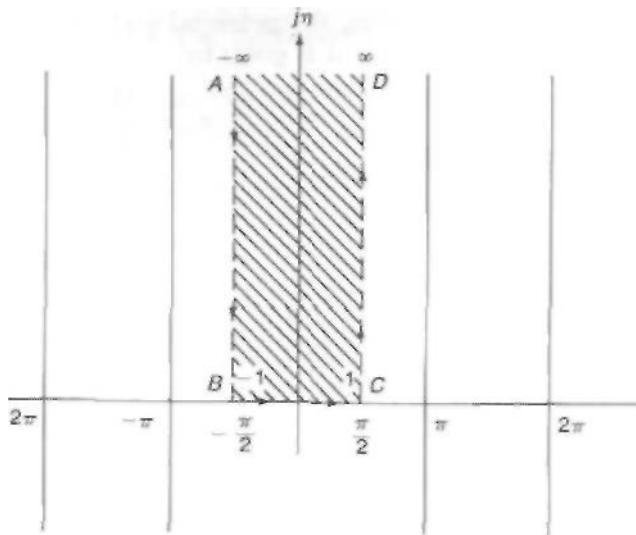


FIGURE III.1
The $\sin f$ function.

then the chosen contour in the f plane maps into the real axis in the W plane as shown in Fig. III.2. All values of f in the cross-hatched region shown in Fig. III.1 map into the upper half of the W plane.

In the W plane, which we will regard as our real physical space, let a total charge Q per meter be placed on each side of the strip extending from $u = -1$ to $u = 1$. Furthermore, let us regard the strip as a conducting strip at constant potential. By symmetry the two boundaries $-\infty < u < -1$ and $1 < u < \infty$ are magnetic walls on which $df/dv = 0$. We can find the solution for Laplace's equation

$$\frac{\partial^2 K}{du^2} + \frac{\partial^2 K}{dv^2} = 0$$

/v

-1

-u

FIGURE III.2
The mapping $W = \sin f$.

in the W plane by solving the potential problem in the ξ plane. In the W plane the potential gradient is given by

$$V<\xi> = \frac{d\xi}{du} a,$$

If we regard u and v as rectangular coordinates, then a and represent new curvilinear coordinates. In the o, 17 coordinate system Laplace's equation has the form

$$\frac{1}{h_a h_{\alpha}} \frac{\partial^2}{\partial a^2} V<\xi> - \frac{1}{h_q h_{\beta}} \frac{\partial^2}{\partial q^2} V<\xi> = 0$$

where k_v and h_{IT} are the metric coefficients. For a coordinate transformation that is obtained through a conformal mapping, the metric coefficients are equal and given by

$$\mathbf{K} = \mathbf{h}_n = \frac{dW}{dC}$$

Consequently, Laplace's equation reduces to the same form as it has in a rectangular coordinate system. Thus a and ξ can be treated as rectangular coordinates when solving for the potential field. However, the gradient of $V<\xi>$ must be evaluated using

$$\nabla^* = \frac{1}{h_{\alpha} da} \frac{d}{da} \mathbf{a}_{\alpha} + \frac{1}{h_{\beta} dr} \frac{d}{dr} \mathbf{a}_{\beta} = \mathbf{V}\xi = \frac{dC}{dW} = \mathbf{V}\xi_w \quad (\text{HI.2})$$

where $\mathbf{V}\xi^*$ represents the gradient in the ξ plane obtained by treating a and ξ as rectangular coordinates. The conformal mapping of the boundary of a polygon in the ξ plane into the real axis in the W plane is called a Schwarz-Christoffel transformation.

For a differential element dW along a contour in the W plane, the corresponding differential element along the mapped contour in the ξ plane is $d\xi = (dt/dW)dW$. The angle that $d\xi$ makes with the real axis is the sum of the angles of $d\xi/dW$ and dW . When the contour in the W plane is the real axis, the angle of dW is zero. When $d\xi/dW$ has the form $i/W - W_1 / JW^*W_2$, then the angle of $jW - W_x$ is $\pi/2$ for $W < W_1$, and zero for $W > W_2$, while that for $1/JW - W_2$ is $-\pi/2$ for $W < W_2$ and zero for $W > W_2$. Hence, as W moves along the real axis, there will be a change of $\pi/2$ in the angle of $d\xi$ as W moves past the point W_1 , and a change of $\pi/2$ when W moves past the point W_2 . This causes the contour in the ξ plane to change direction in a step-like fashion by $+90^\circ$. This property &

used to pick the appropriate mapping functions that are used in this appendix.

We will use the above results to find the charge distribution along the conducting strip in the W plane. In the ξ plane the presence of magnetic walls along $a = \pm iT/2$, $\Pi > 0$ requires that the electric field be uniform and in the $r\backslash$ direction. Hence the charge is uniformly distributed on the lower boundary $-iT/2 < a < iT/2$ with density Q/\sim . Thus the potential gradient is given by

$$\sim \xi \bullet J_r \sim = \frac{Q}{iT}$$

so $\nabla V \sim | = -Q/T \epsilon_0$. From (III.2) we find that in the W plane

$$dt,$$

Upon using $dW/d\xi = \cos \theta = (1 - \sin^2 \theta)^{1/2} = (1 - W^2)^{1/2}$ for $W = u$, we obtain

$$p_s = \frac{Q}{TTVI - u^2} \quad (\text{in-3})$$

for the charge density on one side of an isolated conducting strip two units wide and located in air or completely surrounded by dielectric. This example clearly shows the power of the conformal mapping method of solving potential problems. However, it does depend on our ability to find a mapping function that will transform the boundaries of the physical problem into a configuration for which the solution to Laplace's equation is easily obtained. In the transformed plane the new coordinates may be treated as rectangular coordinates provided we relate the gradients in the two planes by the relation (III.2). Furthermore, the capacitance between conductors remains invariant under a conformal mapping, so that it may be found for the transformed configuration.

2 ELLIPTIC SINE FUNCTION

For the function $W = \sin \xi$, we have $dW/d\xi = \cos \xi = (1 - W^2)^{1/2}$, so

$$\frac{dt}{dw} = \frac{1}{\sqrt{1 - w^2}} \quad (\text{III.4})$$

which gives the inverse function

$$Z = \sin^{-1} W = \frac{w}{\sqrt{1 - w^2}}, \quad dW \quad (\text{III.5})$$

The period along the w axis is obtained from the integral from 0 to 1 which

gives one quarter of the period, i.e.,

$$2 = \frac{dW}{L/I_w} \quad (\text{III.6})$$

A function that is much more useful in the solution of a number of planar-transmission-line problems is a function that is periodic along both the r and $-q$ axis in the complex C plane because it would take on all of its possible values inside a rectangle. A doubly periodic function is the elliptic sine function, which is expressed as

$$W = sn(i, k) \quad (\text{IH.7})$$

The parameter k is called the modulus and determines the two periods. The elliptic sine function has a period of $4K$ along u and $2K'$ along τ . The inverse function is given by

$$\{ = sn^{-1}(W, k) = \int_0^W \frac{dW}{\sqrt{(1 - W^2)(1 - k^2 W^2)}} \quad (\text{in.8})$$

The two quantities K and K' are given by

$$K = \frac{\int_0^{i/k} \frac{dW}{\sqrt{(1 - W^2)(1 - k^2 W^2)}}}{\int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}} \quad (\text{III.9a})$$

$$\text{or } K' = \frac{\int_0^{i/k} \frac{dW}{\sqrt{(W^2 - l)(l - k^2 W^2)}}}{\int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}} \quad (\text{III.9c})$$

By using the substitution $W = \sin \theta$, the expression for K becomes

$$K = K(k) = P' \cdot \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = F(k, \frac{\pi}{2})$$

which is the complete elliptic integral of the first kind. In the expression for K' we can let $k\sqrt{W^2 - 1} = (1 - k^2) \cos \theta$, which reduces the expression to the form

$$K - K(k) = f'^2, \quad \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - f'^2 \sin^2 \theta}} = Flk, \quad J = K(k') \quad (\text{ffl.11})$$

where the complementary modulus $k' = \sqrt{1 - k^2}$. From (III.8) and (III.9) we also see that

$$sn(0, *) = 0 \quad (\text{111.120})$$

$$sn(K, k) = l \quad (\text{ffl.126})$$

$$sn(K + jK', k) = \wedge \quad (\text{ffl.12}^\wedge)$$

A further useful value is

$$\operatorname{sn}(jK', k) = \pm 0 \quad (\text{III.12d})$$

The elliptic sine function is a generalization of the sine function and reduces to the latter when $k = 0$.

Two useful approximations for the K and K' parameters for the cases $k \gg 1$ and $k' \ll 1$ are

$$\operatorname{tf}(A) = *'(*') = - | 1 + T + - * < k < 0.4 \quad (\text{III.13a})$$

$$/T(*) = *(*) = / 1 + - \frac{k^2}{+} - \frac{9}{+} - H \ln -$$

$$\frac{k^{12}}{4} \frac{21}{168} k^* \quad k > 0.65 \quad (\text{III.13b})$$

Note that k and k' can be interchanged. The ratio of $K(k)/K'(k) = K(k)/K(k')$ can be evaluated to an accuracy of one part in 10^5 using the following expression:¹

$$\frac{K}{K'} = \frac{1}{\pi} \operatorname{In} \left(\frac{1 + Jk}{2 - lk} \right)^{\frac{1}{2}} \quad 0.7 < k \leq 1$$

$$= \frac{1}{\pi} \operatorname{In} \left(\frac{1 + Jk}{2 - lk} \right)^{\frac{1}{2}} \quad 0 < A < 0.7 \quad (\text{III.13c})$$

For intermediate values of k and k' the following formula can be used:

$$K' = K(k') = \frac{2}{\pi} \operatorname{In} \left(\frac{1 - k}{1 + k} \right)^{\frac{1}{2}} \quad \text{if } 0 < k < 0.7 \quad (\text{III.13d})$$

where k and k' can be interchanged. When k' is less than 0.707, then $(1 - k)/(1 + k)$ is always less than 0.172 and (III.13a) can be applied. When k' is greater than 0.707, we can use (III.13c).

