FREE

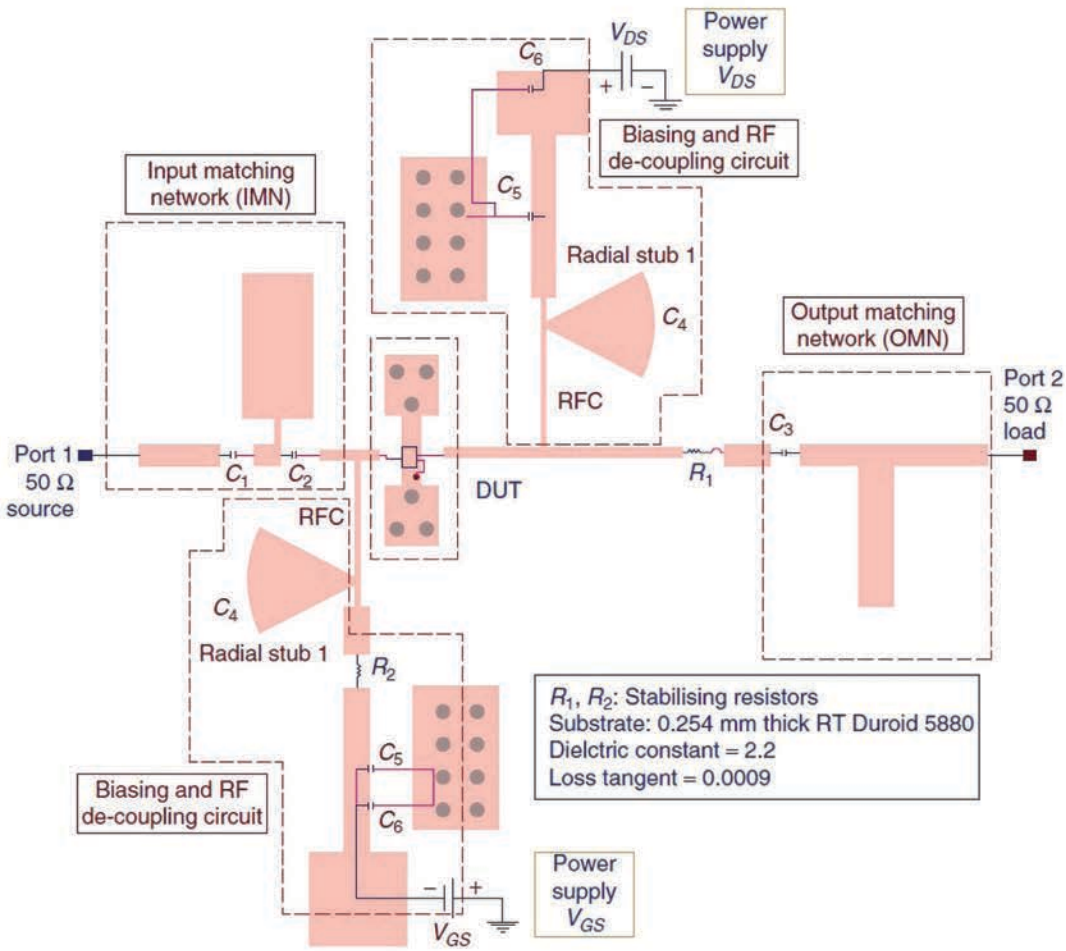
INSTRUCTOR & STUDENT RESOURCES

CRC Press
Taylor & Francis Group

RF AND MICROWAVE DESIGN

CIRCUITS, DEVICES, AND SYSTEMS

SUBAL KAR



RF and Microwave Design

This book deals with the design of circuits, devices, and systems at radio frequency (RF) and microwave frequencies using both the analytical technique and the electronic design automation (EDA) software tools. EDA-based design technique is now used extensively and helps to translate the optimally designed structures directly to actual fabrication.

***Salient features of the book*:**

- Fundamentals of RF and microwave devices and circuits are covered in a concise manner to help students and design engineers refresh and recapitulate some of the basic concepts that are helpful from a designer's point of view.
- Design of RF matching circuit with stub matching technique is provided with an analytical design approach supplemented by design using the Smith chart.
- RF/microwave filter basics are discussed along with the analytical design technique and also with the EDA simulation tool-based design of low-pass, band-pass, and band-stop filters. The design of recently emerging metamaterial-inspired filter is included.
- Basic concepts of RF/microwave oscillator and amplifier are discussed together with the design of IMPact Avalanche Transit Time (IMPATT) oscillator and varactor-tuned voltage controlled oscillator (VCO), low-noise amplifier and power amplifier design using EDA software tool.
- Microwave power combiner and microwave mixer design are given using EDA simulation tool. Discusses RF/microwave antennas with design examples of conventional and metamaterial-inspired microstrip patch antenna.
- Covers RF/microwave system design concepts that are required by students and design engineers to understand system performance, along with some rudimentary design of RF/microwave communication and radar system using EDA simulation tools.

It is primarily written for senior undergraduates, graduate students, and academic researchers and professionals engaged in electrical engineering, electronics and communication engineering, RF and microwave design, nanoelectronics, and physics studies.

RF and Microwave Design

Circuits, Devices, and Systems

Subal Kar

Former Professor and Head
Institute of Radio Physics and Electronics
University of Calcutta
Kolkata, India



CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Designed cover image: © Subal Kar

First edition published 2027
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

© 2027 Subal Kar

Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

For Product Safety Concerns and Information please contact our EU representative GPSR@taylorandfrancis.com. Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany.

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

ISBN: 9781041219798 (hbk)
ISBN: 9781041219804 (pbk)
ISBN: 9781003729969 (ebk)

DOI: 10.1201/ 9781003729969

Typeset in Sabon
by Newgen Publishing UK

Access the [Instructor and Student Resources/Support Material]: www.routledge.com/9781041219798

Dedication

Dedicated to My
Father: Late Sudhirendra Mohan Kar (1959)
and
Mother: Late Amiya Kar (2016)



Sudhirendra Mohan Kar
January/ 1958.

(Sudhirendra Mohan Kar)



Amiya Kar
27.8.63

(Amiya Kar)



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contents

Foreword *xiii*
JOHN CORLETT
Preface *xv*
About the Author *xix*

1 Fundamentals of RF and Microwave Device/Circuit Principles **1**
1.1 Introduction 1
1.2 Low-Frequency vs RF and Microwave Device and Circuit Principles 5
1.2.1 Introduction 5
1.2.2 Device Principles at RF/Microwave Frequencies 7
1.2.3 Circuit Principles at RF/Microwave Frequencies 10

2 Basics of Scattering Matrix Useful for RF and Microwave Device/Circuit Characterization **18**
2.1 Introduction 18
2.2 S-Parameter Formulation of Microwave Network 19
2.3 Properties of S-Parameters 23
2.4 S-Matrix of Various Microwave Passive Components 26
2.4.1 Directional Coupler 26
2.4.2 Magic Tee *29*
2.4.3 Circulator 32
2.5 Signal Flow Graphs 34
2.5.1 Decomposition Rules 35
2.5.2 Mason's Rule 35
2.6 Application of Signal Flow Graph—RF/Microwave Amplifier Gain Formulation in Terms of S-Parameters 37

3 Fundamentals of Smith Chart and Stub Matching Design with Smith Chart 43
3.1 Introduction 43
3.2 Development of Smith Chart 43
3.3 Smith and Jones Chart 48
3.4 Compressed Smith Chart 51
3.5 Stub Matching Design with Smith Chart 53
3.5.1 Introduction 53
3.5.2 Basics of Stub Matching and Analytical Equations for Design of Such Matching Circuits 54
3.5.2.1 Introduction 54
3.5.2.2 Single-Stub Matching 54
3.5.2.3 Double-Stub Matching 57
3.5.3 Design Examples of Stub Matching Using Smith Chart 58

4 Impedance Matching in Microwave Design 67
4.1 Introduction 67
4.2 Impedance Transformers 67
4.2.1 Theory of Small Reflections 69
4.2.2 Stepped Impedance Binomial and Tchebyscheff Transformer 71
4.2.2.1 Binomial Stepped Impedance Transformer 72
4.2.2.2 Tchebyscheff Stepped Impedance Transformer 80
4.2.3 Tapered Impedance Matching Transformers 88
4.2.3.1 Exponential Taper 89
4.2.3.2 Taper with Triangular Distribution 89
4.2.3.3 Klopfenstein Taper 91
4.3 Impedance Matching with Lumped-Component/Planer Distributed-Line Reactive Elements 97

5 Microwave Filter Design: Low-Pass, Band-Pass, Band-Stop, and Metamaterial-Inspired Filter 103
5.1 Introduction 103
5.2 Insertion Loss Method for Low-Pass Prototype Design 105
5.2.1 Introduction 105
5.2.2 Low-pass Prototype Filter Design Technique 110
5.3 Filter Transformations from Prototype 114
5.3.1 Introduction 114
5.3.2 Prototype to Actual Low-pass Transformation 115

5.3.3 *Low-pass Prototype to High-pass Transformation 115*
5.3.4 *Low-pass to Band-pass and Band-stop Transformations 117*
5.4 *Microwave Implementation 119*
5.4.1 *Introduction 119*
5.4.2 *Richard's Transformation 119*
5.4.3 *Kuroda's Identities 122*
5.4.4 *Impedance and Admittance Inverters 126*
5.5 *Practical Microwave Filters—Analysis and Design 129*
5.5.1 *Introduction 129*
5.5.2 *Stepped-Impedance (High-Low) Low-pass Filter 129*
5.5.3 *Coupled-Line Band-pass Filter 139*
5.5.4 *Coupled-Resonator Band-stop Filter 145*
5.6 *Metamaterial-Inspired Filter Design at Microwave Frequency 157*
5.6.1 *Introduction 157*
5.6.2 *SRR-Loaded Waveguide Filter 157*
5.6.3 *CRLH-Type Band-Pass Filter 158*
5.6.3.1 *Design of Unit Cell of CRLH 158*
5.6.3.2 *Microwave Filter Design with CRLH Transmission Line 164*

6 Design of RF/Microwave Oscillators and Amplifiers 169
6.1 *Introduction 169*
6.1.1 *Oscillator/Amplifier with Two-Terminal Microwave Device (Diode) 169*
6.1.1.1 *Fundamental Theoretical Background of RF/Microwave Diode Oscillator/Amplifier 170*
6.1.1.2 *IMPATT Oscillator and VCO Design with 3D Electromagnetic Field Simulator 174*
6.1.2 *Amplifier with Three-Terminal Microwave Device (Transistor) 182*
6.1.2.1 *Fundamental Theoretical Background of RF/Microwave Transistor-based RF/Microwave Amplifier 182*
6.1.3 *Low-Noise Amplifier (LNA) vs Power Amplifier (PA) 195*
6.1.4 *Design of LNA and PA Using EDA Simulation Tool 196*

7 Design of Microwave Power Combiners and Mixers 222
7.1 *Microwave Power Combiners 222*
7.1.1 *Introduction 222*

7.1.2 Different Types of Microwave Power Combiners 223
7.1.2.1 N-Way in Single-Stage/Multiple-stage Power Combiner 223
7.1.2.2 Wilkinson Power Combiner/Divider 225
7.1.2.3 Spatial Power Combining 227
7.1.3 Design of Microwave Power Combiners 228
7.2 Microwave Mixers 246
7.2.1 Introduction 246
7.2.2 Different Types of Microwave Mixers 249
7.2.2.1 Single-Device/Single-Ended Diode Mixer 249
7.2.2.2 Single-Balanced and Double-Balanced Mixers 249
7.2.2.3 Sub-harmonically Pumped Mixer 252
7.2.2.4 Image-Rejection Mixer 253
7.3 Mixer Design Using EDA Simulation Tool 255

8 Microwave Antenna Design 262
8.1 Introduction 262
8.2 Principle of Operation of Microstrip Patch Antenna 264
8.3 Design of Conventional Microstrip Patch Antenna and Metamaterial-Inspired Microstrip Patch Antenna 267
8.3.1 Conventional Microstrip Patch Antenna Design 267
8.3.2 Metamaterial-Inspired Microstrip Patch Antenna Design 274
8.3.2.1 Introduction 274
8.3.2.2 CSRR-Loaded Microstrip Patch Antenna Design 276

9 RF and Microwave System Design 285
9.1 Introduction 285
9.2 Some Technical Aspects Related to RF/Microwave Transmitter and Receiver 287
9.3 Issues Related to RF and Microwave Design 288
9.3.1 Designer's Considerations for RF/Microwave Transmitter and Receiver Design 288
9.3.1.1 Transmitter Parameters for Consideration 288
9.3.1.2 Receiver Parameters for Consideration 290
9.3.2 Link Budget 291
9.3.3 Noise and Distortion 293
9.3.3.1 Introduction 293
9.3.3.2 Noise 293
9.3.3.3 Nonlinearity and Distortion 297
9.4 Design of RF/Microwave Communication Transmitter and Receiver Including Radar System Using EDA Simulation Tool 304

Appendix I *RF and Microwave Band Designation* *329*
Appendix II *MATLAB Programs* *331*
Appendix III *Oscillator Configurations with IMPATT/Gunn Diodes* *343*
Appendix IV *Short-Slot Hybrid and Branch-Line Quadrature Hybrid* *346*
Appendix V *Computational Electromagnetics (CEM), Numerical Techniques, and Simulators* *348*
Appendix VI *Maxwell's Equations and Electromagnetic Wave Equations—A Brief Review* *355*
Appendix VII *Fundamentals of Spectrum Analyzer and Network Analyzer* *364*
Index *377*



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Foreword

John Corlett



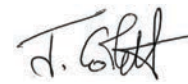
It is my great pleasure to write the foreword for Prof. (Dr.) Subal Kar's book entitled *RF and Microwave Design: Circuits, Devices, and Systems*. I had the pleasure of working with Dr. Kar, from the University of Calcutta, India, during his tenure as Fulbright Scholar (1999–2000) in the Beam Electrodynamics Laboratory of the Center for Beam Physics, Lawrence Berkeley National Laboratory (LBNL), CA, USA, when I was serving as Group Leader for the Beam Electrodynamics Group.

At LBNL Dr. Kar participated in the development of RF structures for a muon collider by assisting in the development and execution of a series of experiments to further our understanding of the performance of thin beryllium (Be) windows as boundaries in high-power microwave cavities. In addition to studies of the Be windows, he contributed to design studies of the high-power RF coupling loop, experimental planning for characterization of cavity thermal properties, measurements, and data analysis. Experimental data thus obtained helped us in the development of cavity designs for application in muon cooling channels required for muon colliders.

More recently, in addition to his teaching duties and research at the University of Calcutta, Dr. Kar has participated as an expert member (since 2007) guiding the scientists of the Society for Applied Microwave Electronics Engineering and Research (SAMEER), Kolkata, India, in the "Establishment of Millimeter-Wave R&D Facility" sanctioned by the Ministry of Communication and Information Technology, Government of India. He also participated as Guide-cum-Consultant (2010–2014) in

the project entitled "Left-Handed Maxwell System Experimental Project" funded by the Board of Research in Nuclear Sciences (BRNS), Government of India, that was implemented at the SAMEER Kolkata Centre.

Dr. Kar's expertise and research experience in RF and microwave engineering, and his teaching at the Institute of Radio Physics and Electronics, University of Calcutta, India, for 33 years (1983–2016), provide an excellent basis for the writing of this book. Targeted to the needs of engineers and scientists wishing to understand the RF/microwave technology, this book will be extremely useful for students, teachers, researchers, and professionals alike who want to learn and work in the design of RF and microwave circuits, devices, and systems. Covering basic principles, techniques, and methods of component and systems characterization; the book has embarked into the design features of critical components and consideration of issues in their integration into systems. I appreciate the effort given by the author to write this book and wish every success for the book.



John Corlett
Lawrence Berkeley National Laboratory (retired)
Berkeley, USA

Preface

Expertise in the design techniques is a must for engineers of all disciplines, with no exception in case of RF and microwave engineers; rather the knowledge of nitty-gritty of RF/microwave design is especially important where the design techniques are quite different from their lower-frequency counterparts. In addition to the simulator-based design, which is the order of the day for all engineering design, side by side the knowledge of long-used analytical design technique in each and every case should not be underestimated as the analytical approach gives the insight of the design technique starting from the first principle of its physical understanding. A mastery of both the simulator-based design and analytical design technique will make a design engineer a truly apt designer from an engineering point of view. A design engineer must also have a clear understanding of the fundamentals of the subject added to the skill earned in simulator-based design and analytical design technique, and this is true for all branches of engineering, including RF and microwaves.

With my long association with the scientists of the Society for Applied Microwave Electronic Engineering and Research (SAMEER) Kolkata Centre, India, as an expert member to help them in setting up their millimeter-wave research facility and as a Guide-cum-Consultant in setting up the experimental facilities for metamaterial research I have gathered some important observations. The present RF and microwave designers can come up with a successful product that will be useful for practical use only if they have strong base and understanding of the nitty-gritty of the design techniques and use of the appropriate simulators for the design which ultimately translates to the actual engineering designed product; of course, knowledge of analytical technique of design makes a design engineer truly confident with the design he/she is making. With all these experiences, I have gained in the last three decades, I planned to write this book so that students, research scholars, and practicing engineers in the RF and microwave domain get benefitted

and earns confidence in designing passive circuits (matching circuits, filters etc.), active devices (oscillators, amplifiers, power combiners, mixers), and a complete system in RF and microwave frequency range.

Chapter 1 deals with the fundamentals of RF and microwave device/circuit principles which will help the user of this book to refresh their ideas/views about the difference between the device and circuit principles at RF and microwave frequencies when compared with the lower-frequency devices and circuits including the circuit analysis. This snap-shot exposer will help the designers to appreciate the approach of design methodology for RF/microwave active device sub-system like oscillators/amplifiers, power combiners, mixers, etc., and different circuit components like matching circuits, filters, etc., and the system as a whole.

Chapter 2 is devoted to the basics of scattering parameters or *S*-parameters useful for characterizing RF and microwave circuits and devices. Unlike at low-frequency, where *Z*, *Y*, and *h*-parameters are used for circuit/device analysis, RF/microwave circuit and device analysis is done with the help of *S*-parameters defined in terms of incident and reflected wave variables. Properties of *S*-parameters and the use of it in deriving various gains in RF/microwave amplifiers have been discussed in this chapter, along with the *S*-matrix of a number of passive RF/microwave components like directional coupler, magic tee, and circulator.

Fundamentals of Smith chart and stub matching design at RF with Smith chart have been discussed in Chapter 3. Smith chart is well-known to the RF/microwave engineers which is a graphical plot of normalized resistance and reactance functions in the complex reflection coefficient plane that helps in understanding how individual variables of device/circuit affects the end result. In modern RF/microwave engineering, this impedance chart has become indispensable in transmission line problem solution and has become an integral part of most of the current computer-aided design software and test equipment. As a matter of fact, most vector network analyzers (useful for RF/microwave circuit characterization) have the Smith chart display as one of the primary screen options.

Chapter 4 deals with the impedance matching circuits in microwave design. Starting with the theory of small reflections the design of impedance transformer both of the binomial and Tchebyscheff stepped impedance transformer has been given. This is followed by the tapered impedance matching transformer of exponential type and with triangular distribution and finally the Klopfenstein taper impedance transformer design is also provided. Impedance matching with lumped component/planer distributed line reactive elements has been included at the end of the chapter which is used in present-day microwave design.

Design of low-pass, band-pass, and band-stop type filter at microwave frequency including the recently emerging metamaterial-inspired filter design is the subject matter of Chapter 5. The chapter begins with

the insertion-loss method of low-pass prototype design followed by filter transformation techniques from low-pass prototype to actual filters of low-pass, high-pass, band-pass, and band-stop type. Then for microwave implementation, Richard's transformation, Kuroda's identity, and impedance/admittance inverters have been discussed. This is followed by practical microwave filter analysis and design for stepped-impedance (high-low) low-pass filter, coupled-line band-pass filter, and coupled-resonator band-stop filter. Finally, metamaterial-inspired filter design for SRR-loaded waveguide filter and CRLH-loaded microstrip filter has been discussed.

Chapter 6 deals with the design of microwave oscillators and amplifiers. The fundamental principle of oscillator/amplifier design with two-terminal device (diode) is discussed first followed by the design of IMPATT oscillator and voltage-controlled oscillator (VCO) with varactor diode has been done using electronic design automation (EDA) tool. Next, the amplifier principle with three-terminal microwave device (transistors like HEMT) has been discussed with special consideration to gain and stability of the amplifier. Finally, low-noise amplifier (LNA) versus power amplifier (PA) characteristics has been discussed, followed by design example for LNA and PA with EDA simulation tool.

Chapter 7 deals with microwave power combiners and mixers those are very useful in microwave transmitters and receivers, respectively. The chapter begins with a brief discussion on different types of power combiners followed by EDA simulation tool-based design examples of short-slot hybrid-based and Wilkinson power combiner/divider. Next different types of microwave mixers have been discussed in brief followed by EDA simulation tool-based design of single-balanced diode mixer.

A brief discussion and design problem related to microwave antenna has been given in Chapter 8. Though there are various RF/microwave antennas used in practice but antenna is a vast subject by itself so we have restricted our discussion only regarding microstrip antenna that is of recent importance for research and in practical use. Both the conventional microstrip patch antenna and metamaterial-inspired microstrip patch antenna have been discussed with design examples.

Chapter 9 is the final chapter of this book. This chapter begins with some technical aspects related to transmitter and receiver followed by different issues related to RF and microwave design like transmitter and receiver parameters to be considered by a designer, link budget, and noise and distortion. Finally, some rudimentary design of RF/microwave communication transmitter and receiver including radar system using EDA simulation tool has been given.

The Appendix covers, in addition to MATLAB programs and other important topics related to different chapters, the fundamentals of computational electromagnetics (CEM) and simulators together with the useful tables for RF/microwave Engineers.

It definitely becomes an honest duty of an author to acknowledge from the bottom of the heart the help, co-operation, and trigger received from all which has ultimately made the writing of this book possible and taking it ultimately to reach the readers for whom the book has been written.

At the very outset, I would like to acknowledge the help and support I have received from my wife Rina and son Sauradeep for writing this book. The quality time they have provided me to write this book with concentrated mind has actually made it possible to translate my dream to reality. Without their active support and encouragement this effort of mine would not have been successful.

I am also thankful to Dr. A L Das, former director of SAMEER Kolkata Centre, who invited me as an expert member to guide their scientists since 2007 to setup "Millimeter-wave R & D Facility" at SAMEER, Kolkata, funded by the Ministry of Communication and Information Technology, Government of India. I am also indebted to some of the scientists of SAMEER Kolkata Centre, namely, Sukhendu Bhanja, Paramita Banerjee, Soumya Chatterjee, Amitesh Kumar, and last but not the least Dr. Arijit Majumder who helped me in solving some of the design problems using Electronic Design Automation (EDA) tools.

Finally, I must thank Mr Gauravjeet Singh Reen, Senior Commissioning Editor of CRC Press, Taylor and Francis Group, and his team. Special thanks also go to Rachael Panthier, Production Editor of Taylor and Francis Group and Sharmila Sasikumar, Project Manager of Newgen KnowledgeWorks including the copyeditor and typesetter for doing their job excellently with patience and meticulousness. I am also thankful to SRS Publishing Services, Puducherry, India, for their assistance in helping me to draw professionally some of the illustrations in this book.

If this book is liked by the inquisitive and interested RF and microwave designers for whom it has been written, then all my efforts will be truly rewarded.

May 27, 2026
Kolkata: 700059, India

Subal Kar

Former Professor and Head
Institute of Radio Physics and Electronics
University of Calcutta, India

About the Author



Subal Kar attended JRB High School, Shillong, India, and earned his B.Sc. degree in Physics (Honours) from St. Edmunds' College, Shillong. Thereafter he joined the Institute of Radio Physics and Electronics (IRPEL), Calcutta University (CU), to complete his B.Tech., M.Tech., and Ph.D. (Tech.) degrees.

Dr. Kar taught microwave engineering and related subjects of wireless technology at both the graduate and postgraduate levels at IRPEL, CU, from 1983 to 2016. In 2016, he retired from IRPEL, CU, as a professor and head of the department. His field of specialization in research covers microwave and terahertz (THz) engineering, metamaterials, and high-energy physics. Dr. Kar has three patents to his credit related to novel developments of IMPATT oscillators, amplifiers, and power combiners. He has published a large number of research papers in peer-reviewed international journals.

Dr. Kar was a visiting scientist at various universities and institutes in the United States, Europe, and Asia, including Lawrence Berkeley National Laboratory (LBNL), USA; Oxford University, UK; Cockcroft Institute, UK; École Polytechnique fédérale de Lausanne (EPFL), Switzerland; and Kyoto University, Japan.

Dr. Kar has worked on "RF structures for a muon collider" at LBNL, during his tenure (1999–2000) as Visiting Fulbright Scholar. He was one of the seven scientists of Berkeley Lab, along with five scientists from Fermilab Lab and other collaborative universities and institutes from the United States, who successfully designed and developed the MUCOOL RF cavities for muon collider.

In 2006, Dr. Kar started research on metamaterials, which was first fabricated at UCSD, USA, in 2001. Dr. Kar and his group at Calcutta University in collaboration with the Society for Applied Microwave Electronics Engineering and Research (SAMEER), Kolkata Centre and

Bhabha Atomic Research Centre (BARC), Mumbai, made the successful design, fabrication, and experimentation with plasmonic metamaterial, first time in India in 2009 (www.nature.com/articles/nindia.2009.273), which could exhibit negative refractive index, a counter-intuitive phenomenon, that is useful in the design of cloaking devices, super-microscope, and so forth. A Board of Research in Nuclear Sciences (BRNS) project to set up a metamaterial research facility at SAMEER, Kolkata Centre, with the title "Left-Handed Maxwell System—Experimental Studies" was successfully completed during 2010–2014, with Dr. Kar as the Guide-cum-Consultant.

Dr. Kar has participated as an expert member (since 2007) for guiding the scientists of SAMEER, Kolkata, in the "Establishment of Millimeter-Wave R&D Facility" sanctioned by the Ministry of Communication and Information Technology, Government of India. During the establishment of this facility, Dr. Kar has guided the scientists of SAMEER Kolkata Centre to design and develop RF, microwave, and millimeter-wave devices, circuits, and systems.

Dr. Kar has authored many books, including *Microwave Engineering—Fundamentals, Design and Applications*, *Physics and Astrophysics—Glimpses of the Progress*, *Metamaterials and Metasurfaces—Basics and Trends, Wireless Technology—Kilohertz to Megahertz, Wireless Technology—Gigahertz to Terahertz, Metamaterials and Metasurfaces Second Edition—Foundations, Trends and Applications* and has also contributed a number of chapters in edited books published by several reputed publications.

Dr. Kar is a reviewer of a number of reputed international scientific journals and a distinguished reviewer of the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, for project proposals related to wireless technology and metamaterials. He is the recipient of the Young Scientist Award of URSI and IEEE MTT, USA, and the Fulbright award from the US government. He is a fellow of the Institute of Electronics and Telecommunication Engineers (IETE) and Vacuum Electronic Devices and Applications (VEDA) Society and a senior member of the IEEE, USA.

Chapter 1

Fundamentals of RF and Microwave Device/Circuit Principles

Fundamentals are essential to delve deep into the subject

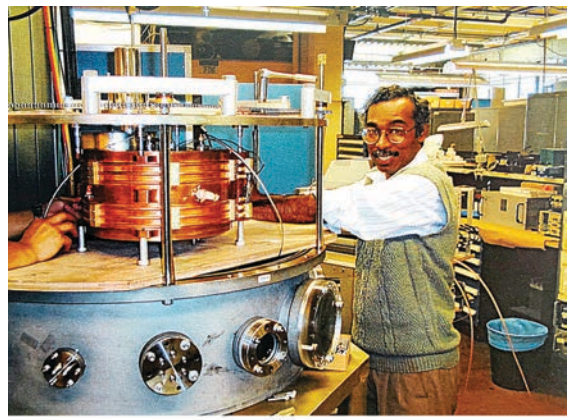
1.1 INTRODUCTION

Radio frequency (RF)/microwave systems drive the requirements of RF/microwave circuits and devices and the capabilities of RF and microwave circuits and devices fuel the evolution of RF/microwave systems. Today, communications are the main driver of RF/microwave system development, leading to RF/microwave technology evolution at an unprecedented pace. Similar relationships exist for radar used in detection and ranging of aircrafts using RF/microwave signal and sensors like Global Positioning System (GPS), thermal cameras, etc., used in drone platform for providing aerial surveillance and real-time threat detection and so on necessary for national security purposes. Other radio systems have less immediate impact on RF/microwave technology but are very important for a smaller number of RF engineers working in the fields of radio-aids for navigation, astronomy, and also RF/microwave heating realized with means other than the age-old fire we have been using from the dawn of civilization.

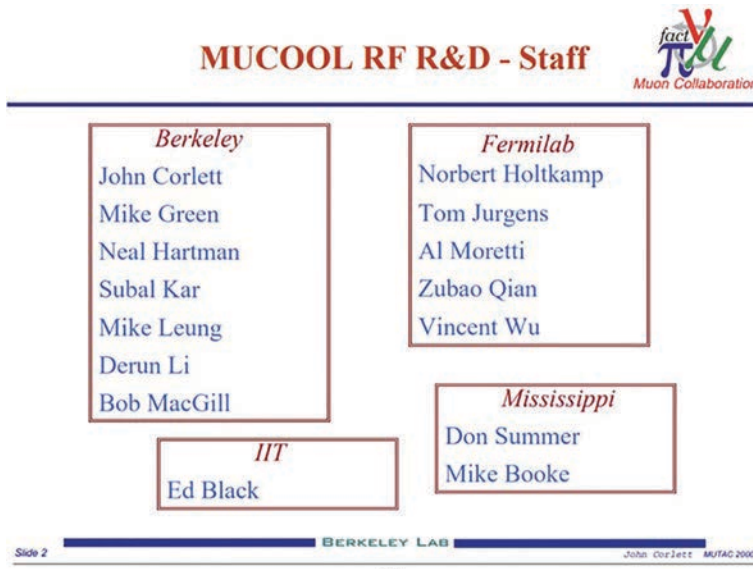
It may be mentioned that in particle accelerators/colliders, used by particle and nuclear physicists around the world, the use of RF cavity is of prime importance to accelerate the particle beams (electron/proton/muon, etc.) to the very high velocities, close to the velocity of light ($c = 3 \times 10^8$ m/s) before they are made to collide in the collider ring of the accelerator. In the Large Hadron Collider (LHC) at CERN, Geneva, Switzerland, the proton beams are accelerated via such a chain of RF cavities to 7 tera-electronvolt (TeV) so as to achieve a velocity of the order of 99.99% of c. I had an experience to work in the design and development of such RF cavities for a Muon Collider at Berkeley Lab during my Fulbright Scholar tenure at Lawrence Berkeley National Laboratory (LBNL), CA, USA, during 1999–2000 (see Figure 1.1a and Figure 1.1b). The RF cavity used in the muon collider is of special design, known as $\pi/2$ interleaved cavity (see the cross-sectional view of the cavity in Figure 1.1c), operating at 805 MHz. This is a pill-box-type cavity with beryllium windows having arrangement for liquid nitrogen (LN2) flow via the cavity to realize very high quality factor (Q = 30,000) of the LN2

DOI: 10.1201/9781003729969-1

cooled cavity; the RF cavity mounted in LN2 cooling chamber is shown in Figure 1.1d. The theoretically simulated unloaded Q of the uncooled cavity being 21,000. The accelerating field generated in the cavity is 30 MeV/m.

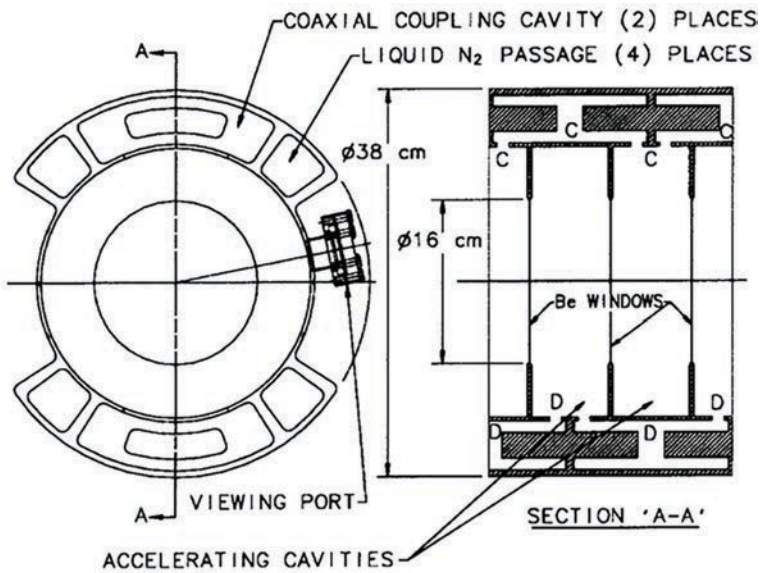


(a)

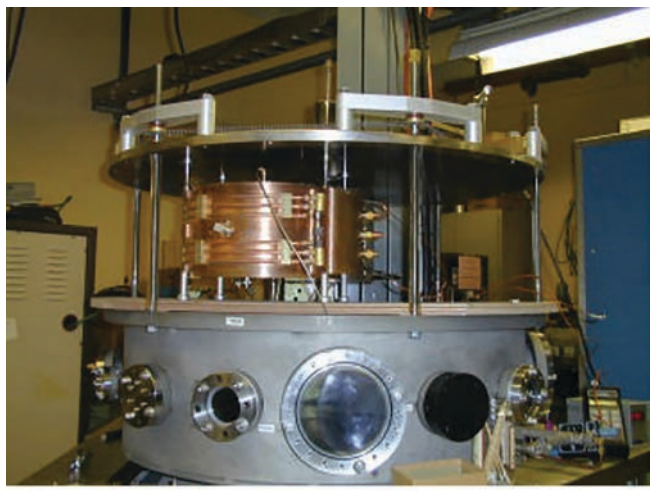


(b)

Figure 1.1 (a) The author is at work with the MUCOOL RF cavity at LBNL, (b) Berkely Lab (US) presentation slide showing the MUCOOL RF R&D staff, (c) Cross-sectional view of the $\pi/2$ interleaved RF cavity, and (d) RF cavity in LN2 cooling chamber. [www.cap.bnl.gov/mumu/conf/nufact00/docs/pdf/wg50525c.pdf].



(c)



(d)

Figure 1.1 (Continued)

Anyway, the prerequisite for RF/microwave system designers is to know about the design of its component parts that are circuit-based like impedance matching circuits, filters and so forth and active-device-based subsystems like oscillators, amplifiers, power combiners, mixers, etc. In this book we have covered the design of most of these circuit-based and active-device-based subsystems and complete communication/radar systems too, built with circuit-based and active-device-based subsystems, at RF/microwave frequencies with design examples using both the analytical and electronic design automation (EDA) tools.

To understand these designs, one has to be conversant with the scattering parameters or *S*-parameters that are used for device and circuit analysis at RF/microwave frequencies and the Smith chart that is exclusively used for microwave device/circuit characterization. Furthermore, RF/microwave system designers need to be exposed to RF/microwave passive components like couplers, isolators, circulators, etc. (which are exclusively used at RF/microwave frequencies; some of which are not at all required in low-frequency systems) and RF/microwave devices (both tube and solid-state devices). However, the details of RF/microwave passive components and device/circuit principles are not in the scope of this book, but are available in standard text books [1–3] and those are being taught in the undergraduate engineering course of electronic and communication (ECE) stream. Anyway, for the benefit of readers we are just glancing through the list of different passive components and active devices (with which the designs of oscillators and amplifiers and so forth are usually done) in a chart form including some basic principles of active device and passive circuit design at RF/microwave frequencies to highlight how they are different from the low-frequency devices and circuits.

Most of the RF/microwave passive components used in modern RF/microwave system design are either waveguide-based or planar transmission line-based—a list of which is shown in Table 1.1.

RF/microwave active devices, required primarily for signal generation and amplification, are available both as tube and solid-state (semiconductor) devices depending on the frequency and power level of the microwave signal required. Different types of RF/microwave tubes are shown in the chartform in Figure 1.2, while the list of solid-state RF/microwave devices is shown Figure 1.3.

Table 1.1 Waveguide and planar transmission line-based RF/microwave passive components

Name of passive components	*Types*
Terminations	Matched Load Variable Short
Tee Junctions and Hybrids	E-Plane Tee and H-Plane Tee Hybrid Tee (Magic Tee), Hybrid Ring (Rat-Race). Wilkinson Power Divider/Combiner, Branch-line Quadrature Hybrid, Short-Slot Hybrid.
Tuners	E-H Tuner Slide-Screw Tuner
Directional Coupler	Multi-hole Directional Coupler, Schwinger Reversed-Phase Coupler, Bethe-hole Coupler. Coupled-line Directional Coupler, Lange Coupler.
Attenuator	Flap Attenuator. Rotary-vane Attenuator.
Isolator	Faraday Rotation Isolator. Resonance Absorption Isolator.
Circulator	Faraday Rotation Circulator—SPDT Switch, Power Divider. Y/T-junction Circulator.
Phase Shifter	Ferrite type, Reflection and Transmission type. Rotary Phase Shifter
Bends and Corners	
Waveguide Twist, Transitions and Adaptor	
Flanges and Connectors	

1.2 LOW-FREQUENCY VS RF AND MICROWAVE DEVICE AND CIRCUIT PRINCIPLES

1.2.1 Introduction

Low-frequency device and circuit characterization is done with Z, Y, and h-parameters which are defined in terms of current and voltage in the device/circuit. But at RF and microwave frequencies current and voltage are derived quantities from magnetic and electric fields, respectively, and the measurable parameters are the scattering parameters or S-parameters which are defined in terms of the incident and reflected wave variables. Hence, designers of RF/microwave circuit, device, and system have to be conversant with the S-parameter and its properties that have been discussed in Chapter 3. Again, for the low-frequency systems the wavelength of operation is very much large compared to the device/circuit dimensions and all analyses at low frequencies are done on the basis of *lumped circuit* analysis. But at RF frequencies, especially at microwave frequencies, device/circuit dimensions become

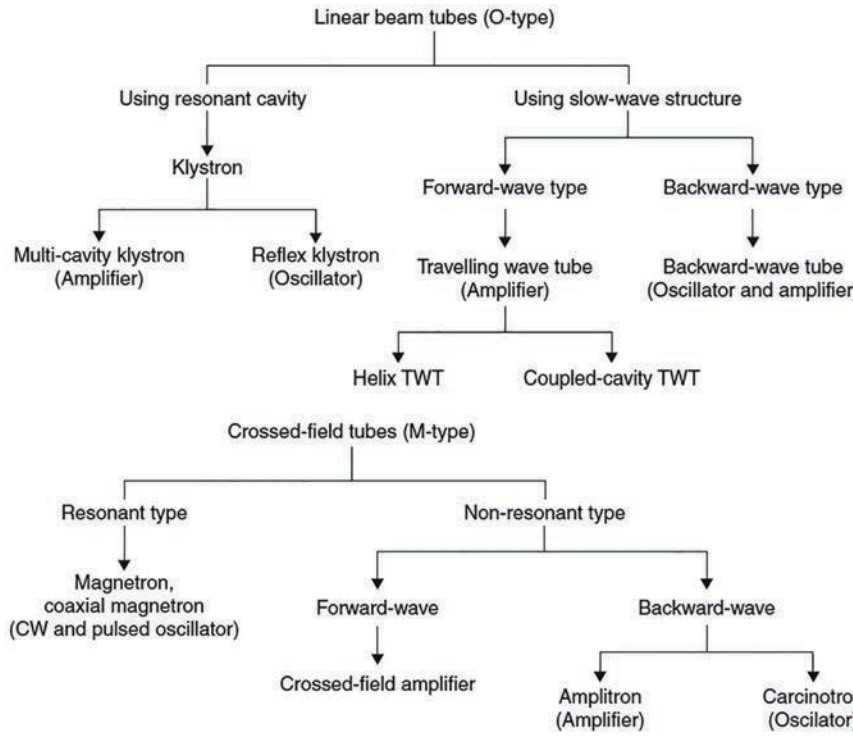


Figure 1.2 RF/microwave tube devices.

comparable to the device/circuit dimensions and thus at RF frequencies we need to do all analysis in terms of *distributed parameters* while at microwave frequencies *field analysis* is resorted to. Thus, RF/microwave designers need to have prerequisite knowledge of all these specialized analysis techniques at RF/microwave frequencies, for which they might refresh their knowledge by going through standard text books [3,4]. However, for the benefit of readers we are briefly glancing through salient features of device and circuit design principles at RF/microwave frequencies, bringing out how they differ from the low-frequency device and circuit design. In this perspective, a quick look through the position of RF and microwaves in the electromagnetic spectrum, in terms of frequency/wavelength (shown in Figure 1.4), may be instructive. A list of RF bands specified by the International Telecommunication Union (ITU) and microwave frequency bands specified by the Institue of Electrical and Electronics Engineers (IEEE) and also those used by military and for commercial use are given in Appendix I.

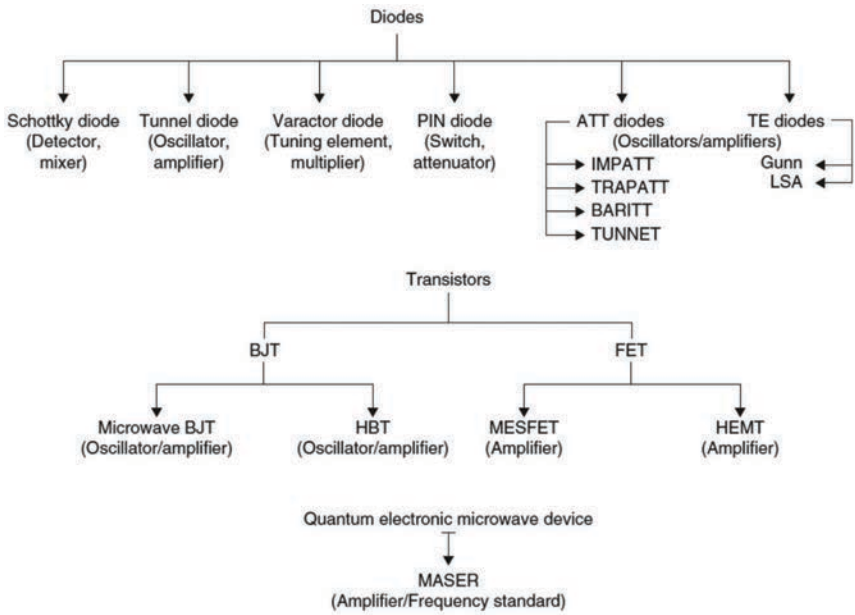


Figure 1.3 RF/microwave solid-state devices.

1.2.2 Device Principles at RF/Microwave Frequencies

When we talk about the RF/microwave device design principle we immediately come up with the issue of "transit time" of the electrons traversing through the device which becomes comparable to the operating frequency, especially at higher RF (i.e., at UHF) and microwave frequency. Thus, the straight forward principle followed in the design of RF/microwave devices is not to fight against the transit time issue but to use it in favor of the operation of the device: following the advice of the pragmatists: "if you cannot beat them join them."

Initial efforts to design UHF, i.e., higher RF, and microwave tube design were to try to reduce the transit time of electrons from cathode to plate, as far as practically possible, in the tube designs like "Disk-sealed tube" and "Light-house tube." But the later successful effort for the design of microwave tube devices like klystron, Traveling-Wave Tube (TWT), and magnetron; transit time of electrons in favor of the operation of the tube was used. This helped in the formation of the so-called "electron bunch," in some form or the other, within the tube (see Applegate diagram in Figure 1.5 for two-cavity klystron), using electron transit time. This has facilitated in the generation/amplification of signal at microwave frequency.

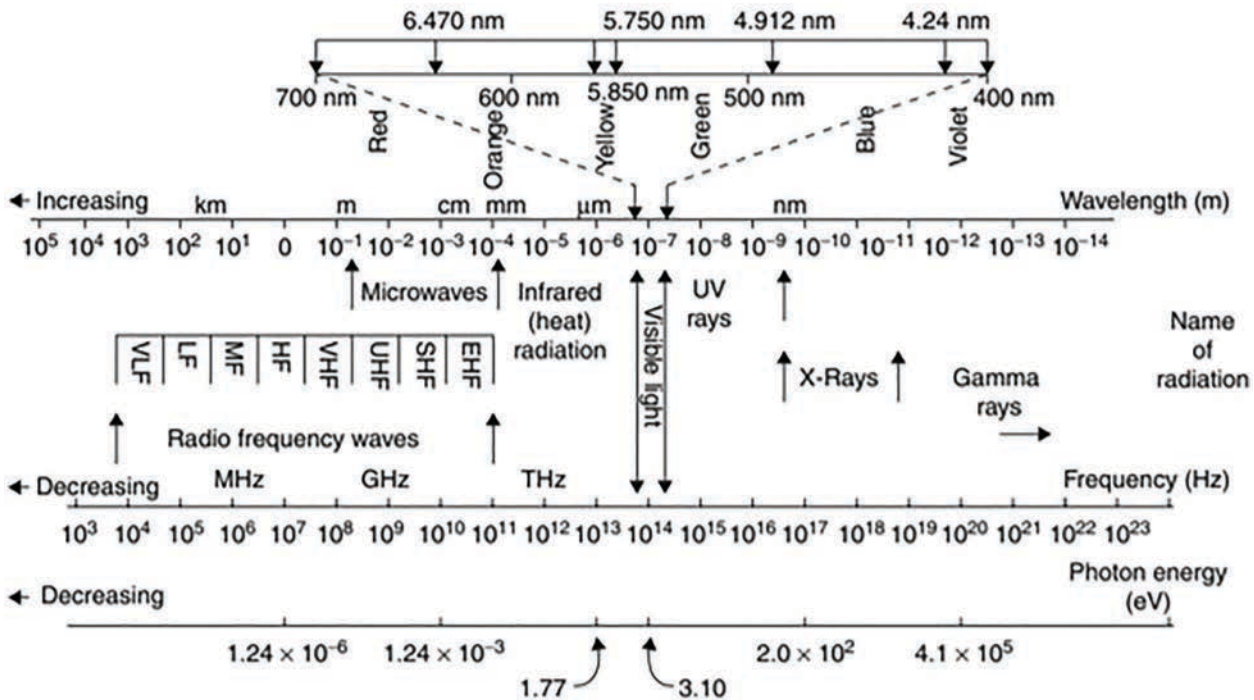


Figure 1.4 Electromagnetic spectrum showing the position of RF and microwave frequency/wavelength range.

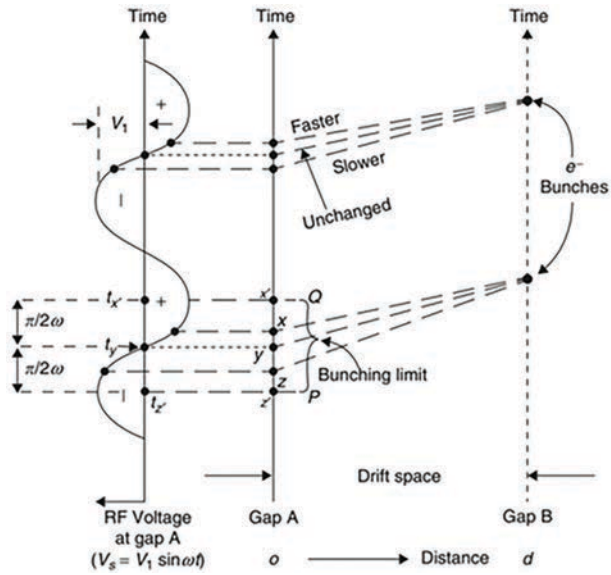


Figure 1.5 Applegate diagram illustrating bunching of electrons in a two-cavity klystron.

In microwave transistor design, various steps were also taken to make it suitable for microwave operation. Since its invention in 1949, the transistor (*trans*fer res*istor*), in which the voltage on one terminal is transferred to control the current through the other two terminals, was confined only to lower RF frequency applications. However, with the advent of many technological breakthrough in semiconductor processing technology such as alloy junction, epitaxial growth, lithography, dry etching, and in other related areas, high frequency transistors (above 1 GHz) have been made possible subsequently for commercial use. However, ultimate approach that was resorted to for the design of Metal-Semiconductor Field-Effect Transistor (MESFET) and High Electron Mobility Transistor/Heterojunction Bipolar Transistor ((HEMT/HBT) was novel and innovative. In MESFET, metal–semiconductor contact (i.e., Schottky-barrier gate-source contact) is used to get over the problem of the so-called minority carrier storage that occurs when a semiconductor–semiconductor junction FET is used, as in low-frequency RF devices. In Schottky junction, there is no question of minority carrier (since here only majority carriers come into play for the operation of the device) and hence no problem of minority carrier storage occurs in MESFET allowing its higher RF and microwave frequency operation.

In HEMT/HBT device for high-frequency operation of the device, the so-called heterojunction principle is used, unlike the homojunction device

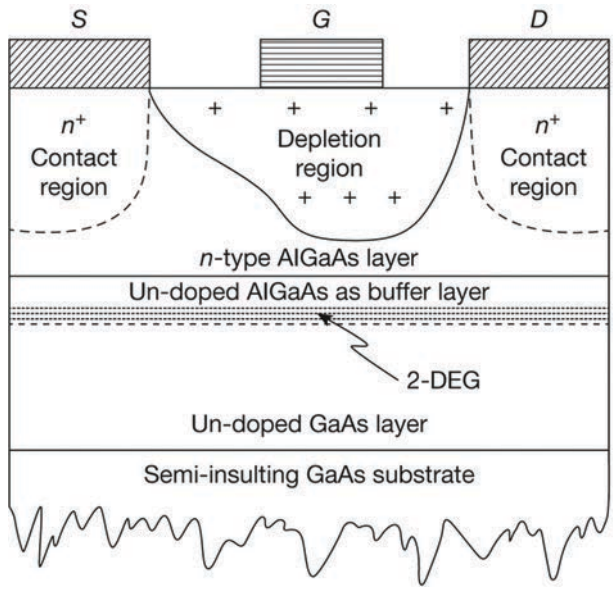


Figure 1.6 Cross-sectional view of HEMT device.

principle used in the lower RF frequency transistor devices. When the transistor junction is joined by two similar semiconductor materials such as silicon to silicon or germanium to germanium, it is a *homojunction* transistor. The transistor junction formed by two different materials, such as Ge to GaAs (semiconductor materials having different energy gaps), is called a *heterojunction* transistor. However, in constructing a heterojunction structure, the lattice constant (a) of the two semiconductor materials must match (a = 5.646 Å for Ge and 5.653 Å for GaAs). This is very important because the lattice mismatch could introduce a large number of interface states and degrade the heterojunction operation. Further, in HEMT, to increase the mobility of electrons within the device (for high frequency operation, millimeter-wave and beyond), some novelty in structural design is resorted to, known as "2D electron gas" formation that separates the immobile atoms from the transporting electrons, thereby allowing higher mobility of the electrons through the device (see the structure of HEMT in Figure 1.6).

1.2.3 Circuit Principles at RF/Microwave Frequencies

At RF and microwave frequencies we cannot use our well-known resistors, inductors, and capacitors, used at low-frequency design, for designing circuit-based RF/microwave subsystems and also the conventional electrical

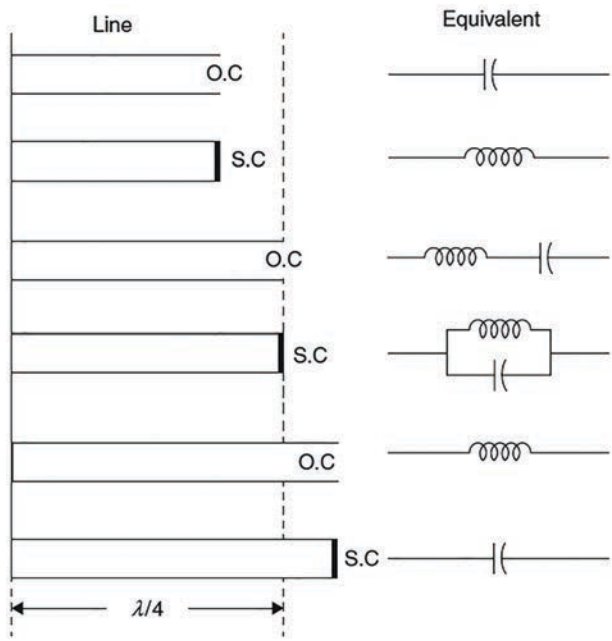


Figure 1.7 Transmission line sections and their LC equivalents.

wires used in our electrical power lines (50/60 Hz) cannot be used for transmission of RF/microwave frequencies. This is because a coil-type inductor at low frequency will start behaving as an undesired resonant circuit with inter-turn capacitance at RF/ microwave frequency. RF transmission line sections with proper terminations are thus used as inductors and capacitors at and above 300 MHz (see Figure 1.7), while waveguide sections are used at microwave frequencies above 3 GHz. Also, at RF/microwave frequency normal electric wire we are familiar with will start radiating as the wirelength will become many wavelengths, especially at higher RF and microwave frequency. Thus, RF transmission lines are used at RF frequencies and waveguides are used at microwave frequency for signal transmission and component and subsystem (oscillator/amplifier) design.

The "transmission line" is a generic term used for describing a structure that is used for the transport of electrical or electromagnetic signals from one point in the system to another. All of us are familiar with the two-wire overhead electrical power transmission lines used for the transport of electric energy (50/60 Hz) from the generating station to households or industrial sites. However, the *twin-wire line* may also be used for voice signal transmission as in voice telephony system (operating in the frequency

range: 300 Hz to 3.4 KHz) or as a flat lead-in line from a roof top Yagi-Uda antenna to a television receiver. The twin-wire line (Figure 1.8a) may be used satisfactorily to transport electromagnetic energy up to 300 MHz (at higher frequencies it becomes difficult to maintain the small distance between the two conductors for reasonable coupling that is required for guided-wave RF signal transmission via such lines). Further, at higher frequencies, the radiation losses become far too much to be tolerated. However, the radiation loss in a two-conductor transmission line may be made negligible with a different configuration where a solid cylindrical conducting rod is placed coaxially with an outer conducting tube (Figure 1.8b), which is known as a *coaxial transmission line*. The fields in the coaxial line remain confined to the annular region between the inner and the outer conductors (which is filled with a dielectric material to keep the conductors in position). Thus, the signal cannot radiate out and, pick-up of spurious signals cannot take place in such "shielded" two-conductor coaxial transmission lines. However, such coaxial line is also not useful at higher RF frequencies as the loss in the dielectric increases with increasing frequency of operation and also such transmission line cannot be used at higher power level as the dielectric softens at higher power level due to the so-called dielectric breakdown. The principal mode of propagation either in twin-wire line or in the coaxial line is transverse electromagnetic (TEM), which is also the mode for electromagnetic wave propagation in free-space, with the electric and magnetic field distribution on the lines as indicated in Figure 1.8c.

A fundamental difference between the signal transmission at power line frequency and that at RF frequency is that, the wavelength of the signal at power line frequency (50/60 Hz) is very much larger than the size of the electrical networks. Thus, the circuit elements in an electrical network are considered to be "discrete" and are usually said to be described by "lumped parameters." It is assumed that currents flowing in lumped-circuit elements do not vary spatially over the elements and that no standing waves exist. But at RF, the component size is a considerable fraction of the wavelength, especially at microwave frequency it may be many wavelengths long. An RF transmission line is thus said to behave as a "distributed-parameter network"—with current, voltage, and impedance varying all along the length of the line.

The coaxial line is used as a transmission structure for electromagnetic signals up to 3 GHz (with a specialized design, it can work up to 110 GHz, but becomes very expensive due to the requirement of very low-loss dielectric) above which the losses in the dielectric (that increases with frequency) and also in the conductor become appreciable and cannot be neglected. The telephone and TV cables and the input cables to microwave precision measuring instruments are generally the coaxial lines. The losses incurred in the coaxial line may be significantly reduced by dispensing with the central conductor and hence the dielectric support, using merely a hollow metal tube of rectangular or circular cross-section (Figure 1.8d) with air dielectric. Such

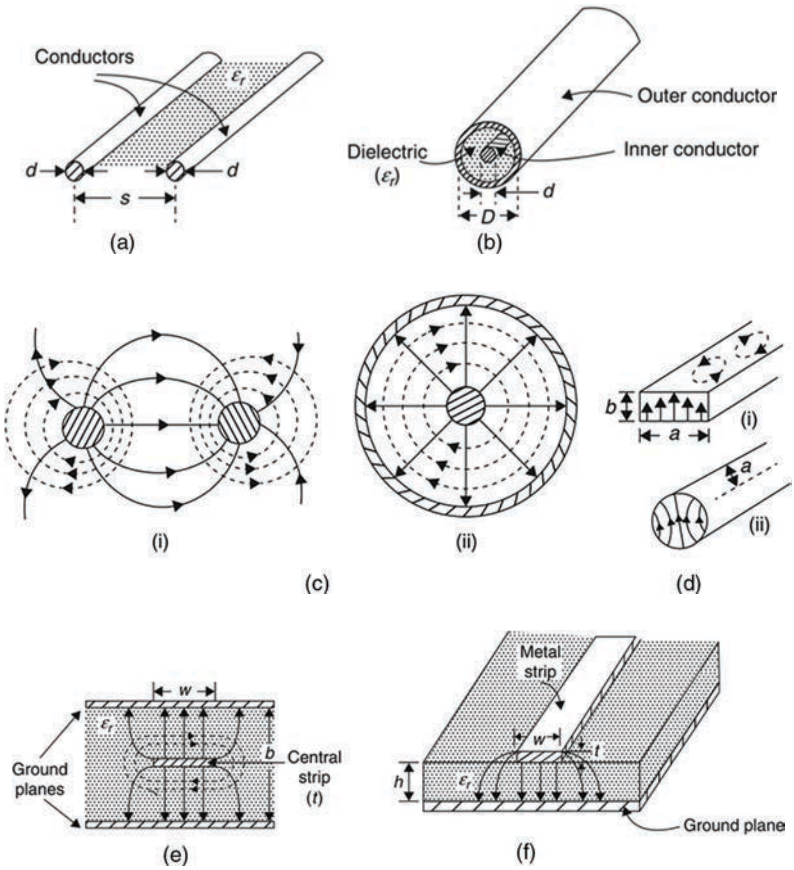


Figure 1.8 Different types of RF/microwave transmission lines: (a) Twin-wire transmission line (s >> d), (b) Coaxial transmission line, (c) Field lines (E → full lines, H → dashed lines) for (i) Twin-wire line, and (ii) Coaxial line, (d) Waveguide: (i) Rectangular (with TE_{10} mode), and (ii) Circular (with TE_{11} mode), (e) Stripline (cross-sectional view with field lines), and (f) Microstrip line.

a system is known as a *waveguide*. A waveguide differs from the twin-wire or the coaxial transmission line (which, in principle, is also a wave-guiding structure for RF signal transmission) in its construction in that it is a single conductor system. The field analysis of the waveguide [4] shows that the boundary conditions to be satisfied by the fields inside a hollow metal waveguide are such that the TEM-mode cannot exist. Instead, either a transverse electric (TE) or transverse magnetic (TM) mode propagates down a waveguide (see Figure 1.9a for TEM, Figure 1.9b for TE, and Figure 1.9c for TM modes, respectively). In TEM mode, the mutual orientation of the electric

and magnetic field vectors and the direction of propagation of uniform plane wave are such that $\vec{E}$ and $\vec{H}$ fields are at right angles to each other and the direction of propagation (z-direction) is perpendicular (i.e., transverse) to the plane (wavefront) containing the E- and H-field vectors in the x-y plane. For TE mode or H-wave in a waveguide, the E-field is absolutely transverse to the direction of propagation (z), i.e., $E_z \equiv 0$, but the magnetic field has a component along the direction of wave propagation (see Figure 1.9b). But for TM mode or E-wave, the H-field is absolutely transverse to the direction of propagation (z), i.e., $H_z \equiv 0$, but the electric field has a component along the direction of wave propagation (see Figure 1.9c).

The attenuation of the electromagnetic signal propagating through an air-dielectric (which is normally the case) waveguide structure is extremely low. Although waveguides are the best transmission line structures at microwave frequencies (covering the range 0.3–300 GHz), so far as the efficiency of power transmission is concerned, they cannot be practically used at lower microwave frequencies. This is because the cross-sectional dimensions of

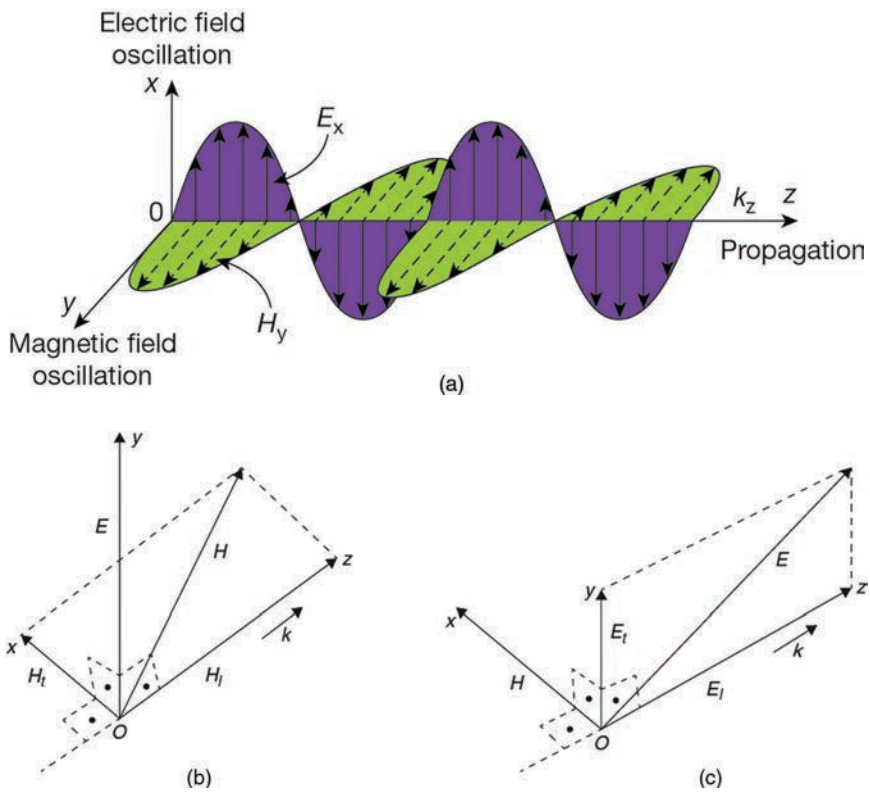


Figure 1.9 (a) TEM mode, (b) TE mode or H-wave, and (c) TM mode or E-wave.

such metal pipes or waveguides have to be directly related to the wavelength of operation (broad dimension approximately half-wavelength and narrow dimension approximately a quarter-wavelength in the case of a rectangular waveguide) for transmission of power to take place through them. It is therefore only when the frequency of operation is a few gigahertz (say > 3 GHz) that the cross-sectional dimensions are small enough for waveguides to be convenient for practical use. So far as the propagation of electromagnetic waves along the waveguides is concerned, in contrast to the coaxial line, the waveguides possess a low-frequency cutoff and exhibit a high-pass property or more specifically, a band-pass filter property, see Figure 1.10, (thus comparatively narrow band), whereas the operation of coaxial line in principle begins from DC and goes up to a few GHz (hence a broadband structure). It may be noted herewith that when an electromagnetic wave emanates from the source, there exist both the propagating (or far-field) waves and evanescent (or near-field) waves—the latter contains the sub-wavelength details (and hence the finer details) of the source. The

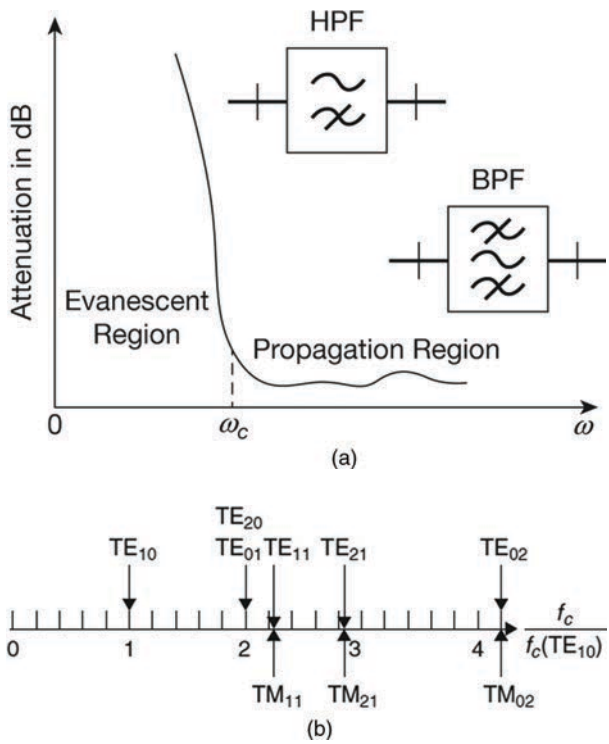


Figure 1.10 (a) Evanescent and propagation regions in waveguide, and (b) Different possible modes in rectangular waveguide with cut-off frequency relative to the dominant TE_{10} mode.

term "evanescence" has been derived from Latin which means "tending to vanish like vapor." A hallmark of an evanescent field is that there is no net energy flow in that region. Since the net flow of electromagnetic energy is given by the average Poynting vector, this means that the Poynting vector in these regions, as averaged over a complete oscillation cycle, is zero. The analysis of waveguides is done with Maxwell's equations with proper boundary conditions. A number of good text books are there for waveguide analysis [1, 4] which the readers might refer to; however, a brief review of Maxwell's equations and electromagnetic wave equations has been given in Appendix VI for the benefit of readers.

Another class of transmission lines that is important for Microwave Integrated Circuit (MIC) is *planar transmission lines*. Phenomenologically, when a coaxial line is thought of as having planar symmetry, we get the fundamental planar transmission line—the strip-line (Figure 1.8e). A common type of practical planar transmission line is the microstrip, in which a thin metal strip is deposited on a dielectric substrate backed by a ground plane (Figure 1.8f). The two main advantages obtained with planar transmission lines are the reduction in size and weight and the ability to use planar integrated circuit technology for their fabrication. However, the power handling capability of such transmission line structures is comparatively low and is thus applicable for low and medium power applications but when demand for miniaturization and integration is a practical need, such planar transmission lines are exclusively used.

A few words about the most important planer transmission line, the microstrip line [5], will be instructive herewith. It consists of a central conducting strip deposited on a dielectric material of dielectric constant ε_r, with the width W of the strip much greater than its thickness t, ($W >> t$) and a ground plane that backs the dielectric from the bottom. A cross-sectional view of a microstrip line is shown in Figure 1.11, together with the typical electric and magnetic field lines. The distribution of electric field lines shows that the pure TEM mode cannot propagate in the microstrip line, rather, the mode that propagates is quasi-TEM.

The quasi-TEM propagation of electromagnetic waves in a microstrip line is due to the fact that a part of the electromagnetic field is traversing the air–dielectric composite in an inhomogeneous medium. The propagation velocity of the electromagnetic wave in the two dielectric media is different and thus an effective dielectric constant has to be considered. However, it is advisable to use a substrate (i.e., the dielectric material) with a high dielectric constant [alumina (ε_r = 8 – 10), sapphire (ε_r = 9.4)] so that the electromagnetic field will be confined more effectively in the substrate dielectric and hence near the strip. Further, substrates with high dielectric constants are advantageous since they reduce the phase velocity (and thus guide wavelength) and consequently the proportional circuit dimensions would be reduced too. The quasi-TEM propagation means a hybrid mode

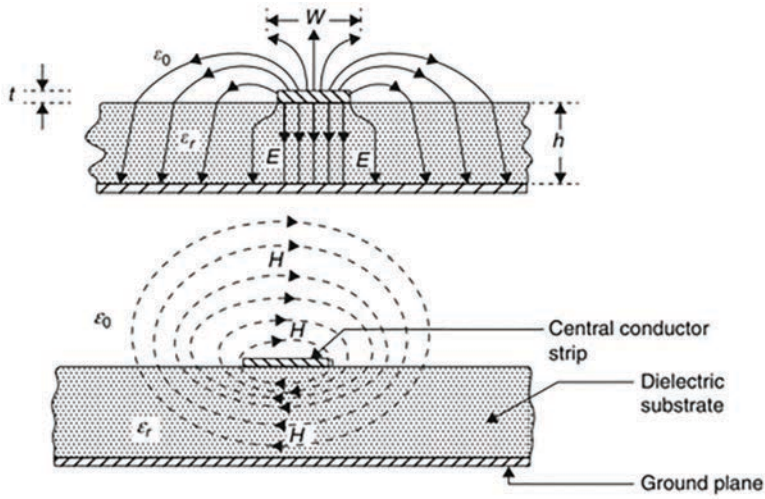


Figure 1.11 Cross-sectional schematic of microstrip line showing the electric (E) and magnetic (H) field pattern.

propagation being characterized by a surface-wave mode propagating in the inhomogeneous air–dielectric media.

REFERENCES

[1] Rizzi P A, *Microwave Engineering Passive Circuits*, Prentice-Hall of India Pvt. Ltd., New Delhi, 1999.
[2] Liao S Y, *Microwave Devices and Circuits*, Third Edition., Pearson Education Inc., India, 2003.
[3] Pozar D M, *Microwave Engineering*, Fourth Edition, Wiley, New York, 2012.
[4] Jordan E, and Balmain K G, *Electromagnetic Waves and Radiating Systems*, Second Edition, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1968.
[5] Edward T and Steer M B, *Foundations for Microstrip Circuit Design*, Fourth Edition, Wiley-IEEE Press, West Sussex, UK, 2016.

Chapter 2

Basics of Scattering Matrix Useful for RF and Microwave Device/ Circuit Characterization

Thoughts get scattered in the material world;
while, in the spiritual world they are always in consolidation

2.1 INTRODUCTION

Devices and circuits in a network at low frequency are characterized in terms of *Z* (impedance), *Y* (admittance), or *h* (hybrid) parameters—the latter being often used to describe transistor characteristics. These parameters are defined in terms of voltage and current in the device/circuit. This is because at low frequency, the device/circuit dimensions are small compared to the wavelength, and thus the devices and circuits may be considered as *lumped* components for which voltages and currents can be defined uniquely at any point in the network. However, at microwave frequency the device/circuit dimensions become comparable to wavelength which makes them *distributed* components, i.e., the phase of the signal changes from one point to another in the network. This calls for field analysis of microwave network. The voltages and currents are no longer measurable quantities (as they are difficult to measure) at RF and microwave frequency but are actually the derived quantities (voltage derived from electric field and current derived from magnetic field). The measurable quantities at RF/microwave frequency are the incident, reflected and transmitted power in the network (the reflected and transmitted powers are being generated due to scattering of the incident signal at the RF/microwave junction). Again, in view of the linearity of the field equations and most devices, the scattered wave amplitudes are linearly related to the incident wave amplitudes. The matrix describing this linear relationship between the reflected and transmitted waves with incident wave amplitude is called the *Scattering Matrix* and the microwave device/circuit is analyzed in terms of the so-called scattering parameters or *S-parameters*.

Unlike the *Z*, *Y*, *h*, and *ABCD*-parameters (the last one being used at low frequencies for cascaded networks), the *S*-parameters do not use short- and open-circuit terminations. From a practical view point this has significance, as it is difficult to obtain perfect open-and short-circuit termination at RF and microwave frequency and also the impedance of short and open varies with frequency and thus defining parameters to characterize RF and

DOI: 10.1201/9781003729969-2

microwave device/circuit should not use them. Further, the presence of short/open eventually causes complete reflection of signal from the load terminal, leading to possible oscillation, and might even cause damage to the source. But the use of matched load in defining *S*-parameters would cause absorption of the incident energy and thus causing no reflection back to the source.

Similar to the low frequency transmission matrix (the *ABCD*-matrix) at microwave frequency, the transmission or *T*-matrix (alternatively known as chain *S*-matrix) is used when dealing with multistage circuits (such as filters and amplifiers) or infinitely long periodic structures such as those used in circuits for traveling wave tubes.

Further, it may be noted that both *S*- and *T*-matrix are useful for linear circuits. But to analyze nonlinear circuits and systems a new class of parameters, known as *X*-parameters, is defined. While *S*-parameters are useful for DC and small-signal analysis in linear circuits the *X*-parameters are useful for large-signal conditions, which is usually the case with nonlinear circuit analysis that we come across in harmonic mixing, nonlinear reflection effects, etc. However, in the limit of small-signal approximation the *X*-parameters reduce to *S*-parameters. The calculations based on *X*-parameters are usually performed within a harmonic balance simulator environment.

S-parameters have many advantages over *Z*, *Y*, and *h*-parameters. They relate to familiar measurements such as gain, loss, and reflection coefficient. The measured *S*-parameters of multiple devices can be cascaded to predict overall system performance. If desired, *Z*, *Y*, and *ABCD*-parameters can be derived from *S*-parameters and vice versa [1]. From RF/microwave designers' point of view, it is also very important that the *S*-parameters can be easily imported and used for circuit simulations in electronic design automation (EDA) simulation tools. *S*-parameters are the shared language between simulation and measurement. Since the 1970s the popularity of *S*-parameters has increased with the advent of the new generation network analyzers which could perform *S*-parameter measurement with ease.

2.2 *S*-PARAMETER FORMULATION OF MICROWAVE NETWORK

Let us consider an *n*-port microwave junction shown in Figure 2.1. Let a wave with an associated equivalent voltage V_1^+ is incident on the junction at terminal plane t_1. This will generate a reflected wave in line 1 being given by: $S_{11}V_1^+ = V_1^-$, where S_{11} is the reflection coefficient, or scattering coefficient, for line 1, with a wave incident at line 1. Waves will also be transmitted, or scattered, out to the remaining $(n - 1)$ lines from line 1 and will have amplitudes proportional to V_1^+. These amplitudes can be expressed as: $V_i^- = S_{i1}V_1^+$ where $i = 2,3,\ldots, n$, (S_{i1} is the transmission coefficient for signal transmitted to *i*th line from line 1).

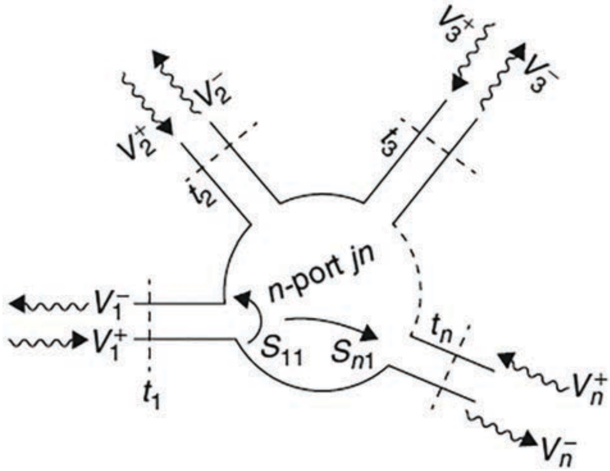


Figure 2.1 An *n*-port microwave junction.

When waves are incident in all the *n* lines, the wave voltage in a particular line has its own reflected wave voltage plus the scattered wave voltage contributions from all the other (*n* -1) lines. Thus, in general, we can write [2]:

$$
\begin{aligned}
V_1^- &= S_{11}V_1^+ + S_{12}V_2^+ + S_{13}V_3^+ + \ldots\ldots\ldots + S_{1n}V_n^+ \\
V_2^- &= S_{21}V_1^+ + S_{22}V_2^+ + S_{23}V_3^+ + \ldots\ldots\ldots + S_{2n}V_n^+ \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
V_n^- &= S_{n1}V_1^+ + S_{n2}V_2^+ + S_{13}V_3^+ + \ldots\ldots\ldots + S_{nn}V_n^+
\end{aligned} \tag{2.1}
$$

which can be written in matrix form as: $[V^-] = [S][V^+]$, where both $[V^-]$ and $[V^+]$ are column matrices and S is an $n \times n$ matrix called scattering matrix given by:

$$
\begin{bmatrix} V_1^- \\ V_2^- \\ \cdot \\ \cdot \\ V_n^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & . & . & S_{1n} \\ S_{21} & S_{22} & S_{23} & . & . & S_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ S_{n1} & S_{n2} & S_{n3} & . & . & S_{nn} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \\ \cdot \\ \cdot \\ V_n^+ \end{bmatrix} \tag{2.2}
$$

The coefficients S_{11}, S_{12}, ……S_{nn} are called the *Scattering parameters* (*S-parameters*) or *scattering coefficients*.

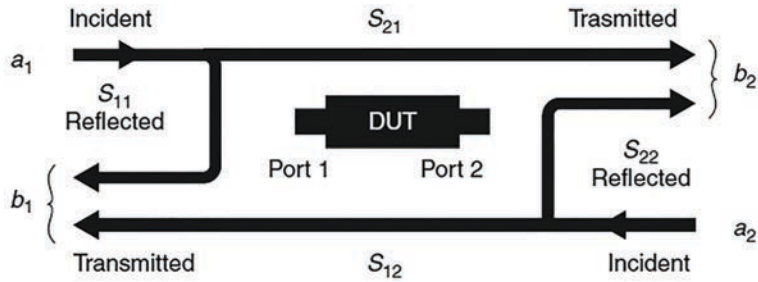


Figure 2.2 A two-port RF/microwave network showing the incident and reflected plus transmitted wave variables and the *S*-parameters.

The scattering matrix definition so far was implicitly based on the assumptions that all ports have the same characteristic impedance. Although this may be the case in many practical situations, for general analysis we need to consider the case when each port has nonidentical characteristic impedance. This will also bring out some significant meaning of the wave variables: the reflected wave variables (b_1, b_2) as dependent variables and the incident wave variables (a_1, a_2) as independent variables (see Figure 2.2 (a two-port network); in terms of which we are going to define the *S*-parameters.

Taking each port's characteristic impedance into account, we define two normalized voltage waves:

$$a_i = \frac{V_i^+}{\sqrt{Z_{0i}}} \tag{2.3a}$$

$$b_i = \frac{V_i^-}{\sqrt{Z_{0i}}} \tag{2.3b}$$

where a_i and b_i represent the normalized incident and reflected voltages, respectively, at the ith port and Z_{0i} is the characteristic impedance of the ith port (which is a real number for a lossless line).

Thus, the total voltage and current at each port can now be written as:

$$V_i = V_i^+ + V_i^- = \sqrt{Z_{0i}}(a_i + b_i) \tag{2.4a}$$

$$I_i = \frac{V_i^+}{Z_{0i}} - \frac{V_i^-}{Z_{0i}} = \frac{(a_i - b_i)}{\sqrt{Z_{0i}}} \tag{2.4b}$$

The average net power delivered to the *i*th port can now be expressed in terms of a_i and b_i (with no concern of the fact that different ports have different characteristic impedance) as:

$$P_i = \frac{1}{2}\mathrm{Re}\left[V_i I_i^*\right] = \frac{1}{2}\mathrm{Re}\left[\left|a_i\right|^2 - \left|b_i\right|^2 + (a_i^* b_i - a_i b_i^*)\right]$$

$$= \frac{1}{2}\left[\left|a_i\right|^2 - \left|b_i\right|^2\right] \tag{2.4c}$$

Since $\left(a_i^* b_i - a_i b_i^*\right) = 0$, when we take real part of it.

This equation is significant, as it clearly shows that the net power delivered to each port is equal to the normalized incident power less by the normalized reflected power. By defining the wave variables in this way, it may be seen that the square of any wave variable becomes proportional to the average power in the wave, and these are now directly measurable quantities at microwave frequency.

The generalized *S*-matrix can now be defined in terms of the normalized voltage wave variables *a* and *b* as:

$$[b] = [S][a] \tag{2.4d}$$

For a two-port network we may thus write:

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{2.4e}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{2.4f}$$

where S_{11} and S_{22} are the reflection coefficients at ports 1 and 2, respectively, and S_{12} and S_{21} are the transmission coefficients, respectively, to port 1 from port 2 and to port 2 from port 1.

Now we try to understand the physical significance of the S-parameters. Referring to Eq. (2.4e) we find that:

$$S_{11} = \left.\frac{b_1}{a_1}\right|_{a_2=0} \quad (\textit{input reflection coefficient}) \tag{2.5a}$$

The variable a_2 is zero when there is no incident wave at port 2. The parameter S_{11} is, therefore, the reflection coefficient at port 1 when port 2 is terminated with a matched load.

Again from Eq. (2.4f) we have:

$$S_{21} = \left.\frac{b_2}{a_1}\right|_{a_2=0} \quad (\textit{forward transmission coefficient}) \tag{2.5b}$$

This gives the ratio of a wave leaving port 2 for a wave entering port 1, i.e., the transmission coefficient from port 1 to port 2 under matched condition at port 2 (for a lossy transmission line or waveguide, S_{21} should be less than 1, and it indicates the amount of attenuation).

The other two parameters can also be found from Eqs. (2.4e) and (2.4f) as follows:

$$S_{22} = \left.\frac{b_2}{a_2}\right|_{a_1=0} \quad (\textit{output reflection coefficient}) \tag{2.5c}$$

$$S_{12} = \left.\frac{b_1}{a_2}\right|_{a_1=0} \quad (\textit{reverse transmission coefficient}) \tag{2.5d}$$

It may be observed that S_{22} is the reflection coefficient at port 2 with port 1 matched and that S_{12} gives the transmission coefficient from port 2 to port 1 under matched condition at port 1.

In general, the scattering parameters are complex numbers having both magnitude and phase.

2.3 PROPERTIES OF *S*-PARAMETERS

Depending on the nature (reflectionless/reciprocal/lossless) of the microwave junction the *S*-parameters will satisfy certain conditions and these properties of *S*-parameters are very useful in simplifying the *S*-matrix of a microwave junction and to draw useful inference regarding the electrical properties of microwave passive circuit components and active devices.

Property I: Reflectionless Junction: We have seen that the parameters S_{11} and S_{22} represent reflection coefficient at ports 1 and 2 under matched condition at ports 2 and 1, respectively, of a two-port network. Therefore, if the reflection coefficients are zero at the respective ports of a two-port network, then $S_{11} = S_{22} = 0$. Most microwave components are designed to have small reflections (VSWR ~ 1.02). Thus, it is often justified to set the diagonal elements of their *S*-matrix to be zero; under this circumstances, the *S*-matrix becomes a *skew-symmetric matrix*. For an ideal *n*-port junction with no reflection at any port under matched condition we thus have:

$S_{ii} = 0$, $i = 1,2,3 \ldots\ldots n$, for matched junction with no reflection.

Property II: Reciprocal Junction: A reciprocal two-port junction has the same transmission characteristics in either direction, i.e., the output of the two-port junction is the same for either port 1 or port 2 and is used as input/output port. Thus, for a two-port reciprocal junction: $S_{12} = S_{21}$. For

an n-port junction, in general, $S_{ij} = S_{ji}$, $i \neq j$. This means that for a reciprocal microwave junction: $[S] = [S]^t$, where $[S]^t$ is the transpose of the matrix S, when the S-matrix becomes a *symmetrical matrix*.

Property III: Lossless Junction: For a lossless junction the total power leaving all the n-ports must equal the sum of the incident power at all those n-ports. The mathematical statement of this power conservation condition is:

$$\sum_n (a_k a_k^* - b_k b_k^*) = 0 \quad \text{or} \quad \sum_n b_k b_k^* = \sum_n a_k a_k^* \tag{2.6}$$

This condition imposes a number of restrictions on the S-matrix coefficients which is discussed below. Following Eq. (2.4) we may write:

$$b_k = \sum_{i=1}^{n} S_{ki} a_i \tag{2.7}$$

(where k is the parameter and i is the running variable). Substituting Eq (2.7) into Eq. (2.6) we have:

$$\sum_{k=1}^{n} \left| \sum_{i=1}^{n} S_{ki} a_i \right|^2 = \sum_{k=1}^{n} a_k a_k^* \tag{2.8}$$

Since a_k's are independent, we may choose all a_k's except a_i to be zero when we get from Eq. (2.8) above:

$$\sum_{k=1}^{n} \left| S_{ki} a_i \right|^2 = a_i a_i^* = \left| a_i \right|^2 \tag{2.9a}$$

or

$$\sum_{k=1}^{n} \left| S_{ki} \right|^2 = \sum_{k=1}^{n} S_{ki} S_{ki}^* = 1 \tag{2.9b}$$

Since the index i is arbitrary, the above equation is valid for all values of i (i = 1,2,3… n). Eq. 2.9 states that for a lossless microwave network, the product of any column of S-matrix with the complex conjugate of this same column equals unity. This is known as *unity property of S-matrix*.

In addition to the above, another constraint may be imposed on the S-parameters. By selecting all a_k's except a_s and a_r to be zero ($r \neq s$) Eq. (2.8) may be written as:

$$\sum_{k=1}^{n}\left|S_{ks}a_s + S_{kr}a_r\right|^2 = \sum_{k=1}^{n}(S_{ks}a_s + S_{kr}a_r)(S_{ks}a_s + S_{kr}a_r)^* = \left|a_s\right|^2 + \left|a_r\right|^2 \qquad (2.10)$$

Expanding the left hand side and using Eq. (2.9a) in a number of terms on the left and canceling we get:

$$\sum_{k=1}^{n}[S_{kr}S_{ks}^*a_ra_s^* + S_{ks}S_{kr}^*a_sa_r^*] = 0 \qquad (2.11)$$

In view of the independent nature of a_s and a_r we may choose in the first instance $a_s = a_r$, this leads to:

$$\left|a_s\right|^2\sum_{k=1}^{n}[S_{kr}S_{ks}^* + S_{ks}S_{kr}^*] = 0 \qquad (2.12a)$$

If instead we choose $a_s = ja_r$, with a_r real we obtain:

$$j\left|a_r\right|^2\sum_{k=1}^{n}[-S_{kr}S_{ks}^* + S_{ks}S_{kr}^*] = 0 \qquad (2.12b)$$

Since neither a_s nor a_r is zero, both Eqs. (2.12a) and (2.12b) are satisfied only if:

$$\sum_{k=1}^{n}S_{kr}S_{ks}^* = 0, r \neq s \qquad (2.13)$$

This equation states that the product of any column of *S*-matrix with the complex conjugate of any other column is zero. This is called the *zero property of S-matrix*.

A matrix with the elements that satisfy the constraints stated in Eqs. (2.9) and (2.13) is called the *unitary matrix* and thus a lossless microwave junction satisfies the unitary property.

Property IV: Change of Port Position: The port position is the reference plane which determines the phase of the signal from that port. A change in the port position will affect the phase of the scattering coefficients. For analysis and practical design applications, sometimes it may be desirable to shift reference planes so as to make the angles of *S*-parameters to be zero, i.e., to make them real quantities (and not complex quantity) so as to be physically realizable.

2.4 S-MATRIX OF VARIOUS MICROWAVE PASSIVE COMPONENTS

In this section we will derive the scattering matrices for some of the microwave passive components like directional coupler (including quadrature branch-line coupler and Lange coupler), Wilkinson coupler, magic tee (including rat-race), and circulator.

2.4.1 Directional Coupler

With reference to Figure 2.3, which is the symbolic representation of a directional coupler, the generalized scattering matrix of this four-port device is:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix} \tag{2.14a}$$

Using electromagnetic properties of directional coupler and the properties of *S*-parameters, the *S*-matrix of directional coupler can be exhibited in a simpler and elegant way which we are going to do now.