Consider now the rectangle shown in the t plane in Fig. III.3. Also shown are the values of $\operatorname{sn}(\xi, k)$ at the points labeled A, B, C, D, E, F . The mapping $W = \operatorname{sn}(\xi, k)$ maps the rectangle into the real W axis and all interior values into the upper half of the W plane. If the segments AB , CDE , and FG are conducting metal boundaries, then the configuration in the W plane is a coplanar transmission line in air. Its distributed capacitance is twice that between the plates CDE and BAF for the ideal parallel-

¹T. S. Gradshteyn and T. M. Ryzhik, "Table of Integrals, Series, and Products," Academic Press, Inc., New York, 1965, p. 925. formula 8.198 and using $q = \exp(-trK'/K)$.

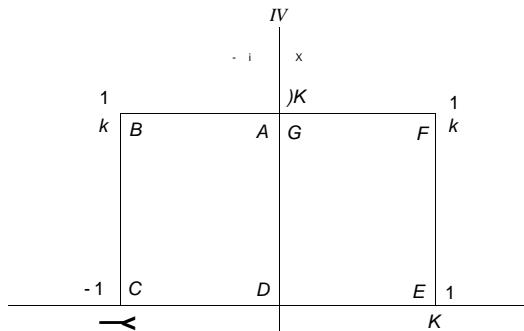


FIGURE m.3
The $\text{sn}(f, k)$ function.

plate capacitor in the f plane and hence is given by

$$C = \frac{2K}{2e_0} = \frac{K}{4\epsilon_0} \quad (111.14)$$

The relevant parameter for the coplanar line is the ratio of the width of the center strip to the spacing between the two ground planes which is given by $2/2u_i = k$, the modulus of the elliptic sine function. Note that in the f plane the boundaries BC and EF are magnetic walls so there is no fringing capacitance.

If we choose the segments BC and EF to be conducting strips, we obtain a coplanar strip transmission line. Its distributed capacitance is that associated with the ideal parallel-plate capacitor with the boundary BC at potential V and EF at potential $-V$; thus

$$C = \frac{K'}{2\epsilon_0 \cdot \frac{K}{2}} = \frac{K'}{2\epsilon_0 K} \quad (111.15)$$

When we let k tend toward zero, the points $\pm u_i$ move out to infinity and we obtain a slot line. However, we cannot let $k = 0$ because the capacitance between two semiinfinite planes separated by a slot becomes infinite.

m.3 CAPACITANCE BETWEEN TWO PARALLEL STRIPS

We will now consider the problem of two parallel strips in air as shown in Fig. III.5. The plate separation is $1H$ and the plates have a width $2w$. The following mapping will map the boundary shown by the dashed fine in Fig. III.5 into the real axis in the W plane shown in Fig. III.4:

$$= \int_o^1 \frac{(1 - k*W^*)}{(1 - W^2)(1 - k^2W^2)}$$

where $k_0 = 1/\epsilon_0$ and $k = 1/u_i$. When $Z = H$ on the interior side of the

/"

- 1

$$u_0 - y - \infty,$$

 $\rightarrow u$

FIGURE m.4
The mapping $W = \operatorname{sn}(f, ft)$.

right-hand strip, $W = 1$; so

$$\begin{aligned} & i - k l w^2 \\ v \operatorname{Ti} & - w^2)(i - k^z w^2) \end{aligned} \quad (\text{III. 17a})$$

At the outer edge of this plate, $Z = H + jw$, and we make this correspond to the unknown point u_{∞} , lying between 1 and u_0 ; thus

$$H = \frac{1 - k^2 W^2}{(1 - W^2)(1 - k^2 W^2)} dW \quad (\text{III. 176})$$

At the point $Z = H$ on the outer side of the strip, we set $W = \infty$; thus

$$H = \frac{1 - k^2 W^2}{j_o - / (1 - W^2)(1 - k^2 W^2)} dW \quad (\text{III. 17c})$$

The distributed capacitance between the parallel strips is equal to that for

iy

/w- n

- "J Lc _____
- H

FIGURE III.5
Two parallel strips and the boundary to be mapped.

the coplanar strip line that results from the mapping and is given by
(111.15).

For the parallel strips the capacitance depends only on the ratio w/H . When we compare (III.176) and (III. 17c), we see that we require

$$jw = -f' \cdot \frac{1 - k^2 W^2}{y/(l - W^2)(1 - k^2 W^2)} dW$$

A similar comparison of (IV.17a) and (IV.176) shows that

$$jw = \frac{f''}{h} \cdot \frac{1 - k^2 W^2}{(i - W^2)(1 - k^2 W^2)} dW$$

By combining these two expressions, we obtain

$$\int_{-k^2 W^2}^{k^2 W^2} \frac{dW}{W^2(l - k^2 W^2)} = 0 \quad (\text{III.18})$$

which is the equation that will determine k_0 and hence α_0 . The integral in (III.18) can be separated into two integrals, and by using (111.9c) to replace the integral that goes from 0 to 1, we find that (111.18) can be expressed in the form

$$K' = k^2 f'^K \int_{1/(W^2 - l)}^{1/(l - k^2 W^2)} dW$$

We can express this integral in terms of complete elliptic integrals. However, in practice it is more expedient to evaluate the integral numerically. By using the substitution $kHW^2 - 1 = (1 - fc^2)\cos^2 0$ as before we obtain

$$T - fejf'^V 1 - *W *# = fMk', j) \quad (\text{III.19})$$

where E is the complete elliptic integral of the second kind.

From the above we get

$$\frac{k^2 K'}{E(k', r/2)} \quad (20)$$

The expression for H given by (III. 17a) can be reduced to the form (we use $W = \sin\theta$)

$$H = K - k \frac{\sin^2 \theta}{\sqrt{1 - \sin^2 \theta}} \frac{\sin^2 \theta}{\sqrt{1 - \sin^2 \theta}} d\theta$$

which gives

$$\begin{aligned} H &= K + \frac{k^2}{2} \int_0^k E(k') dk - \int_0^k F(k') dk \\ &= K[E(k, TT/2) - K] + KE(k', Tr/2) \\ &\quad - E(k', ir/2) \end{aligned}$$

For convenience, the complete elliptic integrals of the second kind will be written in the compact form

$$EJ^*, \wedge) = E(k) = E \quad E(*'.?) - *(**) = E(k) = \mathfrak{E}'$$

We now use $E(k', TT/2) = E'(k, TT/2)$ and the following identity:

$$KE' + K'E - J3T = Z$$

which allows us to express H in the simpler form

$$* = 2 E (* W 2) - 2 \wedge \quad (\text{III-21})$$

The last reduction is the expression for u ; to the form

$$\begin{aligned} w &= \frac{1}{TT} \left/ \frac{rs, kl(l - k^2 siD?6) - k^2}{1/1 - fc'^2 \sin^2 0} \right. \text{cifl} \\ &= \frac{kl}{\tilde{E}(k', 8)} - F(k', 6) \end{aligned} \quad (111.22)$$

where $k'_Q = \sqrt{1 - k^2}$ and

$$0 = \cos \frac{\pi ffc/c ft}{k k_0} \quad (111.23)$$

This equation gives w in terms of incomplete elliptic integrals. Values for the complete and incomplete elliptic integrals are available in the book by Jahnke and Emde.¹

The easiest way to use the above formulas is to choose a value for k , solve (111.20) for k_0 , (111.21) for H , and (111.22) for w . The integrals are readily evaluated numerically. By this means values of w/H and the capacitance C given by (III.15) can be compiled as a function of k . The capacitance between a strip of width $2w$ and spaced a distance H above a ground plane is twice as large as that given by (III.15). Table III.1 gives

¹E. Jahnke and F. Emde, "Tables of Functions," Dover Publications, Inc. New York, 1945.

TABLE III.]
Capacitance between a strip of width $2w$
at a height H above a ground plane

2ic H	C ∞	$2w$ H	C ∞
0.05	1.238	1	2.981
0.1	1.436	1.5	3.621
0.2	1.703	2	4.232
0.3	1.912	2.5	4.822
0.4	2.094	3	5.399
0.5	2.26	4	6.529
0.6	2.415	6	7.63
0.7	2.563	6	8.72
0.8	2.706	8	10.86
0.9	2.845	10	12.98

computed values of C/e_0 as a function of $2w/H$. The integrals were evaluated numerically using Simpson's quadrature formula.

III.4 STRIP TRANSMISSION LINE

The strip transmission line is shown in Fig. III.6. The boundary $A-B-C-D-E$ is mapped into the real axis in the W plane by means of the mapping function

$$W = \sin Z = \sin x \cosh y + j \cos x \sinh y \quad (111.24)$$

In the Z plane H is IT units, so we make the point B be $jirw/H - ir/2$ so as to maintain the same width-to-spacing ratio w/H as in the physical strip line. The real axis in the W plane is mapped into a rectangle in the L plane, shown in Fig. III.6d, by the mapping function

$$W = 1 + 2 \operatorname{sn}(f, A)$$

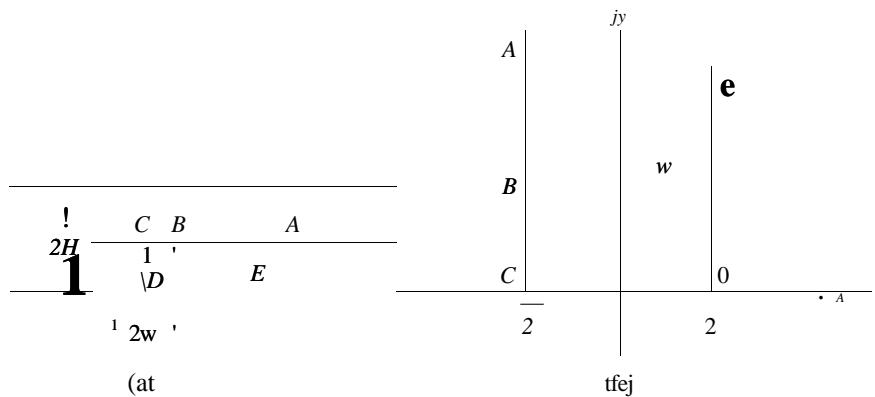
The point B corresponds to $W = -\cosh ITw/H$ and to $\operatorname{sn}(f, k) = \operatorname{sn}(-K + jK', k) = -1/k$. Consequently, we must have

$$\frac{2}{J' I} \cosh \frac{\pi T w}{H} = \frac{rcw}{k}$$

which gives

$$k = \frac{\cosh \pi T w / H}{\operatorname{cosec} \pi T w / H} \quad (111.25)$$

In the C plane the capacitance is given by $e_0 K'/K$ and represents one-quarter of the capacitance of the strip-line configuration. Hence for the strip



>

$$\cosh(jy) - 1$$

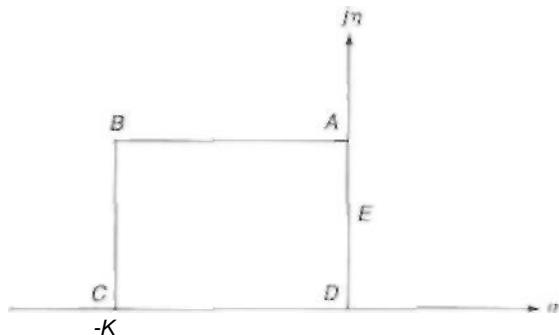
 $|c|$ 

FIGURE III.6
The strip transmission line and its mapping

line

$$\mathbf{C} = 4t \quad K'(k) \quad (\text{III .26})$$

The corresponding characteristic impedance is

$$* = \wedge = mr \frac{K(k)}{K'(k)} \quad (\text{HI .27})$$

m.5 CONDUCTOR LOSS

Since the current density and charge density are related by the continuity equation [see (3.1386)], the current density has the same functional form as the charge density. If we know the current density J_+ , then the power loss is given by

where R_m is the skin-effect surface resistance and the integral is taken around all the contours along the conductor surfaces. If we have infinitely thin conductors with edges, then the current density becomes infinite at the edge inversely proportional to the square root of the distance from the edge. For example, for an isolated strip of width $2w$ the current density, according to (III.3), would be

$$2-rrwy/l - x^2/w^2$$

on each side of the strip. This current density is too singular to allow a finite value for the power loss to be evaluated. Consequently, it will be necessary to take into account the finite thickness of the conductors.

The simplest case to consider is that of an isolated strip of width $2w$ and thickness $2t$ as shown in Fig. III.7. The labeled boundary in the Z plane can be mapped into the real u axis in the W plane using

$$\frac{dZ}{dW} = A \frac{W^2}{1 - k^2 W^2} \quad (111.28)$$

which upon integration gives

$$Z = A \left(\frac{1 - W}{1 - k^2 W^2} \right)^{1/2} dW + jt \quad (\text{III.29})$$

The constant jt is added since $W = 0$ is made to correspond to $Z = jt$. When $W = 1$ we require $Z = w + jt$ and when $W = 1/k$ we want $Z = * u >$. These two conditions provide a solution for A and the modulus k . By following a

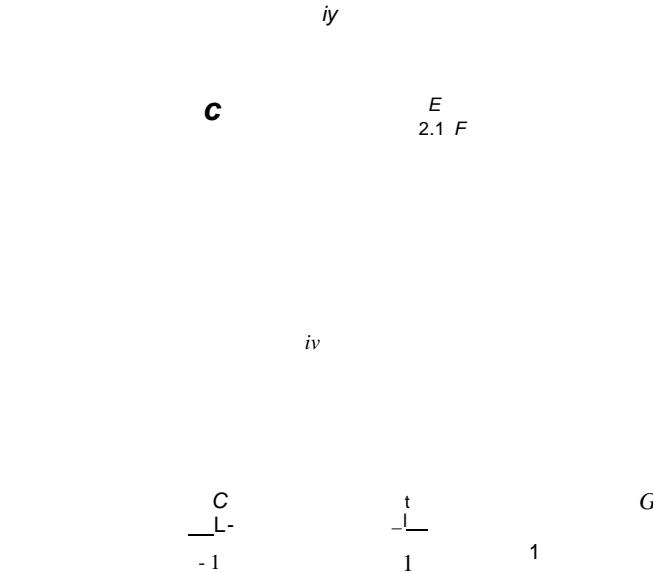


FIGURE III.7

An isolated conducting strip with finite thickness and the mapping of one-half of the strip boundary into the W plane.

procedure similar to that for the parallel strips, we find that

$$A = \frac{wk^2}{E(k)} - k^2 K(k) \quad (\text{III.30a})$$

$$\begin{aligned} t &= E(k') - k^2 K(k') \\ w &= E(k) - k^2 K(k) \end{aligned} \quad (\text{III.30b})$$

In the W plane the conducting strip extends from $-1/k$ to $1/k$ along u . Since it is an isolated infinitely thin strip, the current density in the W plane is of the form

$$J_{\mu\nu} = \frac{A}{A} - k^2 u^2$$

where the factor A has been inserted arbitrarily. In the Z plane the current density is proportional to

$$V4>I_Z = V^*|_w \frac{dW}{dZ}$$

In order to find $J_t(x)$, we need to express u as a function of x , which is not easily accomplished. Fortunately, we will not need this functional relationship.

Let us consider the integration of (111.28) very close to the 90° corner where $Z = w + jt$, of the conducting strip. When $x = w$, $u = 1$; hence

$$\int_{w-x}^{\infty} v \frac{du}{1 - k^2 u^2}$$

For u very close to one, we can use the approximations $1 - u^2 \approx (1 - u)$ and $1 - k^2 u^2 \approx 1 - k^2$. We then obtain

$$w - x = r^A l! \Delta du = 2|T^A(l-u)^{3/2}$$

which gives

$$J_t = \frac{(3A)^{V3} 2^{n/2/3}}{v^A / T^A \cdot \mathbf{f} l! \cdot \mathbf{V}^{-*})^{1/A}} \quad (\text{III.31})$$

This fundamental result shows that close to a 90° corner the current density has a weaker singularity, namely, it is inversely proportional to the cube root of the distance from the corner. This current density, when squared, can be integrated to give a finite result for the power loss. The edge behavior is a local field phenomenon, so that the result obtained is true near any 90° corner.

At some distance away from the corner, the expression $[(1 - u)/(1 - k^2 u^2)]^{1/2}$ can be approximated by unity since k is close to one for a thin strip, i.e., for t/w small. Thus we have

$$\int_0^x dx \sim A \int_0^u du$$

or $x = Au$. Hence away from the corner the current density is given by

$$J_t = \frac{1}{v l} \frac{1}{u^2} \frac{1}{y T} \frac{1}{(x/A)^2} \frac{1}{\mathbf{A}} \sim (x^w Y)$$

since $A \sim w$ then t/w is small. We can summarize the above result with the statement that for a thin conducting strip the current density over most of the strip is the same as for an infinitely thin strip, but as the corner is approached the one over the square root of the distance from the corner behavior changes over to a less singular behavior going as one over the cube root of the distance from the corner.

For the isolated finite thickness strip, we can evaluate the power loss per meter by doing the integration in the W plane. Thus

$$\begin{aligned}
 & \text{P} = -4 / \int_{-w/2}^{w/2} \frac{dx}{l - u^2} \int_{-\infty}^{\infty} \frac{du}{f(l - u^2)(l - k^2 u)} \\
 &= 2R_m A \cdot \frac{du}{\sqrt{f(l - u^2)(l - k^2 u)}} \\
 &= 2R_m A [K(k) + K'(k)]
 \end{aligned} \tag{111.32}$$

The integral gives the loss for one quarter section, so we multiplied by a factor of 4 to get the total loss. The total current on the strip is given by

$$\begin{aligned}
 I = 4 \int_{-w/2}^{w/2} \frac{dx}{l - u^2} \frac{du}{du} &= 4Af \int_{-\infty}^{\infty} \frac{du}{f(l - u^2)} \\
 &= 2\pi A
 \end{aligned} \tag{III.33}$$

The integral was evaluated by using the substitution $ku = \sin \theta$.

The series-distributed resistance R per meter for the conductor is defined by the relationship

$$m^* = p,$$

From the derived expressions we obtain

$$R = \frac{k^2 R_m}{z_A} [K + K'] \tag{111.34}$$

In practice, the ratio t/w is small. For example, a board plated with 1-oz copper has a metalization thickness of 0.036 mm; so a microstrip of width 1 mm will have $t/w = 1/28$. When $t/w < 0.05$ the following approximations are valid since $k \sim 1$:

$$E(k') \sim \frac{1}{4} E(k)$$

$$* < * > - \frac{\pi}{4} | 1 + \frac{it}{4} K(h) \sim ht -$$

By using these approximations in (III.30), we obtain

$$A = k^2 w \quad \text{and} \quad \frac{it}{TTW}$$

We now find that the expression for H reduces to the simple form

$$R = \frac{R}{2TT^2 W} \quad \text{In } \frac{ATTW}{TTW} \tag{111.35}$$

In this expression w is one-half of the strip width and t is one-half of the strip thickness. The part $R_m/2vw$ is the contribution from the currents that flow on the two end faces of the strip. From the expression used for J we find that the current density at the center of the broad face where $u = f_j$ is unity. At the center of the end face where $u = 1/k$, the current density is $k/k' = \sqrt{2w/2t}$ and hence is much larger. It is for this reason that there is a significant contribution to the power loss from the current on the end faces even though t is very small.

If the conductor cross section was approximated by an ellipse with major axis equal to $2w$ and minor axis equal to $2t$, it would be found that the series resistance is

$$R = \frac{y}{TT^2W} \ln \frac{t}{t}$$

when t/w is small. For very small values of t/w , this resistance is twice as large as that for a rectangular bar of width $2w$ and thickness $2t$. The reason is that for the elliptical cross section the current density at the narrow ends is a factor w/t greater than that at the center of the broad face and thus exhibits a more singular behavior than that for the current at the center of the narrow face for a rectangular bar.