In a directional coupler all the four ports are matched; thus, all the ports are reflectionless and hence we can put all the diagonal elements of the *S*-matrix given in Eq. (2.14a) to be zero (i.e., $S_{ii} = 0$) when we get:

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} & S_{14} \\ S_{21} & 0 & S_{23} & S_{24} \\ S_{31} & S_{32} & 0 & S_{34} \\ S_{41} & S_{42} & S_{43} & 0 \end{bmatrix} \tag{2.14b}$$

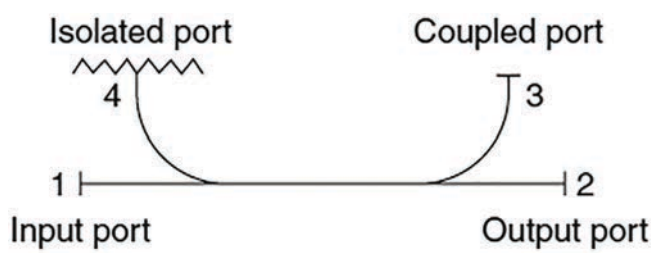


Figure 2.3 Symbolic representation of a directional coupler.

A signal fed at port 1 will not reach at port 4 and similarly a signal fed at port 2 will not reach at port 3 (for a waveguide directional coupler), thus:

$$S_{14} = S_{41} = S_{23} = S_{32} = 0 \tag{2.14c}$$

When Eq. (2.14b) can be written as:

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} & 0 \\ S_{12} & 0 & 0 & S_{24} \\ S_{13} & 0 & 0 & S_{34} \\ 0 & S_{24} & S_{34} & 0 \end{bmatrix} \tag{2.14d}$$

To arrive at Eqs. (2.14c and 2.14d) we have used the reciprocal nature of directional coupler (i.e., $S_{ij} = S_{ji}$).

Next, we apply lossless property of directional coupler at microwave frequency to simplify further the *S*-matrix given by Eq. (2.14d).

Using unity property of *S*-matrix we can write from Eq. (2.14d):

$$|S_{12}|^2 + |S_{13}|^2 = 1 \tag{2.14e}$$

$$|S_{12}|^2 + |S_{24}|^2 = 1 \tag{2.14f}$$

$$|S_{13}|^2 + |S_{34}|^2 = 1 \tag{2.14g}$$

$$|S_{24}|^2 + |S_{34}|^2 = 1 \tag{2.14h}$$

While using zero property of *S*-matrix we get from Eq. (2.14d):

$$S_{12}S_{24}^* + S_{13}S_{34}^* = 0 \tag{2.14i}$$

$$\text{Now from Eqs. (2.14e and 2.14f)}: |S_{13}| = |S_{24}| \tag{2.14j}$$

$$\text{While from Eqs. (2.14f and 2.14h)}: |S_{12}| = |S_{34}| \tag{2.14k}$$

We may choose the reference planes in respective ports (i.e., 1 with respect to (w.r.t.) 2 and 3 w.r.t. 4) in such a way that S_{12} and S_{34} become real quantities. Hence, we may write:

$$S_{12} = S_{34} = \alpha \tag{2.14l}$$

where α is a real quantity and physically α^2 represents the power to the through port, i.e., the output port and hence the transmission factor.

Applying the result of Eq. (2.14l) in Eq. (2.14i) we get:

$$\alpha(S_{24}^* + S_{13}) = 0$$

or, $S_{24}^* + S_{13} = 0$ …… (2.14m) (since $\alpha \neq 0$)

Thus, $S_{13} = -S_{24}^* = -j\beta$ (2.14n)

where β is a real quantity and physically β^2 represents the coupling factor. The negative sign preceding β indicates that the signal flow from 1 to 3 or from 2 to 4 basically represents the attenuation of the signal (coupling a fraction of the input signal to the coupled port).

Substituting Eqs. (2.14l and 2.14n) into Eq. (2.14d) we have:

$$[S] = \begin{bmatrix} 0 & \alpha & j\beta & 0 \\ \alpha & 0 & 0 & j\beta \\ j\beta & 0 & 0 & \alpha \\ 0 & j\beta & \alpha & 0 \end{bmatrix} \quad (2.14o)$$

From Eqs. (2.14l and 2.14n) and any of the Eqs. (2.14e to 2.14h) it can be seen that:

$$\alpha^2 + \beta^2 = 1 \quad (2.14p)$$

This is also justified from the conservation of power. With simple calculation, for a 30 dB directional coupler the coupling factor $\beta^2 = 0.001$ and transmission factor $\alpha^2 = 0.9994$.

There is an alternative way to present the *S*-matrix of directional coupler. Since the transmission factor α^2 and the coupling factor β^2 represent the average power of the electromagnetic signal while the *S*-parameters are expressed in equivalent voltages we can replace α and β by $(1/\sqrt{2})$ in Eq. (2.14o) when we have:

$$[S] = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & j & 0 \\ 1 & 0 & 0 & j \\ j & 0 & 0 & 1 \\ 0 & j & 1 & 0 \end{bmatrix} \quad (2.14q)$$

It may further be noted that in some texts ports 2 and 4 are interchanged when compared with the port numbering used in this text, and in that case the *S*-matrix of directional coupler comes out as:

$$[S] = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & j & 1 & 0 \\ j & 0 & 0 & 1 \\ 1 & 0 & 0 & j \\ 0 & 1 & j & 0 \end{bmatrix} \tag{2.14r}$$

The directional coupler we have been discussing so far is a 90^0 hybrid and is similar to the quadrature (i.e., 90^0) branch-line coupler/hybrid and also the Lange coupler. All these couplers are known as symmetrical couplers. The *S*-matrix for all these types of quadrature coupler/hybrid will be represented by Eqn. (2.14q) or Eq. (2.14r) depending on the port numbering as mentioned above. But for other hybrid structures like magic tee and rat-race which are 180^0 hybrid (also known as anti-symmetrical coupler) the *S*-matrix is different from above and we are going to derive it next.

The Wilkinson coupler (used as power divider/combiner) is a 3-port RF device whose *S*-matrix can directly be written from its physical principle of operation and is given by:

$$[S] = \frac{j}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{2.14s}$$

2.4.2 *Magic Tee*

With reference to Fig. 2.4, which is the symbolic representation of a magic tee, we find that it is also a four-port device and hence its general *S*-matrix is similar to that given in Eq. (2.14a).

We can apply the electromagnetic properties of magic tee and the properties of *S*-parameters to get the *S*-matrix of magic tee in a simple and elegant form.

If ports 1 and 2 are identically terminated then a signal at port 3 reaches port 1 and 2 with equal magnitude and at same phase, and hence:

$$S_{13} = S_{23} \tag{2.15a}$$

Also, a signal fed at port 4 will be coupled to port 1 and 2 with equal magnitude but with opposite phase, thus:

$$S_{14} = -S_{24} \tag{2.15b}$$

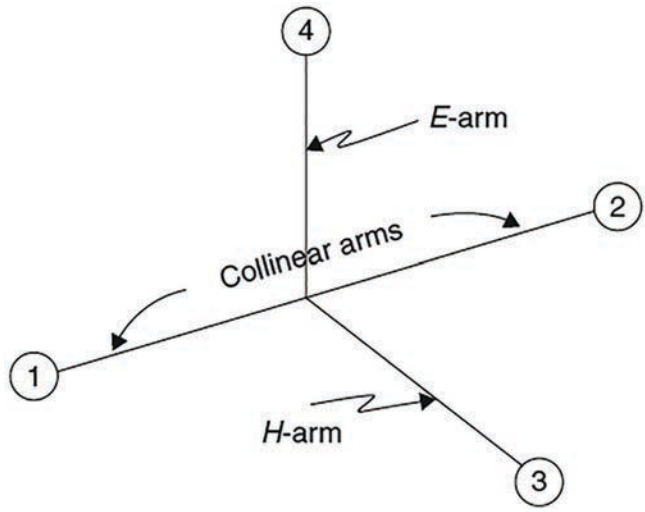


Figure 2.4 Symbolic representation of a magic tee.

The magic tee is reciprocal in the sense that two signals of equal magnitude and of same phase fed at the collinear ports 1 and 2 will add and couple the signal to port 3, also two signals of equal magnitude and opposite phase fed at ports 1 and 2 will add and couple the signal to port 4. So, for reciprocal junction we can apply the symmetry property of *S*-matrix, i.e., $S_{ij} = S_{ji}$, $i \neq j$, thus:

$$S_{21} = S_{12} \quad S_{32} = S_{23}$$

$$S_{31} = S_{13} \quad S_{42} = S_{24} \tag{2.15c}$$

$$S_{41} = S_{14} \quad S_{43} = S_{34}$$

Further, a signal fed at port 3 will not be coupled to port 4 and vice versa (as these ports are de-coupled ports), thus:

$$S_{34} = S_{43} = 0 \tag{2.15d}$$

Substituting the results obtained in Eqs. 2.15 (a to d) into Eq. (2.14a) we get:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{12} & S_{22} & S_{13} & -S_{14} \\ S_{13} & S_{13} & S_{33} & 0 \\ S_{14} & -S_{14} & 0 & S_{44} \end{bmatrix} \tag{2.15e}$$

Matching elements that do not destroy the symmetry of the junction may be placed in the E-plane arm and the H-plane arm so as to make:

$$S_{44} = S_{33} = 0 \tag{2.15f}$$

Using this, *S*-matrix becomes:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{12} & S_{22} & S_{13} & -S_{14} \\ S_{13} & S_{13} & 0 & 0 \\ S_{14} & -S_{14} & 0 & 0 \end{bmatrix} \tag{2.15g}$$

Now we use the unitary property of *S*-matrix (i.e., assume that the structure is lossless). Putting the sum of the products of each term of the first column with its conjugate equal to unity (i.e., applying the unity property) we get:

$$|S_{11}|^2 + |S_{12}|^2 + |S_{13}|^2 + |S_{14}|^2 = 1 \tag{2.15h}$$

And similarly, the expression from the second column of Eq. (2.15g) is:

$$|S_{12}|^2 + |S_{22}|^2 + |S_{13}|^2 + |S_{14}|^2 = 1 \tag{2.15i}$$

Subtracting (2.15i) from (2.15h) we get:

$$|S_{11}|^2 - |S_{22}|^2 = 0 \quad \text{or} \quad |S_{11}| = |S_{22}| \tag{2.15j}$$

Again, from columns 3 and 4 of Eq. (2.15g) using unity property:

$$2|S_{13}|^2 = 1 \text{ or } |S_{13}| = 1/\sqrt{2}, \quad \text{and} \quad 2|S_{14}|^2 = 1 \text{ or } |S_{14}| = 1/\sqrt{2} \tag{2.15k}$$

Using this relation of Eq. (2.15k) in Eq. (2.15h):

$$|S_{11}|^2 + |S_{12}|^2 + (1/2) + (1/2) = 1 \quad \text{or} \quad |S_{11}|^2 + |S_{12}|^2 = 0 \tag{2.15l}$$

This sum can equal zero only when both S_{11} and S_{12} are individually zero, i.e.:

$$S_{11} = S_{12} = 0 \tag{2.15m}$$

Thus, from Eq. (2.15j):

$$S_{22} = 0 \tag{2.15n}$$

Substituting the results from Eqs. (2.15m) and (2.15n) into Eq. (2.15g) we have:

$$[S] = \begin{bmatrix} 0 & 0 & S_{13} & S_{14} \\ 0 & 0 & S_{13} & -S_{14} \\ S_{13} & S_{13} & 0 & 0 \\ S_{14} & -S_{14} & 0 & 0 \end{bmatrix} \tag{2.15o}$$

By proper choice of terminal planes in the E- and H-arms, we can make both S_{13} and S_{14} as real quantities. Thus, the S-matrix of a magic tee may be exhibited in the form:

$$[S] = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} \tag{2.15p}$$

2.4.3 *Circulator*

Let us consider a three-port circulator (see symbolic representation in Figure 2.5).

The generalized *S*-matrix for a three-port device is:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \tag{2.16a}$$

If all the three ports are perfectly matched, i.e., there is no reflection at the respective ports, then:

$$S_{11} = S_{22} = S_{33} = 0 \tag{2.16b}$$

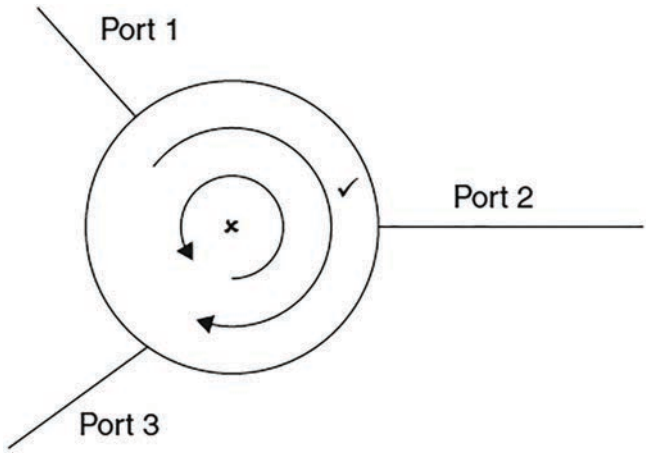


Figure 2.5 Symbolic representation of a 3-port circulator.

The device is nonreciprocal, thus:

$$S_{12} = 0,\ S_{23} = 0,\ S_{31} = 0 \tag{2.16c}$$

$$\text{Also:}\quad S_{12} \neq S_{21},\ S_{13} \neq S_{31},\ \text{and}\ S_{23} \neq S_{32} \tag{2.16d}$$

Thus, applying Eqs. (2.16b), (2.16c), and (2.16d) into Eq. (2.16a) we get:

$$[S] = \begin{bmatrix} 0 & 0 & S_{13} \\ S_{21} & 0 & 0 \\ 0 & S_{32} & 0 \end{bmatrix} \tag{2.16e}$$

Assuming the circulator to be lossless we get from unity property:

$$\left|S_{21}\right|^2 = \left|S_{32}\right|^2 = \left|S_{13}\right|^2 = 1 \tag{2.16f}$$

Thus, from Eq. (2.16e) we have the *S*-matrix for a perfect three-port circulator to be:

$$[S] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{2.16g}$$

2.5 SIGNAL FLOW GRAPHS (SFGs)

SFGs are very useful for the analysis of microwave networks in terms of incident wave (independent) variable *a* and the reflected wave (dependent) variable *b*. This graphical representation implicitly indicates the energy or power flow in the microwave network in terms of *S*-parameters. In fact, the SFGs, through the use of graphical diagrams, provide a physical insight into the relationship between the cause (incident wave variable *a*) and the effect (reflected wave variable *b*) in the system. In essence, an SFG is an alternative and yet a simpler method to the well-known "block diagram" method in representing a complicated linear system. The advantage of signal flow diagram method over the block diagram method is the availability of a flow-graph gain formula, which provides the relation between the system variables.

In SFG the wave variables are shown as nodes and *S*-parameters as branches connecting the nodes, the directions of branches are from "*a*" node to "*b*" node (see Figure 2.6), a two-port network. The nodes are shown by small circles and the branches have arrow on the line indicating the direction of signal flow and the letter near the arrow indicates the multiplication factor.

The SFG gives an intuitive graphical illustration of the network behavior. For example, a wave incident at port 1 at node a_1 will split, with a part going as reflected wave (S_{11}) generated at port 1 and as a transmitted wave (S_{21}) from port 1 to port 2. Again, at node b_2, a wave goes out through port 2 and if a load with nonzero reflection coefficient is connected at that port, then a reflected wave (S_{22}) will be generated at port 2 and a part of the signal from node a_2 will be transmitted (S_{12}) from port 2 to port 1.

Once a microwave network has been represented with SFG, the evaluation of the ratio of any combination of wave amplitudes at the input-output of the network can be done in two ways. One can use the *decomposition rules* or alternatively one may use the *Mason's rule* of control theory. In the following sections we discuss the decomposition rules (without going into

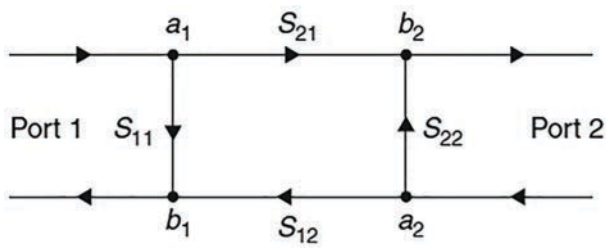


Figure 2.6 Signal flow graph of a two-port network.

mathematical derivations) and also the Mason's rule in a nutshell. However, it may be noted that the use of decomposition rule is much easier than the Mason's rule as for complicated graphs the use of Mason's rule may be quite intricate.

2.5.1 Decomposition Rules

Rule I (*Series or Multiplication rule*): This rule states that two series branches (whose common node 2 has only one incoming and one outgoing wave) may be combined into a single branch from node 1 to node 3 with effective transmission coefficient (multiplication factor) equal to the product of the transmission coefficient from node 1 to node 2 with that from node 2 to node 3 as shown in Figure 2.7a; mathematically: $S_{31} = S_{21}S_{32}$.

Rule II (*Shunt or Addition rule*): This rule states that two or more parallel connecting branches (going from one node to another) may be replaced by a single branch with effective transmission coefficient equivalent to the sum of the transmission coefficients of the individual branches as shown in Figure 2.7b; mathematically: $S_{\text{eff}} = S_1 + S_2$.

Rule III (*Loop rule*): (*i*) *Self-loop*: When a node has a self-loop (a branch that begins and ends on the same node, see Figure 2.7ci) of coefficient S, the self-loop can be eliminated by multiplying the transmission coefficients of the branches feeding that node by $1/(1 - S)$, thus: $S_{31} = S_{21}S_{32}/(1 - S_{22})$.

ii. *Feedback-loop*: A feedback loop (see Figure 2.7c*ii*) may be eliminated by dividing the input transmission coefficient by 1 minus the transmission factor around the loop which is the product $S_{23}S_{32}$ (by Rule I), thus: $S_{31} = S_{21}/(1 - S_{23}\,S_{32})$.

Rule IV (*Splitting rule*): Any node can be split into two separate nodes where each of the two nodes are connected only once to the incoming and outgoing nodes with the same coefficient, as shown in Fig. 2.7d.

2.5.2 Mason's Rule

It states that the transfer function, i.e., the ratio of a dependent to independent variable is given by [3]:

$$T = \frac{P_1\left[1 - \sum L(1)^{(1)} + \sum L(2)^{(1)} - \ldots.\right] + P_2\left[1 - \sum L(1)^{(2)} + \ldots..\right] + P_3\left[1 - \ldots.\right]}{1 - \sum L(1) + \sum L(2) - \sum L(3) + \ldots.} \tag{2.17}$$

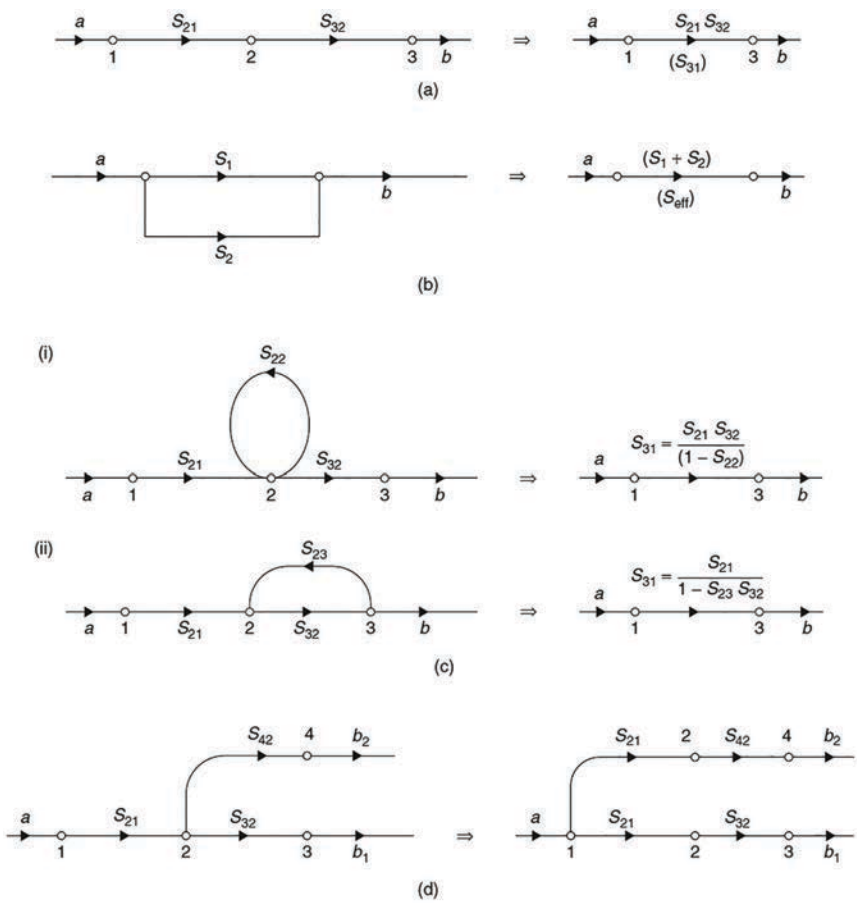


Figure 2.7 Decomposition rule: (a) Series rule, (b) Shunt or addition rule, (c) Loop rule: (i) self-rule, and (ii) Feedback loop, and (d) Splitting rule.

Where: P_1, P_2, P_3, …. etc., are the values of the various paths connecting the dependent and independent variables whose transfer function is to be determined. A path is a branch starting at an initial node and following the directed arrow of the branches to the final node in question such that no node is touched more than once. The value of the path is the product of all the branch values (coefficients) covered along the particular path. It is to be remembered while traversing along a path no node is touched more than once.

$\Sigma L(1)$, $\Sigma L(2)$, …. etc., are the sum of all *first order loop* products, *second order loop* products, …. respectively.

$\Sigma L(1)^{(1)}$, $\Sigma L(2)^{(1)}$, …. etc., are respectively the sum of all first order loop products, second order loop products, ….., that do not touch the *first* path between the nodes representing the two variables.

$\Sigma L(1)^{(2)}$, $\Sigma L(2)^{(2)}$, …. etc., are respectively the sum of all first order loop products, second order loop products, ….., that do not touch the *second* path between the nodes representing the two variables.

It is to be noted that a first order loop is defined as the product of all branches encountered in a round trip as we move from a node in the direction of arrow of the branch or branches back to the initial node. While a second order loop is defined as the product of any two non-touching first order loops. And a third order loop is product of any three non-touching first order loops.

2.6 APPLICATION OF SIGNAL FLOW GRAPH—RF/ MICROWAVE AMPLIFIER GAIN FORMULATION IN TERMS OF *S*-PARAMETERS

In Fig. 2.8 we have shown a two-port network that represents a microwave amplifier with load Z_L and signal source V_S having source impedance Z_S. With the help of SFG, we will derive (a) input reflection coefficient (Γ_{IN}), and (b) output reflection coefficient (Γ_{OUT}) in terms of *S*-parameters. Finally, we will derive the expressions for different gains of RF/microwave amplifier like the transducer gain, available power gain and operating power gain in terms of *S*-parameters.

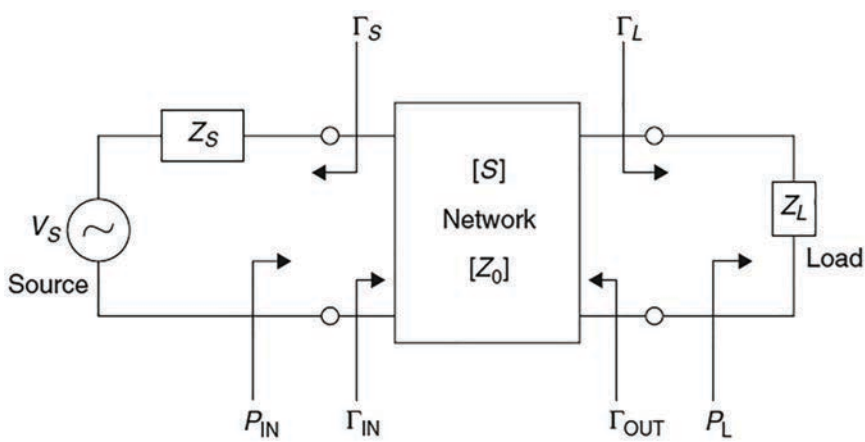


Figure 2.8 A representative microwave amplifier with source and load.

(a) To determine the SFG we first dissect the amplifier network block diagram of Fig. 2.8 into three parts: (*i*) the signal generator section, (*ii*) two port network section, and (*iii*) the load section.

(i) The schematic of signal generator section together with its SFG is shown in Fig. 2.9a, for which:

$$b_S = \frac{V_S Z_0}{(Z_S + Z_0)} \cdot \frac{1}{\sqrt{Z_0}} = \frac{V_S \sqrt{Z_0}}{(Z_S + Z_0)} \quad \text{and} \quad \Gamma_S = \frac{Z_S - Z_0}{Z_S + Z_0}$$

(ii) The scheme for two port network is reproduced from Fig. 2.8 in Fig 2.9b.

(iii) The scheme for the load section is shown in Fig. 2.9c together with its SFG, for which:

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}$$

On combining the SFG of three sections mentioned above, we have the SFG for the two-port amplifier with the associated signal generator and load as depicted in Fig. 2.9d.

(b) To obtain the input reflection coefficient $\Gamma_{IN} (= b_1/a_1)$ or output reflection coefficient $\Gamma_{OUT} (= b_2/a_2)$, we may use either the decomposition rule (method I) or the Mason's rule (method II).

Method I: To find the input reflection coefficient $\Gamma_{IN} = b_1/a_1$, we redraw the necessary part of the SFG of Fig. 2.9d in Fig. 2.9e. Using decomposition rule in steps (see Fig. 2.9e(i -iv) we can derive the expression for Γ_{IN}, given by:

$$\Gamma_{IN} = S_{11} + \frac{\Gamma_L S_{12} S_{21}}{1 - S_{22}\Gamma_L} = \frac{S_{11} - \Delta\Gamma_L}{1 - S_{22}\Gamma_L}$$

Where $\Delta = S_{11}S_{22} - S_{12}S_{21}$.

Similarly, we can also obtain $\Gamma_{OUT} = b_2/a_2$, using the SFG shown in Fig. 2.9f and using decomposition rule. The expression for Γ_{OUT} is given by:

$$\Gamma_{OUT} = S_{22} + \frac{S_{12} S_{21} \Gamma_S}{1 - S_{11}\Gamma_S} = \frac{S_{22} - \Delta\Gamma_S}{1 - S_{11}\Gamma_S}$$

Method II: To find Γ_{IN} we again refer to Fig 2.9e and applying Mason's rule we note: $P_1 = S_{11}$, $P_2 = S_{21}\Gamma_L S_{12}$, $\Sigma L(1) = S_{22}\,\Gamma_L$, and $\Sigma L(1)^{(1)} = S_{22}\,\Gamma_L$. Now using Eq. (2.17) we have:

$$\Gamma_{IN} = \frac{S_{11}(1 - S_{22}\Gamma_L) + S_{21}\Gamma_L S_{12}}{1 - S_{22}\Gamma_L} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}$$

Similarly, Γ_{OUT} may also be found out.

Derivation of Expressions for the Gains of the Amplifier

To derive expressions for different gains (viz. *transducer gain*, *available power gain* and *operating power gain*) of the amplifier in terms of the *S*-parameters of the network (S_{11}, S_{22}, S_{21}, S_{12}) and the source/load reflection coefficients (Γ_S/Γ_L) and input/output reflection coefficients (Γ_{IN}/Γ_{OUT}) we reproduce the SFG of Fig. 2.9d in Fig. 2.9g. To do this we need to define different power quantities in the amplifier system.

The *input power* (P_{IN}) to the network is given by:

$$P_{IN} = \frac{1}{2}|a_1|^2 - \frac{1}{2}|b_1|^2 = \frac{1}{2}|a_1|^2(1 - |\Gamma_{IN}|^2), \quad \text{since} \quad |\Gamma_{IN}| = \frac{|b_1|}{|a_1|}.$$

The *available power from the source* (P_{AVS}), defined as the input power delivered by the source to conjugately matched input impedance, is given by:

$$P_{AVS} = P_{IN}\big|_{\Gamma_{in} = \Gamma_S^*} = \frac{1}{2}|b_g|^2 - \frac{1}{2}|a_g|^2 = \frac{1}{2}|b_g|^2(1 - |\Gamma_S|^2) = \frac{b_S}{2\{1 - |\Gamma_S|^2\}}$$

Since, $a_g = \Gamma_S^* b_g$ and $b_g = b_S + \Gamma_S a_g = \dfrac{b_S}{1 - |\Gamma_S|^2}$

Power delivered to the load (P_L) is given by:

$$P_L = \frac{1}{2}|b_2|^2 - \frac{1}{2}|a_2|^2 = \frac{1}{2}|b_2|^2(1 - |\Gamma_L|^2), \quad \text{since} \quad |\Gamma_L| = \frac{|b_2|}{|a_2|}$$

The *power available from the network* (P_{AVN}), defined as the power delivered to the load when the load is conjugate matched to the network, is given by:

$$P_{AVN} = P_L\big|_{\Gamma_L = \Gamma_{OUT}^*} = \left[\frac{1}{2}|b_2|^2 - \frac{1}{2}|a_1|^2\right]_{\Gamma_L = \Gamma_{OUT}^*} = \frac{1}{2}|b_2|^2(1 - |\Gamma_{OUT}|^2)$$

Now, the *Transducer Power Gain* (G_T) is defined as the ratio of P_L to P_{AVS}, when:

$$G_T = \frac{|b_2|^2}{|b_S|^2}(1 - |\Gamma_L|^2)(1 - |\Gamma_S|^2)$$

The value of b_2/b_S may be determined from Mason's rule referring to Fig. 2.9g. It may be seen that there are three first order loops having values: $S_{11}\Gamma_S$, $S_{22}\Gamma_L$, and $S_{21}\Gamma_L S_{12}\Gamma_S$ while only one second order loop having value $S_{11}\Gamma_S S_{22}\Gamma_L$. Thus, using Mason's rule we can write:

$$\frac{b_2}{b_S} = \frac{S_{21}}{1 - S_{11}\Gamma_S - S_{22}\Gamma_L - S_{21}\Gamma_L S_{12}\Gamma_S + S_{11}\Gamma_S S_{22}\Gamma_L}$$

Substituting this value of b_2/b_S in the expression for G_T we get:

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

This may also be expressed as:

$$G_T = \frac{1 - |\Gamma_S|^2}{|1 - \Gamma_{IN}\Gamma_S|^2} |S_{21}|^2 \frac{1 - |\Gamma_L|^2}{|1 - S_{22}\Gamma_L|^2}$$

Or, we may write it in yet an alternative way:

$$G_T = \frac{1 - |\Gamma_S|^2}{|1 - S_{11}\Gamma_S|^2} |S_{21}|^2 \frac{1 - |\Gamma_L|^2}{|1 - \Gamma_{OUT}\Gamma_L|^2}$$

Next, the *Available Power Gain* (G_{AV}) is defined as the ratio of P_{AVN} to P_{AVS}, when:

$$G_{AV} = \frac{|b_2|^2}{|b_S|^2}(1 - |\Gamma_S|^2)(1 - |\Gamma_{OUT}|^2)$$

The expression for b_2/b_s with $\Gamma_L = \Gamma^*_{OUT}$ becomes after some manipulation as:

$$\frac{b_2}{b_S} = \frac{S_{21}}{(1 - S_{11}\Gamma_S)(1 - |\Gamma_{OUT}|^2)}$$

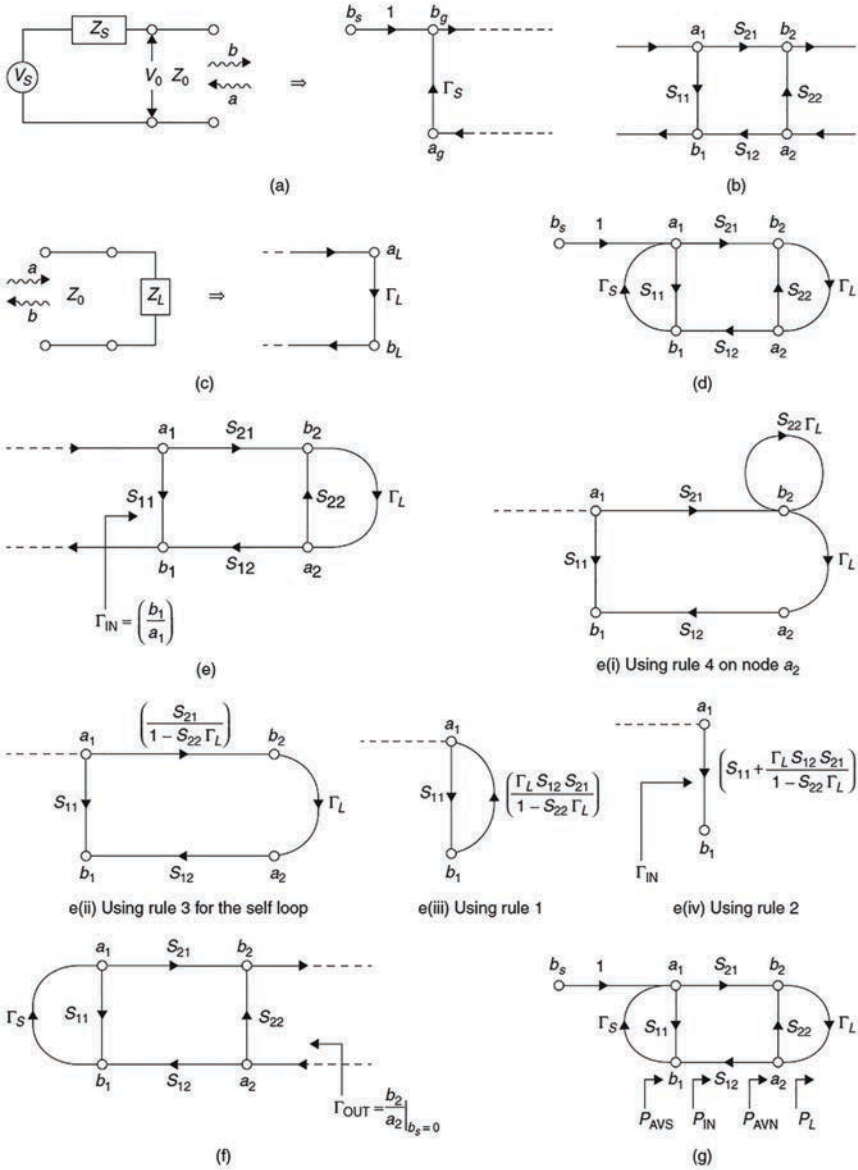


Figure 2.9 (a–g) represent the different signal flow graphs for RF/microwave amplifier gain formulation in terms of *S*-parameters.

When G_{AV} can be expressed as:

$$G_{AV} = \frac{1-\left|\Gamma_S\right|^2}{\left|1-S_{11}\Gamma_S\right|^2}\left|S_{21}\right|^2 \frac{1}{1-\left|\Gamma_{OUT}\right|^2}$$

Finally, the *Operating Power Gain* or simply *Power gain* (G_P) is defined as the ratio of P_L to P_{IN}, when:

$$G_P = \frac{\left|b_2 / b_S\right|^2}{\left|a_1 / b_S\right|^2}\left\{\frac{1-\left|\Gamma_L\right|^2}{1-\left|\Gamma_{IN}\right|^2}\right\}$$

Using Mason's rule, it can be obtained that:

$$\frac{a_1}{b_S} = \frac{1-S_{22}\Gamma_L}{1-S_{11}\Gamma_S - S_{22}\Gamma_L - S_{21}\Gamma_L S_{12}\Gamma_S + S_{11}\Gamma_S S_{22}\Gamma_L}$$

Thus, G_P can be written as:

$$G_P = \frac{1}{1-\left|\Gamma_{IN}\right|^2}\left|S_{21}\right|^2 \frac{1-\left|\Gamma_L\right|^2}{\left|1-S_{22}\Gamma_L\right|^2}$$

REFERENCES

[1] Pozar D M, *Microwave Engineering*, Fourth Edition, Wiley, New York, 2012.
[2] Collin R E, *Foundations for Microwave Engineering*, Second Edition, John Wiley & Sons, New Jersey, 2007.
[3] Rizzi P A, *Microwave Engineering Passive Circuits*, Prentice-Hall of India Pvt. Ltd., New Delhi, 1999.

Chapter 3

Fundamentals of Smith Chart and Stub Matching Design with Smith Chart

A disciplined life always follows a routine-chart for everyday action

3.1 INTRODUCTION

Radio frequency (RF) and microwave design problems may, in general, be solved by evaluating analytical equations which often involve tedious manipulation of complex numbers. To avoid this, a number of well-structured graphical techniques were developed to solve transmission line impedance matching and other problems of which the best known and most widely used one is the *Smith chart*, developed by Philips H. Smith of Bell Telephone Laboratories, US, in 1939. Essentially the chart is the complex impedance plane mapped onto complex reflection coefficient plane. Thus, the Smith chart is a graphical plot of normalized resistance and reactance functions in the complex reflection coefficient plane, which provides an intuitive understanding of how individual variables affect the end result. Instead of having a separate Smith chart for transmission lines with different characteristic impedances, it is preferable to have just one (where all impedances are normalized with respect to the characteristic impedance Z_0) that can be used for any line. In modern microwave engineering, this impedance chart has become indispensable in transmission line problem solution and has become an integral part of most of the current computer-aided design software and test equipment. As a matter of fact, most vector network analyzers (useful for RF/microwave circuit characterization; the fundamentals of which are given in Appendix VII) have the Smith chart display as one of the primary screen options (see Figure 3.1). Smith chart has many applications in solving RF/microwave circuit problems including the solution of stub matching problems, amplifier design, and impedance measurement at RF/microwave frequency.

3.2 DEVELOPMENT OF SMITH CHART

The development of Smith chart begins with the fundamental equation for reflection coefficient, i.e.:

DOI: 10.1201/9781003729969-3



Figure 3.1 Vector network analyzer (VNA) depicting the impedance locus of a microwave circuit on Smith chart (Photograph: Courtesy SAMEER Kolkata Centre, India).

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{3.1}$$

This may be solved for Z_L, which when normalized with respect to the characteristic impedance Z_0 yields:

$$z_L = \frac{Z_L}{Z_0} = \frac{1+\Gamma}{1-\Gamma} \tag{3.2}$$

Putting $z_L = r + jx$ and $\Gamma = u + jv$, we get from Eq. (3.2):

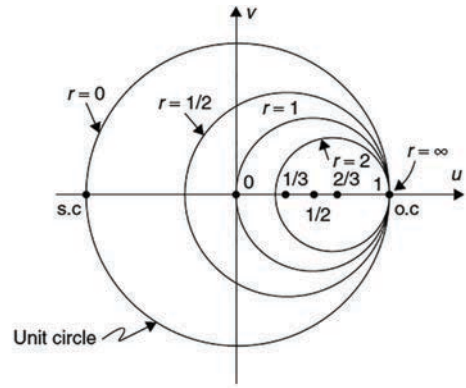
$$r + jx = \frac{1+u+jv}{1-u-jv} \tag{3.3}$$

Rationalizing the right-hand side of Eq. (3.3) and equating the real and imaginary parts on both sides we have:

$$r = \frac{(1-u)^2 - v^2}{(1-u)^2 + v^2} \tag{3.4a}$$

r	Centre at		Radius
	u	v	
0	0	0	1
1/2	1/3	0	2/3
1	1/2	0	1/2
2	2/3	0	1/3
∞	1	0	0

(a)



(b)

Figure 3.2 Mapping of constant *r*-circles in the complex Γ-plane: (a) Center and radius values for different values of *r*, and (b) *r*-circles in complex ($u - v$), i.e., Γ-plane.

$$x = \frac{2v}{(1-u)^2 + v^2} \tag{3.4b}$$

These equations can be rearranged so that by completing squares we obtain:

$$\left(u - \frac{r}{1+r}\right)^2 + v^2 = \frac{1}{(1+r)^2} \tag{3.5a}$$

$$(u-1)^2 + \left(v - \frac{1}{x}\right)^2 = \frac{1}{x^2} \tag{3.5b}$$

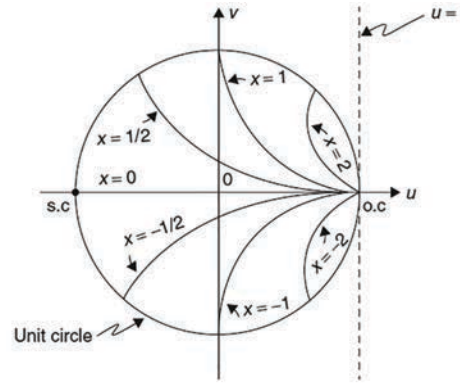
These two equations generate r and x circles on the complex ($u - v$) plane. The r-circles have radii equal to $1/(1 + r)$ and center at $[u = r/(1 + r), v = 0]$ while the x-circles have radii equal to $1/|x|$ and center at $[u = 1, v = 1/x]$. Since x is normalized reactance, x can have either positive (+ve) or negative (–ve) values. Figures 3.2(a) and 3.3(a), respectively, show tables for several values of r and x, while Figure 3.2b and Figure 3.3b show corresponding r and x circles in the complex ($u - v$) plane, i.e., in Γ plane.

We may now make a few observations.

i. The center of all r-circles lies on the u-axis, the $r = 0$ circle (having a radius of 1 and centered at the origin) is the largest. The r-circles become progressively smaller as r increases from 0 (i.e., short-circuit

x	Centre at		Radius
	u	v	
0	1	∞	∞
0.5	1	2	2
1	1	1	1
2	1	1/2	1/2
∞	1	0	0

(a)



(b)

Figure 3.3 Mapping of constant *x*-circles in the complex Γ-plane: (a) Center and radius values for different values of *x*, and (b) *x*-circles in complex ($u - v$), i.e., Γ-plane.

[s.c.]) toward ∞, ending at the ($u = 1$ and $v = 0$) point for open circuit (o.c). All *r*-circles pass through this point marked as o.c.

ii. The center of all *x*-circles lies on the $u = 1$ line; those for $x > 0$ (inductive reactance) lie above the *u*-axis, and those for $x < 0$ (capacitive reactance) lie below the *u*-axis; the $x = 0$ circle is along the *u*-axis. The *x*-circles become progressively smaller as $|x|$ increases from 0 toward ∞, ending at the ($u = 1$ and $v = 0$) point for open-circuit, and all *x*-circles pass through this o.c. point.

iii. $r = 0$ means pure reactive load (when $z_L = \pm jx$), i.e., complete reflection ($\Gamma = 1$). Thus, the Smith chart is bounded by $r = 0$ circle.

iv. The point $r = 1$, $x = 0$ (when $Z_L = Z_0$, i.e., perfectly matched condition) corresponds to $u = 0$ and $v = 0$, i.e., the origin of the Γ plane, i.e., $|\Gamma| = 0$.

Essentially a Smith chart is thus a chart of *r*- and *x*-circles in the complex Γ-plane for $|\Gamma| \leq 1$ as shown in Figure 3.4. The *r*- and *x*-circles are everywhere orthogonal to each other.

The intersection of an *r*-circle and *x*-circle defines a point *P* that represents a normalized load impedance $z_L = r \pm jx$, which when multiplied by the characteristic impedance of the line give the actual load impedance. For all practical purposes, a Smith chart is marked with polar coordinates, such that every point in the Γ-plane is specified by a magnitude $|\Gamma|$ and a phase angle θ_Γ. If we draw a circle centered at the origin through the load point *P* (which is a constant Γ-circle), the fractional distance (compared to unity radius to the edge of the chart) from the centre to the marked point *P* is equal to the magnitude of reflection coefficient (i.e., $|\Gamma|$) and the angle that

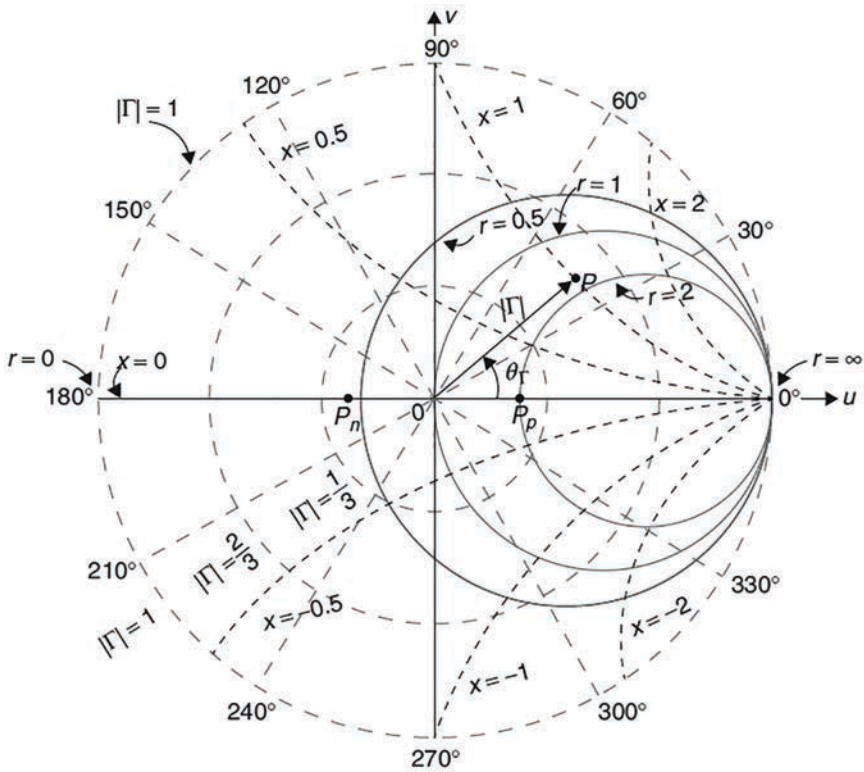


Figure 3.4 Smith chart in polar coordinate with *r*- and *x*-circles shown together in the Γ-plane.

the line to the point makes with the real axis is the phase of the reflection coefficient (i.e., θ_Γ).

Each $|\Gamma|$-circle intersects the real axis at two points, the point P_p on the positive real axis and the point P_n on the negative real axis, as shown in Figure 3.4. Since $x = 0$ along the real axis, P_p and P_n both represent situations with a purely resistive load, $Z_L = R_L$. Thus we observe that, $R_L > R_0$ at P_p, where $r > 1$; and $R_L < R_0$ at P_n, where $r < 1$. We know that for purely resistive load, when $R_L > R_0$, the voltage standing wave ratio (VSWR) $\rho = R_L/R_0$, which also happens to be equal to r. This implies that the value of the r-circle passing through the point P_p is numerically equal to the standing-wave ratio, ρ. Similarly, the value of the r-circle passing through the point P_n on the negative real axis is numerically equal to $1/\rho$.

So far, we have discussed the development of Smith chart based on defining the reflection coefficient at the load ($z = 0$). However, it may be needed (and

actually we need in the solution of many transmission line problems) to move either toward generator or toward load from a generalized point $z \neq 0$, say, at $z = l$ on the transmission line. In such situations it may be recalled that the generalized reflection coefficient is given by $\Gamma_l = |\Gamma| e^{j(\theta - 2\beta l)}$, where z is +ve, i.e., we move toward the generator as per our convention. This indicates that the magnitude of the reflection coefficient at any location on the line remains same as that at the load; only the phase at the load, θ_Γ, is decreased in this situation by $2\beta l$, which corresponds to a rotation in a clockwise direction on the Smith chart. If z is –ve, i.e., if we move from a generalized point $z \neq 0$ on the line toward load, then θ_Γ will be increased by $2\beta l$, and we need to move in an anticlockwise direction on the Smith chart to account for this phase increase of the reflection coefficient. Thus, in commercial Smith chart the perimeter of the Smith chart has three concentric scales. The innermost scale is labeled with angle of reflection coefficient in degrees and two more scales are also there. The outermost scale around the perimeter of the Smith chart is called the *wavelengths toward generator* (WTG), which denote movements on the transmission line toward the generator (and this movement is clockwise) while a third scale (in between θ_Γ and the WTG scale) represents *wavelengths toward load* (WTL), which denote movements on the transmission line toward load (and this movement is anticlockwise). Noting that a complete rotation around the Smith chart is equal to a phase change of 2π, the length l of transmission line corresponding to this phase rotation of 2π would be equal to $\lambda/2$. The WTG and WTL scales are expressed in terms of the length of transmission line (l) normalized with respect to wavelength (λ) and thus the scales start from 0 and end at 0.5 for a complete phase rotation of 2π, as shown in Figure. 3.5.

3.3 SMITH AND JONES CHART

In solving certain types of transmission line problems (for example in stub matching), it is often more convenient to work with admittances rather than with impedances. The Smith chart can be used for normalized admittance ($y_L = g \pm jb$) in the same way that it is used for normalized impedance ($z_L = r \pm jx$), and can also be used to convert between impedance and admittance. It may be observed that, in the normalized form, the input impedance of a line connected to a load via a $\lambda/4$ line is given by: $z_{in} = 1/z_L$ (which follows from the concept of quarter-wave impedance transformation), that has the effect of converting normalized impedance to normalized admittance. Since a complete revolution around the Smith chart corresponds to a length of $\lambda/2$, a $\lambda/4$ transformation is equivalent to rotating the chart by 180^0; this effectively means we can image a given impedance (or admittance) point across the center of the chart to obtain the corresponding admittance (or impedance) point. Thus, the Smith chart can be used either

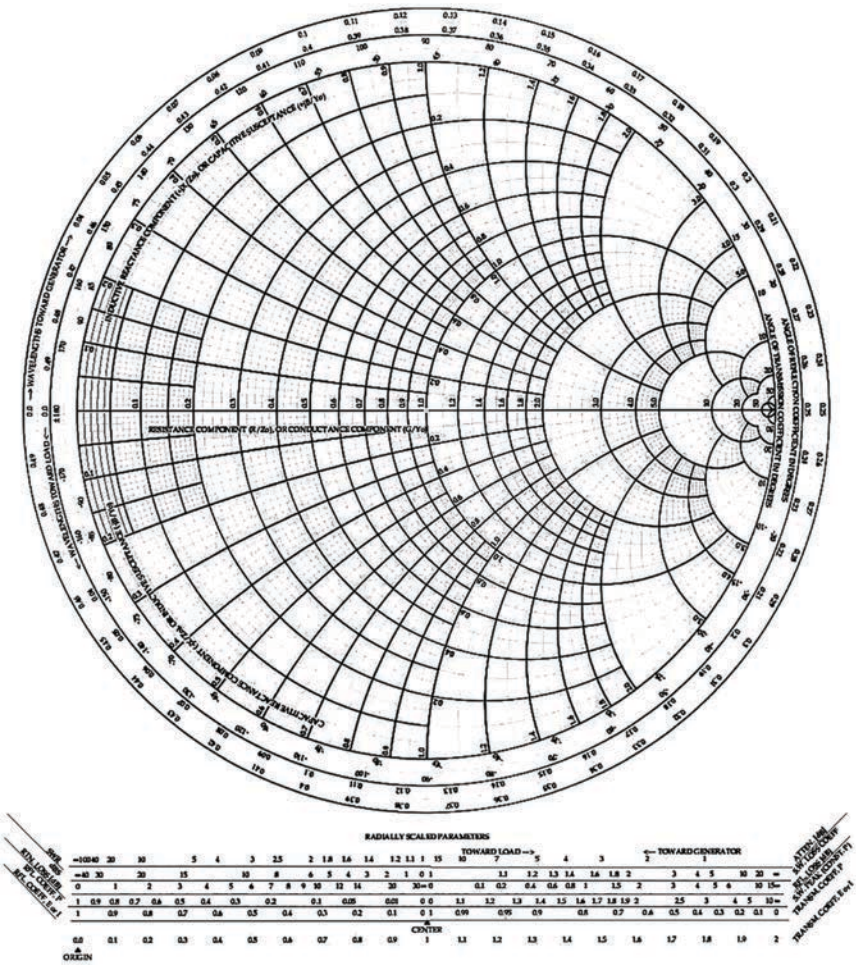


Figure 3.5 A typical Smith chart (impedance chart) showing wavelength toward generator (WTG) and wavelength toward load (WTL) and radially scaled parameters.

as an impedance or as an admittance chart. As an impedance chart, the Smith chart consists of r_L and x_L circles, representing the normalized resistance and reactance of a normalized load impedance z_L. When used as an admittance chart, the r_L circles become g_L circles and x_L circles become the b_L circles, where g_L and b_L are, respectively, the normalized conductance and susceptance of the normalized load admittance y_L. However, in solution of some typical transmission line problems it may be necessary at different stages of the solution to use the impedance and admittance

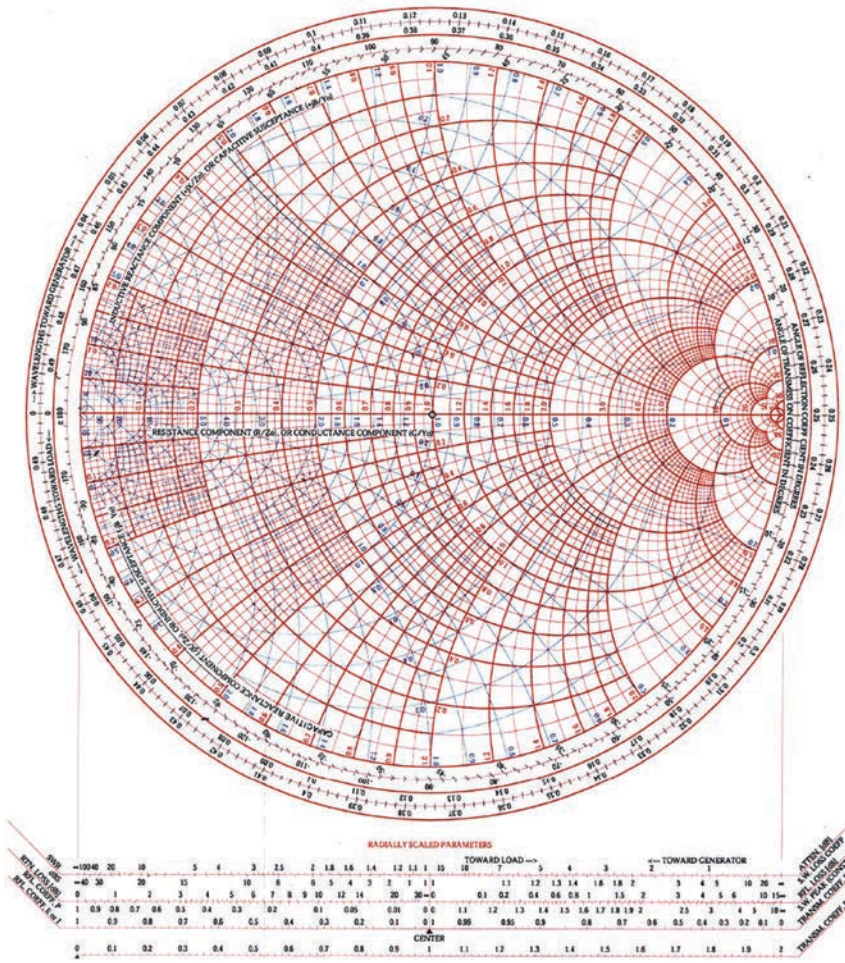


Figure 3.6 Smith and Jones chart (i.e., impedance-admittance Smith chart in dual grid: Z-Y format).

charts simultaneously; in such cases an *impedance-admittance Smith chart* needs to be used where the regular impedance chart is superimposed with another Smith chart that has been rotated 180^0 (which represents the admittance chart) and two scales have normally two different colors, the standard impedance chart is in "red" and the admittance chart is in "blue", as shown in Figure 3.6. This dual grid format ($Z - Y$) chart is also referred to as the *Smith and Jones chart*.

3.4 COMPRESSED SMITH CHART

So far, we have discussed the use of Smith chart for loss-less transmission line, in which the magnitude of the reflection coefficient term does not change with the line length and that we can find input impedance of the line from the load impedance, and vice versa, simply by moving along the $|\Gamma|$ circle by an angle equal to $2\beta l$. But on a lossy line the reflection coefficient at any point on the line will be given by: $\Gamma_l = |\Gamma| e^{-2\alpha l + j(\theta - 2\beta l)}$, in which the loss term $2\alpha l$ is now nonnegligible. Hence, to find z_{in} from z_L, we cannot simply move along the $|\Gamma|$ circle; auxiliary calculations are necessary to account for the $e^{-2\alpha l}$ factor. Implication of this loss factor is that as we now move from the load toward the generator the magnitude of the reflection coefficient decreases continuously and, hence, we now need to move from the load point to the generator point in a spiraling path directed toward the center. Many practical charts have convenient scales (such as the transmission loss scale in the radially scaled parameters of the Smith chart shown in Figure 3.5) attached to them, so that the amount of spiraling is readily obtained.

When working on the design of oscillators/amplifiers using negative resistance device, the magnitude of reflection coefficient happens to be more than unity, i.e., $|\Gamma| > 1$. Under these circumstances we cannot use the usual form of Smith chart discussed so far which is used for networks where $\mathrm{Re}(Z) \geq 0$ which corresponds to mapping of only the positive half of the impedance plane into a circle in the reflection coefficient plane with radius $|\Gamma| \leq 1$ (as shown in Figure 3.5); in which r takes the value from 0, i.e., s.c. to ∞, i.e., o.c. But to cater for negative resistance values we need a more general form/open form of Smith chart known as *compressed Smith chart*, which takes care of both positive and negative values of the real part of impedance as shown in Figure 3.7. The compressed Smith chart is obtained when the entire impedance plane is mapped in one-to-one basis onto the reflection coefficient plane. This mapping of impedance with both positive and negative real parts in the Γ plane helps to include the region of $|\Gamma| > 1$. In the open form, the compressed chart is of infinite extent, however, for all practical purposes it is a standard practice to bind the chart by imposing a limit on the maximum value of reflection coefficient that will exist in the design. A value of 3.16 is normally selected for the reflection coefficient as this allows a significant portion of the negative resistance contours that is usually encountered in practical design problems.

Finally, a few words about the "radially scaled parameters" at the bottom of the Smith chart (see Figures 3.5 and 3.6) may be mentioned now. The four scales on the right side fall under the general heading "reflection." The top scale marked "coefficient" is for $|\Gamma| = |V^- / V^+|$, the magnitude of reflection coefficient, and the adjacent scale gives the power reflection coefficient that

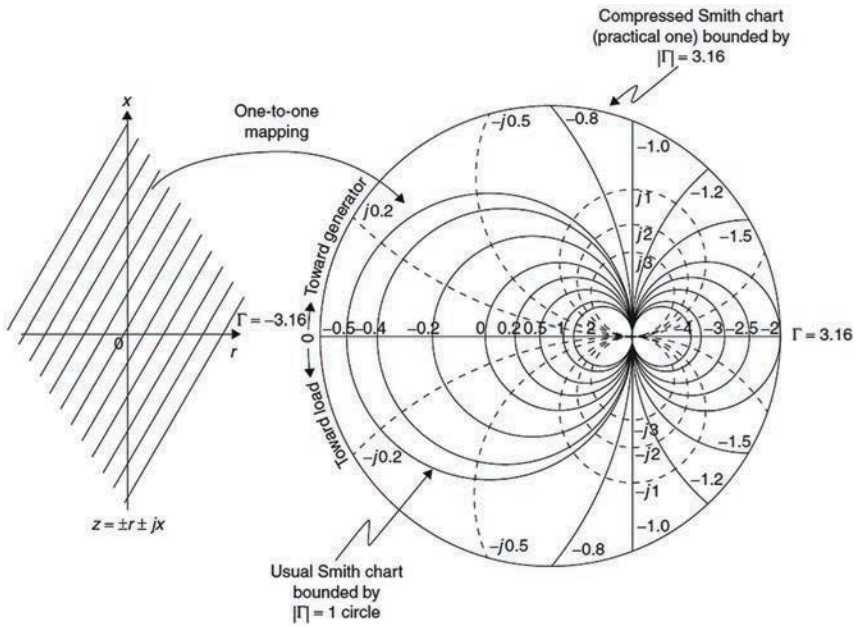


Figure 3.7 Compressed Smith chart.

is $|\Gamma|^2 = |V^- / V^+|^2$; both the scales start from the center of the Smith chart (where the magnitude of reflection coefficient is zero, *perfectly matched condition*) and end at the periphery of the chart (where the magnitude of reflection coefficient is maximum, i.e., equal to 1, *complete reflection*). The "loss in dB" scale under the same heading "reflection" gives the return and reflection losses. The "return loss" scale starts from 0 (corresponding to $|\Gamma| = 1$) at the periphery of the chart and approaches infinity at the center of the chart (where $|\Gamma| = 0$, *perfectly matched*), which implies that with better load matching reflection from load will be less and hence the return loss will be higher. The "reflection loss" scale starts with zero at the center of the chart (meaning that for perfectly matched condition, i.e., with $|\Gamma| = 0$, total power is absorbed in the load) and approaches infinity as $|\Gamma|$ approaches unity at the periphery of the chart (meaning that for *complete reflection*: $|\Gamma| = 1$, there is no absorbed signal by the load). The four scales on the left side have "standing wave" as one subheading expressed either in voltage ratio or in dB (with unity at the center for ratio scale and zero for dB scale as for perfectly matched condition minimum value of SWR is 1 and approaches to higher and higher values as one reaches toward the periphery of the chart when mismatch increases—leading to infinity for perfect mismatch

as for open, short, and purely reactive termination. The other subheading on the left side scale is for "transmission loss," which takes care of the line losses incurred in a lossy transmission line due to mismatched load causing VSWR. This scale starts from zero at the center of the Smith chart for matched load condition and increases as the VSWR increases. Further, when one moves from generator toward the load the loss incurred in a lossy line increase. Thus, the label "Towards load" is for moving away from the center of the chart toward the periphery (corresponding to progressively higher $|\Gamma|$ or higher VSWR), while the label "Towards generator" is for moving from the periphery of the chart toward the center (corresponding to progressively lower $|\Gamma|$ or lower VSWR). In this context it may be noted that the step size of "transmission loss in 1 dB steps" scale happens to be one-half the corresponding step size values on the "return loss in dB" scale. This is expected since transmission loss pertains to one-way attenuation loss while return loss is for two-way (up and down) attenuation through the transmission line.

3.5 STUB MATCHING DESIGN WITH SMITH CHART

3.5.1 Introduction

Impedance matching is a very important issue related to the design of an RF/microwave system. When a transmission line/waveguide is connected to a generator, a load, or another transmission line/waveguide, the impedances may, in general, be mismatched. This results in undue reflection at the location of the mismatched impedances and hence a standing wave gets set up on the line/guide, which is definitely undesirable. It is thus often necessary to match a transmission line/waveguide to a load or to generator, for the purpose of eliminating such undesirable standing waves on the line/guide. Similarly, if a discontinuity exists, such as the connection of unmatched line/guide sections, it is often necessary to eliminate this mismatch before the lines/guides can be connected to each other. The results of mismatch on a high frequency line/guide can be disastrous, as a large amount of reactive power may travel along the line/guide which can easily damage circuitry, especially generators.

A transmission line or waveguide is said to be matched to a load if the load impedance is equal to the characteristic impedance of the line. Similarly, if the line impedance is equal to the generator impedance, the two are matched. To match a load to a line (or a generator for that matter), a matching network is connected between the line and the load. There are two types of impedance matching networks that are particularly useful at RF and microwave frequencies. One matching technique, known as *stub matching*, employs short-circuited (or open-circuited) sections of transmission line (a *stub*) as reactive elements, which may be connected either

across (i.e., in parallel with) or along (i.e., in series with) the parent transmission line at a suitable distance from the load. The stub matching is applicable where the need is for narrow band matching (as the length of the stub is frequency dependent). The other method of impedance matching is based on the principle of impedance transformation with stepped impedance transformer or exponential transformer and so forth (normally used with waveguides), which can be made broadband, that will be discussed in Chapter 4.

3.5.2 Basics of Stub Matching and Analytical Equations for Design of Such Matching Circuits

3.5.2.1 Introduction

The idea of stub matching is to connect open- or short-circuited "stub" (sections of transmission lines) either in shunt (i.e., in parallel) or in series with the parent transmission line. Stub matching (or tuning, as it effectively tunes out the impedance mismatches) may be realized either with one piece of stub (single-stub matching) or with two/three pieces of stubs (double-stub or triple-stub matching). The idea of stub matching relies on the fact that the impedance of the stub and/or location of it on the parent transmission line is so chosen that the combined impedance of the parent line and the stub (or stubs) equals the characteristic impedance of the parent line.

In single-stub matching, the stub is placed at a suitable distance from the load and the susceptance (for shunt stub) or reactance (for series stub) presented by the stub cancels out the load susceptance at the location of the stub, thereby terminating the parent transmission line in its characteristic impedance. A disadvantage with single-stub matching is that every load requires a new stub position. Thus, if matching is desired for a variable impedance load, then single-stub matching is difficult to use in practice. The use of two stubs spaced by a fixed amount and located at a fixed distance from the load may be used to overcome this disadvantage. However, a double-stub matching system of this type may not be able to match all possible values of load admittance. A triple-stub matching arrangement can overcome this problem of double-stub tuner; in fact, a triple-stub tuner may be considered to be two double-stub tuners in series.

3.5.2.2 Single-Stub Matching

In single-stub matching (see Figure 3.8), the two adjustable parameters are the distance d from the load to the stub position, and the stub length l which determines the value of the susceptance/reactance provided by the shunt/series stub. Figures 3.8(a) and 3.8(b) shows single-stub matching, respectively, with shunt stub and series stub. In the diagrams: $y = g + jb$, where g

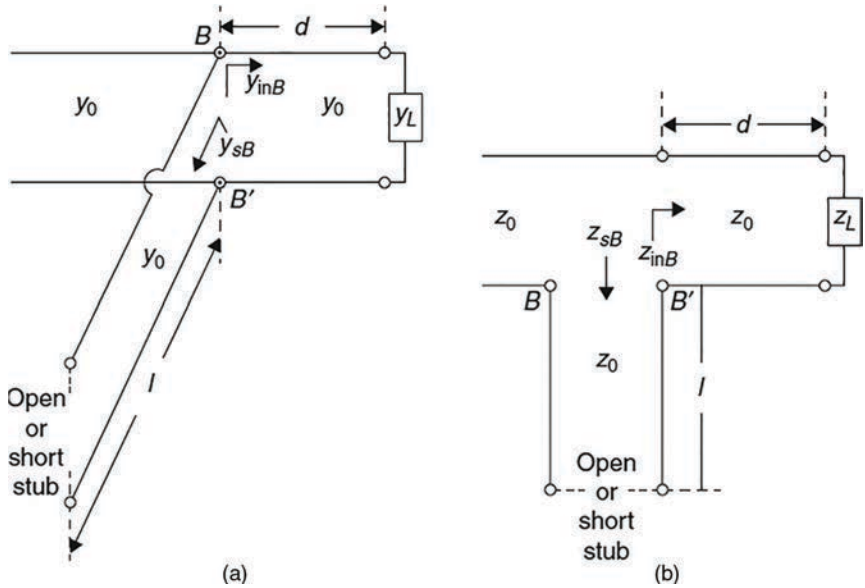


Figure 3.8 Single-stub matching with: (a) Shunt stub, and (b) Series stub.

is the conductance and b is the susceptance, while $z = r + jx$, where r is the resistance and x is the reactance.

Analytical equations for single-stub matching using shunt/series stub are given next [1,2].

The distance d can be obtained with:

$$d = \frac{\lambda}{2\pi}\tan^{-1}(t), \quad \text{for} \quad t \geq 0$$

$$= \frac{\lambda}{2\pi}[\pi + \tan^{-1}(t)], \quad \text{for} \quad t < 0$$

where: $t = \dfrac{1}{(r_L - 1)}\left\{x_L \pm \sqrt{r_L[(1 - r_L)^2 + x_L^2]}\right\}$, for $r_L \neq 1$ and $t = -\dfrac{x_L}{2}$, *for* $r_L = 1$
[*for shunt stub*]

$t = \dfrac{1}{(g_L - 1)}\left\{b_L \pm \sqrt{g_L[(1 - g_L)^2 + b_L^2]}\right\}$, for $g_L \neq 1$ and $t = -\dfrac{b_L}{2}$, for $g_L = 1$
[*for series stub*]

Length l can be obtained with:

A. When using of shunt stub:

i. *With open-circuited shunt stub*:

$$l_{oc} = \frac{\lambda}{2\pi}\tan^{-1}(b_{SB}) = \frac{\lambda}{2\pi}\tan^{-1}(-b_B), \quad \text{for } b_B \geq 0$$

$$= \frac{\lambda}{2\pi}[\pi + \tan^{-1}(b_{SB})] = \frac{\lambda}{2\pi}[\pi + \tan^{-1}(-b_B)], \quad \text{for } b_B < 0$$

ii. *With short-circuited shunt stub*:

$$l_{sc} = \frac{\lambda}{2\pi}\tan^{-1}\left(-\frac{1}{b_{SB}}\right) = \frac{\lambda}{2\pi}\tan^{-1}\left(\frac{1}{b_B}\right), \quad \text{for } b_B \geq 0$$

$$= \frac{\lambda}{2\pi}\left[\pi + \tan^{-1}\left(-\frac{1}{b_{SB}}\right)\right] = \frac{\lambda}{2\pi}\left[\pi + \tan^{-1}\left(\frac{1}{b_B}\right)\right], \quad \text{for } b_B < 0$$

(where b_B is the susceptance of the loaded line looking into the load from the stub position and b_{SB} is the stub susceptance).

B. When using series stub:

i. *With open-circuited series stub*:

$$l_{oc} = \frac{\lambda}{2\pi}\tan^{-1}\left(\frac{1}{x_{SB}}\right) = \frac{\lambda}{2\pi}\tan^{-1}\left(-\frac{1}{x_B}\right), \quad \text{for } x_B \geq 0$$

$$= \frac{\lambda}{2\pi}\left[\pi + \tan^{-1}\left(\frac{1}{x_{SB}}\right)\right] = \frac{\lambda}{2\pi}\left[\pi + \tan^{-1}\left(-\frac{1}{x_B}\right)\right], \quad \text{for } x_B < 0$$

ii. *With short-circuited series stub*:

$$l_{sc} = \frac{\lambda}{2\pi}\tan^{-1}(-x_{SB}) = \frac{\lambda}{2\pi}\tan^{-1}(x_B), \quad \text{for } x_B \geq 0$$

$$= \frac{\lambda}{2\pi}[\pi + \tan^{-1}(-x_{SB})] = \frac{\lambda}{2\pi}[\pi + \tan^{-1}(x_B)], \quad \text{for } x_B < 0$$

(where x_B is the reactance of the loaded line looking into the load from the stub position and x_{SB} is the stub reactance).

Thus, for a given load impedance, both d and l values can be calculated using the above equations.

3.5.2.3 Double-Stub Matching

In double-stub matching (see Figure 3.9), the first stub is located at the load plane marked as AA and the second stub is at a distance d from the first stub toward generator and located at BB. In practice d is chosen to be $\lambda/8$ (or $3\lambda/8$ at higher frequency).

The stub lengths for open- and short-circuited shunt stubs are:

$$\textit{For open-circuited stub: } l_{oc} = \frac{\lambda}{2\pi}\tan^{-1}(b)$$

and

$$\textit{For short-circuited stub: } l_{sc} = \frac{\lambda}{2\pi}\tan^{-1}\left(-\frac{1}{b}\right)$$

where $b = b_{SA,}$ and b_{SB} is the stub susceptance of the first stub and the second stub, respectively,

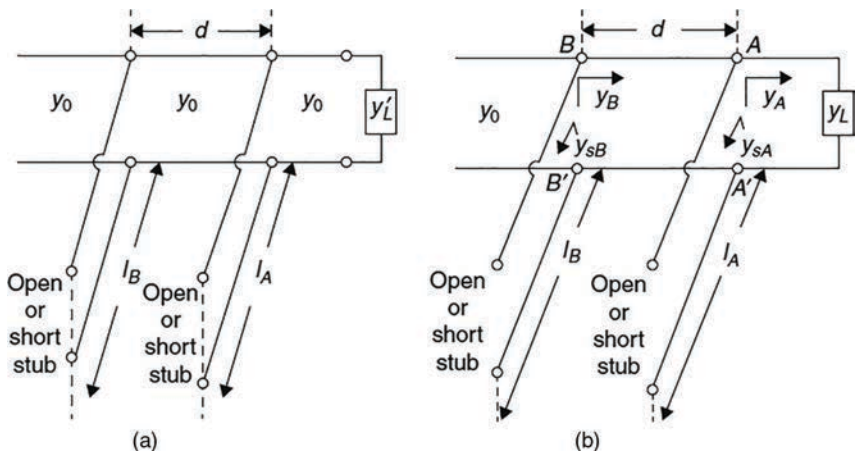


Figure 3.9 Double-stub matching: (a) Representative practical situation, and (b) Transformed load for ease of analysis.

given by:

$$b_{SA} = -b_L + \frac{1 \pm \sqrt{(1+t^2)g_L - g_L^2 t^2}}{t}$$

$$\text{while } b_{SB} = \frac{\pm\sqrt{g_L(1+t^2) - g_L^2 t^2 - g_L}}{g_L t}$$

$$\text{in which } g_L = \frac{1+t^2}{2t^2}\left[1 \pm \sqrt{1 - \frac{4t^2(1 - b_L t - b_{SA} t)^2}{(1+t^2)^2}}\right], \text{ with } t = \tan\beta d.$$

The analytical solution of stub matching problem may be used to obtain accurate results by solving the analytical equations through computer programs; however, such accuracies are seldom required in actual problem solutions and thus stub matching problem may be solved more easily by using Smith chart. A number of such design problems of single- and double-stub matching with shunt/series stub are given in the following section.

3.5.3 Design Examples of Stub Matching Using Smith Chart

Example 3.1 A lossless line of characteristic impedance $Z_0 = 50\Omega$ is to be matched to a load impedance $Z_L = 50/[2 + j(2 + \sqrt{3})]\Omega$ by means of a short-circuited shunt stub. The characteristic impedance of the stub is 100Ω. Determine the stub position and the stub length so that a match is obtained between the line and the load.

Solution:

Step 1: The shunt stub problems are convenient to solve by computing the normalized load admittance and entering it in the Smith chart used as an admittance chart. From the given data, normalized load admittance is:

$$y_L = Z_0 / Z_L = 50[2 + j(2 + \sqrt{3})] / 50 = 2 + j3.732.$$

Step 2: Next, we draw an admittance circle through the point y_L which intersects the normalized unit conductance circle (i.e., $g = 1$ circle) at two points y′ and y″, where y′ = 1.0 – *j*2.70 and y″ = 1.0 + *j*2.70 (see Figure 3.10). Each one of these provides one possible solution.

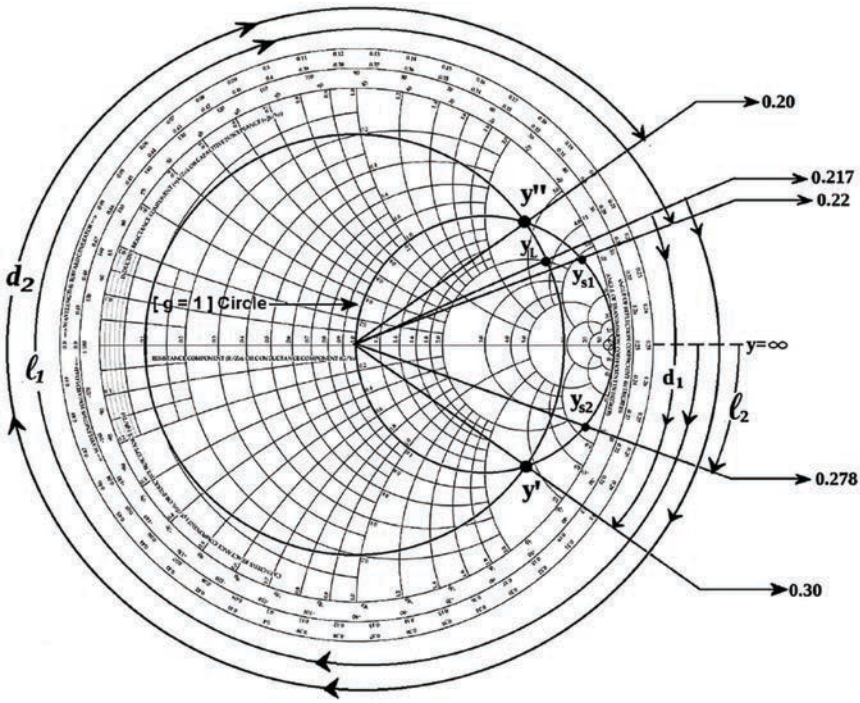


Figure 3.10 Smith chart solution of shunt type single-stub matching problem (Example 3.1).

Step 3(a): Let us first choose y′. We now move from the load point (y_L) on the admittance circle toward generator (i.e., WTG scale) to reach this point y′. As shown in Figure 3.10, the lines extended from the center of the chart through y_L and y′ intersect the WTG scale at 0.217λ and 0.30λ, respectively. Thus, the distance d_1 of the stub position from the load plane (see Figure 3.11) is obtained from Figure 3.10 as: $d_1 = (0.30 - 0.217)\lambda = 0.083\lambda$.

Step 3(b): Now at the junction corresponding to the point 1-1′ of the line (see Figure 3.11) the condition for impedance matching at the junction with the stub and the load requires that:

$$Y_{11} = Y' + Y_s, \quad \text{or} \quad y_{11}Y_0 = y'Y_0 + y_sY_{0s}$$

where Y_s is the susceptance that a stub of length l will contribute at 1-1′ plane and Y' is the line admittance. Y_0 and Y_{0s} are the characteristic impedances of the line and the stub, respectively.

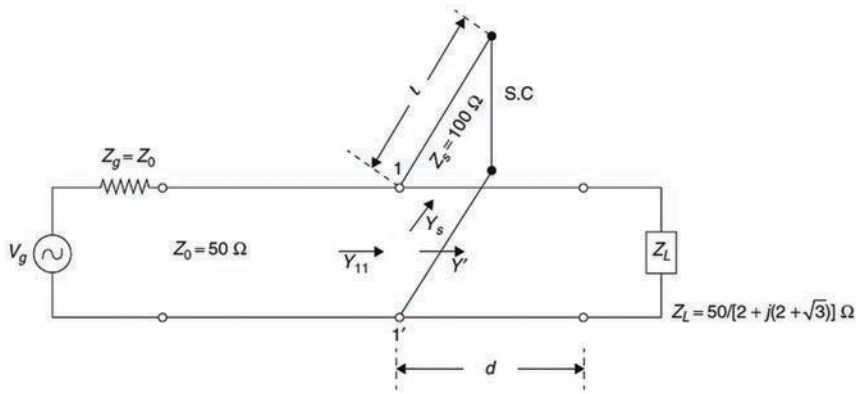


Figure 3.11 Scheme for short-circuited shunt type single-stub matching (Example 3.1).

Thus, $y_s = (y_{11} - y')Y_0/Y_{0s} = [1.0 - (1.0 - j2.7)]100/50 = j5.40$, which is the normalized stub susceptance (capacitive) for matching at 1–1′ junction.

[Note: Since the characteristic impedance of the stub is different from that of the line, y_s is different from $-b'$ (since $y' = g - jb'$). But if the characteristic impedances were equal then $y_s = -b'$, as the stub is added to cancel the susceptance of the loaded line at 1–1′ plane.]

Step 3(c): To obtain the length of the stub, we start at $y = \infty$ (i.e., at $z = 0$, the short-circuit point which corresponds to the right side of the real axis in the admittance chart, as shown in Figure 3.10). Since the stub contributes a susceptance of $+j5.40$, we enter this on the Smith chart and starting at $y = \infty$ we move along the WTG circle until we reach the point $y_{s1} = +j5.40$. The lines extended from the center of the chart through $y = \infty$ and $y_{s1} = +j5.40$ are drawn that intersect the WTG circle at 0.250λ and 0.22λ, respectively. Thus, the stub length is:

$l_1 = 0.25\lambda + 0.22\lambda = 0.47\lambda$

Step 4: If we would have chosen y″, then $y_{s2} = -j5.40$ (i.e., the required normalized stub susceptance is inductive) and it may be seen from Figure 3.10 that on moving from the load point (y_L) on the admittance circle toward generator (i.e., WTG scale) we reach the point y″ by covering a distance of d_2 given by:

$d_2 = 0.50 - (0.217 - 0.20) = 0.50 - 0.017 = 0.483\lambda$

In this case the stub length is obtained by starting at $y = \infty$ and moving along the WTG circle until we reach the point $y_{s2} = -j5.40$, when we get:

$l_2 = 0.278 - 0.25 = 0.028\lambda$

[Note: Out of the two possible solutions it is advisable to choose the one that has the stub position closest to the load (in this case choose the solution with d_1). This helps to improve bandwidth of the match and also reduces the possibility of losses that might otherwise be caused by large standing wave ratio on the line between the stub and the load.]

Example 3.2 A transmission line of characteristic impedance $Z_0 = 75\Omega$ is to be matched to a load impedance $Z_L = (15 + j60)\Omega$ by means of a short-circuited series stub having same characteristic impedance as the line. Determine the stub position and the stub length so that a match is obtained between the line and the load.

Solution:

Step 1: The series stub problems are convenient to solve by computing the normalized load impedance and entering it in the Smith chart used as an impedance chart. From the given data, normalized load impedance is: $z_L = Z_L/Z_0 = (15+ j60)/75 = (0.2 + j0.8)$.

Step 2: Next, we draw an impedance circle through the point z_L which intersects the normalized unit resistance circle (i.e., $r = 1$ circle) at two points z′ and z″, where $z' = (1.0 + j2.55)$ and $z'' = (1.0 - j2.55)$, as shown in Figure 3.12. Each one of these provides one possible solution.

Step 3(a): Let us first choose z′. We now move from the load point (z_L) on the impedance circle toward generator (i.e., WTG scale) to reach this point z′. As shown in Figure 3.12, the lines extended from the center of the chart through z_L and z′ intersect the WTG scale at 0.109λ and 0.196λ, respectively. Thus, the distance d_1 of the stub position from the load plane (see Figure 3.13) is obtained from Figure 3.12 as:

$d_1 = (0.190 - 0.109)\lambda = 0.081\lambda$

Step 3(b): The normalized line reactance at z′ is $+j2.55$. The stub length must be such that its normalized input impedance is $-j2.55$. This required impedance is marked as point z_{s1} in the Smith chart. Now starting at z = 0 (short-circuit point, which corresponds to the left side

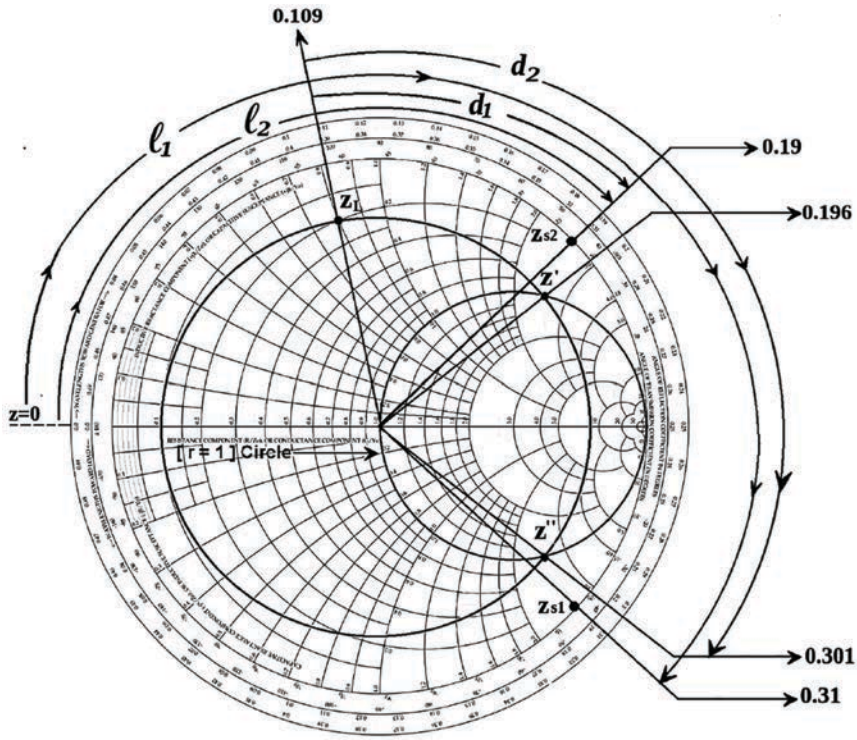


Figure 3.12 Smith chart solution of series type single-stub matching problem (Example 3.2).

of the real axis in the impedance chart, as shown in Figure 3.12) we move along the WTG circle until we reach the point $z_{s1} = -j2.55$. Lines extended from the center of the chart through z = 0 and $z_{s1} = -j2.55$ are drawn that intersect the WTG circle at 0.00λ and 0.31λ, respectively. Thus, the stub length is:

$$l_1 = 0.31\lambda - 0.00\lambda = 0.31\lambda$$

Step 4: If we would have chosen z″, then $z_{s2} = +j2.55$ and it may be seen from Figure 3.12 that on moving from the load point (z_L) on the impedance circle toward generator (i.e., WTG scale) we reach the point z″ by covering a distance of d_2 given by:

$$d_2 = (0.31\lambda - 0.109\lambda) = 0.301\lambda$$

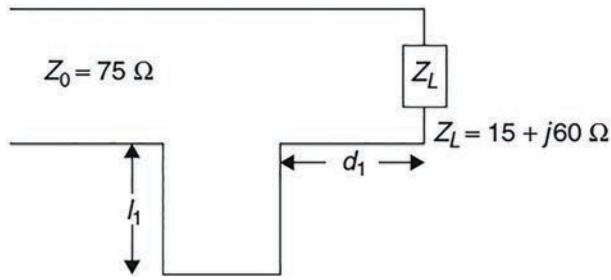


Figure 3.13 Scheme for series type short-circuited single-stub matching (Example 3.2).

In this case the stub length is obtained by starting at z = 0 and moving along the WTG circle until we reach the point $z_{s2} = +j2.55$, when we get:

$l_2 = 0.190\lambda - 0.00\lambda = 0.190\lambda$

Either solution is correct; however, from a practical point of view the closest stub to the load (i.e., the first solution) needs to be chosen.

Example 3.3 A 75Ω double-stub shunt type tuner is used to match a load $Y_L = (5.70 + j1.50)$ mS at 1.0 GHz to a 75 Ω lossless air dielectric line. The first stub is located at the load and the separation between the stubs is 6 cm. Find the lengths of the shorted shunt stubs for the matched condition.

Solution

Step 1: With the given data the normalized load admittance can be found to be: $y_L = Y_L R_0 = (5.70 + j1.50) \times 10^{-3} \times 75 = (0.427 + j\ 0.112)$, which is located on the Smith chart.

Step 2: To carry out this stub matching problem we assume that stub 2 will match the load admittance given above (forgetting for the moment the presence of stub 1) acting as short-circuited single-stub and then moving by a distance of 0.2λ (corresponding to the separation between the two stubs: (λ/5) = 6 cm—λ being equal to 30 cm as per given data in the problem) toward the load to reach the position of the first stub and the final match is then realized (Figure 3.14). To do this we first draw the unit conductance circle g_2 for stub 2 and then

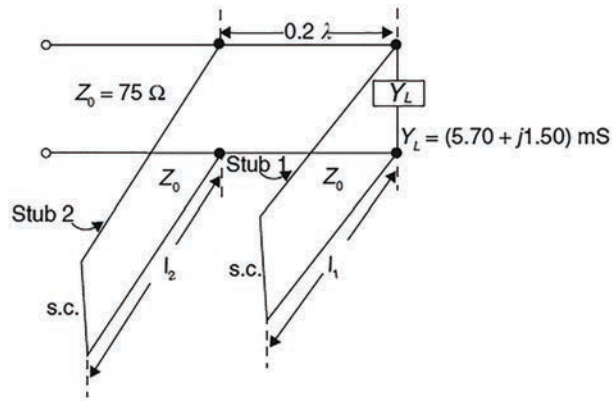


Figure 3.14 Scheme for the double-stub matching with shunt type short-circuited stubs (Example 3.3).

moving 0.2λ toward load (WTL scale) we draw the unit conductance circle g_1 for stub 1 as indicated in Figure 3.15.

Step 3: The solution process starts with stub 1 at the load. The stub's impedance is purely imaginary and thus it can only change the susceptance of the combined stub and load while the conductance remains the same. To find the combined admittance on the unit circle (g =1) for stub 1, we start from the load point y_L and move toward generator on the constant conductance circle (with $g_L = 0.427$) which cuts the g_1 circle at two points—y'_1 and y''_1. The admittances at y'_1 and y''_1 are, respectively, $(0.427 + j0.79)$ and $(0.427 - j0.22)$. Thus, the susceptance contributed by stub 1 is: $y'_{1a} = y'_1 - y_L = (0.427 + j0.79) - (0.427 + j0.112) = j0.678$ and $y''_{1b} = y''_1 - y_L = (0.427 + j0.22) - (0.427 + j0.112) = j0.108$. These two values are shown as S'_1 and S''_1 in Figure 3.15. The possible stub lengths are found by moving from $y = \infty$ point (the short-circuit point in the admittance chart) toward the generator (WTG scale) to the points S'_1 and S''_1, when we get: $l_{1a} = (0.25 + 0.095)\lambda = 0.345\lambda$ and $l_{1b} = (0.25 + 0.018)\lambda = 0.268\lambda$.

Step 4: Now we consider the admittances at y'_1 and y''_1 as the new load admittances. From here on, we treat the problem as a single-stub matching for each of these admittances and the distance between the load and stub 2 being known and equal to 0.2λ. We draw the admittance circle through y'_1, and starting at y'_1 as we move toward the generator by 0.2λ, we intersect the unit conductance circle g_2

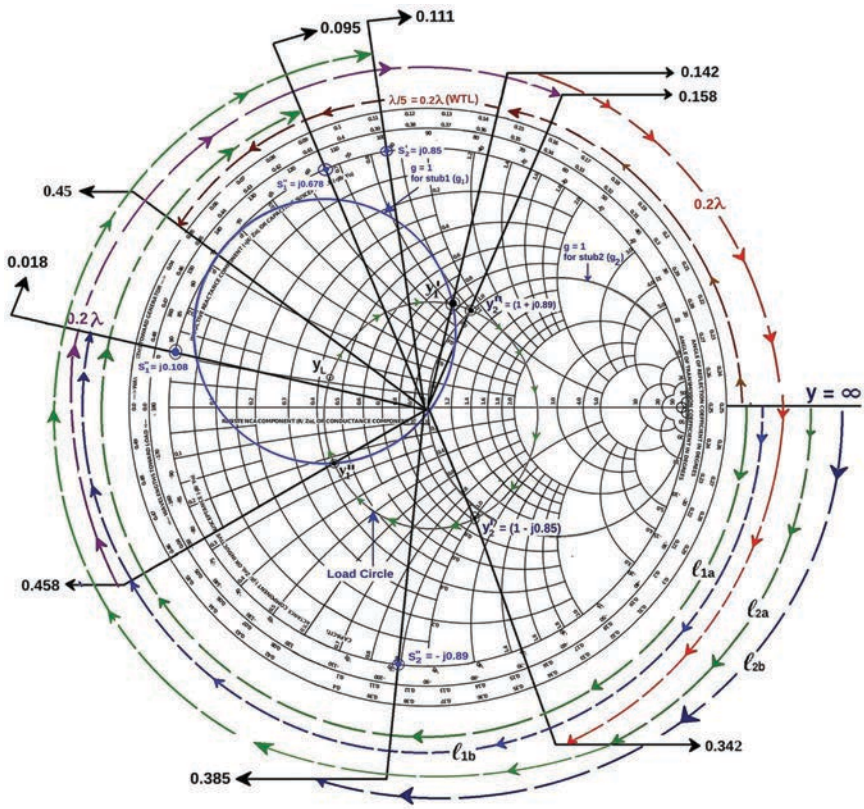


Figure 3.15 Smith chart showing the solution of the double-stub matching problem (Example 3.3).

for stub 2 at y'_2. Although we cut g_2 at another point symmetrically located above the real axis, this intersection cannot be used since stub 2 must be at a distance of 0.2λ from the load. At y'_2, the line admittance is: $(1 - j0.85)$. Thus, the stub must have an admittance of $+j0.85$ so that the line susceptance is cancelled at the location of stub 2. The latter is marked as S'_2. The stub length that will accomplish this is the distance between the short-circuit point $y = \infty$ and S'_2. This is: $l_{2a} = (0.25 + 0.111)\lambda = 0.361\lambda$. Similarly for the point y''_1, we draw the admittance circle and move 0.2λ toward the generator, to point y''_2. The line admittance at y''_2 is: $(1 + j0.89)$. The required stub susceptance is $-j0.89$, which is marked as point S''_2. The stub length is the distance between $y = \infty$ and this point, when we get $l_{2b} = (0.385 - 0.25)\lambda = 0.135\lambda$.

Since the two stubs with first solution (having l_{1a} = 0.345 λ and l_{2a} = 0.361λ) have larger lengths of the stub compared to the second solution (l_{1b} = 0.268λ and l_{2b} = 0.135λ), the first solution will thus have narrow bandwidth of operation compared to the second solution.

REFERENCES

[1] Collin R E, *Foundations for Microwave Engineering*, Second Edition, John Wiley & Sons, New Jersey, 2007.

[2] Liao S Y, *Microwave Devices and Circuits*, Third Edition, Pearson Education Inc., India, 2003.

Chapter 4

Impedance Matching in Microwave Design

If your actions match your words—
Know, that is a test of your genuineness

4.1 INTRODUCTION

In Chapter 3 we have discussed the impedance matching technique at RF frequencies, known as *stub matching*, that employs short-circuited or open-circuited sections of transmission line (a stub) as reactive elements, which may be connected either in shunt or in series with the parent transmission line at a suitable distance from the load. The stub matching is applicable where the need is for narrow band matching.

The other method of impedance matching is based on the principle of impedance transformation, commonly used in waveguides and also in microstrip transmission lines. The theory of small reflection at the junction of multiple quarter-wave impedance transformers is used to design broad-band impedance transformers (since the single quarter-wave transformer is narrow band as its length is frequency dependent). The *stepped* and the *tapered impedance transformers* are commonly used for impedance matching at microwave frequencies.

Incorporation of an inductive/capacitive window (iris) or introduction of a tuning screw in a waveguide is also used as a means for impedance matching in certain applications for the microwave circuit design. Further, a simple yet practical technique is to use an L-section network formed with a lumped component/planar distributed line reactive elements for impedance matching, used for input/output matching in the design of RF/microwave amplifiers.

4.2 IMPEDANCE TRANSFORMERS

Impedance transformers at microwave frequency are necessary in order to match two dissimilar impedances which may arise due to connection of two transmission lines/waveguides of different characteristic impedances (for example, when connecting two waveguides of different heights, an empty waveguide to a dielectric-loaded waveguide, and a lens antenna to free-space). The section of a quarter-wave (λ/4) transmission line is known

DOI: 10.1201/9781003729969-4

to act as an impedance transformer, known as *quarter-wave impedance transformer*, at RF/microwave frequency; however, it will perform as a very narrow band transformer (as its length is dependent on the wavelength). The frequency response for a single-section quarter-wave transformer is shown in Figure 4.1(a), which represents the variation of $|\Gamma|$ with θ (= βl = $f\pi/2f_0$,), where $|\Gamma|$ is the magnitude of the reflection coefficient and $\theta = \beta l$ is the electrical length of the transmission line. Further, a plot of the reflection coefficient magnitude versus normalized frequency is shown in Figure 4.1(b) for various mismatched loads and it may be observed that larger bandwidth is realizable with smaller impedance ratio (Z_L/Z_1) between the two impedances Z_L and Z_1 to be matched [1]. In other words, the bandwidth of single-section quarter-wave transformer increases as the impedances to be matched (Z_L and Z_1 in the present case) become closer to each other (for less mismatched load). It must also be noted that a quarter-wave transformer can only match real load impedance. Although complex load impedance can always be transformed to real impedance either by using an appropriate length of transmission line between the load and the transformer (or with an appropriate series or shunt reactive stub, as

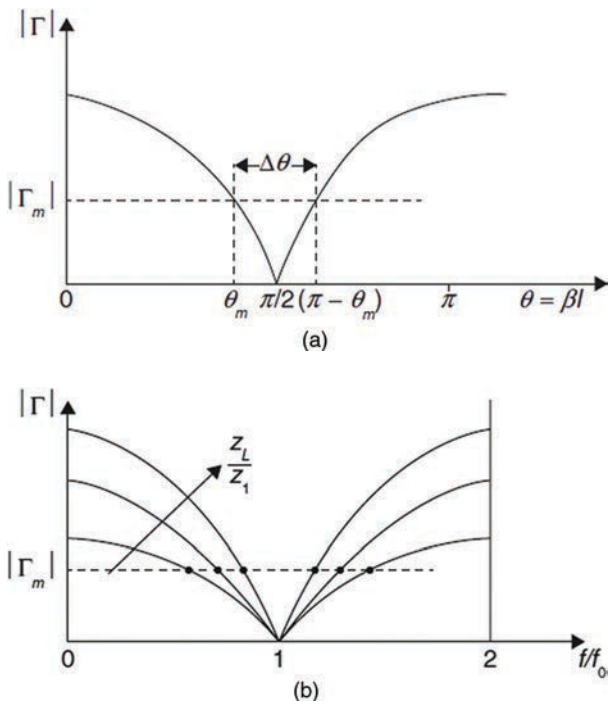


Figure 4.1 (a) Frequency response of a single-section quarter-wave transformer, and (b) Effect of load mismatch on the bandwidth of single-section quarter-wave transformer.

discussed in Chapter 3), these techniques will usually alter the frequency dependence of the equivalent load, which often has the effect of reducing the bandwidth of match.

In order to realize impedance matching over a broad band of frequencies, a multi-section quarter-wave transformer is used. This is done in principle by cascading multiple sections of quarter-wave lines (waveguide or microstrip line) having progressively increasing/decreasing characteristic impedance. The analysis and design of such an impedance transformer is done on the basis of the *theory of small reflections* [1] in which the discontinuity (small impedance mismatch) between the junctions of two consecutive quarter-wave sections is considered to exhibit a "small reflection" (with VSWR, of the order of, say, 1.05). The theory of small reflection also finds application in the design of other RF/microwave devices and circuits, such as directional couplers, filters, and antenna arrays.

4.2.1 Theory of Small Reflections

The overall reflection coefficient at the input of any structure having multiple discontinuities (causing reflection) may be evaluated by considering all those multiple reflections taking place at different discontinuities (junctions) of the structure. It will be seen that for small reflections, the total reflection coefficient can be obtained only from the first-order reflections.

Let us first consider a single-section quarter-wave transformer having two discontinuities at which the reflection coefficient (Γ) and transmission coefficient (T) are indicated by appropriate suffices in Γ and T (Figure 4.2(a)). These are expressed by:

$$\Gamma_1 = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \quad \Gamma_2 = -\Gamma_1, \quad \Gamma_3 = \frac{Z_3 - Z_2}{Z_3 - Z_2} \tag{4.1}$$

$$T_{21} = 1 + \Gamma_1, \quad T_{12} = 1 + \Gamma_{21} = 1 - \Gamma_1 \tag{4.2}$$

Let us assume that a unit amplitude wave is incident at the input. Inspection of Figure 4.2(b) shows that the net reflected wave amplitude due to multiple partial reflections is:

$$\begin{aligned}\Gamma &= \Gamma_1 + T_{21}T_{12}\Gamma_3 e^{-j2\theta} + T_{12}T_{21}\Gamma_3^2\Gamma_2 e^{-j4\theta} + \ldots. \\ &= \Gamma_1 + T_{12}T_{21}\Gamma_3 e^{-j2\theta}\sum_{n=0}^{\infty}\Gamma_2^n\Gamma_3^n e^{-j2n\theta} \\ &= \Gamma_1 + \frac{T_{12}T_{21}\Gamma_3 e^{-j2\theta}}{1 - \Gamma_2\Gamma_3 e^{-j2\theta}}\end{aligned} \tag{4.3}$$

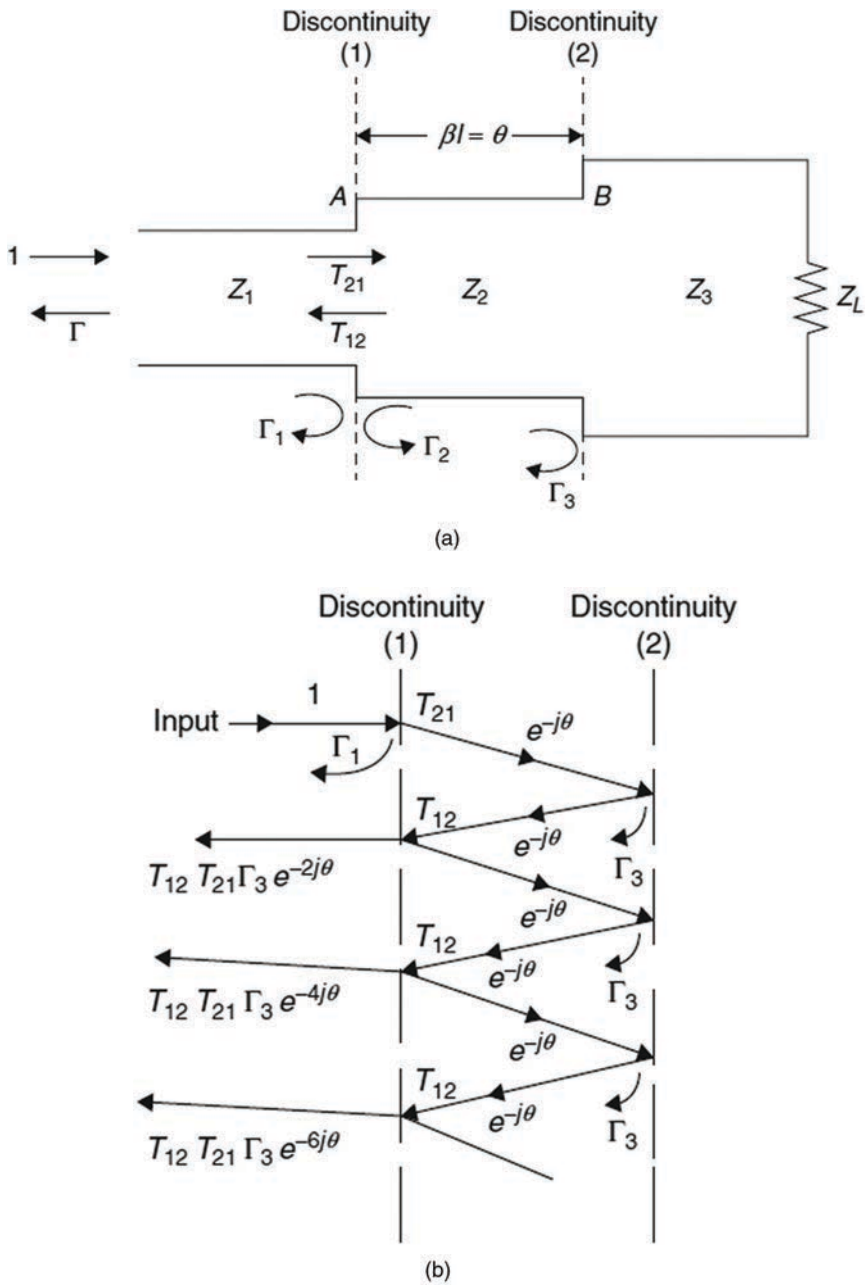


Figure 4.2 (a) Single-section quarter-wave transformer with two discontinuities, and (b) Multiple partial reflections of waves at the two discontinuities (reflecting surfaces).

where we have used the geometric series:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \text{ for } |x| < 1$$

Using $\Gamma_2 = -\Gamma_1$, $T_{21} = 1 + \Gamma_1$ and $T_{12} = 1 - \Gamma_1$ from Eqs. (4.1) and (4.2), Eq. (4.3) may be expressed as:

$$\Gamma = \frac{\Gamma_1 + \Gamma_3 e^{-j2\theta}}{1 + \Gamma_1 \Gamma_3 e^{-j2\theta}} \tag{4.4}$$

Assuming small reflections, $|\Gamma_1|$ and $|\Gamma_3|$ << 1, we can approximate Eq. (4.4) as:

$$\Gamma \approx \Gamma_1 + \Gamma_3 e^{-j2\theta} \tag{4.5}$$

This result states that the total reflection is dominated by the reflection from the initial discontinuity between Z_1 and Z_2, i.e., Γ_1 and the first reflection from the discontinuity between Z_2 and Z_L, i.e., $\Gamma_3 e^{-2j\theta}$. The $e^{-2j\theta}$ term accounts for the phase delay when the incident wave travels up and down the line/guide. Thus, for small reflections, the total reflection coefficient can be obtained only from the first-order reflection, e.g., with $\Gamma_1 = \Gamma_3 \leq 0.2$ the maximum error is only 4%.

4.2.2 Stepped Impedance Binomial and Tchebyscheff Transformer

Let us consider the problem of matching two impedances Z_1 and Z_{n+1} using $(n - 1) = N$ sections, i.e., n steps of the quarter-wave transformers (note that the number of sections N is always less than the number of steps by 1; in Figure 4.2(a) we had single section and two step transformer). These n steps will generate n reflection coefficients $\Gamma_1, \Gamma_2, \ldots\ldots, \Gamma_n$. Because of the discontinuities at the junctions, each component wave will undergo multiple reflections back and forth, between any two successive steps. Therefore, for $(n - 1)$ sections, the total reflection coefficient Γ_T at any frequency f, can be written as:

$$\Gamma_T = \Gamma_1 + \Gamma_2 e^{-j2\theta} + \Gamma_3 e^{-j4\theta} + \ldots\ldots \Gamma_n e^{-j2(n-1)\theta} \tag{4.6}$$

where $\Gamma_i = (Z_{i+1} - Z_i)/(Z_{i+1} + Z_i)$, the reflection coefficient at the junction between $(i + 1)$th and ith sections for small reflection.

Assuming that the transformer is symmetrical with respect to a transverse reference plane at the center (this however does not mean that Z_n's are symmetrical):

$$\Gamma_1 = \Gamma_n, \Gamma_2 = \Gamma_{n-1}, \Gamma_3 = \Gamma_{n-2} \text{ etc.} \tag{4.7}$$

Then,

$$\Gamma_T = 2[\Gamma_1 \cos(n-1)\theta + \Gamma_2 \cos(n-3)\theta + \Gamma_3 \cos(n-5)\theta + ...] \tag{4.8}$$

$$= 2\sum_{k=1}^{m} \Gamma_k \cos\{n-(2k-1)\}\theta;\ m = n/2,\ [\text{for } n \text{ even}] \tag{4.9}$$

$$= 2\sum_{k=1}^{m} \Gamma_k \cos\{n-(2k-1)\}\theta + \Gamma_{(n+1)/2};\quad m = (n-1)/2,\ [\text{for } n \text{ odd}] \tag{4.10}$$

It may be observed that when $\theta = \pi/2$,

$$\Gamma_T = 0;\ \text{when } n \text{ is even} \tag{4.11}$$

$$= \Gamma_{(n+1)/2} - 2\Gamma_{(n-1)/2} + 2\Gamma_{(n-3)/2} + ... \pm 2\Gamma_1;\ \text{when } n \text{ is odd} \tag{4.12}$$

By properly choosing the reflection coefficient Γ_k, k = 1, 2, ….. n, a variety of pass-band characteristics of a multi-section transformer can be realized. The impedances Z_k's for various sections are calculated from Γ_k's. In one approach the Γ_k's are set proportional to binomial coefficients, i.e., to the coefficients of $(a + b)^{n-1}$. In this approach, the magnitude of Γ_T is zero at f_0 and increases slowly and monotonically on either side of f_0. The curve at f_0 is the *flattest* realizable and thus is popularly known as *Maximally Flat/binomial* transformer. In another approach, Γ_T is made proportional to the Tchebyscheff (an alternative spelling of this Russian name is Chebyshev) polynomial $T_{n-1}(x)$, where x is a function of both actual frequency and bandwidth. In this approach, Γ_T oscillates in magnitude over the pass-band frequency but remains less (in magnitude) than a prescribed ripple value. This type of impedance transformer is known as *Equal-Ripple/Tchebyscheff transformer*.

4.2.2.1 Binomial Stepped Impedance Transformer

Binomial transformer analysis and design consists in letting the Γ_k's to be proportional to the coefficients of $(a + b)^{n-1}$, i.e.:

$$\Gamma_k = \Gamma_1 C_{k-1}^{n-1} \tag{4.13}$$

The reflection coefficient Γ_T can now be expressed as:

$$\Gamma_T = 2\Gamma_1 \sum_{k=1}^{m} C_{k-1}^{n-1} \cos\{n-(2k-1)\}\theta; \quad m = n/2, \text{ [for } n \text{ even]} \tag{4.14}$$

$$= 2\Gamma_1 \sum_{k=1}^{m} C_{k-1}^{n-1} \cos\{n-(2k-1)\}\theta + C_{(n-1)/2}^{n-1};$$

$$m = (n-1)/2, \text{ [for } n \text{ odd]} \tag{4.15}$$

Expanding cos $p\theta$ in terms of cos θ, where $p = n - (2k - 1)$, we get:

$$\cos p\theta = 2\cos(p-1)\theta \cos\theta - \cos(p-2)\theta \ldots\ldots \tag{4.16}$$

It may be verified that, when the expansion of cos $p\theta$ in terms of cos θ is substituted into the expression of Γ above, only one term involving $\cos^{(n-1)}\theta$ with coefficient $2^{(n-2)}$ does not cancel out, i.e.:

$$\Gamma = 2\,\Gamma_1\, 2^{(n-2)} \cos^{(n-1)}\theta = \Gamma_1\, 2^{(n-1)} \cos^{(n-1)}\theta \ldots\ldots \tag{4.17}$$

Since all the derivatives of Γ at $f = f_0$, where $\theta = \pi/2$ up to the $(n-2)^{th}$ term vanishes, the binomial distribution provides a *maximally flat* response as shown in Figure 4.3.

In the design of binomial transformer, the problem is stated in terms of a tolerable (maximum) reflection coefficient $|\Gamma| = |\Gamma_m|$ over a specified band of

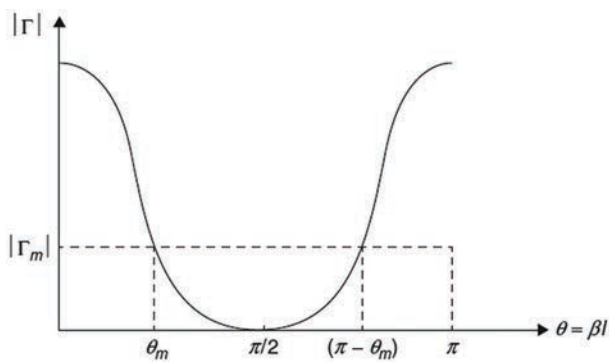


Figure 4.3 Maximally flat response.

frequencies (i.e., pass-band) from f_1 to f_2 with $f_0 = (f_1 - f_2)/2$ realizable with n number of sections.

The maximum value of Γ_1 that occurs at the band edges may then be computed from equation:

$$|\Gamma_1| \le \frac{|\Gamma_m|}{2^{n-1}\cos^{n-1}\theta_1} \quad \text{and} \quad |\Gamma_2| \le \frac{|\Gamma_m|}{2^{n-1}\cos^{n-1}\theta_2} \tag{4.18}$$

where θ_1 and θ_2 correspond to the total phase shift per section at f_1 and f_2, which are, respectively, the lower and higher ends of the frequency band. It may be noted that for symmetry of the response curve, $\cos\theta = |\cos\theta_2|$, or $\theta_2 = (\pi - \theta_1)$.

Now, let us put together some of the design formulas for binomial impedance transformer [2,3]. Let r_w be the ratio of the wavelengths at the band-edge frequencies f_1 and f_2. Then:

$$r_w = \frac{\lambda_{g_1}}{\lambda_{g_2}} = \frac{\theta_2}{\theta_1} = \frac{\pi - \theta_1}{\theta_1} = \frac{\theta_2}{\pi - \theta_2} \tag{4.19}$$

$$\theta_1 = \frac{\pi}{1 - r_w} = \frac{2\pi l}{\lambda_{g_1}} = \frac{\pi\lambda_{g_0}}{2\lambda_{g_1}} \tag{4.20}$$

$$\theta_2 = \frac{\pi r_w}{1 + r_w} = \frac{2\pi l}{\lambda_{g_2}} = \frac{\pi\lambda_{g_0}}{2\lambda_{g_2}} \tag{4.21}$$

$$l = \frac{\lambda_{g_0}}{4} = \frac{\theta_1\lambda_{g_1}}{2\pi} = \frac{\theta_2\lambda_{g_2}}{2\pi} \tag{4.22}$$

$$\lambda_{g_0} = \frac{2\lambda_{g_1}\lambda_{g_2}}{\lambda_{g_1} + \lambda_{g_2}} \tag{4.23}$$

If n is not specified, it is to be selected by trial-and-error method to ensure that $|\Gamma| \le |\Gamma_m|$ over the desired pass-band.

The characteristic impedances of the sections are determined from the relation:

$$\Gamma_k = \frac{Z_{k+1} - Z_k}{Z_{k+1} + Z_k} = C_{k-1}^{n-1} \tag{4.24}$$

where $Z_{(i+1)}$ is the characteristic impedance of the ith section and Γ_i is the reflection coefficient at the input of ith section. Further, if the reflections are small (in all practical designs, $\Gamma_k \leq 0.35$ when error is less than 4%) we can write:

$$\ln\left(\frac{Z_{k+1}}{Z_k}\right) = 2\left[\frac{Z_{k+1}-Z_k}{Z_{k+1}+Z_k} + \frac{1}{3}\left(\frac{Z_{k+1}-Z_k}{Z_{k+1}+Z_k}\right)^3 + ...\right]$$

$$= 2\left[\Gamma_k + \frac{1}{3}\Gamma_k^3\right]$$

$$\approx 2\Gamma_k$$

$$\text{Thus, } \Gamma_k = \frac{1}{2}\ln\frac{Z_{k+1}}{Z_k} = C_{k-1}^{n-1}\frac{1}{2}\ln\frac{Z_2}{Z_1} \tag{4.25}$$

$$\text{When } \ln\left[\frac{Z_{k+1}}{Z_k}\right] = C_{k-1}^{n-1}\ln\frac{Z_2}{Z_1} \tag{4.26}$$

Therefore, the impedances can be computed from the corresponding reflection coefficients:

$$\Gamma_1 = \frac{1}{2}\ln\left(\frac{Z_2}{Z_1}\right), \quad Z_2 = Z_1[anti\ln(2\Gamma_1)]$$

$$\Gamma_2 = \frac{1}{2}\ln\left(\frac{Z_3}{Z_2}\right), \quad Z_3 = Z_2[anti\ln(2\Gamma_2)].$$

$$\Gamma_k = \frac{1}{2}\ln\left(\frac{Z_{k+1}}{Z_k}\right), \quad Z_{k+1} = Z_k[anti\ln(2\Gamma_k)] \tag{4.27}$$

By adding up to n terms:

$$\Gamma_T = \Gamma_1 + \Gamma_2 + ... + \Gamma_n = \frac{1}{2}\ln\left(\frac{Z_{n+1}}{Z_1}\right) \tag{4.28}$$

This gives the total reflection coefficient of n steps or $(n-1) = N$ section transformer.

Example 4.1 Design a two-section binomial impedance transformer to match a 225Ω load (typical characteristic impedance of an X-band rectangular waveguide) to a 75Ω line (characteristic impedance of a coaxial line) at 10GHz. What bandwidth is obtained for maximum tolerable reflection coefficient of 0.05? Evaluate the design dimensions of the transformer and characterize the transformer by optimizing the design using 3D electromagnetic field simulator.

Solution:

Given that the number of sections, $N = 2$. Thus, the number of steps, $n = (2 + 1) = 3$.

Here: $Z_1 = 75\Omega$, $Z_{n+1} = Z_4 = 225\Omega$, $\Gamma_m = 0.05$, $f_0 = 10$ GHz ($\lambda_0 = 3.0$ cm).

Since $n = 3$, we have from Eq. (4.28):

$$\Gamma_T = \Gamma_1 + \Gamma_2 + \Gamma_3 = \frac{1}{2}\ln\left(\frac{Z_4}{Z_1}\right) = \frac{1}{2}\ln 3 = \frac{1.099}{2} = 0.549$$

As the transformer is symmetrical, $\Gamma_1 = \Gamma_3$; and $\Gamma_2 = \Gamma_1 C_{(2-1)}^{(3-1)} = 2\Gamma_1$ (see Eq. 4.13)

Therefore, $\Gamma_1 + 2\Gamma_1 + \Gamma_1 = 0.549$, or $\Gamma_1 = 0.137$ and $\Gamma_2 = 0.274$

Now, since: $\Gamma_k = \frac{1}{2}\ln\left(\frac{Z_{k+1}}{Z_k}\right)$; $Z_{k+1} = Z_k[anti\ln(2\Gamma_k)]$

So, $Z_2 = Z_1[anti\ln(2\Gamma_1)] = 75[anti\ln(2\times 0.137)] = 98.64\Omega$

$Z_3 = Z_2[anti\ln(2\Gamma_2)] = 98.64[anti\ln(2\times 0.274)] = 170.63\Omega$

Length of the sections is $\lambda_g/4 = 0.993$ cm (since $\lambda_0 = 3$ cm and $\lambda_g = 3.975$ cm).

Now: $\Gamma_m = \Gamma_1 2^{n-1}\cos^{n-1}\theta_1 = \Gamma_1 2^{n-1}\cos^{n-1}\theta_2$ (Eq. 4.17)

Thus: $0.05 = 0.137 \times 2^2 \times \cos^2\theta_1$; or, $\cos^2\theta_1 = 0.091$; when, $\cos\theta_1 = 0.302$;

with, $\theta_1 = 71.33^0$; and, $\theta_2 = (180 - \theta_1) = 108.67^0$.

Therefore, $\theta = (\theta_2 - \theta_1) = (108.67^0 - 71.33^0) = 37.34^0$.

When: $r_w = \theta_2/\theta_1 = 108.67/71.33 = 1.523$

For TEM line the bandwidth can be calculated as:

$\Delta f = (2f_0/\pi)(\theta_2 - \theta_1) = (2 \times 10/180)(37.34) = 4.148$ GHz.

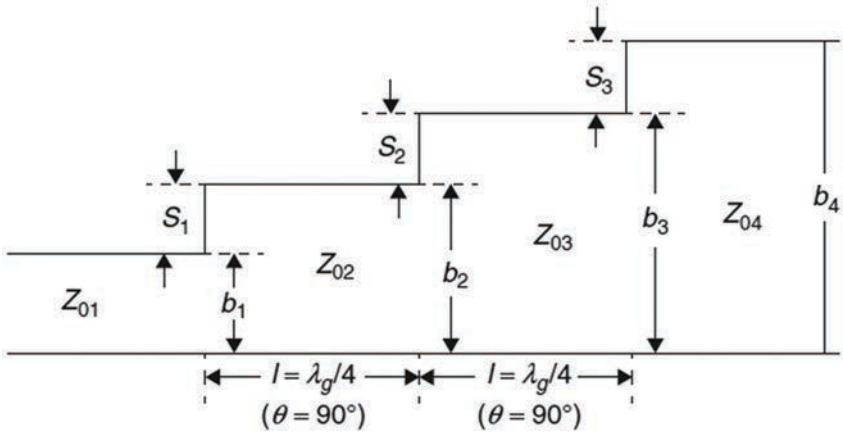


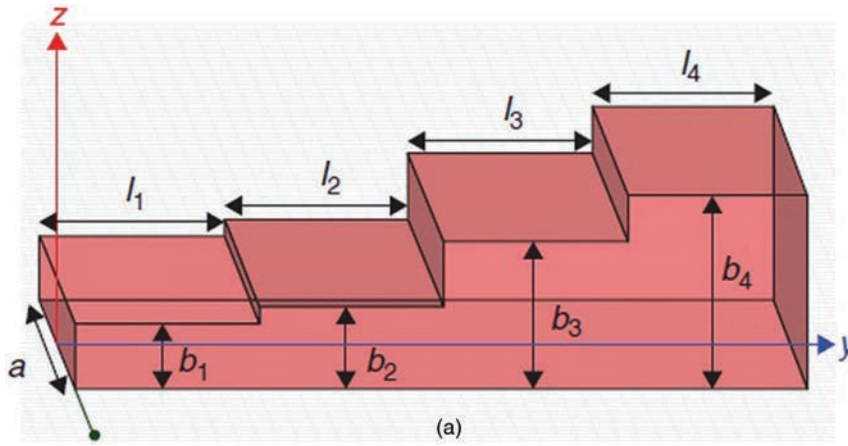
Figure 4.4 Schematic of stepped impedance transformer [n = 3 (steps), N = 2 (section)] for Examples 4.1 and 4.3.

The design dimensions of the transformer may be evaluated as follows. Referring to Figure 4.4, the physical length of each section of the transformer is $\lambda_g/4$ (electrical length $\theta = 90^0$). The step heights are given by: $s_1 = (b_2 - b_1)$, $s_2 = (b_3 - b_2)$, and $s_3 = (b_4 - b_3)$, where b is given by: $b = \dfrac{Z_0 a}{377}\sqrt{1 - \left(\dfrac{\lambda_0}{2a}\right)^2}$ (which follows from the expression for characteristic impedance of rectangular waveguide).

For X-band waveguide a = 0.9″ = 2.286cm and with λ_0 = 3cm, λ_g = 3.975cm. Here Z_{01} = 75Ω, Z_{02} = 98.64Ω, Z_{03} = 170.625Ω, Z_{04} = 225Ω. It may be calculated that b_1 = 0.3431cm, b_2 = 0.4513cm, b_3 = 0.7807cm, b_4 = 1.0295cm. Thus, s_1 = 0.1082cm, s_2 = 0.3294cm, and s_3 = 0.2488cm.

The impedance transformer designed through the synthesis process (discussed above) is simulated in 3D electromagnetic (EM) field simulator to realize optimized design with little trimming of the physical dimensions of the transformer. The optimized simulated structure of the transformer along with different dimensions and a comparison of the simulation result and synthesis result are shown in Figure 4.5(a–c).

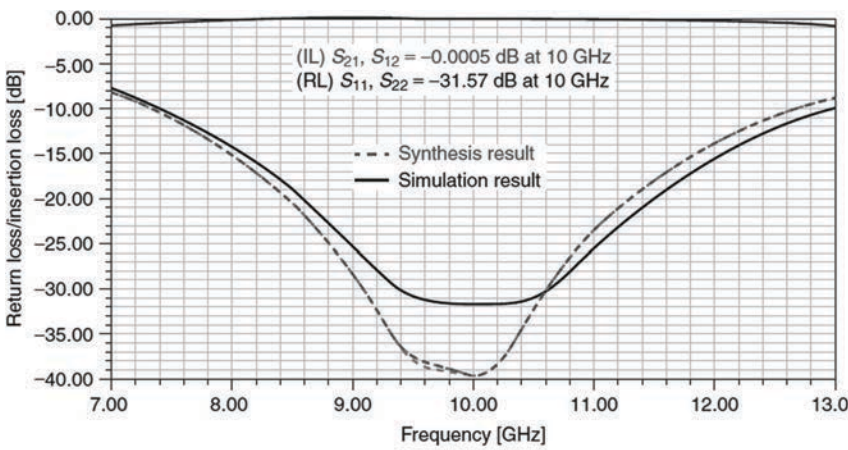
Example 4.2 Design a five-section quarter-wave stepped impedance transformer at 10GHz to match a 100Ω line to an 800Ω load. Assume binomial distribution pass-band response of the reflection coefficients for the band-edge frequency ratio of 1.75. Find the maximum VSWR at the input of the transformer.



(a)

Parameter	Synthesis dimensions	Simulation dimensons
a	22.86	22.86
b_1, (l_1)	3.431, (9.5)	3.431, (9.5)
b_2, (l_2)	4.513, (9.5)	4.50, (9.4)
b_3, (l_3)	7.807, (9.5)	7.80, (9.1)
b_4, (l_4)	10.295, (9.5)	1.0295, (9.5)

(b)



(c)

Figure 4.5 3D EM field simulator-based design of binomial transformer: (a) Structure, (b) Design dimensions (mm), and (c) Comparison of the simulation result and synthesis result.

Solution:

Given that the number of sections $N = 5$. Thus, the number of steps $n = (5 + 1) = 6$.

$Z_1 = 100\ \Omega$, $Z_{n+1} = Z_7 = 800\ \Omega$

Assuming small reflection at each step:

$$\Gamma_T = \Gamma_1 + \Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5 + \Gamma_6 = \frac{1}{2}\ln\left(\frac{Z_7}{Z_1}\right) = \frac{1}{2}\ln\left(\frac{800}{100}\right) = \frac{2.079}{2} = 1.039$$

For binomial distribution: $\Gamma_k = \Gamma_1 C_{(k-1)}^{(n-1)} = \Gamma_1 C_{k-1}^5$

Thus: $\Gamma_1 = \Gamma_1 C_0^5 = \Gamma_6$; $\Gamma_2 = \Gamma_1 C_1^5 = 5\Gamma_1 = \Gamma_5$; $\Gamma_3 = \Gamma_1 C_2^5 = 10\Gamma_1 = \Gamma_4$

where the symmetry of the binomial transformer is used.

Therefore, $2(\Gamma_1 + 5\,\Gamma_1 + 10\,\Gamma_1) = 1.039$ or $32\Gamma_1 = 1.039$

When $\Gamma_1 = 0.03246 = \Gamma_6$, $\Gamma_2 = 0.1623 = \Gamma_5$, and $\Gamma_3 = 0.3246 = \Gamma_4$

The characteristic impedances of the sections are:

$$Z_2 = Z_1[anti\ln(2\Gamma_1)] = 100[anti\ln(2\times 0.03246)] = 106.70\Omega$$

$$Z_3 = Z_2[anti\ln(2\Gamma_2)] = 106.70[anti\ln(2\times 0.1623)] = 147.61\Omega$$

$$Z_4 = Z_3[anti\ln(2\Gamma_3)] = 147.61[anti\ln(2\times 0.3246)] = 282.52\Omega$$

$$Z_5 = Z_4[anti\ln(2\Gamma_4)] = 282.52[anti\ln(2\times 0.3246)] = 540.74\Omega$$

$$Z_6 = Z_5[anti\ln(2\Gamma_5)] = 540.74[anti\ln(2\times 0.1623)] = 748.11\Omega$$

At the band-edge,

$$\theta_1 = \frac{\pi}{1+r_w} = \frac{\pi}{1+1.75} = \frac{\pi}{2.75} = 1.142 \text{ radian} = 65.44^0$$ (since $r_w = 1.75$, given).

Therefore, $\cos\theta_1 = 0.4156$ and $\cos^5\theta_1 = 0.01239$

Now, $\Gamma_m = \Gamma_1\, 2^{n-1} \cos^{n-1}\theta_1 = 0.03246 \times 2^5 \times \cos^5\theta_1 = 0.012869$

Thus, maximum $VSWR = \dfrac{1+|\Gamma_m|}{1-|\Gamma_m|} = \dfrac{1+0.012869}{1-0.012869} = 1.02607$

[Note: It may be noted that for practical stepped impedance transformer design more than 2–3 steps are not required because the impedance ratios met with are normally 3–5 and hardly as large as 8 in the present case. However; if that happens, then stepped impedance transformation is used to take care of matching impedance ratio up to 4 or 5 and further impedance transformation is done by other novel techniques.]

4.2.2.2 Tchebyscheff Stepped Impedance Transformer

One may also let Γ_k's to be proportional to $T_{n-1}(x)$, where $T_n(x)$ is the Tchebysheff polynomial of order n given by:

$$T_n(x) = \cos(n \cos^{-1} x);\ -1 < \times < 1 \tag{4.29a}$$

$$T_n(x) = \cosh(n \cosh^{-1} x);\ x < -1,\ x > 1$$

While x is a parameter for change of variable which is a function of frequency being given by: $x = \cos\theta / \cos\theta_1 = \cos\theta / \left|\cos\theta_2\right|$. Instead of maximally flat pass-band characteristics of binomial transformer the Tchebyscheff transformer exhibits an *equal-ripple* characteristic in which the maximum tolerable reflection coefficient $\left|\Gamma_m\right|$ vanishes at as many different frequencies in the pass-band as there are transformer sections (N). In contrast with the binomial matching transformer, the Tchebyscheff transformer optimizes bandwidth at the expense of pass-band ripple. If such a pass-band characteristic can be tolerated, the Tchebyscheff transformer can provide a substantially wider bandwidth compared to the binomial transformer, for a given number of sections.

The nth order Tchebyscheff polynomial is a polynomial of degree n and is denoted by $T_n(x)$. The first four Tchebyscheff polynomials are:

$$\begin{aligned} T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \end{aligned} \tag{4.29b}$$

Higher order polynomials can be found by using the following recurrence formula:

$$T_k(x) = 2x\, T_{k-1}(x) - T_{k-2}(x) \ldots\ldots \tag{4.29c}$$

The first four Tchebyscheff polynomials are plotted in Figure 4.6 from which the following important properties of Tchebysheff polynomials may be noted:

i. For $-1 \le x \le 1$, $\left|T_k(x) \le 1\right|$, i.e., all the polynomials are limited within the range (1, -1), which means that the Tchebyscheff polynomials oscillate between ±1 and hence it exhibits equal-ripple characteristics that are mapped in the pass-band of the transformer.
ii. For $|x| > 1$, $\left|T_k(x) \ge 1\right|$, which falls outside the pass-band and $\left|T_k(x)\right|$ increases faster with x as k increases.

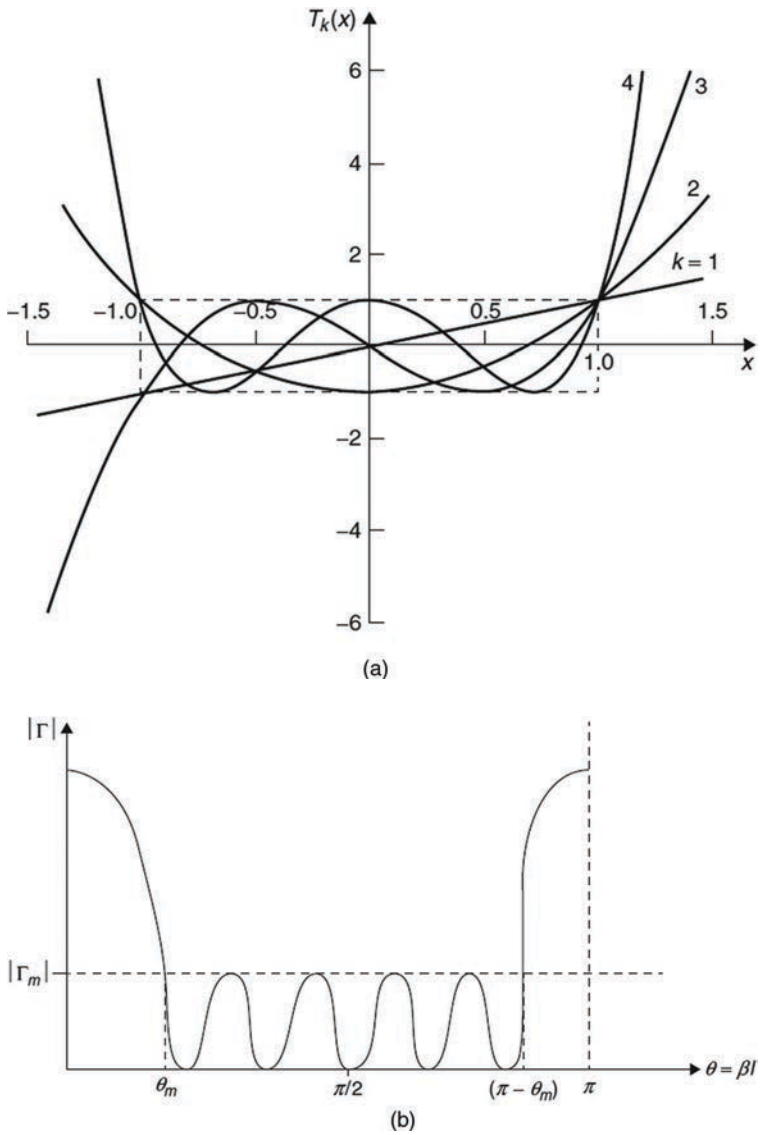


Figure 4.6 (a) (a) Tchebyscheff polynomials of different order (here *k* is up to 4), and (b) Equal-ripple characteristics for a five-section transformer.

For the design of Tchebyscheff impedance matching transformer, on specifying a maximum tolerable reflection coefficient $|\Gamma_m|$ we can write with reference to a transverse plane at the center by using the total reflection coefficient expressions given by Eqs. (4.9) and (4.10):

$$\Gamma_T = |\Gamma_m| T_{n-1}(x) = 2\sum_{k=1}^{m} \Gamma_k \cos\{n-(2k-1)\}\theta; \quad m = n/2, \text{ [for } n \text{ even]} \tag{4.30}$$

$$= 2\sum_{k=1}^{m} \Gamma_k \cos\{n-(2k-1)\}\theta + \Gamma_{(n+1)/2}; \quad m = (n-1)/2, \text{ [for } n \text{ odd]} \tag{4.31}$$

This ensures that $|\Gamma_m T_{n-1}(x)| \le |\Gamma_m|$, since $|T_{n-1}(x)| \le 1$. The Γ_k's are obtained by expanding the right-hand side of Eqs. (4.30) and (4.31) in terms of powers of cos θ and substituting x cos θ_1 for cos θ and equating the coefficient of like powers of x with those on left-hand side. The Tchebyscheff transformer gives the largest bandwidth for a specified $|\Gamma_m|$ and number of sections and for a given bandwidth it gives the lowest ripple $|\Gamma_m|$.

Example 4.3 Design a two-section Tchebyscheff impedance transformer to match a 225 Ω load to a 75 Ω line at 10 GHz. The maximum VSWR allowed is 1.1. Find the corresponding bandwidth. Evaluate the step dimensions of the transformer. Using 3D electromagnetic field simulator characterize the designed transformer.

Solution:

Given that the number of sections $N = 2$. Thus, the number of steps $n = (2 + 1) = 3$.

$Z_1 = 75\ \Omega$, $Z_{n+1} = Z_4 = 225\ \Omega$

$$|\Gamma_m| = \frac{1.1-1}{1.1+1} = \frac{0.1}{2.1} = 0.0476 \text{ (since maximum VSWR is specified as 1.1)}$$

Total reflection coefficient for symmetrical transformer (n odd) is:

$\Gamma = |\Gamma_m| T_{n-1}(x) = 0.0476\ T_2(x) = 2\ [\Gamma_1 \cos 2\theta] + \Gamma_2$ (see Eq. 4.31)

or, $0.0476\ [2x^2 - 1] = 2\ [\Gamma_1\ (2\cos^2\theta - 1)] + \Gamma_2$
$= 2\ [\Gamma_1\ (2x^2 \cos^2\theta_1 - 1)] + \Gamma_2$

Equating the coefficients of x^2, x, and so on

$2 \times 0.0476 = 4\ \Gamma_1 \cos^2\theta_1$; or, $\Gamma_1 = 2 \times 0.0476 / 4 \cos^2\theta_1$
and $-0.0476 = -2\ \Gamma_1 + \Gamma_2$; or, $\Gamma_2 = 2\ \Gamma_1 - 0.0476$

Now at zero frequency, θ (= βl) = 0, and we can write for general reflection coefficient:

$(Z_4 - Z_1)/(Z_4 + Z_1) = \Gamma_m T_{n-1} (\sec \theta_1)$; [since $x = \cos\theta/\cos\theta_1 = \sec\theta_1$]

or, $(225 - 75)/(225 + 75) = 0.0476T_2 (\sec \theta_1)$
or, $150/300 = 0.0476 (2 \sec^2 \theta_1 - 1)$
or, $\sec^2 \theta_1 = 5.752$, when $\cos^2 \theta_1 = 0.1738$ and $\cos \theta_1 = 0.4168$, with $\theta_1 = 65.367^0 = 0.36315\pi$ radian.

Therefore, $\Gamma_1 = (2 \times 0.0476) / (4 \times 0.1738) = 0.1369$
$\Gamma_2 = 2\Gamma_1 - 0.0476 = 2 \times 0.136 - 0.0476 = 0.22627$

Now, $\Gamma_1 = (Z_2 - Z_1) / (Z_2 + Z_1)$, and $\Gamma_2 = (Z_3 - Z_2) / (Z_3 + Z_2)$

When $Z_2 = \dfrac{(1+\Gamma_1)}{(1-\Gamma_1)} Z_1 = \dfrac{1+0.1369}{1-0.1369} \times 75 = 98.79\Omega$

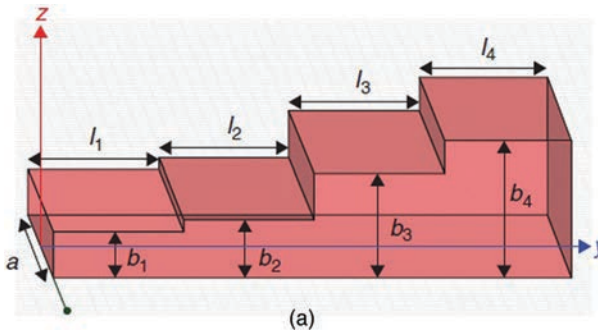
$$Z_3 = \frac{(1+\Gamma_2)}{(1-\Gamma_2)} Z_2 = \frac{1+0.22627}{1-0.22627} \times 98.79 = 156.57\Omega$$

The fractional bandwidth: $\Delta f/f_0 = 4/\pi (\pi/2 - \theta_1)$
$= 4/\pi (\pi/2 - 0.36315\pi)$
or, $\Delta f/f_0 = 0.5474$
Thus, the bandwidth is: $\Delta f = 0.5474 \times 10$ GHz $= 5.4744$ GHz.

[Note: When compared with the result of Example 4.1, it is observed that Tchebyscheff transformer is capable of providing larger bandwidth than binomial transformer].

The step dimensions of the transformer can be calculated using the necessary formula and the data values as appropriate (following Example 4.1) and in this design we obtain: $b_1 = 0.3431$ cm, $b_2 = 0.45193$ cm, $b_3 = 0.7162$ cm, $b_4 = 1.0295$ cm and hence $s_1 = 0.10883$ cm, $s_2 = 0.26427$ cm, and $s_3 = 0.3133$ cm (where $Z_{01} = 75\ \Omega$, $Z_{02} = 98.79\ \Omega$, $Z_{03} = 156.57\ \Omega$, and $Z_{04} = 225\ \Omega$). The schematic of transformer designed with the synthesis process is similar to that shown in Figure 4.5.

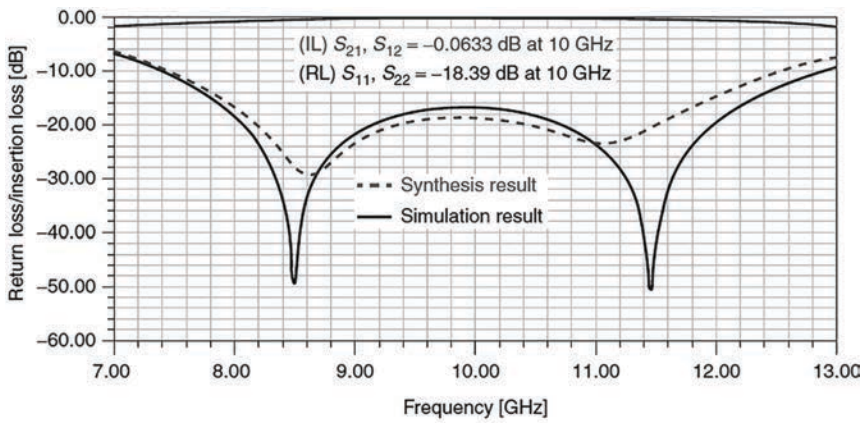
The impedance transformer designed through the synthesis process (discussed above) is simulated in 3D electromagnetic field simulator to realize optimized design. The optimized simulated structure of the transformer along with different dimensions and a comparison of the simulation result and synthesis result are shown in Figure 4.7(a–c).



(a)

Parameter	Synthesis dimensions	Simulation dimensons
a	22.86	22.86
b_1, (l_1)	3.431, (9.5)	3.431, (9.5)
b_2, (l_2)	4.519, (9.5)	4.80, (9.5)
b_3, (l_3)	7.162, (9.5)	7.3, (9.5)
b_4, (l_4)	10.295, (9.5)	10.00, (9.5)

(b)



(c)

Figure 4.7 3D EM field simulator-based design of Tchebyscheff transformer: (a) Structure, (b) Design dimensions, and (c) Comparison of the simulation result and synthesis result.

Example 4.4 Design a three-section binomial and also Tchebyscheff transformer at 10 GHz with the following data: Impedances to be matched are 100 Ω to 50 Ω, allowable maximum reflection coefficient $\Gamma_m = 0.04$. Show that the bandwidth realizable with Tchebyscheff transformer is larger than that with binomial transformer but with a penalty of non-zero ripple in the pass-band of the Tchebyscheff transformer.

Repeat the design for 1, 2, and 4 section transformers in each case and with the plot of reflection coefficient magnitude versus frequency show that greater bandwidth is realizable using a greater number of sections for both the impedance transformers.

Solution:

Binomial Transformer:

Given that the number of sections $N = 3$. Thus, the number of steps $n = (3 + 1) = 4$.

Here: $Z_1 = 50\ \Omega$, $Z_{n+1} = Z_5 = 100\ \Omega$, $\Gamma_m = 0.04$, f_0 = center frequency = 10 GHz.

Since $n = 4$, we have from Eq. (4.28):

$$\Gamma_T = \Gamma_1 + \Gamma_2 + \Gamma_3 + \Gamma_4 = \frac{1}{2}\ln\left(\frac{Z_5}{Z_1}\right) = \frac{1}{2}\ln 2 = 0.3465$$

As the transformer is symmetrical, $\Gamma_1 = \Gamma_4$; and $\Gamma_2 = \Gamma_3 = \Gamma_1 C_{(3-1)}^{(4-1)} = 3\Gamma_1$ (see Eq. 4.13).

Therefore, $2\Gamma_1 + 6\Gamma_1 = 0.3465$, or $\Gamma_1 = 0.0433$, and $\Gamma_2 = \Gamma_3 = 0.1299$

Now since: $\Gamma_k = \frac{1}{2}\ln\left(\frac{Z_{k+1}}{Z_k}\right)$; $Z_{k+1} = Z_k[anti\ln(2\Gamma_k)]$

Also: $\Gamma_m = \Gamma_1 2^{(n-1)} \cos^{(n-1)} \theta_1 = \Gamma_1 2^{(n-1)} \cos^{(n-1)} \theta_2$ (see Eq. 4.17)

Thus: $0.04 = 0.0433 \times 2^3 \times \cos^3 \theta_1$; or, $\cos^3 \theta_1 = 0.1154$; when, $\cos \theta_1 = 0.4868$;

with, $\theta_1 = 60.86^0$; and, $\theta_2 = (180 - \theta_1) = 119.14^0$.

Therefore, $\theta = \theta_2 - \theta_1 = (119.14^0 - 60.86^0) = 58.28^0 = 0.3237\pi$ radian.

When: $r_w = \theta_2 / \theta_1 = 119.14/60.86 = 1.957$

For TEM line, the bandwidth can be calculated as:

$\Delta f = (2f_0/\pi)(\theta_2 - \theta_1) = (2 \times 0.32327)f_0 = 0.6475f_0$ GHz = 6.475 GHz.

Tchebyscheff Transformer:

Given that the number of sections, $N = 3$. Thus, number of steps, $n = (3 + 1) = 4$.

Here: $Z_1 = 50\Omega$, $Z_{n+1} = Z_5 = 100\Omega$, $\Gamma_m = 0.04$, f_0 = Centre frequency = 10GHz.

Total reflection coefficient for symmetrical transformer (n even) is:

$$\Gamma = \left|\Gamma_m\right| T_{n-1}(x) = 0.04\, T_3(x) = 2\,[\Gamma_1 \cos 3\theta + \Gamma_2 \cos\theta] \text{ (see Eq. 4.31)}$$

$$\text{or, } 0.04\,[4x^3 - 3x] = 2\,[\Gamma_1 (4\cos^3\theta - 3\cos\theta) + \Gamma_2 \cos\theta]$$

$$= 2\,[\Gamma_1 (4x^3 \cos^3\theta_1 - 3x\cos\theta_1) + \Gamma_2 x \cos\theta_1]$$

Equating the coefficients of x^4, x^2, and so on:

$$4 \times 0.04 = \Gamma_1 \cos^3\theta_1 \times 8; \text{ or, } \Gamma_1 = 0.04/2 \times \cos^3\theta_1$$

$$\text{and } (-3\Gamma_1 + \Gamma_2) \times 2\cos\theta_1 = -3 \times 0.04$$

Now at zero frequency, $\theta = 0$, and we can write for general reflection coefficient:

$$(Z_5 - Z_1)/(Z_5 + Z_1) = \Gamma_m T_{n-1}(\sec\theta_1) \text{ [since } x = \cos\theta/\cos\theta_1 = \sec\theta_1]$$

$$\text{or, } (100 - 50)/(100 + 50) = 0.04 T_3 (\sec\theta_1)$$

$$\text{or, } 50/150 = 0.04\,(4\sec^3\theta_1 - 3\sec\theta_1)$$

or, $\sec\theta_1 = 1.47$, when $\cos\theta_1 = 0.68$; with $\theta_1 = 47.156^0 = 0.2619\,\pi$ radian.

The fractional bandwidth: $\Delta f/f_0 = 4/\pi\,(0.5 - 0.2619)\pi$

$$\text{or, } \Delta f/f_0 = 0.9524$$

Thus, the bandwidth is: $\Delta f = 0.95 f_0$ GHz = 9.524 GHz.

Therefore, the bandwidth obtained with a Tchebyscheff transformer (Δf = 9.524 GHz) is greater than that realizable with a binomial transformer (Δf = 6.475 GHz); however, for Tchebyscheff transformer this increase of bandwidth is obtained with a penalty of tolerating passband ripple.

The plots of reflection coefficient magnitude versus frequency for binomial transformer and Tchebyscheff transformer are shown in Figures 4.8(a) and 4.8(b), respectively, which clearly indicates that greater bandwidth is realizable using a greater number of sections for both the impedance transformers. It may also be observed that for a given maximum reflection coefficient and for a given number of sections the Tchebysheff transformer can provide larger bandwidth than binomial transformer. It may further be observed that for a Tchebyscheff transformer, the number of dips in the ripple is equal to the number of sections (N) and one less than the number of steps n ($=N+1$), i.e., the order of the filter.

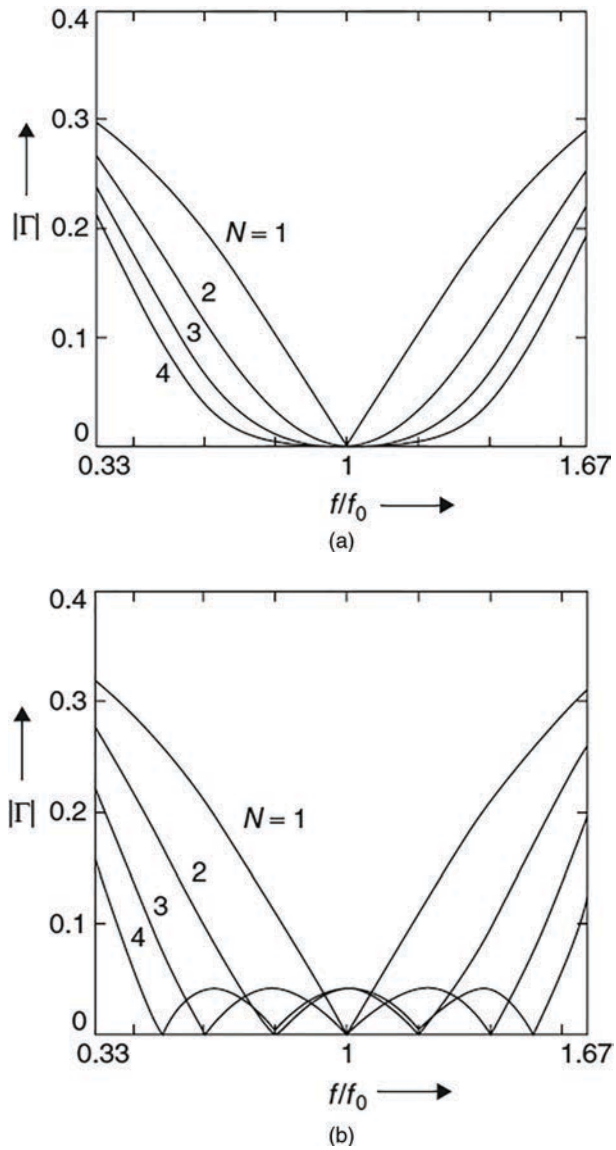


Figure 4.8 Reflection coefficient magnitude versus frequency for different number of sections of the matching transformer: (a) Binomial type, and (b) Tchebyscheff type.

4.2.3 Tapered Impedance Matching Transformers

In the previous section we have seen how two transmission media (waveguide or microstrip line) of different characteristic impedances can be matched by interposing a number of discrete steps to change the impedance level. As the number of discrete sections increases, the step changes in characteristic impedance between the sections of the transformer become smaller, and we practically approach a continuously tapered impedance matching transformer as shown in Figure 4.9. By changing the type of taper, we can obtain different pass-band characteristics [1].

It may be observed from Figure 4.9 that the line has a continuous tapering of the transverse dimension in the longitudinal direction z and thus the impedance of the line is a function of z. Assuming that the taper is gradual, it can be approximated by a large number of elementary sections dz with incremental impedance change dZ and a differential reflection coefficient:

$$d\Gamma = \frac{(Z+dZ)-Z}{(Z+dZ)+Z} = \frac{dZ}{2Z+dZ} \approx \frac{dZ}{2Z}$$

$$= \frac{1}{2}d(\ln Z) = \frac{1}{2}\frac{d}{dz}(\ln Z)dz \tag{4.32a}$$

Since dZ << Z

The input reflection coefficient (the total reflection coefficient at $z = 0$) for a tapered line of length L can now be found (following the theory of small reflections) by summing all partial reflections with their appropriate phase shifts as:

$$\Gamma_i = \frac{1}{2}\int_0^L \frac{d}{dz}(\ln Z)dz.e^{-j2\beta z} \tag{4.32b}$$

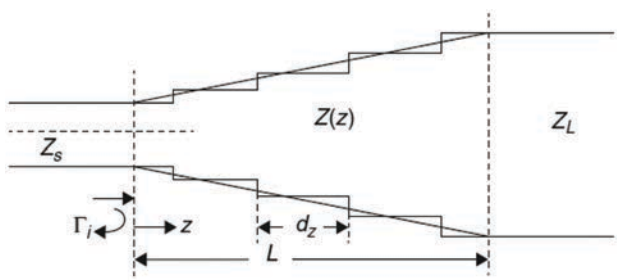


Figure 4.9 Tapered transmission line matching section.

Thus, if Z(z) is known, Γ_i can be found as a function of frequency. Alternatively, for a given response characteristic of Γ_i, $Z(z)$ can be determined. The latter approach is difficult and is generally avoided in practical design. Now we shall consider three special cases of $Z(z)$ impedance tapers and evaluate the resulting responses.

4.2.3.1 Exponential Taper

For exponential taper, $Z(z)$ varies exponentially. We can write:

$$Z(z) = Z_0 e^{pz}, \quad \text{for } 0 < z < L \tag{4.33a}$$

Where the exponential taper control parameter p can be evaluated from:

$$p = \frac{1}{L}\ln\left(\frac{Z_L}{Z_0}\right) \tag{4.33b}$$

In this impedance matching Z_0 is matched to Z_L using exponential impedance matching transformer of length L (inset in Figure 4.10).

Now following Eq. (4.32b) and using Eqs. (4.33a and 4.33b) we have:

$$\Gamma_i = \frac{1}{2}\int_0^L e^{-j2\beta z}\frac{d}{dz}(\ln e^{pz})dz = \frac{\ln(Z_L/Z_0)}{2l}$$
$$\int_0^L e^{-j2\beta z}dz = \frac{\ln(Z_L/Z_0)}{2}\left(\frac{\sin\beta L}{\beta L}\right)e^{-j\beta l} \tag{4.33c}$$

In this derivation it is assumed that $\beta = 2\pi/\lambda$ is independent of z (considering the TEM case).

The magnitude of the reflection coefficient in Eq. (4.33c) has been sketched in Figure 4.10. It may be noted that the peaks in $|\Gamma_i|$ decrease with increasing length, as expected, and that the length should be greater than $\lambda/2$ ($\beta L > \pi$) to minimize the mismatch at low frequencies.

4.2.3.2 Taper with Triangular Distribution

For triangular taper $Z(z)$ is given by:

$$Z(z) = Z_0 e^{2(z/L)^2\ln(Z_L/Z_0)}, \quad 0 \le z \le L/2$$
$$= Z_0 e^{[4z/L - 2z^2/L^2 - 1](\ln Z_L/Z_0)} \quad L/2 \le z \le L \tag{4.34a}$$

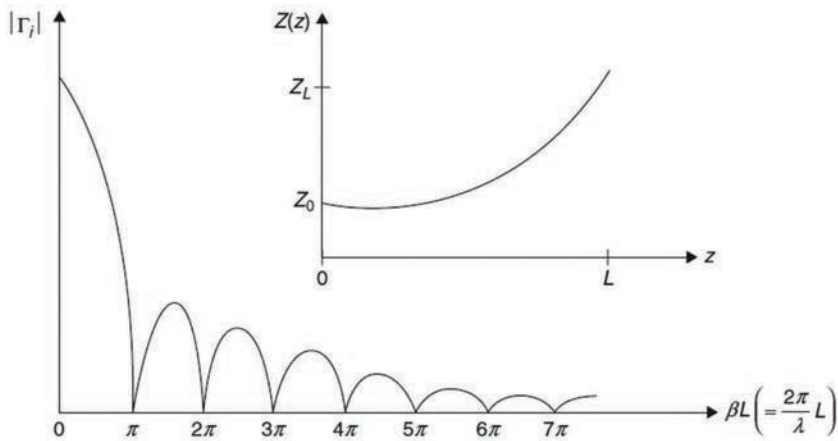


Figure 4.10 Exponential taper with reflection coefficient magnitude response (impedance response is in the inset).

Using Eq. (4.34a) we can write:

$$\frac{d(\ln Z/Z_0)}{dz} = (4z/L^2)\ln(Z_L/Z_0) \quad 0 \le z \le L/2$$

$$= (4/(L-4z)/L^2)\ln(Z_L/Z_0) \quad L/2 \le z \le L \tag{4.34b}$$

Substituting (4.34a) into Eq. (4.32b) and integrating we get:

$$\Gamma_i = \frac{1}{2}e^{-j\beta l}\ln\left(\frac{Z_L}{Z_0}\right)\left[\frac{\sin(\beta L/2)}{(\beta L/2)}\right]^2 \tag{4.34c}$$

The magnitude of $|\Gamma_i|$ has been plotted in Figure 4.11. It may be noted that compared to the exponential taper, the percentage of the minor lobe amplitude to that of the major lobe is much smaller for this taper. Although this small value of reflection coefficient occurs for taper length of twice that required for the exponential taper (first null for triangular taper occurs at $\beta L = 2\pi$, whereas for the exponential taper it occurs at $\beta L = \pi$); however, this taper will be preferable in many practical situations as much smaller values of $|\Gamma_i|$ can be obtained for all frequencies in the pass-band.

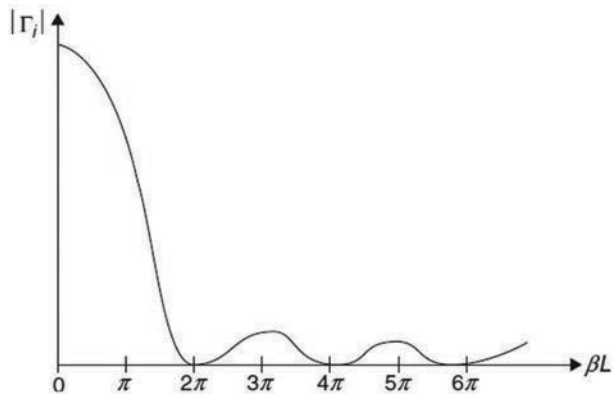


Figure 4.11 Triangular taper with reflection coefficient magnitude response.

4.2.3.3 Klopfenstein Taper

Looking into the infinite number of possibilities (at least theoretically) of designing tapered impedance transformer by choosing the impedance function $Z(z)$ with relative advantages and disadvantages as far as the performance characteristics and size of the transformer are concerned; a question naturally arises whether we can design a transformer that gives minimum reflection in the pass-band yet with a minimum size—whether an optimized design is possible. This end has been realized with *Klopfenstein taper* transformer design in which the shortest length between two impedances to be matched with continuous tapering sections can be realized with lowest possible reflection coefficient over the desired pass-band. In other words, from a designer's point of view, with a given maximum tolerable reflection coefficient specification in the pass-band, the length of the matching transformer section is shortest for the *Klopfenstein taper* when compared with other taper distribution functions like exponential/triangular taper.

With a pre-designated minimum reflection coefficient (maximum allowable ripple) in the pass-band Γ_m, taper length L, input impedance Z_0, and load impedance Z_L, the logarithm of the characteristic impedance variation for the Klopfenstein taper is given by [4]:

$$\ln Z(z) = 0.5 \ln Z_0 Z_L + \Gamma_m A^2 \phi\left(\frac{2z}{L} - 1, A\right) \quad 0 \le z \le L \tag{4.35a}$$

where $\Gamma_m = \dfrac{\Gamma_0}{\cosh A}$, $A = \cosh^{-1}\left(\dfrac{\Gamma_0}{\Gamma_m}\right)$, and $\Gamma_0 = 0.5 \ln\left(\dfrac{Z_L}{Z_0}\right)$.

The function $\phi(z, A)$ is defined as:

$$\phi(z, A) = -\phi(-x, A) = \int_0^z \frac{I_1(A\sqrt{1-y^2})}{A\sqrt{1-y^2}} dy \quad \text{for } |z| \le 1 \tag{4.35b}$$

where I_1 is the modified Bessel function which has the following special values:

$$\phi(0, A) = 0,\ \phi(z, 0) = 0.5z, \text{ and } \phi(1, A) = (\cosh A - 1)/A^2$$

The resulting input reflection coefficient is given by:

$$\Gamma_i = \Gamma_0 e^{-j\beta L} \frac{\cos\sqrt{(\beta L)^2 - A^2}}{\cosh A} \text{ for } \beta L > A$$

$$= \Gamma_0 e^{-j\beta L} \frac{\cos\sqrt{A^2 - (\beta L)^2}}{\cosh A} \text{ for } \beta L < A \tag{4.35c}$$

The comparative performance characteristics and design of exponential taper, triangular taper, and Klopfenstein taper will be clear in the solved example given below.

Example 4.5 Design an exponential taper, triangular taper, and a Klopfenstein taper in a waveguide system with maximum allowable ripple in the pass-band to be Γ_m = 0.025 to match a reduced height waveguide with characteristic impedance 58 Ω to a full-height waveguide having characteristic impedance 222 Ω (typical for an X-band waveguide with λ_0 = 3 cm). Show the plot of impedance variation with normalized length of the matching waveguide and the variation of the resulting reflection coefficient magnitude with the electrical length βl of the guide. Design all three types of tapered transformers with 3D electromagnetic field simulator including performance characterization.

Solution:

Exponential taper: From Eq. (4.33a) the impedance variation is given by:

$$Z(z) = Z_0 e^{pz}, \text{ for } 0 < z < L, \text{ with } p = \frac{1}{L}\ln\left(\frac{Z_L}{Z_0}\right) = 1.342/L\,.$$

The reflection coefficient response is given by Eq. (4.33c) as:

$$\left|\Gamma_i(\theta)\right| = \frac{\ln(Z_L/Z_0)}{2}\left(\frac{\sin\beta L}{\beta L}\right)$$

Triangular Taper: From Eq. (4.34a) the impedance variation is given by:

$$Z(z) = Z_0 e^{2(z/L)^2 \ln(Z_L/Z_0)} \quad 0 \le z \le L/2$$

$$= Z_0 e^{[4z/L - 2z^2/L^2 - 1](\ln Z_L/Z_0)} \quad L/2 \le z \le L$$

With $Z_0 = 58\ \Omega$ and $Z_L = 222\ \Omega$. The reflection coefficient response is given by Eq. (4.34c) as:

$$\left|\Gamma_i(\theta)\right| = \frac{1}{2}\ln\left(\frac{Z_L}{Z_0}\right)\left[\frac{\sin(\beta L/2)}{(\beta L/2)}\right]^2$$

Klopfenstein Taper: We have in this case:

$$\Gamma_0 = 0.5\ln\left(\frac{Z_L}{Z_0}\right) = 0.671, \text{ and}$$
$$A = \cosh^{-1}\left(\frac{\Gamma_0}{\Gamma_m}\right) = \cosh^{-1}\left(\frac{0.671}{0.025}\right) = 3.98$$

The impedance taper characteristics need to be evaluated numerically using Eq. (4.35a). The reflection coefficient response being determined from Eq. (4.35c):

$$\Gamma_i(\theta) = \Gamma_0 e^{-j\beta L}\frac{\cos\sqrt{(\beta L)^2 - A^2}}{\cosh A}$$

The pass-band for *Klopfenstein taper* is defined as $\beta L > A = 3.98 = 1.267\pi$. The MATLAB plotted curves showing impedance variation Z(z) with z/L and resulting reflection coefficient (magnitude) $|\Gamma|$ with L for all the three types of taper transformers have been shown in Figures 4.12(a) and 4.12(b), respectively (see Appendix II for the MATLAB program). It may be observed that the Klopfenstein taper gives the desired response with $|\Gamma| \le \Gamma_m = 0.025$ for electrical length $\beta L = 1.3\pi$

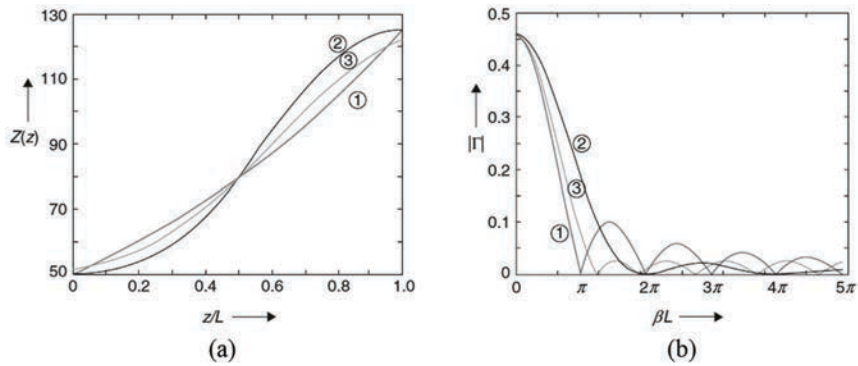


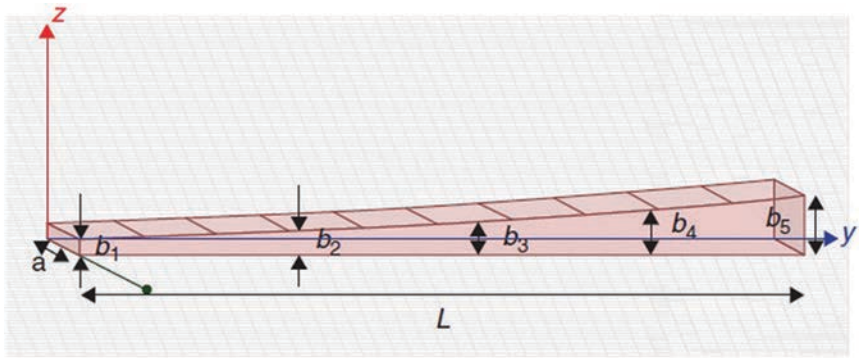
Figure 4.12 (a) Impedance variation with normalized taper length for (1) Exponential, (2) Triangular, and (3) Klopfenstein taper, and (b) Reflection coefficient (magnitude) variation with frequency for those three types of tapers.

Table 4.1 Data for drawing the exponential, triangular, and Klopfenstein taper structures

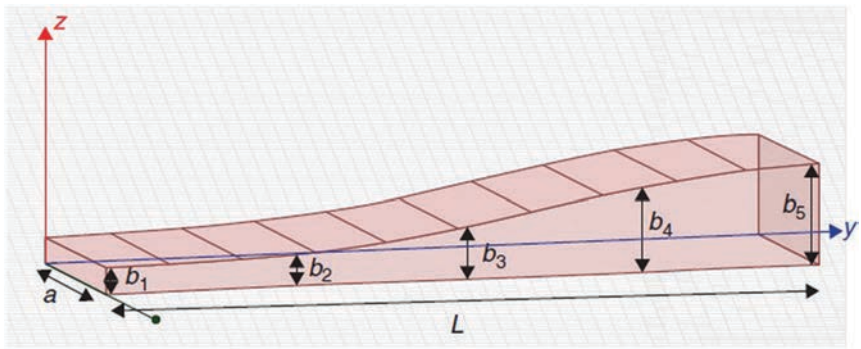
Parameters (a, b, and L in mm; Z in Ω)	*Exponential*	*Triangular*	*Klopfenstein*
a	22.86 mm	22.86 mm	22.86 mm
L	126.3 mm	74.5 mm	23.68 mm
b_1, L_1 (Z_1)	2.65 mm, 0.0 mm (58 Ω)	2.65 mm, 0.0 mm (58 Ω)	2.65mm, 0.0 mm (58 Ω)
b_2, L_2 (Z_2)	3.67 mm, 30.618 mm (80.31 Ω)	3.10 mm, 18.06 mm (67.91 Ω)	3.10 mm, 5.741 mm (67.83 Ω)
b_3, L_3 (Z_3)	5.15 mm, 62.512 mm (112.7 Ω)	5.12 mm, 36.87 mm (112 Ω)	5.13mm, 11.72 mm (112.2 Ω)
b_4, L_4 (Z_4)	7.32 mm, 95.682 mm (160.3 Ω)	8.56 mm, 55.68 mm (187.1 Ω)	8.54 mm, 17.7 mm (186.8 Ω)
b_5, L_5 (Z_5)	10.16 mm, 126.3 mm (222 Ω)	10.16 mm, 74.5 mm (222 Ω)	10.16 mm, 23.68 mm (222 Ω)

($\approx 1.267\pi$), which is shorter than either the triangular or exponential taper responses. The response of Klopfenstein taper in the pass-band resembles the equal-ripple characteristics similar to Tchebysheff-type impedance matching transformer.

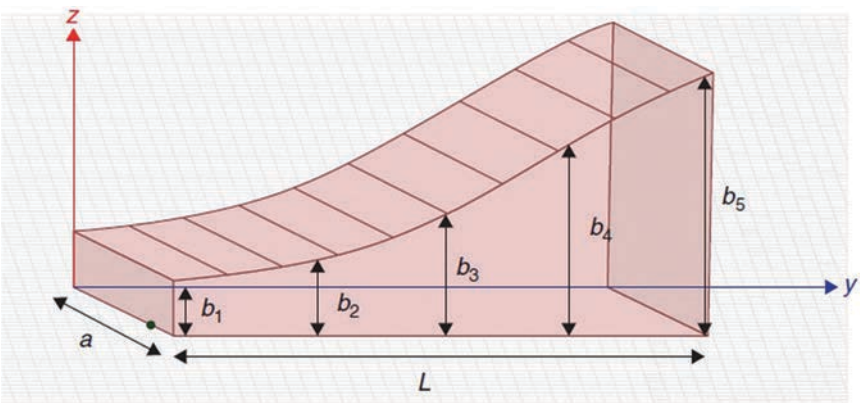
The curve in Figure 4.12(a) may be used to design the transformers to evaluate their physical dimensions using Table 4.1 for the three types of tapered transformers.



(a)(i)

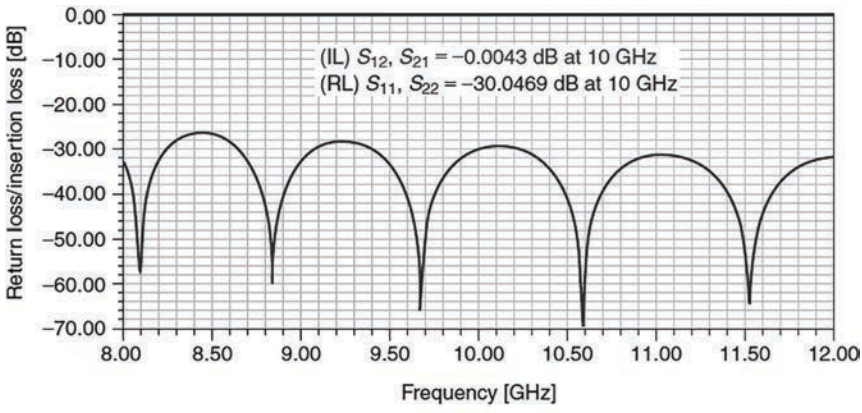


(a)(ii)

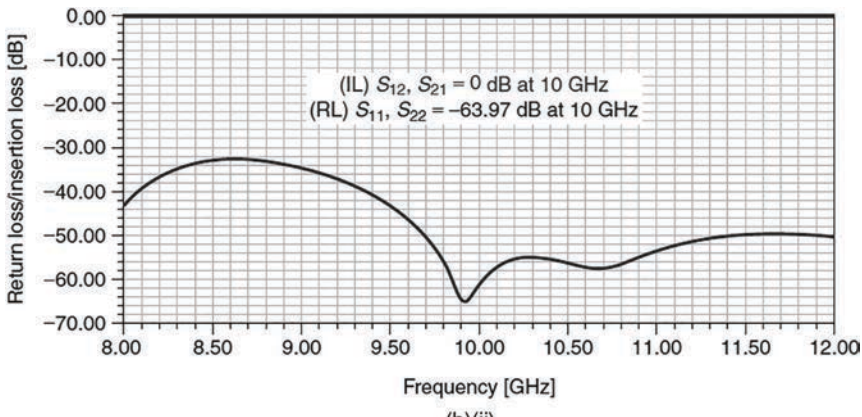


(a)(iii)

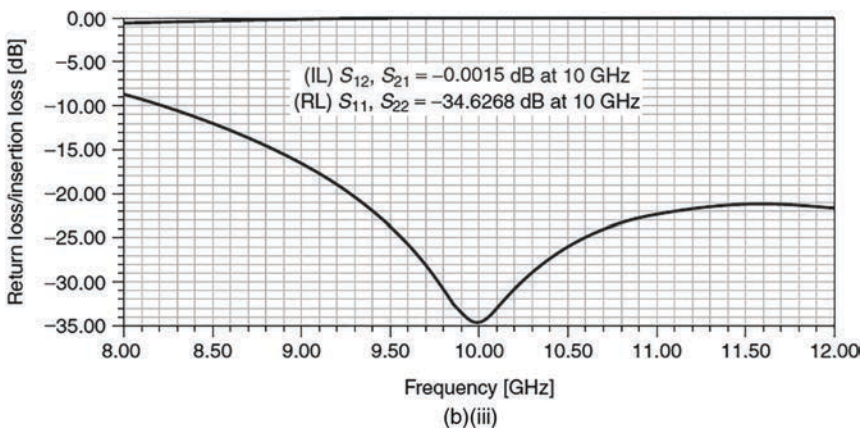
Figure 4.13 (a) Contours and, (b) Performance characteristics of: (i) Exponential taper, (ii) Triangular taper, and (iii) Klopfenstein taper.



(b)(i)



(b)(ii)



(b)(iii)

Figure 4.13 (Continued)

Data given in Table 4.1 is used to draw the transformer contours in 3D electromagnetic field simulator (Figure 4.13(a)) and to characterize their performance in terms of the simulated return loss and insertion loss (see Figure 4.13(b)). The values of different b and L dimensions are obtained from the table of data generated while drawing Figure 4.12(a) with MATLAB program.

From Figure 4.13(b) it may be observed that for exponential taper the return loss variation repeats every 0.95 GHz which is expected to take place every π radian (equivalent to every 1.18 GHz) if the plot was $|\Gamma|$ vs βl (Figures 4.9 and 4.12(b). But for triangular taper this repetition is seen to be almost at double that frequency as expected.

4.3 IMPEDANCE MATCHING WITH LUMPED-COMPONENT/ PLANER DISTRIBUTED-LINE REACTIVE ELEMENTS

A simple yet practically useful technique of impedance matching in certain applications in microwave circuit design is to use either lumped-component or planer distributed-line reactive elements (inductance and capacitance). In fact, in the design of microwave amplifiers the source and load impedances are to be matched, respectively, with the input and output impedance of the active device and this is done by such reactive element impedance matching. When two reactive elements are used to realize the impedance matching, the network is called L-section network; while three element matching networks are either T or Π-section type. The names are descriptive of the configuration (L, T, Π) formed by the reactive elements. Lumped-element inductance and capacitance may be used with limited values and at lower frequencies; however, for MMIC applications planer distributed-line (microstrip line) is used to realize the reactive elements. We would discuss here the L-section matching technique with Z-Y Smith chart, i.e., *Smith and Jones chart*.

With two different reactive elements L and C there are 2^2 (four) possible configurations for the L-section network as shown in Figure 4.14.

We shall now discuss with the help of some examples how Smith and Jones chart (Z-Y Smith chart) can be used to design L-section matching networks.

Example 4.6 Illustrate with the use of Smith and Jones chart the effect of adding series/shunt inductance/capacitance to impedance/admittance in the following cases:

i. A series inductor L ($z_L = +j0.7$) to an impedance $z = (0.2 - j0.2)$.
ii. A series capacitor C ($z_C = -j0.7$) to an impedance $z = (0.2 - j0.2)$.
iii. A shunt inductor L ($y_L = -j2.6$) to an admittance $y = (1.7 + j1.7)$.
iv. A shunt capacitor C ($y_C = +j3.6$) to an admittance $y = (1.7 + j1.7)$.

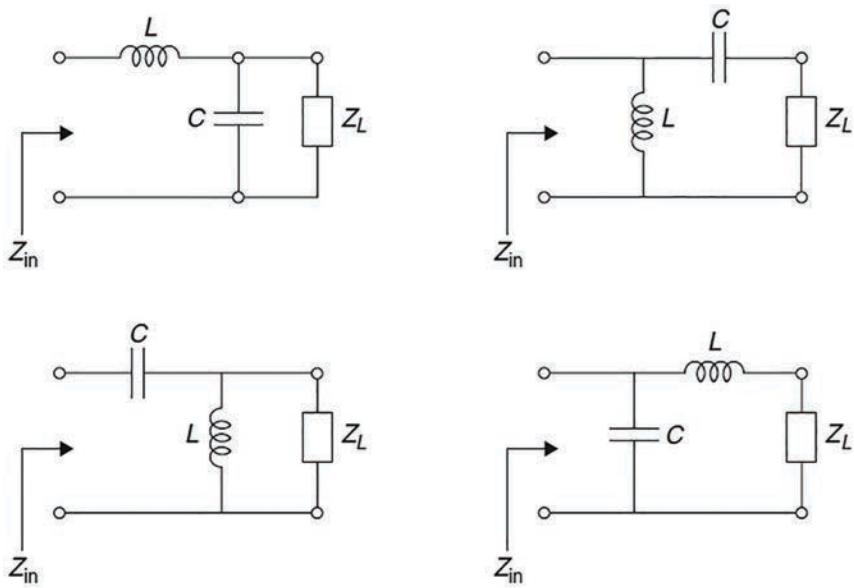


Figure 4.14 Four possible configurations of L-section network for impedance matching.