A useful approximate way to find the series resistance of a thin conducting strip is to use the current density for an infinitely thin strip but to integrate the square of the current density up to a point a distance d from the edge. The distance d is chosen so that the same resistance as given by (III.35) is obtained. Thus for the isolated strip we use

$$P = 2f_1 4T \int_{-x}^{x-d} \frac{dx}{y w} - * \left(\frac{w+x}{w-x} \right)^{\frac{w-d}{w}} * R_m w \ln \frac{2w}{d}$$

The total current on the strip is $2mv = I$; so by equating $I^2R/2$ to P , we obtain

$$R = \frac{R_m J}{217 \cdot 10} \ln \frac{2w}{d}$$

We now equate this expression to (III.35) and solve for d , which gives

$$d = \frac{2t}{Aire''} \quad 290.8$$

In the derivation we expressed v as $\ln e^\sim$. This method of finding the loss resistance was proposed by Lewin. This method is based on the fact tha

the effect of finite thickness is to change the local behavior of the current near the edge but leaving the current density on the major portion of the conducting strip relatively unchanged.

ni.6 CONDUCTOR LOSSES FOR A MICROSTRIP TRANSMISSION LINE

Consider a microstrip fine with a strip of width $2u$, thickness $2t$, and spaced a distance H above a ground plane. In order to apply the rule established above, we need to know the division of the total current on the upper and lower faces of the strip, which is affected by the presence of the ground plane. Apart from an unequal division of the total current, the density of current is very nearly the same as on an isolated strip. From the mapping function (III.16), we find that the current density is proportional to

$$J_z \propto v \operatorname{Im} \left| \frac{dW}{dz} \right| \cdot \operatorname{Re} \left| \frac{dW}{dw} \right| \propto i - k l w^2$$

At the center of the upper face $u = u_c = 1/k$ and at the center of the lower face $u = 1$. Hence the ratio of the currents on the two faces is

$$J_{z2} : J_{z1} \propto k^2 / k^2 = 1$$

The fraction of the total current on the upper face is p , where p is given by

$$\frac{J_{z1}}{J_{z1} + J_{z2}} = \frac{k^2(l-k^2)}{(1-k^2)^2} = \frac{E-k^2 K'}{(1-k^2)K'} \quad [\text{III.36}]$$

where (III.20) has been used to eliminate K' . The power loss on the upper face is proportional to $(1-p)^2$. The total power loss is proportional to $p^2 + (1-p)^2$. For an isolated strip $p = 1/2$, so that the series resistance of the microstrip in the presence of the ground plane is given by, multiplying R by the ratio $[p^2 + (1-p)^2]/1/0.5$; thus

$$R_x = 2(2p^2 - 2p + 1)R \quad (111.37)$$

where p depends only on the ratio $2w/H$ and R is the same as in (III.35). When $p = 0.25$ we find that $R_x = 1.25R$ which shows that an unequal division of the total current increases the resistance.

The ground plane corresponds to the y axis in Fig. III.5, which maps into the u axis in the VW plane. Thus the current density will be proportional to the gradient along the v axis in the W plane which is given by

$$J_z \propto \frac{1}{1 + 0.5 W^2}$$

Let J_r be chosen as

$$J_r = i + k'y$$

where j_0 is to be found so that the total current will equal I . The total current on the ground plane is given by

$$\begin{aligned} &= 2j_0 / \left[1 + k^2 v^2 \right] \frac{dZ}{dW} dv \\ &= 2j_0 / \left[J(l + k^2 v^2)(l + v^2) \right] dv \end{aligned}$$

The integral is easily done in the g plane by using

$$dv = \frac{|dW|}{dC} d_v = y'a + AV)(i + v^2) d-q$$

and the limits 0 to $K'(k)$ for $-q$. Thus we find that $j_0 = 2I_0 K'(k)$.

The power loss in the ground plane is given by

$$\begin{aligned} P_s &= R_s \frac{I^2}{4K} \int_{-f}^{f} \frac{dZ}{dW} dv \\ &= R_s I^2 \int_{-f}^{f} \frac{dv}{4A''^2 j_0 (1 + k^2 v^2) y / (l + k^2 v^2)(l + v^2)} \end{aligned}$$

The series resistance R_s of the ground plane is obtained by equating the power loss to $T^2 R_s / 2$. The series resistance will depend on the absolute dimensions of the microstrip line. Thus we will multiply P_s by $(2w/H)$ and divide by this same factor but use (111.21) for H . We thereby obtain

$$R_s = \frac{R_m}{2w} \frac{T^2 2w}{4 \sim H} \frac{1}{E^l K^2 jh} \frac{|f|^*}{(l + k^2 v^2) J(l + k^2 v^2)(l + v^2)} dv \quad (III.38)$$

The normalized resistance $2wR_s/R_m$ is a function only of the ratio $2w/H$ and thus only needs to be evaluated once. The relevant expressions for the loss ratio

$$LR = 2(2p^2 - 2p + 1) \quad (111.39)$$

and the normalized resistance $2wR_s/R_m$ have been evaluated numerically as a function of $2w/H$. We have fitted simple polynomial functions to the resultant data, so that for application purposes the loss ratio and normalized resistance can be computed from the following expressions to an

accuracy of 2 percent or better:

$$\text{LR} = 1 - \frac{2w}{2w + 5.8H} < 0.5 \quad (\text{III.40a})$$

$$\text{LR} = 0.94 + 0.134 \frac{2w}{2w/H} - 0.0062 \frac{2w^2}{H} \quad 0.5 < \frac{2w}{H} < 10 \quad (\text{III.406})$$

$$\frac{2w}{2w/H} = \frac{2JO/H}{5.8 + 0.03H/2LV} \quad 0.1 < \frac{2w}{H} < 10 \quad (\text{m.41})$$

This formula states that the resistance of the ground plane is that of a strip of width $2w + 5.8H$ and with a uniform current density. It is remarkable that this relationship holds for such a wide range of $2w/H$ ratios. In practice, surface roughness increases the resistance by 10 to 50 percent so an accuracy greater than a few percent in the theoretical formulas is not needed.

III.7 ATTENUATION FOR A COPLANAR LINE

We consider a coplanar line as shown in Fig. III.8a. The ground planes and center conductor have a thickness $2t$. The strip width is $2a$ and the ground plane separation is $2b$. The upper half of the Z plane will be mapped into the upper half of the W plane with the contour labeled 1-2—8 mapping into

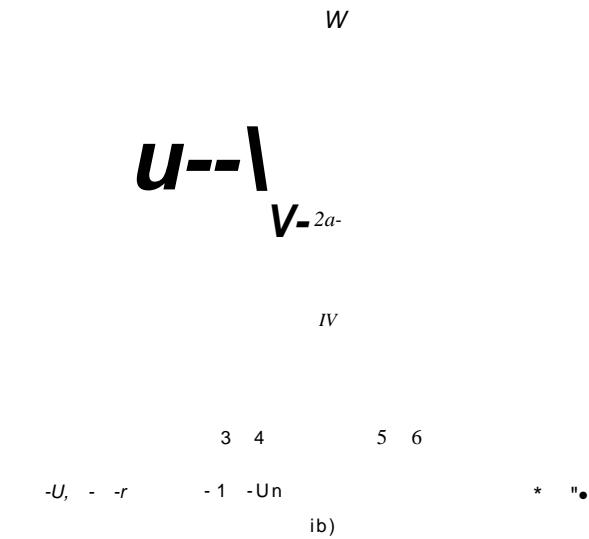


FIGURE m.8

Conformal mapping of the cross section of a coplanar transmission line with finite thickness conductors.

the real axis in the W plane as shown in Fig. III.86. The required mapping function is

$$w = \frac{(l-k^2 W^2)(l+k^2 W^2)}{A} \quad (\text{HI .42})$$

where $k_0 = l/u_0$ and $k_l = 1/\epsilon_r$. The parameters A , k_0 , k , and k_l are determined by requiring that the four points $Z = a + jt$, a , b , and $b + jt$ map into the points $W = u_0$, 1 , l/k , and u_v . These requirements lead to the following four equations:

$$a = A(\overset{U}{\circ} F(W)dW) \quad (\text{III.43a})$$

$$t = jApF(W)dW \quad (\text{III.43b})$$

$$b = -A(\overset{''}{F}(W)dW) \quad (\text{III.43c})$$

$$t = -\frac{jA}{h/k}(\overset{U}{\circ} F(W)dW) \quad (\text{HI.43d})$$

where $F(W)$ is the integrand shown in (111.42). The real axis in the W plane can be mapped into the rectangle in the ξ plane using (III.8). The boundary-value problem in the C plane is that of an ideal parallel-plate capacitor. Hence V_{∞} is a constant which we set equal to one.

The current density J_z is thus given by

$$J_z = V_{\infty} \frac{d\xi}{dZ} = \left| \frac{d\xi}{dZ} \right| = \left| \frac{\mathbf{u}}{dW} \right| \frac{dW}{dZ}$$

Consequently,

$$J_z dZ = \left| \frac{dC}{dW} \right| \frac{dW}{dZ} = \left| \frac{dC}{dW} \right| \mathbf{d}W$$

The total current on the center conductor is

$$I = \int_0^1 \left| \frac{d\xi}{dW} \right| dW = \int_0^1 \frac{dW}{(1 - W^2)(1 - k^2 W^2)} = 4K(k)$$

The power loss on the center conductor is given by

$$\begin{aligned} \int_Z^1 |J_z|^2 dZ &= 2R_m J_c^2 \int_0^1 \frac{dW^2}{(1 - W^2)(1 - k^2 W^2)} dW \\ &= \frac{2R_m J_c^2}{a} \int_0^1 \frac{du}{(1 - u^2)(1 - k^2 u^2)(1 - k^2 u^2)(1 - k^2 u^2)} \end{aligned} \quad (\text{III.44o})$$

while that on the two ground planes is given by

$$\int_{-R/l}^{2/a} \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)(l-k^2u^2)}} = T \quad (111.44)$$

The integrals in (111.43) and (III.44) can be evaluated numerically but iterative technique is necessary in order to find the required values of k_0 and k_x . We will only consider the case when the thickness t is very small; say $t < 0.05a$. For this important special case, a number of approximate can be made that will lead to relatively simple expressions for the pov loss.