Solution:

a. When a *series inductor* with normalized reactance +j0.7 is added to an impedance (0.2 – j0.2), we need to move in a constant-resistance circle in the Smith and Jones chart starting from the initial reactance value of –j0.2 to the final value of +j0.5, which makes *clockwise rotation* as shown by curve "a" in Figue 4.15.
b. When a *series capacitor* with normalized reactance –j0.7 is added to an impedance (0.2 – j0.2), we need to move in a constant-resistance circle in the Smith and Jones chart starting from the initial reactance value of –j0.2 to the final value of –j0.9, which makes *anticlockwise rotation* as shown by curve "b" in Figure 4.15.
c. When a *shunt inductor* with normalized susceptance –j2.6 is added to an admittance (1.7 + j1.7), we need to move in a constant-conductance circle in the Smith and Jones chart starting from the initial susceptance value of +j1.7 to the final value of –j0.9, which makes *anticlockwise rotation* as shown by curve "c" in Figure 4.15.

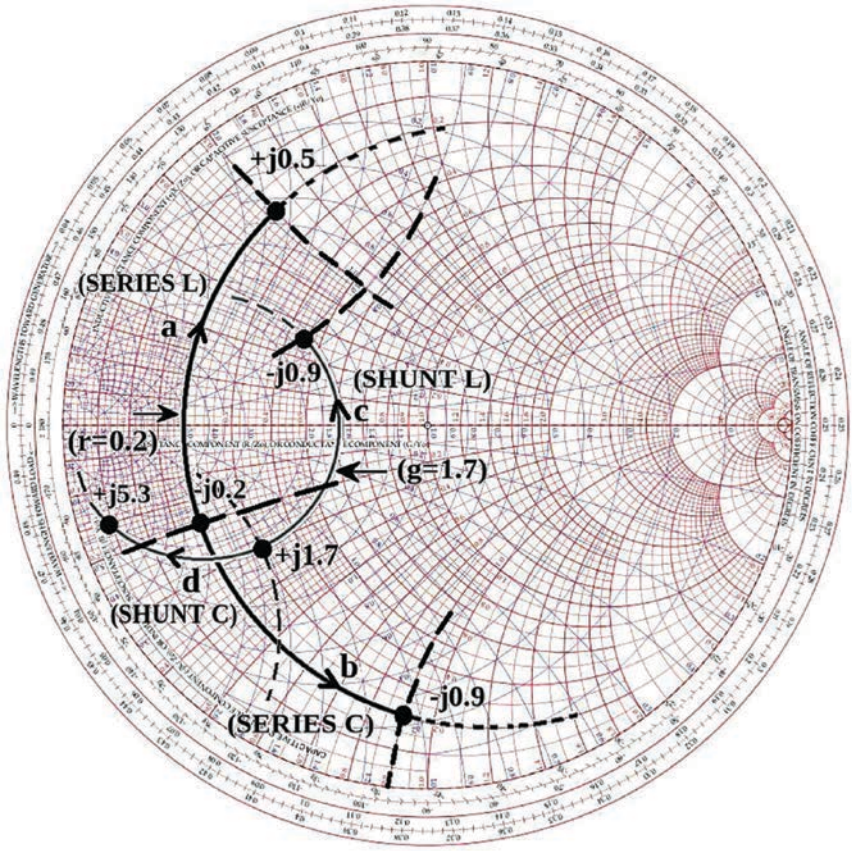


Figure 4.15 Smith and Jones chart showing the impedance and admittance locus for Example 4.6.

d. When a *shunt capacitor* with normalized susceptance +j3.6 is added to an admittance (1.7 + j1.7), we need to move in a constant-conductance circle in the Smith and Jones chart starting from the initial susceptance value of +j1.7 to the final value of +j5.3, which makes *clockwise rotation* as shown by curve "d" in Figure 4.15.

Example 4.7 Design two matching networks at 1.0 GHz one with series L and shunt C while the other one with series C and shunt L in order to match a load impedance of (15 + j15) Ω to a 75 Ω line. Use Smith and Jones chart to solve the problem.

Solution

The normalized load impedance is: $z_L = (0.2 + j0.2)$, which is first located in the Smith and Jones chart marked by point A (Figure 4.16).

Matching with Series L and Shunt C: When we add a *series L* with normalized inductive reactance of +j2, say, we need to move along the constant-resistance circle of 0.2 in *clockwise direction* to reach at an impedance of (0.2 + j0.4), which is incidentally an admittance of (1 – j2) located at point B (Figure 4.16). Now the final matching of the load would need a *shunt capacitive* susceptance of +j2 (capacitive reactance of 1/j2 = –j5), moving along the unit constant-conductance circle in *clockwise direction* to reach the origin of the Smith and

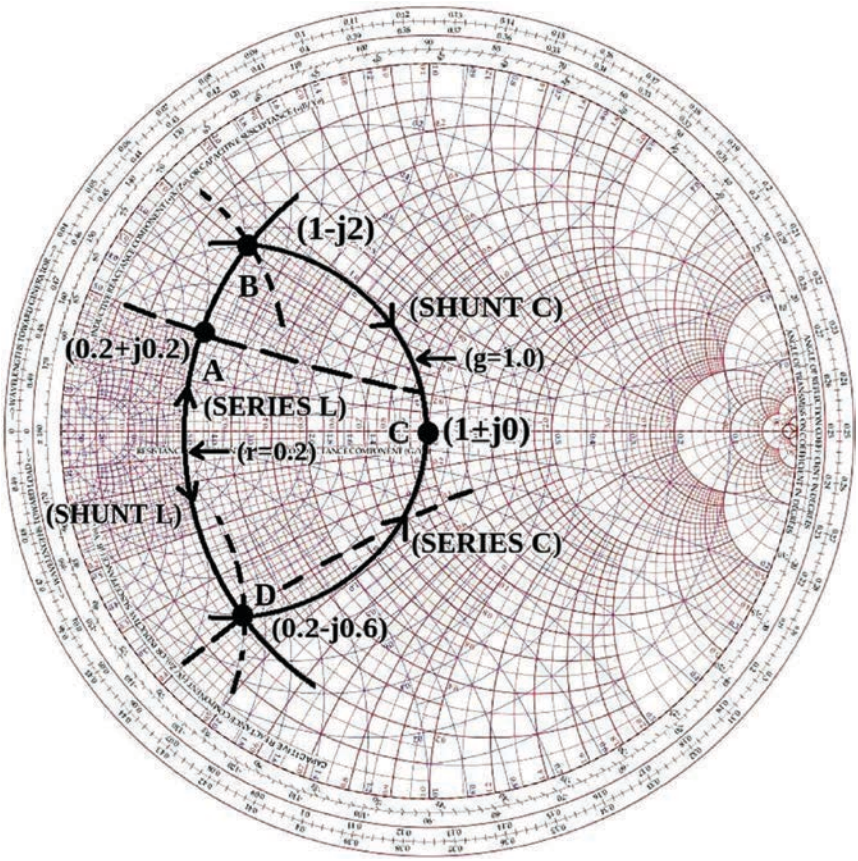


Figure 4.16 Smith and Jones chart showing the impedance and admittance locus of Example 4.7.

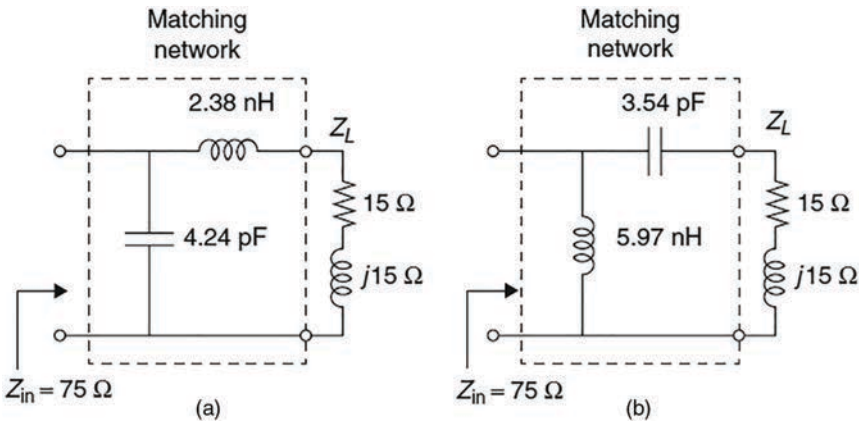


Figure 4.17 Matching network for Example 4.7: (a) With series L and shunt C, and (b) With series C and shunt L.

Jones chart, which is the location for perfect matching. Thus, with a series inductance L = 0.2 × 75/2π × 10^9 = 2.38 nH and shunt capacitance C = 1/0.5 × 75 × 2π × 10^9 = 4.24 PF the load impedance can be matched with the 75 Ω line—the matching network is shown in Figure 4.17(a).

Matching with Series C and Shunt L: However, when we add a *series* C with normalized capacitive reactance of –j0.6, say, we need to move along the constant-resistance circle of 0.2 in *anticlockwise direction* to reach at an impedance of (0.2 – j0.4), which is incidentally an admittance of (1 + j2) located at point D (Figure 4.16). Now the final matching of the load would need a *shunt inductive* susceptance of –j2 (inductive reactance of 1/–j2 = +j0.5), moving along the unit constant conductance circle in *anticlockwise direction* to reach the origin of the Smith and Jones chart, which is the location of perfect matching. Thus, with a series capacitance C = 1/0.6 × 75 × 2π × 10^9 = 3.54 PF and shunt inductance L = 0.5 × 75/2π × 10^9 = 5.97 nH, the load impedance can be matched with the 75 Ω line (Figure 4.17(b)).

(Note: In both cases of matching, our aim of choosing the initial value of the first reactive element was to move finally into a unit conductance circle to affect the final matching with the second reactive element.)

REFERENCES

[1] Collin R E, *Foundations for Microwave Engineering*, Second Edition, John Wiley & Sons, New Jersey, 2007.
[2] Altman J L, *Microwave Circuits*, D Van Nostrand Company Inc., Princeton, NJ, 1964.
[3] Das A, and Das S K, *Microwave Engineering*, Second Edition, Tata McGraw-Hill Publishing Company Limited, New Delhi, India, 2009.
[4] Klopfenstein R W, A Transmission Line Taper of Improved Design, *Proc. IRE,* Vol. 44, pp. 31–35, 1956.

Chapter 5

Microwave Filter Design

Low-Pass, Band-Pass, Band-Stop, and Metamaterial-Inspired Filter

We should respond to our conscience
which acts as a filter to lead our path towards goodness

5.1 INTRODUCTION

A filter is a two-port network which is ideally required to provide perfect transmission for all frequencies in certain *pass-band* regions and exhibit infinite attenuation in the *stop-band* region. Such ideal characteristics cannot be realized in practice, and the goal of the filter designer is to approximate the ideal requirements to within an acceptable tolerance. Filters are usually categorised by their frequency response characteristics—low-pass, high-pass, band-pass, and band-stop/band-reject (Figure 5.1). The application of filters can be found virtually in any type of communication, radar, and test/measurement systems.

At low frequencies, inductors and capacitors have simple frequency characteristics, and a very general and complete synthesis procedure has been developed for the design of filters at low frequency using them. However, filter design at RF/microwave frequencies (where distributed parameter elements must be used) is much more complicated due to complex frequency behaviour of RF/microwave circuit elements. In spite of these added complications at RF/microwave frequencies, a number of useful techniques have been developed for the design of RF/microwave filters. The design of a narrowband filter (the term *narrowband* is generally used to indicate bandwidths of less than 10%, while *wideband* usually denotes bandwidths in excess of 40%) at RF/microwave frequencies is particularly straight forward since many RF/microwave elements will have frequency characteristics essentially like those of an ideal inductive or capacitive reactance over a limited frequency range. In such cases, a low frequency prototype filter may be used as a model. The RF/microwave filter is realized by replacing all inductors and capacitors by suitable RF/microwave circuit elements (like lengths of transmission line sections suitably terminated at the load end) that have similar frequency characteristics over the frequency range of interest. During the past few years, lumped inductors and

DOI: 10.1201/9781003729969-5

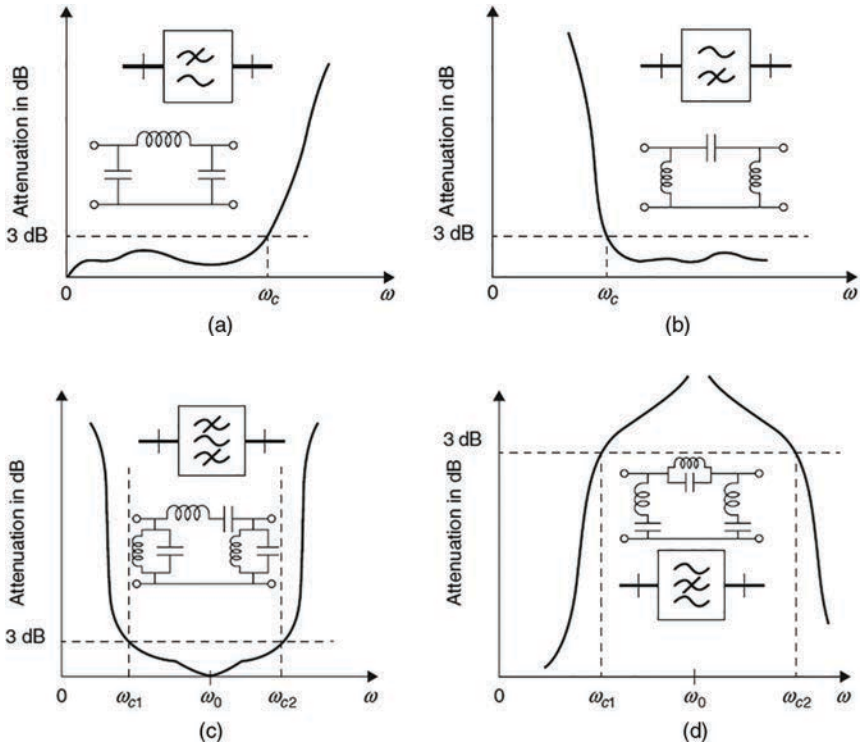


Figure 5.1 Filter characteristics with equivalent electrical circuit: (a) Low-pass filter, (b) High-pass filter, (c) Band-pass filter, and (d) Band-stop filter.

capacitors have been fabricated and used at RF/microwave frequencies (but with limited values and up to 60 GHz or so). If lumped elements are used for the filter circuit at RF/microwave frequencies, the RF/microwave filter design becomes identical to the low frequency filter design.

The RF/microwave filter theory began with the pioneering work by Darlington, Fano, Lawson, and Richards just before World War II. The image parameter method of filter design was developed in the late 1930s and was useful for low frequency filters in radio and telephony. In the early 1950s, G Matthaei, L Young, E Jones, S Cohn, and others of the Stanford Research Institute became very active in RF/microwave filter and coupler development. A voluminous handbook on filters and couplers resulted from their work and is still a valuable reference [1]. Most RF/microwave filter designs these days is done with computer-aided design (CAD) packages based on electronic design automation (EDA) tool.

5.2 INSERTION LOSS METHOD FOR LOW-PASS PROTOTYPE DESIGN

5.2.1 Introduction

A perfect filter should have zero insertion loss in the pass-band, infinite attenuation in the stop-band, and a linear phase response (to avoid signal distortion) in the pass-band. However, such filters do not exist in practice. Thus, compromise must be made; and here is the art of filter design.

The design of RF/microwave filter with insertion loss method is done in following steps:

i) A low-pass prototype filter is first designed in terms of normalized impedance and frequency with desired pass-band characteristics.
ii) This prototype network is then transformed to the required filter type (low-pass, high-pass, band-pass, and band-stop) with proper scaling of the prototype circuit elements in terms of impedance and frequency.
iii) Finally, the filter network is realized in RF/microwave form with distributed circuit elements which essentially uses sections of terminated (open/short) RF/microwave transmission line of suitable length representing the desired reactance.

In the insertion loss method, a filter response is defined by its insertion loss, or power loss ratio, P_{LR}:

P_{LR} = (Power available from source)/(Power delivered to the load)

$$= \frac{P_i}{P_L} = \frac{P_i}{P_i - P_r} = \frac{1}{1 - |\Gamma(\omega)|^2} \tag{5.1a}$$

where the difference of the power available from the source (P_i) and the power reflected from the filter has been taken as the power delivered to the load (P_L) with the assumption that the filter is lossless.

P_{LR} expressed in dB is *insertion loss* (IL) given by

$$\text{IL} = 10 \log P_{LR} \tag{5.1b}$$

It may be noted that $|\Gamma(\omega)|^2$ is an even function of ω and can therefore be expressed as polynomial in ω^2. Thus, we can write

$$|\Gamma(\omega)|^2 = \frac{X(\omega^2)}{X(\omega^2) + Y(\omega^2)} \tag{5.2}$$

Substituting this in Equation (5.1a), we have

$$P_{\text{LR}} = 1 + \frac{X(\omega^2)}{Y(\omega^2)} \tag{5.3}$$

where X and Y are real polynomials in ω^2. Thus, for a filter to be physically realizable, its power loss ratio must be of the form depicted in Equation (5.3).

In principle, there can be an unlimited number of physical networks for which power loss ratio can be specified but most of them would be very complex to realize in practice. Most useful power loss ratios for RF/microwave filter design are those that give a *maximally flat* pass-band response and those that give an *equal-ripple* response in the pass-band. Such pass-band response characteristics correspond to those of the binomial (maximally flat) and Tchebysheff (equal-ripple) multi-section quarter-wave transformers discussed in Chapter 4.

Case I Maximally Flat/Butterworth Filter

The power loss ratio for a maximally flat/Butterworth filter is specified by

$$P_{\text{LR}} = 1 + A_{p^2}(\Omega)^{2n} \tag{5.4}$$

where $\Omega = (\omega/\omega_c)$ represents the normalized frequency variable of the prototype filter and n is the order of the filter; ω and ω_c being the running frequency variable and cut-off frequency of the actual filter. Referring to Figure 5.2a, it may be observed that the pass-band extends from $\Omega = 0$ ($\omega = 0$) to $\Omega = \Omega_p$ ($\omega = \omega_c$) with the maximum value of P_{LR} in the pass-band being $(1 + A_p^2)$; where A_p determines the attenuation tolerance in the pass-band. Normally, A_p is chosen to be equal to 1 when the band-edge power loss ratio $(1 + A_p^2)$ happens to be equal to –3 dB (Figure 5.2a). For $\Omega > 1$, the attenuation increases without bound with the increase of Ω and the rate of increase is more (sharper) with order n of the filter. While for $\Omega < 1$, P_{LR} shows progressive closeness to the Ω axis as n increases. For $\Omega >> 1$ ($\omega >> \omega_c$), $P_{LR} \approx A_p^2\left(\omega/\omega_c\right)$, which indicates that the insertion loss incurred by the filter far-off the cut-off frequency increases at the rate of $6n$ dB/octave (or $20n$ dB/decade).

This type of filter is known as maximally flat filter because its power loss ratio given by Equation (5.4) has the maximum number of $(2n - 1)$ zero derivatives at $\Omega = 0$; approximating the ideal low-pass filter response at $\Omega = 0$ but with gradual deviation from that as Ω approaches the cut-off frequency Ω_p (Figure 5.2a).

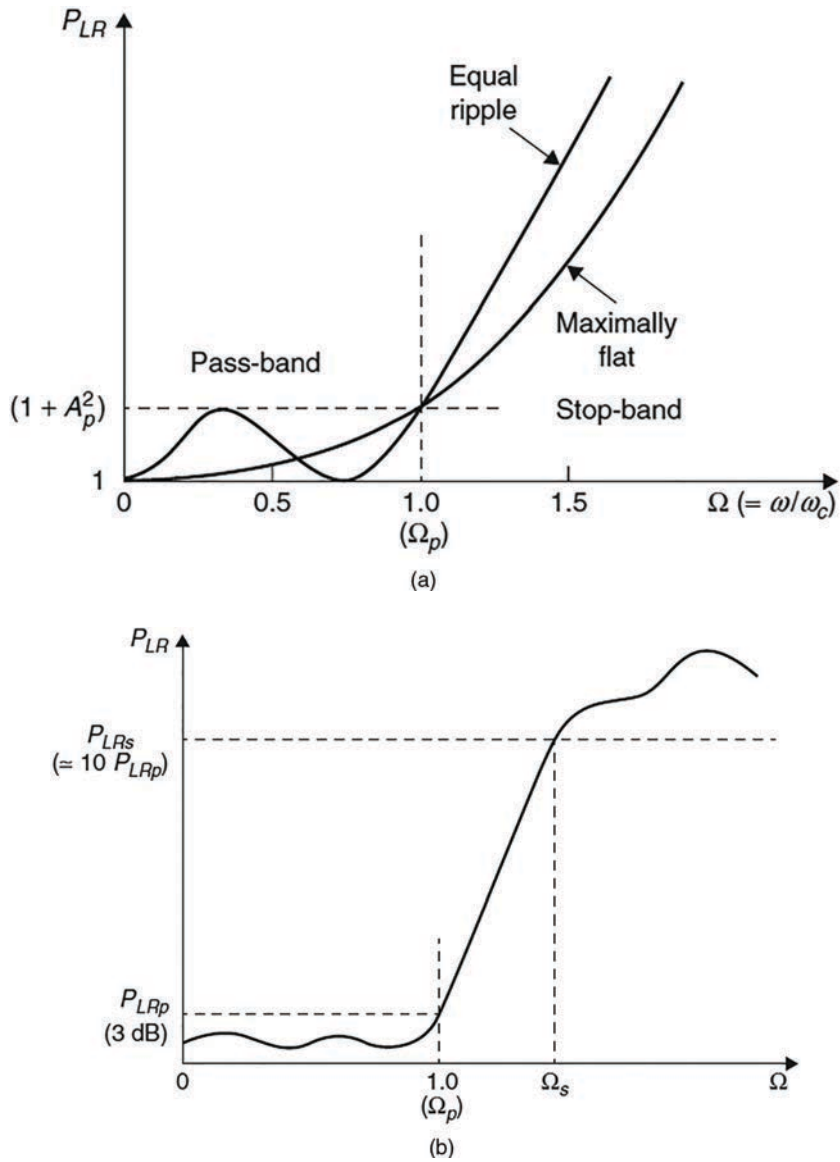


Figure 5.2 (a) Maximally flat and equal-ripple low-pass filter response, (b) Low-pas filter characteristics showing the pass-band edge frequency (Ω_p) and stop-band edge frequency (Ω_s) and corresponding power loss ratios ($P_{LRp,\,s}$), and (c) Elliptic function low-pass filter.

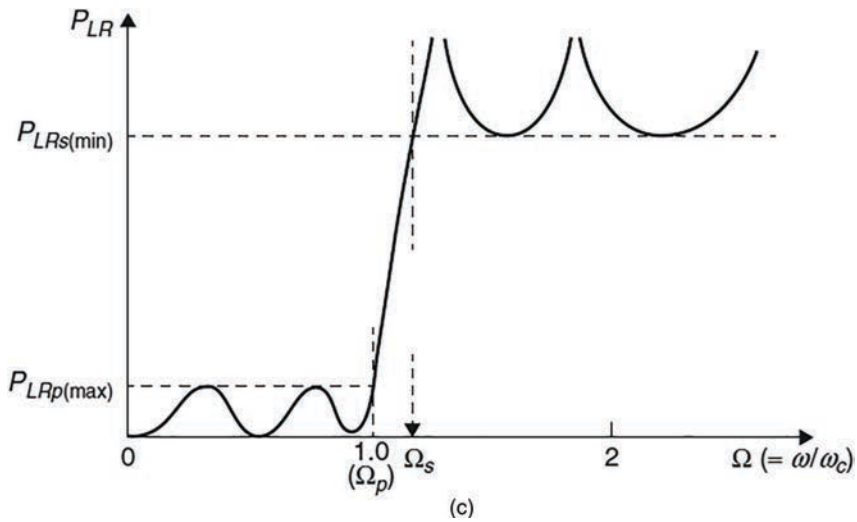


Figure 5.2 (Continued)

Case II Equal-Ripple/Tchebyscheff Filter

The power loss ratio of equal-ripple/Tchebyscheff filter (that has some tolerable ripples of equal amplitude unlike the maximally flat one) is represented by

$$P_{LR} = 1 + A_p^2 T_n^2(\Omega) \tag{5.5}$$

where $T_n(\Omega)$ is the Tchebyscheff polynomial of first kind having order n (see Section 4.4.2), which is given below

$$T_n(\Omega) = \cos(n \cos^{-1} \Omega) \text{ for } |\Omega| \leq 1$$

$$= \cosh(n \cosh^{-1} \Omega) \text{ for } |\Omega| \geq 1 \tag{5.6}$$

The pass-band will have equal ripples of amplitude $(1 + A_p^2)$ (Figure 5.2a), since $T_n(x)$ varies between ± 1, for $x \leq 1$. For large values of x, $T_n(x) \approx \frac{1}{2}(2x)^n$, so that for $\Omega >> 1$ ($\omega >> \omega_c$), $P_{LR} \approx (A_p^2/4)[2(\omega/\omega_c)^{2n}]$; i.e., insertion loss incurred by this filter also increases at the rate of $6n$ dB/octave (or $20n$ dB/decade) far-off the cut-off frequency, but it is $2^{(2n-2)}$ times larger than that for maximally flat filter. Thus, equal-ripple filter will have sharper rate of rise above cut-off with better stop-band characteristics (approximately a

"brick-wall" characteristics outside the pass-band) compared to maximally flat filter, though with a penalty of ripple within the pass-band. However, for a specified rate of increase of the power loss ratio beyond the cut-off frequency, the equal-ripple characteristics can provide smallest pass-band tolerance (A_p). In this sense, the equal-ripple filter represents an optimum design.

It may further be observed that for a given insertion loss, the equal-ripple filter needs fewer elements (number of elements is equal to the order, n, of the filter; as will be seen shortly) than maximally flat filter which may be established as follows. Let the power loss ratio at the pass-band edge frequency Ω_p (=1) be P_{LRp} while the same at stop-band edge frequency Ω_s be P_{LRs} (Figure 5.2b); then insertion loss at Ω_p and Ω_s are respectively: $IL_p = 10 \log P_{LRp}$ and $IL_s = \log P_{LRs}$ (since $IL_s = 10\ IL_p$ for a practical design). Then for maximally flat filter, it can be derived from Equation (5.4) that

$$n \geq \frac{\log[(P_{LRs} - 1)/(P_{LRp} - 1)]^{1/2}}{\log(\Omega_s/\Omega_p)} \tag{5.7a}$$

While for an equal-ripple filter, it can be derived by using Equation (5.5) and the expression for $T_n(x)$ from Section (4.4.2.2) that

$$n \geq \frac{\cosh^{-1}[(P_{LRs} - 1)/(P_{LRp} - 1)]^{1/2}}{\cosh^{-1}(\Omega_s/\Omega_p)} \tag{5.7b}$$

From a comparison of the two expressions given by Equations (5.7a) and (5.7b), it may be concluded that given the same requirements for the insertion loss characteristics in the stop-band (which determines the roll-off of the filter) equal-ripple filter requires fewer elements (lower order of the filter) compared to the maximally flat filter. In other words, with the same order of the filter, equal-ripple filter will have better roll-off beyond cut-off.

Case III Elliptic Function Filter

The maximally flat and equal-ripple responses both have monotonically increasing attenuation in the stop-band. In many applications, it is required to specify a minimum stop-band attenuation, when a better cut-off rate can be obtained. Such filters are called elliptic function filters and have equal-ripple response in the pass-band as well as in the stop-band as shown in Figure 5.2c. The maximum attenuation in the pass-band $P_{LRp(max)}$ as well as the minimum attenuation in the stop-band $P_{LRs(min)}$ is specified. The synthesis for elliptic function filter is quite involved and beyond the scope of this text; however, details are available in suitable reference books [2].

5.2.2 Low-pass Prototype Filter Design Technique

The low-pass prototype filter design is done in terms of normalized frequency and impedance. This normalization simplifies the design of filters for arbitrary frequency, impedance, and type of the filter. The low-pass prototypes are then scaled to the desired frequency and impedance.

Let the low-pass prototype be a ladder network of capacitance and inductance as shown in Figure 5.3a or its dual in Figure 5.3b. The load impedance is chosen to be equal to unity (1 ohm) and the source impedance R_s (Figure 5.3a), and $G_s = 1/R_s$ (Figure 5.3b) are chosen to be unity. The element values are denoted by g_k and are same for both the circuits; it is capacitance in farads (for k odd) and inductance in Henries (for k even) in Figure 5.3a; while it is inductance in Henries (for k odd) and capacitance in farads (for k even) in Figure 5.3b.

The element values of the prototype low-pass filter with equal-ripple and maximally flat responses for the normalized value of cut-off angular frequency ($\Omega = 1$) can be calculated from Bennett's formula [3]:

Case I For an Equal-Ripple Low-Pass Prototype Filter:

$$g_0 = 1, \text{ for all } n \quad \text{and} \quad g_1 = \frac{2a_1}{\gamma}$$

$$g_k = \frac{4a_{k-1}a_k}{b_{k-1}g_{k-1}}; \quad \text{with } k = 2,3,4 \ldots n$$

$$g_{n+1} = 1 \qquad \text{for } n \text{ odd}$$

$$= \coth^2(\beta/4) \quad \text{for } n \text{ even} \ldots \tag{5.8}$$

$$\text{where } a_k = \sin\frac{2k-1}{2n}\pi, \; b_k = \gamma^2 + \sin^2\frac{k\pi}{n}, \; \beta = \ln\left(\coth\frac{A_{pm}}{17.37}\right),$$

$$\gamma = \sinh\left(\frac{\beta}{2n}\right)$$

with $k = 2,3, \ldots n$.

[Note A_{pm} is the maximum attenuation in dB in the pass-band (i.e., the maximum ripple height) or in other words $IL_{\text{at } \omega = \omega c}$, i.e., $10\log_{10}(1 + A_p^2)$ since at $\omega = \omega_c$, $T_n(x) = 1$ (Eq. 4.29a)].

Case II For Maximally Flat Low-Pass Filter:

$$g_0 = 1 = g_{n+1}, \quad \text{for all } n$$

$$g_k = 2 \sin[(2k-1)\pi/2n], \quad k = 1,2, \ldots. n \ldots \tag{5.9}$$

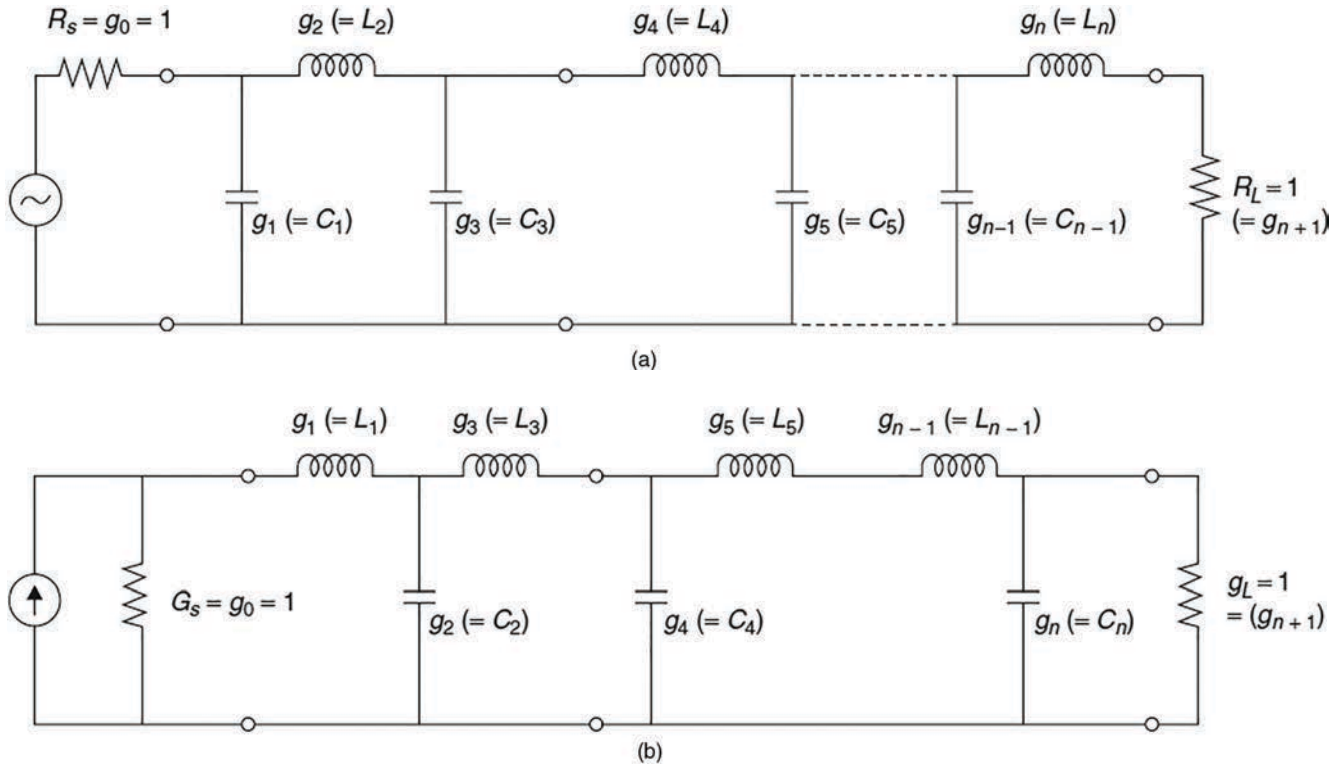


Figure 5.3 Low-pass prototype filter network: (a) Ladder network with $g_k = C_k$ (when k odd), $g_k = L_k$ (when k even), and (b) Dual of (a) with $g_k = L_k$ (when k odd), $g_k = C_k$ (when k even).

This analytical approach is used to evaluate the element values for prototype low-pass filter. Such tables are available in standard text [1]; however, we are also providing this for ready reference in Figure 5.4(a) and 5.4(b) in tabular form for equal-ripple and maximally flat cases, respectively. The tables have been generated by using MATLAB program (see Appendix II). In these tables, the values of g_k for different order (n) of Tchebysheff filter with maximum tolerable ripple height $A_{pm} = IL_{at}\ \omega = \omega_c = 10\ \log_{10}(1+A_p^2)$ as variable parameter are given in Figure 5.4(a): (i) to (v); while values of g_k for different order (n) of Butterworth (maximally flat) low-pass prototype filter for IL = 3 dB when $\Omega\ (=\omega/\omega_c) = \Omega_p = 1$ are given in Figure 5.4(b)

(i) A_{pm} = 0.1 dB

n	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}
1	0.3052	1.0000									
2	0.8430	0.6220	1.3554								
3	1.0315	1.1474	1.0315	1.0000							
4	1.1088	1.3061	1.7703	0.8180	1.3554						
5	1.1468	1.3712	1.9750	1.3712	1.1468	1.0000					
6	1.1681	1.4039	2.0562	1.5170	1.9029	0.8618	1.3554				
7	1.1811	1.4228	2.0966	1.5733	2.0966	1.4288	1.1811	1.0000			
8	1.1897	1.4346	2.1199	1.6010	2.1699	1.5640	1.9444	0.8778	1.3554		
9	1.1956	1.4425	2.1345	1.6167	2.2053	1.6167	2.1345	1.4425	1.1956	1.0000	
10	1.1999	1.4481	2.1444	1.62651	2.2253	1.6418	2.2046	1.5821	1.96281	0.8853	1.3554

(a)(i)

Figure 5.4 (a) [(i) to (v)] Table showing the values of g_k for different order (n) of Tchebysheff filter with maximum tolerable ripple height A_{pm} as parameter: 0.1 dB to 0.5 dB. (b) Table showing the values of g_k for different order (n) of Butterworth (maximally flat) low-pass prototype filter for IL = 3 dB when $\Omega\ (=\omega/\omega_c) = \Omega_p = 1$.

(ii) A_{pm} = 0.2 dB

n	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}
1	0.4342	1.0000									
2	1.0378	0.6745	1.5386								
3	1.2275	1.1525	1.2275	1.0000							
4	1.3028	1.2844	1.9761	0.8468	1.5386						
5	1.3394	1.3370	2.1660	1.3370	1.3394	1.0000					
6	1.3598	1.3632	2.2394	1.4555	2.0974	0.8838	1.5386				
7	1.3722	1.3781	2.2756	1.5001	2.0756	1.3781	1.3722	1.0000			
8	1.3804	1.3875	2.2963	1.5217	2.3413	1.4925	2.1349	0.8972	1.5386		
9	1.3860	1.3938	2.3093	1.5340	2.3728	1.5340	2.3093	1.3938	1.3860	1.0000	
10	1.3901	1.3983	2.3181	1.5417	2.3904	1.5536	2.3720	1.5066	2.514	0.9034	1.5386

(a)(ii)

(Continued)

(iii) A_{pm} = 0.3 dB

n	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}
1	0.5349	1.0000									
2	1.1800	0.6966	1.6967								
3	1.3706	1.1383	1.3729	1.0000							
4	1.4450	1.2540	2.1270	0.8541	1.6967						
5	1.4810	1.2995	2.3090	1.2999	1.4852	1.0000					
6	1.5009	1.3221	2.3784	1.4026	2.2428	0.8881	1.6967				
7	1.5130	1.3349	2.4124	1.4407	2.4126	1.3355	1.5191	1.0000			
8	1.5210	1.3429	2.4318	1.4591	2.4744	1.4342	2.2786	0.9010	1.6967		
9	1.5264	1.3483	2.4440	1.4695	2.5039	1.4695	2.4443	1.3491	1.5342	1.0000	
10	1.5304	1.3521	2.4521	1.4760	2.5203	1.4862	2.5032	1.4462	2.2944	0.9076	1.6967

(a)(iii)

(iv) A_{pm} = 0.4 dB

n	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}
1	0.6213	1.0000									
2	1.2983	0.7054	1.8435								
3	1.4902	1.1184	1.4926	1.0000							
4	1.5642	1.2228	2.2534	0.8509	1.8435						
5	1.5999	1.2635	2.4308	1.2639	1.6042	1.0000					
6	1.6195	1.2836	2.4978	1.3558	2.3658	0.8818	1.8435				
7	1.6315	1.2950	2.5306	1.3896	2.5308	1.2956	1.6377	1.0000			
8	1.6394	1.3022	2.5492	1.4059	2.5904	1.3838	2.4003	0.8936	1.8435		
9	1.6447	1.3069	2.5609	1.4151	2.6186	1.4151	2.5613	1.3077	1.6528	1.0000	
10	1.6486	1.3103	2.5687	1.4208	2.6343	1.4298	2.6178	1.3944	2.4156	0.8997	1.8435

(a)(iv)

(v) A_{pm} = 0.5 dB

n	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}
1	0.6986	1.0000									
2	1.4029	0.7071	1.9841								
3	1.5963	1.0967	1.5963	1.0000							
4	1.6703	1.1926	2.3661	0.8419	1.9841						
5	1.7058	1.2296	2.5408	1.2296	1.7058	1.0000					
6	1.7254	1.2479	2.6064	1.3137	2.4758	0.8696	1.9841				
7	1.7372	1.2583	2.6381	1.3444	2.6381	1.2583	1.7372	1.0000			
8	1.7451	1.2647	2.6564	1.3590	2.6964	1.3389	2.5093	0.8796	1.9841		
9	1.7504	1.2690	2.6678	1.3673	2.7239	1.3673	2.6678	1.2690	1.7504	1.0000	
10	1.7543	1.2721	2.6754	1.3725	2.7392	1.3806	2.7231	1.3485	2.5239	0.8842	1.9841

(a)(v)

(Continued)

n	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	g_{11}
1	2.000	1									
2	1.413	1.4159	1								
3	0.999	2.000	1.0023	1							
4	0.765	1.8473	1.84875	0.7679	1						
5	0.617	1.6175	2.0000	1.6193	0.6206	1					
6	0.517	1.4136	1.9315	1.9323	1.4159	0.5205	1				
7	0.444	1.2464	1.8014	2.0000	1.8028	1.2489	0.4479	1			
8	0.390	1.1107	1.6624	1.9613	1.9619	1.6642	1.1133	0.3931	1		
9	0.347	0.9995	1.5315	1.8790	2.0000	1.8801	1.5336	1.0023	0.350	1	
10	0.312	0.9076	1.4137	1.7815	1.9752	1.9756	1.7830	1.4159	0.910	0.315	1

(b)

5.3 FILTER TRANSFORMATIONS FROM PROTOTYPE

5.3.1 Introduction

The low-pass filter prototype of previous section was normalized design (with $R_L = R_s = 1$ and $\Omega = 1$ at the edge of the pass-band). However, the prototype design may be used as the basis for design of actual low-pass, high-pass, band-pass, and band-stop filters at any desired impedance and frequency. To do this, we need to scale the prototype circuit elements in terms of frequency and impedance, the procedure for which is given below.

The frequency transformation/mapping basically transforms or maps a response (Tchebysheff/Butterworth) of the low-pass prototype in the frequency domain Ω (as shown in Figure 5.2a) into the frequency domain ω of an actual/practical filter in which it is expressed. The frequency transformation will have effect on all reactive components but having no effect on the resistive elements of the actual filter.

The impedance scaling is required to remove the $g_0 = 1$ normalization of the source impedance as a practical filter might need to work for any desired value of source impedance, say $Z_s = Z_0$. For impedance scaling, a prototype reactive element in the series branch (i.e., appearing in the impedance branch) needs to be multiplied by the impedance scaling factor I_t while a prototype reactive element in the shunt branch (appearing in the admittance branch) needs to be divided by the scaling factor I_t. Thus, with reference to Figure 5.3a, the series inductances (L_k) and shunt capacitances (C_k) of the prototype will be represented by

$$L'_k = I_t L_k \quad \text{and} \quad C'_k = C/I_t \tag{5.10}$$

where $I_t = Z_0/g_0$.

The impedance scaling accounts for the actual impedance of the practical filter, and since it is independent of frequency same transformation/scaling holds for any type of filter.

Now, we would discuss transformation from low-pass prototype to actual low-pass, high-pass, band-pass, and band-stop filter.

5.3.2 Prototype to Actual Low-pass Transformation

For frequency scaling from low-pass prototype to practical low-pass filter, the latter having cut-off frequency ω_c in the angular frequency axis ω, we need to change $\Omega_p = (\omega/\omega_c) = 1$, to the actual band-edge cut-off frequency ω_c. This demands the frequency transformation from Ω to Ω', where $\Omega' = \left(\frac{\Omega_p}{\omega_c}\right)\omega$, for the prototype reactive elements, allowing us to write the following equations: $j\Omega' L = j\left(\frac{\Omega_p}{\omega_c}\right)\omega L_k = j\omega L'_k$ and $j\Omega' C_k = j\left(\frac{\Omega_p}{\omega_c}\right)\omega C_k = j\omega C'_k$,

whence:

$$L'_k = \Omega_p L_k/\omega_c, C'_k = \Omega_p C_k/\omega_c \tag{5.11}$$

When both impedance and frequency scaling are required, the results in Equations (5.10) and (5.11) can be combined to give

$$L'_k = \left(\frac{\Omega_p}{\omega_c}\right) L_k I_t \tag{5.12a}$$

$$C'_k = \left(\frac{\Omega_p}{\omega_c}\right)\frac{C_k}{I_t} \tag{5.12b}$$

The effect of the frequency transformation observed here basically thus represents a stretching out of the original low-pass prototype filter pass-band which shifts the cut-off from unity (Ω_p = 1, for prototype) to $\omega_c > 1$ (as desired by the actual low-pass filter pass-band) in the frequency scale; compare Figure 5.5b with 5.5(a).

5.3.3 Low-pass Prototype to High-pass Transformation

A conversion from low-pass to high-pass demands that in the frequency scale ω = 0 should map to $\omega = \pm\infty$ and vice versa (see Figure 5.5c) and inductance and capacitance of low-pass configuration be interchanged in the high-pass configuration. To implement this, we need to use the fre

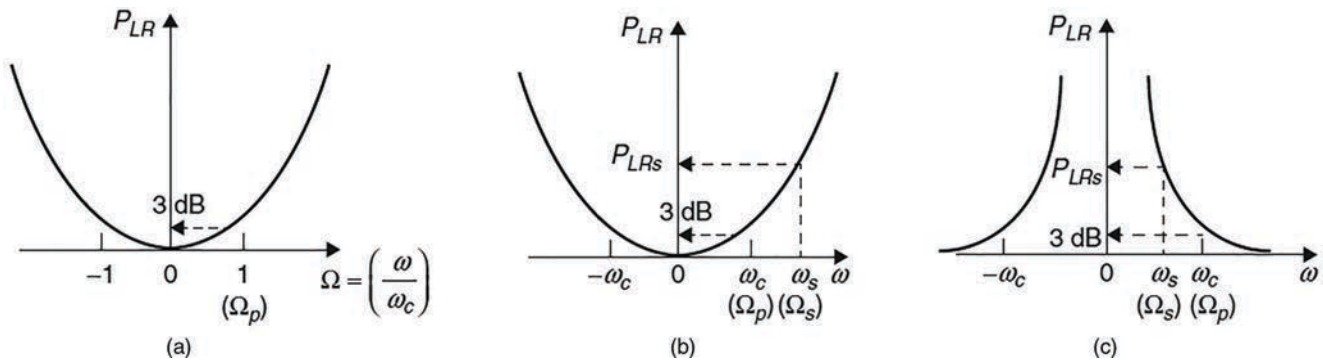


Figure 5.5 Filter response: (a) Low-pass prototype, (b) Actual low-pass with frequency scaling, and (c) High-pass with frequency scaling.

quency transformation: $\Omega' = -\left(\frac{\Omega_p}{\omega}\right)\omega_c$ to get the actual high-pass filter elements and the negative sign converts inductors/capacitors of low-pass prototype to capacitors/inductors in the actual high-pass filter. Having done this frequency transformation and including the impedance transformation, the actual high-pass filter elements can be derived as

$$C'_k = \left(\frac{1}{\omega_c \Omega_p}\right)\frac{1}{L_k I_t} \text{ and } L'_k = \left(\frac{1}{\omega_c \Omega_p}\right)\frac{I_t}{C_k} \tag{5.13}$$

This shows that the low-pass prototype to actual high-pass transformation effectively replaces all inductances L_k by capacitances C'_k and the C_k must be replaced by inductances L'_k.

5.3.4 Low-pass to Band-pass and Band-stop Transformations

The low-pass prototype filter design may also be transformed to exhibit band-pass and band-stop responses as illustrated in Figure (5.6a) and (5.6b). The low-pass response can be converted to a band-pass response by means of the following frequency transformation (remembering that band-pass is a combination of low-pass and high-pass characteristics):

$$\Omega' = \frac{\Omega_p}{\Delta}\left\{\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right\} \tag{5.14}$$

where $\Delta = (\omega_{c2} - \omega_{c1})/\omega_0$ is the fractional bandwidth of the pass-band with ω_{c1} and ω_{c2} denoting the edges of the pass-band (they correspond to $\Omega = \pm 1$

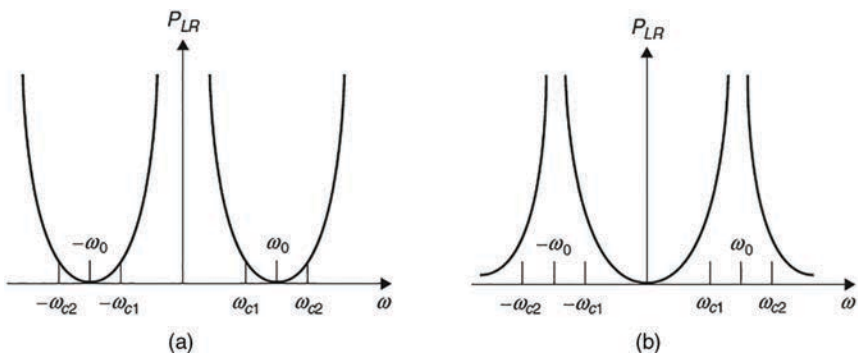


Figure 5.6 Filter response with frequency scaling: (a) Band-pass, and (b) Band-stop.

in the low-pass prototype). The mid-band frequency ω_0 is given by $\sqrt{\omega_{c1}\omega_{c2}}$. The mapping principle may be understood following standard texts [4,5].

The element values for the band-pass filter (BSF) can be determined from low-pass prototype values using the frequency transformation given by Equation (5.14) as follows. The series reactance of the prototype with frequency transformation yields

$$\frac{j\Omega_p}{\Delta}\left\{\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right\}L_k = j\frac{\Omega_p\omega L_k}{\Delta\omega_0} - j\frac{\Omega_p\omega_0 L_k}{\Delta\omega} = j\omega L_k' - j\frac{1}{\omega C_k'}$$

This shows that a series inductor L_k of low-pass prototype is transformed into a series LC circuit in BSF with element values:

$$L_{kse}' = \left(\frac{\Omega_p}{\Delta\omega_0}\right) I_t L_k \quad \text{and} \quad C_{kse}' = \left(\frac{\Delta}{\Omega_p\omega_0}\right)\frac{1}{I_t L_k} \tag{5.15}$$

(where impedance transformation is also included)

Similarly, the shunt susceptance of the prototype with frequency transformation yields

$$\frac{j\Omega_p}{\Delta}\left\{\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right\}C_k = j\frac{\Omega_p\omega C_k}{\Delta\omega_0} - j\frac{\Omega_p\omega_0 C_k}{\Delta\omega} = j\omega C_k' - j\frac{1}{\omega L_k'}$$

This shows that a shunt capacitor C_k of low-pass prototype is transformed to a shunt LC circuit in BSF with element values (including the impedance transformation):

$$L_{ksh}' = \left(\frac{\Delta}{\Omega_p\omega_0}\right)\frac{I_t}{C_k} \quad \text{and} \quad C_{ksh}' = \left(\frac{\Omega_p}{\Delta\omega_0}\right)\frac{C_k}{I_t} \tag{5.16}$$

The low-pass filter elements are thus converted to series resonant circuits (low impedance at resonance) in the series arms, and parallel resonant circuits (high impedance at resonance) in the shunt arms; however, both series and parallel resonator elements have the same resonant frequency, ω_0.

To realize the band-stop response (Figure 5.5b), we need to use inverse transformation:

$$\Omega' = \frac{\Omega_p}{\Delta}\left\{\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right\}^{-1} \tag{5.17}$$

where Δ and ω_0 have the similar meaning as above. Then, the series inductors of low-pass prototype are converted to shunt LC circuits in band-stop filter having element values:

$$L'_{ksh} = \left(\frac{\Omega_p \Delta}{\omega_0}\right) I_t L_k \quad \text{and} \quad C'_{ksh} = \left(\frac{1}{\Omega_p \Delta \omega_0}\right) \frac{1}{I_t L_k} \tag{5.18}$$

The shunt capacitors of the low-pass prototype are converted to series LC circuits in band-stop filter having element values:

$$L'_{kse} = \left(\frac{1}{\Omega_p \Delta \omega_0}\right) \frac{I_t}{C_k} \quad \text{and} \quad C'_{kse} = \left(\frac{\Omega_p \Delta}{\omega_0}\right) \frac{C_k}{I_t} \tag{5.19}$$

The element transformations from low-pass prototype to actual low-pass, high-pass, band-pass, and band-stop filter (with both frequency and impedance scaling) are summarized in a tabular form in Figure 5.7.

5.4 MICROWAVE IMPLEMENTATION

5.4.1 Introduction

So far, we have discussed lumped-element filters that generally work well at low frequencies but have limited scope to be used directly at RF/microwave frequencies. Though lumped components may be implemented at RF/microwave frequency, they are available only up to certain frequency and with limited range of values and thus RF/microwave filters are mostly designed with distributed component inductors and capacitors. In addition, at RF/microwave frequencies, the distance between filter components sometimes become negligibly small. To convert lumped elements to distributed components with transmission line sections *Richard's transformation* is used, while *Kuroda's identities* can be used to separate filter elements by using transmission line sections. Such additional transmission line sections normally do not affect the filter response; however, it is possible to design RF/microwave filters that take advantage of these sections to improve the filter response.

5.4.2 Richard's Transformation

This transformation was introduced by Paul I Richard in 1948 to synthesize an LC network using open- and short-circuited transmission line sections or stubs. Here, an inductance can be replaced by a short-circuited stub of length l and characteristic impedance Z_k, and a capacitance can be replaced

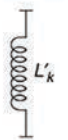
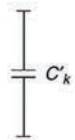
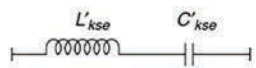
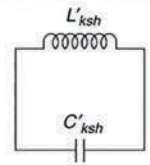
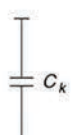
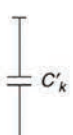
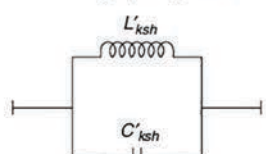
Low-pass prototype	Actual low-pass	High-pass	Band-pass	Band-stop
				
$g_k = L_k$	$L'_k = \left(\frac{\Omega_p}{\omega_c}\right) L_k I_t$	$C'_k = \left(\frac{1}{\omega_c \Omega_p}\right) \frac{1}{L_k I_t}$	$L'_{kse} = \left(\frac{\Omega_p}{\Delta \omega_o}\right) I_t L_k$ $C'_{kse} = \left(\frac{\Delta}{\Omega_p \omega_o}\right) \frac{1}{I_t L_k}$	$L'_{ksh} = \left(\frac{\Omega_p \Delta}{\omega_o}\right) I_t L_k$ $C'_{ksh} = \left(\frac{1}{\Omega_p \Delta \omega_o}\right) \frac{1}{I_t L_k}$
				
$g_k = C_k$	$C'_k = \left(\frac{\Omega_p}{\omega_c}\right) \frac{C_k}{I_t}$	$L'_k = \left(\frac{1}{\omega_c \Omega_p}\right) \frac{I_t}{C_k}$	$L'_{ksh} = \left(\frac{\Delta}{\Omega_p \omega_o}\right) \frac{I_t}{C_k}$ $C'_{ksh} = \left(\frac{\Omega_p}{\Delta \omega_o}\right) \frac{C_k}{I_t}$	$L'_{kse} = \left(\frac{1}{\Omega_p \Delta \omega_o}\right) \frac{I_t}{C_k}$ $C'_{kse} = \left(\frac{\Omega_p \Delta}{\omega_o}\right) \frac{C_k}{I_t}$

Figure 5.7 Summary of prototype to actual filter transformation: [$\Delta = (\omega_{c2} - \omega_{c1})/\omega_0$; $\Omega_p = 1$; $I_t = Z_0/g_0$ (with $g_0 = 1$)].

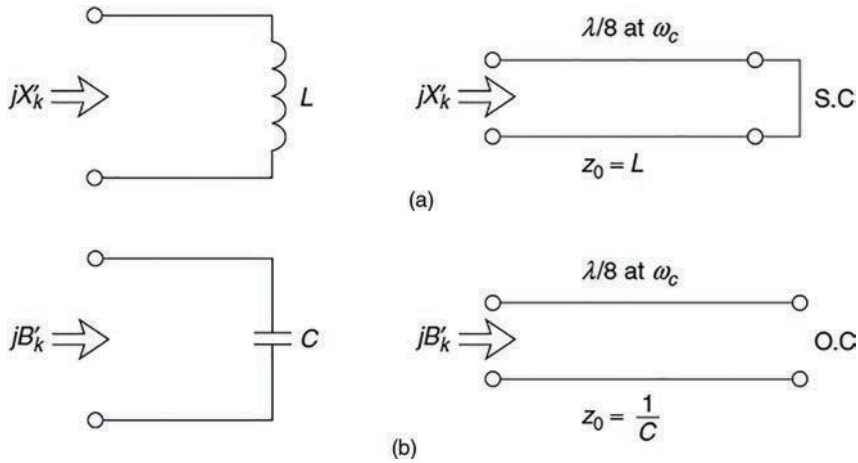


Figure 5.8 Richard's transformation: (a) For an ideal inductor to a short-circuited stub, and (b) For a capacitor to an open-circuited stub.

by an open-circuited stub of length l and characteristic impedance Y_k. The new reactance and susceptance functions now become:

$$jX_k' = jZ_k \tan\beta l = jZ_k \tan(\omega l/c) = jZ_k \tan\theta = jg_k \tan\theta \tag{5.20a}$$

$$jB_k' = jY_k \tan\theta = jg_k \tan\theta \tag{5.20b}$$

The power loss ratio for the filter is given by

$$P_{LR} = 1 + P(\tan^2\theta) \tag{5.20c}$$

The frequency transformation $\Omega' = \tan\theta$ maps the whole Ω' axis periodically into intervals of length π in the θ domain. Hence, the inductors and capacitors of a lumped-element filter can be replaced with short-circuited and open-circuited stubs as illustrated in Figure 5.8. The reason why the length of the line should be $\lambda/8$ at ω_c is that for a short-circuited transmission line when length of the line is made $\lambda/8$ the characteristic impedance of the line represents inductive impedance at ω_c while an open-circuited transmission line of length $\lambda/8$ represents a capacitive impedance at ω_c. This equivalence of transmission line length representing lumped inductance and capacitance by short-circuited and open-circuited line, respectively, of length $\lambda/8$ is valid not only at ω_c but also holds approximately good for any ω far-off the cut-off frequency. This is because, at $\omega = \omega_c$, for short-circuited stub with $l = \lambda/8$, $\omega_c L = Z_0$; while for $\omega << \omega_c$, when $\tan(\pi/4) \approx \pi/4$, $\omega L = 0.8Z_0$.

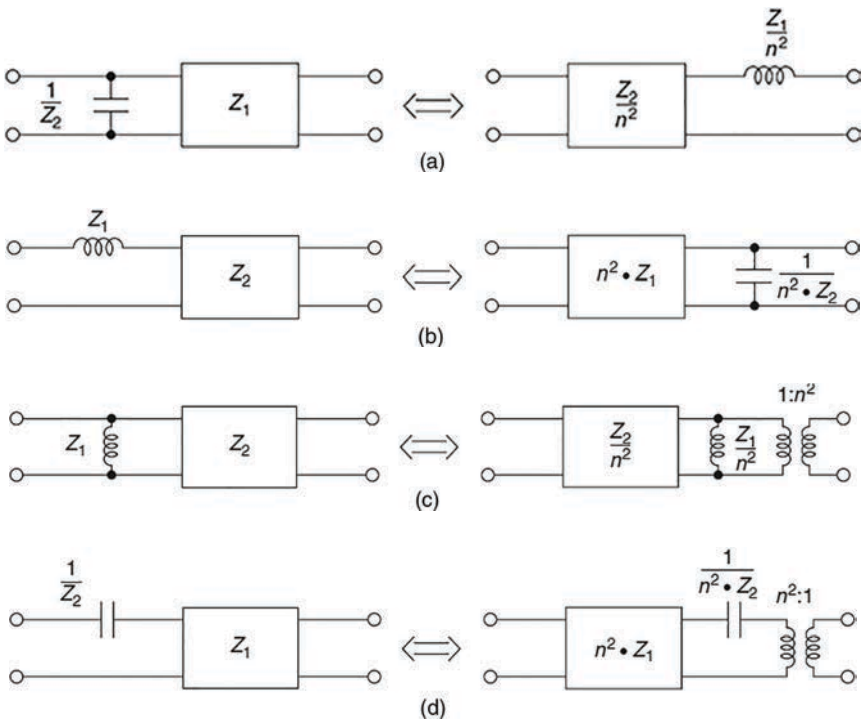


Figure 5.9 Kuroda's identities: (a) Shunt capacitor input, (b) Series inductor input, (c) Shunt inductor input, and (d) Series capacitor input; [$n^2 = (1 + Z_2/Z_1)$].

5.4.3 Kuroda's Identities

Kuroda's identities (transforms), proposed by Kunikatsu Kuroda in 1955 originally in his PhD thesis, are now used to convert a section of transmission line with a shorted series stub into an electrically equivalent section of transmission line with an open parallel stub and vice versa. As a result, an identical *ABCD*-parameter matrix is produced that performs the same function. These identities may also be used to separate physically the transmission line stubs and also to change impractical characteristics impedances into more realizable ones. To be precise, these identities are used to achieve a more practical RF/microwave filter implementation.

There are four such network equivalences (see Figure 5.9); the equivalence is only valid in the S-plane defined as

$$S = j\Omega' = j\tan\beta l = j\tan\theta \tag{5.21}$$

Each box in Figure 5.9 represents a unit element or a transmission line of indicated characteristic impedance and having length $\lambda/8$ at $\omega = \omega_c$. The

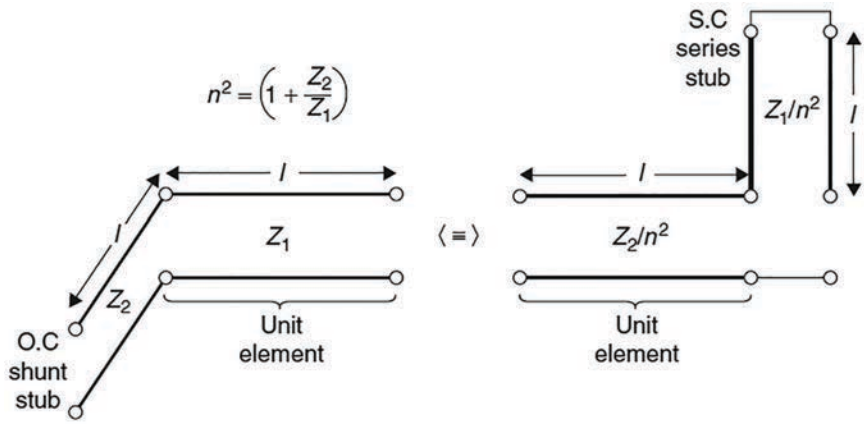


Figure 5.10 Equivalent circuit illustrating Kuroda's identity (a) in Figure 5.9.

inductors and capacitors represent short-circuited and open-circuited stubs, respectively. The two circuits of identity (a) in Figure 5.9 can be redrawn as in Figure 5.10, which are equivalent.

Example 5.1 Design a third-order low-pass filter in a 50Ω microstrip transmission line system with substrate material having ε_r = 10 and substrate height equal to 0.65 mm. The cut-off frequency of the filter is desired to be at 1.8GHz. Assume 0.5dB equal-ripple pass-band characteristics.

Solution:

Given the order of the filter (n) = 3 (odd); $g_{n+1} = g_4 = g_0 = 1.0$.

Referring to table (a)(v) of Figure 5.4, the normalized low-pass prototype element values are given by

$g_1 = 1.5963$ (= L_1), $g_2 = 1.0967$ (= C_2), $g_3 = 1.5963$ (= L_3).

The corresponding lumped-element circuit is shown in Figure 5.11a. Now we use Richard's transformation to convert series inductors in the form of series stubs and shunt capacitors in the form of shunt stubs, as shown in Figure 5.11b. According to Equation (5.20), characteristic impedance of a series stub (inductor) is $Z_{0se} = L$, and the characteristic impedance of a shunt stub (capacitor) is $Z_{0sh} = 1/C$. Length of each stub will be $\lambda/8$ at ω_c.

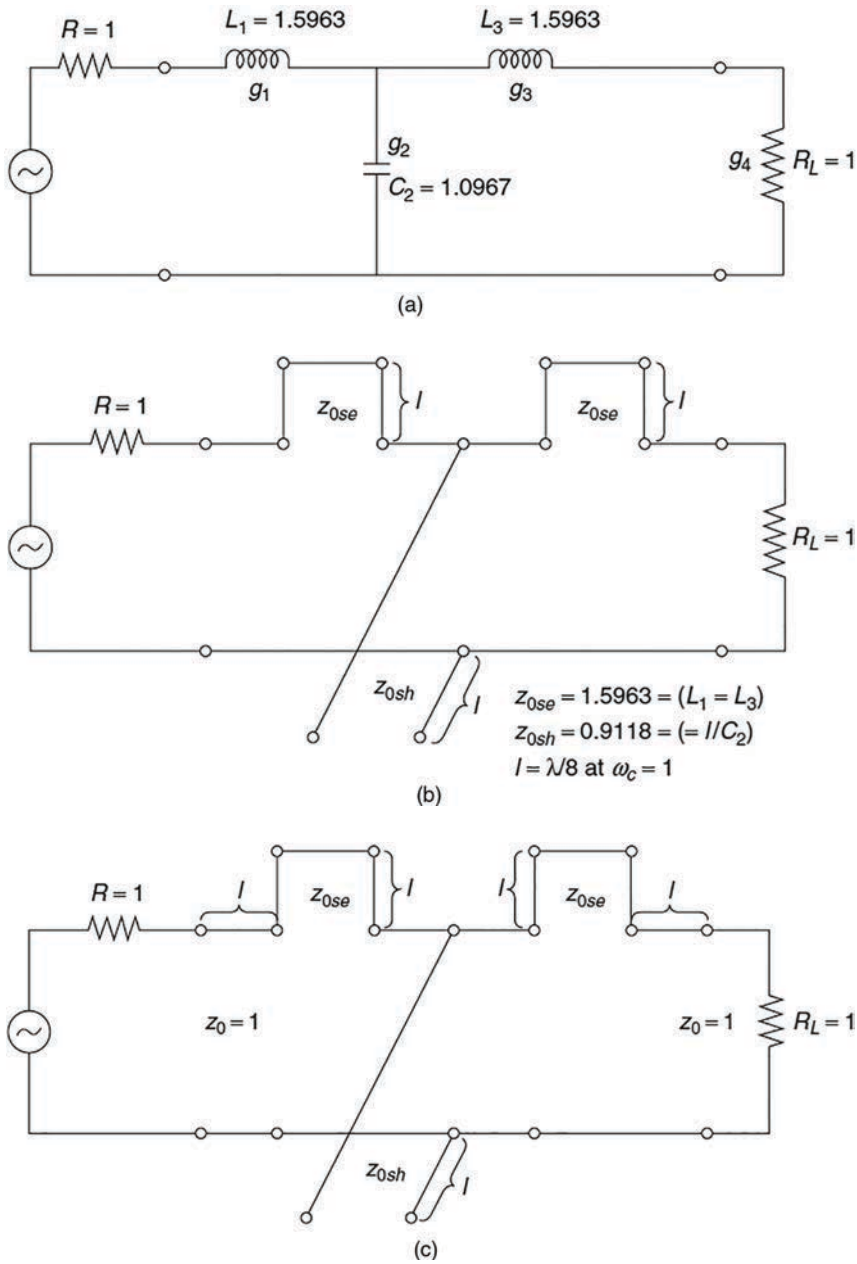


Figure 5.11 (a) Lumped element low-pass filter prototype, (b) Richard's transformation applied to convert L and C to series and shunt stubs, (c) Unit elements added to the ends of the filter, (d) Application of Kuroda's identity, (e) After impedance and frequency scaling, and (f) Microstrip implementation of the filter (all dimensions in mm).

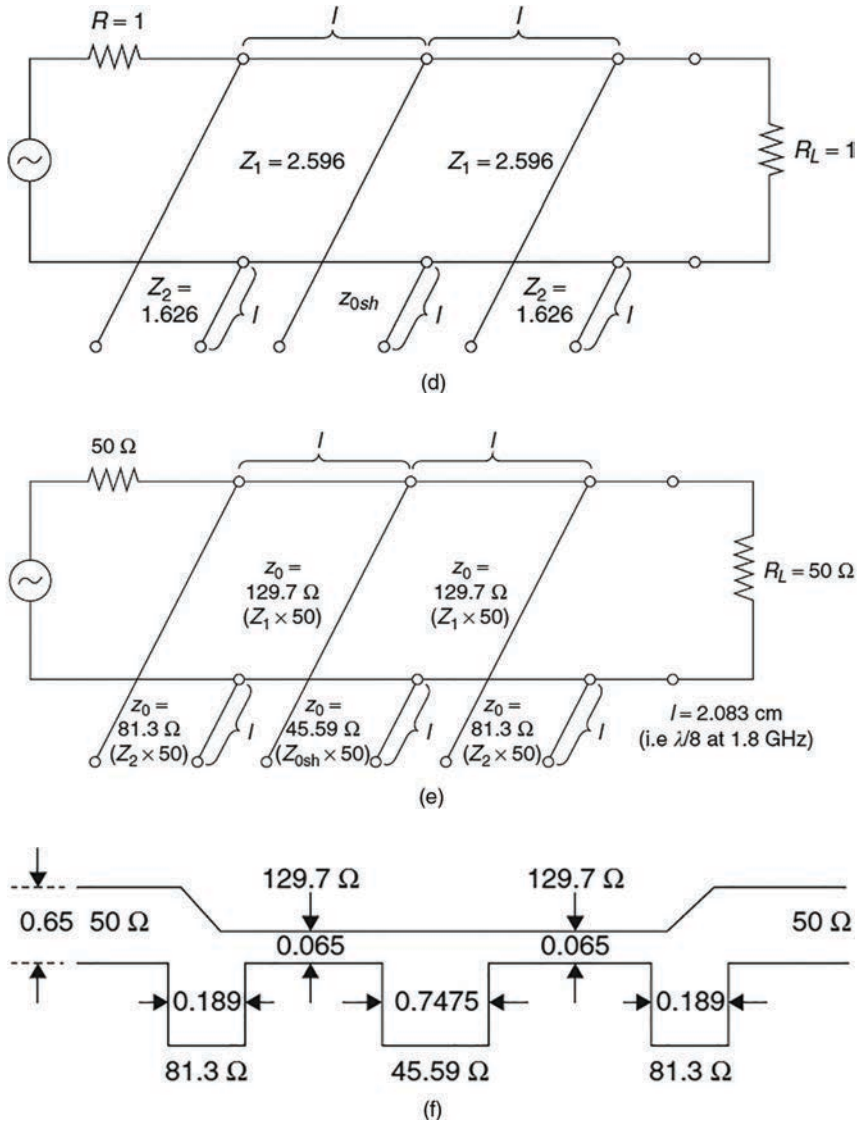


Figure 5.11 (Continued)

As the series stubs of Figure 5.11b would be very difficult to implement in microstrip form, we will use one of the Kuroda's identities to convert these to shunt stubs. First, we add unit elements at either end of the filter, as shown in Figure 5.10c. These redundant elements do not affect the filter performance since they are matched to the source and load (Z_0 = 1 = R = R_L). Now we can apply Kuroda's identity (b) shown in Figure 5.9 to both ends of the filter. In both cases, we have

$$n^2 = 1 + \frac{Z_2}{Z_1} = 1 + \frac{1}{1.5963} = 1.626$$

(Since $Z_1 = Z_{0se} = 1.5963\ \Omega$ and $Z_2 = 1\ \Omega$)

The result is shown in Figure 5.11d, which follows from Figure 5.11c where we use $Z_2/n^2 = 1$ (when $Z_2 = 1.626$) and $Z_1/n^2 = Z_0$ (1.5963 Ω) for shunt element to convert it to series element.

Finally, we need to scale the circuit in terms of impedance and frequency, which simply involves multiplying the normalized characteristic impedances by 50 Ω and choosing the line and stub lengths to be $\lambda/8$ at 1.8 GHz (l = 2.083 cm). The final circuit is shown Figure 5.11e, with microstrip layout in Figure 5.11f. It may be observed that lines with smaller characteristic impedance have larger width as the characteristic impedance of a microstrip line is inversely proportional to the strip width.

5.4.4 Impedance and Admittance Inverters

At RF/microwave frequency it is desirable, from fabrication point of view, to use either all series or all shunt elements to implement filter networks. If a combination of both series and shunt elements occurs (as we have seen in Example 5.1), then it is advisable to convert the series elements to shunt elements and vice versa. Kuroda's identity can be used to realize such conversions; however, another practical approach is to use impedance (K) or admittance (J) inverters. Such inverters are especially useful for the design of band-pass and band-stop filters at RF/microwave frequency.

The impedance and admittance inverters, in essence, are ideal quarter-wave transformers; as a load connected at one end of such a transformer is seen to be inverted with respect to the characteristic impedance squared at the input. There is no basic difference between the inverting properties of impedance/admittance inverters; the only distinction is that the symbol K is used to denote the characteristic impedance of impedance inverter and the symbol J is for the characteristic admittance of admittance inverter. The K inverter help to transform a shunt element into a series element while a J inverter transforms a series element into a shunt element having the same dependence on frequency and an arbitrary impedance level. Thus, the impedance and admittance inverters enable us to use identical resonators, either series or parallel tuned, throughout a filter network.

For simple illustration, let us consider (Figure 5.12a) a parallel admittance element Y_p connected by an ideal impedance inverter of characteristic impedance K (=1) on both sides of it. The input impedance Z_{in} is given by

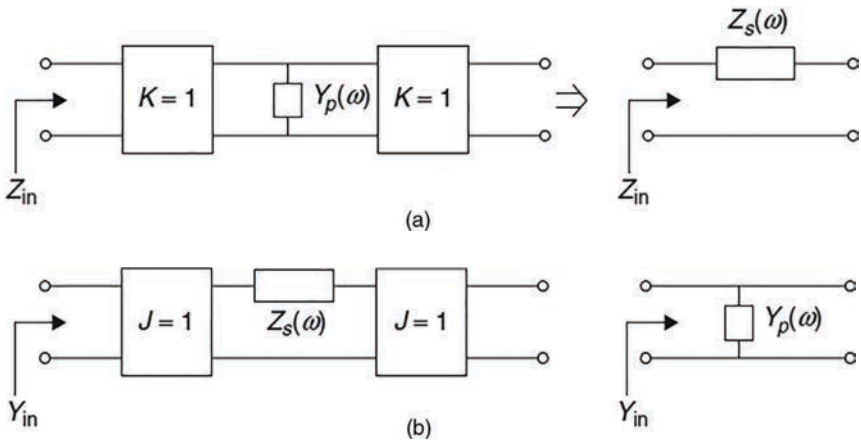


Figure 5.12 (a) Impedance (K) inverter, and (b) Admittance (J) inverter.

$$Z_{in} = \frac{K^2}{Z_P} = K^2 Y_P = Y_P = Z_S \tag{5.22}$$

Thus, the shunt element with two impedance inverters on both sides of it converts the shunt admittance into an equivalent series impedance $Z_s(\omega) = Y_p(\omega)$.

If Y_p is a parallel-tuned resonator with $Y_P = j\omega C - (j/\omega L) = j\omega C(1 - \omega_0^2/\omega^2)$, it will be converted into a series-tuned circuit with $Z_S = j\omega L(1 - \omega_0^2/\omega^2)$. The inductance L in Henries will have the same numerical value as the capacitance C in farads. If we want to convert an admittance $Y_1 = j\omega C_1(1 - \omega_0^2/\omega^2)$ into a particular series-tuned element with arbitrary value of inductance L, then *K* should be chosen on the basis of:

$$K^2 j\omega C_1(1 - \omega_0^2/\omega^2) = j\omega L(1 - \omega_0^2/\omega^2) \tag{5.23}$$

$$\text{or} \quad K = \sqrt{\frac{L}{C_1}}$$

Thus, we can transform a shunt element into a series element having the same dependence on frequency but with an arbitrary impedance level.

Figure 5.12b is shown a series element $Z_s(\omega)$ with admittance inverter having characteristic admittance *J* (=1) connected on both sides of it. The input admittance will be

$$Y_{in} = \frac{J^2}{Y_S} = J^2 Z_S = Y_P \tag{5.24}$$

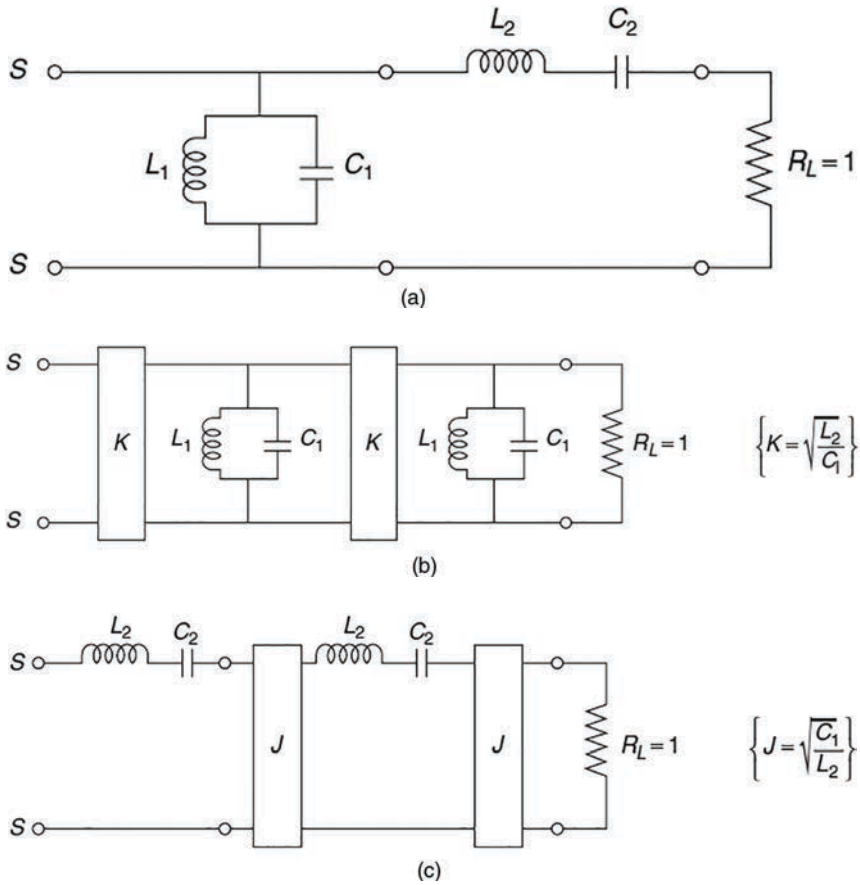


Figure 5.13 (a) A section of band-pas filter, (b) Conversion to all series resonator by K-inverter, and (c) Conversion to all parallel resonator by J-inverter.

Thus, the series element with the two admittance inverters on both sides of it will appear from the input terminals to be equivalent to a shunt admittance.

We can use inverters to convert the circuit shown in Figure 5.13a into one with two identical series-tuned resonators (using K inverters) or one with two identical parallel-tuned resonators (using J inverters).

For Figure 5.13b, we choose K so that

$$K^2 j\omega C_1(1-\omega_0^2/\omega^2) = j\omega L_2(1-\omega_0^2/\omega^2) \tag{5.25}$$

i.e., $K = \sqrt{\dfrac{L_2}{C_1}}$

While for Figure 5.13c, J is chosen so that

$$J^2 j\omega L_2(1-\omega_0^2/\omega^2) = j\omega C_1(1-\omega_0^2/\omega^2) \tag{5.26}$$

i.e., $J = \sqrt{\dfrac{C_1}{L_2}}$

5.5 PRACTICAL MICROWAVE FILTERS—ANALYSIS AND DESIGN

5.5.1 Introduction

We shall now take-up the analysis and design of common type of low-pass, band-pass, and band-stop microwave filters those are used in practice. Low-pass filters may be conveniently constructed using TEM structures like coaxial lines, strip-line/microstrip line (waveguide form of low-pass filters cannot be realized since waveguides basically exhibit high-pass characteristics). These low-pass filters are realized with high-Z (representing series L of the prototype filter) and low-Z (representing shunt C of the prototype filter). Popularly, they are known as high-low or stepped-impedance filters. The band-pass and band-stop filters require elements that behave as series or parallel resonant circuits (see the table in Figure 5.6). Multiple sections of coupled transmission lines acting as coupled resonators are used to design BSFs. Parallel (edge) coupled lines are preferred over end-coupled lines because of stronger coupling thereby providing relatively larger bandwidth (though still falls in the narrow-band regime) of the filter. Interdigital coupled lines, however, provide compact design. Coupled-line type BSF can be implemented in strip-line/microstrip line. In general, band-stop filters for practical use are concerned only with narrow frequency bands, and such filters can be designed with coupled resonators either of strip-line/microstrip line type or waveguide type.

5.5.2 Stepped-Impedance (High-Low) Low-pass Filter

Now we would discuss the design of high-low or stepped-impedance low-pass filter, which is commonly used in RF/microwave circuit design using microstrip transmission line.

Following the basic transmission line theory, we can use short length ($\beta l << 1$) of lossless transmission line at RF/microwave frequency to realize inductive and capacitive elements of the low-pass filter. For such short length of RF/microwave transmission line, we know from transmission line theory: $\omega L = R_{0l}\beta l_L$, which yields

$$l_L = f_c L\lambda/R_{0L} \quad \text{or} \quad R_{0L} = f_c L\lambda/l_L \ \ldots\ldots \tag{5.27a}$$

While, $1/\omega C = R_{0C}/\beta l_C$, which yields:

$$l_C = f_c C \lambda R_{0C} \quad \text{or} \quad R_{0C} = l_C/f_c C\lambda \ \ldots\ldots \tag{5.27b}$$

where f_c is the cut-off frequency of the filter.

It may be noted that L and C in Equations (5.27a) and (5.27b) are total inductance and capacitance of the line and not the per unit length of the line. The line providing inductive impedance will represent a line having high characteristic impedance, while the line providing capacitive impedance will represent a line having low characteristic impedance.

Example 5.2 Design a stepped-impedance low-pass filter having a Tchebysheff response with 0.3dB ripple and cut-off frequency of 2.4GHz. It is desired to have at least 35dB insertion loss at 4.0GHz in the stop-band. Realize the filter in microstrip transmission line having a dielectric substrate of thickness 0.65mm with ε_r = 10 at 2.4GHz. Assume 50Ω terminations at both load and source ends and the highest practical line impedance are 120Ω and the lowest being 25Ω.

Solution:

This problem will be solved both by analytical synthesis process and also using EDA tool.

METHOD I: Analytical Synthesis Process:

For the given ripple magnitude, 0.3 dB (A_{pm}) = 10 $\log_{10}(1+ A_p^2)$ or, $A_p^2 = 0.07152$

The power loss ratio of actual low-pass filter having equal-ripple, i.e., Tchebysheff response is given by

$P_{LR} = 1 + A_p^2 T_n^2(\Omega')$; with $\Omega' = \omega_s/\omega_c$ (see Figure 5.5b), (since $\Omega_p = 1$).

Now, $T_n(\Omega') = \cosh(n \cosh^{-1}\Omega')$; where $\Omega' = \omega_s/\omega_c = 4/2.4 = 1.67 > 1$ (as per the given data).

Substituting respective values (P_{LRs} = 35dB and Ω' = 1.67) in the expression for P_{LR} and solving for n we get n = 5.48, which can be taken approximately as 5.

[**Note:** $[1 + 0.0715\cosh^2(n\cosh^{-1}1.67)] = 10^{35/10}$ or

$$[\cosh(n\cosh^{-1}1.67)] = [(10^{3.5} - 1)/0.715]^{1/2} = 210.27 \text{ or}$$
$$(n\cosh^{-1}1.67)] = \cosh^{-1}(210.27) = 6.04;$$

whence $n = 6.04/\cosh^{-1}1.67 = 5.48 \approx 5$].

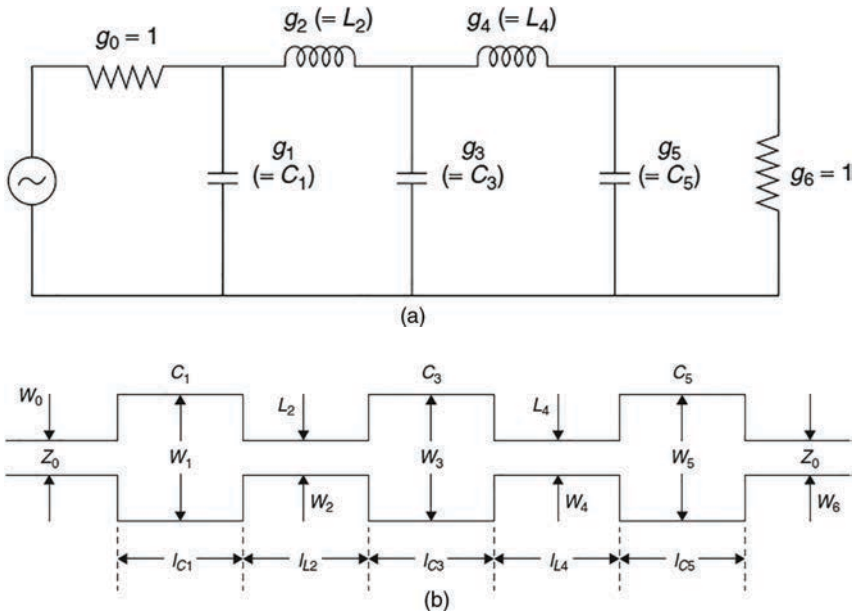


Figure 5.14 (a) Low-pass prototype corresponding to Example 5.3, and (b) Microstrip implementation scheme.

Therefore, a five-section low-pass filter is required.

Since n is odd, $g_{n+1} = g_6 = g_0 = 1.0$ while $g_1 = 1.4810 = g_5$, $g_2 = 1.299 = g_4$, $g_3 = 2.3090$ (which follows from table (a)(iii) of Figure 5.4). The low-pass prototype filter is shown in Figure 5.14a.

Next, we use impedance and frequency scaling to find the inductor and capacitor values of the prototype (referring to Eq. (5.11a) and Eq. (5.11b)), where $R_L = R_s = Z_0$ (say) = 50 Ω and $\omega_c = 2\pi \times 2.4$GHz, when: $C_1 = C_5 = 1.965$ pF, $C_3 = 3.0639$pF and $L_2 = L_4 = 4.3093$ nH.

[Note: $\Omega_p = 1$ and $I_t = Z_0/g_0 = Z_0$ since $g_0 = 1$].

These lumped elements can be realized with sections of microstrip lines having length and width having corresponding suffices as shown in Figure 5.14b. The capacitive elements will be the one with lowest line impedance, i.e., $Z_{0C} = 25\Omega$, while the inductive elements will be with highest line impedance, i.e., $Z_{0L} = 120\Omega$ and the impedance on the load and source side being $Z_0 = 50\Omega$. As the filter will be constituted with low- and high-impedance sections, they are also known as low-high-low filters. It is to be noted that low impedance section will have larger strip width while high impedance section will have smaller strip width (since for microstrip line $Z_0 \; \alpha \; 1/W$). Referring to Figure 5.15 (computed plots for (a) Z_0 vs W/h, and (b) ϵ_{reff} vs W/h for different parametric values of dielectric constant (ϵ_r) of the substrate material)

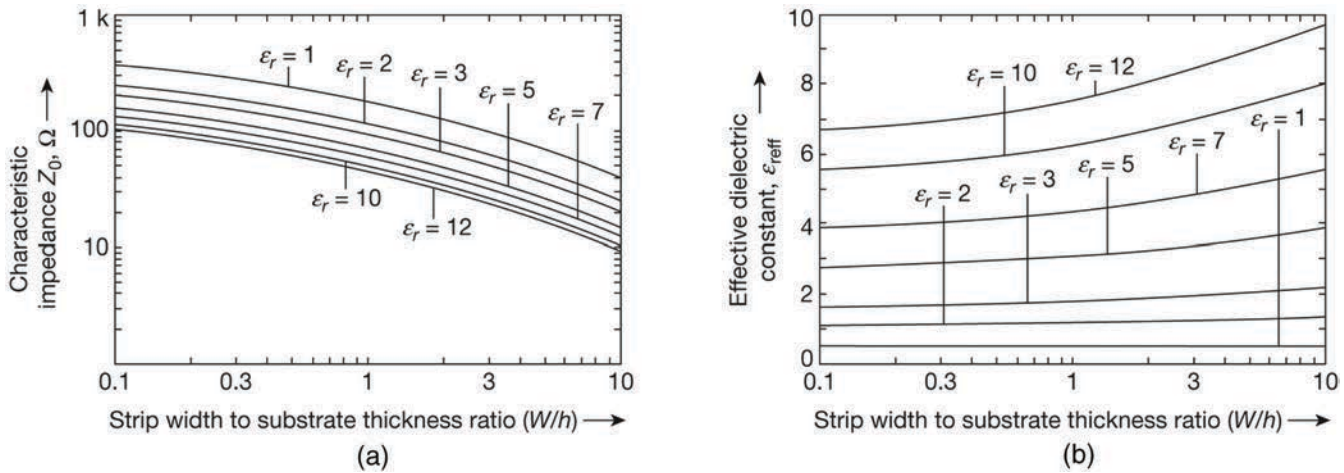


Figure 5.15 Computed plots for (a) Z_0 vs W/h, and (b) ϵ_{reff} vs W/h for different parametric values of dielectric constant (ϵ_r) of the substrate material.