When t/a is small k_n is approximately equal to one and k_x approximately equal to k . Thus the intervals u_0 to 1 and l/k to u , i small. We also note that if $t=0$, then $k_0 = 1$, $k_x = k$ and (111.42) gives $Z = AW$ from which we obtain $a = A$ and $b = A/k$ or $k = a/b$. We will ignore these values for A and k for the case of small but finite values of I . (III.436) we can set u equal to 1 in all but the critical factors $1 - u$ a $1 - k_0 u$, since the interval of integration is very small. Thus we find tr.

$$t = aJ \cdot \frac{1}{2(1 - \frac{1}{k_0})} \int_{-1}^{-1} \frac{du}{\sqrt{1 - u^2}}$$

This integral can be evaluated in terms of elementary functions and give;

$$t = \frac{\pi}{4} \left(\frac{k_0}{1 - k_0} \right)^{1/2}$$

Hence we obtain

$$k_0 = 1 + \frac{2}{t} \quad (\text{in.4.1})$$

which verifies that k_0 is close to one.

In (III.43rf) we put $u = l/k$ in all factors except the critical ones $1 - ku$ and $1 - Ajt$, since the interval of integration in this equation is already small. An integral of the same type as found above is obtained and readily leads to the result

which shows that k_x is close to $k = a/b$.

The power-loss terms involve the following four integrals:

$$7_1 = \int_0^l f'''G(u)du \quad /., \quad \sim \int_{u_0}^l G(u)du$$

$$I_3 = \int_{i/k}^l G(u)du \quad 7_4 = \int_{i/k}^l f'''G(u)du$$

where $G(u) = l[(1 - u^2 X_1 - k^2 u H_l - k^2 u^2)(l - k^2 u^2)V^{1/2}]$. The integrals 7_2 and 7_4 will give us the power loss on the conductor edges and will be considered first. For the evaluation of I_2 we use the approximation $[(1 - k^2 u H_l - k^2 u)]^{1/2} = (1 - k^2)^{1/2}$ and make the substitution $k_0 u = A$, $k_a = l/k_0$ to obtain

$$7_2 = M \frac{\int_0^1 h \cdot fitf}{1 - *0} - 1)(1 - k^2 A^2) \quad *0(1 - k^2)$$

Since k_a is close to one, we can use $K'(fe_a) \ll T/2$ and $fe_0 = 1$ to get

$$\mathbf{h} = \frac{d}{2(1 - a^2/6^2)} \quad (\text{III.47a})$$

The same approach is used to evaluate 7_4 to obtain

$$\mathbf{h} \sim \frac{k}{\int_{-a}^a V(A^2 - 1) \{1 - fe_0 A^2\}} \quad \frac{1 - *0^2}{1 - *0^2} \quad (\text{III.476})$$

since $k_h = k_x/k$ is close to one.

In order to evaluate 7_1 , we first make the substitution $k_n u = \cos \theta$ to obtain

$$7_1 = \int_{-\pi/2}^{\pi/2} \frac{\lambda^2 r f f l}{y \{k^2 - \cos^2 \theta\} (\lambda^2 - k^2 \cos^2 \theta) (\lambda^2 - *f \cos^2 \theta)}$$

We now split the integral into one over a small interval 0 to 0, plus an integral from 0, to $\pi/2$. Over the first interval the expression under the square-root sign is approximated by

$$(K + 1)(*o - *i)(*o - *i)(*o - \cos B) \\ = (1 - *0^2)^2 [(fc_o - 1)2 + 0^2]$$

The resultant integral is an elementary one and easily integrated to give

$$\begin{aligned} /_{10} &= \frac{r''}{4} \int \frac{k_0}{i - ik^2} \frac{dO}{y/2(k_O - l) + 6^2} \wedge \frac{k^2}{1 - k^2} \wedge \ln[0] + j2(k_0 - l) + 0^* \\ &\quad 1 - k^2 \wedge \frac{4t/ira}{y/4t/va} \quad 1 - k^2 \quad \text{In } ij4t/TTO \end{aligned}$$

Even if $4t/ira$ is of the same order as 0^2 , the error in the last approximation is small. For $At/a = ft^2$ it is $\ln(1.207) = 0.207$, which is small relative to other terms that occur for P_n .

In the integral from 0_t to $7r/2$, we can assume that $t = 0$ so $k_{l0} = 1$, $k_x = k$, and we then obtain

$$\frac{j}{16} \frac{\text{pr}/8}{l_h} \frac{dO}{\sin 0 (1 - e^2 \cos^2 0)}$$

By using $A = \cos 0$ this integral becomes

$$\frac{r \cos 0}{(1 - A^2)(1 - A^2 A^2)} dK$$

and is readily evaluated. We obtain

$$\frac{r}{2(1 - A^2)} \frac{dO}{8?} \frac{4 - t^2}{A} \frac{2(1 + A) - Atf}{2(1 - k) + eff}$$

upon using $\cos 0 = 1 - 0^2/2$. If we choose $0 = 0.25$, then a number of terms involving 0^2 can be dropped. When we combine I_{la} and I_{lb} , we get

$$I_{.} = \frac{1}{2(1 - A^2)} \frac{Arra}{t} \frac{1 + k}{\sqrt{1 - k}} \quad (111.48a)$$

which has the nice feature that it does not depend on $0.$

For the final integral $/_3$, we make the substitution $k_x u = \cosh 0$ and again split the integral into one over the interval 0 to 0 , plus one over the interval 0 , to 1 . In the first interval $\cosh 0$ is replaced by unity in the noncritical terms and by $1 + 0^2/2$ in the $1 - k^2 u^2 = 1 - i.k^2/k^2 \cosh^2 0$ term. In the second integral we use $k_0 = 1$, $k = k$. We then find that

$$\frac{k}{2(1 - A^2)} \frac{4TT6}{t} \frac{1}{k} \frac{1 + k}{\sqrt{1 - k}} \quad (\text{III.486})$$

The series resistance of the center conductor is $i?$, and is defined by the relationship

$$U^2 R_s = P_n = \frac{2}{-R_m(l_x + I_2)}$$

The series resistance R_2 of the ground planes is defined by a similar

relationship, namely,

$$\frac{W^2 f l_2}{2} = P, \quad 2 = \frac{r_i}{a} \cdot \frac{1}{m} \left(\frac{1}{l_3} + \frac{1}{l_4} \right)$$

The total current I equals $4K(k)$, so we find that

$$\begin{aligned} R_x &= \frac{8aK^2(k)(l - k^2)}{l} \left[\frac{7r + \ln \frac{4ira}{t}}{k} \ln \frac{1+k}{1-k} \right] & (III.49a) \\ R_{xx,k} &= \frac{- + \ln \frac{4-n-b}{l-k}}{\ln \frac{1+k'}{l-k}} & (III.496) \end{aligned}$$

where $k = a/b$. The attenuation is given by

$$\alpha = R_x + R_{xx,k} \quad (111.50)$$

where Z_r is the characteristic impedance of the coplanar line.

When t is very small and the ground planes are widely separated, the expression for R_x reduces to that for an isolated strip, as one would expect. Apart from the factor $1/(1 - k^2)$ the ground-plane losses are approximately what one would find for an isolated conductor of width $2b$. The formulas derived above are estimated to be accurate to within 10 percent for $t < 0.05a$ and $k < 0.8$. More accurate approximations can be made but the resultant formulas would be more complex. The expressions obtained above can also be derived with fewer steps by using Lewin's method. The alternative derivation was chosen so as to provide an example of a more complete derivation and to more fully show the approximations involved.

APPENDIX IV

PHYSICAL CONSTANTS AND OTHER DATA

IV.1 PHYSICAL CONSTANTS

Permittivity of vacuum = $\epsilon_0 = 8.854 \times 10^{-12} = (1/36\pi) \times 10^{-9}$ F/m

Permeability of vacuum = $\mu_0 = 4\pi \times 10^{-7}$ H/m

Impedance of free space = $Z_0 = 376.7 = 120\pi$ ohms

Velocity of light = $c = 2.998 \times 10^8$ m/s

Charge of electron = $e = 1.602 \times 10^{-19}$ C

Mass of electron = $m_e = 9.107 \times 10^{-31}$ kg

,, $=e/m = 1.76 \times 10^{11}$ C/kg

Mass of proton = $M_p = 1.67 \times 10^{-27}$ kg

Boltzmann's constant = $k = 1.380 \times 10^{-23}$ J/K

Planck's constant - $h = 6.547 \times 10^{-34}$ J • s

10^7 ergs = 1 J

1 J = 0.6285×10^{19} eV

1 eV = energy gained by an electron in accelerating through a potential of iV

Energy of 1 eV = equivalent electron temperature of 1.15×10^4 K

Electron plasma frequency $f_p = \frac{e / N}{2\pi \sqrt{m_e J}} = 8.97N^{1/2}$ Hz, where iV is

the number of electrons per cubic meter

Electron cyclotron frequency $f_c = eB/2vm = 28,000 B$ MHz for B in

webers per square meter; $f_c = 2.8 B$ MHz for B in gauss

10^{-8} G = 1 Wb/m²

IV.2 CONDUCTIVITIES OF MATERIALS

Material	Conductivity, S/m	Material	Conductivity, S/m
Copper (annealed)	5.8×10^7	Steel	$0.5-1.0 \times 10^7$
Aluminum	3.54×10^7	Water (distilled)	2×10^{-4}
Silver	6.14×10^7	Sea water	3-5
Nickel	1.28×10^7	Quartz (fused)	$< 2 \times 10^{-17}$

IV.3 DIELECTRIC CONSTANTS OF MATERIALS

Material	Frequency, MHz	*'AO	Loss tangent ϵ''/ϵ'
Polystyrene	3,000	2.54	0.00025-0.0016
Polystyrene (foam)	3,000	1.05	0.00003
Lucite	10,000	2.56	0.005
Teflon	10,000	2.08	0.00037
Fused quartz	10,000	3.78	0.0001
Ruby mica	3,000	5.4	0.0003
Titanium dioxide	10,000	90	0.002
Mahogany wood	10,000	1.7	0.021