[6], we clearly find that for $Z_{0C} = 25\ \Omega$, $W/h > 1$ and for $Z_{0L} = 120\ \Omega$, $W/h < 1$.

The length of each section will be calculated using the following formulas:

$L_{C1} = l_{C5} = f_c \lambda_{gC} R_{0C} C_1$, and $l_{C3} = f_c \lambda_{gC} R_{0C} C_3$. While $l_{L2} = l_{L4} = f_c\ L_2\ \lambda g_L / R_{0L}$.

In the above, value of λ_g will be calculated from $\lambda_g = \lambda_0 / \sqrt{\varepsilon_{reff}}$; where [6]

$$\varepsilon_{\text{reff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[\left(1 + 12\frac{h}{W}\right)^{-1/2} + 0.04\left(1 - \frac{W}{h}\right)^2\right] \quad \text{for} \quad (\text{W/h}) \le 1$$

$$= \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + 12\frac{h}{W}\right)^{-1/2} \quad \text{for } (\text{W/h}) > 1$$

The values for W/h can be calculated from [6]:

$$\frac{W}{h} = \frac{8e^A}{e^{2A} - 2} \quad \text{for } (\text{W/h}) \le 2$$

$$= \frac{2}{\pi}\left[B - 1 - \ln(2B - 1) + \frac{\varepsilon_r - 1}{2\varepsilon_r}\left\{\ln(B - 1) + 0.39 - \frac{0.61}{\varepsilon_r}\right\}\right] \quad \text{for} \quad (\text{W/h}) \ge 2$$

where $A = \dfrac{Z_0}{60}\sqrt{\dfrac{\varepsilon_r + 1}{2}} + \dfrac{\varepsilon_r - 1}{\varepsilon_r + 1}\left(0.23 + \dfrac{0.11}{\varepsilon_r}\right)$ and $B = \dfrac{377\pi}{2Z_0\sqrt{\varepsilon_r}}$

The expression for Z_0 for microstrip line is given by [6]

$$Z_0 = \frac{60}{\sqrt{\varepsilon_{reff}}}\ln\left(\frac{8h}{W} + \frac{W}{4h}\right) \quad \text{for } (W/h) \le 1$$

$$= \frac{120\pi}{\sqrt{\varepsilon_{reff}}\left[\left(\dfrac{W}{h} + 1.393\right) + 0.667\ln\left(\dfrac{W}{h} + 1.444\right)\right]} \quad \text{for } (W/h) > 1$$

Necessary calculation (either done analytically or by writing MATLAB program (see Appendix II) yields

$W_1 = W_5 = W_3 = 2.00347$ mm, $W_2 = W_4 = 0.0392$ mm.

While, $W_0 = W_6 = 0.6216$ mm, with $Z_0 = 50\ \Omega$ and $\varepsilon_r = 10$)

The cut-off frequency of the filter being at 2.4 GHz, so $\lambda_0 = 12.5$ cm, when

$\varepsilon_{reff} = 5.977$ for $(W/h) \leq 1$ and $\varepsilon_{reff} = 7.534$ for $(W/h) > 1$.

Thus, $\lambda_{gL} = 5.1128$ cm, and $\lambda_{gC} = 4.5539$ cm.

Whence $l_{C1} = l_{C5} = 6.028$ mm, $l_{L2} = l_{L4} = 4.406$ mm, and $l_{C3} = 9.399$ mm.

METHOD II: Using EDA Tool: The characteristics of stepped-impedance low-pass filter obtained with EDA tool is depicted in Figure 5.16a.

The curve 1 in Figure 5.16a is obtained with ideal (i.e., lossless having infinite Q value) lumped components as evaluated from analytical synthesis. However, all practical circuit elements will have a finite loss (finite Q value). Curve 2 is obtained with a Q value of 169 (say) that shows some finite insertion loss in the pass-band of the filter; but the filter response otherwise is similar to the ideal filter since no parasitic effect has been considered so far. The distributed realization using microstrip configuration may introduce additional losses arising from surface wave propagation in the quasi-TEM structure (as indicated by curve 3 in Figure 5.16a, obtained with circuit simulator). The complete electromagnetic (EM) simulation (using Method of Moments, MoM) takes care of extra parasitic due to the cross-coupling between elements along with above mentioned non-ideal characteristics. This is manifested in the form of a reduced bandwidth and changed filter skirt in the result (curve 4 in Figure 5.16a). The return loss is seen to be nearly –15 dB down from 0 dB reference (in fact 10 dB down is fine, however, for good design –20 dB down is desirable which is done by optimization), this may be observed from Figure 5.16b showing the reflection characteristics. The screen-shot of the schematic for ideal lumped component and practical lumped component LPF is shown in Figure 5.16b. Microstrip realization with circuit simulator is shown in Figure 5.16c while advanced design system (ADS) layout of EM simulation of the stepped-impedance LPF is shown in Figure 5.16d.

[NOTE: A few observations may be made with respect to the design of the stepped-impedance filter either with analytical synthesis method or by using CAD-based EDA tool. As per the problem, the design data demands $W_2 = W_4 = 0.0392$ mm which is a critically small dimension from fabrication point of view. Also, microstrip realization either with circuit simulator or EM simulation shows that the cut-off frequency is shifted from the design specification. For design optimization of a

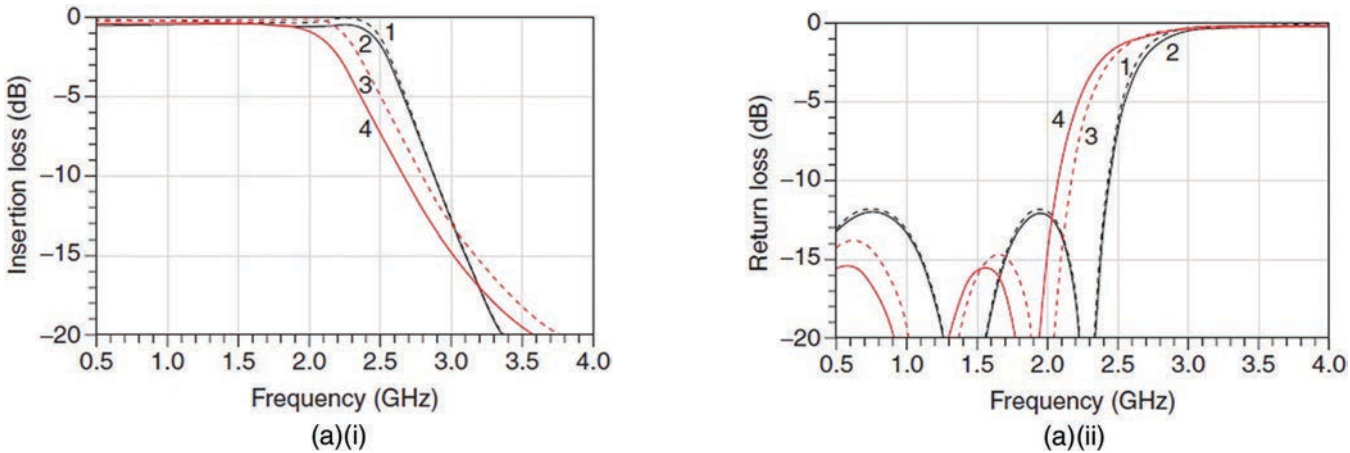


Figure 5.16 (a) Stepped-impedance low-pass filter (LPF) characteristics obtained with EDA tool: (i) Transmission characteristics, and (ii) Reflection characteristics, (b) Screenshot of the schematic for ideal lumped component and practical lumped component LPF, (c) Screenshot of microstrip realization (with circuit simulator) of stepped-impedance LPF, and (d) Layout of EM simulation for stepped-impedance LPF corresponding to Example 5.3.

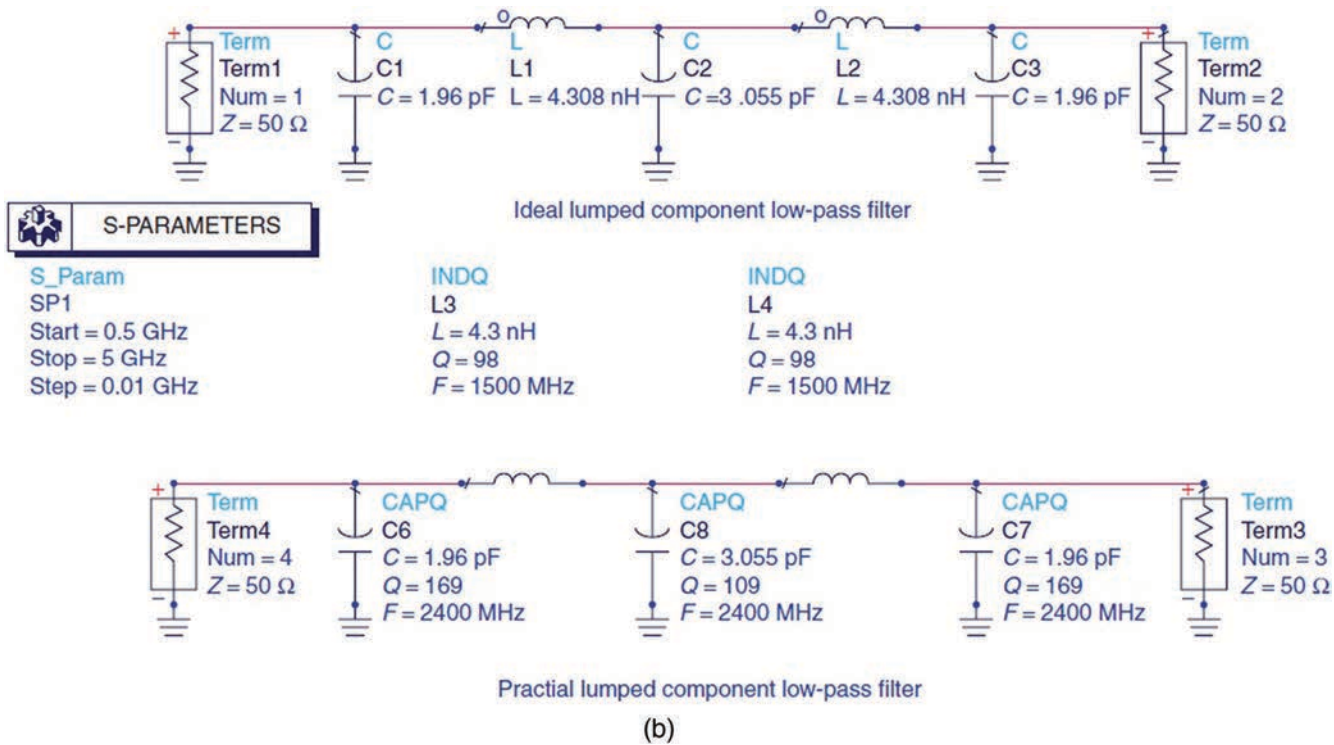


Figure 5.16 (Continued)

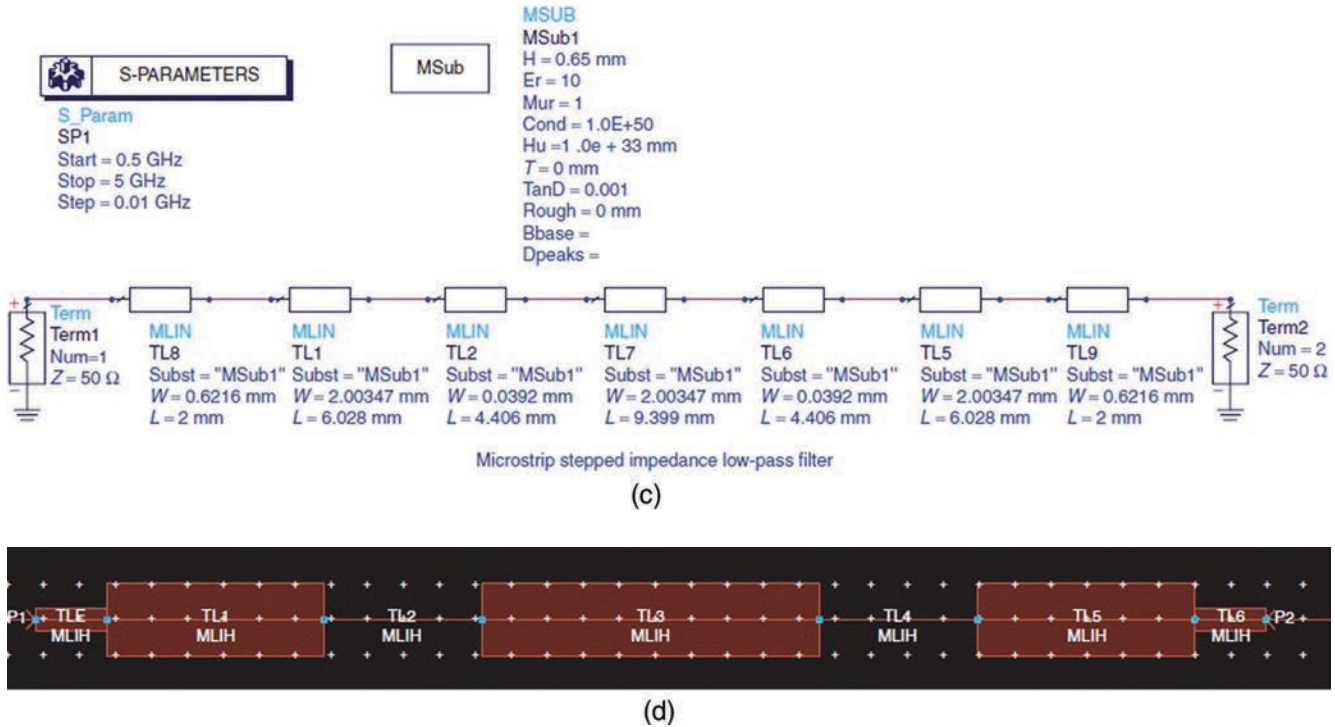


Figure 5.16 (Continued)

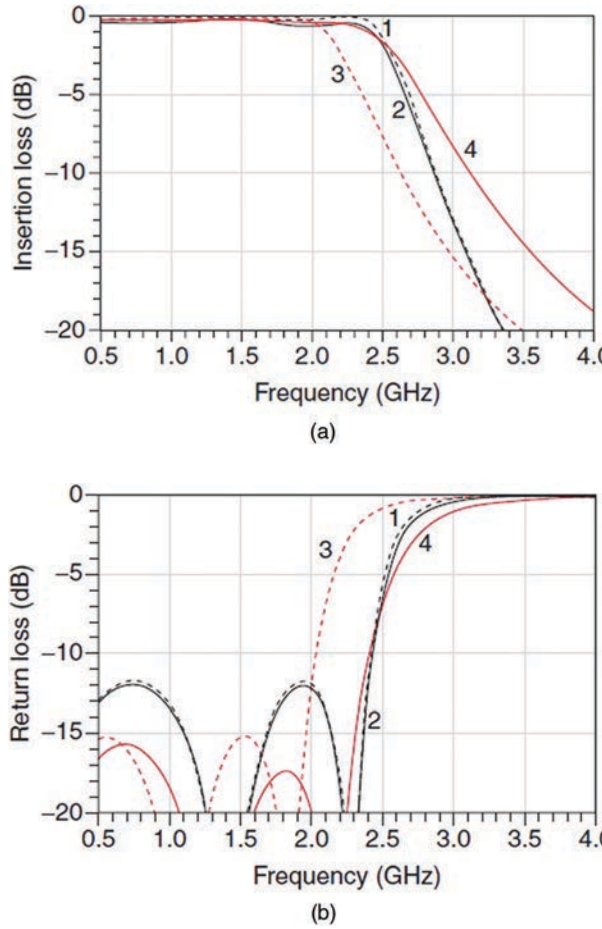


Figure 5.17 Stepped-impedance low-pass filter (LPF) characteristics obtained after optimization with EDA tool: (a) Transmission characteristics, and (b) Reflection characteristics.

particular filter, one often starts with a given set of filter specifications and an initial design is realized through analytical synthesis. Filter performance is then evaluated with the initial design data obtained above either with circuit simulator or EM simulation. If the results fail to satisfy the desired specifications/realizable dimensions, the desirable (optimization) parameters of the filter are altered in a systematic manner. The sequence of filter analysis, comparison with the desired performance, and modification of desirable parameters is performed iteratively until optimum performance of the filter is achieved. This process is known as *optimization*. The result of optimization carried out in layout (to have the design frequency close to the specified one) is

shown in Figure 5.17 (a) and (b); where Optimized: $l_{C1} = l_{C5} = 4.9$ mm, $l_{L2} = l_{L4} = 3.3$ mm, and $l_{C3} = 8$ mm. The return loss may be improved to bring it down to –20 dB that can be done with further iterations; this is left to the students as an exercise. We shall use such CAD tool based optimum design approach also for band-pass and band-stop filter design in the following sections.]

5.5.3 Coupled-Line Band-pass Filter

The band-pass filter (BPF) circuit can be derived from low-pass prototype (see Figure 5.6a and the table in Figure 5.7) which is shown in Figure 5.18a. It consists of series and parallel resonant circuits. Thus, basic design unit for such filter will be resonators which need to be multiple-coupled structure. When strip-line/microstrip line is used for the design, the resonators can be realized with parallel-coupled microstrip line as shown in Figure 5.18b [6]. The resonator elements are half-wavelength long at the centre frequency

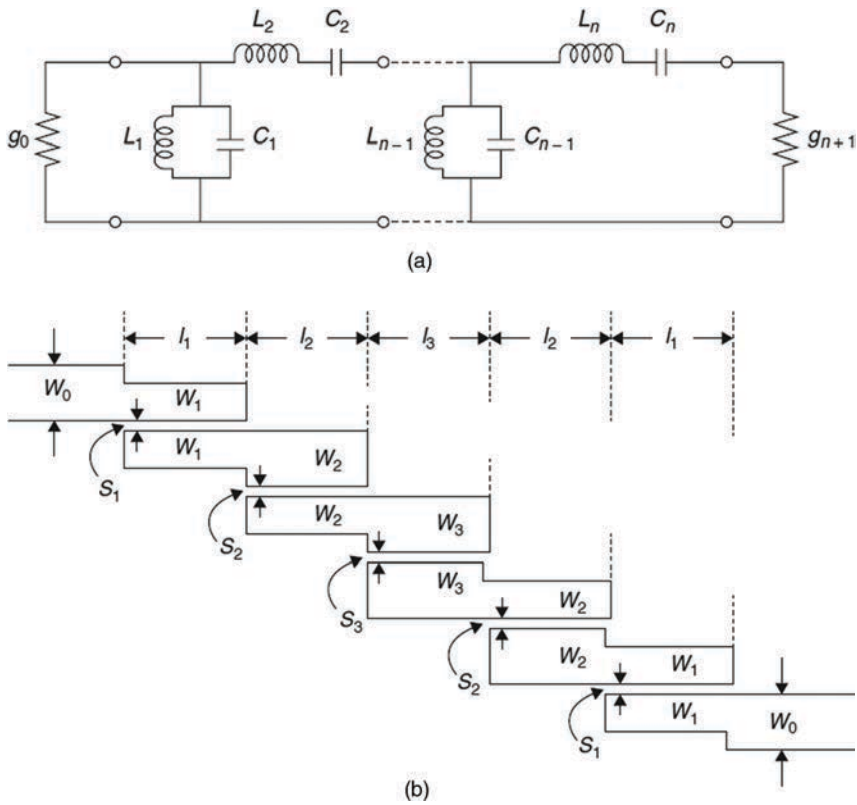


Figure 5.18 (a) Band-pass filter circuit derived from low-pass prototype, and (b) General microstrip configuration for 5-section parallel-coupled band-pass filter.

and open at both ends and maximum coupling occur between the resonators when the overlapping regions between the resonators, i.e., l_1, l_2, l_3, ….. l_n is quarter-wave (= $\lambda g/4$) long. The detailed derivation for the coupling and impedance of such parallel coupled filter may be obtained by using even- and odd-mode excitations which are summarized below for Tchebysheff response of n section BPF.

The BPF circuit (5.18a) consisting of parallel resonant circuit in the shunt arm and series resonant circuit in the series arm should be converted to either cascade of all parallel resonant networks together with impedance inverters K or to cascade of all series resonator networks together with admittance inverters J as shown in Figure 5.13.

The useful formulas for the admittances of the inverters when all series network configurations with admittance inverters J are used may be summarized as follows [1]:

For the first coupling section:

$$\frac{J_{10}}{Y_0} = \left\{\frac{\pi\Delta}{2g_0g_1}\right\}^{1/2} \tag{5.28}$$

For the final coupling section:

$$\frac{J_{n+1,n}}{Y_0} = \left\{\frac{\pi\Delta}{2g_ng_{n+1}}\right\}^{1/2} \tag{5.29}$$

For the intermediate coupling section:

$$\frac{J_{k+1,k}}{Y_0} = \left\{\frac{\pi\Delta}{2\omega_c' g_n\sqrt{g_kg_{k+1}}}\right\}^{1/2}, k = 1 \text{ to } (n-1) \tag{5.30}$$

where $\Delta = (\omega_{c2} - \omega_{c1})/\omega_0$ is the fractional bandwidth.

In the microstrip filter, the line dimensions are determined from the knowledge of even- and odd-mode coupled line impedances which are given in terms of admittances of the inverters.

$$(Z_{0e})_{k+1,k} = Z_0(1 + J_{k+1,k}Z_0 + J_{k+1,k}^2Z_0^2) \tag{5.31a}$$

and

$$(Z_{0o})_{k+1,k} = Z_0(1 - J_{k+1,k}Z_0 + J_{k+1,k}^2Z_0^2) \tag{5.31b}$$

where Z_0 is the characteristic impedance of the input/output lines of the filter.

Several workers described the variation of Z_{0o} and Z_{0e} as functions of W/h, S/h, and ε_r to evaluate the microstrip widths (W) and separations (S) of the parallel-coupled resonant lines. A suitable method is that of Akhtarzad et al. [7] synthesis process which starts with the even- and odd-mode impedance for single microstrip namely, Z_{0es} (= $Z_{0e}/2$) and Z_{0os} (= $Z_{0o}/2$) and evaluates $(W/h)_{es}$ and $(W/h)_{os}$ using standard expression for Z_0 with W/h for microstrip line. Then, the S/h and W/h ratios for the coupled lines are determined from the expressions:

$$\frac{S}{h} = \frac{2}{\pi}\cosh^{-1}\left[\frac{\cosh\left\{\frac{\pi}{2}(W/h)_{es}\right\} + \cosh\left\{\frac{\pi}{2}(W/h)_{os}\right\} - 2}{\cosh\left\{\frac{\pi}{2}(W/h)_{os}\right\} - \cosh\left\{\frac{\pi}{2}(W/h)_{es}\right\}}\right] \tag{5.32a}$$

$$(W/h)_{se} = \frac{2}{\pi}\cosh^{-1}\left(\frac{2d-g+1}{g+1}\right) \tag{5.32b}$$

$$(W/h)_{so} = \frac{2}{\pi}\cosh^{-1}\left(\frac{2d-g-1}{g-1}\right) \tag{5.32c}$$

$$\text{where } g = \cosh\left(\frac{\pi}{2}.\frac{S}{h}\right) \tag{5.32d}$$

$$\text{and } d = \cosh\left(\frac{\pi W}{h} + \frac{\pi S}{2h}\right) \tag{5.32e}$$

Using Equation (5.32a), one can calculate S/h when $(W/h)_{es}$ and $(W/h)_{os}$ are known. Finally using either Equations (5.32b) or (5.32c), one can determine W/h.

The physical lengths of the resonators and the coupled regions are, respectively, half- and quarter-wavelengths long. The coupled region length is determined in terms of the mid-band wavelength: $\lambda_{gm} = (\lambda_{go} + \lambda_{ge})/2$, where λ_{go} and λ_{ge} are the odd-mode and even-mode guide wavelengths, respectively, being calculated from:

$$\lambda_{go} = \frac{300}{f(GHz)}\frac{Z_{0o}}{Z_{0oair}}mm, \lambda_{ge} = \frac{300}{f(GHz)}\frac{Z_{0e}}{Z_{0eair}}mm \tag{5.33}$$

where $Z_{0e\ air}$, $Z_{0o\ air}$ are the even- and odd-mode characteristic impedances of the air-spaced coupled lines which can be calculated in terms of the

corresponding characteristic impedance of the single line viz. $Z_{0es\ air}$, $Z_{0os\ air}$; using: $Z_{0e\ air} = 2\ Z_{0es\ air}$ and $Z_{0o\ air} = 2\ Z_{0os\ air}$. The values for $Z_{0es\ air}$ and $Z_{0os\ air}$ can be calculated by using the standard plots for Z_0 vs W/h (with $\varepsilon_r = 1$), see Figure 5.15a.

Example 5.3 Design a parallel-coupled microstrip BPF having Tchebysheff characteristics with pass-band ripple of 0.2dB between the pass-band edge frequencies of 2.27GHz and 2.53GHZ. It is desired to have at least 30dB insertion loss at 2.65GHz in the stop-band. The filter impedance is 50 Ω. Consider a substrate of dielectric constant $\varepsilon_r = 10$ and height $h = 0.65$ mm.

Solution:

This problem will be solved both by analytical synthesis process and also by using EDA tool.

METHOD I: Analytical Synthesis Process:

STEP 1: At first, we determine the number of sections for the filter from the specified attenuation characteristics. The center frequency: $f_0 = \sqrt{f_{c1} f_{c2}} = \sqrt{2.27 \times 2.53} = 2.4$GHz and the fractional bandwidth: $\Delta = (f_{c2} - f_{c1})/f_0 = (2.53 - 2.27)/2.4 = 0.108$.

Given, equal-ripple magnitude (A_{pm}) as: $10 \log_{10}(1 + A_p^2) = 0.2$ dB, when $A_p^2 = 0.047$

The power loss ratio of the BPF having equal-ripple, i.e., Tchebysheff response is given by

$$P_{LR} = 1 + A_p^2 T_n^2(\Omega'); \text{ where } T_n(x) = \cosh(n \cosh^{-1} \Omega').$$

$$\text{Here: } \Omega' = \frac{\Omega_p \omega_0}{\omega_{c2} - \omega_{c1}} \left(\frac{\omega_s}{\omega_0} - \frac{\omega_0}{\omega_s} \right) = \frac{2.4}{2.53 - 2.27} \left(\frac{2.65}{2.4} - \frac{2.4}{2.65} \right) =$$

$$1.83 > 1; \text{ (with } \Omega_p = 1)$$

Substituting respective values ($P_{LRs} = 30$ dB and $\Omega' = 1.83$) in the expression for P_{LR} and solving for n, we get $n = 4.679$, which can be taken approximately as 5.

Therefore, a five-section BPF is required.

STEP 2: Now we need to determine the values of the prototype elements to realize the specifications.

For 0.2dB ripple with n = 5, the low-pass prototype elements are: $g_1 = 1.3394 = g_5$, $g_2 = 1.337 = g_4$, and $g_3 = 2.166$ (see the table (a)(ii) in Figure 5.4); while $g_0 = g_6 = 1.0$ (since n is odd).

STEP 3: Next, we determine the inverter values (say we use the J inverter, i.e., with cascaded series resonators):

As per symmetry, for the first and final coupling, we have

$$\frac{J_{10}}{Y_0} = \frac{J_{6,5}}{Y_0} = \sqrt{\frac{\pi\Delta}{2g_0g_1}} = \sqrt{\frac{\pi\Delta}{2g_5g_6}} = \sqrt{\frac{\pi\times 0.108}{2\times 1\times 1.3394}} = 0.356$$

For the intermediate coupling section (k = 1 to 2)

$$\frac{J_{2,1}}{Y_0} = \frac{\pi\Delta}{2\sqrt{g_1g_2}} = \frac{J_{5,4}}{Y_0} = \frac{\pi\times 0.180}{2\sqrt{(1.3394\times 1.337)}} = 0.127$$

$$\frac{J_{3,2}}{Y_0} = \frac{\pi\Delta}{2\sqrt{g_2g_3}} = \frac{J_{4,3}}{Y_0} = \frac{\pi\times 0.180}{2\sqrt{(1.337\times 2.166)}} = 0.0997$$

STEP 4: Then we have to obtain the even- and odd-mode coupled line characteristic impedances Z_{0e} and Z_{0o} (refer to Eq. (5.31a) and Eq. (5.31b)).

Looking into the symmetry, only k = 0 to 2 are required to be computed, when k = 3 and k = 4 will, respectively, take the values of k = 1 and k = 0. Following Equations (5.31a) and (5.31b), we have

For $k = 0$, $(Z_{0e})_{1,0} = Z_0(1 + J_{1,0}Z_0 + J_{1,0}^2Z_0^2) = 50\ [1 + 0.356 + (0.356)^2] = 74.14\ \Omega$

$(Z_{0o})_{1,0} = Z_0(1 - J_{1,0}Z_0 + J_{1,0}^2Z_0^2) = 50\ [1 - 0.356 + (0.356)^2] = 38.54\ \Omega$

For $k = 1$, $(Z_{0e})_{2,1} = Z_0(1 + J_{2,1}Z_0 + J_{2,1}^2Z_0^2) = 50\ [1 + 0.127 + (0.127)^2] = 57.16\ \Omega$

$(Z_{0o})_{2,1} = Z_0(1 - J_{2,1}Z_0 + J_{2,1}^2Z_0^2) = 50\ [1 - 0.127 + (0.127)^2] = 44.46\ \Omega$

For $k = 2$, $(Z_{0e})_{3,2} = Z_0(1 + J_{3,2}Z_0 + J_{3,2}^2Z_0^2) = 50\ [1 + 0.0997 + (0.0997)^2] = 55.48\ \Omega$

$(Z_{0o})_{3,2} = Z_0(1 + J_{3,2}Z_0 + J_{3,2}^2Z_0^2) = 50\ [1 - 0.0997 + (0.0997)^2] = 45.51\ \Omega$

STEP 5: Next, we need to determine the microstrip widths (W) and separations (S) of the parallel-coupled resonant lines. To do this, we use Akhtarzad et al. synthesis approach when the even- and odd-mode impedances for single microstrip for different values of k are given by

For $k = 0$, $(Z_{0es})_{1,0} = (Z_{0e})_{1,0}/2 = 74.14/2 = 37.07\ \Omega$

$$(Z_{0os})_{1,0} = (Z_{0o})_{1,0}/2 = 38.54/2 = 19.27\ \Omega$$

For $k = 1$, $(Z_{0es})_{2,1} = 57.1/2 = 28.58\ \Omega$, $(Z_{0os})_{2,1} = 44.46/2 = 22.23\ \Omega$

For $k = 2$, $(Z_{0es})_{3,2} = 55.48/2 = 27.74\ \Omega$, $(Z_{0os})_{3,2} = 45.51/2 = 22.75\ \Omega$

Then using, Z_0 (in terms of W/h) and inserting value of ε_{reff} in it [6], we can get

$$Z_0 = \frac{60\ln\left[\frac{8h}{W} + \frac{W}{4h}\right]}{\left[\frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left\{\left(1 + \frac{12h}{W}\right)^{-1/2} + 0.04\left(1 - \frac{W}{h}\right)^2\right\}\right]^{1/2}} \quad \text{for } (\text{W/h}) \le 1$$

$$= \frac{120\pi}{\left[\frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + \frac{12h}{W}\right)^{-1/2}\right]^{1/2}\left[\frac{W}{h} + 1.4 + 0.667\ln\left(\frac{W}{h} + 1.444\right)\right]} \quad \text{for } (\text{W/h}) > 1$$

Now for every Z_{0es} and Z_{0os}, we calculate the corresponding $(W/h)_{es}$ and $(W/h)_{os}$ from above equation when we get:

$$\left[(W/h)_{es}\right]_{1,0} = 1.67\,,\ \left[(W/h)_{es}\right]_{2,1} = 2.512\,,\ \left[(W/h)_{es}\right]_{3,2} = 2.63$$

$$\left[(W/h)_{os}\right]_{1,0} = 4.41\,,\ \left[(W/h)_{os}\right]_{2,1} = 3.66\,,\ \left[(W/h)_{os}\right]_{3,2} = 3.52$$

Next the corresponding S/h values can be calculated from Eq. (5.32a), when we have

For $k = 0$, $S/h = 0.138$, $k = 1$, $S/h = 0.538$, $k = 2$, $S/h = 0.684$.

Then, the W/h ratio for the coupled line can be calculated from any of the two relations given in Eq. (5.32b) and Eq. (5.32c) which yields

For: $k = 0$, $W/h = 0.77$, $k = 1$, $W/h = 1.044$, $k = 2$, $W/h = 1.06$.

STEP 6: Finally, we calculate (following Eq. (5.33)) the physical lengths of the resonators and the coupled regions which are respectively

Table 5.1 Final design data for the resonators

Coupled lines	*W/h*	*W (mm)*	*S/h*	*S (mm)*	$l = \lambda_{gm}/4$ *(mm)*
$k = 0$	0.945	0.61425(W_1)	0.125	0.08125(S_1)	11.214 (l_1)
$k = 1$	1.157	0.75205(W_2)	0.422	0.2743 (S_2)	11.391 (l_2)
$k = 2$	1.17	0.7605 (W_3)	0.5006	0.32539(S_3)	11.3945(l_3)
$k = 3$	1.157	0.75205(W_2)	0.422	0.2743 (S_2)	11.391 (l_2)
$k = 4$	0.945	0.61425(W_1)	0.125	0.08125(S_1)	11.214 (l_1)

half- and quarter-wavelengths long. The final design data for the resonators is given in Table 5.1 (note: $W_0 = 0.65$ mm, with $Z_0 = 50\ \Omega$ and $\varepsilon_r = 10$, see Figure 5.15a). These data values given in the table may be obtained either with analytical calculations or may be evaluated by writing a MATLAB program (see Appendix II).

METHOD II: Using EDA Tool: The EDA tool generated BPF filter characteristics with microstrip realization (curve 1 is obtained with circuit simulator) and with complete EM simulation (curve 2) has been shown in Figure 5.19a. The layout for coupled microstrip BPF obtained with EM simulation is also shown in Figure 5.19b together with the screen shot of optimization in layout is depicted in Figure 5.19c.

The optimized design data of the EM simulated BPF shown in Figure 5.18b is given in Table 5.2 (see Figure 5.18b for *W*, *S*, and *l*).

5.5.4 Coupled-Resonator Band-stop Filter

The band-stop filter (BSF) circuit can be derived from low-pass prototype (see Figure 5.6b and the table in Figure 5.7) which is shown in Figure 5.20(a). It consists of parallel and series resonant circuits and hence similar to the case of BPF. Here also we need to use either all series or all parallel resonators with the application of J/K inverter concept. When this is done, the lumped element ladder network of Figure 5.20(a) may be equivalently replaced by all series resonators with quarter-wave coupling network as shown in Figure 5.20(b) or with all shunt resonators with quarter-wave resonators as shown in Figure 5.20(c). The design procedure for this type of filter using microstrip line is summarized below [1].

Let Z_0 is the characteristic impedance of the main microstrip line, and Z_1 is the characteristic impedance of the quarter-wave coupling networks (i.e., of the inverters) then:

$$Z_1 = Z_0, \quad \text{for } n \text{ odd} \quad \text{and} \quad Z_c = Z_1(1/g_0 g_{n+1})^{1/2}, \quad \text{for } n \text{ even}$$

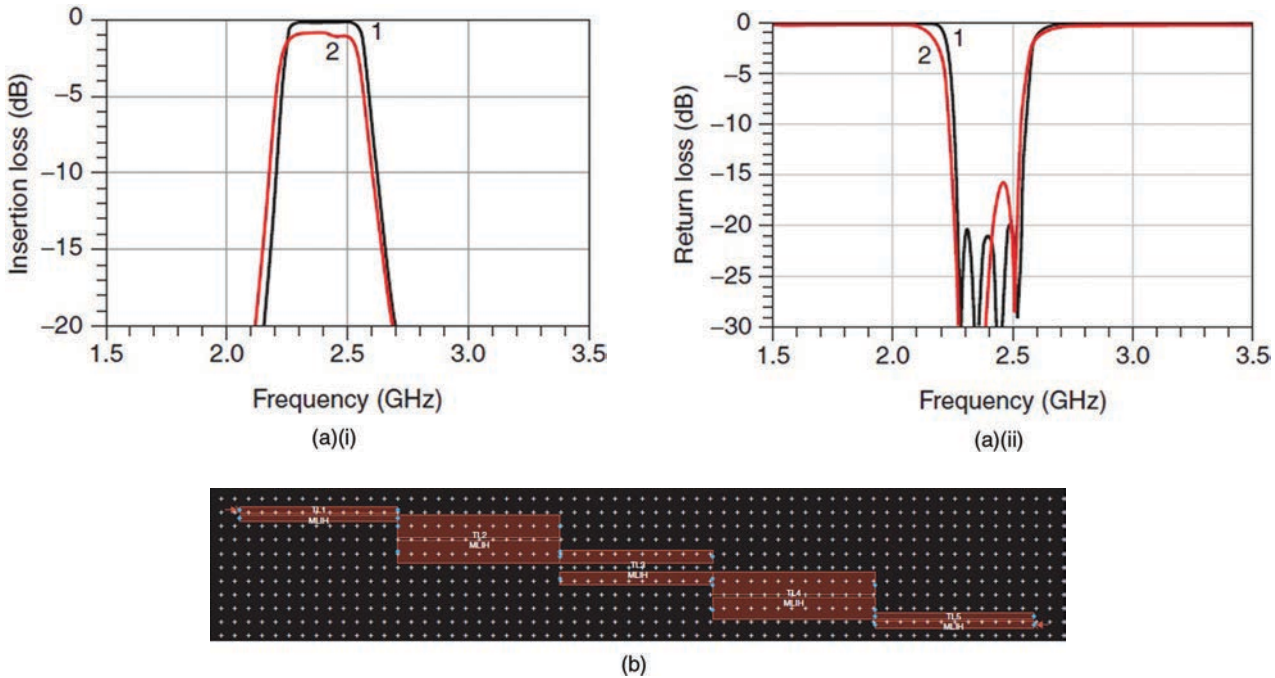


Figure 5.19 (a) Coupled-line microstrip band-pass filter (BPF) characteristics obtained with EDA tool: (i) Insertion loss characteristics, and (ii) Return loss characteristics, (b) Layout of EM simulation for coupled microstrip BPF corresponding to Example 5.3, and (c) Screenshot of BPF design optimization in layout.

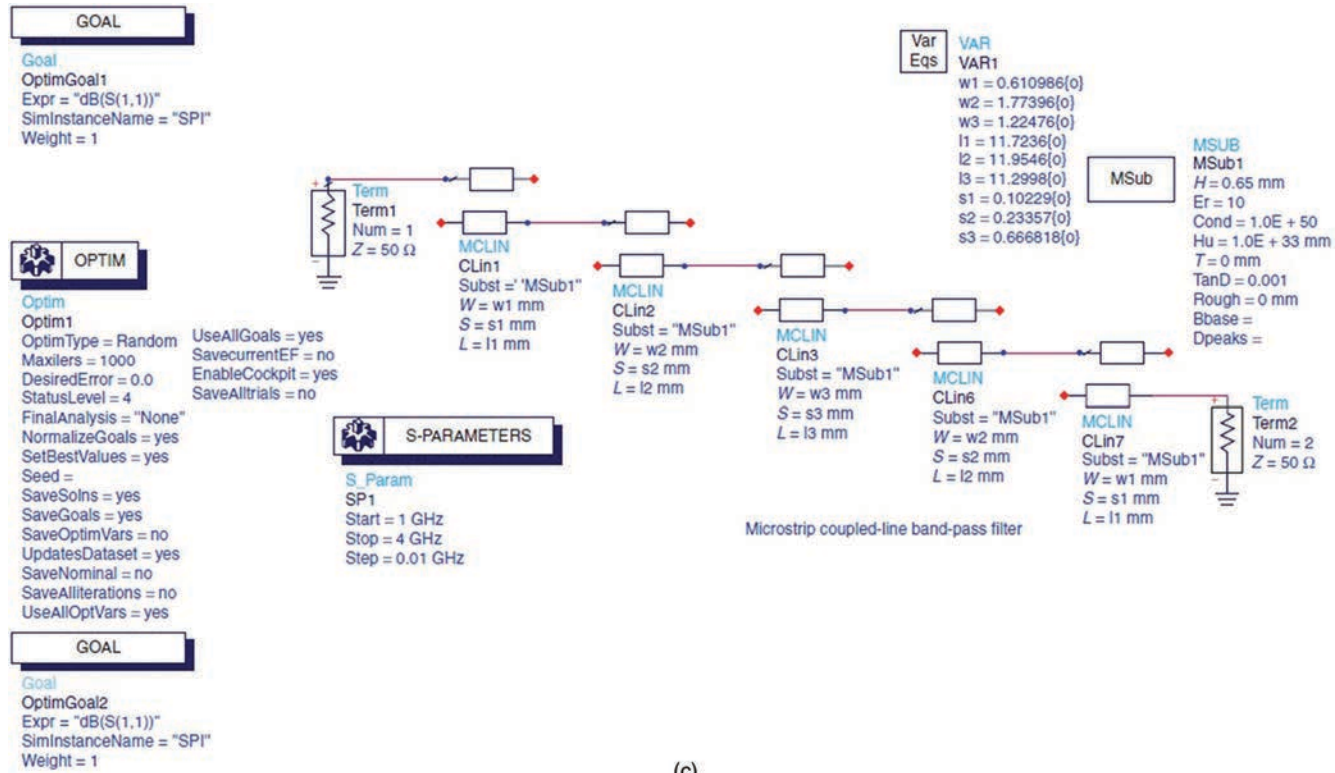


Figure 5.19 (Continued)

Table 5.2 Optimized design data of the EM simulated BPF

W (mm)	*S (mm)*	*l (mm)*
$W_1 = 0.61$	$S_1 = 0.102$	$l_1 = 11.72$
$W_2 = 0.77$	$S_2 = 0.23$	$l_2 = 11.95$
$W_3 = 1.22$	$S_3 = 0.67$	$l_3 = 11.3$

The filter design is carried out by defining *reactance slope parameters*: $x = \frac{\omega_0}{2}\frac{dX}{d\omega}\Big|_{\omega=\omega_0}$, with $X = (\omega L - 1/\omega C)$ of the resonators given by

$$\frac{x_1}{Z_0} = \frac{1}{g_0 g_1 \Delta}, \frac{x_k}{Z_0}\Big|_{k(even)} = \left(\frac{Z_I}{Z_0}\right)^2 \frac{g_0}{g_k \Delta}, \frac{x_k}{Z_0}\Big|_{k(odd),\neq 1} = \left(\frac{Z_I}{Z_0}\right)^2 \frac{1}{g_0 g_k \Delta}$$

Where $\Delta = (\omega_{c2} - \omega_{c1})/\omega_0$, is the fractional bandwidth.

In strip-line/microstrip realization, the inductances in series/parallel resonant circuits can be realized with stub (less than quarter-wave long if realized with short-circuited stub and less than half-wavelength long if realized with open-circuited stub) and the capacitance by means of a gap (g) between the main microstrip line and the stub (Figure 5.20d). The coupling between the main microstrip line and the resonators may be electrical as in Figure 5.20d or may also be magnetic (with hair-pin type resonator) as shown in Figure 5.20e. For practical purposes (as will be seen while solving Example 5.4) to realize significant capacitance through the gap (g), L-type resonators (that combine both electrical and magnetic coupling) may be used (Figure 5.20f) [2]. Anyway, when electric coupling as shown in Figure 5.20d is used, the length of stub is determined from: $l_k = \phi_l/\beta_{eff} = \phi_l \lambda_{eff}/2\pi$, where φ_l is the phase delay incurred by the signal in propagating via the line. Evaluation of φ_l is done by using the formula: $F(\phi_l) = \phi_l \sec^2 \phi_l + \tan \phi_l$, which is normally evaluated numerically and a table for $F(\phi_l)$ vs ϕ_l is given in Table 5.3. If short-circuited stub is used then $F(\phi_l) = 2x_k/Z_{c(stub)}$ but if open-circuited stub is used then $F(\phi_l) = x_k/Z_{c(stub)}$; where $Z_{c(stub)}$ is the characteristic impedance of the line forming the stub. The series gap capacitance is determined from: $C_{gk} = 1/(\omega_0 Z_{c(stub)} \tan \phi_l)$.

Design of such filter using waveguide is available in advanced references [1].

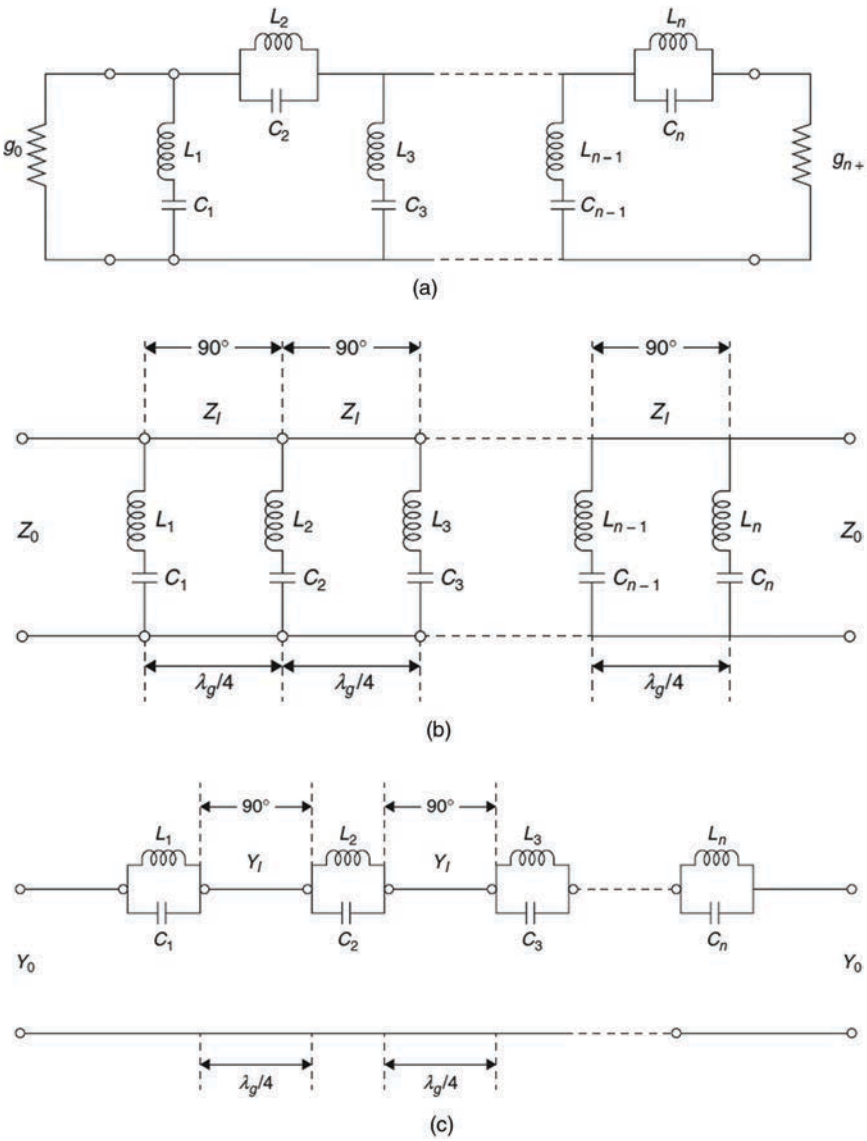


Figure 5.20 (a) Band-stop filter (BSF) circuit derived from low-pass prototype, (b) BSF equivalent circuit with series resonators, (c) BSF equivalent circuit with shunt resonators, (d) Electric coupling of stub resonators with main microstrip line, (e) Magnetic coupling of hair-pin type resonators with main microstrip line, and (f) L-type resonators coupled with the main microstrip line.

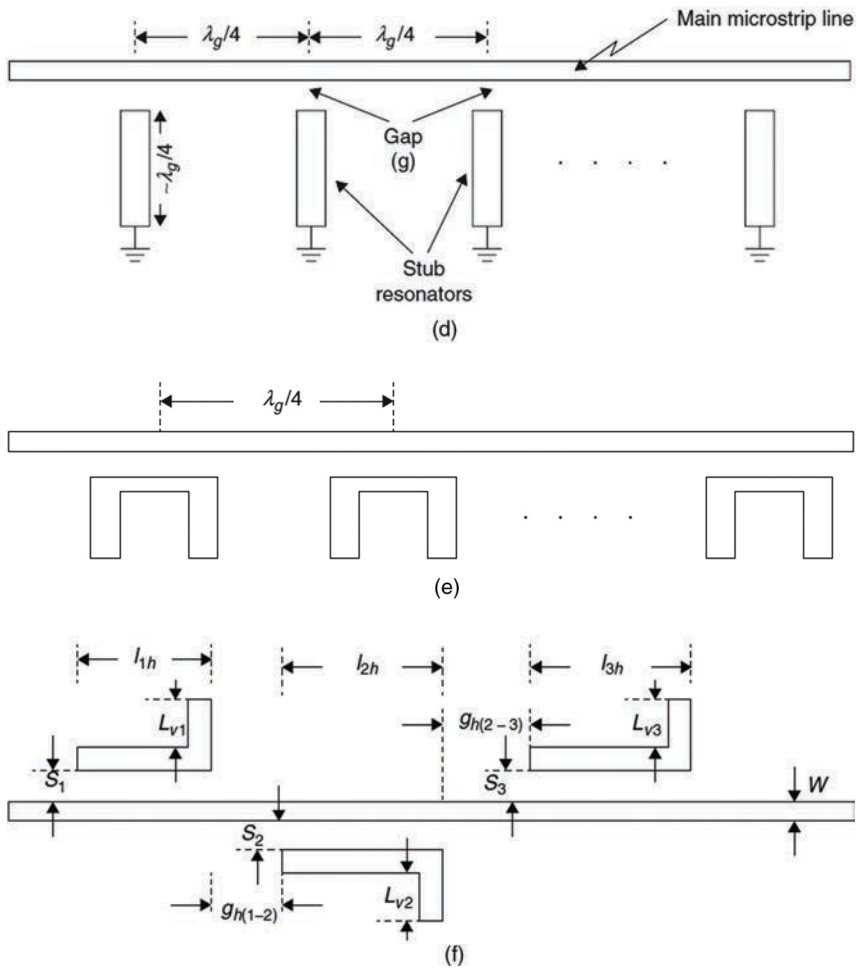


Figure 5.20 (Continued)

Example 5.4 Design a strip-line based band-stop filter with Tchebysheff response having frequency of infinite attenuation at 2.4GHz, fractional bandwidth 0.05 (i.e., narrow-band type), pass-band ripple 0.1dB having 20dB minimum attenuation occurring at 2% of the center frequency. Given the characteristic impedance of strip-line is 50 Ω with dielectric substrate of ε_r = 10 and height h = 0.65 mm. Show calculations using open-circuited stub and also short-circuited stub lines having nominal characteristic impedance of 70Ω to realize the inductance of the series resonators.

Solution:

This problem will be solved both by analytical synthesis process and also by using EDA tool.

METHOD I: Analytical Synthesis Process:

STEP 1: At first, we determine the number of sections for the filter.

Given pass-band ripple magnitude (A_{pm}) as: 10 $\log_{10}(1 + A_p^2)$ = 0.1 dB, when $A_p^2 = 0.023$.

The power loss ratio of the band-stop filter having equal-ripple, i.e., Tchebysheff response is given by

$$P_{LR} = 1 + A_p^2 T_n^2(\Omega'); \text{ where } T_n(x) = \cosh(n \cosh^{-1} \Omega').$$

$$\text{Here } \Omega' = \frac{\Omega_p \omega_0}{\omega_{c2} - \omega_{c1}} \left(\frac{\omega_s}{\omega_0} - \frac{\omega_0}{\omega_s} \right)^{-1} = 0.05 \left(\frac{2.424}{2.4} - \frac{2.4}{2.424} \right)^{-1}$$

$$= 2.5 > 1; \text{ (with } \Omega_p = 1)$$

(Since the frequency at 20 dB attenuation is f_s = (2.4 GHz + 0.5 × 2% of 2.4 GHz) = 2.424 GHz) and the fractional bandwidth $\Delta = (\omega_{c2} - \omega_{c1})/\omega_0 = 0.05$).

Substituting respective values (P_{LRs} = 20 dB and $\Omega' = 2.5$) in the expression for P_{LR} and solving for n we get $n = 3.108$, which can be taken approximately as 3.

Therefore, a three-section band-stop filter is required.

STEP 2: Now we need to determine the values of the prototype elements to realize the specifications.

For 0.1dB ripple with $n = 3$, the low-pass prototype elements are: $g_1 = 1.0315 = g_3$ and $g_2 = 1.1474$ (see table (a)(i) of Figure 5.4); while $g_0 = g_4 = 1.0$.

STEP 3: We can now proceed for RF/microwave realization of the filter using strip-line circuit. Here $n = 3$ (odd), thus the impedance of the quarter-wave long coupling lines will be $Z_I = Z_0$. The reactance slope parameter of the series resonators is thus given by

$$\frac{x_1}{Z_0} = \frac{1}{g_0 g_1 \Delta} = \frac{x_3}{Z_0} = \frac{1}{1 \times 1.0315 \times 0.05} = 19.32$$

$$\frac{x_2}{Z_0} = \left(\frac{Z_I}{Z_0} \right)^2 \frac{g_0}{g_k \Delta} = \frac{1}{1.1474 \times 0.05} = 17.43 \left(\text{Since } Z_I = Z_0 \right)$$

Table 5.3 Values of $F(\varphi_l)$ for different values of φ_l (in degrees) obtained by numerical solution of equation: $F(\phi_l) = \phi_l \sec^2 \phi_l + \tan \phi_l$

φ_l	$F(\varphi_l)$	φ_l	$F(\varphi_l)$	φ_l	$F(\varphi_l)$	φ_l	$F(\varphi_l)$
15	0.5485	35	1.6106	55	4.3459	75	23.2721
16	0.5890	36	1.6865	56	4.6082	76	26.6739
17	0.6302	37	1.7660	57	4.8935	77	30.8879
18	0.6722	38	1.8493	58	5.2051	78	36.1958
19	0.7153	39	1.9368	59	5.5462	79	43.0132
20	0.7593	40	2.0288	60	5.9208	80	51.9730
21	0.8044	41	2.1256	61	6.3336	81	64.0787
22	0.8507	42	2.2277	62	6.7903	82	80.9982
23	0.8982	43	2.3356	63	7.2974	83	105.6711
24	0.9471	44	2.4498	64	7.8628	84	143.6788
25	0.9975	45	2.5708	65	8.4961	85	207.7040
26	1.0495	46	2.6993	66	9.2088		
27	1.1031	47	2.8360	67	10.0151		
28	1.1586	48	2.9817	68	10.9322		
29	1.2160	49	3.1373	69	11.9819		
30	1.2755	50	3.3038	70	13.2301		
31	1.3372	51	3.4824	71	14.5948		
32	1.4014	52	3.6743	72	16.2921		
33	1.4683	53	3.8810	73	18.1752		
34	1.5378	54	4.1043	74	20.4861		

Therefore $x_1 = x_3 = 19.32 \times 50 = 966\Omega$ and $x_2 = 17.43 \times 50 = 871.5\Omega$

STEP 4: Now, if we use *open-circuited* stub lines (with characteristic impedance of the line forming the stub $Z_{c(stub)} = 70\Omega$) to realize the inductance of the series resonators then:

$$F(\phi_{l1}) = F(\phi_{l3}) = \phi_{l1} \sec^2 \phi_{l1} + \tan \phi_{l1} = \frac{x_1}{Z_{c(stub)}} = \frac{966}{70} = 13.8$$

$$F(\phi_{l2}) = \phi_{l2} \sec^2 \phi_{l2} + \tan \phi_{l2} = \frac{x_2}{Z_{c(stub)}} = \frac{871.5}{70} = 12.45$$

From the $F(\varphi_l)$ vs φ_l values (Table 5.3), we find:

$\varphi_{l1} = \varphi_{l3} = 70.5^0 = 1.2309$ radian, $\varphi_{l2} = 69.5^0 = 1.2134$ radian

Thus, the stub lengths are given by

$$l_1 = l_3 = \phi_{l1,3} \lambda_{eff} / 2\pi = 1.2309 \times 3.972 / 2\pi = 0.7781\text{cm}$$

$$l_2 = \phi_{l2}\lambda_{eff}/2\pi = 1.2134 \times 3.972/2\pi = 0.767\text{cm}$$

(Since $\lambda_{eff} = \lambda_0/\sqrt{\varepsilon_r} = (30/2.4)/\sqrt{9.9} = 3.972\,\text{cm}$, and $\lambda_{eff}/2 = 1.986\,\text{cm}$).

STEP 5: If we use *short-circuited* stub lines to realize the inductance of series resonator then:

$$F(\phi_{l1}) = F(\phi_{l3}) = \phi_{l1}\sec^2\phi_{l1} + \tan\phi_{l1} = \frac{2x_1}{Z_{c(stub)}} = \frac{2\times 966}{70} = 27.6$$

$$F(\phi_{l2}) = \phi_{l2}\sec^2\phi_{l2} + \tan\phi_{l2} = \frac{2x_2}{Z_{c(stub)}} = \frac{2\times 871.5}{70} = 24.9$$

From the $F(\varphi_l)$ vs φ_l values (Table 5.3), we find:

$\varphi_{l1} = \varphi_{l3} = 76.2^0 = 1.3304$ radian, $\varphi_{l2} = 75.5^0 = 1.3182$ radian

Thus, the stub lengths are given by

$$l_1 = l_3 = \phi_{l1,3}\lambda_{eff}/2\pi = 1.3304 \times 3.972/2\pi = 0.8410\text{ cm}$$

$$l_2 = \phi_{l2}\lambda_{eff}/2\pi = 1.3182 \times 3.972/2\pi = 0.8333\text{ cm}$$

STEP 6: The gap capacitances can be calculated from: $C_{gk} = 1/(\omega_0 Z_{c(stub)}\tan\phi_l)$, where $Z_{c(stub)} = 70\Omega$, assumed to be equal for all the stubs. Thus, when *open-circuited* stub is used:

$$C_{g1,3(o.c)} = (1/2\pi \times 2.4\times 10^9 \times 70 \times \tan 70.5^0) = 0.335\text{ pF};$$

$$\text{and } C_{g2(o.c)} = (1/2\pi \times 2.4\times 10^9 \times 70 \times \tan 69.5^0) = 0.354\text{ pF};$$

While if *short-circuited* stub is used:

$$C_{g1,3(s.c)} = (1/2\pi \times 2.4\times 10^9 \times 70 \times \tan 76.2^0) = 0.232\text{ pF};$$

$$\text{and } C_{g2(o.c)} = (1/2\pi \times 2.4\times 10^9 \times 70 \times \tan 75.5^0) = 0.245\text{ pF};$$

STEP 7: Finally, we need to determine the line widths of the strip-line transmission line to complete our design. For the main line characteristic impedance is 50Ω and for the resonator line it is 70Ω; while the dielectric constant of the substrate material for the strip-line is $\varepsilon_r = 10$ in either case. Referring to Figure 5.15a, we can calculate the strip-width as follows.

For 50 Ω line, $\frac{W}{h} = 1.0$, thus $W = 1.0 \times 0.65 = 0.65$ mm

For 70 Ω line, $\frac{W}{h} = 0.49$, thus $W = 0.49 \times 0.65 = 0.3185$ mm.

METHOD II: Using EDA Tool: The microstrip realization of the coupled resonator band-stop filter (with circuit simulator of the EDA tool) using the dimensions obtained through analytical synthesis is shown in Figure 5.21a (curve 1), which practically does not depict any band-stop characteristics whatsoever. This is because the realizable gap capacitance between the host microstrip line and individual stub resonators is very small. However, the use of the calculated capacitance values corresponding to the gap (g) shows significant band-stop performance when simulated in the circuit simulator of the EDA tool (curve 2 in Fig 5.21a). The curve 3 in Figure 5.21a obtained with EM simulation of the filter structure when the resonators are of L-type (that provides much larger capacitance between the host microstrip line and the resonators) shows pretty good band-stop characteristics when compared with conventional stub resonators (curve 1 in Fig 5.21a) as is used in the analytical synthesis. The layout of EM simulation with L-type resonators is shown in Figure 5.21b together with the screenshot of the optimization in layout as shown in Figure 5.21c.

The optimized design data of the EM simulated BSF shown in Figure 5.21b is given in Table 5.4; (see Figure 5.20f for l_h, l_v, g_h, S).

The width of the host transmission line of characteristic impedance 50 Ω is given by $W = 0.65$ mm (see Figure 5.15a, with $Z_0 = 50$ Ω and $\varepsilon_r = 10$).

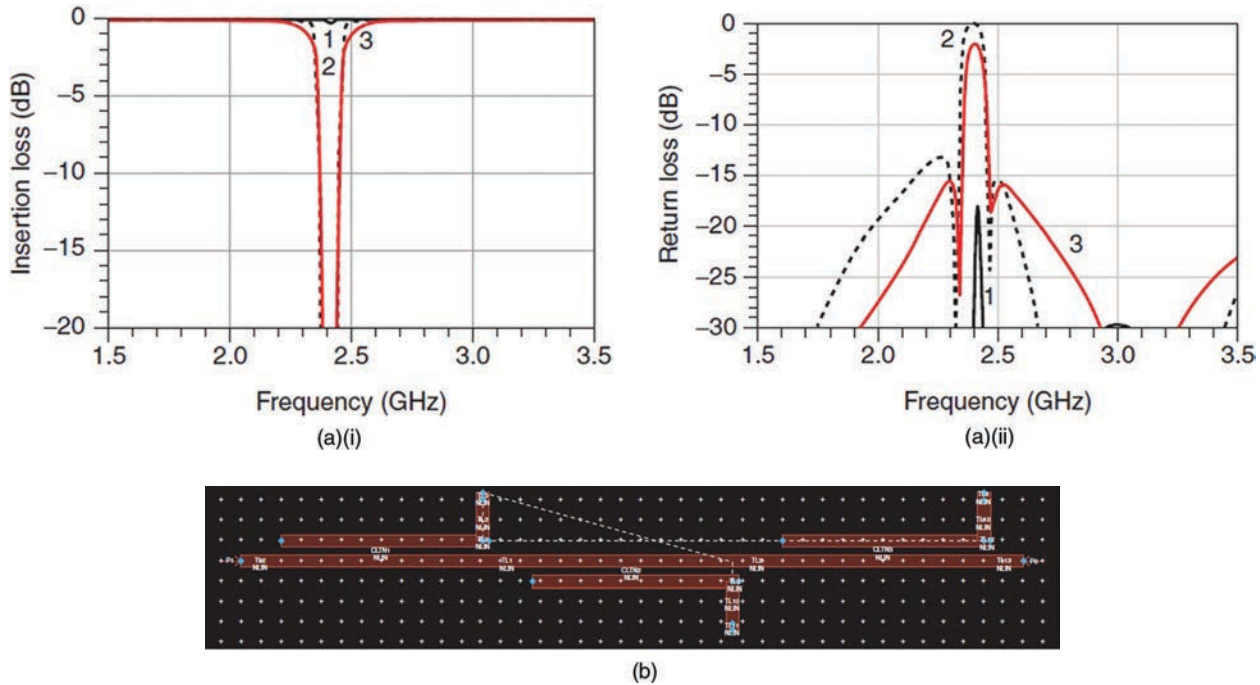


Figure 5.21 (a) Microstrip band-stop filter (BSF) characteristics: (i) Insertion loss characteristics, and (ii) Return loss characteristics, (b) Layout of EM simulation with L-shaped resonator for coupled resonator BSF corresponding to Example 5.4, and (c) Screenshot of BSF design optimization in layout.

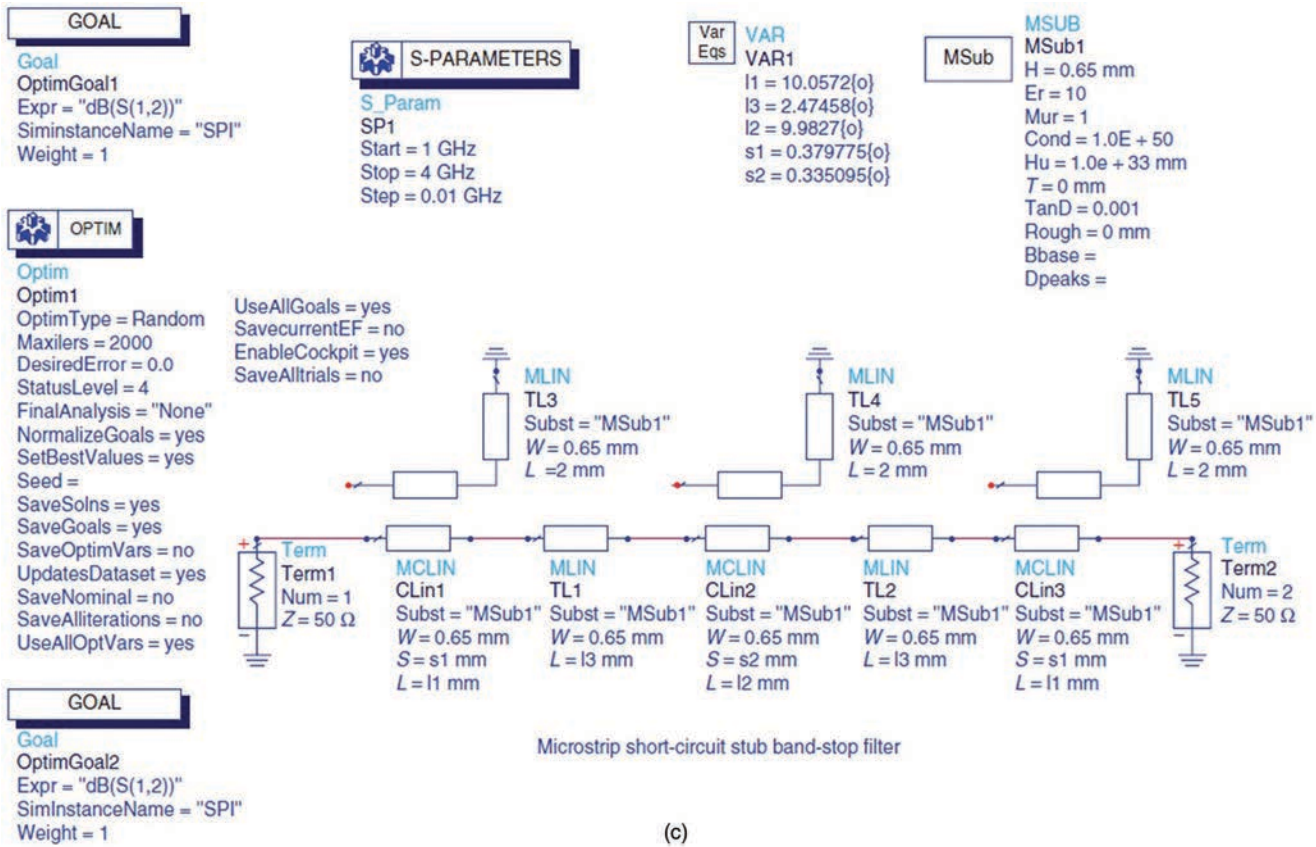


Figure 5.21 (Continued)

Table 5.4 Optimized design data of the EM-simulated BSF

l_h (mm)	l_v (mm)	g_h (mm)	S (mm)
$l_{h1} = 10 = l_{h3}$ $l_{h2} = 9.98$	$l_{v1} = l_{v2} = l_{v3} = 2.5$	$g_{h(1-2)} = 2.47$ $g_{h(2-3)} = 2.47$	$S_1 = 0.38 = S_3$ $S_2 = 0.335$

5.6 METAMATERIAL-INSPIRED FILTER DESIGN AT MICROWAVE FREQUENCY

5.6.1 Introduction

Metamaterials are artificially structured material whose properties depend not on the chemistry of the material with which it is made of but on the structural arrangement of the artificial "electric atoms" and "magnetic atoms" of which it is constituted. Metamaterials or left-handed materials (LHM) are known to exhibit counter-intuitive properties, not found in nature (hence the name meta-material, in Greek "meta" means beyond nature); like reversal of Snell's law, reversal of Doppler Effect, and reversal of Cherenkov Effect. It has many exotic applications like cloaking of objects, super-microscope, and so forth. However, metamaterial inclusions can also be used to miniaturize with improved performance in the design of RF/microwave components like antennas, filters etc. Metamaterials can be either plasmonic type or transmission line type [8,9]; the former one can be designed with a combination of split-ring resonator (SRR) as magnetic inclusion structure and thin-wire (TW) as electric inclusion structure. Transmission line metamaterial can be designed with periodic loading (with series capacitance and shunt inductance) of host transmission line with unit cell dimension much smaller than the wavelength of the signal. A transmission line metamaterial (usually designed with the host microstrip transmission line, which is a natural material or the so-called right-handed material, RHM) is thus, in general, a composite right-/left-handed (CRLH) line [9]. In the following we would discuss the design of SRR-loaded waveguide filter and CRLH-type band-pass filter (BPF).

5.6.2 SRR-Loaded Waveguide Filter

SRR-loaded waveguide filter is a new type of RF/microwave filter [10] in which negative permeability magnetic inclusion structures, SRR, conventionally used in metamaterial design, have been used judiciously to realize sharp-cut-off notch filter [11]. A schematic diagram showing rectangular waveguide loaded with SRR matrices in 1D structure is shown in Figure 5.22a together with an SRR unit cell in Figure 5.22b. Such loading of waveguide with μ-negative structures would allow us to introduce selective

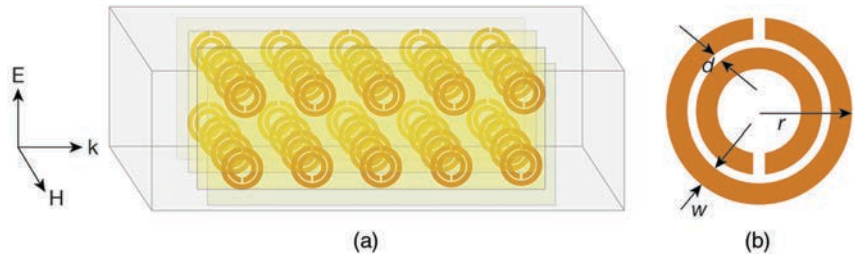


Figure 5.22 (a) Rectangular waveguide loaded with sheets of 1-D SRR matrices, and (b) SRR unit cell.

absorption, i.e., stop-band in the pass-band of normal waveguide propagation and selective transmission, i.e., pass-band in the stop-band of the waveguide propagation.

Referring to Figure 5.23a, it may be observed that in the transmission characteristics it is possible to introduce a stop-band in the normal pass-band of a rectangular waveguide having a cut-off frequency of 2.5GHz. The stop-band of the metamaterial-inclusion-loaded waveguide can be placed at 2.56 GHz, 2.95 GHz, and 4.84 GHz with respective bandwidths 0.04 GHz, 0.04 GHz, and 0.02 GHz, which has been realized with SRR outer radius of 5.6 mm, 5.0 mm, and 3.4 mm, respectively. Similarly, it may be seen from Figure 5.23b that in the stop-band of a rectangular waveguide a pass-band may be introduced by loading the waveguide with SRR metamateral inclusion. The pass-band has been introduced at 9.38 GHz, 2.95 GHz, and 1.95 GHz with respective bandwidths 0.25 GHz, 0.78 GHz, and 2.16 GHz using SRR radius of 2.1 mm, 5 mm, and 7 mm, respectively, for the rectangular waveguide having a cut-off frequency of 10 GHz.

Thus, the bandwidth and location of the pass-band and the stop-band caused by loading of the SRR matrices in a rectangular waveguide is controllable by changing the geometrical dimensions of the SRR elements. Such SRR-loaded waveguide filter which can provide sharp stop-band in the propagation band of normal waveguide may be used as notch-type band-reject filter or BSF at microwave frequency for communication applications.

5.6.3 CRLH-Type Band-Pass Filter

5.6.3.1 Design of Unit Cell of CRLH

RF/microwave passive components like filter, phase shifter, coupler etc., and antennas can be designed with CRLH using multiple unit cells of CRLH to realize improved performance characteristics and size reduction of the component. For all these applications, the design methodology of unit cell

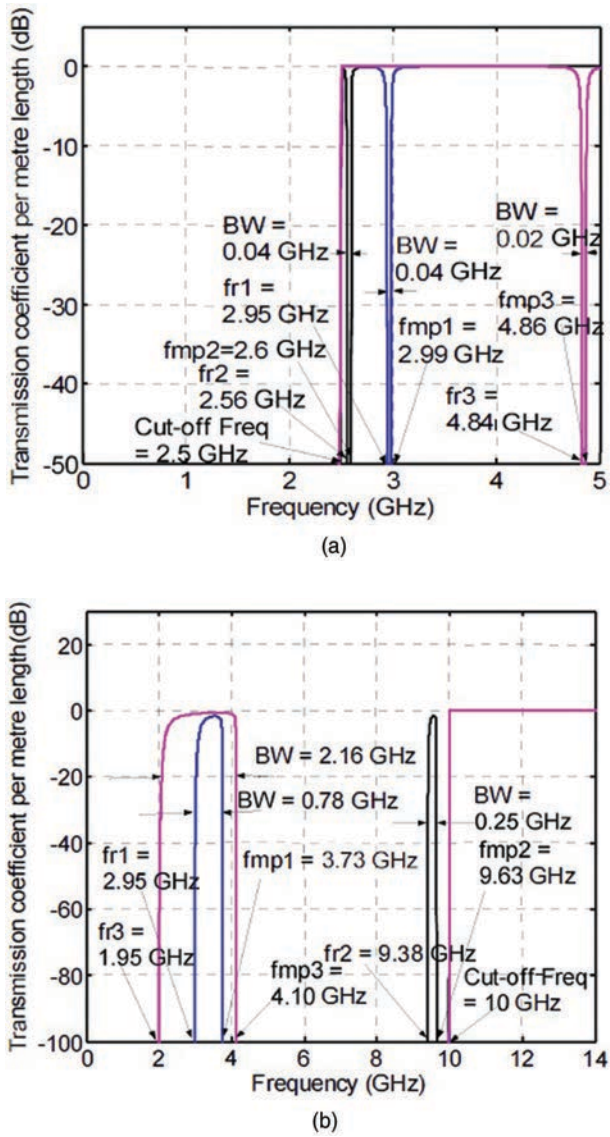


Figure 5.23 (a) Transmission coefficient vs frequency with LHM stop-band, and (b) Transmission coefficient vs frequency with LHM pass-band.

of CRLH using interdigital capacitor and stub inductor is to be understood first. The methodology is discussed in the following steps [8].

Step 1: At first, we select an appropriate transition frequency, ω_0, which normally represent the centre of the operational bandwidth given by

$$\omega_0 = \sqrt{\omega_{se}\omega_{sh}} = \frac{1}{\sqrt[4]{L_R C_R L_L C_L}} = \sqrt{\omega_R \omega_L} \tag{5.34a}$$

where $\omega_R = \sqrt{L_R C_R}$ and $\omega_L = \sqrt{L_L C_L}$

Step 2: Matching condition is applied to the ports of the CRLH structure making:

$$Z_c = Z_R = \sqrt{\frac{L_R}{C_R}} \tag{5.34b}$$

$$\text{and } Z_c = Z_L = \sqrt{\frac{L_L}{C_L}} \tag{5.34c}$$

where Z_c is the characteristic impedance of the CRLH transmission line network.

It may be observed that we have so far three equations, such as Equations (5.34a), (5.34b), and (5.34c) to determine four unknowns L_R, C_R, L_L, C_L.

Step 3: The fourth equation required to determine all the four unknowns mentioned above may be the bandwidth (a designer's constraint) which extends from high-pass LH cut-off ω_{cL} to low-pass RH cut-off ω_{cR} and in the balanced case is given by

$$\omega_{cL} = \omega_R \left| 1 - \sqrt{1 + \frac{\omega_L}{\omega_R}} \right| \text{ and } \omega_{cR} = \omega_R \left(1 + \sqrt{1 + \frac{\omega_L}{\omega_R}} \right) \tag{5.34di}$$

It may be noted that the cut-offs do not depend on the number of cells but only on the value of LC parameters. The fractional bandwidth may be written as

$$FBW = 2\frac{(\omega_{cR} - \omega_{cL})}{(\omega_{cL} + \omega_{cR})} \tag{5.34 dii}$$

An appropriate number of cells N need to be specified to realize a desired performance from the component to be designed.

Step 4: The network can now be implemented by practically realizing the inductances L_R, L_L, and capacitances C_R, C_L with a suitable technology. For instance, in microstrip technology, inductors can be implemented in the

form of spiral inductor strips or simply stub strips; while capacitors may be implemented by metal-insulator-metal (MIM) structures or with interdigital structure.

Step 5: The characteristic parameters of the effectively homogeneous CRLH transmission line network is finally determined by the relations:

$$L'_R = \frac{L_R}{p}, \quad C'_R = \frac{C_R}{p}, \quad L'_L = L_L.p, \quad C'_L = C_L.p \tag{5.34e}$$

where p represents the physical length of the unit cell (this is required as L_R, C_R, L_L, C_L were defined in terms of the quantities designated in the incremental model of CRLH as shown in Figure 5.24a together with its physically realizable planer, i.e., microstrip (μ-strip) LH structure with interdigital capacitor and stub inductor in Figure 5.24b.

A typical microstrip-based CRLH transmission line of nine unit-cells is shown in Figure 5.24c in which each unit cell consists of *interdigital capacitors* and *stub inductors* shorted to ground plane by *via* connection.

The unit cells may be of asymmetric or symmetric type as shown in Figure 5.25; where W_{IDC}, l_{IDC}, g_{IDC}, l_{stub}, and W_{stub}, are the finger width of interdigital lines, finger length of interdigital lines, gap between the fingers, length of the grounded stub, and the width of the grounded stub, respectively.

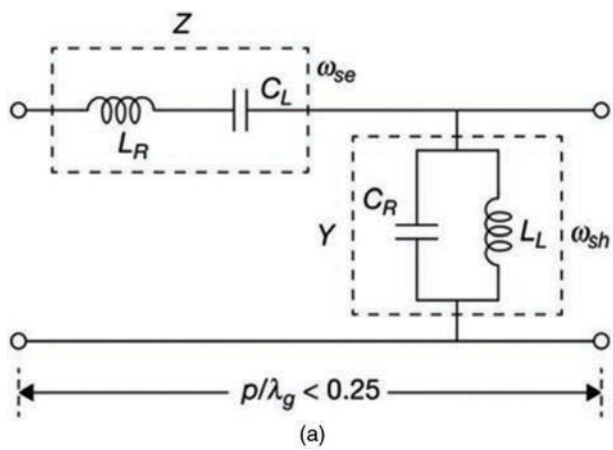


Figure 5.24 (a) Incremental model of CRLH, (b) Physically realizable planer, i.e., microstrip (μ-strip) LH structure with interdigital capacitor and stub inductor, and (c) A microstrip-based 9 unit cell CRLH transmission line using interdigital capacitors and shorted stub inductors.

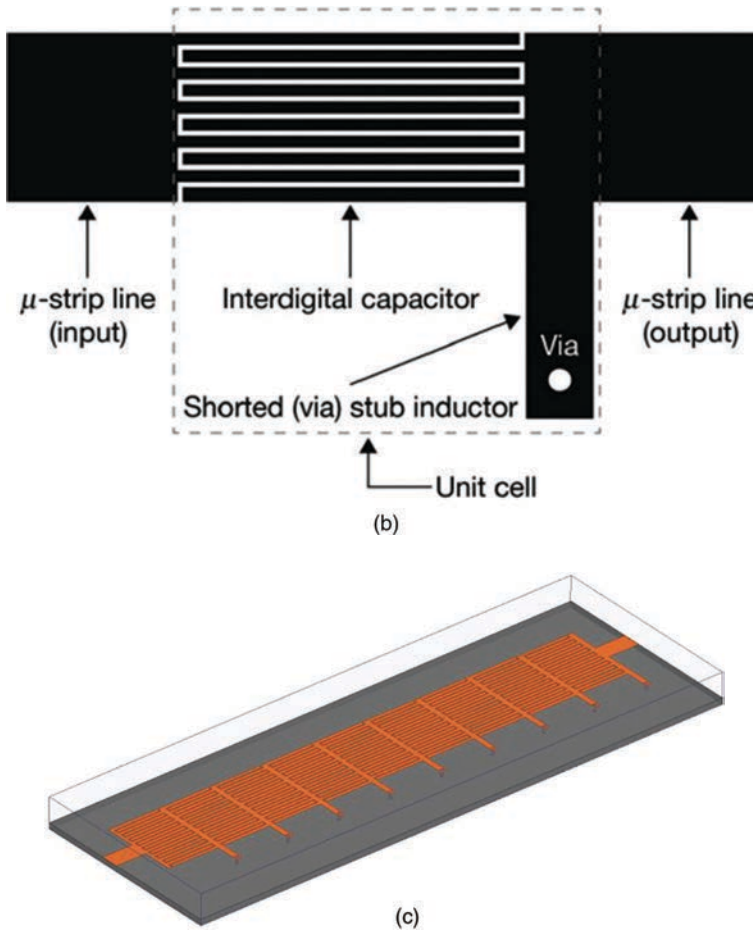


Figure 5.24 (Continued)

The equivalent circuit of a symmetric type unit cell of CRLH transmission line is shown in Figure 5.26. This is a T-network having two impedance branches with capacitance $2C_L$ and inductance $L_R/2$ and an admittance branch with inductance L_L and capacitance C_R. The contributions L_L and C_L are due to the stub inductor and interdigital capacitor and L_R and C_R arises due to parasitic reactance caused by the magnetic flux generated by the currents flowing along the fingers of the interdigital capacitor and the voltage gradient that exist between the fingers of the said capacitor with the ground plane. The parasitic reactance has an increasing effect with increasing frequency.

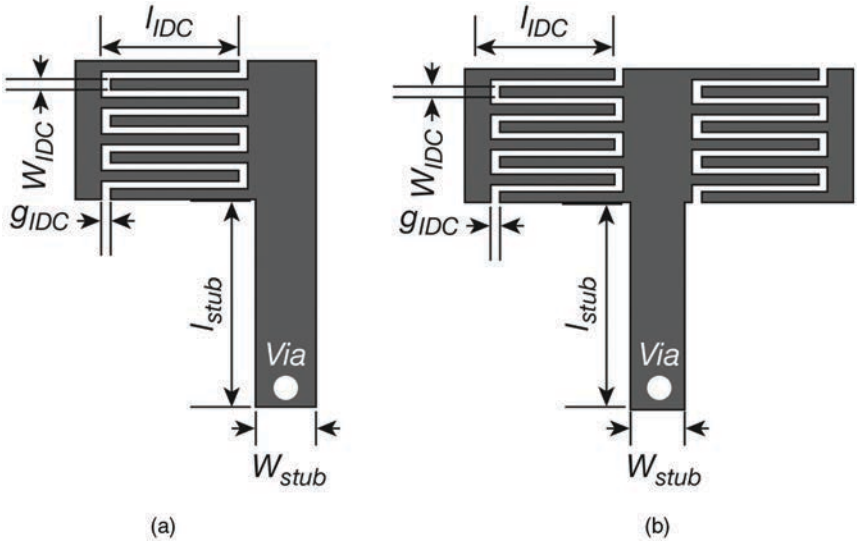


Figure 5.25 Unit cells of CRLH transmission line: (a) Asymmetric type, and (b) Symmetric type.

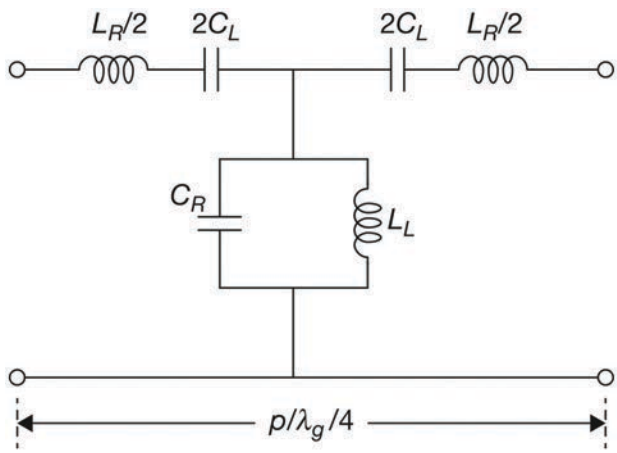


Figure 5.26 T-equivalent circuit of symmetric type unit cell of CRLH transmission line.

Approximate formulas for calculating the inductance L_L and capacitance C_L may be obtained as follows. The shorted stub can be realized with a shorted transmission line section whose input impedance is an inductive reactance [8] when

$$L_L \approx \frac{R_0}{\omega}\tan(\beta l) \tag{5.34f}$$

where R_0 is the characteristic impedance of the host transmission line, β the phase constant, and l is the length of the stub.

The capacitance of the interdigital capacitor may be calculated from [2]

$$C = (\varepsilon_r + 1)l[(N-3)A_1 + A_2]\text{pF} \tag{5.34g}$$

$$\text{With, } A_1 = 4.41\tanh\left\{0.55\left(\frac{h}{w}\right)^{0.45}\right\}10^{-6}\ \text{pF}/\mu\text{m} \tag{5.34h}$$

$$A_2 = 9.92\tanh\left\{0.52\left(\frac{h}{w}\right)^{0.5}\right\}10^{-6}\ \text{pF}/\mu\text{m} \tag{5.34i}$$

where l, w, and h represent the finger length (usually $\leq \lambda/4$), finger width, and finger height of the substrate, respectively, and N is the number of fingers of the interdigital capacitor.

The approximate formulas to calculate L_L and C_L is just the starting guess for the design, however, for optimization full-wave simulation technique is used which are discussed in advanced texts [8].

In the design of CRLH-based RF/microwave passive components and antennas extraction of the equivalent circuit parameters L_R, C_R, L_L, and C_L of CRLH is essential; the details of this are given in advanced text [8] and has not been repeated here for brevity of presentation.

5.6.3.2 Microwave Filter Design with CRLH Transmission Line

Here, three CRLH resonators are used in the design of microwave BPF in which the first and third resonators are of the asymmetric type (see Figure 5.25a) while the second one is symmetric type CRLH resonator (see Figure 5.25b).

The target specification of the filter to be designed [11] is given in Table 5.5.

Table 5.5 Target specification of the filter to be designed

Parameter	*Specification*
Centre Frequency	2.45 GHz
Bandwidth	100 MHz
Order	3
Insertion loss	< 1.0 dB
Pass-band ripple	< 0.1dB
Return loss	< 10 dB
Substrate	Roger RT/Duroid 5880
Dielectric Constant	2.2
Height of the substrate	0.508 mm
Loss tangent	0.0009

Table 5.6 Final dimensions of CRLH-based BPF at 2.45GHz

Items	*Value (mm)*	
	Resonators 1 and 3	*Resonator 2*
W_{IDC}	0.466	0.2
l_{IDC}	5.6	3
Number of fingers	2	2
Gap between the fingers (g_{IDC})	0.18	0.36
W_{stub}	1	0.95
l_{stub}	8.8	12.67
Width of the port coupling	0.3	
Length of the port coupling	4.74	

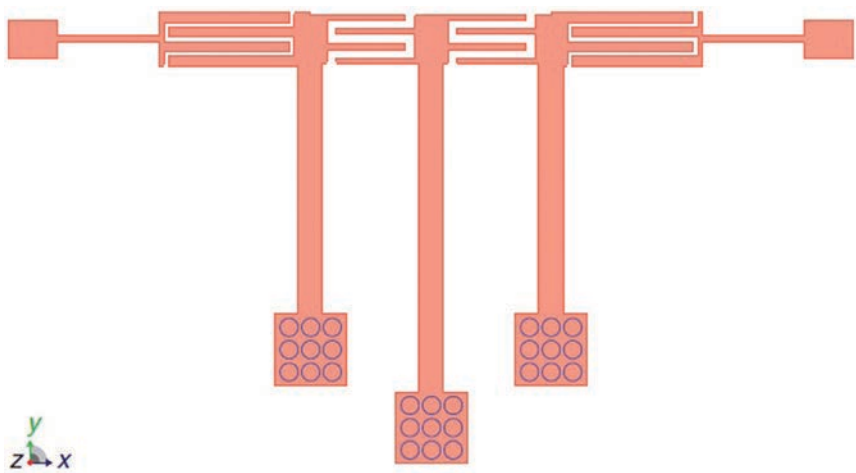


Figure 5.27 Final layout of CRLH TL BPF at 2.45GHz.