IV.4 SKIN DEPTH IN COPPER

Frequency, Hz	in	60	to^a	10^a	10^4	10^8
Skin depths,, cm	2.08	0.85	0.66	0.208	6.6×10^{-6}	$6.6 \times KT^4$

$$S_s = ft/io/ur \quad " \quad 6.6 f^{-1/2} \text{ cm for copper (} a = 5.8 \times 10^7 \text{ S/m).}$$

AUTHOR INDEX

- Afsar, M. N., 130
Alexopolous, N. G., 180
Allison, J., 190
Alseyab, S. A., 647
Altschuler, E. E-, 351
Anderson, T. N., 421
Ayres, W. P., 456
- Baden Fuller, A. J., 219
BabJ, J., 166, 219, 413, 432, 479, 500, 875
Bailey, A. E., 16
Balanis, C. A., 16
Beck, A. H. W., 712
Belevitch, V., 596
Benedek, P., 491
Benson, F. A., 190
Bethe, H. A., 284, 416
Bevensee, R. M., 647
Bhartia, P., 166, 219, 413, 432, 479, 500, 875
Blackwell, L. A., 830
Blight, R. E., 480
Bobroff, D., 668
Bolinder, F., 374, 393
Bosma, H., 475, 480
Bostian, C. W., 16
Bowman, F., 885
Branch, G. M., 660
Brewer, G. R., 712
Brillouin, L., 647, 653
- Brown, J., 346
Burton, M. N., 207, 209
Button, K. J., 130, 480
- Carlin, H. J., 471
Carson, R., 798
Caulton, M., 322
Chang, K., 210, 219, 413-414, 479, 722, 724, 798, 875
Chang, K. K. N., 830
Chew, W. C. 149
Chodorow, M., 672, 712
Chow, K. K., 549
Chu, K. R., 712
Chu, L. J., 668
Clarricools, P., 480
Cohn, S. B., 173, 393, 346, 444, 590, 596, 639, 647
Coleman, J. T., 703
Collin, R. E., 16, 43, 165, 171, 208, 241, 277, 284, 286, 340, 346, 380, 393, 417, 418, 525, 544, 888
Collins, G. B., 712
Comstock, R. L., 480
Copson, E. T., 229
Costanzo, A., 850
Courtoise, L., 549
Cullen, A. L., 830

- Dalman, G. C., 219
 Daly, D. A., 322
 Dicke, R. H., 220, 245, 302, 416, 479,
 Dormann, J. L., 549
 Dow, W. G., 653
 Drobot, A. T., 712
- Edwards, T. C., 219
 Ekholdt, R., 322
 Elliott, R. S., 16
 Emde, F., 895
- Fay, C. E., 480
 Ferguson, P. E., 703
 Fox, A. G., 409
 Fung, A. K., 16
- Gastine, M., 549
 Gentile, C., 798
 Ghione, G., 176
 Gilden, M., 830
 Ginzton, E. L., 16
 Gonzalez, G., 798
 Gopinath, A., 368
 Goubau, G., 549
 Gradshteyn, T. S., 891
 Guillemin, E. A., 229, 590
 Guillou, P., 549
 Gunston, M. A. R., 150
 Gupta, C., 368
- Ha. T. T., 798
 Hahn, W. C., 663, 712
 Hamilton, D. R., 712
 Hammerstad, E. O., 149, 151
 Harrington, R. F., 63, 144
 Hartwig, C. P., 156
 Harvey, A. F., 479, 577
 Haskal, H., 224
 Haus, H. A., 668
 Hayt, W. H., Jr., 70
 Heilmeier, G. H., 830
 Held, D. N., 875
 Helszajn, J., 500
 Hensperger, E. S., 346
 Heuer, H. J., 818
 Hirshfield, J. L., 712
 Honey, R. C., 830
 Hopfer, S., 207, 209
 Horner, J. B., 712
 Homo, M., 430
 Howe, H., 130, 173
 Hutter, R. G. E., 576, 712
- Ishii, T. K., 219
 Ivanek, F., 16
- Jackson, R. W., 179
 Jahnke, E., 895
 James, D. S., 500
 Johnk, C. T. A., 70
 Johnson, R. C., 393
 Jones, E. M. T., 364, 434, 830
- Kajfez, D., 549
 Kales, M. L., 465
 Kaul, R., 219
 Kerns, D. M., 302
 Kerr, A. R., 875
 Kleen, W. J., 712
 Klopenstein, R. W., 393
 Kluyver, J. W., 668
 Knight, S. P., 322
 Knipp, J. K., 712
 Knoppik, N., 497, 499
 Kobayashi, M., 162
 Kollberg, E. L., 875
 Komatsu, Y., 515
 Kong, J. A., 70, 149
 Kotzebue, K. L., 830
 Kraus, J. D., 16, 70
 Krauss, H. L., 712
 Kuhn, N., 264
 Kurokawa, K., 268, 549
- Lange, J., 434
 Laverghetta, T., 130
 Lax, B., 480
 Lewin, L., 340, 342, 902
 Lewis, J., 712
 Li, Q. F., 712
 Liboff, R. L., 219
 Liechti, C. H., 717
- Makimoti, M., 647
 Malherbe, J. A. G., 647
 Manley, J. M., 804
 Marcuvitz, N., 207, 302, 339-340, 479, 552
 Mason, S. J., 261
 Mass, S. A., 875
 Massee, D. J., 156
 Matsumaru, K., 393
 Matthaei, G. L., 364, 434, 630, 830
 Matthews, H., 647
 McDonald, N. A., 286
 McLachlan, N. W., 885
 Medina, F., 430

- Meixner, J., 43
 Melchor, J. L., 456
 Mihran, T. G., 660
 Mobbs, C. I., 647
 Montgomery, C. G., 16, 220, 245, 302, 416,
 479, 548
 Moore, R. K., 16
 Morich, M., 166
 Moynihan, R. L., 449
 Mumford, W. W., 636
 Murakami, Y., 515
- Nakatani, A., 180
 Naldi, C. U., 176
 Neri, A., 850
- Okress, E. C., 16
 Ordung, P. F., 712
- Pannenborg, A. E., 302
Parad, L. I., 449
 Park, S. Y., 712
 Pavio, A- M., 719, 798, 875
 Penfield, P., 830
 Pengelley, R. S., 722, 798
 Pierce, J. R., 650, 653, 712
 Pierpont, J., 229
 Poh, S. Y., 149
 Pozar, D. M.. 219, 479
 Pratt, T., 16
 Presser, A., 434
 Pucel, R. A., 156
 Purcell, E. M., 220, 245, 302. 416, 479. 548
- Rafuse, R. P., 830
 Ragan, G. L., 397, 479, 548, 592
 Ramo, S., 219, 712
 Read, M. E., 712
 Reich, H. J.. 712
 Rhode, U. L., 719, 798, 875
 Rhodes, J. D., 647
 Riblet, H. J., 393
 Rigrod, W., 712
 Rizzi, P. A., 219
 Rizzoli, V.. 850
 Roberts, J., 480
 Rodrique, G. P., 476
 Rosenbaum, F. J., 475, 480
 Rowe, H. E., 804
 Ryzhik, T. M., 891
- Schelkunoff, S. A., 218
 Schloemann, E., 480
- Schneider, M. V., 150
 Schwarz, S. E., 70
 Sensiper, S., 585
 Shen, L. C., 70
 Sheng, N. H., 717
 Silverstein, J. D., 712
 Sylvester, P., 491
 Skalnik, J. K., 712
 Skolnik, M. I., 16
 Slater, J. C., 548. 647, 712
 Solymar, L., 393
Soohoo. R. F., 458, 480
 Spangenberg, K. R., 650, 672, 712
 Sprangle, P., 712
 Stratton, J. A., 70
 Stutzman, W. L., 16
 Suskind, C, 672, 712
 Symons, R. S., 703
- Thiele, G. A., 16
 Troetsehe), W. O., 318
- Uenohara. M., 823
 Ulaby, F. T., 16
- Valier, G., 703
 Van Bladel, J., 549
 Van der Zeil, A., 799
 Van Duzer, T., 219
 Van Trier, A. A. Th. M.. 465
 Vartanian, P. H.. 456
 Vendelin, G. D., 719, 798, 875
 Villeneuve, A. T.. 63
- Wait, J. R., 70
 Watkins, D. A., 585, 647
 Watson, G. N.. 885
 Weale, J. R., 150
 Weinberg, L., 596
 Wenzel, R. J.. 647
 Wheeler, H. A., 149
 Whinnery, J. R., 219
 Wilkinson, E., 443
 Williams, A. E., 647
 Wolff, E., 219
 Wolff, I., 497, 499
 Wu, Y. S., 475, 480
- Yamashita, S., 647
 Young, L., 346, 360, 364, 393, 434
- Zitelli, L., 712

SUBJECT INDEX

- Admittance
characteristic, of transmission line, 76
electronic, in klystron, 688
input, for transmission line, 93
inverters, in niters, 603-614
- Amplification
of klystron, 685
of parametric amplifier, 813, 815-820
of traveling-wave tube, 698
- Amplifier,
design of, 755-759, 780-795
double stage, 788-793
low noise, 773-776, 787
single stage, 781-788
gain of, 274, 728-735
stability of, 736-744
- Angular momentum, 451
- Anisotropic media, 26-28
- Antenna, pi-obe in waveguide, 276-281
- Aperture
coupling by, in waveguide, 284-294
polarizability of circular, 285
in rectangular cavity, 517-523
- Attenuation
for circular waveguide, 196-197
for coaxial transmission line, 111, 117
for coplanar line, 178-180
for microstrip line, 153-157, 163-164
for rectangular waveguide, 188-189
for strip line, 171-173
for transmission line, 108-111
- Attenuator
electronic, 400-404
rotary, 397-400
- Babinet's principle, 580
- Backward-wave oscillator, 709
- Bandwidth
of matching network, 325-330
of resonant circuit, 483
- Beam, electron (see Electron beam)
- Beam coupling parameter, 672
- Bessei functions, 195, 581-583, 881-885
spherical, 510-511
- Bethe directional coupler, 416-419
- Bilinear transformation, 716, 725-726
- Binomial quarter-wave transformer, 350-352
- Bloch wave, 556
impedance of, 556-557
(*see also* Periodic structures)
- Boundary conditions
at conducting edge, 43-44
at conducting surface, 41-43
for electromagnetic field, 39-44
at infinity, 44

- Branch line directional coupler, 432-434
 Brillouin flow, for electron beam, 653, 701
- Capacitance
 distributed
 of coaxial line, 115
 of coplanar line, 176
 of microstrip line, 147-151
 of strip line, 896-898
 of transmission line, 72-73
 of microstrip gap, 493
 of microstrip open end, 492
 of microstrip step, 368
- Capacitors, for microstrip circuits, 322
- Carcinotrons, 709
- Cavity
 coupling parameter for, 496, 521-523
 cylindrical, 504-508
 mode chart for, 507
 Q of, 507
 resonant frequency of, 506
 degenerate modes in, 536-538
 excitation of, 538-541, 683-686
 field expansion in, 525-533
 filter, 635-641
 loop-coupled, 523-525
 oscillations in, 533-536
 perturbation of, 541-545
 rectangular, 500-504
 aperture coupled, 517-523
 Q of, 503-504
 resonant frequency of, 501-502
 Chebyshev filters, 593-598
 Chebyshev polynomials, 353, 355
 Chebyshev quarter-wave transformer,
 352-360
 Chebyshev tapered transmission line.
 380-383
 Choke joint, 397
 in variable short circuits, 395-397
 Circles, constant
 gain, 744-755
 mismatch, 776-780
 noise figure, 772-776
 stability, 736-744
 load, 736
 source, 739
 Circular polarized field, 405-407, 452
 Circulator, four-port, 468-471
 for parametric amplifier, 816-817
 scattering matrix for, 471-472
 three-port, 471-476
 Coaxial transmission line
 attenuation in, 111, 117
 characteristic impedance of, 115
 distributed parameters for, 115-116
 fields in, 106-108
 Conformal mapping, 886-889
 and conductor loss, 898-910
 for coplanar line, 905-910
 for microstrip line, 903-905
 for coplanar line, 892
 for microstrip line, 892-896
 for slot line, 892
 for strip line, 896-898
 Constitutive relations, 23-28
 Continuity equation for current, 20
 Coplanar transmission line, 126-127,
 175-180
 attenuation in, 178-180
 impedance of, 176-178
 Corrugated plane as periodic structure,
 571-577
 Coupled microstrip line, 126-127
 for directional coupler, 427-432
 Coupled strip line, 173-174
 Coupling
 in directional coupler, 414
 of modes in lossy cavity, 536-538
 Coupling coefficient, for coupled microstrip
 line, 166
 Coupling parameter, for cavity, 496, 521-523
 Current, equivalent, in waveguide, 221-223
 linear, excitation of waveguide by, 281-283
 loop, in waveguide, 283-284
 normalized, 223
 on transmission line, 106
 Cutoff frequency (see Waveguide, circular;
 Waveguide, rectangular)
 Cyclotron frequency, 701, 704
- Damping of cavity, 484
- Delta function, 59-60
- Diaphragm
 capacitive, in rectangular guide, 341-342
 inductive, in rectangular guide, 340-341
- Dielectric constant, 25
- Dielectric resonator, 508-517
 cylindrical, 515-516
 hemispherical, 509-515
 Q of, 513
- Directional coupler
 Bethe type, 416-419
 branch line, 432-434
 Chebyshev, 422-427
 coupled line, 427-432
 coupling in, 414
 directivity of, 414
 Lange, 434-435

- multielement, 422-427
- scattering matrix for, 414-416
- Schwinger reverse phase, 420
- two-hole, 419-420
 - Moreno crossed guide, 421
 - Riblet T-slot, 421
- Disk resonator, 496-500
- Dispersion
 - in microstrip line, 158-163
 - of signal in waveguide, 198-204
- Double-stream amplifier, 708
- Double-stub tuner, 312-317
 - for waveguide, 342-343

- E* mode, 102-104
 - in circular guide, 194-196
 - in rectangular guide, 193
- E-H* tuner, 342-343
- Electron beam
 - ac power relations for, 667-670
 - with axially confined flow, 651
 - beam coupling parameter for, 672
 - Brillouin flow for, 653, 701
 - dc conditions for, 650
 - ion-neutralized, 650-651
 - kinetic-power theorem for, 670
 - perveance of, 650
 - space-charge waves on, 654-667
 - velocity modulation of, 670-678
 - (see also Space-charge waves)
 - Electron precession in ferrite, 451-460
 - Electronic admittance of reflex klystron, 688
 - Elliptic sine function, 889-891
 - Energy
 - electric, 34-36
 - magnetic, 34-36
 - velocity of, in free space, 48
 - in periodic structures, 566-571
 - in waveguides, 204-205
 - Excitation
 - of cavity, 538-541, 683-686
 - of waveguides, 281-294
 - Exponential taper for transmission line, 372

 - Faraday rotation in ferrites, 460-464
 - Faraday's law, 18
 - Ferrite
 - electron precession in, 451-460
 - Faraday rotation in, 460-464
 - magnetic permeability of, 455, 457-459
 - in microwave devices, 464-476
 - plane-wave propagation in, 459-460
 - Filling factor, 155

 - Filters
 - cavity
 - direct-coupled, 639-642
 - quarter-wave-coupled, 635-639
 - frequency transformations in, expansion, 599
 - low-pass to bandpass, 600-602
 - low-pass to high-pass, 599-600
 - periodic, 602-603
 - half-wave, 360-370, 617-626
 - image-parameter design of, 587-590
 - impedance inverters in, 603-615
 - insertion-loss design of, 591-592
 - low-pass designs for, 595-598
 - parallel coupled, 626-635
 - power loss ratio in, 592-594
 - for Chebyshev, 593
 - for maximally flat, 593
 - Fin line, 208-210
 - Floquet's theorem, 569-571
 - Foster's reactance theorem, 230-232
 - Frequency bands, 2-3
 - Fresnel reflection coefficient, 51-52
 - Fresnel transmission coefficient, 51-52

 - Gain
 - definitions of
 - available, 274, 728
 - maximum, 274, 728
 - power, 274, 728-735
 - transducer, 273-274, 728
 - of klystron, 685
 - of parametric amplifier, 813, 815-820
 - of traveling-wave tube, 698
 - Gauss' law, 19
 - Group velocity
 - in periodic structures, 566-571
 - in waveguide, 204-205
 - Gunn oscillator, 832-837
 - Gyrator, 464-465
 - Gyrotron, 701-708

 - H* modes, 98, 100-102
 - in circular guide, 196-198
 - in rectangular guide, 182-192
 - Half-wave filter, 360-370, 617-626
 - Half-wave plate, 405
 - Hankel functions, 881-885
 - spherical, 510-511
 - Helix
 - general properties of, 583-585
 - sheath, 580-583
 - dispersion equation for, 583
 - in traveling-wave tube, 693

- Helmholtz's equation, 32, 97
- Helmholtz's theorem, 19, 525
- HEMT transistor, 722
- Hybrid junction
 - as balanced mixer, 865-866
 - branch line coupler as, 432-434
 - magic T as, 435-437
 - ring circuit as, 437-442
 - scattering matrix for, 436-437, 441
- Image parameters of filters, 587-590
- IMPATT oscillator, 837-840
- Impedance
 - characteristic
 - of capacitively loaded transmission line, 556
 - of coaxial line, 115
 - of coplanar line, 176-178
 - of microstrip line, 150-153
 - of strip line, 171
 - of transmission line, 76
 - general definition of, 38
 - input, even and odd properties of, 232-233
 - input, on transmission line, 93
 - matching, with lumped elements, 319-330
 - with stubs, 309-319
 - (*see also* Quarter-wave transformers;
 - Transmission line, tapered)
 - matrbc
 - imaginary property of, 236-237
 - symmetry of, 235-236
 - normalized, 90, 237-238
 - surface, 56
 - wave
 - for circular guide, 196-197
 - for TE waves in rectangular guide, 185, 190
 - for TM waves in rectangular guide, 189
 - of waveguide elements, 339-342
- Impedance inverters in filters, 603-615
- Impedance mismatch factor, 334
 - invariance of, 334-339
- Impedance termination, design of, 330-334
- Inductance, distributed, for transmission line, 72-73
- Inductor, for microstrip circuits, 320-322
- Insertion loss in filters, 591-592
- Interdigital line, 577-579
- Isolator, 466-468
- Johnson noise, 762
- k_0-p diagram, 564-566
- Kinetic power theorem for electron beam, 670
- Kinetic voltage, 670
- Klystron
 - reflex
 - electronic admittance in, 688
 - oscillation conditions for, 688
 - tuning curves for, 688-689
 - two-cavity, 678-686
 - equivalent circuit for, 684-685
 - excitation of fields in, 683-686
 - gain of, 685
- Laplace's equation, 29
- Larmor frequency, 452
- Lorentz condition, 57, 133
- Lorentz force, 17-18
- Lorentz reciprocity theorem, 62-64
- Loss tangent, 26
- Magic T, 435-437, 865-868
- Magnetic permeability, 18, 27
 - for ferrite, 455, 457-459
- Magnetic susceptibility, 27
- Magnetron, 690-692
- Manley-Rowe relations, 804-807
- Matching network
 - design of
 - for amplifier, 330-334, 338-339
 - lumped element, 319-330
 - Q of, 325-330
 - with transmission line stubs, 309-319
- Maxwell's equations, 21
- Meander line, 577-579
- MESFET, 721
- MIC circuit, 714
- Microstrip line, 125-128, 130-169
 - attenuation of, 153-157, 163-164
 - coupled, 164-170
 - dispersion in, 158-163
 - effective dielectric constant for, 149-152
 - impedance of, 150-153
 - inverted-suspended, 126-127
- Microstrip resonator, 490-496
 - disk, 496-500
 - Q of, 499
- Mixer, 856-868
 - balanced, 865-868
 - compression in, 862-863
 - harmonic balance method for, 869-873
 - image-enhanced, 868
 - image-rejection, 868
 - intermodulation in, 863-864
 - noise figure, 864-865
 - subharmonic, 868