The resonators have been optimized separately at 2.45 GHz in EDA tool. At first, the design of resonators need to be simulated in circuit simulator of the EDA tool (using approximate formulas for calculating the initial values of the inductance and capacitance of the interdigital capacitor and stub inductor as detailed under the heading "design of unit cell of CRLH" in section 5.6.3.1) and then need to be optimized using the suitable optimization technique available in the EDA simulation tool. The optimized schematic is to be converted next to layout and final optimization is done using electromagnetic (EM) simulation tool of the simulator. Doing all these steps one can get the final dimensions of the CRLH-based BPF at 2.45 GHz (see Table 5.6) while the final layout looks like the one given in Figure 5.27.

A typical CRLH transmission line BPF and a conventional (parallel coupled/edge-coupled) microstrip BPF are given in Figure 5.28a with their performance characteristics in Figure 5.28b(i) and Figure 5.28b(ii). The performance and size comparison given in Table 5.7 shows the superiority of CRLH-based filter both in terms of miniaturization and improved performance.

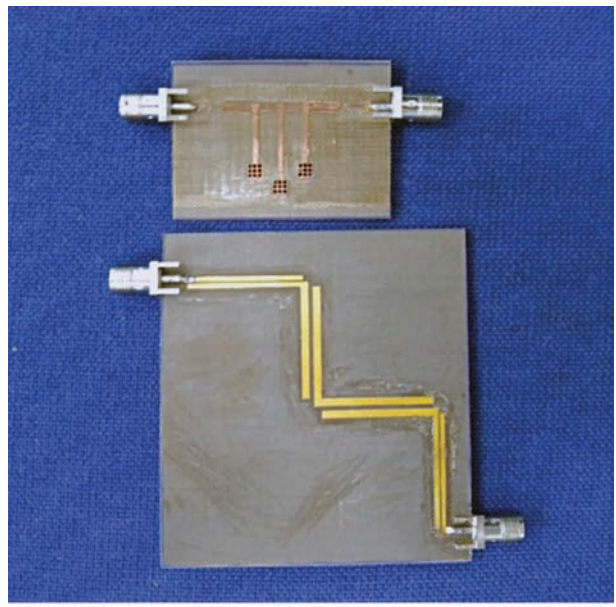


Figure 5.28 (a) Fabricated CRLH-based BPF (top) and conventional (edge-coupled/parallel coupled) microstrip BPF (bottom), and (b) Comparisons of insertion loss (S_{21}) and return loss (S_{11}) of CRLH BPF with conventional BPF: (i) Simulation results, and (ii) Measurement results.

Table 5.7 Performance characteristics and size comparison of CRLH based and conventional (parallel coupled/edge-coupled) microstrip BPF

Performance characteristics/Size	*CRLH BPF*	*Conventional BPF*
Insertion loss	1.6 dB	2.3 dB
Return loss	15 dB	18dB
Harmonics suppression	At least up to 10 GHz	No harmonic suppression
Size*	4.3 cm × 3 cm	6.5 cm × 6 cm

* CRLH-type BPF shows 67% size reduction (area wise) compared to the conventional one.

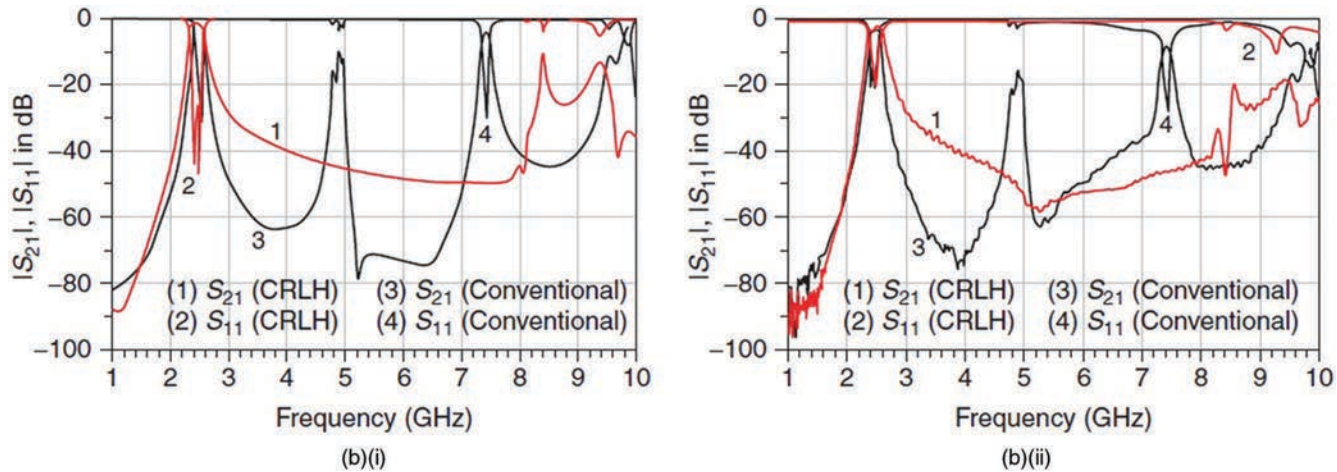


Figure 5.28 (Continued)

REFERENCES

[1] Metthai G, Young L, and Jones E M T, *Microwave Filters, Impedance Matching Networks and Coupling Structures*, Artech House, Norwood, MA, 1980.

[2] Hong J G, and Lancaster M J, *Microstrip Filters for RF/Microwave Applications*, Wiley, New York, 2001.

[3] Altman J L, *Microwave Circuits*, D Van Nostrand Company Inc., Princeton, NJ, 1964.

[4] Collin R E, *Foundations for Microwave Engineering*, Second Edition, John Wiley & Sons, New Jersey, 2007.

[5] Pozar D M, *Microwave Engineering*, Fourth Edition, Wiley, New York, 2012.

[6] Edward T and Steer M B, *Foundations for Microstrip Circuit Design*, Fourth Edition, Wiley-IEEE Press, West Sussex, UK, 2016.

[7] Akhtarzad S, Rowbotham T R, and Jones P B, "The design of coupled microstrip lines", *IEEE Transactions*, *MTT, Vol 23*, No. 6, pp. 486–492, 1975.

[8] Caloz C, and Itoh T, *Electronic Metamaterials: Transmission line Theory and Microwave Applications*, Wiley Interscience, Wiley, New Jersey, 2006.

[9] Kar S, *Metamaterials and Metasurfaces—Basics and Trends*, IOP Publishing, Bristol, 2023.

[10] Kar S, Sikdar S S, and Saha T P, "Metamaterial-loaded waveguide characterization – Some basic studies", *IEEE/OSA/IAPR International Conference on Informatics, Electronics & Vision (ICIEV)*, pp. 1202–1205, 2012.

[11] Kar S, Kumar A, Majumder A, Ghosh S K, Saha S, Sikdar S S, and Saha T K, "CRLH and SRR based microwave filter design useful for microwave communication", *International Journal of Computer, Electrical, Automation, Control and Information Engineering,* Vol. 10, No. 4, pp. 629–633, 2016.

Chapter 6

Design of RF/Microwave Oscillators and Amplifiers

Let good thoughts get amplified in our heart for delivering good service to mankind

6.1 INTRODUCTION

Active devices like oscillators, amplifiers, power combiners, and mixers are in the heart of any communication system (transmitter-receiver) design. Oscillators are known to generate the signal at desired frequency of operation while amplifiers are used to amplify the signal amplitude. Power amplifiers are normally used as output stage of the transmitter for boosting the power of the transmitted signal while low-noise amplifiers are used at the front end of microwave receivers. Power combiners are used to realize large amount of power when a single diode is not capable to generate that desired amount of power; there can be circuit-level or device-level power combining. In fact, with the increase of frequency the signal power available from a single diode is low (following the law: $P_{\text{Availasble}}$ is proportional to the inverse of the frequency of operation) and hence power combiners are essential in microwave system design, especially for the transmitter system. Mixers are used to down/up convert the RF/microwave signal. At the receiver of an RF/microwave system, the mixer down converts the frequency of the incoming signal (RF/microwave) to a desired intermediate frequency (IF) when it is convenient to amplify the received signal in the receiver with multi-stage amplifiers operating at the IF frequency thereby ensuring the stability, reliability, dynamic range, and cost-effectiveness of the system being deigned. In this chapter, we would discuss the fundamentals of RF/Microwave oscillator and amplifier along with practical design examples while the fundamentals of power combiner and mixer along with design examples will be discussed in the next chapter.

6.1.1 Oscillator/Amplifier with Two-Terminal Microwave Device (Diode)

Diodes (two terminal devices) like tunnel diode, Gunn diode, and IMPATT diode (all of which are negative resistance device) are commonly used for the design of oscillators and amplifiers at RF/microwave frequency. Tunnel

DOI: 10.1201/9781003729969-6

diodes are used for the design of low-gain amplifiers while Gunn diode is used for the design of low- to medium-power oscillators and IMPATT diodes are used for the design of medium and high-power microwave oscillators (both CW and pulsed type). IMPATT diode is also used in the design of injection-locked amplifiers. In this section, we would first discuss fundamental theoretical background of such oscillators and amplifiers followed by different oscillator/amplifier configurations used for IMPATT/Gunn diode-based design and finally design examples will be given to illustrate the methodology for the design of IMPATT oscillator and also varactor-tuned IMPATT oscillator with 3D electromagnetic field simulator.

6.1.1.1 Fundamental Theoretical Background of RF/Microwave Diode Oscillator/Amplifier

Let us understand this with respect to IMPATT diode (though the general principle is true for any negative resistance device). The simplest equivalent circuit of an IMPATT diode has a negative deice resistance (R_D) in series with a capacitive reactance (X_D). The circuit part of the oscillator/amplifier represented by Z_C includes the impedance matching circuit (that matches the device having impedance Z_D with the load R_L), and the diode package parasitic is represented by a positive resistance R_C in series with an inductive reactance X_C. In other words, the driving point impedance of the device is given by: $Z_D = (-R_D - jX_D)$; where magnitude of $R_D > 0$ and Z_D is a function of frequency (f), DC current (I_{DC}), RF current (I_{RF}), and temperature (T). The circuit impedance looking into the device terminals is given by $Z_C = (R_C + jX_C)$; where Z_C is a function of frequency (f) only. A generalized oscillator/amplifier scheme is shown in Figure 6.1a, and the condition for stable oscillation is

$$|R_D| = |R_C| \quad \text{and} \quad |X_D| = |X_C| \tag{6.1}$$

The stability of the oscillator is ensured by the net zero resistance of the circuit-diode combination, and the frequency of oscillation is determined by the resonant frequency of the system consisting of device and the circuit; the total reactance of which is zero at resonance. In general, the two conditions in Equation 6.1 will not be satisfied simultaneously in all conditions, since the negative resistance of the diode will not in general be equal to the circuit resistance at the resonant frequency. However, stable oscillation is realized in the following way. The oscillation is started, as usual by a random noise fluctuation caused by the switching transient, which will grow if the total resistance of the device-circuit combination system is negative (for positive resistance there will be decay of the signal). This is possible when with proper circuit design the magnitude of R_D is made to be larger than R_C.

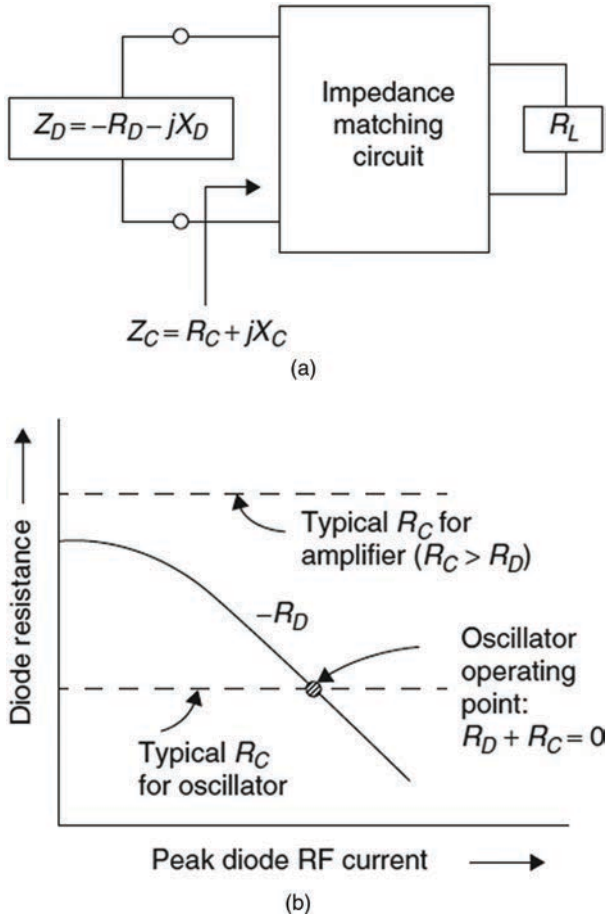


Figure 6.1 (a) Generalized oscillator/amplifier scheme, and (b) Variation of the magnitude of negative resistance of diode with RF signal amplitude.

Now, the inherent property of IMPATT diode (or any negative resistance device) is that the magnitude of the negative resistance of the diode decrease as the RF signal amplitude generated in the device increases (Figure 6.1b). Thus, with device-circuit interaction the oscillation continues to grow for several cycles until the magnitude of the diode negative resistance decreases to the value of the positive resistance of the circuit, when stable oscillation is achieved. Thus at oscillation frequency f_0 [1,2],

$$R_C(f_0) = R_D(f_0, I_0, I_{RF}, T) \tag{6.2a}$$

$$X_C(f_0) = -X_D(f_0, I_0, I_{RF}, T) \tag{6.2b}$$

The output power of the oscillator is governed by Equation (6.2a) while the oscillation frequency is determined from Equation (6.2b).

The *reflection type amplifier* is a common configuration for signal amplification with two terminal negative resistance devices; see Figure 6.2a. A circulator is used to separate the input signal P_i from the amplified output power P_o. The power gain (G_P) of the amplifier is given by

$$G_p = \frac{P_o}{P_i} = |\Gamma|^2 = \left|\frac{Z_D - Z_C}{Z_D + Z_C}\right|^2 = \frac{(R_C + R_D)^2}{(R_C - R_D)^2}, \ \left(\text{assuming } X_D = X_C\right) \tag{6.2c}$$

For stable amplification, choice should be $R_C > R_D$ and hence gain can be controlled by proper choice of R_C; however, when $R_C = R_D$ the gain becomes infinite and the device oscillates.

In a number of applications, we need highly stabilized low-noise oscillators/amplifiers at microwave frequency. This can be done either by *cavity stabilization* or by *injection-locking* of the oscillator.

A scheme for cavity-stabilized Gunn oscillator is shown in Figure 6.2b. Here, the Gunn cavity is coupled to a high Q (> 10^4) cylindrical cavity made of invar (which has a small temperature expansion coefficient), and the output power is extracted from the Gunn cavity. The Q of the stabilizing cavity determines the stability of the cavity-stabilized oscillator. The degree of frequency stability is measured by *stabilization factor S* given by

$$S = \frac{Q_{\text{LS}}}{Q_{\text{LUS}}} \tag{6.2d}$$

Substantial FM noise reduction (>20dB) can be readily achieved using cavity stabilization, however, tunabilty of such oscillator is limited to a very narrow band.

The scheme for injection-locked oscillator/amplifier is shown in Figure 6.2c. Here, an input signal (also known as locking signal which is a highly stable signal) of frequency f_i and having power P_i is injected through a circulator into a free-running oscillator whose frequency is f_0 with power output P_o. By bringing f_i close to f_0, the free-running oscillator can be injection-locked by the externally applied (input) signal at frequency f_i when all the power P_o of the free-running oscillator appears at f_i and assumes the stability of the injected signal over a locking range of $2\Delta f$; given by

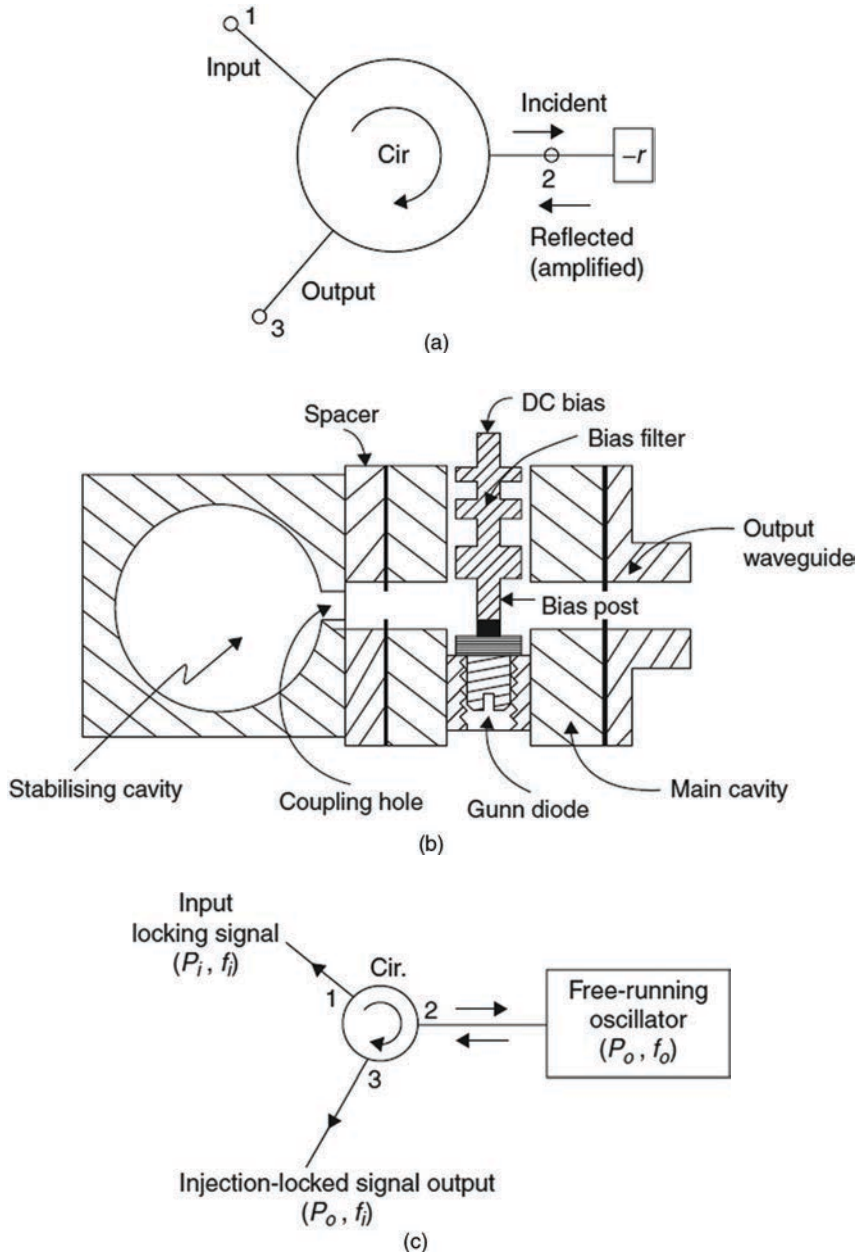


Figure 6.2 (a) Reflection-type amplifier with negative resistance diode, (b) Scheme for cavity-stabilized Gunn oscillator, and (c) Scheme for injection-locked oscillator/amplifier.

$$2\Delta f = \frac{2f_0}{Q_e}\left(\frac{P_o}{P_i}\right)^{-1/2} \tag{6.2e}$$

where Q_e is the external Q of the oscillator circuit.

This phenomenon is helpful to reduce the noise of an oscillator by locking it to an external low-noise source. Its use as an amplifier is also important since within the locking range $2\Delta f$, a small-signal input at f_i manifests as a large output signal at f_i. The gain of such injection-locked amplifier is higher compared to the usual stable-mode amplifiers and is given by

$$G_p = \frac{P_o}{P_i} = \frac{4}{Q_e^2}\left(\frac{f_0}{2\Delta f}\right)^2 \tag{6.2f}$$

Elaborate phase-lock loop (PLL) techniques are also there which can get over the limitations of cavity stabilization and injection-locking techniques. The cavity stabilization is limited in tunability and demand tight dimensional constraints at higher microwave frequencies while injection-locking needs highly stable reference oscillator and requires lower Q_e for large locking bandwidth worsening the FM noise performance of the locked oscillator.

6.1.1.2 IMPATT Oscillator and VCO Design with 3D Electromagnetic Field Simulator

We would discuss here the design approach of an IMPATT oscillator using standard 3D electromagnetic field simulator which is followed these days for professional engineering design of such oscillators. The design approach given here is just a rudimentary one which might help students/designers to get a preliminary flavor of design steps while designing such oscillators. A design approach for electronic tuning of such single diode IMPATT oscillator with varactor diode will also be given which is commonly used as a voltage-controlled oscillator (VCO) at microwave frequency.

Example 6.1 A microwave oscillator has to be designed at 10GHz using X-band (8 – 12) GHz IMPATT diode mounted in three different types of circuit configurations—bias-post in full-height waveguide, bias-post in reduced-height waveguide, and resonant-cap in full-height waveguide.

(i) Design the necessary bias filter and characterize it,
(ii) Design the oscillator on the basis of device-circuit interaction using 3D electromagnetic field simulator and compare the performance

characteristics of three oscillator configurations in terms of device-circuit matching, and

(iii) Design the impedance transformer necessary to match the reduced-height waveguide to the full-height waveguide in case of reduced-height waveguide cavity type oscillator.

Solution:

Bias-filter design and its characterization: The bias-filter (Figure 6.3a) is a low-high-low (stepped-impedance) coaxial-type low-pass filter that is used to allow the DC power supply voltage (from the DC port) to bias the active device (IMPATT diode) and prevents the RF signal generated by the oscillator at the RF port to reach the DC power supply. Thus, the design of this filter should be done on one hand to restrict the maximum frequency of operation to f_{max} = 8 – 12 GHz (X-band), typically say 10GHz for TEM mode (the principal mode) of propagation via the coaxial line, and on the other hand the cut-off frequency of the filter should be less than $0.5f_{max}$. The condition to be met for TEM mode propagation via the coaxial line without allowing higher order modes to propagate can be derived from field analysis as: $(a_2 + b) < 2c/\pi f_{max}$ (Figure 6.3a). With f_{max} = 10 GHz (λ_{max} = 3cm), $(a_2 + b)$ = 1.91cm; and to satisfy the condition mentioned above we may choose $(a_2 + b)$ to be 10% less than the calculated value, i.e., we choose $(a_2 + b)$ = 17.2mm. The insulating material (Teflon with ε_r = 2.5) is of thickness t_i = 0.85mm when we have $(b - a_2) = 2t_i$ = 1.7mm. Thus, we find b = 9.45mm, when a_2 = 7.75mm. The Z_{low} represented by this section of length l_1 (= $\lambda_{max}/4\sqrt{\varepsilon_r}$) = 0.474cm (acting as coaxial line) can be calculated from: $Z_0 = \frac{138}{\sqrt{\varepsilon_r}}\log_{10}\left(\frac{D}{d}\right)$; when we have Z_{low} = 7.505Ω. The high impedance Z_{high} to be represented by the section of the low-pass filter having length l_2 (= $\lambda_{max}/4$) = 0.75cm and having diameter a_1 is determined by choosing $Z_{high} = 10\,Z_{low}$. Thus, with Z_{high} = 75.05Ω, we can calculate a_1 = 2.701mm. With these dimensions of b, a_1, and a_2 we may calculate $C = \frac{2\pi\varepsilon_0\varepsilon_r}{\ln(D/d)}$ for line of length l_1 and $L = \frac{\mu_0}{2\pi}\ln\frac{D}{d}$ for line of length l_2 (the formula for C and L are per unit length of the coaxial line). When this is done, we get: $C = 341.28\times10^{-14}$ F and $L = 187.5\times10^{-11}$H and the cut-off frequency: $f_c = 1/\pi\sqrt{2LC}$ = 2.81GHz. It may be observed from the simulation characteristics of the filter (Figure 6.3b) that

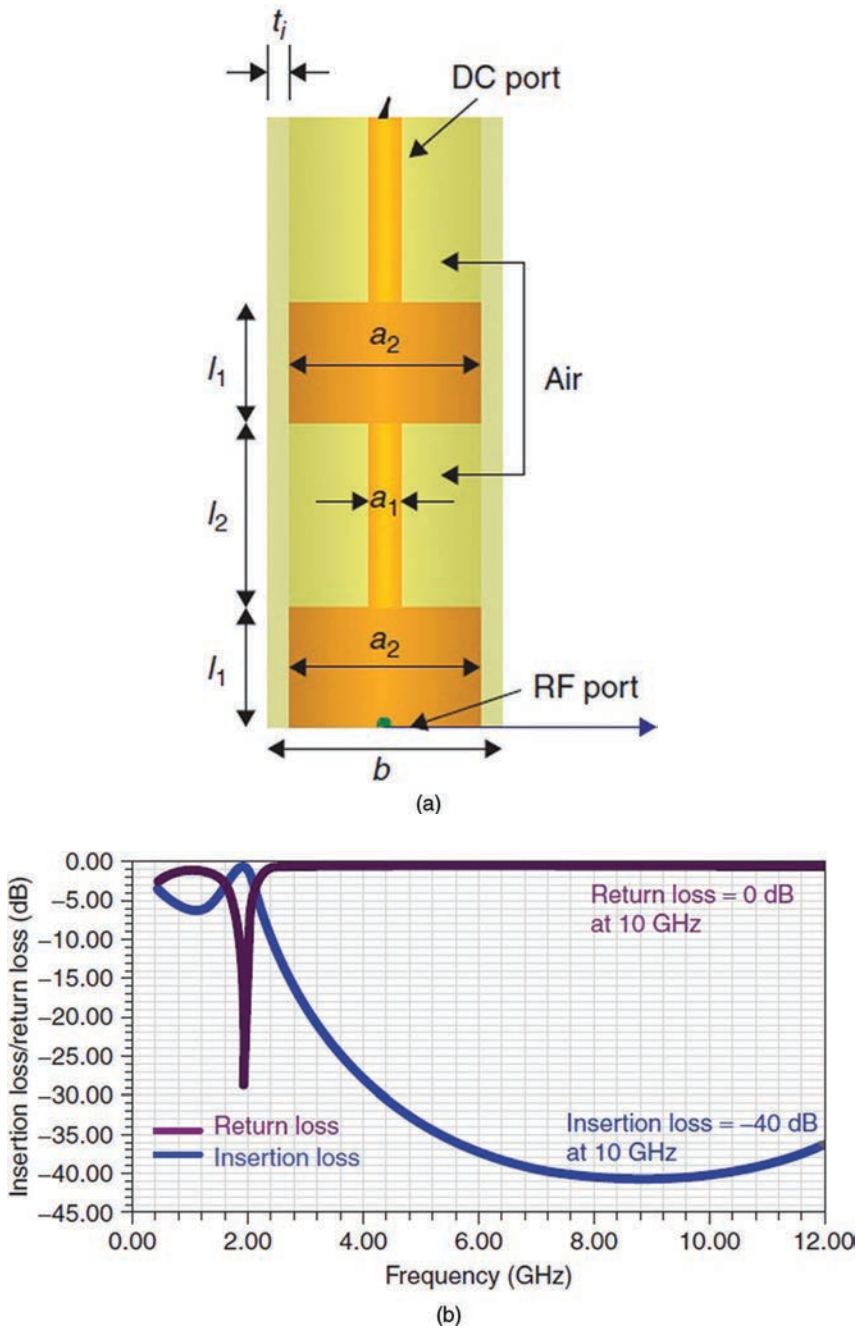


Figure 6.3 Low-hi-low coaxial type low-pass filter: (a) Schematic, and (b) Simulation characteristics.

around 2 GHz there is a dip in the return loss and a peak of insertion loss above which frequency the return loss and insertion loss sharply approach to lower and higher values, respectively, and becomes 0 dB and –40 dB at and around 10 GHz.

(ii) 3D *electromagnetic field simulator-based Oscillator design approach and characterization*: The oscillator system with bias-filter, bias-post in reduced-height waveguide cavity and integral post-cum-resonant-cap cavity in full-height waveguide along with the diode has been shown in Figure 6.4a and 6.4c, respectively (use of only bias-post in full-height waveguide is not able to realize device-circuit impedance match and was thus not used). The IMPATT device equivalent circuit is also shown in Figure 6.4b with

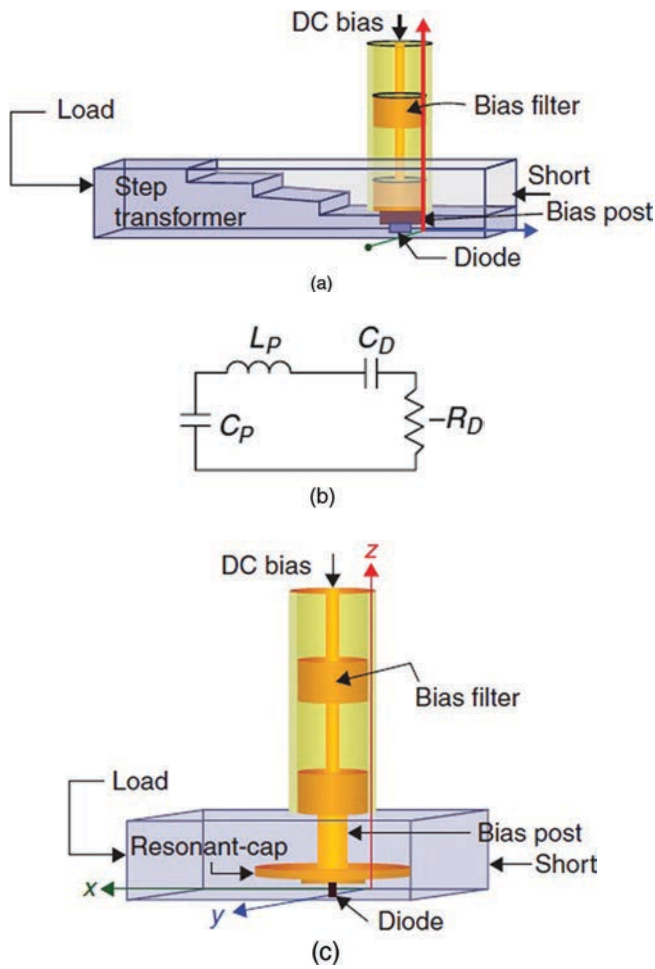


Figure 6.4 (a) Reduced-height waveguide type oscillator, (b) Equivalent circuit of packaged diode, and (c) Resonant-cap type oscillator.

different diode parameters as R_D = –0.7 Ω, C_D = 0.2 pF, L_P = 0.6 nH, and C_P = 0.3 pF. The oscillator system has been designed in 3D electromagnetic field simulator considering device-circuit interaction to get the resonance at the desired freqency in which the effects of bias-post/ resonant-cap cavity, short (which is normally set at $\lambda_g/4$ from the device plane) and the active device has been included. The bias-post diameter in case of post in reduced-height wavegide is chosen to be 6.0 mm, and the resonant-cap diameter is 1.98 cm (with post above it with diameter = 4 mm). The

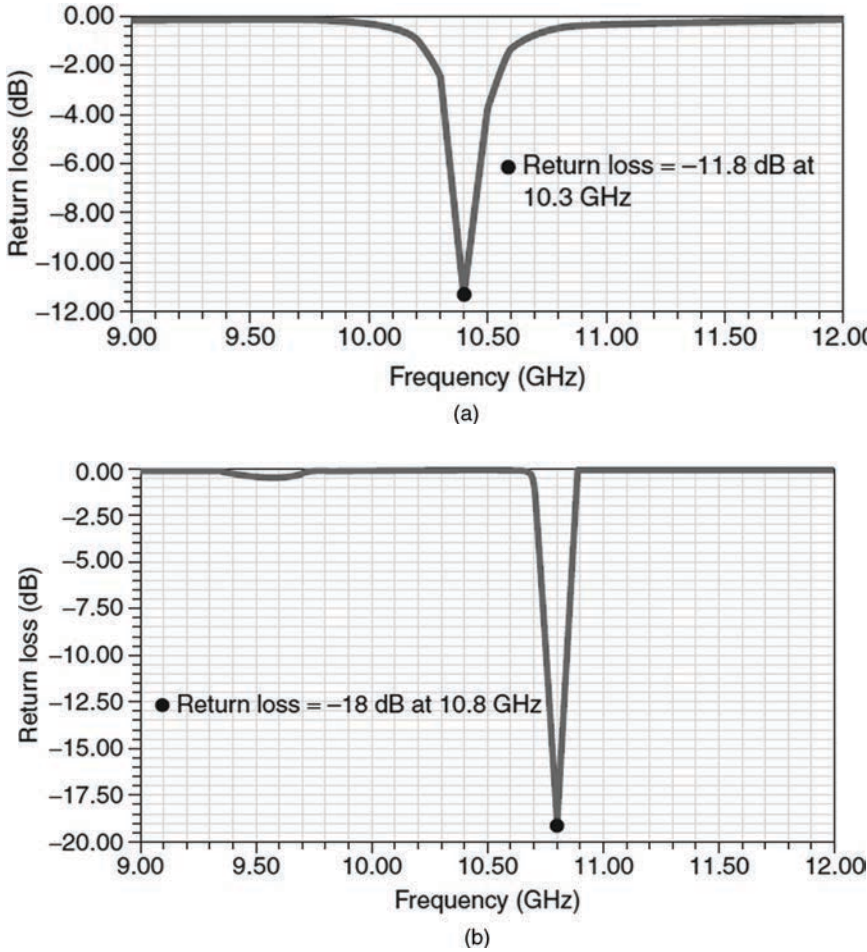
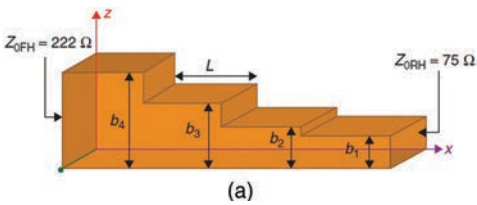


Figure 6.5 Plot of frequency vs return loss of oscillator system: (a) For diode embedded in reduced-height waveguide cavity, and (b) For diode embedded in resonant-cap cavity.

simulation result shown in Figure 6.5 indicates that the resonance is around 10.3 GHz with a return loss of –11.8 dB for reduced-height waveguide-type oscillator and for resonant-cap-type oscillator resoanace freqency and return loss are seen to be 10.8 GHz and –18dB, respectively. It may be observed that the resoant-cap type oscillator provides a better impedance match due to device-circuit interaction. Howevere, both the oscillators can be further optimized for impedance matching and resonance freqency (this is left to the students/designers as an exercise).

(iii) The design of 2-section stepped impedance transformer required in the reduced-height cavity type oscillator (Figure 6.4a) may be done by following the procedure given in Section 4.2.2 or using simulator design. The transformer structure with design data of simulation and the simulation characteristics of the designed transformer are given in Figure 6.6.

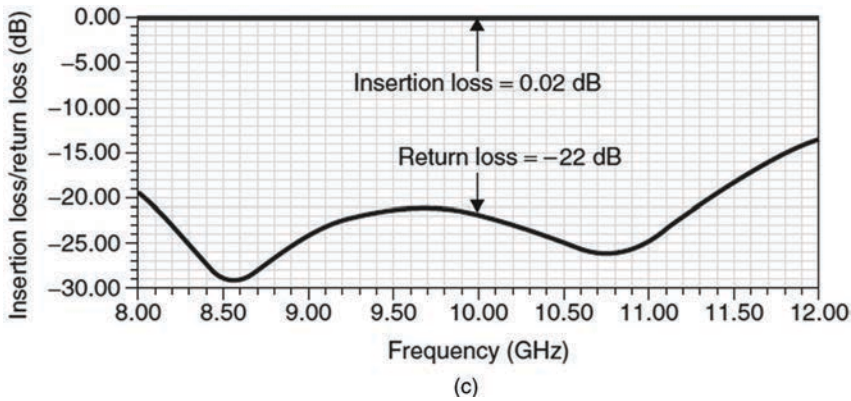
[*NOTE*: Oscillator Configurations with IMPATT/Gunn Diodes have been discussed in Appendix III.]



(a)

Imdepance Z (Ω)	Step height (mm)
Z_{ORH} = 75	3.43 (b_4)
Z_2 = 86.77	3.97 (b_3)
Z_3 = 134.35	6.14 (b_2)
Z_{OFH} = 222	10.16 (b_1)

(b)



(c)

Figure 6.6 Two-section stepped impedance transformer: (a) 3D schematic, (b) Simulation design data, and (c) Simulation chracteristics of the transformer.

Example 6.2 Design a varactor-tuned IMPATT oscillator (i.e., a voltage-controlled oscillator, VCO) at X-band and show its electronic tuning characteristics with the following data.

The packaged IMPATT diode parameters are: R_D = –0.7Ω, C_D = 0.2pF, L_P = 0.6nH, and C_P = 0.3pF; while the varactor diode parameters are: L_p = 0.3nH, C_p = 0.15pF, R_S = 1.2Ω, $C_j(V_{rb})$ = 3pF @ 0V and $C_j(V_{rb})$ = 0.2pF @ 20V. The bias voltage for varactor diode can be changed from 0 volt to 20 volt.

Solution:

The varactor-tuned oscillator using IMPATT/Gunn diode may be of two types, such as, parallel-tuned (staggered) configuration and series-tuned (antipodal) configuration (Figure 6.7). In staggered configuration, the distance between the two posts is $\lambda_g/2$ and the distance between the left side post and the short-circuited plane is $\lambda_g/4$. But in antipodal configuration, the electrical length between the IMPATT and varactor is zero; thus, wideband VCOs are generally designed with antipodal configuration (as there is no λ dependence in this configuration). Our design in this problem will be with the parallel-tuned configuration.

The oscillation condition for varactor-tuned IMPATT oscillator can be written with reference to Figure 6.8a where block marked as equivalent one port power device is the IMPATT oscillator itself consisting of the bias-posts in waveguide and the IMPATT diode; while the varactor diode is represented by its junction capacitance C_j (V_{rb}) and series resistance R_S.

The oscillation condition for the VCO is thus given by Equation (6.1):

$$R_S + R_D(\omega, I) = 0 \text{ and } -\frac{1}{\omega C_j(V_{rb})} + X_D(\omega, I) = 0 \tag{6.3}$$

To design the VCO in 3D electromagnetic field simulator, we refer to Figure 6.8 (b) where equivalent circuit of dual-post waveguide structure in conjunction with the equivalent packaged device (both IMPATT and varactor) impedances are shown. The device-circuit interaction is optimized in the simulator to realize good impedance matching between the device and the circuit. It may be observed from Figure 6.9a that with zero varactor bias voltage (V_{rb} = 0V) the oscillation frequency of VCO is at 10.09GHz with good device-circuit impedance matching (return loss = –22dB). The varactor bias voltage change causes the change of its capacitance and hence the frequency change of the VCO and with varactor reverse bias set to V_{rb} = –20Volts the oscillation frequency is at 10.50GHz (with return loss = –20 dB);

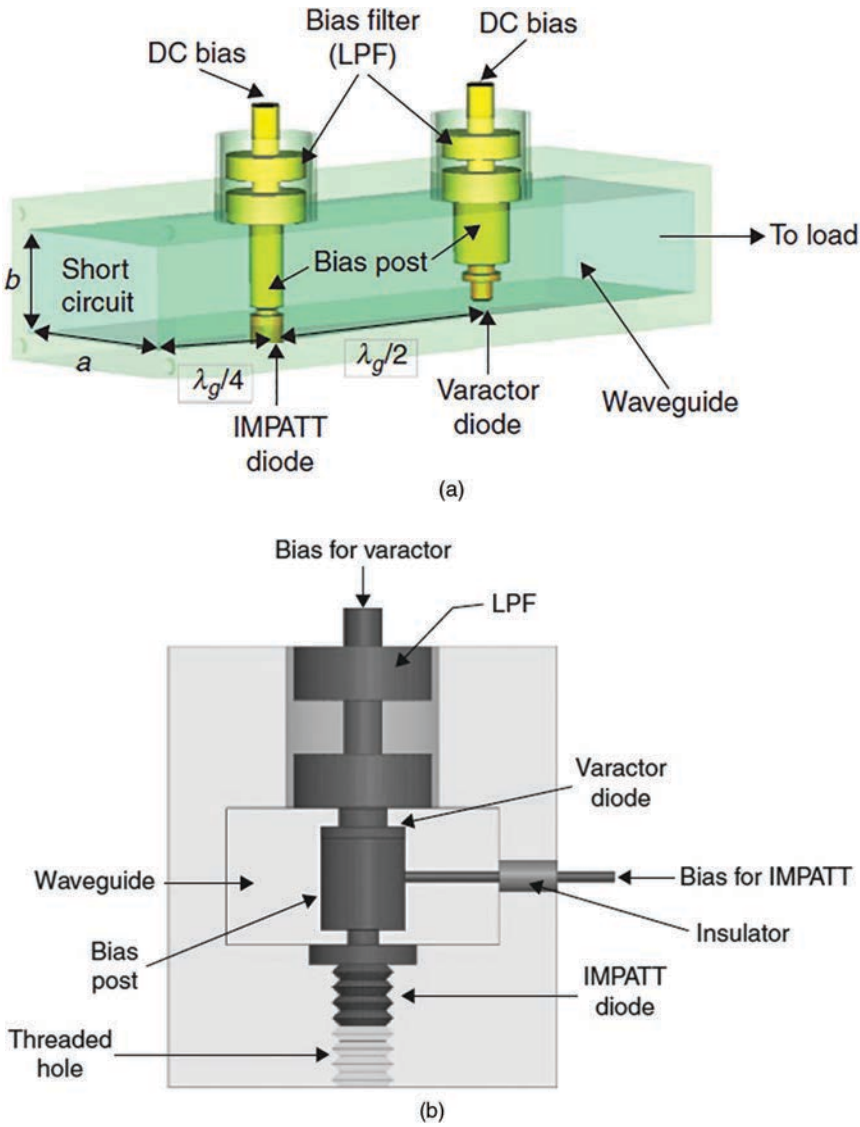


Figure 6.7 Varactor-tuned oscillator using: (a) Staggered configuration, and (b) Antipodal configuration.

see Figure 6.9b. The electronic tuning of the IMPATT oscillator with the change of varactor bias voltage of the VCO is seen to be almost linear with a frequency change of 410MHz for varactor bias variation from 0 to 20 Volts; see Figure 6.9c.

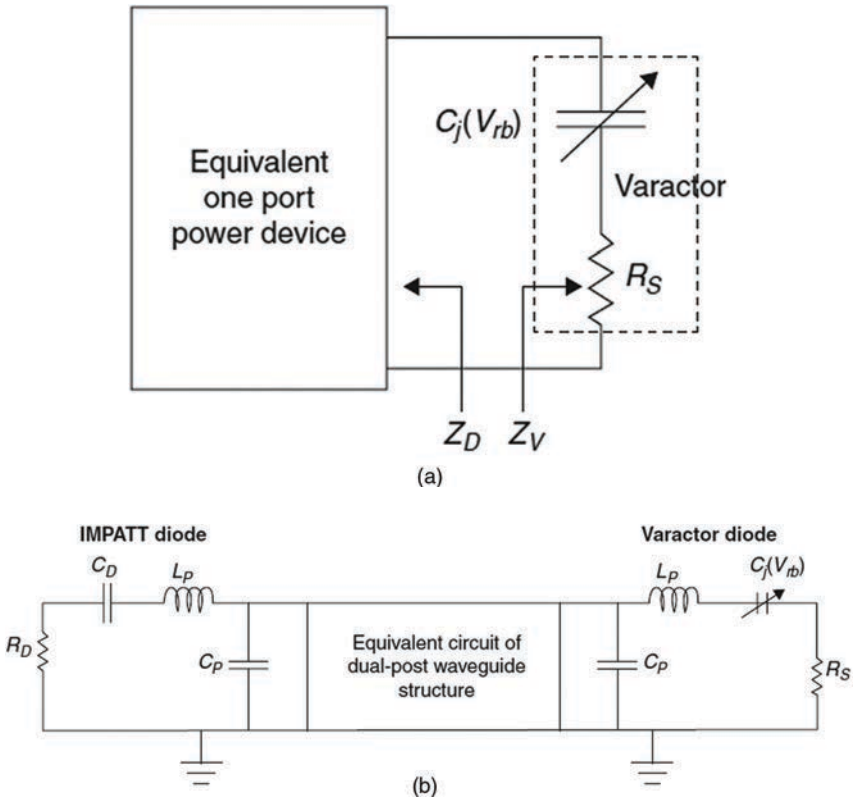


Figure 6.8 (a) Power device with varactor diode for oscillation condition of VCO, and (b) Device-circuit representation for 3D electromagnetic simulation.

6.1.2 Amplifier with Three-Terminal Microwave Device (Transistor)

Microwave transistors (i.e., three terminal devices) like MESFET, HEMT are used for the design of amplifiers at microwave frequency. However, HEMT devices are preferred over MESFET devices for high frequency amplifier design with better noise performance. In this section, we would discuss at first fundamental theoretical background of transistor amplifier at RF/microwave frequency followed by low-noise amplifier (LNA) design and power amplifier (PA) design with HEMT device using EDA simulation tool.

6.1.2.1 Fundamental Theoretical Background of RF/Microwave Transistor-based RF/Microwave Amplifier

With the advent of various electronic design automation (EDA) tools, modern design procedures for transistor amplifiers at RF/microwave

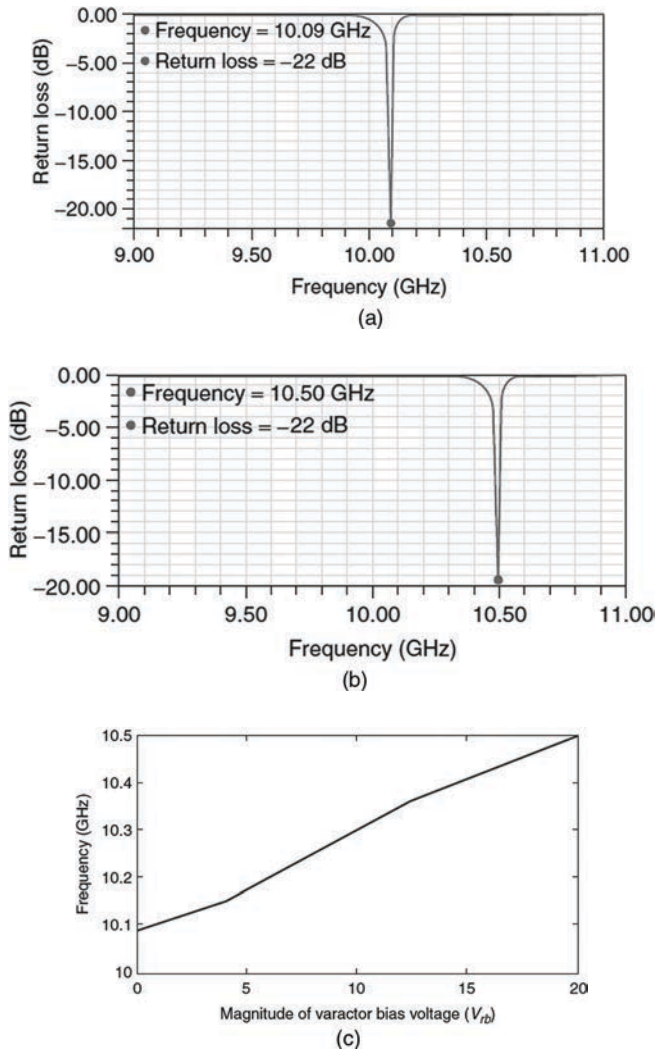


Figure 6.9 Varactor-tuned IMPATT oscillator (VCO) characteristics: (a) With zero varactor bias (V_{rb} = 0V), and (b) With V_{rb} = 20V; while (c) VCO electronic tuning.

frequency are now carried out using such computer-aided design tools; however, understanding of the fundamental theoretical background is essential even when one desire to apply such design tools. The analysis and design of RF/microwave transistor amplifiers are based on the *S*-parameters of the transistor. The scheme for RF/microwave transistor amplifier is given in Figure 6.10.

In the amplifier configuration, the power available from the source is delivered to the input of the transistor through the input matching network.

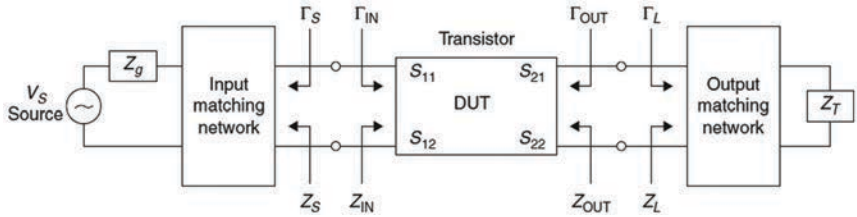


Figure 6.10 Scheme for an RF/microwave transistor amplifier.

This is designed to transform the generator impedance Z_g (usually 50Ω) to the source impedance Z_S (as seen looking from the active device terminals towards the generator). The active device together with the matched generator is expected to deliver maximum power to the load. To realize this, output matching network is used to transform the terminating impedance Z_T (usually 50Ω) to the load impedance Z_L. These matching networks are designed to realize the required stability, gain, noise performance, and bandwidth of the amplifier. Matching networks may be L, Π, or T-type of which L-section networks (see section 4.3) are simple to design and practically very common.

The amplifiers may be designed for *unilateral* case when $|S_{12}| = 0$, (i.e., the transmitted signal from load to the source is neglected) or for *bilateral* case when $|S_{12}| \neq 0$ (i.e., effect of load on source need to be considered).

The input matching network (IMN) and the output matching network (OMN) are used on the one hand to match the source (generator) impedance $Z_g = 50\Omega$ and terminating (load) impedance $Z_T = 50\Omega$, respectively, and on the other hand to match the input and output impedances Z_{IN} and Z_{OUT} of the transistor, device under test (DUT), to Z_S of IMN and Z_L of OMN. For maximum power transfer from IMN to DUT, we need: $\Gamma_S = \Gamma^*_{IN}$ or $Z_S = Z^*_{IN}$ while for maximum power transfer from DUT to OMN we need: $\Gamma_L = \Gamma^*_{OUT}$ or $Z_L = Z^*_{OUT}$; in other words, we need *conjugate matching* for these complex impedances. In case of unilateral amplifier (for eg., LNA), $S_{12} = 0$, i.e., Γ_{IN} is not effected by Γ_{OUT} and vice versa, then input and output matching is done seperately. But for bilateral amplifier (eg., power amplifier, PA), $S_{12} \neq 0$, i.e., when the effect of load on source need to be considered then input and output sections are to be matched simultaneously. It may be noted that for LNA design, we resort to the matching condition at input as: $\Gamma_S = \Gamma_{opt}$ or $Z_S = Z_{opt}$ (instead Γ_{IN} or Z_{IN} we use Γ_{opt} or Z_{opt}, where Γ_{opt} or Z_{opt} represents the optimum reflection coefficient or optimum impedance at the input of the device for minimum realizable noise as supplied by the device manufacturer); the output is however conjugately matched. But for PA design one has to find out Γ_{IN} and Γ_{OUT} or Z_{IN} and Z_{OUT} for maximum power output

(P_o) or maximum power-added efficiency (PAE) and the matching condition at input is $\Gamma_S = \Gamma_{\text{IN}[P_{o(\max)}/\text{PAE}_{(\max)}]}$ or $Z_S = Z_{\text{IN}[P_{o(\max)}/\text{PAE}_{(\max)}]}$, while the output match condition is $\Gamma_L = \Gamma^*_{\text{OUT}[P_{0(\max)}/\text{PAE}_{(\max)}]}$ or $Z_L = Z^*_{\text{OUT}[P_{0(\max)}/\text{PAE}_{(\max)}]}$, a conjugate match; but input and output are simultaneously matched for maximum power output or maximum power-added efficiency.

The transistor oscillator design approach at RF/microwave frequency is similar to that of the amplifiers except that the circuit is purposely designed to be unstable. It will be seen shortly that for amplifier we need $K > 1$ while for oscillator $K < 1$, where K is known as the stability factor. The operating principle of three-terminal (i.e., transistor) oscillator is similar to that for the two-terminal (i.e., diode) oscillator discussed in section 6.1.1.1.

Now we would discuss the formulations and graphs necessary to understand the principle of operation of RF/microwave amplifier and its design based on *S*-parameters of the active device.

6.1.2.1.1 Amplifier Gain

Many *power gain* equations are derived in terms of the *S*-parameters and the reflection coefficient at the input and output of RF/microwave amplifiers. While discussing *S*-parameters in Chapter 2, we discussed how to derive different gain expressions of an RF/microwave amplifier in section 2.6, namely, *transducer power gain*, *available power gain*, and *operating power gain* or simply *power gain*. These are repeated here for convenience.

Transducer Power Gain (G_T) is given by

$$G_T = \frac{1-|\Gamma_S|^2}{|1-\Gamma_{IN}\Gamma_S|^2}|S_{21}|^2\frac{1-|\Gamma_L|^2}{|1-S_{22}\Gamma_L|^2} \tag{6.3a}$$

Or alternatively as

$$G_T = \frac{1-|\Gamma_S|^2}{|1-S_{11}\Gamma_S|^2}|S_{21}|^2\frac{1-|\Gamma_L|^2}{|1-\Gamma_{OUT}\Gamma_L|^2} \tag{6.3b}$$

$$\text{where: } \Gamma_{IN} = S_{11} + \frac{\Gamma_L S_{12}S_{21}}{1-S_{22}\Gamma_L} \text{ and } \Gamma_{OUT} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1-S_{11}\Gamma_S} \tag{6.3c}$$

Some special cases of transducer power gain may also be mentioned herewith.

(i) When both input and output networks are matched to source impedance and load impedance then we have $\Gamma_S = \Gamma_L = 0$; and hence, the gain is *matched transducer power gain* given by

$$G_{TM} = |S_{21}|^2 \tag{6.3d}$$

(ii) If it is assumed that the amplifier is unilateral, i.e., $|S_{12}| = 0$, then $\Gamma_{IN} = S_{11}$ and $\Gamma_{OUT} = S_{22}$ and thus *unilateral transducer power gain* is given by

$$G_{TU} = \frac{1-|\Gamma_S|^2}{|1-S_{11}\Gamma_S|^2}|S_{21}|^2\frac{1-|\Gamma_L|^2}{|1-S_{22}\Gamma_L|^2} \tag{6.3e}$$

(iii) One can optimize the input reflection coefficient Γ_S and output reflection coefficient Γ_L making $\Gamma_S = S_{11}^*$ and $\Gamma_L = S_{22}^*$ to realize *maximum unilateral transducer power gain* given by

$$G_{TU_{MAX}} = \frac{|S_{21}|^2}{(1-|S_{11}|^2)(1-|S_{22}|^2)} \tag{6.3f}$$

Available Power Gain (G_{AV}) is given by

$$G_{AV} = \frac{1-|\Gamma_S|^2}{|1-S_{11}\Gamma_S|^2}|S_{21}|^2\frac{1}{1-|\Gamma_{OUT}|^2} \tag{6.3g}$$

Operating Power Gain or simply *Power Gain* (G_P) is given by

$$G_P = \frac{1}{1-|\Gamma_{IN}|^2}|S_{21}|^2\frac{1-|\Gamma_L|^2}{|1-S_{22}\Gamma_L|^2} \tag{6.3h}$$

However, it may be mentioned that, practically the gain of an amplifier we talk about is the *operating power gain* or simply the *power gain*.

6.1.2.1.2 Stability Considerations for Amplifier

The design of an amplifier should ensure its *stability* in the sense that it does not oscillate either unconditionally or conditionally. The amplifier stability depends on the *S*-parameters of the active device (transistor), the matching networks, and the terminations. The two-port amplifier network

(see Figure 6.10) is said to be unconditionally stable at a given frequency if the real part of the input and output impedances of the active device are positive (i.e., *Re* [Z_{IN}] > 0 and *Re* [Z_{OUT}] > 0) for all passive load and source impedances. If the two-port network is not unconditionally stable, it is potentially unstable or conditionally stable when the real part of input and output impedances are negative (i.e., *Re* [Z_{IN}] < 0 and *Re* [Z_{OUT}] < 0) for some passive load and source impedances at a specific frequency.

Thus, for unconditionally stable case at a given frequency, we can write in terms of reflection coefficient:

$$\left|\Gamma_S\right| \le 1; \left|\Gamma_L\right| \le 1 \tag{6.4a}$$

$$\left|\Gamma_{IN}\right| = \left|S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}\right| < 1 \tag{6.4b}$$

$$\left|\Gamma_{OUT}\right| = \left|S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S}\right| < 1 \tag{6.4c}$$

While for potentially unstable (or conditionally stable) case, we have in terms of reflection coefficient:

$$\left|\Gamma_{IN}\right| > 1 \text{ or } \left|\Gamma_{OUT}\right| > 1 \tag{6.5}$$

The conditions for amplifier stability can be derived on the basis of graphical analysis of Equation (6.4b) and (6.4c) as follows. When we set $\left|\Gamma_{IN}\right|$ and $\left|\Gamma_{OUT}\right|$ equal to unity, a boundary will be obtained beyond which the amplifier is unstable. It can be shown that solving for Γ_L from the solution of equation $\left|\Gamma_{IN}\right| = 1$ yields circle in the Γ_L plane known as *output stability circle* while solving for Γ_S from the solution of equation $\left|\Gamma_{OUT}\right| = 1$ yields circle in the Γ_S plane known as *input stability circle* [3]. The equations representing these circles are given by

$$\left|\Gamma_L - \frac{(S_{22} - \Delta S_{11}^*)^*}{\left|S_{22}\right|^2 - \left|\Delta\right|^2}\right| = \left|\frac{S_{12}S_{21}}{\left|S_{22}\right|^2 - \left|\Delta\right|^2}\right| \tag{6.6a}$$

[For output stability circle]

$$\left|\Gamma_S - \frac{(S_{11} - \Delta S_{22}^*)^*}{\left|S_{11}\right|^2 - \left|\Delta\right|^2}\right| = \left|\frac{S_{12}S_{21}}{\left|S_{11}\right|^2 - \left|\Delta\right|^2}\right| \tag{6.6b}$$

[For input stability circle]

where $\Delta = (S_{11}S_{22} - S_{12}S_{21})$ (6.6c)

The radii and centres of output stability and input stability circles are given by

With $|\Gamma_{IN}| = 1$, the *output stability circle* in the Γ_L plane has

$$\text{Radius } (r_L) = \left| \frac{S_{12}S_{21}}{|S_{22}|^2 - |\Delta|^2} \right| \tag{6.7a}$$

$$\text{Center } (C_L) = \frac{(S_{22} - \Delta S_{11}^*)^*}{|S_{22}|^2 - |\Delta|^2} = \frac{C_L^*}{|S_{22}|^2 - |\Delta|^2} \tag{6.7b}$$

where $C_L^* = (S_{22} - \Delta S_{11}^*)^*$.

With $|\Gamma_{OUT}| = 1$, the *input stability circle* in the Γ_S plane has

$$\text{Radius } (r_S) = \left| \frac{S_{12}S_{21}}{|S_{11}|^2 - |\Delta|^2} \right| \tag{6.8a}$$

$$\text{Center } (c_S) = \frac{(S_{11} - \Delta S_{22}^*)^*}{|S_{11}|^2 - |\Delta|^2} = \frac{C_S^*}{|S_{11}|^2 - |\Delta^2|} \tag{6.8b}$$

where $C_S^* = (S_{11} - \Delta S_{22}^*)^*$.

Knowing the *S*-parameters of the active device at a specified frequency, the stability circles can be plotted on the Smith chart (using Eq. (6.7) and Eq. (6.8)) to investigate the stable and unstable regions of operation of the amplifier in the Γ_L plane and Γ_S plane.

The stability circles are shown in Figure 6.11 when $|\Gamma_{IN}| = 1$ and $|\Gamma_{OUT}| = 1$; on one side of the stability circle boundary, in the Γ_L plane, we will have $|\Gamma_{IN}| < 1$ and on the other side $|\Gamma_{IN}| > 1$. Similarly, in the Γ_S plane on one side of the stability circle boundary, we will have $|\Gamma_{OUT}| < 1$ and on the other side $|\Gamma_{OUT}| > 1$. It may be noted that whenever $|\Gamma_{IN}|$ or $|\Gamma_{OUT}| < 1$, we have stable operation of the amplifier while for the case $|\Gamma_{IN}|$ or $|\Gamma_{OUT}| > 1$, unstable operation of the amplifier is resulted.

To identify these zones of stable and unstable operation on the Smith chart, we take typical example. With $Z_L = Z_0$, $\Gamma_L = 0$ and hence $|\Gamma_{IN}| = |S_{11}|$ (Eq. (6.3c)). Now with $|S_{11}| < 1$, $|\Gamma_{IN}| < 1$; we have stable region of operation

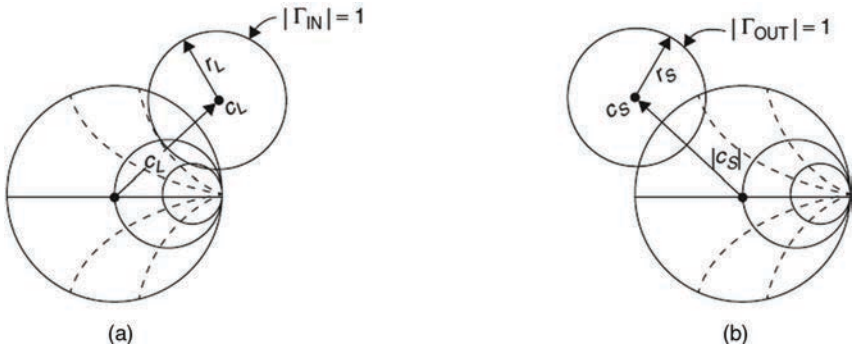


Figure 6.11 Stability circle construction in the Smith chart: (a) Plot in Γ_L plane, and (b) Plot in Γ_S plane.

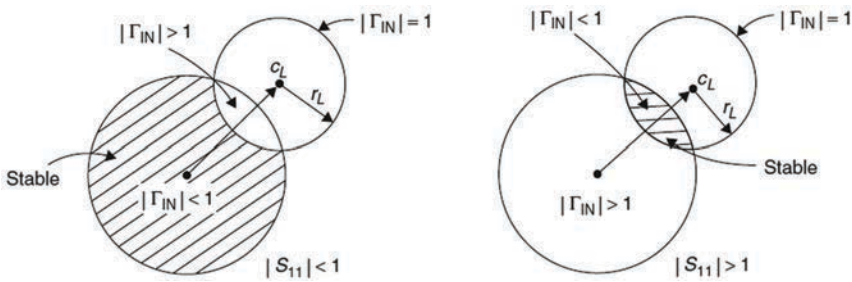


Figure 6.12 Smith chart showing the stable and unstable regions in the Γ_L plane for conditional stability (i.e., for potentially unstable case). [$|\Gamma_{IN}| = |S_{11}|$, for $\Gamma_L = 0$].

while with $|S_{11}| > 1$, $|\Gamma_{IN}| > 1$; we have unstable region of operation (see Figure 6.12) in which the shaded region indicates the stable region of operation). Similar zones of stable and unstable regions of operation on Smith chart in the Γ_S plane are shown in Figure 6.13.

The above discussion relates to the potentially stable (i.e., conditionally stable) case of amplifier from which it is clear that for unconditionally stable case the stability circle should be outside the Smith chart (see Figure 6.14).

The condition for unconditional stability for all passive loads and sources can be expressed as

In Γ_L plane:

$$\left|\left|c_L\right| - r_L\right| > 1 \quad \text{for} \quad |S_{11}| < 1 \tag{6.9a}$$

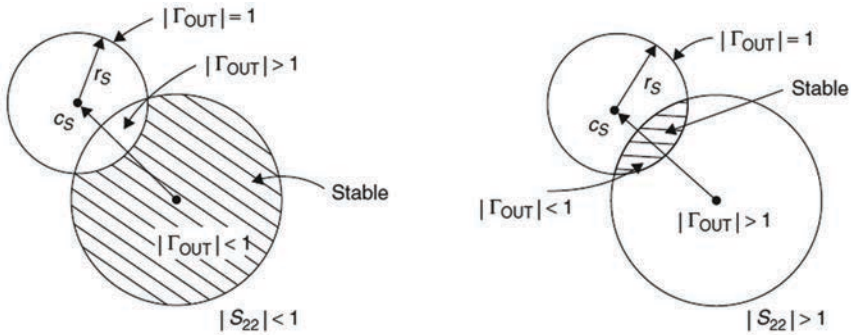


Figure 6.13 Smith chart showing the stable and unstable regions in the Γ_S plane for conditional stability (i.e., for potentially unstable case). [$|\Gamma_{OUT}| = |S_{22}|$, for $\Gamma_S = 0$].

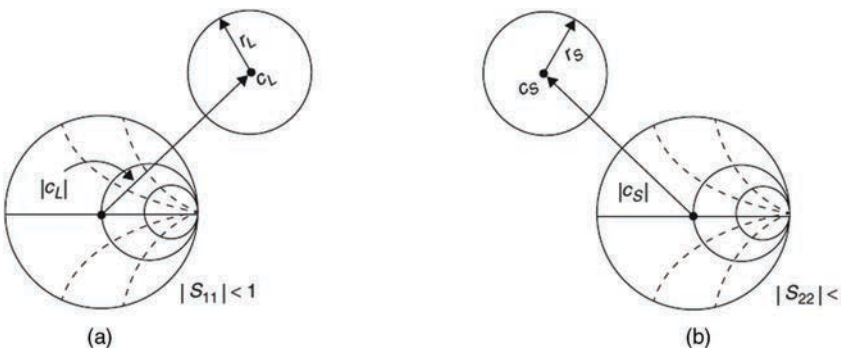


Figure 6.14 Stability circles with respect to the Smith chart for unconditional stability: (a) Plot in Γ_L plane, and (b) Plot in Γ_S plane.

In Γ_S plane:

$$\left|\left|c_S\right| - r_S\right| > 1 \quad \text{for} \quad \left|S_{22}\right| < 1 \tag{6.9b}$$

If either $\left|S_{11}\right| > 1$ or $\left|S_{22}\right| > 1$, the network cannot be unconditionally stable since the termination $\Gamma_L = 0$ or $\Gamma_S = 0$ will make $\left|\Gamma_{IN}\right| > 1$or $\left|\Gamma_{OUT}\right| > 1$ (see Eq. 6.3c).

The necessary and sufficient conditions for an amplifier to be unconditionally stable can be derived by manipulating the Eqs. (6.4a – 6.4c) [3] and is given by

$$K > 1 \quad \text{and} \quad \left|\Delta\right| < 1 \tag{6.10a}$$

$$\text{where } K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} \tag{6.10b}$$

$$\text{and } \Delta = (S_{11}S_{22} - S_{12}S_{21}) \tag{6.10c}$$

It may be noted that for all other cases when K < 1, $|\Delta| < 1$; K < 1, $|\Delta| > 1$; K > 1, $|\Delta| > 1$; the amplifier is potentially unstable or conditionally stable. Manufacturers do not fabricate transistors with K > 1, $|\Delta| > 1$; however, certain terminations or feedback network added to the transistor can sometimes create a potentially unstable amplifier with K > 1, $|\Delta| > 1$.

Example 6.3 The scattering parameters of two different microwave transistors are given below:

(TR_1) $S_{11} = 0.752\angle -138^0$, $S_{12} = 0.092\angle 8.3^0$, $S_{21} = 1.52\angle 40.5^0$, $S_{22} = 0.5\angle -100^0$

(TR_2) $S_{11} = 0.775\angle -165^0$, $S_{12} = 0.030\angle 34^0$, $S_{21} = 6.01\angle 90^0$, $S_{22} = 0.360\angle -35^0$

(a) Determine the nature of stability of amplifiers made with these transistors.
(b) For the potentially unstable amplifier draw, the input and output stability circles on the Smith chart.

Solution:

(a) To know the nature of stability of the amplifiers designed with the above-mentioned three transistors, we need to know the values of Δ and K.

When the design is done with TR_1 we have:

$$\Delta = S_{11}S_{22} - S_{12}S_{21}$$

$$= (0.752\angle -138^0 \times 0.5\angle -100^0) - (0.092\angle 8.3^0 \times 1.52\angle 40.5^0)$$

$$= 0.376\angle -238^0 - 0.139\angle 48.8^0$$

$$= (-0.199 + j0.3188) - (0.0915 + j0.10458)$$

$$= -0.2905 - j0.4238$$

Thus $|\Delta| = \sqrt{(0.2905)^2 + (0.42338)^2} = 0.51337$

and

$$\phi = \tan^{-1}\left(\frac{-0.42338}{0.2905}\right) = \angle -124.46^0$$

Hence, $\Delta = 0.51337\angle -124.46^0$ with $|\Delta| = 0.51337$

$$\text{Now } K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} = \frac{1 - 0.5655 - 0.25 + 0.2635}{0.27968} = 1.6018$$

Thus, for amplifier design with TR_1 will make $K > 1$ and $|\Delta| < 1$; making the designed amplifier unconditionally stable.

When the design is done with TR_2 we have:

$$\begin{aligned}\Delta &= S_{11}S_{22} - S_{12}S_{21}\\ &= (0.775\angle -165^0 \times 0.360\angle -35^0) - (0.03\angle 34^0 \times 6.01\angle 90^0)\\ &= (0.279\angle -200^0 - 0.1803\angle 124^0)\\ &= (0.262 - j0.0954) - (0.1008 - j0.1494)\\ &= 0.1612 + j0.054\end{aligned}$$

Thus $|\Delta| = \sqrt{(0.1612)^2 + (0.054)^2} = 0.1699$

and

$$\phi = \tan^{-1}\left(\frac{0.054}{0.1612}\right) = \angle -161.48^0$$

Hence, $\Delta = 0.1699\angle -161.48^0$ with $|\Delta| = 0.1699$

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

$$= \frac{1 - 0.6006 - 0.1296 + 0.02886}{0.3606} = 0.8282$$

Thus, for amplifier design with TR_2 will make $K < 1$ and $|\Delta| < 1$; making the designed amplifier conditionally stable, i.e., potentially unstable.

(b) The transistors TR_2 is found to be potentially unstable. Thus for this transistor, we need to draw the input and output stability circles by evaluating the centre and radius of the circles using necessary equations as follows.

From Eq. (6.7), we have for *output stability circle*:

$$C_L^* = (S_{22} - \Delta S_{11}^*)^* = S_{22}^* - \Delta^* S_{11}$$
$$= 0.360\angle 35^0 - (0.1699\angle -161.48^0 \times 0.775\angle -165^0)$$
$$= 0.360\angle 35^0 - 0.1316\angle -326.48^0$$
$$= (0.2948 + j0.2064) - (0.1097 + j0.726)$$
$$= 0.1851 + j0.1338$$

$$|C_L^*| = \sqrt{(0.1851)^2 + (0.1338)^2} = 0.2283$$
$$\phi = \tan^{-1}\left(\frac{0.1338}{0.1851}\right) = \angle 35.85^0$$

$$\text{Thus: } c_L = \frac{C_L^*}{|S_{22}|^2 - |\Delta|^2} = \frac{0.2283\angle 35.85^0}{0.1296 - 0.0288} = 2.264\angle 35.85^0$$

$$\text{And } r_L = \left|\frac{S_{12}S_{21}}{|S_{22}|^2 - |\Delta|^2}\right| = \frac{0.030 \times 6.01}{(0.360)^2 - (0.1699)^2} = 1.786$$

Again, referring to Eq. (6.8) we have for *input stability circle*:

$$C_S^* = (S_{11} - \Delta S_{22}^*)^* = S_{11}^* - \Delta^* S_{22} = 0.775\angle 165^0$$
$$- (0.1699\angle -161.48^0 \times 0.360\angle -35^0)$$
$$= 0.775\angle 165^0 - 0.061\angle 3.52^0$$
$$= (-0.7485 + j0.2005) - (0.0608 + j0.0037)$$
$$= -0.687 + j0.1968$$

$$\left|C_S^*\right| = \sqrt{(0.687)^2 + (0.1968)^2} = 0.7146$$

$$\phi = \tan^{-1}\left(\frac{0.1968}{0.687}\right) = \angle 164.02^0$$

$$\text{Thus:} \quad c_S = \frac{C_S^*}{\left|S_{11}\right|^2 - \left|\Delta\right|^2} = \frac{0.7146\angle 164.02^0}{0.6006 - 0.288} = 1.249\angle 164.02^0$$

$$\text{And} \quad r_S = \left|\frac{S_{12}S_{21}}{\left|S_{11}\right|^2 - \left|\Delta\right|^2}\right| = \frac{0.030 \times 6.01}{(0.775)^2 - (0.1699)^2} = 0.315$$

Smith chart plot showing the input and output stability circles are given in Figure 6.15.

NOTE: A few considerations for RF/microwave amplifier design may be discussed briefly in the following. In the design of RF/microwave amplifiers, it is not possible to realize maximum gain and minimum

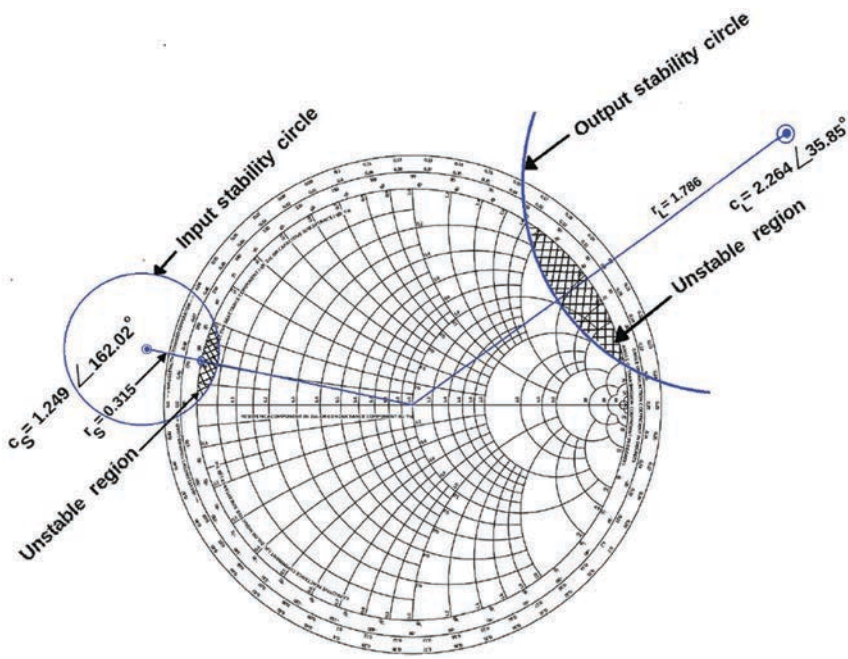


Figure 6.15 Smith chart plot showing the input and output stability circles for Example 6.3.

noise figure at the same time as the two requirements demand different input/output impedance matching. In fact, in the design of LNA, a compromise is done to realize the desired noise figure at the cost of trading-off the gain of the amplifier. In some other applications, to improve bandwidth, the amplifier is operated at less than the possible maximum gain. Thus, in the design of amplifiers, some optimum (i.e., best under given condition/requirement) input/output matching is used to realize the goal of the design as per specification. The analytical approach is to use *constant gain circle* and *constant noise figure circle* [3,4] from which to find the optimum input/output matching impedance. But in modern amplifier design with EDA tool, a designer chooses such optimum load from the data sheet of the active device provided by the device manufacturer.

6.1.3 Low-Noise Amplifier (LNA) vs Power Amplifier (PA)

LNA is used in the front-end of receivers (i.e., as first stage of the receiver amplifier chain) where the power level of the incoming signal is very low. LNA is operated in *small-signal* condition, and its primary aim is to boost the low-level received signal. The noise figure of the LNA primarily determines the overall noise figure of the receiver system (this follows from Friss' formula: $F_n = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + ... + \frac{F_n - 1}{G_1 G_2 ... G_{n-1}}$).

PA on the other hand is used in a transmitter system which is required to transmit an RF/microwave signal with a moderate to high power level. A PA operates under *large-signal* condition. The RF or microwave signal generated in a transmitter is amplified by more than one amplifier stages and its power level is gradually increased. These intermediate amplifiers act as *gain blocks* and *drivers*. The last stage amplifier which is driven by the former ones is the *high-power amplifier*. It finally outputs the required high-power signal to the transmitting antenna system.

The difference between a LNA, a gain block amplifier, and a driver or high-power amplifier (HPA) is tabulated in Table 6.1.

1. ***Class A amplifier***: This is an amplifier in which the active device is so biased that the output current flows throughout the whole of the 360° of the sinusoidal excitation signal. The voltage amplifiers falls in this class of amplifier and are said to be small-signal amplifier.
2. ***Class B and AB amplifier***: A *class B* amplifier is one in which the input bias is approximately the cut-off value of the active device so that with no input signal the output current is zero; however, the output current flows for half (or 180°) of the input signal cycle. A *class AB* amplifier is one in which the output current flows via the

Table 6.1 Difference between LNA, gain block amplifier, and driver or high-power amplifier

	LNA	*Gain Block*	*Driver/ HPA*
Operating signal level	Small-signal condition (–100 dBm to –80 dBm)	Small-signal Condition (-80 dBm to -20 dBm)	Large-signal condition (0 dBm to +15 dBm)
Bias region	Linear	Linear	Saturation
Matching (between the device and the matching network), [see Fig. 6.10]	Input: Noise match: $[Z_S = Z_{\text{opt}(for\,\text{min}.noise)}]$ Output: Conjugate match: $[Z_L = Z^*_{\text{OUT}}]$ (Unilateral case)	Conjugate match: Input: $[Z_S = Z^*_{\text{IN}(at S11)}]$ Output: $[Z_L = Z^*_{\text{OUT}(at S22)}]$	Input: Power match $[Z_S = Z_{\text{IN}(for P_o\,\text{max})}]$ Output: Conjugate power match: $[Z_L = Z^*_{\text{OUT}(for P\,\text{max})}]$ (Bilateral case)
Requirements for matching	Noise Figure of device and noise circles along with *S*-parameters or equivalent linear device model	*S*-parameters or equiv. linear device model	Load-pull data for device along with nonlinear device model
Class of operation†	A	A	B, C, E, F
Power compression	Low P_{1dB}	Low or moderate P_{1dB}	High P_{1dB}; sometimes P_{3dB} is also considered

NOTE†: It would be worthwhile to say herewith a few words about *classification of amplifiers* in terms of the position of the quiescent point, i.e., biasing condition of the active device (transistor/tube) used in the design of the amplifier and according to that we have *Class A*, *Class B* and *Class AB*, and *Class C* amplifier, see Figure 6.16.

active device for appreciably more than half but less than the entire cycle of the input signal.

3. ***Class C amplifier***: If the bias condition of the active device is such that the angle of current flow via the device is appreciably less than 180°, *Class C* operation is involved.

Higher classes, such as D, E, F uses the active device as a switch to pump a highly resonant tank circuit, and may achieve very high efficiencies.

6.1.4 Design of LNA and PA Using EDA Simulation Tool

Let us now discuss the design of the low-noise amplifier (LNA) and power amplifier (PA) using the standard electronic design automation (EDA) simulation tool which is followed these days for professional engineering design of RF/microwave amplifiers. The design approach given here is rudimentary,

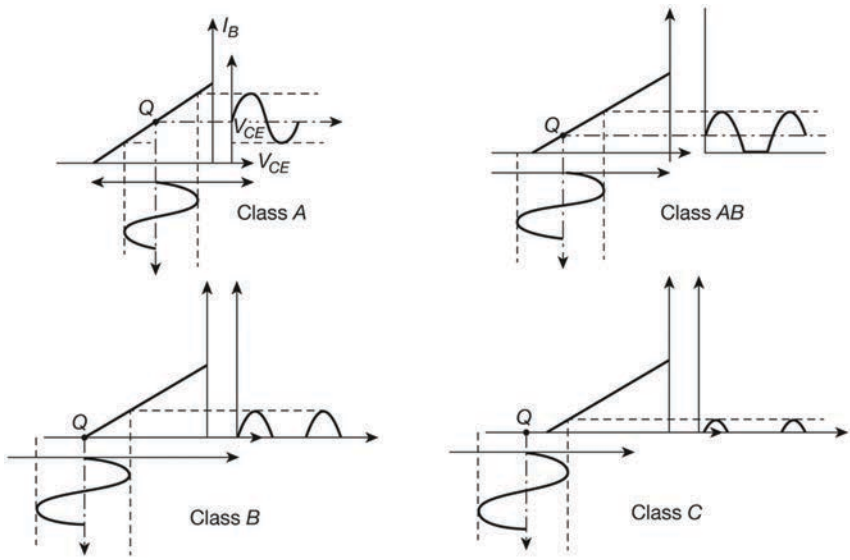


Figure 6.16 Amplifier classification in terms of the position of the quiescent point of the device.

which might help students/designers to get a preliminary flavor of the design steps while designing such amplifiers.

Example 6.4 Design a single-stage pHEMT-based low-noise microwave amplifier for satellite up-link system with the following specifications:

Frequency of operation: 14 ± 0.2 GHz
Small-signal gain: 10 ± 0.5dB
Noise Figure: < 3.5 dB
Input return loss: < 10 dB
Output Return Loss: < 15 dB.

Solution:

The design of the amplifier can be done in the following steps using EDA simulation tool.

Step 1: *Selection of the device:* The selection of the pHEMT device is done on the basis of the cut-off frequency of the device, maximum available gain, and the noise figure desired from the amplifier. The device manufacturer provides the small-signal model of the device

that contains the values of the intrinsic and extrinsic elements of the packaged device (Figure 6.17). The designer is required to rebuild the model using the given circuit elements in the circuit simulator prior to the initiation of the design.

The device manufacturers may also provide the S2P file for the device which is the 2-port *S*-parameter file that contains the raw 2-port S-parameter data and noise data of the device over a given operating frequency range and for a particular bias point.

Step 2: *Biasing:* The biasing of a FET requires positive voltage at the drain terminal ($+V_{DS}$) and a negative voltage on the gate ($-V_{GS}$), and the source is usually kept grounded. The biasing scheme and typical I–V characteristics of the pHEMT device used for the design is shown in Figure 6.18.

Bias de-coupling circuit is an essential part of amplifier design to prevent the RF signal at the amplifier from entering into the path of the DC supply. To realize this, inductor acting as RF choke and a few bypass capacitors to ground are used (Figure 6.19a). The combination basically acts as band-reject filter at the RF frequency. The inductive choke may be either a surface mount inductor which provides high impedance to RF signal, or a short-circuited transmission line stub that behaves as open circuit to the RF signal. The capacitor may be either a surface mount capacitor or an open circuit stub that behaves as a short circuit for the RF signal. The lumped components are used from the component library of the simulation system and the distributed

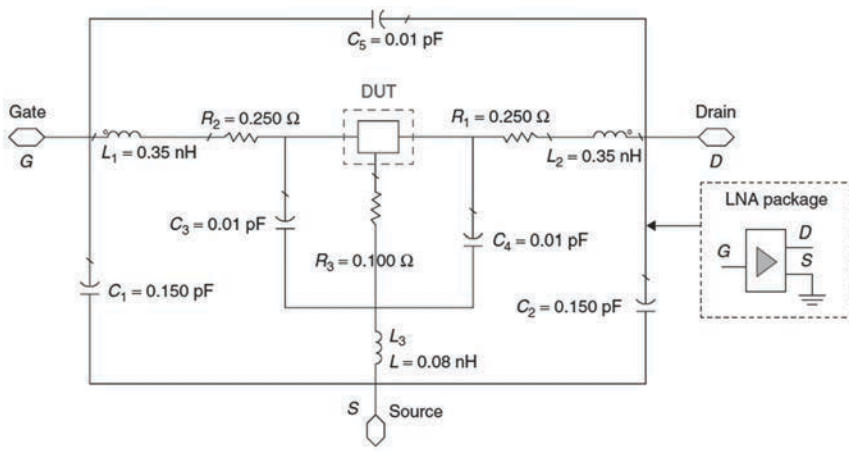


Figure 6.17 Small-signal model (equivalent circuit) of packaged pHEMT device.

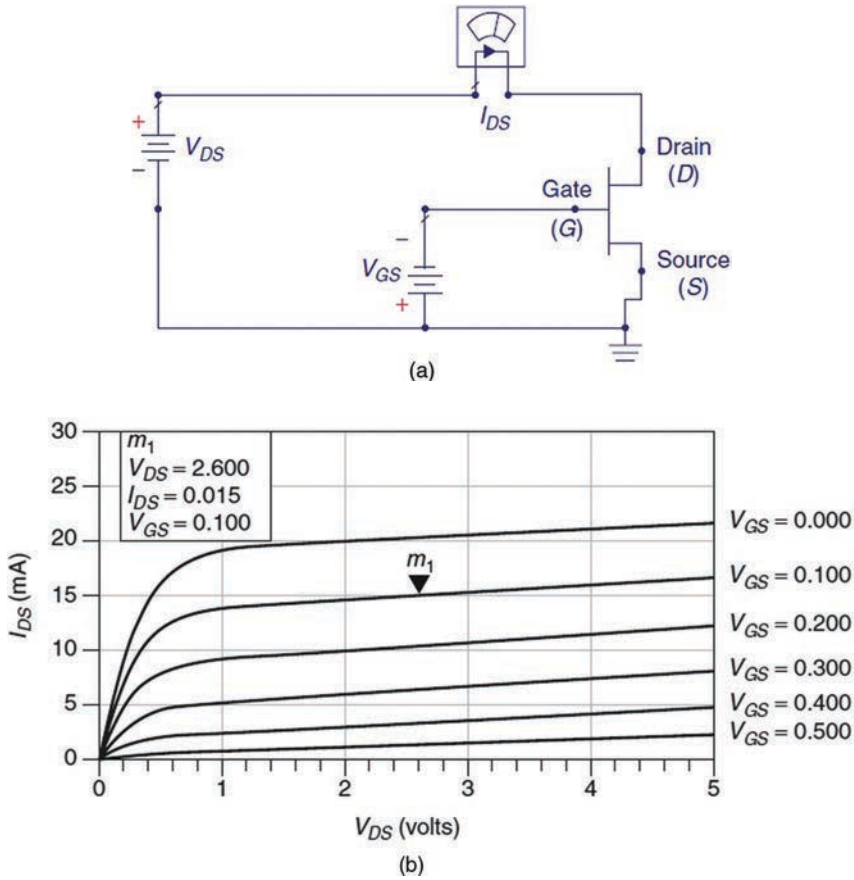


Figure 6.18 (a) Biasing scheme, and (b) I–V curves for pHEMT device used in the design.

components are designed with 10mil (0.025mm) thick Rogers 5880 substrate for microstrip-based transmission line stub design. The bias de-coupling circuit used is shown in Figure 6.19b, while the simulation result shown in Figure 6.19c indicates that S_{21} is below –43 dB (ensuring complete isolation of RF port from the DC port) and S_{11} is 0.999 (indicating complete reflection of the signal at the RF port). A stabilising resistor is usually introduced in the de-coupling section for the gate terminal that ensures small-signal stability and prevents any possible oscillation to build-up at low frequencies.

Step 3: *Input and output matching circuits:* In Section 6.1.2.1, we discussed the importance of input matching network (IMN) and output matching network (OMN) and what matching conditions are to be used for the LNA.

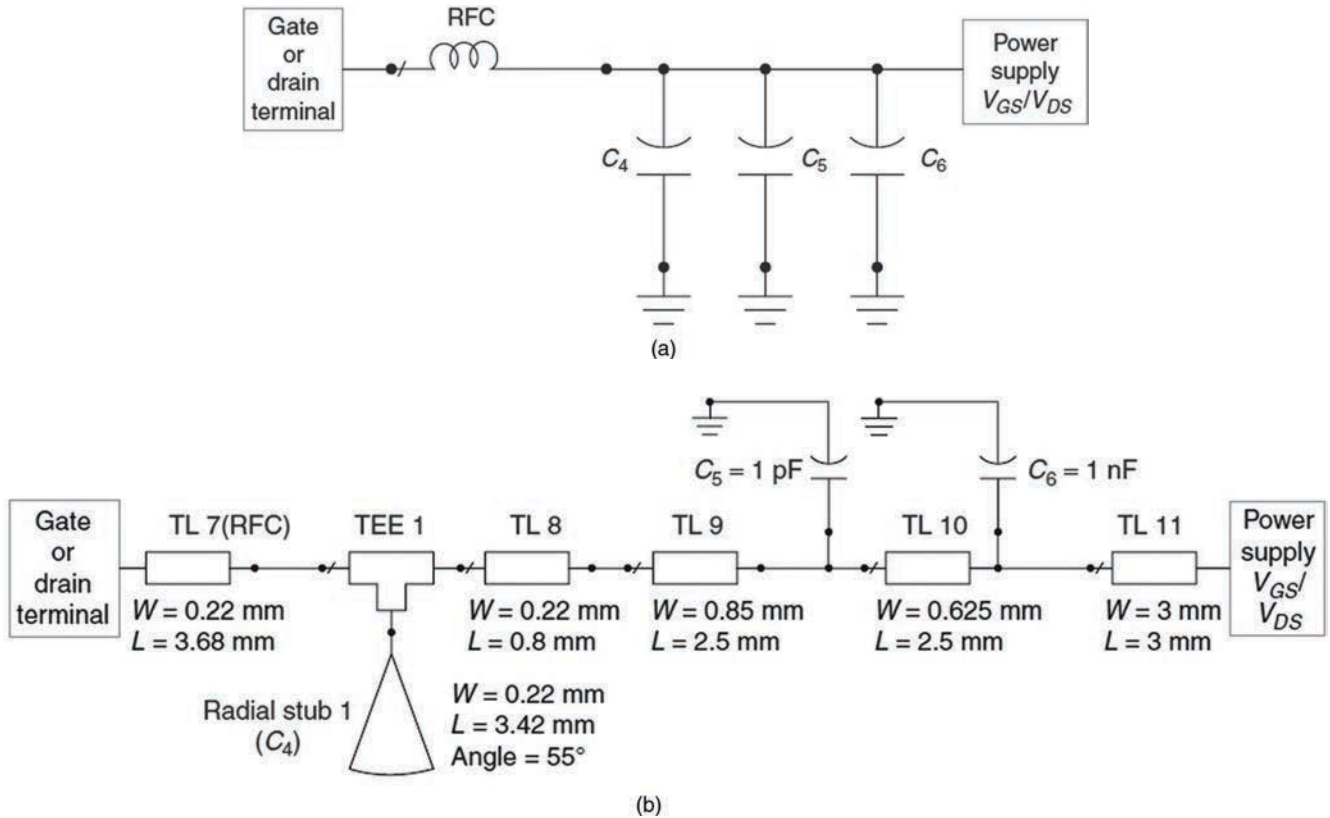
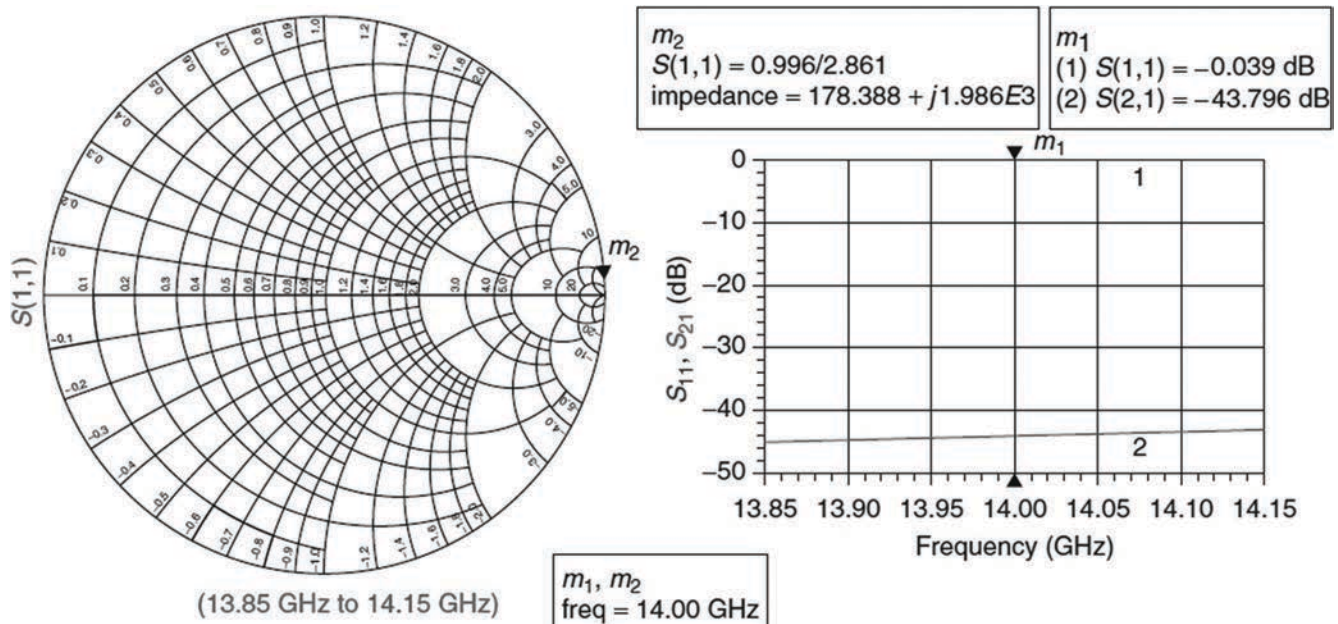


Figure 6.19 Bias decoupling circuit: (a) Lumped circuit schematic, (b) Transmission line realization, and (c) Simulation result of bias de-coupling circuit.



(c)

Figure 6.19 (Continued)

In the present design, Z_{in}, Z_{out} and Z_{opt} are, respectively, given by (44.26 + j85.93)Ω, (4.09 – j58.29)Ω, and (55.42 – j147.28) Ω; which are either obtained from the 2-port S-parameter data (supplied by the manufacturer) or form S-parameter response of the transister model with calculation or with the EDA tool (ADS), the later technique being explained in the next paragragph. Anyway, for conventional amplifier design where gain has to be maximized, conjugate match has to be done, i.e., $Z_S = Z^*_{in}$ and $Z_L = Z^*_{OUT}$ (for practical bilateral case, $S_{12} \neq 0$). However, for LNA design the input matching is done with $Z_S = Z_{opt}$, sacrificing little bit of the gain of the amplifier.

How to get Z_{in}, Z_{out} and Z_{opt} with ADS simulator for the device used and the bias point to which the device is biased to is explained below. The FET device is biased at the selected DC bias point and checked for its RF response to evaluate the maximum available gain, stability, and input/output impedances over the frequency range of its operation. For that purpose, the FET is attached to 50Ω terminations as shown in Figure 6.20, with DC blocks at input and outputs and proper DC bias in the drain and gate terminals using an ideal DC feed. This template gives the unmatched S-parameter response over frequency range of interest. The Z_{in}, Z_{out}, and Z_{opt} pallets can be used in ADS tool to evaluate the values of input, output, and optimum noise impedances over the operating frequency range. These values are then utilized for simultaneous conjugate matching of the input and output of the transistors to the source and load impedances of 50Ω, respectively. The Z_{opt} is used to achieve noise match at the input end of the FET.

It may be noted that L/∏/T-section impedance matching techinique can be used for impedance matching both at the input and output of the device. The inductance is realized with short-circiuted stub and capacitor with open-circuited stub in microstrip transmission line. The matching sections are shown in Figure 6.21. Initial matching of the input impedance was done between Z_{opt} and Z_S then the matching has been slightly detuned to achieve good input match (in terms of S_{11}) at the cost of increasing the noise figure within permissible limits. DC blocking capacitors have also been included as part of the matching circuitry. The values of S_{11}, S_{22}, and S_{opt} before and after impedance matching are shown in Figure 6.22. After incorporating above impedance matching sections, the curves for S_{11}, S_{22}, and S_{opt} are observed to approach the 50Ω point; however, S_{11} and S_{opt} are seen to be largely seperated from each other indicating a trade-off between optimum noise figure and good input match that is used for LNA design.

Step 4: *Schematic/circuit-level design:* The circuit schematic of a single-stage LNA can be built by combining different blocks of Figure 6.23

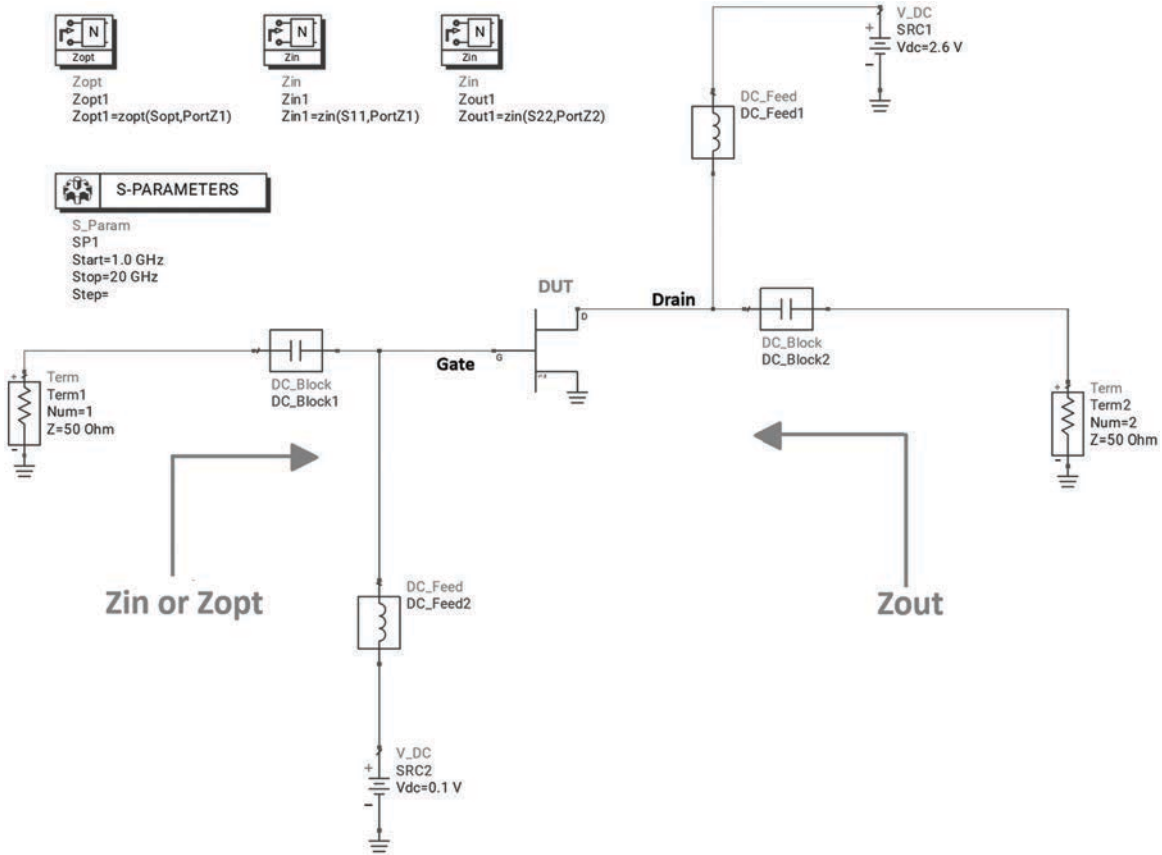


Figure 6.20 Representative diagram to indicate as how to get Z_{in}, Z_{out}, and Z_{opt} with ADS simulator for the device used and the bias-point to which the device is biased to.

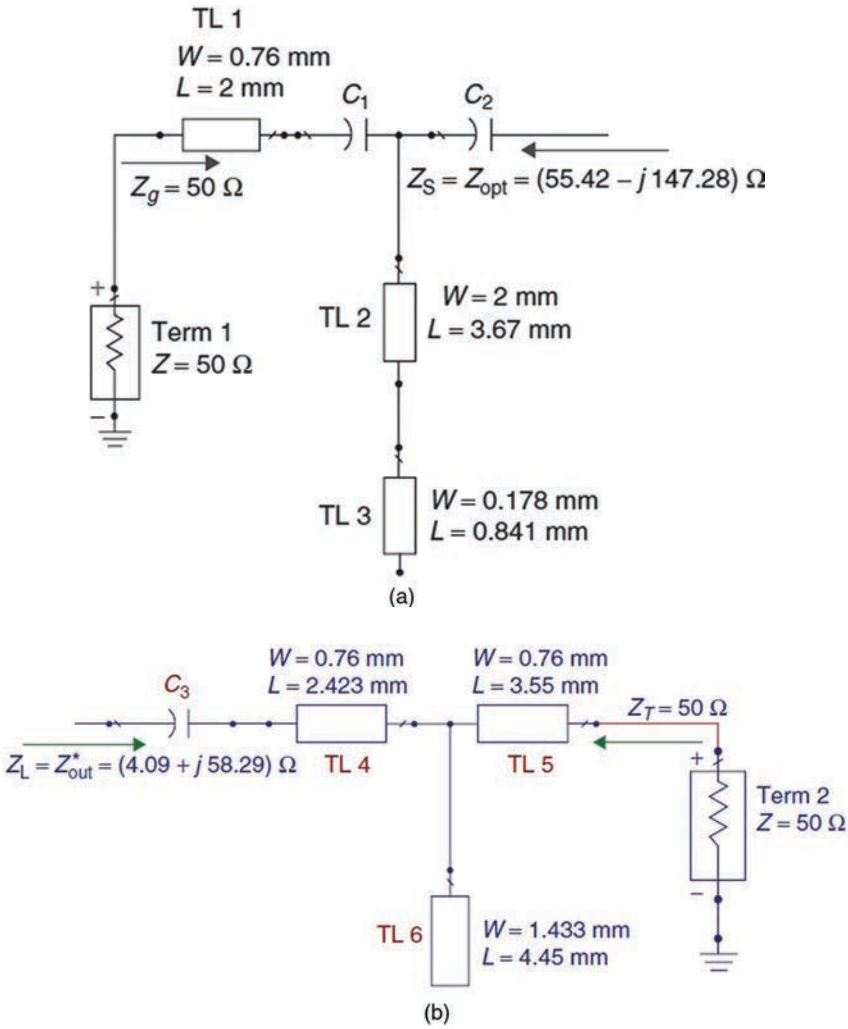


Figure 6.21 Amplifier matchning sections: (a) Input matching section, and (b) Output matchning section.

discussed in earlier steps. The initial amplifier design is done using the *circuit simulator* of the EDA simulation tool. However, to include the parasitic effects, the system has to be finally optimised in *layout* using the Method of Moments (MoM) solver that performs the electromagnetic (EM) simulation.

Step 5: *Layout design and optimization:* The final optimized layout developed for a single-stage LNA is shown in Figure 6.24 together

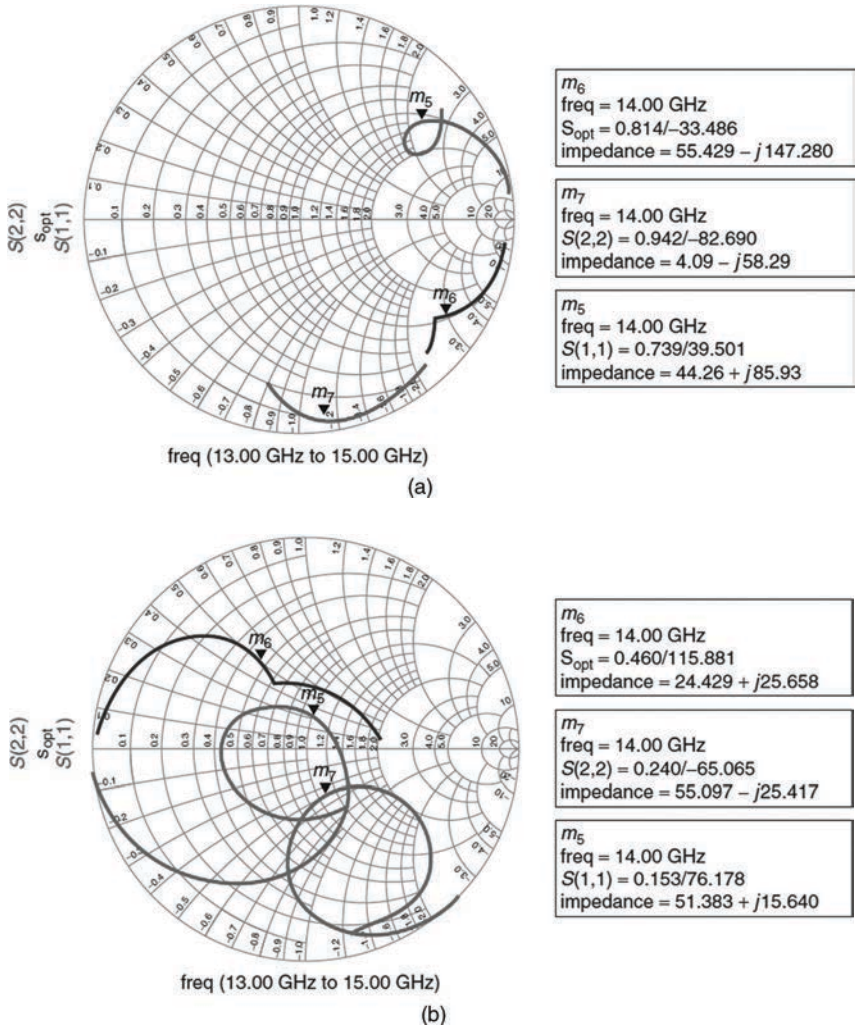


Figure 6.22 Smith chart showing S_{11}, S_{22}, and S_{opt}: (a) Before input-output matching, and (b) After input–output matching.

with electromagnetic (EM) simulation result for gain, noise figure, and stability of the amplifier in Figure 6.25.

A final gain of 11.10 dB has been achieved with input and output return loss of better than 14dB and 20dB, respectively, at a frequency of 14GHz (see Figure 6.25). The noise figure achieved is 3.3 dB at the design frequency, and the amplifier is seen to be unconditionally stable, as the stability circles are outside the Smith chart.

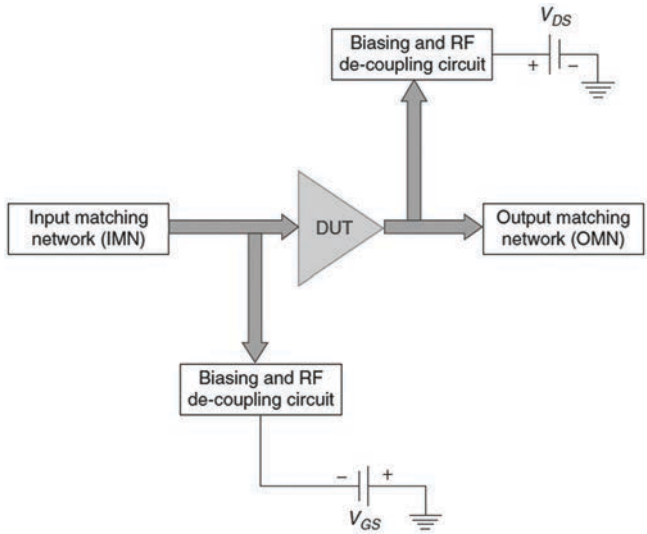


Figure 6.23 Block diagarm showing the circuit schematic of single-stage LNA.

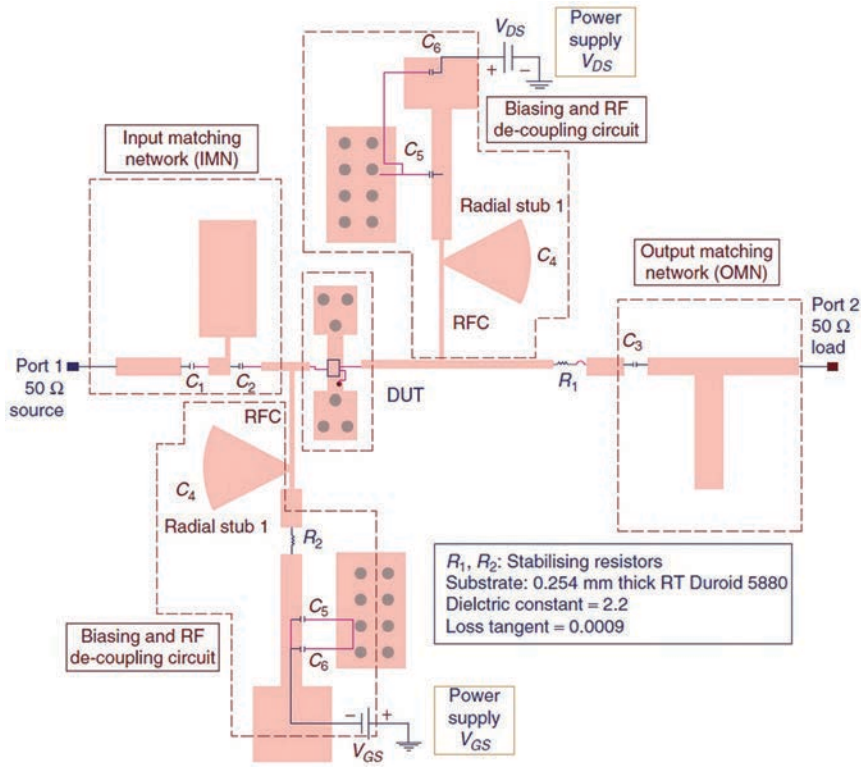


Figure 6.24 Designed layout for single-stage LNA simulated in EDA tool.

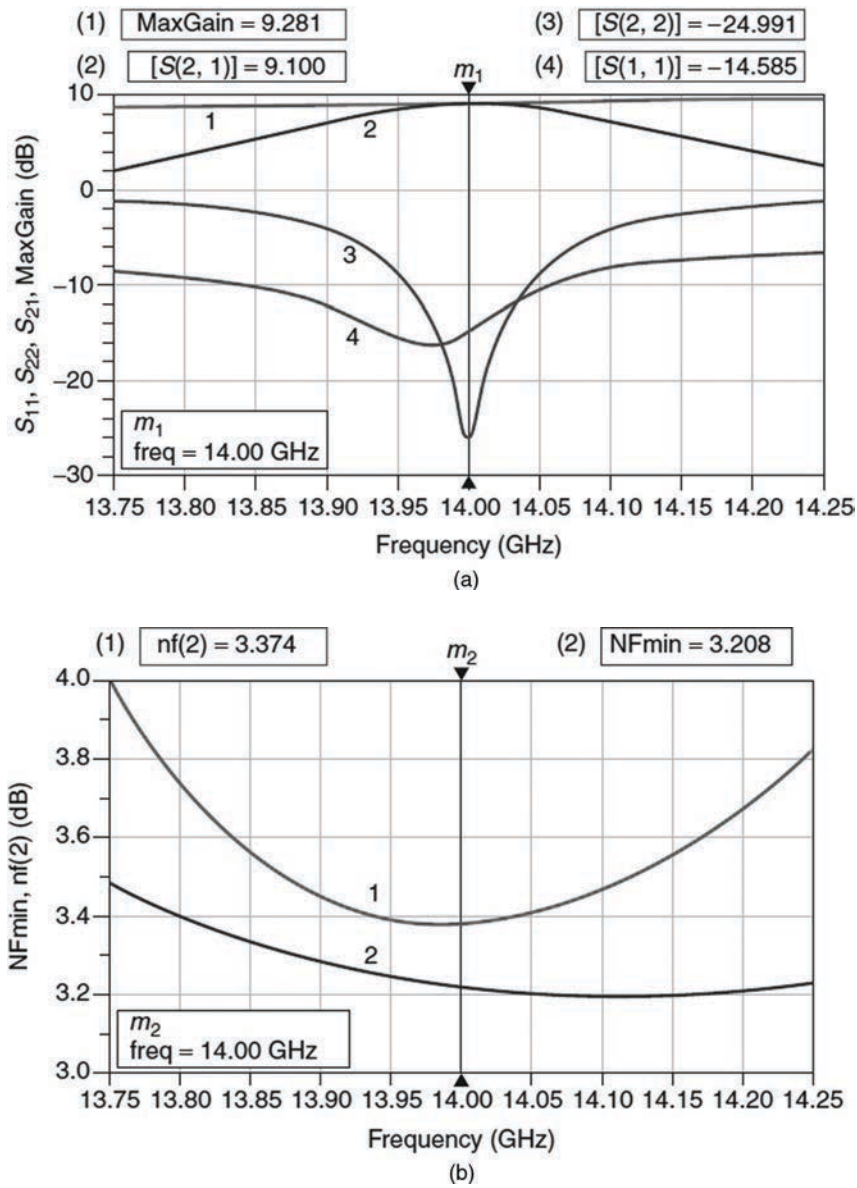


Figure 6.25 EM-simulated result optimized in layout showing: (a) Gain, (b) Noise figure, and (c) Stabilty of the designed amplifier.

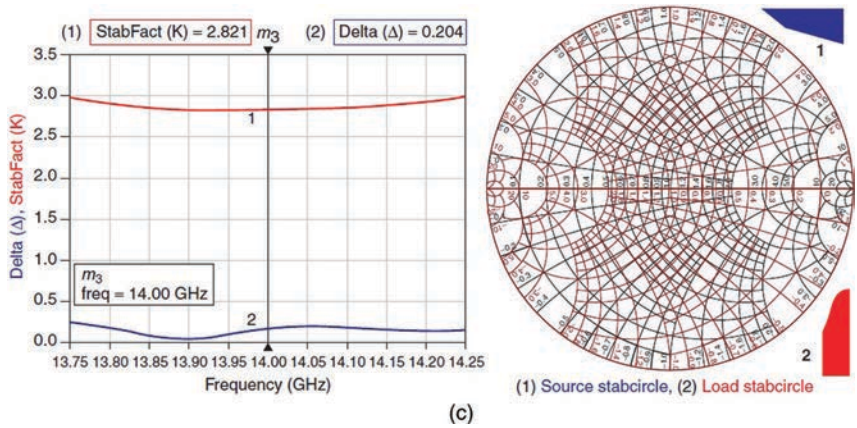


Figure 6.25 (Continued)

Example 6.5 Design a single-stage power amplifier at 6GHz having an output power of 15 watts; the design specifications being:

Frequency: (6 ± 0.2)GHz
Small-Signal Gain: (15 ± 0.5)dB
Output Power@ 1dB Compression: (42 ± 1)dBm
Input Return Loss: < 15dB
Output Return Loss: < 10dB

Solution

The design of the amplifier can be done in the following steps using the EDA simulation tool.

Step 1: *Device model and bias point selection:* The device chosen is T2G6003028 which is a gallium nitride (GaN)-based HEMT device that can provide a power output of up to 25 watts at 6GHz. Common-emitter configuration is used for performing DC simulation (and also in consequent load-pull analysis) in which the amplifier is driven by a source with impedance of 50Ω at the fundamental frequency (which is also assumed to be the same at higher harmonics). The biasing circuit and the I–V characteristics of the device are shown in Figure 6.26.

Step 2: *Load-pull test:* Power amplifiers are generally operated with input power level comparable to or greater than IP_{1dB}, where the nonlinearity of the HEMT device sets in. Under such *large-signal* operation of the amplifier, the measured scattering parameters will depend on the input power level and the output termination impedance. To characterize the large-signal behaviour of a transistor, the contours of constant power output are plotted on a Smith chart as a function of the load reflection coefficient, known as *load-pull contours*. Thus, for a large-signal operation of power amplifiers, load-pull test is required to be done to obtain the optimum load impedance at which maximum power is delivered by the device to its output load or maximum power-added efficiency [PAE = $(P_{out} - P_{in})_{(RF)}/P_{(DC)}$] is achieved (whichever is desirable based on the application). The load-pull test is based on *harmonic balance* simulation method. Usually, load pull is a measurement methodology by which a device's large-signal model can be extracted. Sometimes only its load impedance contours for different power outputs or for optimum PAE are collected through this measurement for the design of power amplifiers. In the present design, such a device model shared by Modelithics [www.modelithics.com] has been used.

The simulation template (see Figure 6.27) provided in the model helps to extract and view the load-pull data of the device. A series of power and PAE contour lines have been plotted on the Smith chart, representing the load reflection coefficients with similar output power level. At the optimum load impedance, the power amplifier will deliver maximum power to the load (unlike small-signal operation of amplifiers where maximum gain is desired).

A load tuner block (see Figure 6.28) is utilized, by which the load to the device is varied and the power contours for all possible load combinations are plotted. The source used is a one-tone (fundamental frequency) power source. As shown in Figure 6.28, this example uses a source that delivers a power of +23dBm if a load of 50Ω is connected to it; in other words, P_{in} = +23 dBm is the *available power*.
Optimum load impedance can be determined in the following steps:

(i) Set the source power at a starting value, say +23dBm. The input impedance at the fundamental frequency is initially kept at 5Ω.
(ii) Perform the load-pull test. Find the load impedance for maximum output power from the power contours.
(iii) Check the source impedance for that maximum power output.

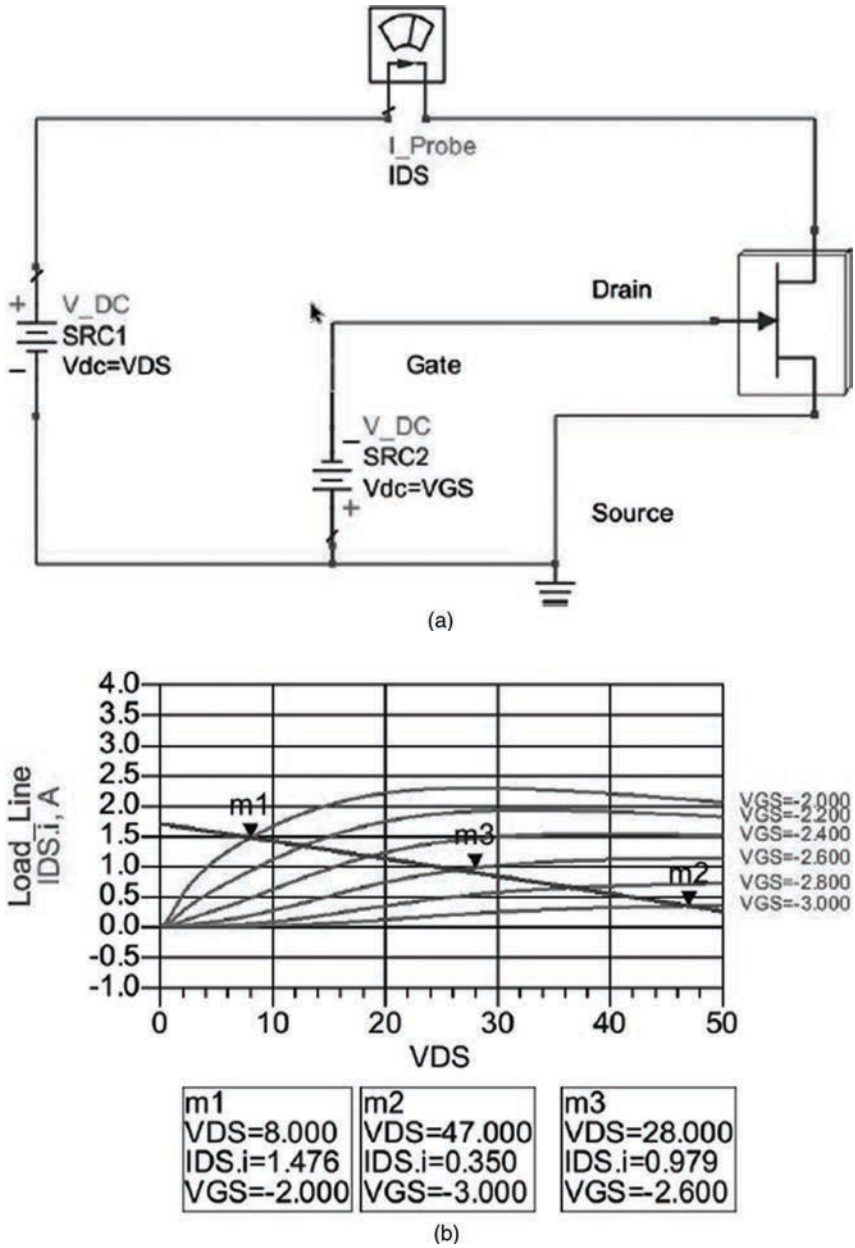


Figure 6.26 (a) Biasing scheme, and (b) I–V characteristics of the device.

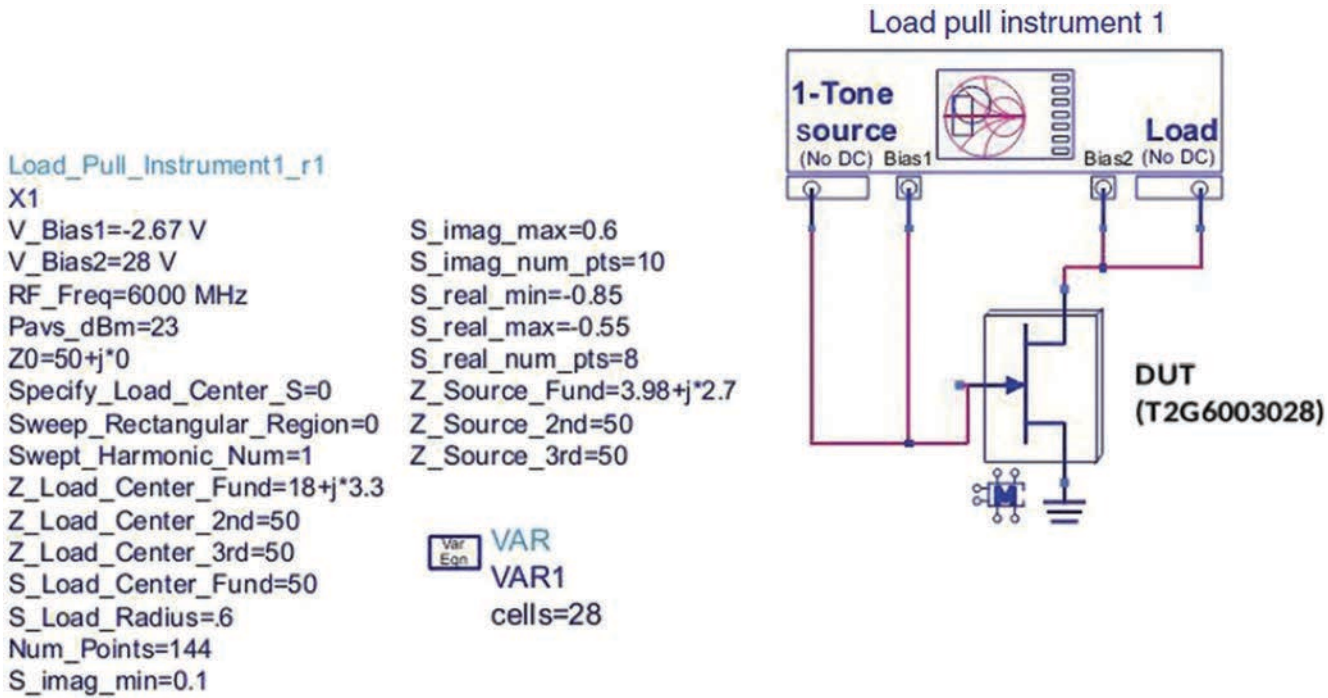


Figure 6.27 Screenshot of the load-pull simulation template.

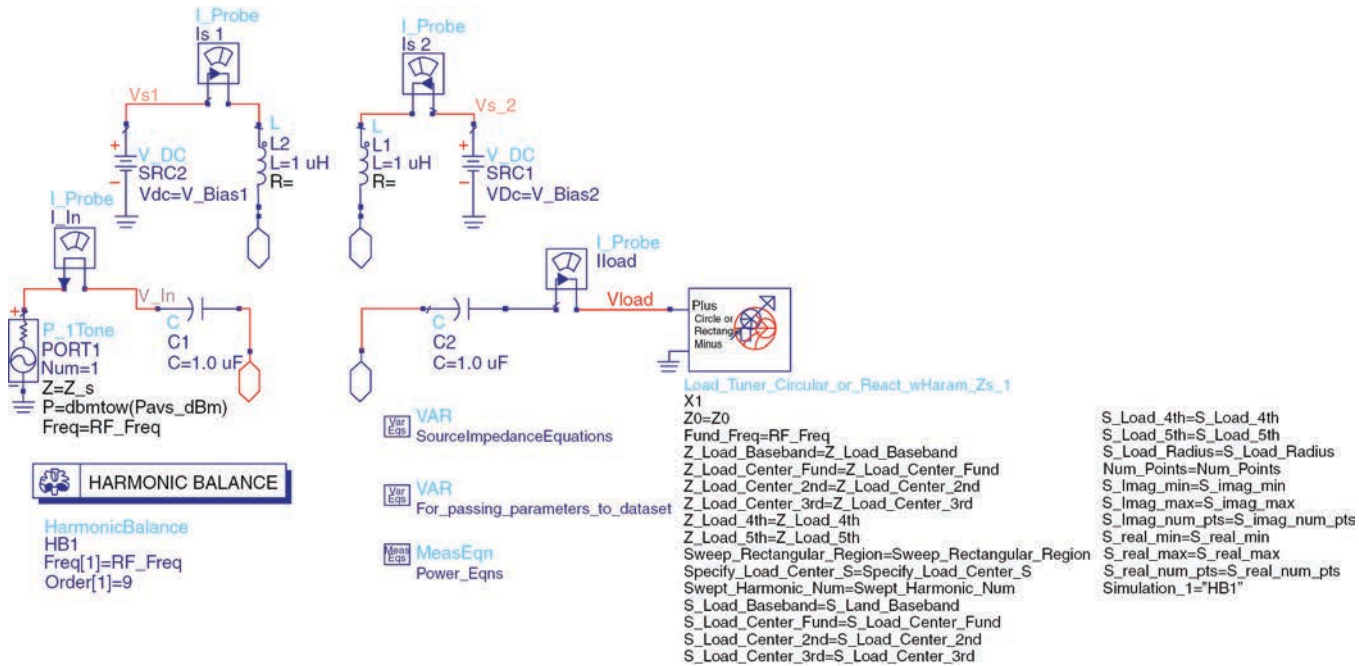


Figure 6.28 Screenshot of the load-pull tuner used in the template above.

(iv) Use the derived source impedance value in the template by replacing the initial setting of 5Ω.
(v) Re-simulate the load-pull engine.
(vi) Two or three such iterations will show that a particular value for load and source impedance is obtained for the maximum output power.
(vii) The load impedance for maximum output power is the optimum load impedance.

The load-pull test result is given in Figure 6.29.

From Figure 6.29, it may be observed that the maximum power that the power amplifier can deliver without significant non-linear distortion is 41.14 dBm (13 watt).

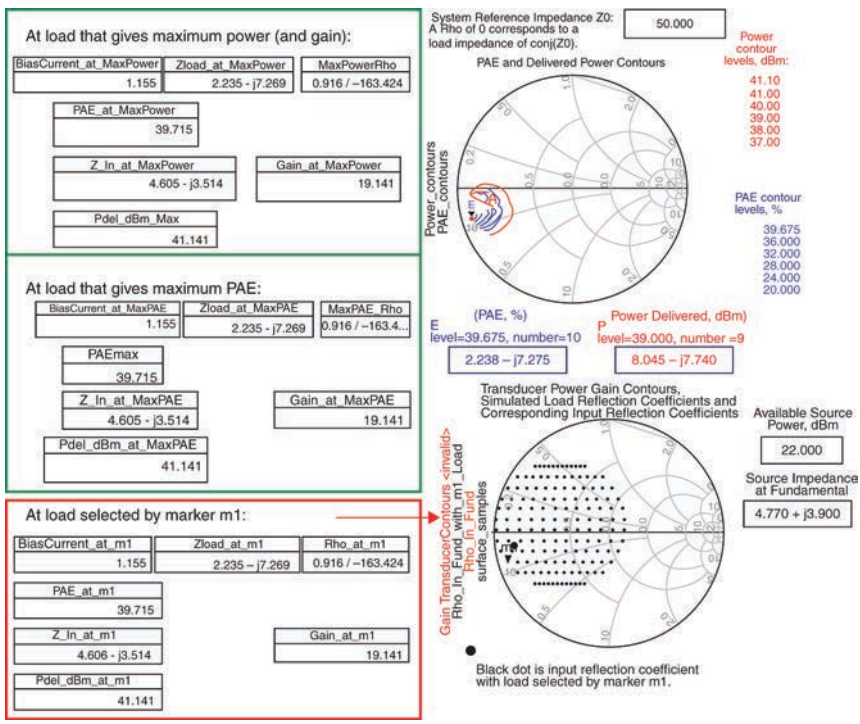


Figure 6.29 Screenshot of the load-pull test result.

Step 3: *Input and output impedance matching:* In Section 6.1.2.1, the matching conditions to be used for the power amplifier (PA) design is given.

Input matching: The optimum input impedance of the device is obtained from load-pull test as $Z_{in} = (4.605 - j3.514)$; see Figure 6.30a. This value of input impedance corresponds to maximum power output of 41.141 dBm for an input frequency of 6 GHz and a power drive level of 23 dBm. Hence if the input matching circuit (IMN) is matched to this optimum value of Z_{in} then maximum power is achievable at the output. The input matching network is shown in Figure 6.30b and the input return loss obtained is shown in Figure 6.30c.

Output matching: The optimum output impedance of the device is obtained from the load-pull test as $Z_{OUT} = 2.235 - j7.269$ (see Figure 6.31a). This value of output impedance corresponds to maximum power output of 41.141 dBm for an input frequency of 6 GHz and a power drive level of 23 dBm. Hence, if the output matching circuit is matched to the complex conjugate of this optimum value of Z_{OUT} then maximum power is achievable at the output. The output

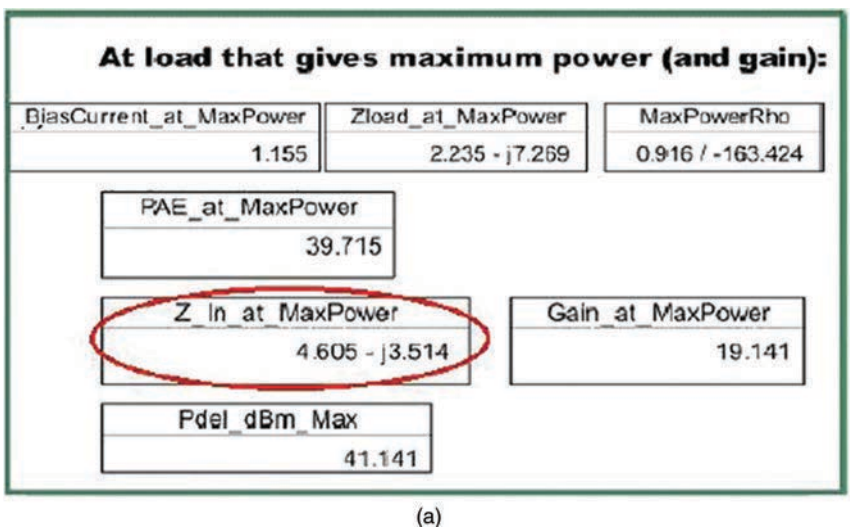


Figure 6.30 Input matching (a) Optimum impedance for input match, (b) Screenshot of input matching network (IMN) and bias section for gate, and (c) Input return loss.

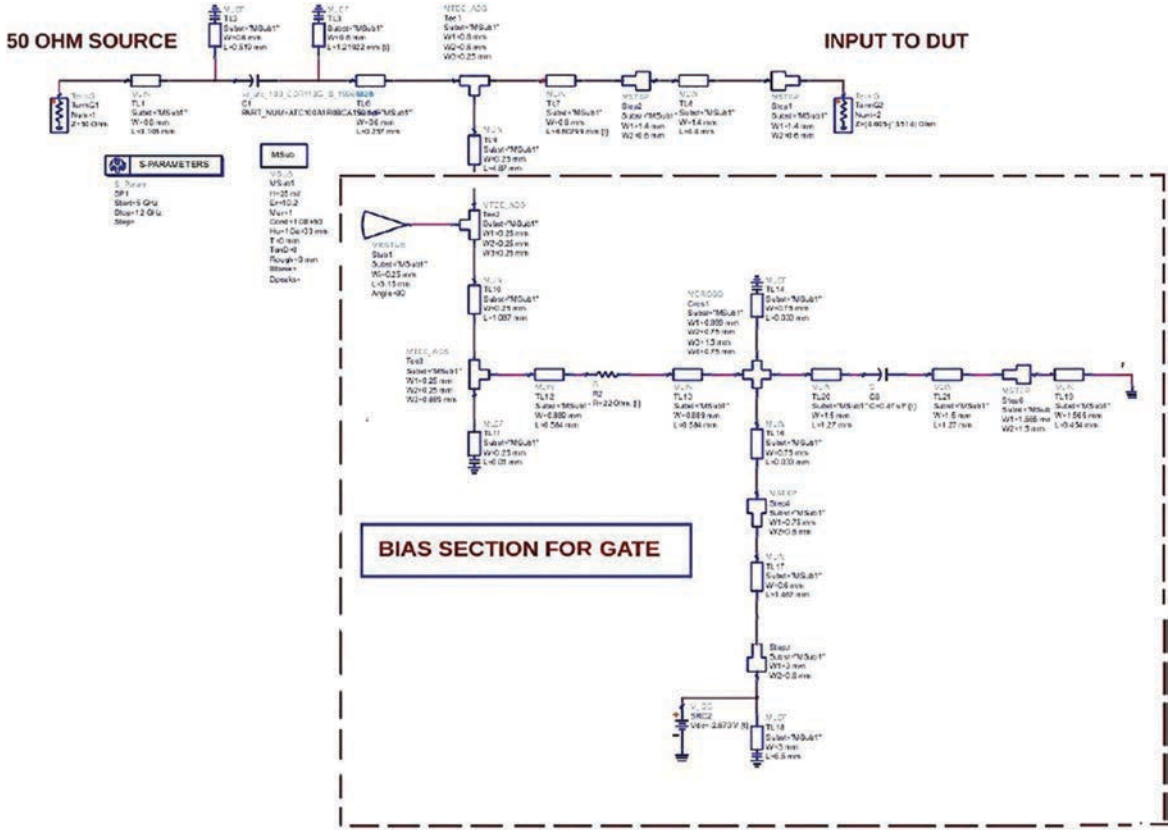


Figure 6.30 (Continued)

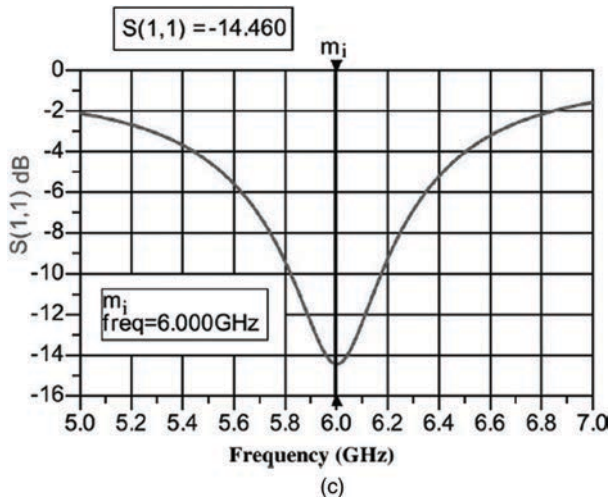


Figure 6.30 (Continued)

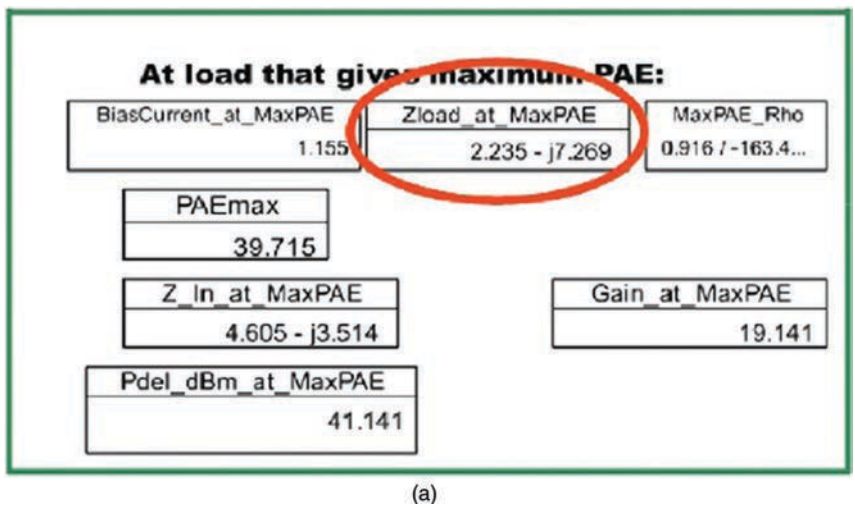


Figure 6.31 Output matching: (a) Optimum impedance for output match, (b) Screenshot of output matching network (OMN) and bias section for drain, and (c) Output return loss.

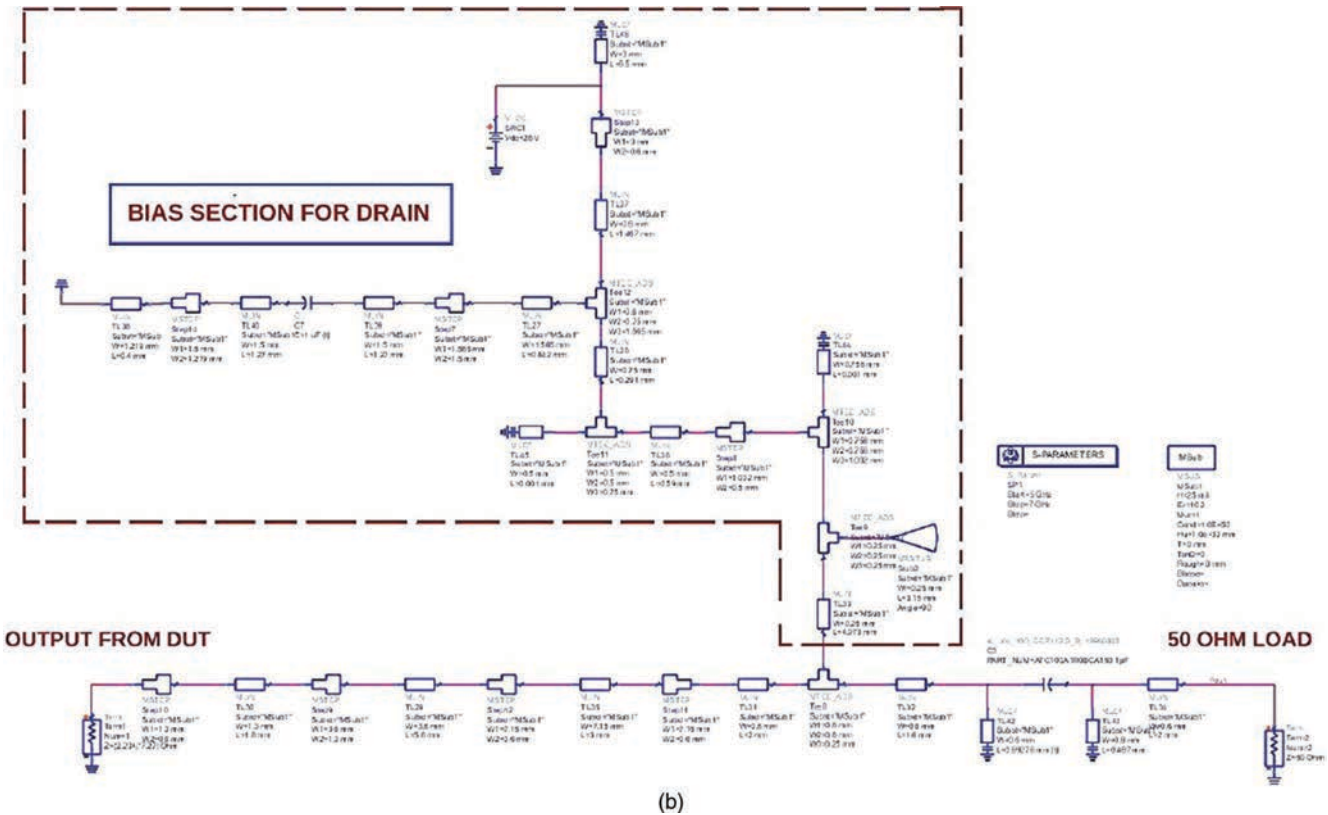


Figure 6.31 (Continued)

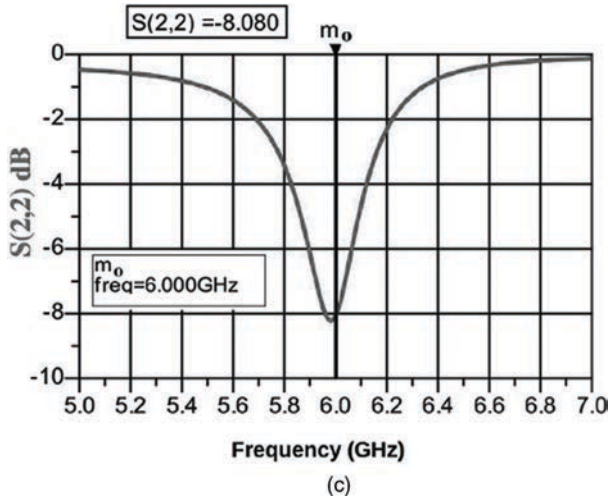


Figure 6.31 (Continued)

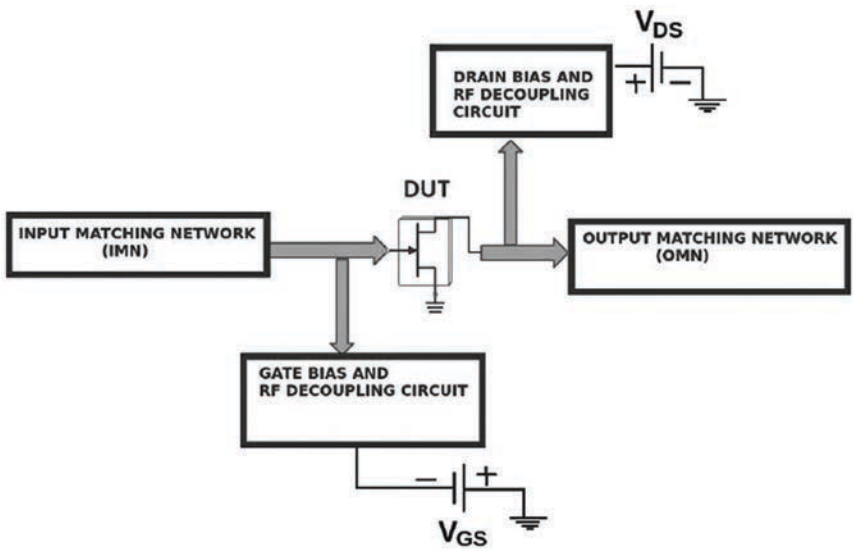


Figure 6.32 Block diagram showing the different component parts of the power amplifier.

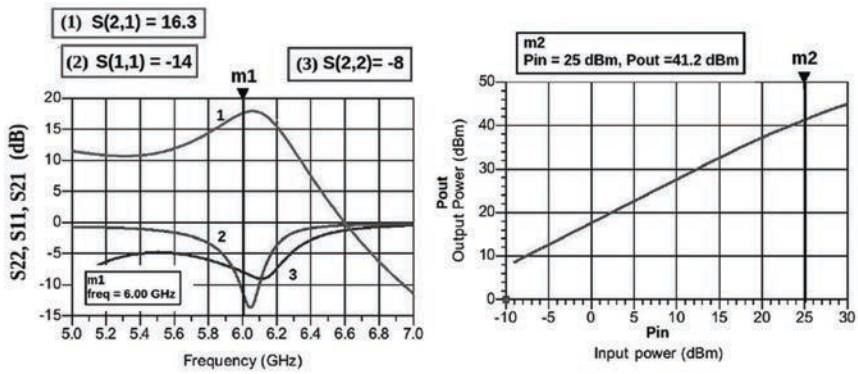


Figure 6.33 Simulation result showing the characteristics of the power amplifier.

matching network is shown in Figure 6.31b, and the output return loss obtained is shown in Fig ure 6.31c.

Step 4: *Complete Amplifier circuit in block diagram:* The complete amplifier circuit in terms of block diagram is shown in Figure 6.32.

Step 5: *Simulation result of the designed power amplifier:* The simulation result showing the characteristics of the power amplifier is given in Figure 6.33.

Step 6: *Layout of the designed power amplifier:* The layout of the designed power amplifier is shown in Figure 6.34.

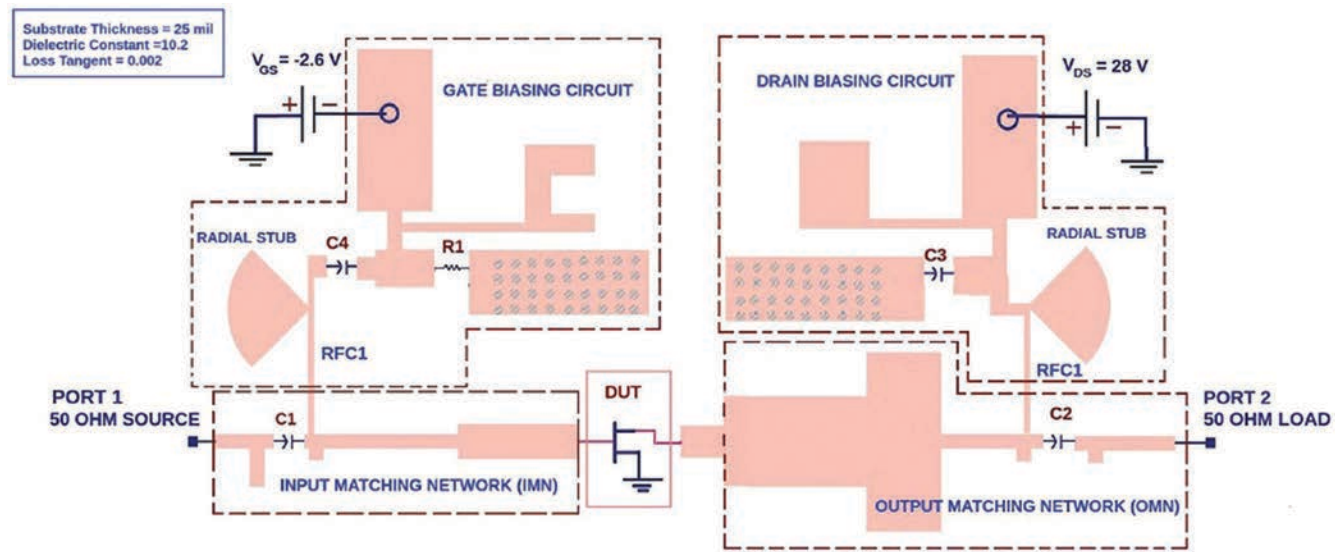


Figure 6.34 Layout of the designed power amplifier.

REFERENCES

[1] Holzman E, and R S Robertson, *Solid-State Microwave Power Oscillator Design*, Artech House Inc., Boston, 1992.
[2] Ludwig R, *RF Circuit Design: Theory and Application*, Second Edition., Pearson Education, India, 2011.
[3] Gonzalez G, *Microwave Transistor Amplifiers Analysis and Design*, Second Edition, Prentice Hall Inc, New Jersey, 1996.
[4] Liao S Y, *Microwave Circuit Analysis and Amplifier Design*, Prentice Hall Inc, New Jersey, 1987.

Chapter 7

Design of Microwave Power Combiners and Mixers

Mix with like-minded people and you will never be hurt

7.1 MICROWAVE POWER COMBINERS

7.1.1 Introduction

It is known that the power output (P_O) realizable from a device decreases rapidly with the inverse square of the frequency of operation (f_{op}): ($P_O \alpha f_{op}^{-2}$). Thus, to achieve high power output at microwave and millimetre-wave frequencies with solid-state devices like IMPATT/Gunn, power combiners are needed to combine the power from several individual devices. This is the demand of the day since almost all radar transmitters for airborne and space applications, including military applications, prefer to use solid-state power sources in order to realize power amplification in the transmitter stage. A microwave power combiner, in principle, desires to add the microwave power of several devices (as in device-level combining) or of several single-device oscillators (as in circuit-level combining) as efficiently as possible. Device-level combining is accomplished by clustering the devices in a region whose extent is small compared with a wavelength and is thus limited in the number of devices that can be efficiently combined. However, circuit-level combining is not limited to placing the diodes in an area small compared to a wavelength. Circuit-level combiners are more practical and are, in general, of two categories—those which combine the output of N devices in a single stage (known as N-way in single-stage) and those which combine the output of N devices in multiple stages (known as N-way in multiple-stage). The latter category is simpler and more widely used; however, the former category has the possibility of higher power combining efficiency since power generated does not have to pass through several stages of combining. The performance of a power combiner is characterized in terms of power combining efficiency defined as $\eta_c = P_{OT}/N.\ P_{OS}$, where P_{OT} is the net power at the output of the combiner using N number of active devices and P_{OS} is the power output expected from single device oscillator using the same devices as in the combiner. Typical power combining efficiency of a well-designed power combiner may be around 90%.

DOI: 10.1201/9781003729969-7

7.1.2 Different Types of Microwave Power Combiners

7.1.2.1 N-Way in Single-Stage/Multiple-stage Power Combiner

The basic *N-way in single-stage* power combiner designed with resonant cavity combining structure was due to Kurokawa and Magalhaes [1] in 1971. Here, the sum of output powers from a number of devices is obtained by coupling their outputs to a single resonator (Figure 7.1). It basically represents pairs of double coaxial-entry waveguide systems in which each coaxial line of a pair is located close to the side wall of the rectangular waveguide where the magnetic field maxima exist. The active device is mounted at one end of the coaxial line and the central conductor of the line helps to feed the necessary DC bias to the device. Since along the propagation direction of the rectangular waveguide the magnetic field maxima repeats at every $\lambda_g/2$, the coaxial line pairs mounted in the E-plane of the guide are spaced by $\lambda_g/2$ in the z-direction (Figure 7.1). The combiner output is taken via an iris from one end of the guide (say at $z = l$), and the other end (at $z = 0$ plane) the guide is closed by a sliding short.

Another simple, compact, yet elegant duo-diode power combiner can be a twin-cap combiner in which IMPATT (or Gunn) diodes embedded in radial-line cavities combine their power at the electric field maximum of the waveguide in which the twin-cap structure is mounted [2] (Figure 7.2). The twin-cap structure constitutes coupled radial-line cavities resembling a couplet of the conventional resonant-cap cavity used in the design of single-diode oscillator (see Appendix III). This duo-diode configuration may be

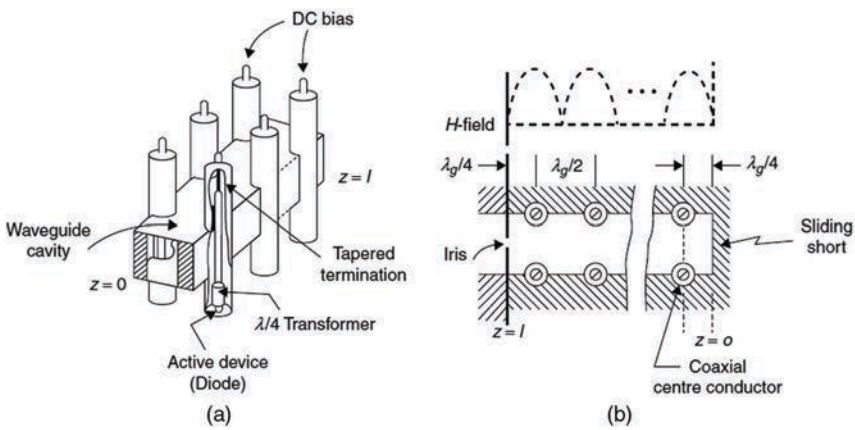
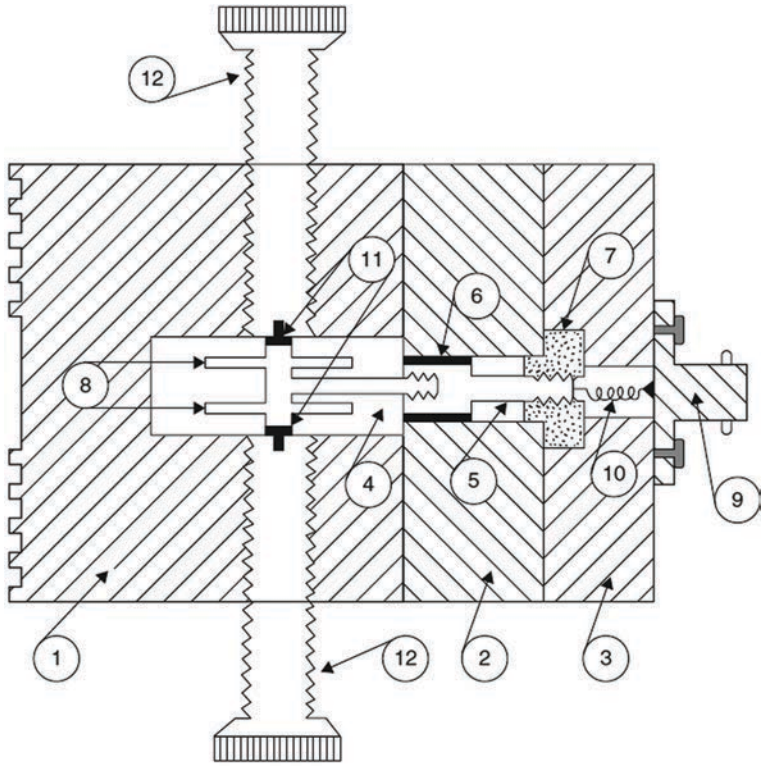


Figure 7.1 Rectangular waveguide resonant cavity combiner (Kurokawa–Magalhaes type): (a) Schematic view of the combiner, and (b) Sectional view of the combiner as seen from the top.



Parts list

(1,2,3) Copper blocks,
(4) Rectangular waveguide cavity,
(5) Bias filter,
(6) Insulating material,
(7) Insulating guard-ring
(8) Twin-cap structure,
(9) BNC connector,
(10) Gold spring,
(11) IMPATT diode,
(12) Threaded diode holder.

Figure 7.2 Twin-cap IMPATT power combiner: (a) Vertical cross-sectional view, and (b) Parts list.

extended similar to the scheme used by Kurokawa and Magalhaes [1] to design a multi-diode power combiner.

N-way in multiple stages may be realized in many ways. One common approach is to use short-slot hybrid or branch-line quadrature (90°) hybrid, basic principle of operation of short-slot hybrid, and branch-line quadrature hybrid are given in Appendix IV. The power combiner scheme in which four short-slot hybrids are used with a hybrid coupling adapter and four IMPATT sources (operating in reflection-mode amplifier) whose combined power will be delivered in the output load is shown in Figure 7.3.

The operation of such four-way power combiner using short-slot hybrid may be understood as follows. The input power fed to the port 1 of the short-slot hybrid marked QH I feed power to the through port 2 and coupled port 3 with the phase of the signal: 90° and 180° respectively, while no signal is ideally coupled to the isolated port 4. The signal from port 2 of QH I enters the input port 1 of second short-slot hybrid QH II and reaches the IMPATT amplifiers marked A1 and A2 with the phase of the signal: 180° and 270°, respectively. Assuming that these signals are reflected without experiencing any phase change from the amplifiers will reinforce (being of same phase) the signals at port 4 of QH II, but the reflected signals reaching back from the two amplifiers at port 1 of QH II will cancel out (being of phase opposition). Similarly, the signal from port 3 of QH I after entering the port 1 of third short-short hybrid QH III will reinforce signals reflected from the IMPATT amplifiers marked A3 and A4 and will not deliver back any signal to the port 1 of QH III (as shown). Finally, the signals from port 4 of QH II and QH III will give additive signals at port marked as 4 of the short-slot hybrid QH IV (where the output is taken) and will ideally cancel out at the port marked as 1 of QH IV. Thus, the combined power output of the four IMPATT sources A1, A2, A3, and A4 can be obtained with this short-slot hybrid power combiner. The hybrid combiner is suitable for cascading (in fact in Figure 7.3 it may be noted that the input is received from a preceding hybrid combiner); this way, the power output of many oscillators can be combined. In a hybrid combiner, the short-slot hybrid provides the necessary isolation between the active sources connected to its two output ports and hence the modeing problem (that is, chance of oscillating at multiple frequencies close-by), which the resonant cavity combiner is prone to, is minimized here. However, injection-locked sources need to be used in this case (when synchronization of frequency and phase of the two sources can be assured), but since the diodes are in close proximity with strong coupling between them, they are inherently locked in resonant combiners.

7.1.2.2 Wilkinson Power Combiner/Divider

Yet another type of microwave power combiner/divider is *Wilkinson power combiner/divider* [3] that was proposed by Ernest J Wilkinson in 1960.

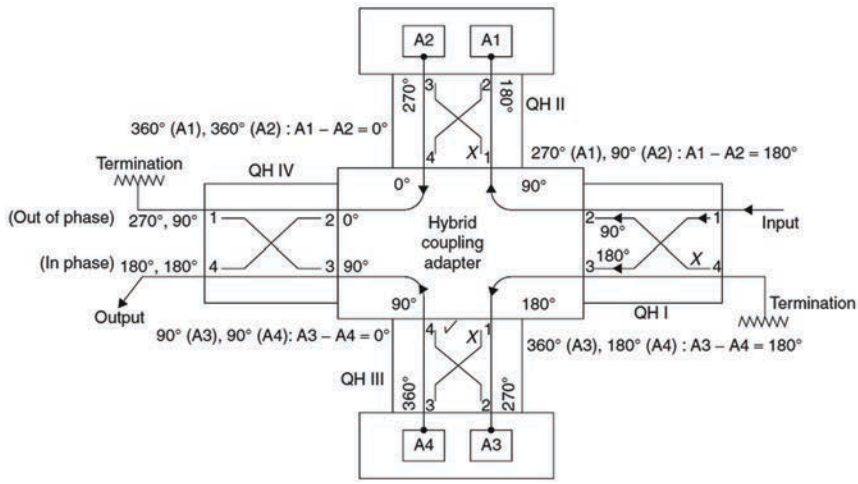


Figure 7.3 Power combiner using short-slot hybrids with phase relationship shown at different ports.

Wilkinson power divider is a specific class of power divider circuit that can achieve isolation between the output ports while maintaining a matched condition on all ports. The Wilkinson design can also be used as a power combiner because it is made up of passive components and hence is reciprocal. A Wilkinson power divider/combiner is a three-port network that divides the input power at port 1 equally between the two output ports 2 and 3 (which are also in phase). Essentially, it acts as a 3 dB power divider or coupler. Reversing the direction of flow of signal, that is, feeding inputs to ports 2 and 3 will result in a signal at port 1 which is the phasor sum of the input signals at ports 2 and 3; thus, making the Wilkinson junction to act also as a power combiner (which is in fact true for any 3 dB power divider junction, such as E-, H-plane tee and so on). A schematic diagram of a 3 dB Wilkinson power divider/combiner is shown in Figure 7.4.

It may be noted that the characteristic impedance of each of the $\lambda/4$ line should be $Z_{0(\lambda/4)} = \sqrt{Z_0.2Z_0} = \sqrt{2}Z_0$; as these two matched $\lambda/4$ lines appear in shunt so their combined input impedance (each $2Z_0$) must result in Z_0 of the input line. Further, it is assumed that the output ports 2 and 3 are perfectly matched, which seldom occurs. To take care of mismatch at any of the output ports, a resistor $R = 2Z_0$ is connected across the output arm; thus, when a mismatch occurs at one of the output ports, the reflected signal will get split equally between the $\lambda/4$ line and the resistor R [as the resistor $R = 2Z_0$ and input impedance ($2Z_0$) of each $\lambda/4$ line are equal]. Under this condition, the two signals reaching the input port will be in phase while to

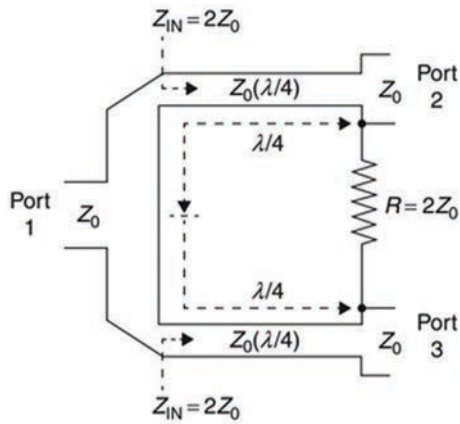


Figure 7.4 Schematic of Wilkinson power combiner/divider.

the other output port will see phase cancellation. The Wilkinson power divider/combiners are easily constructed in strip-line (and not in waveguides) because the impedance of various lines (which are different) can be easily adjusted by controlling the strip width (the characteristic impedance of a strip-line is inversely proportional to the strip width).

7.1.2.3 Spatial Power Combining

Spatial power combining (SPC) techniques have become one of the most reliable and efficient techniques of millimetre-wave power combination reaching high powers in small volumes with high fault tolerances. A common form of SPA is solid-state power amplifiers (SSPA) which are perfect candidates to replace travelling-wave tube (TWT) in space communication as the later need high bias voltage and being heavy in wait compared to the former. Exploiting the SPA technique, it is possible to combine the RF signals from a number of solid-state amplifiers to produce a single, stronger beam by physically arranging amplifier components in "space" (like an array) and combining their output energy; see Figure 7.5 (which shows a rudimentary scheme for an X-band SSPA transmitter). It consists of three-stage amplifications with GaAs and GaN power FET devices. RF input will be amplified by a pre-driver amplifier with 30 dBm of output power followed by driver amplifier. Two-way power divider and combiner topology has been used at the final stage. Sometimes focussing elements (lenses/mirrors) are used for beam-forming [4]. With such power combing technique 10s to 100s of watts of power can be realized with solid-state technology.

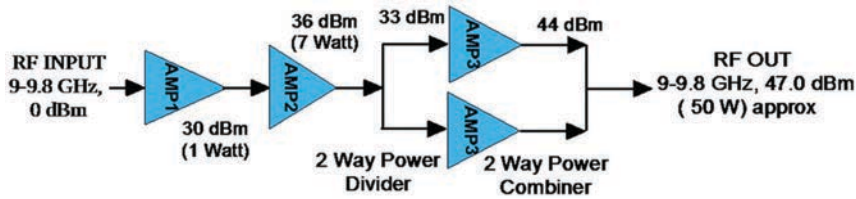


Figure 7.5 Block diagram of an X-band SSPA transmitter section.

7.1.3 Design of Microwave Power Combiners

We shall discuss here design of two types of RF/Microwave power combiners. One is the power combiner using short-slot hybrid while the second one is Wilkinson power combiner.

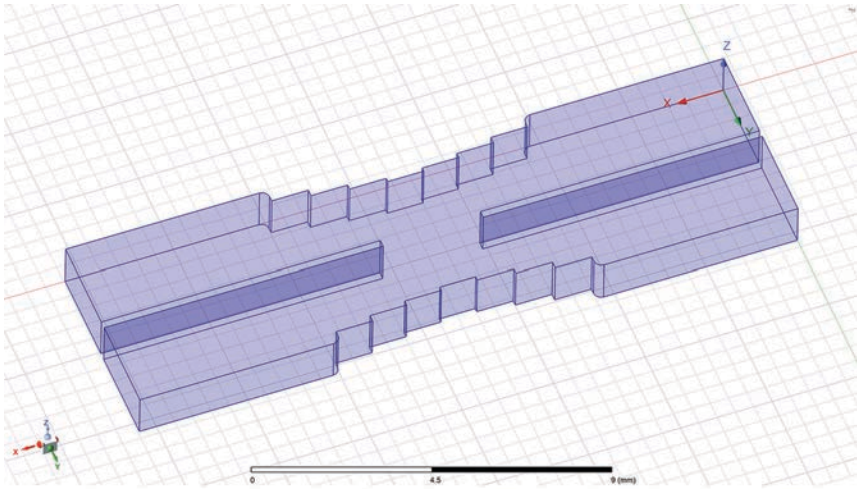
Example 7.1 Design and characterize power combiners/dividers designed with short-slot hybrid of two types.

(A) Two-way or 4-port short-slot hybrid.
(B) Four-way or 8-port short-slot hybrid.

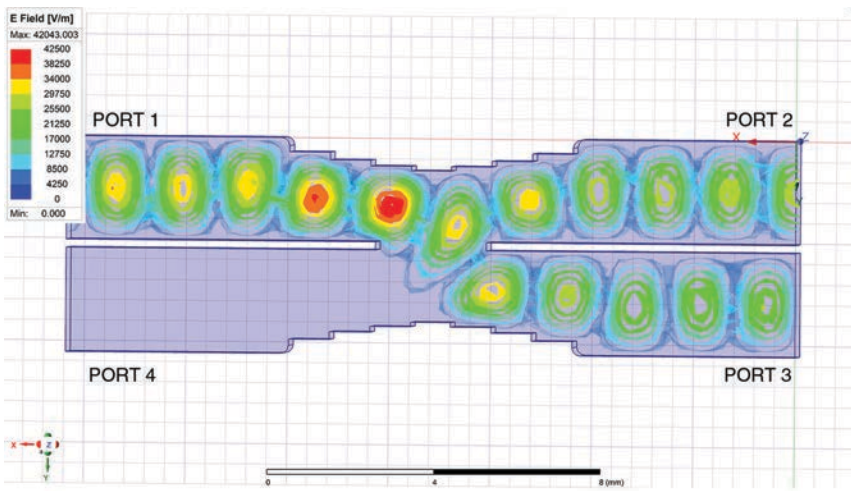
Frequency of operation is a 94 GHz (W-band). For simulation-based design use High Frequency Structure Simulator (HFSS) and Automatic Design System (ADS) software as and where/when necessary

Solution

(A) The structure of a two-way or 4-port short-slot hybrid is shown in Figure 7.6a together with the magnitude of the E-field in Figure 7.6b(i) and port details in Figure 7.6b(ii). The magnitude and angle of S-parameters for this structure are shown in Figure 7.6c and 7.6d respectively. To realize *power combining* action, input powers to be combined are applied at P_{thru} and P_{coup} (ports 2 and 3, respectively) and power gets added at P_{in} (port 1). To perform this, the *S* parameters (.s4p file) of the 3D structure simulated in HFSS is imported in ADS and time domain simulation is carried out. The schematic and the simulation result for the power combiner are shown in Figure 7.6e and 7.6f, respectively. To realize *power divider* action, input power is applied at P_{in} (port 1) and divided equally at P_{coup} and P_{thru} (ports 2 and 3, respectively). The schematic and the simulation result for the power divider are shown in Figure 7.6g and 7.6h, respectively.



(a)



(b)(i)

PORT DETAILS

- Port 1 : INPUT PORT
- Port 2 : THROUGH PORT
- Port 3 : COUPLED PORT
- Port 4 : ISOLATED PORT

(b)(ii)

Figure 7.6 The design and simulation-based characterization results for two-way or 4-port short-slot hybrid-based power combiner/divider: (a) Structure of a two-way or four-port short-slot hybrid, (b) (i) Magnitude of the E-field, and (ii) Port details, (c & d) Magnitude and angle of S-parameters for the structure, (e & f) Schematic and the simulation result for the power combiner, and (g & h) Schematic and the simulation result for the power divider.