- MMIC circuits, 714
- Mode chart for cylindrical cavity, 507

- Negative-resistance amplifier, 814-821
- Noise, conductance
 - equivalent, 767
 - equivalent temperature of, 762
 - figure, 768-773
 - circles for, 772-776
 - of cascaded stages, 770-772
 - of mixer, 864-865
 - optimum source impedance for minimum, 769-770
 - of parametric amplifier, 821-829
- Johnson or Nyquist, 762
- temperature
 - of amplifier, 771
 - of system, 771-772
- resistance, equivalent, 767
- theory of, 760-765
 - in two-ports, 766-767
- Normalized current, 223
- Normalized load impedance, 90
- Normalized voltage, 223
- TV-port circuits, 233-235

- Oscillators, design of, 851-856
 - Gunn, 832-837
 - IMPATT diode, 837-840
 - three-port scattering matrix for, 843-849
 - transistor, 840-856
 - O-type traveling-wave tube, 692-699

- Parallel plate transmission line, 117-125
- Parametric amplifier
 - linearized equations for, 807-809
 - Manley-Rowe relations for, 804-807
 - negative resistance, 814-821
 - gain of, 815-820
 - gain-bandwidth product for, 821
 - noise in, 823-825
 - noise figure, of degenerate negative resistance, 825-829
 - of negative resistance, 823-825
 - of up-converter, 821-823
 - p-n* junction diodes for, 800-802
 - up-converter, 809-814
 - gain of, 813
- Periodic structures
 - Bloch-wave impedance for, 555-556
 - Bloch waves in, 556

- energy flow velocity in, 566-571
- and filters, 587-590
- Floquet's theorem for, 569-571
- group velocity in, 566-571
- $k_0 p$ diagram for, 564-566
- matching of, 563-564
- spatial harmonics in, 569-571
- terminated, 560-563
- for traveling-wave tube, corrugated plane, 571-577
 - helix, general properties of, 583-585
 - interdigital line, 577-579
 - meander line, 577-579
 - sheath helix, 580-583
 - tape ladder line, 577-578
- unsymmetrical two-ports in, 559-560
- Permeability, 18, 27
 - for ferrite, 455, 457-459
- Perveance of electron beam, 650
- Phase shifter, electronic, 409-413
 - rotary, 404-409
- Phase velocity, 47, 198-199
 - in waveguides, 182
- Physical constants, 911-912
- PIN diode, 401-403
- Plane waves, 44-48
- Plasma frequency, 653
 - effective, 659
- Poisson's equation, 29
- Polarization
 - circular, 405-407, 452
 - of circular aperture, 285
 - in dielectric, 23-27
- p-n* junction diode, 800-802
- Post
 - capacitive, in waveguide, 342
 - inductive, in waveguide, 341
- Potential
 - scalar, dynamic, 57
 - static, 28
 - vector, dynamic, 57
 - static, 30
- Power, in circular guide, 197
 - for TE waves in rectangular guide, 186-187
- Power added efficiency, 842
- Power divider, 442-450
 - Wilkinson, 443-450
- Power gain, 274, 728-735
- Power loss ratio
 - in filter, 591-594
 - in quarter-wave transformer, 356-357
- Power orthogonality, in waveguides, 186
- Power waves, scattering matrix for, 268-276
- Poynting vector, 38-39
 - complex, 37

- Probe, radiation resistance of. in waveguide, 281
- Pulse propagation, on transmission line, 78-85
- Quality factor or Q , 325. 503-504
of cylindrical cavity, 507
of dielectric resonator, 513
of disk resonator, 499
external, 483
loaded, 483
of matching network, 325-330
of rectangular cavity. 503-504
unloaded, 483
- Quarter-wave plate, 405
- Quarter-wave transformers
Chebyshev, exact results. 356-360
three-section, 359-360
two-section, 356-358
N-section, approximate theory for. 348-350
binomial, 350-352
Chebyshev, 352-356
prototype circuit for filter, 360-370
single-section, 343-346
- Reactive elements in waveguide. 339-343
shunt capacitive, 341 -342
shunt inductive, 340-341
stubs as, 342-343
- Reciprocity theorem, 62-64
- Reflection
from conducting plane. 53-56
from dielectric surface
parallel polarization, 49-52
perpendicular polarization, 52-53
small, theory of. 348-350
- Reflection coefficient, current, 91
for tapered transmission line, 371
and Riccati equation. 383-386
for terminated transmission line. 90-91
voltage, 90
- Reflex klystron. 686-689
- Resistance, radiation
of probe in waveguide, 281
of transmission line. 114
- Resistance-wall amplifier, 708
- Resonant circuits
bandwidth of. 482-483
damping of, 484
 Q of. 482-484
transmission line
antiresonant. 488-490
open circuited, 487-488
short-circuited. 485-487
- Return loss, 329
- Riccati equation for tapered transmission line 383-386
- Ridge waveguide, 205-207
- Ring circuit, 437-442
- Scalar potential
dynamic, 56-59
static, 28
- Scattering matrix
of circulator, 471-476
of directional coupler, 414-416
of hybrid junction, 436-437, 441
for lossless junction, 251-253
for power waves, 268-276
symmetry of, 250-251
and transformation of terminal planes, 249-250
for transistor, 843-849
for two-port junction, 254-257
unitary property of, 253
- Schwinger directional coupler, 420
- Separation constant, 45
- Separation of variables method, 44, 183
- Sheath helix, 580-583
in traveling-wave tube, 693
- Short circuit
choke-type, 397
variable, in waveguide, 395-397
- Signal flow graphs, 260-268
- Signal velocity, 200-204
- Skin depth, 54
- Slot line, 127
- Smith chart. 304-308
- Snell's law. 50
- Space-charge reduction factor, 659-660
- Space-charge waves
ac power relations for, 667-670
and kinetic-power theorem, 670
and kinetic voltage. 670
on axially confined beam, 654-661
dc propagation constant for. 656
effective plasma frequency for, 659-660
fast and slow, 658
reduction factor for. 659-660
on unfocused beam, 661-667
- Spatial harmonics in periodic structures, 569-571
- Stability, of amplifier, 735-744
- Stability circles. 736-744
load. 736
source, 739
- Standing wave ratio, 92
- Standing waves, on transmission line, 91-92

- Static fields, 28-30
- Strip line, i TO-174
 - attenuation on, 171-173
 - coupled, 173-174
 - impedance of, 171
- Stub
 - matching with, 309-319
 - double, 312-317
 - single, 309-312
 - triple, 317-319
 - in waveguide, 342-343
- Substrate, properties of, 130
- Surface impedance, 56
- Surface wave, 124
- Susceptibility
 - electric, 25
 - magnetic, 27
- TE waves, 98, 100-102
- TEM waves, 98-100
- Termination, waveguide, 394-397
- TM waves, 98, 102-104
- Transducer gain, 273-274, 728
- Transmission coefficient, 51
- Transmission line
 - capacitively loaded, 551-557
 - Bloch waves in, 556
 - characteristic impedance of, 556
 - circuit analysis of, 551-557
 - eigenvalue equation for, 554
 - $k_r p$ diagram for, 564-566
 - wave analysis of, 557-559
- distributed circuit analysis of, 86-89
- field theory of coaxial line, 106-108, 111
 - lossless line, 106-108
 - lossy coaxial line, 111
 - with small loss, 111
- parallel-plate, with dielectric, 117-125
- parameters of capacitance, 112, 117
 - characteristic impedance, 113, 117
 - coaxial line, 115, 117
 - conductance, 115, 117
 - inductance, 115, 117
 - resistance, 116, 117
- resonant circuit, 485-490
 - antiresonant, 488-490
 - open-circuited, 487-488
 - short-circuited, 485-487
- tapered, Chebyshev, 380-383
 - exponential, 372
 - reflection coefficient on, approximate equation, 371
- reflection coefficient on, Riccati equation for, 385
- synthesis of, 373-380
 - triangular, 372-373
- terminated, 89-96
- Transmission matrix, for cascade network
 - voltage-current, 257-259
 - wave-amplitude, 259-260
- Transverse resonance method, 206-208
- Traveling-wave tube
 - M-type, 699-701
 - O-type, 692-699
 - gain of, 698
 - periodic structures for, 571-585
- Two-port junctions, 238-248
 - equivalent circuits for, 245-248
- Vector formulas, 876-880
- Vector potential
 - dynamic, 56-59
 - solution for, 59-62
 - static, 30
- Velocity
 - energy flow
 - in periodic structures, 566-571
 - for plane waves, 48
 - in waveguides, 204-205
 - group
 - in periodic structures, 566-571
 - in waveguides, 200-204
 - phase
 - for plane waves, 47-48
 - in waveguides, 182
 - signal, in waveguides, 199-204
 - wavefront, in waveguides, 199
- Velocity-jump amplifier, 708
- Velocity modulation, of electron beam, 670-678
 - beam coupling parameter in, 672
- Voltage, equivalent, in waveguides, 221-224
 - normalized, 223
- Voltage standing wave ratio, 91-93
- Wave
 - classification of, 96-99
 - impedance
 - of TE mode, 185, 190
 - of TM mode, 189
 - plane, 44-48
 - reflection of, from conducting plane, 53-56
 - reflection of, from dielectric surface, 49-53
 - TE, 98-102
 - TEM, 98-100
 - TM, 98, 102-104

Wave (Cont'd)

transmission matrix, 259-260

(see also Periodic structures; Space-charge waves; Transmission line; Waveguide I)

Wave equation, 31

Wave number, 32

Waveguide

capacitive diaphragm in, 341-342

capacitive post in, 342

capacitive rod in, 342

circular, attenuation in, 196-197

solutions for, 194-197

TE waves in, 196-197

TM waves in, 194-196

equivalent current and voltage for, 221-224

excitation of, by aperture, 284-294

by current loop, 283-284

by linear current element, 281-283

inductive diaphragm in, 340-341

inductive post in, 341

properties of, 180-182

rectangular, attenuation in, 188

cutoff frequency of, 184

dominant TE mode in, 190-194

power in, 186-187

solutions for, 189

TE waves in, 182-190

TM waves in, 193

wave impedance for, 185, 189, 197

ridge, 205-207

termination, 394-397

velocity in energy, 204-205

group, 200-204

phase, 182

signal, 199-204

wavefront, 199