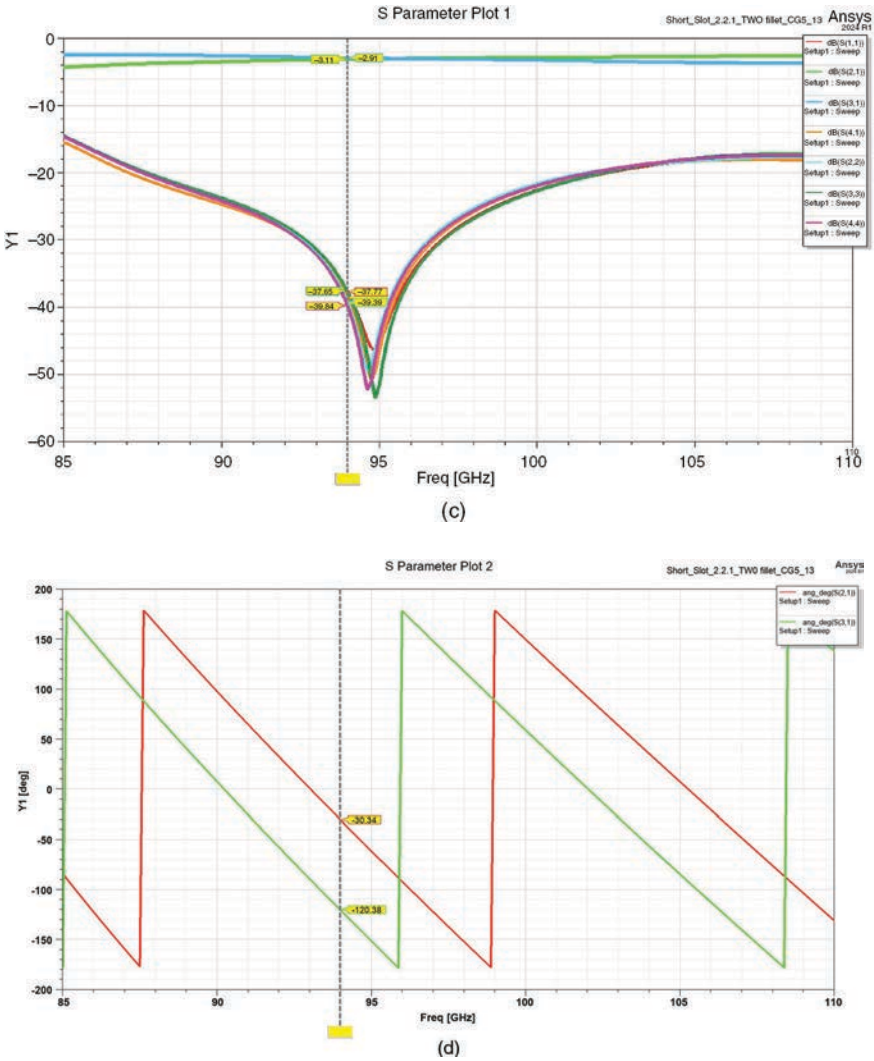
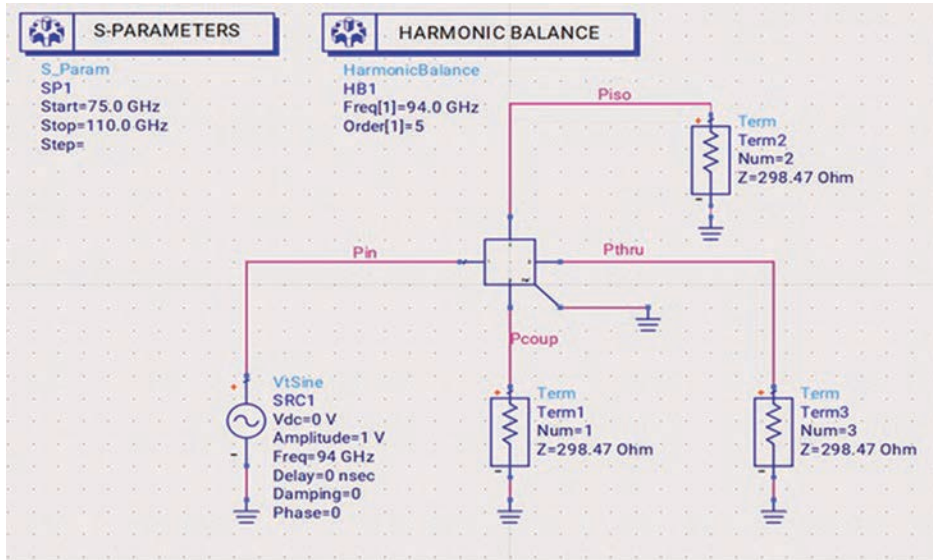


Figure 7.6 (Continued)



(e)

Figure 7.6 (Continued)

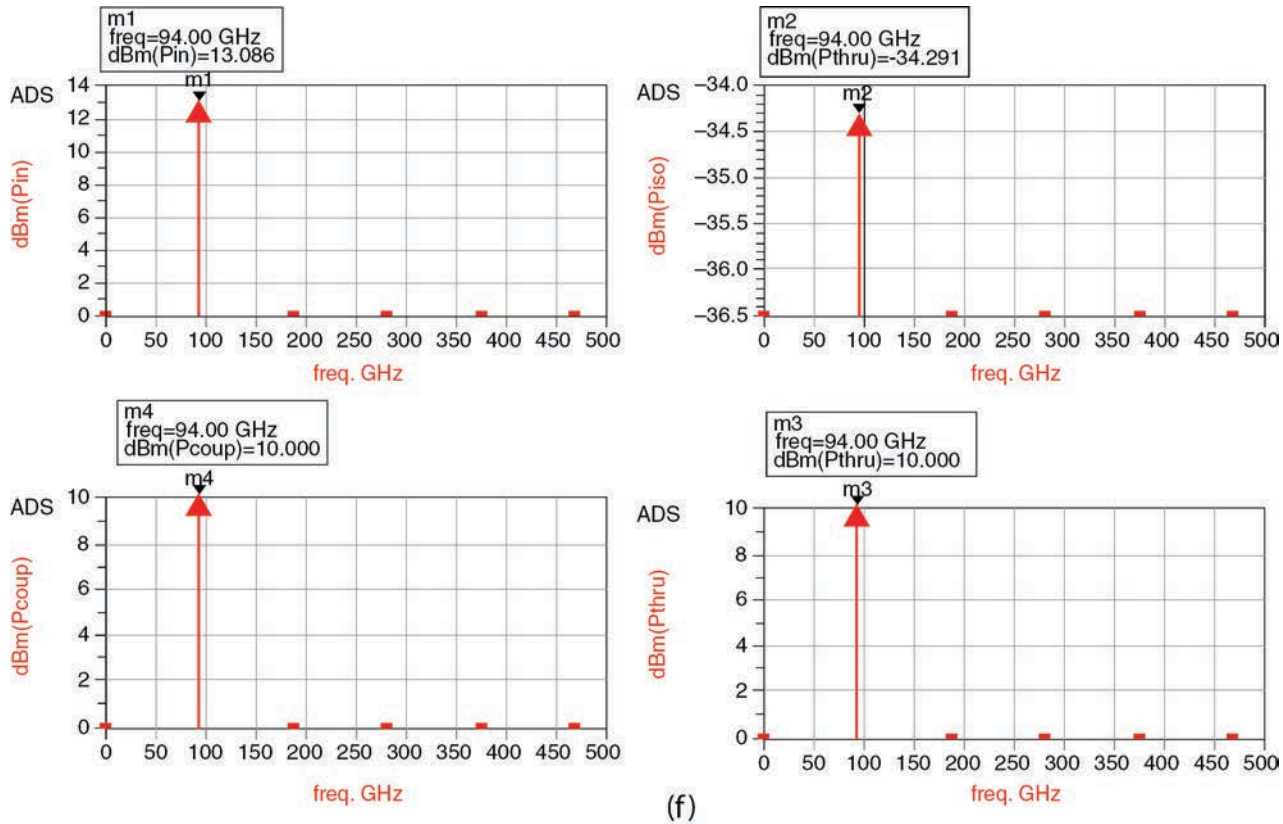


Figure 7.6 (Continued)

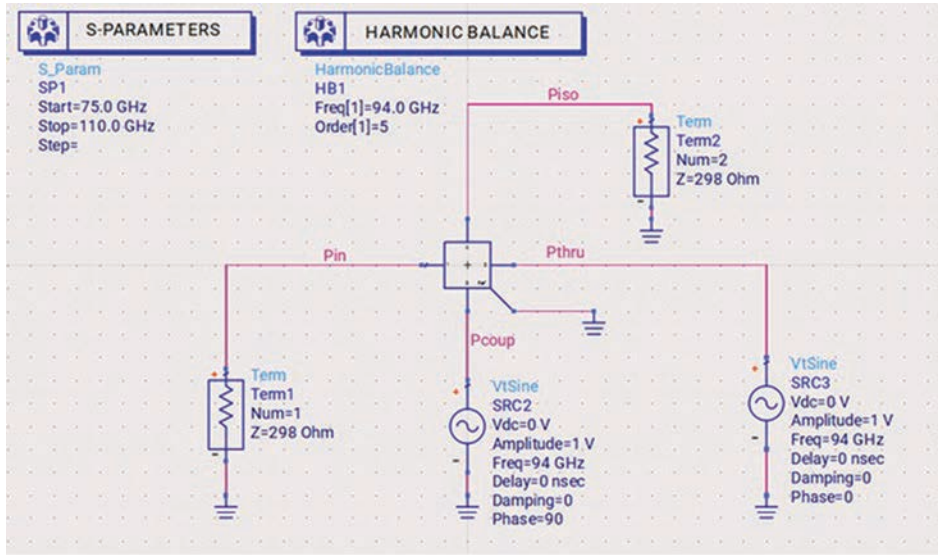


Figure 7.6 (Continued)

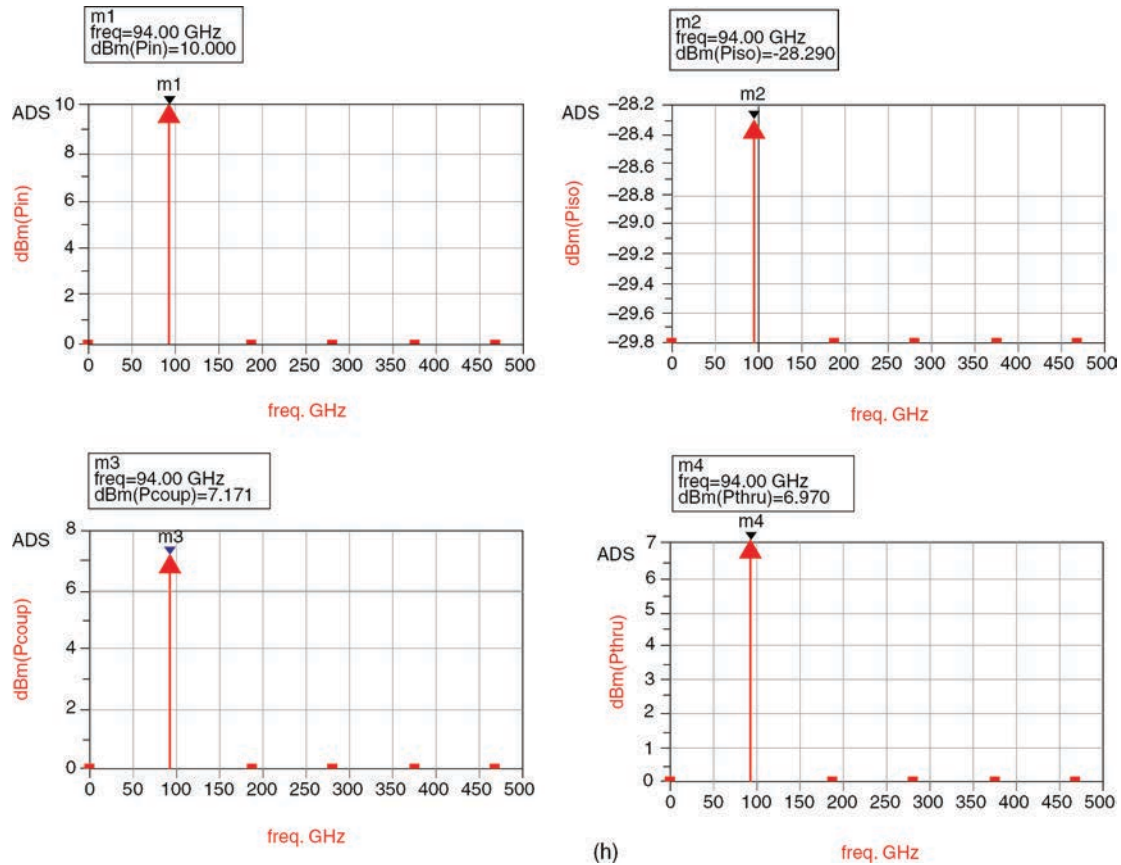
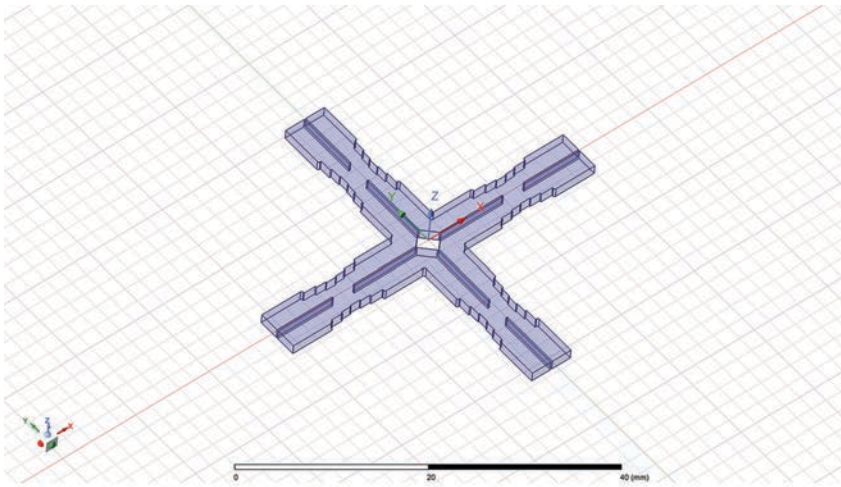
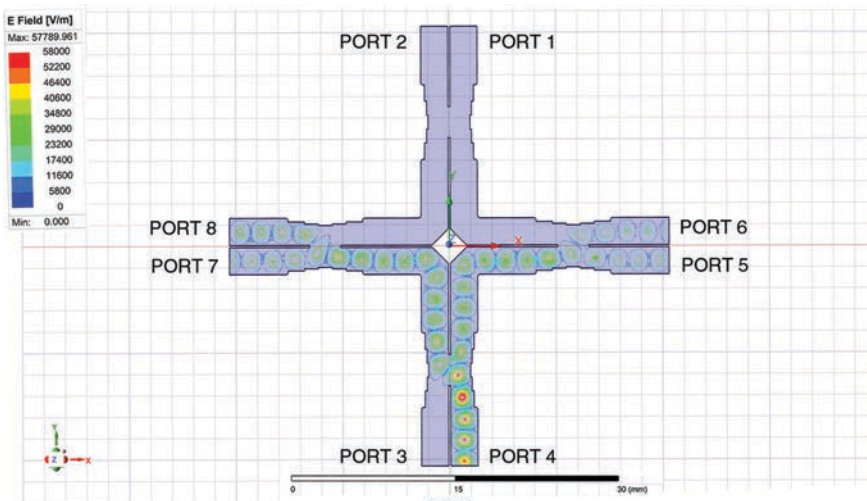


Figure 7.6 (Continued)

(B) The structure of a four-way or eight-port short-slot hybrid is shown in Figure 7.7a together with the magnitude of the E-field in Figure 7.7b(i), while the port designation with phase in the respective ports is shown in Figure 7.7b(ii). The S-parameter plot



(a)

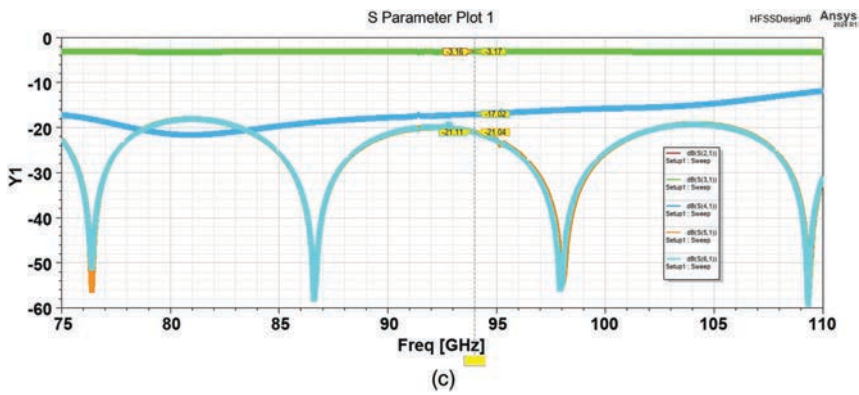


(b)(i)

Figure 7.7 The design and simulation-based characterization results for four-way or 8-port short-slot hybrid-based power combiner: (a) Structure of a four-way or 8-port short-slot hybrid, (b) (i) Magnitude of the E-field, and (ii) Port designation with phase in the respective ports, (c) S-parameter plot of the 8-port short-slot hybrid, (d) Schematic of 8-port short-slot hybrid, and (e) Simulation result: (i) power-combined output, and (ii) Input applied power.

Source (PORT)	Magnitude	Phase
1:1	0W	0deg
2:1	0W	0deg
3:1	0W	0deg
4:1	0W	0deg
5:1	1W	270deg
6:1	1W	0deg
7:1	1W	90deg
8:1	1W	0deg

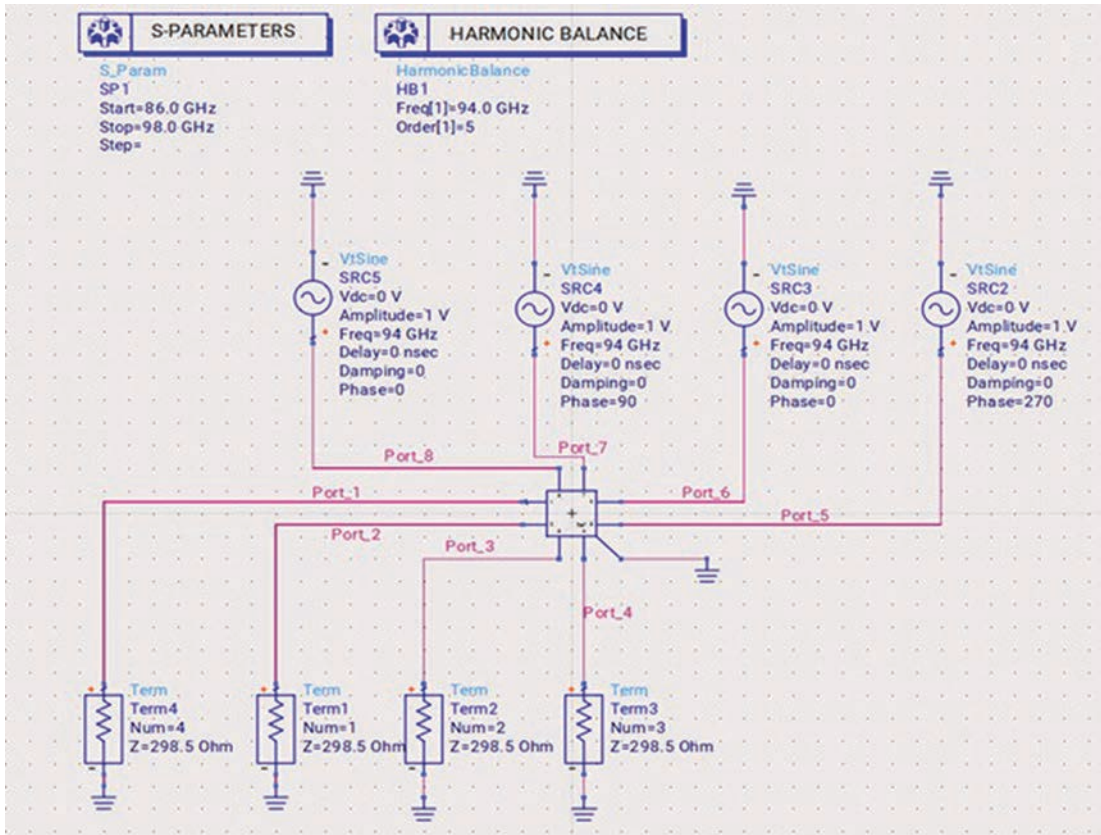
(b)(ii)



(c)

Figure 7.7 (Continued)

of the eight-port short-slot hybrid is shown in Figure 7.7c. To realize *power combining* action, input power (+10 dBm or 10mW) is applied at port 5 (phase 270 or -90), port 6 (phase 0), port 7 (phase 90), and port 8 (phase 0), all phase being in degrees, while the combined power output observed at port 1 and port 4 are +15.294 dBm (or 36.4 mW) and +15.614 dBm (or 36.4 mW), respectively. To perform this, the S parameters (.s8p file) of the 3D structure simulated in HFSS is imported in ADS and time domain simulation is carried out. The schematic of eight-port short-slot hybrid are shown in Figure 7.7d, while the simulation result is



(d)

Figure 7.7 (Continued)

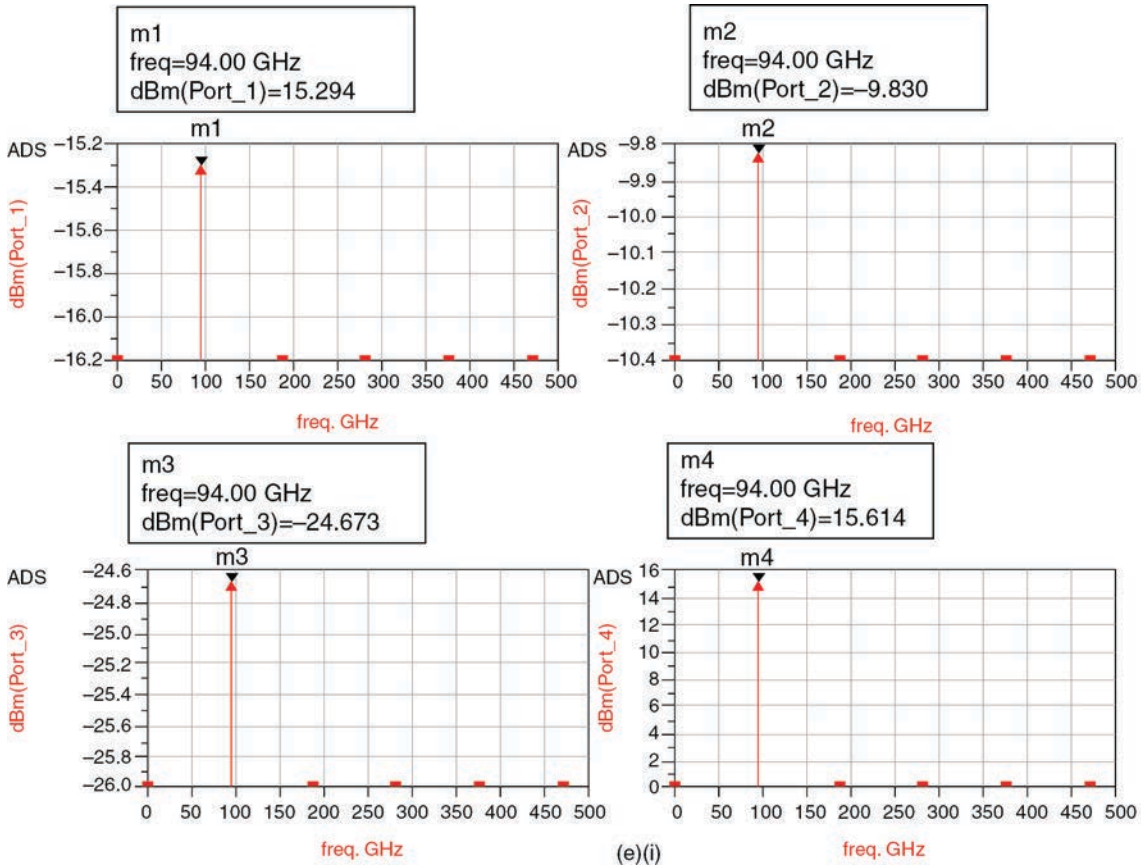


Figure 7.7 (Continued)

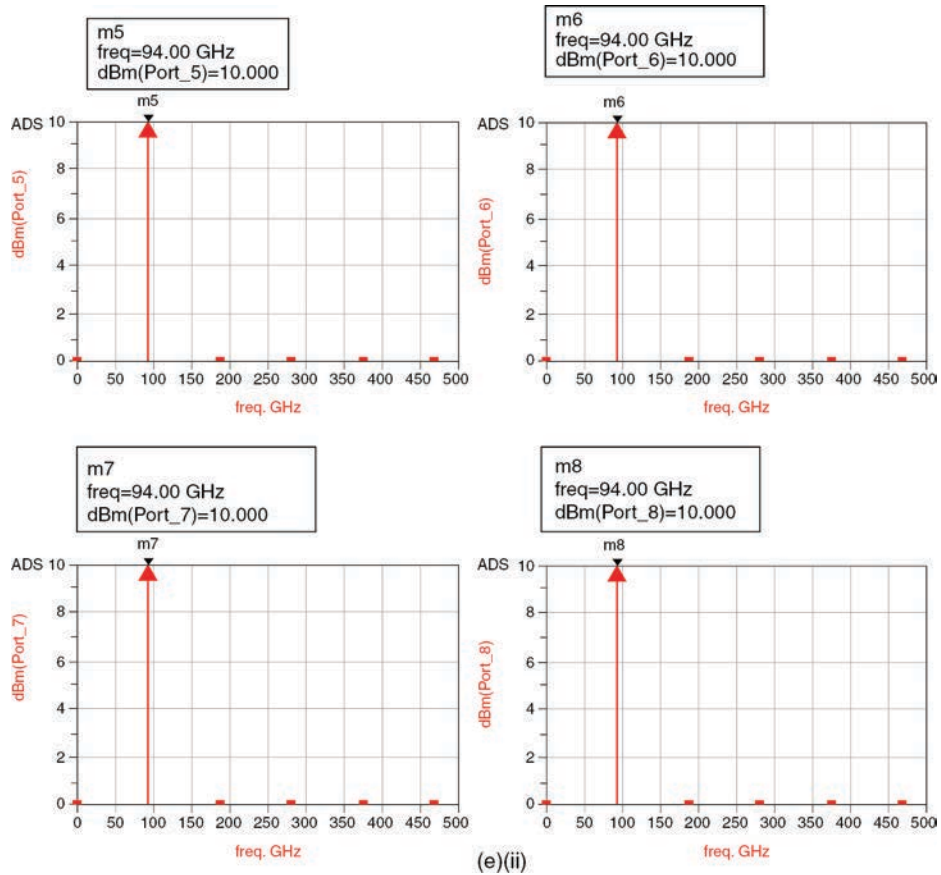


Figure 7.7 (Continued)

shown in Figure 7.6e; in the latter figure, the power-combined output is shown in Figure 7.e(i) and input applied power is shown in Figure 7.7e(ii).

Example 7.2 Design a 4-way Wilkinson power combiner at 10 GHz with input power of 10 dBm (i.e., 10 mW). Use FR4 as substrate material (with dielectric constant ε_r = 4.6) for the microstrip line. Perform the design with Automatic Design System (ADS) software.

Solution:

In this solution, a two-way and a four-way Wilkinson power combiner is presented with uniform impedance line. The Wilkinson power combiner we are going to design uses all of the transmission lines at input with characteristic impedance $Z_{0(in)}$ = 50 Ω and output transmission line with characteristic impedance $Z_{0(out)}$ = 70.7 Ω (= $Z_{0(\lambda/4)} = \sqrt{Z_0.2Z_0} = \sqrt{2}Z_0$).

For this combiner, the microstrip line is designed with FR4 as substrate material with dielectric constant ε_r = 4.6. The height (*h*) of the substrate of microstrip line is 30 mil and thickness (*t*) of the microstrip line is 0.015 mm; see Figure 1.8f in Chapter 1. The width (W) of the strip = 1.412820 mm, and length (L) = 3.970530 mm with $Z_{0(in)}$ being set to 50 Ω in the input line. While W = 0.732511 mm, and L = 4.098360 mm with $Z_{0(out)}$ being set to 70.7 Ω in the output line. The calculations for the microstrip line can be done either by using LineCalc or by using formulations available in standard text books [1]

The circuit schematic of a two-way Wilkinson power combiner (WPC) is shown in Figure 7.8.

Simulated plot for return loss [Return loss (dB) = -20 log $|S_{11}|$], insertion loss [Insertion loss (dB) = -20 log $|S_{21}|$, isolation [Isolation (dB) = -20 log $|S_{32}|$], and the phase difference between two output ports are shown in Figure 7.9.

The layout of the two-way Wilkinson power combiner is shown in Figure 7.10. In the layout, we used a resistor pad of value 100Ω that has length = 1.412820 mm and width = 1 mm.

The simulated plot of two-way power combiner is shown in Figure 7.11. In each input port 10 mW (10 dBm) of power is applied, and at the final stage (i.e., with 2-way combining) 20 mW (13 dBm) is obtained.

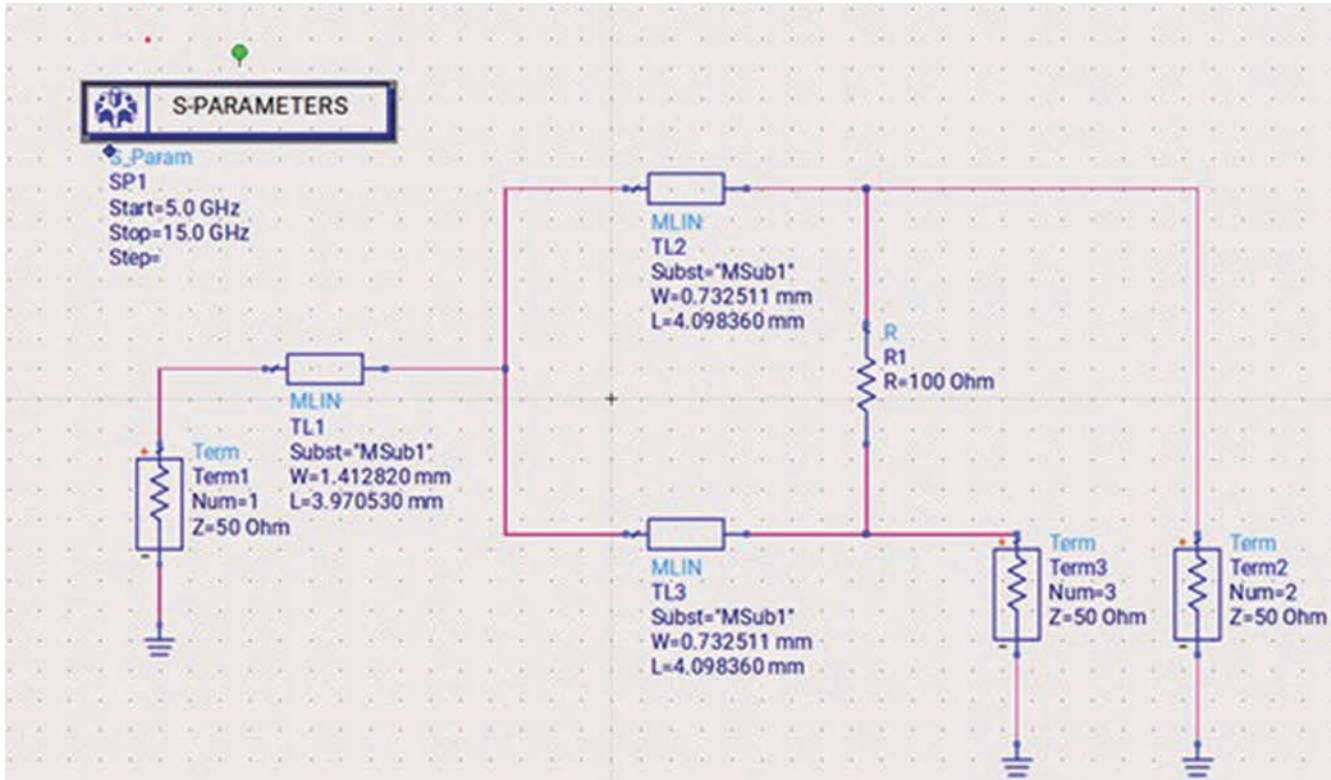


Figure 7.8 Circuit schematic of a two-way Wilkinson power combiner.

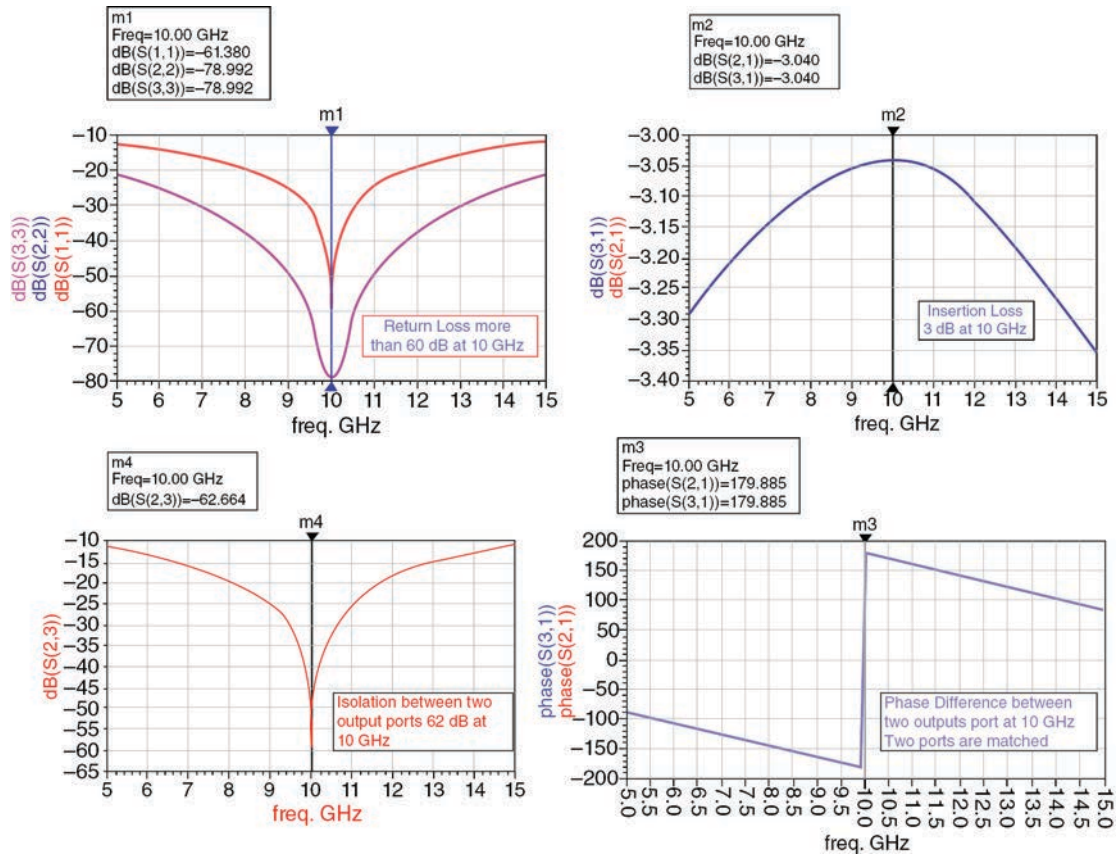


Figure 7.9 Simulation plot of two-way power combiner (Schematic Result).

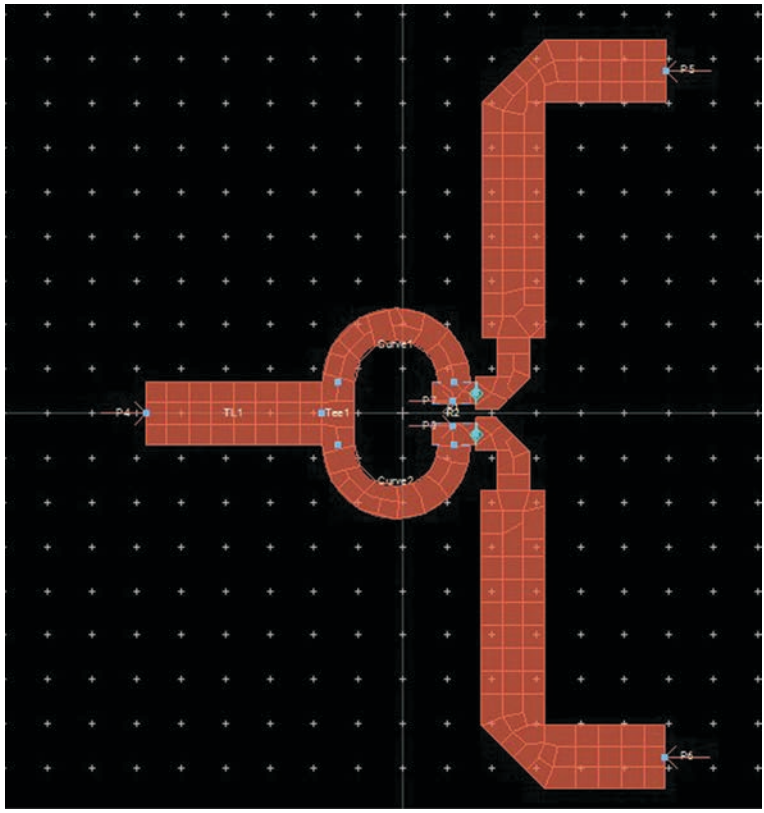


Figure 7.10 ADS-simulated layout of two-way Wilkinson power combiner.

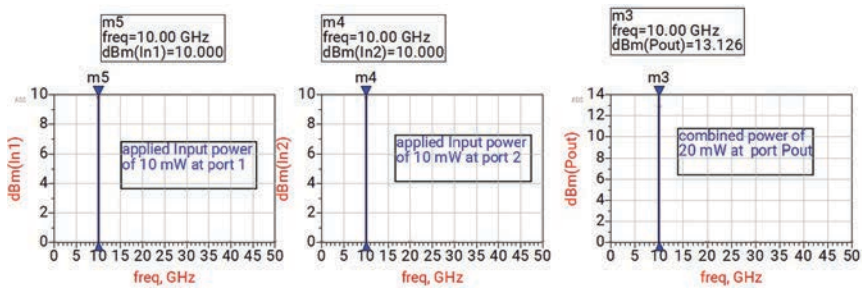


Figure 7.11 Simulation plot of two-way Wilkinson power combiner.

Four-way Wilkinson power combiner is an extension of the twoway WPC. The values of W and L are the same. We need to optimize the length of the branched lines for the expected results. 100 Ω resistors are used for the isolation from each other and to make the output port well matched.

The layout of 4:1 (i.e. 4way) power combiner obtained with ADS simulation is shown in Figure 7.12.

The simulated plot of four-way power combiner is shown in Figure 7.13. In each input port 10 mW (10 dBm) of power is applied, at the first stage (i.e., with two-way combining) 20 mW (13 dBm) is obtained and at the final stage with four-way combining power at the output port, P_{out}, is 40 mw (16 dBm), see Figure 7.13.

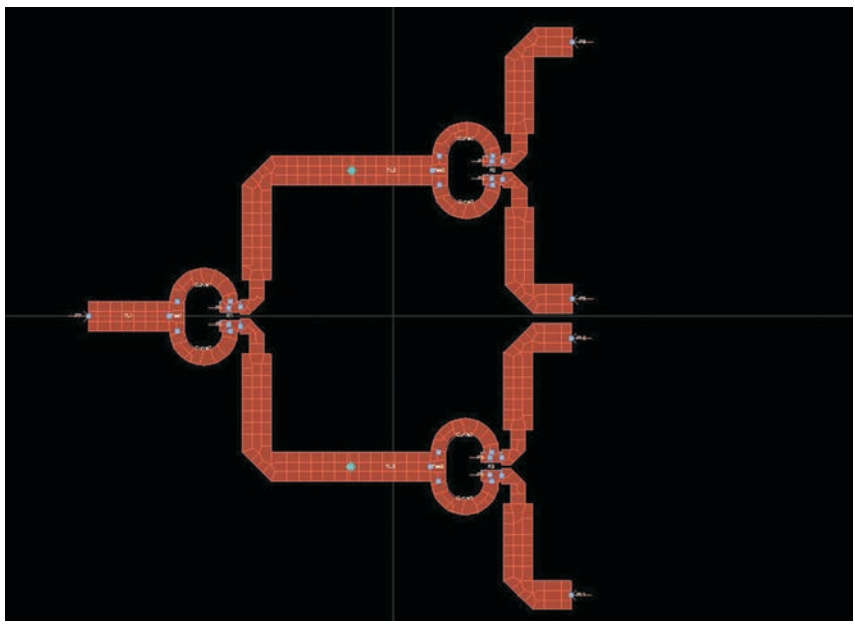


Figure 7.12 ADS-simulated layout of 4:1 power combiner.

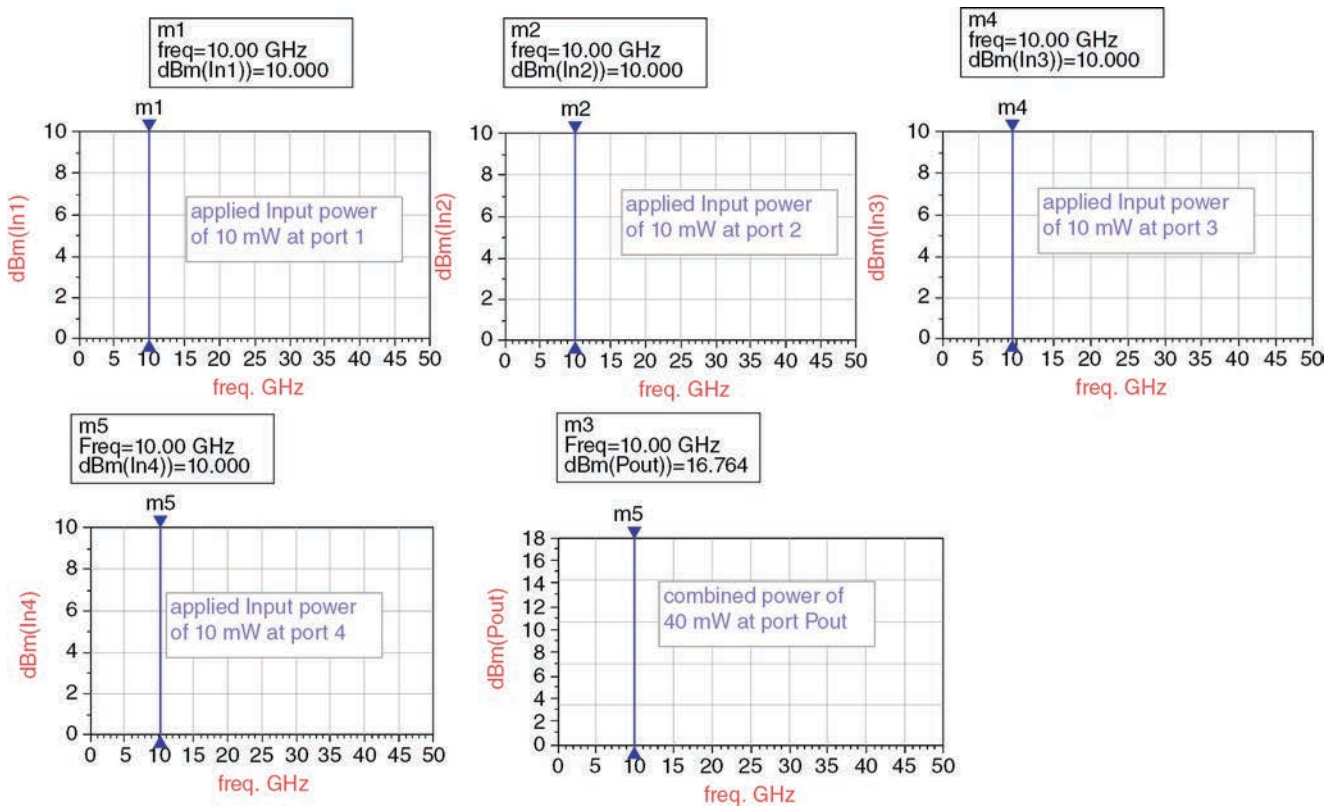


Figure 7.13 Simulation plot of four-way Wilkinson power combiner.

7.2 MICROWAVE MIXERS

7.2.1 Introduction

Mixers are used in communication receivers to achieve frequency conversion (down/up) of an input RF signal; usually to translate down the carrier frequency of the modulated RF signal to an intermediate frequency (IF). A nonlinear device, most commonly a diode or a transistor, is used to realize this frequency translation in which the input RF signal and the output IF signal are linearly related with each other. The nonlinear device generates the necessary integer multipliers, or harmonics, of the signals fed to the device (the RF signal and a local oscillator signal) that are necessary for the process of frequency conversion/translation. The general equation representing the process of frequency conversion and also other related aspects of a mixer is given by

$$f_{\text{mn}} = mf_s + nf_p \tag{7.1}$$

where f_{mn} represents the frequency converted (down/up) signal coming out of the mixer; for down conversion ($f_s - f_p$), it is generally known as the *intermediate frequency* (IF) signal. Here, f_s is the frequency of input RF signal (f_{RF}) and f_p is frequency of the driving signal (also known as *pump signal*) or *local oscillator* (LO) signal (f_{LO}). It may be noted that m and n are integers denoting the harmonics of the input RF signal and the LO signal, respectively; m, n = –2, –1, 0, +1, +2, … For conventional or *fundamental mixers*, $|m| = 1$ and $|n| = -1$. Thus, the conventional mixers use the fundamental frequencies of the RF and LO signal to perform the frequency conversion, and these mixers require an LO signal with a frequency of the same order as the frequency of the RF signal; although, in general, it is the convention to choose $f_{\text{LO}} > f_{\text{RF}}$. The fundamental mixing product with $m = 1$ and $n = -1$ is, in most cases, the desired *intermediate frequency* (IF), while the second order mixing product with $m = -1$ and $n = 2$ is termed as the *image frequency*.

In a mixer, there is always a possibility of getting *image frequency* being generated which will reach the IF port of the mixer as an undesired/confusing signal if the RF pass-band of the receiver is not chosen carefully. It may be noted that the desired IF frequency is $f_{\text{IF}} = (f_{\text{LO}} - f_{\text{RF}})$, while the image frequency is an RF signal given by $f_{\text{IM}} = (-f_{\text{RF}} + 2f_{\text{LO}}) = (f_{\text{LO}} + f_{\text{IF}}) = (f_{\text{RF}} + 2f_{\text{IF}})$. Thus, if an RF signal equal to the image frequency is received at the input of the receiver, it will produce after mixing with LO signal in the mixer an IF which is just the mirror image of the desired IF with respect to the desired RF signal and these two IF signals will be indistinguishable (since the Fourier spectrum of any real signal is symmetric about zero frequency

and thus contains negative as well as positive frequencies). Unfortunately, this spurious IF signal (originated due to image frequency) will also be amplified by the IF stage and therefore will behave as an interference as this has the effect of two stations being received simultaneously giving rise to the so-called "*double-spotting*" especially in the short-wave frequency band.

Hence, in the receiver design, care should be taken in the RF stages of the receiver to pre-select signals only within the desired RF frequency band. Otherwise, the RF signal at image frequency will be indistinguishable from the desired RF signal at the IF stage. It may be noted that the choice of which RF frequency is the desired one and which one is image frequency is arbitrary, depending on whether the LO frequency is above or below the desired RF frequency. Most practical receivers, however, prefer to choose $f_{\mathrm{LO}} = (f_{\mathrm{RF}} + f_{\mathrm{IF}})$, as mentioned above, because this demands a lower LO tuning ratio when the receiver must select RF signals over a given band.

A mixer is theoretically a linear device, translating a signal from one frequency to another mostly with a nonlinear device (though in recent times a time-varying linear resistor realizable with FET device is also used as mixer element). The presence of a nonlinear device in mixer, to which two signal frequencies f_s and f_p are being fed to generate the f_{IF}, there is every possibility of the production of the so-called *intermodulation products*. A linear time-varying element used for mixer operation will, however, practically have no such distortion; as in resistive FET mixer. Anyway, any receiver that has a nonlinear device-based mixer suffers from intermodulation distortion (IMD); typical values of IIP_3 (input power level for third-order intercept point) of mixers ranges from 15 to 30 dBm. Its contribution can be minimized by using minimal gain in the low-noise RF amplifiers preceding the mixer. Details of IMD and its effect in mixer performance will be discussed in section 9.3.3.3 of Chapter 9; where it will be seen that third-order IMD is a serious issue of concern rather than the second-order IMD. Hence, special consideration has to be taken in the design of RF/microwave receiver system to get-rid-of/minimize the effect of third-order intermodulation distortion.

A mixer may also be viewed as a multiplier, and in fact, it is denoted by a multiplication sign (Figure 7.14). When two sinusoidal signals: $A(t)\cos 2\pi f_{RF}t$ (the input RF signal) and $A(t)\cos 2\pi f_{LO}t$ (the LO signal) are applied to an ideal analogue multiplier (a mixer). The output is found to consist of modulated components at the sum and difference frequencies. The sum frequency can be rejected by the filter tuned to the desired frequency, leaving the IF, the difference frequency.

In addition to frequency conversion in communication receivers, microwave mixers are also used in microwave instrumentation (vector network analyzer) to translate down the microwave signal to a suitable lower radio frequency where it is more convenient for amplification and signal

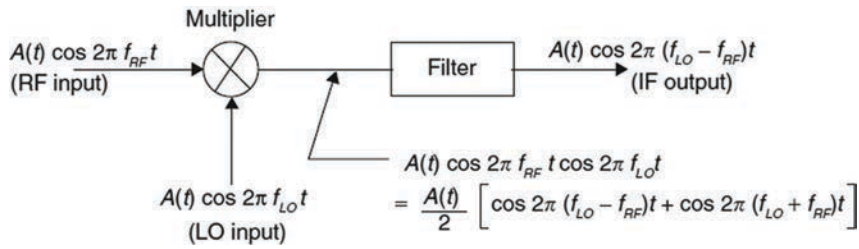


Figure 7.14 Schematic of a mixer which is fundamentally a multiplier.

processing for deriving the amplitude and phase information of the input microwave signal. Further, the mixers find application in phase detectors and demodulators too in which such functions are realized with minimal addition of noise and distortion.

While discussing mixers, we need to talk about *conversion loss* in the case of diode-based mixers and *conversion gain* in the case of transistor-based mixers. An important figure of merit of a mixer is the conversion loss (L_c) which is defined as the ratio of available RF input power to the available IF output power expressed in decibels.

$$L_c = 10\log_{10}[(\text{available RF input power})/(\text{available IF output power})] \geq 0 \text{ dB} \quad (7.2a)$$

$$\text{In other words,} \quad L_c[\text{dB}] = P_{RF}[\text{dB}] - P_{IF}[\text{dB}] \quad (7.2b)$$

Conversion loss accounts for resistive losses in a mixer as well as loss in the frequency conversion process from an RF signal to an IF signal. The conversion loss is usually specified for a specific frequency and for a specific LO power level. Practical diode mixers have conversion losses between 6 dB and 8 dB in the frequency range (1–14) GHz. Transistor mixers have lower conversion loss, and their major advantage over diode mixers is their ability to provide a few decibels of conversion gain.

The last but not the least important point for mixer design is that the RF and LO ports in a mixer have to be properly isolated. Otherwise, some of the power from the LO port will be unduly radiated via the antenna in the RF port. The isolation between the LO and RF port is highly dependent on the type of coupler being used for diplexing these two input ports with typical values ranging from 20 dB to 40 dB.

There are various types of microwave mixers based on diode and transistors including the harmonic mixer and image-reject mixer. Let us discuss the physical principle of some of those mixers [5].

7.2.2 Different Types of Microwave Mixers

7.2.2.1 Single-Device/Single-Ended Diode Mixer

Such mixers provide the simplest way of frequency conversion, and the analysis and synthesis of the single-ended diode mixers form the building blocks for most multi-diode mixers such as single-balanced and double-balanced mixers. The simplicity and minimal components make the single-device mixer the only truly effective mixer at millimetre-wave frequency; however, at lower microwave frequencies, they are outperformed by other improved configurations. The schematic of a single-ended diode mixer is shown in Figure 7.15 where the RF and LO inputs are fed to the diode via a diplexer which may be a hybrid junction to provide signal combining (RF + LO) and also to provide isolation between these two signal ports. The diode is biased from the DC source which needs to be decoupled from the RF signal paths and is realized by incorporating DC blocking capacitors on both sides of the diode and an RF choke between the diode and the DC voltage source. For optimal design of the mixer, matching networks need to be provided for RF/LO source impedance that matches the diode's input impedance and IF load impedance that matches the diode's output impedance. The AC output of the diode is passed through a low-pass filter to provide the desired IF output signal. The spectrum of the down converting single-ended mixer is shown in Figure 7.16.

7.2.2.2 Single-Balanced and Double-Balanced Mixers

The diode-based balanced mixers have an advantage over the single-ended mixers in terms of better isolation between the RF and LO ports, LO noise cancellation at IF port, better power handling capability including the rejection of all even order intermodulation products. There are, in general, two types of balanced mixers—single-balanced mixer based on 180° hybrid (magic tee/rat-race) and 90° (quadrature) hybrid, and double-balanced mixers, which use diodes in a ring or star configuration. Another

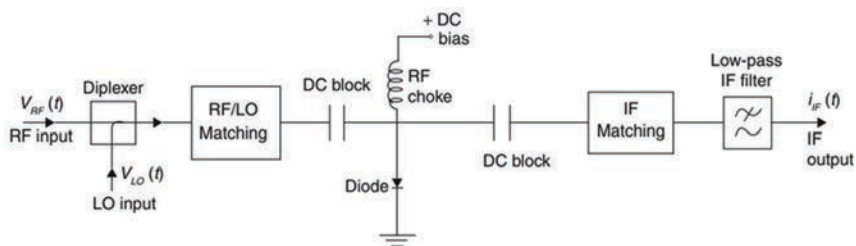


Figure 7.15 The schematic arrangement for a single-ended diode mixer.

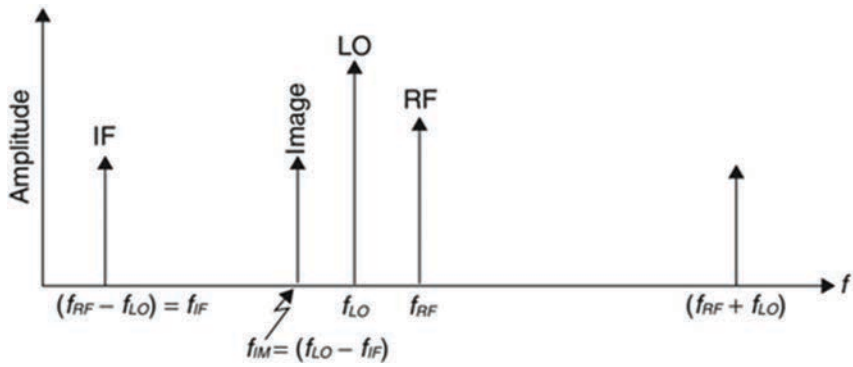


Figure 7.16 The spectrum of the down converting single-ended mixer.

balanced-type mixer, which consists of antiparallel diode pair is the sub-harmonically pumped mixer that uses an LO at half the desired LO frequency and the second harmonic of LO signal generated within the system mix with the RF input signal; the latter may be at a higher millimeter-wave frequency.

The *single-balanced diode mixer designed with* 180° *hybrid* (magic tee) is shown in Figure 7.17. Microwave crystal diodes/Schottky diodes are located at the two ends of the collinear arms of the magic tee. The local oscillator (LO) signal is fed into the H-plane arm and will arrive at the diodes in phase. The input microwave (RF) signal is fed into the E-plane arm and will arrive at the two diodes in anti-phase. The IF signals (the difference frequency signal produced by the mixing of the LO and RF signals) generated in the two diodes will be 180° out of phase. However, the LO noise will be in phase at the diodes. By feeding the diode outputs into a balanced IF input (push–pull arrangement), the local oscillator noise is seen to be cancelled out but the IF signals add-up in phase giving the full strength of IF signal to the IF amplifier input. The use of such a balanced microwave mixer using magic tee and push–pull arrangement at the IF port helps to separate the RF and LO ports so that LO noise is not unduly radiated through the antenna at the RF port (as E and H arm are cross-coupled with identical termination of the collinear arms) and also noise of the LO is not allowed to reach the IF port.

A single-balanced diode mixer designed with 90° *(quadrature) hybrid* is shown in Figure 7.18. Such a balanced mixer basically consists of two single-diode mixers connected to two mutually isolated ports of the hybrid, while the other two ports (also mutually isolated) are used for the RF and LO signal inputs. Typically, the IF outputs from the two mixers are connected in parallel, although they sometimes are combined by another hybrid.

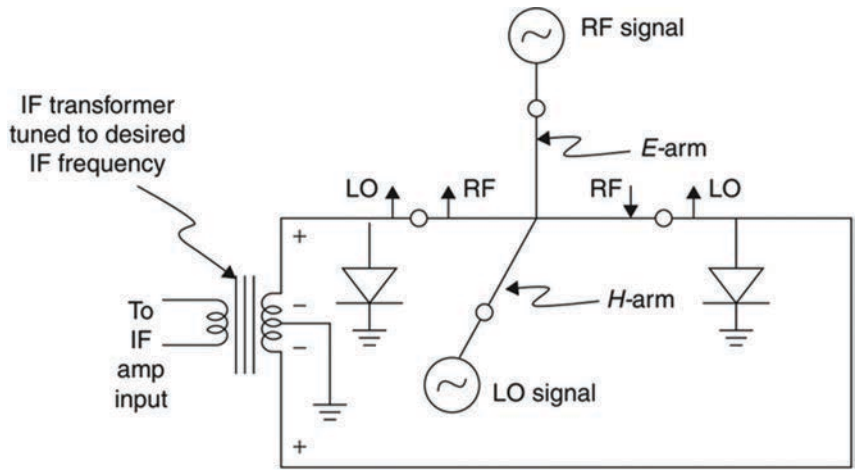


Figure 7.17 Single-balanced diode mixer using hybrid tee.

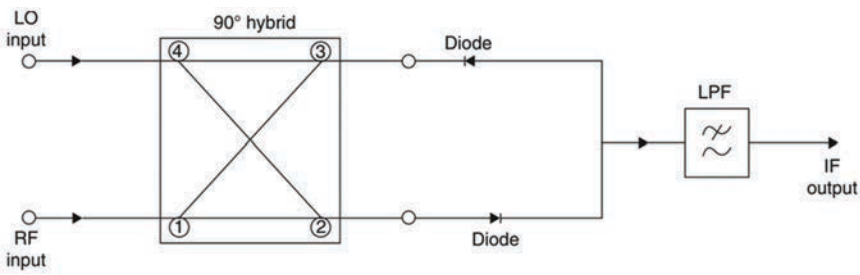


Figure 7.18 Single-balanced diode mixer designed with 90° (quadrature) hybrid.

The *double-balanced diode mixer* uses four diodes either in the ring or star configuration and is the most common type of mixer used in the frequency range 1–18 GHz. Though such mixers need more diodes and hybrids, greater LO power, and have higher conversion loss; they have better isolation between all ports, rejection of LO noise and spurious signals including certain intermodulation products with broadband operation. The schematic arrangement for a *double balanced ring-type diode mixer* is shown in Figure 7.19. It consists of two hybrid transformers/BALUNs (BALanced to UNbalanced unit) and a ring of four identical diodes. The RF signal is applied to the primary of one transformer and the LO is applied to the primary of the other. The center-tap of the LO transformer's secondary is grounded while the center-tap of the RF secondary serves as the IF output.

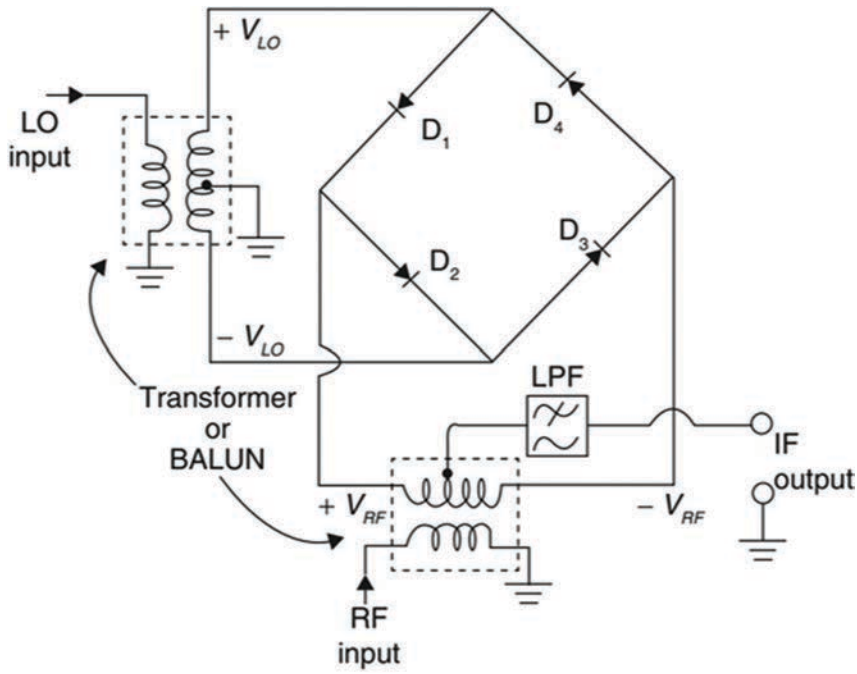


Figure 7.19 Schematic of double-balanced ring-type diode mixer.

7.2.2.3 Sub-harmonically Pumped Mixer

Sub-harmonically pumped mixer (SHPM) finds application at higher millimetre-wave frequency where a fundamental frequency LO becomes expensive, noisy, and the power level is inadequate. In such cases, it may be wise to use a mixer that uses a pump or LO signal at half the desired frequency and whose second harmonic, called the "virtual LO", generated by the mixer itself (resulted by the diode's junction conductance waveform), is mixed with the input RF signal. The SHPM has a little higher conversion loss but possesses better immunity to second order intermodulation distortion compared to other mixer configurations. Figure 7.20 shows a sub-harmonically pumped mixer using an antiparallel diode pair. If the diodes are identical, there is no fundamental mixing response. The nonlinearity of the back-to-back diodes acts as a resistive frequency multiplier (doubler) thereby converting the input LO signal to its second harmonic (2LO) which eventually mixes with the input RF signal to generate the desired IF signal. The RF and LO voltages are applied to the diode pair and filters tuned to indicated frequency are used to separate the different frequencies. It may be mentioned that harmonic mixers, in general, essential

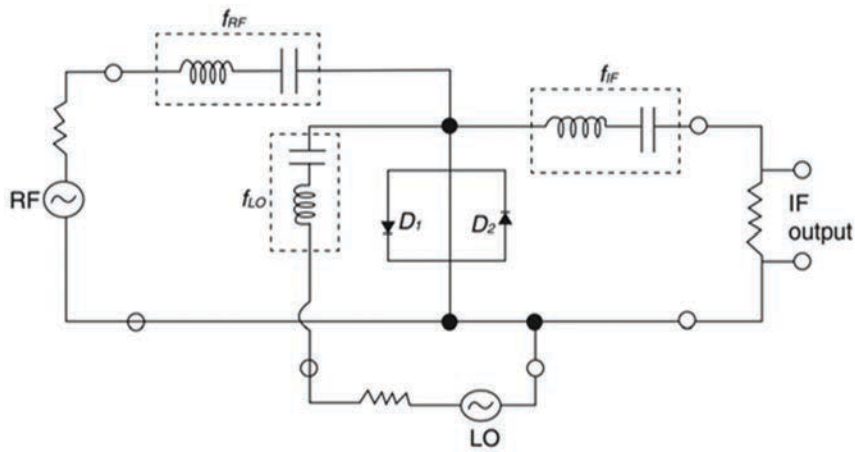


Figure 7.20 Sub-harmonically pumped mixer using an antiparallel diode pair.

for higher millimeter-wave and submillimeter-wave frequencies, resemble the sub-harmonically pumped mixer but differ in the sense that it employs higher multiples (greater than the second) of the injected LO signal.

7.2.2.4 Image-Rejection Mixer

If the input RF signal fed to the RF stage in the receiver is not pre-selected within the desired RF frequency band, then two distinct RF input signals at frequencies $f_{RF} = f_{LO} \pm f_{IF}$ will be down converted to the same IF frequency (in magnitude) when mixed with the LO signal f_{LO}. The desired IF will be either the lower or the upper sideband (–IF or + IF), respectively, of this DSB (double sideband) signal depending on $f_{LO} > f_{RF}$ or $f_{LO} < f_{RF}$; the other one being due to the image frequency. If the IF frequency band happens to be small, then the image signal giving a "confusing IF" will be very close to the RF or LO signal frequency. In such cases, it would become difficult to filter out the image at the IF stage; thus, the necessity of image-rejection mixer. The image-rejection mixer will help to isolate the two sidebands (one due to the desired IF and another due to the image) of the RF. It consists of two double-balanced mixers connected to quadrature (90°) hybrids at their inputs and outputs (see Figure 7.21). The LO is split via a conventional power divider and applied in phase to the two mixers. The RF signal above the LO frequency, the USB (upper side band) appears at one output of the IF hybrid, while the LSB (lower side band) appears at the other output. The image-rejection mixer described here does not achieve any performance improvement as the energy of the image is not "recycled". However, it

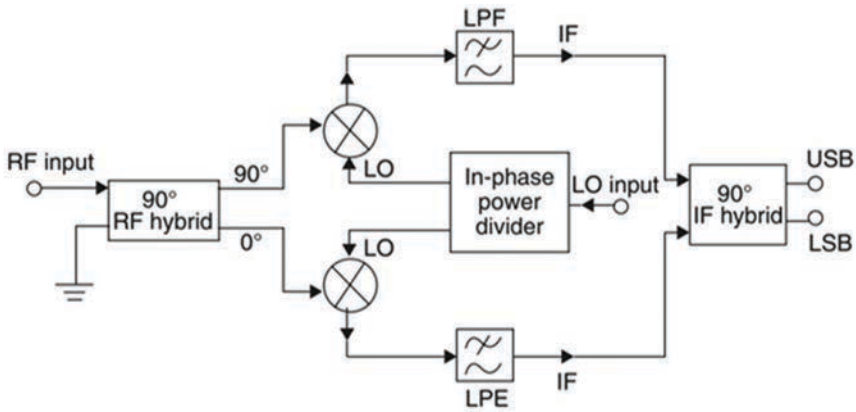


Figure 7.21 Schematic of image-rejection mixer.

is possible to design image-rejection mixers with such recycling leading to image enhancement; such mixers are often called "image-recovery mixers".

Example 7.3 A digital cellular telephone system using IS-136 TDMA technology has a receiver frequency band of 1975 MHz to 2035 MHz for 1900 MHz base station with a first IF frequency of 193 MHz and the specified channel bandwidth of 30 kHz. What are the two possible ranges for the LO frequency? If the upper LO frequency range is used determine the image frequency range. Does the image frequency fall within the receiver frequency band?

Solution:

The two possible LO frequency ranges are given by

$$f_{LO} = f_{RF} \pm f_{IF} = (1975 \text{ to } 2035) \pm 193 = (2168 \text{ to } 2228) \text{ MHz and } (1782 \text{ to } 1842) \text{ MHz.}$$

Using the upper LO frequency range: (2168 to 2228)MHz, we find that the IF frequency is

$$f_{IF} = (f_{RF} - f_{LO}) = (1975 \text{ to } 2035)\text{MHz} - (2168 \text{ to } 2228)\text{MHz} = -193\text{MHz}.$$

Then $f_{IM} = (-f_{RF} + 2f_{LO}) = (f_{LO} - f_{IF}) = (2168 \text{ to } 2228)\text{MHz} + 193\text{MHz} = (2361 \text{ to } 2421)\text{MHz}.$

Thus, the image frequency is seen to be well outside the receiver pass-band.

7.3 MIXER DESIGN USING EDA SIMULATION TOOL

In the current mixer design, engineers use electronic design automation (EDA) simulation tools for professional designing of microwave mixers. A rudimentary approach for that has been explained with Example 7.4.

Example 7.4 Design a single-balanced diode mixer for FM-CW radar receiver using a hybrid ring with the following specifications:

f_{RF} = 14.5 GHz, f_{IF} = 1.5 GHz. Use electronic design automation (EDA) simulation tool to realize the design.

Solution:

Step 1: *Rat-race design and characterization*: The scheme of a singly-balanced diode mixer designed with hybrid ring (rat-race), implemented in microstrip transmission line, along with a pair of Schottky diode and a low-pass filter is shown in Figure 7.22. The rat-race has a ring type guide with characteristic impedance $\sqrt{2}Z_0$ and four side arms each with characteristic impedance Z_0. Based on the choice of the substrate material, the line width (W) of the ring structure and the side arms of the rat-race can be calculated. To do this, we refer to Figure 5.15a of Chapter 5 and observe that both for Z_0 = 50 Ω (for side arms) and Z_0 = 70.71 Ω (for the ring type guide); *W*/*h* is more than 2 and hence we need to use appropriate equation for (*W*/*h*) ≥ 2 [6], see Example 5.2. With due calculation, it can be found that W_{ring} = 0.7075 mm and $W_{side\ arm}$ = 1.2292 mm. For the present design, Rogers substrate with ε_r = 2.2 and having substrate height (h) = 10 mil (= 0.254 mm) is used. The radius of the ring can be calculated from: $r_{ring} = 1.5\lambda_{eff(ring)}/2\pi$ = 3.939 mm (with f = 14.5 GHz); and λ_{eff} (=1.6569 cm); calculated from appropriate equation with (W/h) > 1 [6], see Example 5.2.

The simulated structure for the rate-race in EDA simulation tool is shown in Figure 7.23, and the simulated results for the *S*-Parameters are shown in Figure 7.24.

Step 2: *The diode and its characteristics*: For the present design Schottky diode with part number SMS3925-040LF is used. The SPICE (Simulation Program with Integrated Circuit Emphasis) model of the diode with the diode parameter values are shown in Figure 7.25.

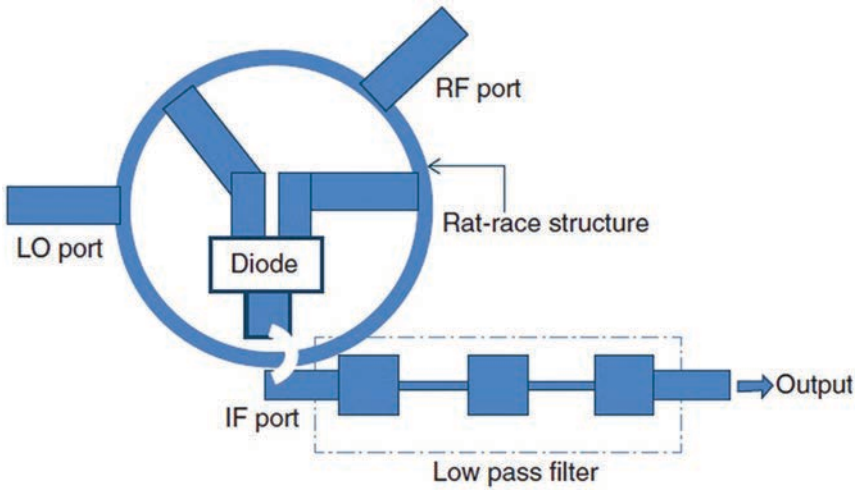


Figure 7.22 Schematic of the single-balanced diode mixer based on rat-race.

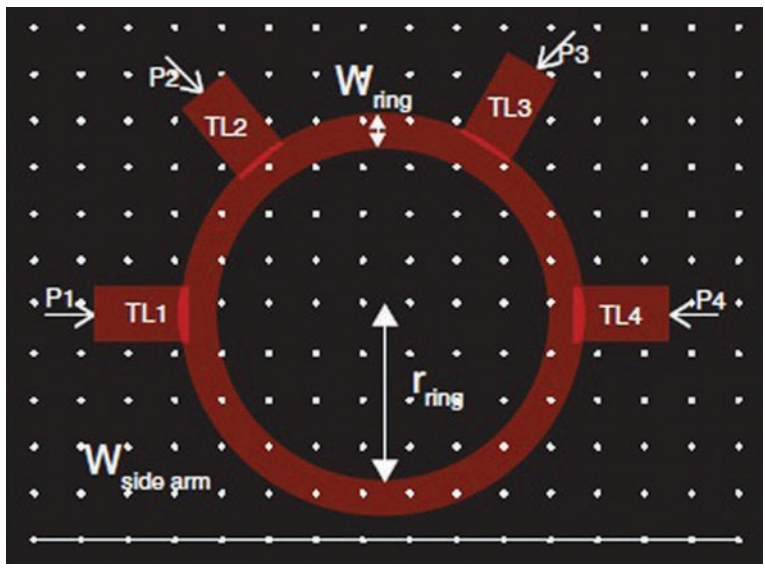


Figure 7.23 Simulated structure of the rat-race.

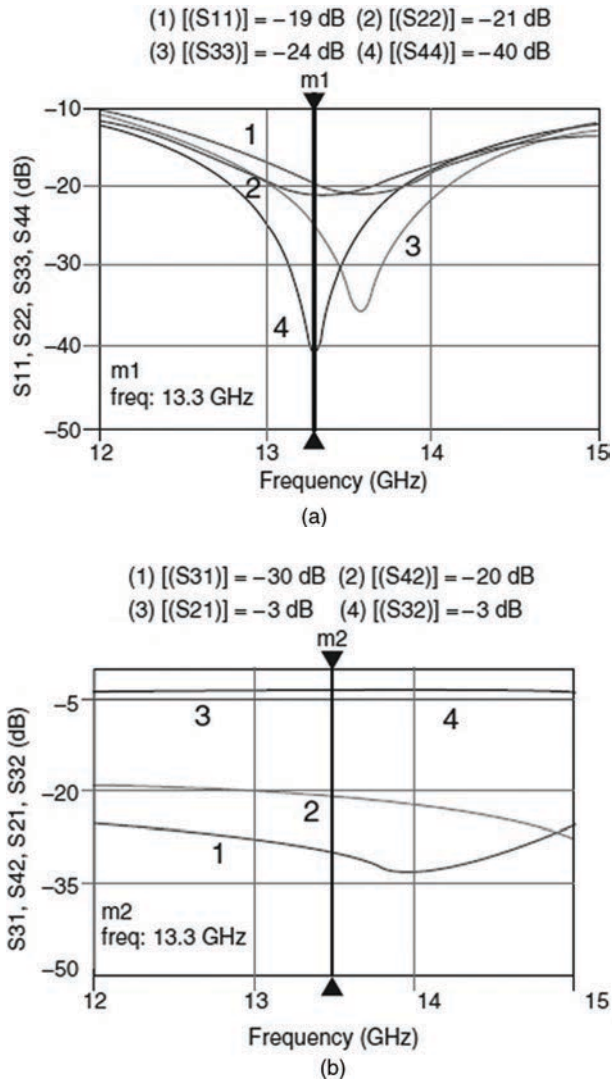
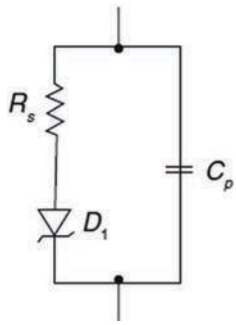


Figure 7.24 Simulation results for the rate-race structure showing: (a) Return loss for all ports, and (b) Isolation between ports 3 and 1 and ports 4 and 2 and insertion loss between ports 2 and 1 and ports 3 and 2.



Parameter	Units	Description	SMS3925-040LF
R_s	Ω	Series Resistance	5.4
C_{j0}	pF	Junction Capacitance	0.36
V_j	V	Junction Potential	0.8

Figure 7.25 The SPICE model of the diode with the parameter values.

Step 3: *Mixer simulation and simulated result*: The schematic of the total mixer circuit with the diode is shown in Figure 7.26 while the simulated result of the mixer output is given in Figure 7.27.

Simulation result shown in Figure 7.27 indicates that the mixer has a conversion loss: (RF Power – IF Power) = [–20dBm – (–29.7dBm)] i.e., 9.7dB at IF frequency of 1.5GHz. The mixer conversion loss can further be improved by using matching elements in the RF and LO ports. The LO rejection (Input LO Power – Leakage LO Power at IF Port) at 13GHz is [10dBm – (– 17.6dBm)], i.e., 27.6dB. The LO rejection can further be improved at the IF port by using a low-pass filter (stepped-impedance type, section 5.5.2). The layout of the filter structure used here is shown in Figure 7.28 with the dimensions in Table 7.1 (which is same as the one designed in Example 5.2); the cut-off frequency of the filter being 2.4 GHz.

It may be observed from Figure 7.29 that the second harmonic of IF (m2) has been practically eliminated by the use of LPF while the LO rejection has significantly improved. It is now 44.07dB [10dBm – (–34.07dBm)]. The amplitude of the image frequency signal has also been reduced.

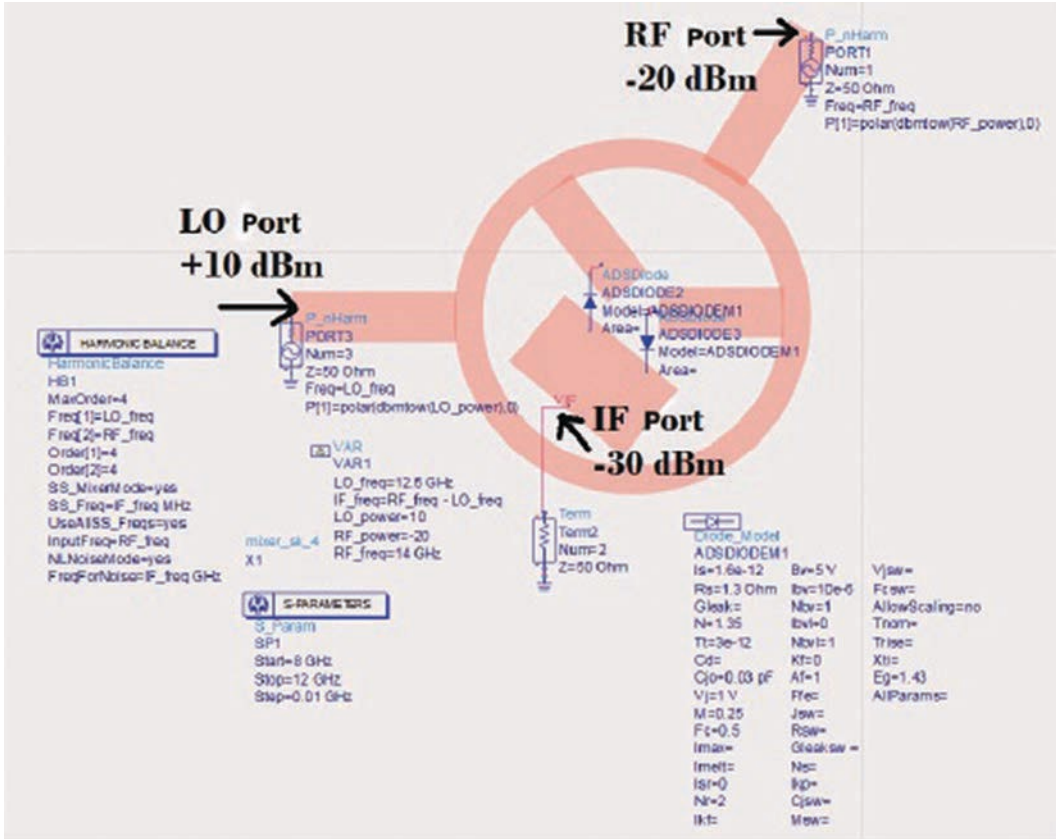


Figure 7.26 Simulated structure of the total mixer.

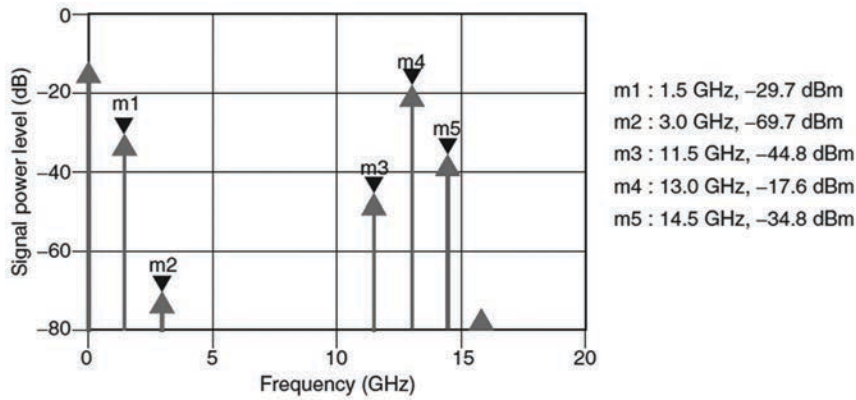


Figure 7.27 Simulation result of the mixer output [m1→ IF, m5 → RF, m4 → LO, m2 → second harmonic of IF, m3 → image frequency].

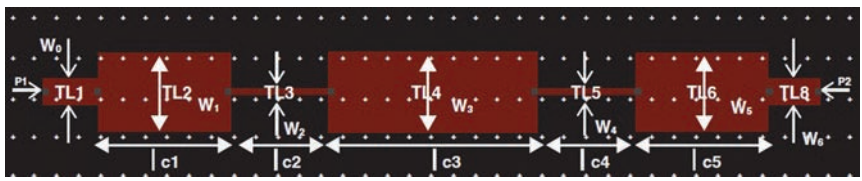


Figure 7.28 Layout of the low-pass filter.

Table 7.1 Different dimensions of the low-pass filter used in the design

Parameters	W_1, W_3, W_5	W_0, W_6	W_2, W_4	l_{c1}, l_{c5}	l_{c2}, l_{c4}	l_{c3}
Values (mm)	2.00347	0.6216	0.0392	6.028	4.406	9.399

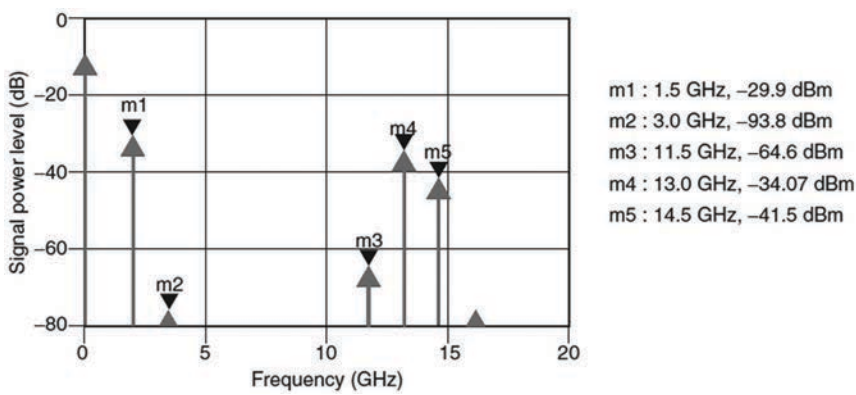


Figure 7.29 Final output at the IF port of the mixer.

REFERENCES

[1] Kurokawa K, and Magalhaes F M, "An X-band 10W multiple diode oscillator", *Proceedings of the IEEE,* Vol. 61, pp. 242–249, February 1973.
[2] Kar S (sole inventor), *A Twin-cap IMPATT Power Combiner, Indian Patent No.: 18414*, issued in 2003.
[3] Wilkinson E J, "An N-way power divider", *IRE Transactions on Microwave Theory and Techniques*, Vol. 8, pp. 116–118, January 1960.
[4] Fantauzzi S, Valletti L, and Di Paolo F D, "High power density spatial combiner for the q-Band, ready for space applications", *Progress In Electromagnetics Research M*, Vol. 109, pp. 163–177, 2022. doi:10.2528/PIERM21120903
[5] Mass S A, *Microwave Mixers,* Second Edition, Artech House Inc., MA, 1993.
[6] Edward T and Steer M B, *Foundations for Microstrip Circuit Design*, Fourth Edition, Wiley-IEEE Press, West Sussex, UK, 2016.

Chapter 8

Microwave Antenna Design

Something that emerges apparently from invisibility is creativity

8.1 INTRODUCTION

Antenna (US)/Ariel (UK) is a transducer that converts electrical signal to electromagnetic waves in a transmitter and vice versa in a receiver of a radio/wireless system. In fact, the term "radio" has been derived from the *radiation* of electromagnetic waves from antenna. Antenna is an integral part of a wireless communication system in RF and microwave communication and radar systems.

Antennas can, in general, be of two types: *wire/rod type* antennas as the radio frequency antennas like loop antenna, dipole antenna, rhombic antenna, and so forth (used in kHz – MHz frequency range) and *diffraction-type* antennas like horn antenna, parabolic dish antenna and so forth that are used at higher RF microwave frequencies, i.e., in GHz (microwave) frequency range. Apart from that there are some special antennas of recent origin like microstrip antenna, dielectric resonator antenna useful at microwave and millimeter-wave frequencies.

In the higher RF and microwave frequency range, the wavelength of the electromagnetic signal becomes comparable to or even smaller than the size of the antenna and thus diffraction phenomena are dominant as in optical systems (such as reflectors, lenses). Microwave antennas can be made relatively high gain compared to any lower frequency RF antenna since its aperture size (D) with respect to wavelength (λ) of operation can be made equal to or larger than unity, $D/\lambda \geq 1$. This is absolutely necessary as power available from microwave active devices is much smaller than that available with lower frequency RF devices (as $P_{\text{available}} \; \alpha \; 1/f$). Thus, effective radiated power from the antenna (which is proportional to the product of antenna gain and power available from transmitter) remains practically the same whether operating at lower frequency RF or at higher RF and microwave frequency.

At lower radio frequency (RF), we use separate transmitting and receiving antennas (for example, a vertical mast antenna is used for AM/FM transmission while a whip antenna is common for FM reception, turnstile antenna is

 DOI: 10.1201/9781003729969-8

used for TV transmission while a Yagi–Uda antenna is used for TV reception and so forth). But at microwave frequencies, the same type of antenna can be used for both transmission and reception. For example, a parabolic dish antenna is used for both transmission and reception in point-to-point microwave communication and satellite communication. In fact, in a monostatic pulsed radar operating at microwave frequency, the same (single) antenna is used for both transmission and reception via a circulator so as to separate (isolate) the transmitted and received signals. Wire antennas of various forms (loop antenna, rhombic antenna) are extensively used at lower RF applications, while at microwave frequencies, wire antennas are rare. However, a variant of the wire antenna known as helical antenna (which can selectively generate right-handed or left-handed circularly polarized electromagnetic signals) is used in some microwave applications demanding broad band of operation.

At present, many different forms of microwave antennas are used (dependent on diffraction principle) whose principle of operation is more or less similar and differ significantly from antennas operating at lower RF. The differences between such diffraction principle-based microwave antennas, however, lie in the design features, feeding methods, and so on. In general, there are five types of such microwave antennas—horn, reflector, slot, lens, and surface-wave (dielectric rod) antenna. Each of these types has several sub-types with a large number of variants [1,2].

Apart from these conventional diffraction-type microwave antennas, many civil, military, and air-and space-borne applications demand planar (printed) antennas such as microstrip patch antenna and its variants which are useful for low- to medium-power applications. In recent times, dielectric resonator antenna is also gaining importance for millimeter-wave applications.

A microstrip patch antenna (MPA) is a patch of conductive material that is etched on a dielectric substrate. This dielectric material is mounted on a ground plane made up of metal with high conductivity (usually copper). This type of antenna operates at high frequencies. To achieve higher efficiency and larger bandwidth, the dielectric material which has low permittivity should be selected. MPAs have become extremely useful because of they can be directly printed on the circuit board, which makes it compact in size and weight. They can be easily fabricated and hence have wide range of applications in mobile communication, satellite communication, and even in global positioning system (GPS).

Detailed analysis of all types of diffraction type antennas is available in a number of standard text books [1,2] and those for microstrip antenna and dielectric resonator antenna are also available in literature [3,4]. However, it must be noted that antennas form a subject by itself [2] including the microwave antennas [1] and hence we would restrict here to the discussion of only

microstrip antennas for brevity of presentation; including metamaterial-inspired microstrip patch antenna as these are of recent origin that leads significant size miniaturization compared to conventional MPA together with better performance useful for mobile phone applications and air-and space-borne applications of recent times.

8.2 PRINCIPLE OF OPERATION OF MICROSTRIP PATCH ANTENNA

Historically, in the 1970s, the military needs primarily demanding antenna that do not interfere with the aerodynamics of airplanes and missiles led to the development of microstrip patch antenna. However, their planar profile and miniature size was later on found to be attractive for commercial applications as well. Further, these antennas can be designed to have dual-polarization and multi-frequency operations which are the demand of modern communication systems. The microstrip patch antenna finds application from space shuttle imaging radar to commercial cell phone receivers and so on. Though the basic microstrip patch antenna is a narrow band antenna with practically realizable bandwidth between 2% and 5%, however, various bandwidth-widening techniques have been developed which can achieve a bandwidth of the order of even 50% to 60% at the sacrifice of increased antenna physical volume. The realizable gain of these antennas is 5 to 6 dB. Due to the small separation between the metal patch and the ground plane of such a structure of the antenna, a microstrip patch antenna can handle relatively low RF/microwave power when compared to conventional antennas; the average power handling capacity being few tens of watts.

The basic structure of a microstrip patch antenna consists of a thin metallic patch (may be rectangular, square, circular, or triangular in shape) etched on a thin (thickness being small compared to wavelength) dielectric substrate with printed circuit technology (for this reason it is also known as *printed antenna*), the substrate being backed by a ground plane. The dielectric substrate is preferably of low dielectric constant may be 2 to 3 (commonly used one is PTFE, polytetrafluoroethylene) for good radiation efficiency (as lower the dielectric constant, less is the confinement of electromagnetic signal in the substrate dielectric material allowing larger fringing field and higher radiation capability).

The operation of a microstrip patch antenna may be understood with reference to Figure 8.1a which shows a rectangular patch of length L and width W printed on a dielectric substrate of thickness/height h (which is basically a truncated microstrip of length L and width W). The region between the patch and the ground plane may be treated as a cavity that is surrounded by magnetic walls around the periphery (open boundary) and by electric walls

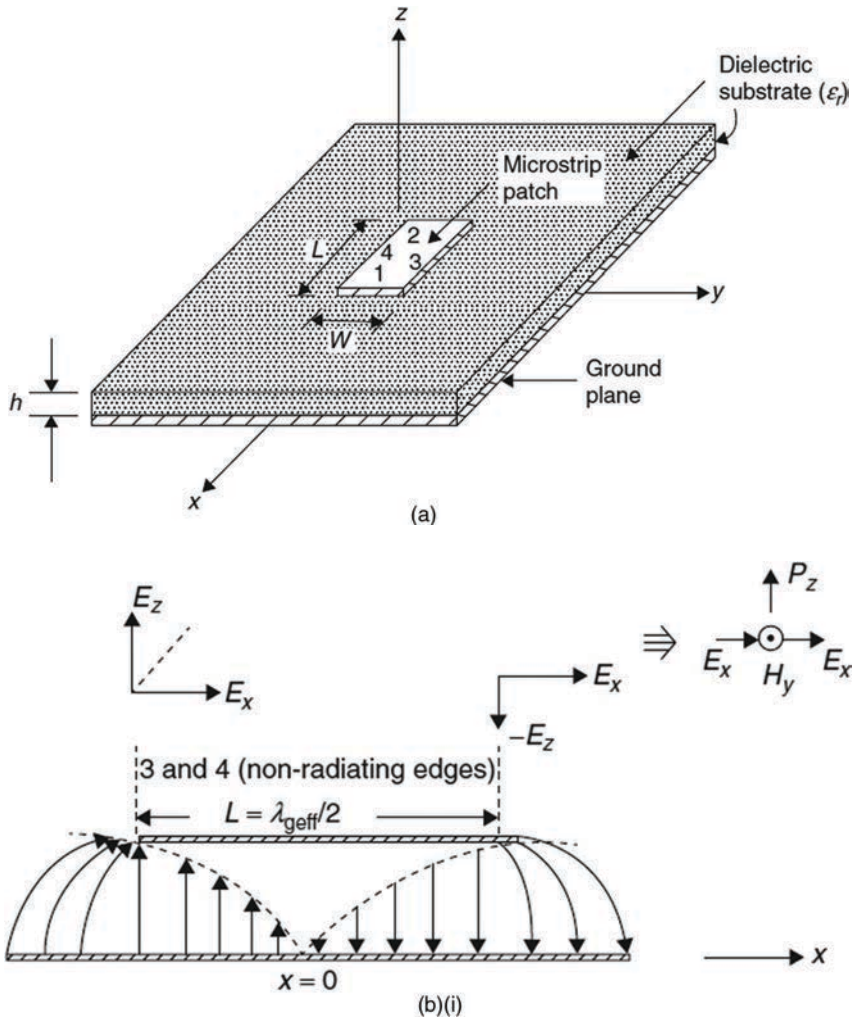


Figure 8.1 Microstrip patch antenna: (a) Basic structure of the antenna showing radiating and non-radiating edges, and (b) (i) Electric field variation along non-radiating edges, and (ii) Fringing electric field across the radiating edges.

at the top and bottom sides (short circuit boundary). Since thin substrates are used, the field inside the cavity is uniform along the thickness/height of the substrate. The length L of the patch is chosen to be approximately half guide-wavelength ($\lambda_{geff}/2$, λ_{geff} being the effective guide wavelength of the microstrip line whose truncated version is the patch antenna), and the cavity mode is of TM type (unlike the quasi-TEM in microstrip line). With

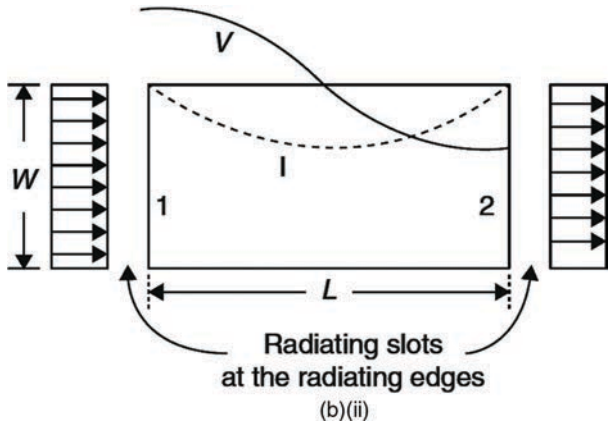


Figure 8.1 (Continued)

the fundamental TM_{10} mode in the open boundary cavity (for a rectangular patch), the electric field varies one $\lambda_{geff}/2$ along the length (L), and there is no variation along the width (W) of the patch, Figure 8.1b. The field is seen to be ideally zero near the patch center $x = 0$, gradually increasing on either side of the center (with phase opposition) and becomes maximum near the patch edges $x = \pm L/2$ (Figure 8.1b(i)), causing a pair-wise cancellation of the electric field on either side of the patch center along L. Thus, the lengthwise edges 3 and 4 act as nonradiating edges. The fringing electric field across the truncated edges W of the microstrip has a nonvanishing electric field E_x (Figure 8.1b (ii)) which are in phase, making W the radiating edge (edges 1 and 2).

The radiated field (caused by the fringing fields along the edge W indicated by radiating slots in the figure) finally comes out from the open microstrip resonator in the broadside direction z, which is the direction of the Poynting vector for the corresponding electric and magnetic fields along E_x and H_y, respectively. Excitation of other modes such as, TM_{01}, TM_{20} and so on, would lead to different radiation patterns. It may be noted that L is chosen in terms of the resonator length ($\lambda_{geff}/2$) while choice of a larger value of W would increase the radiated field and also the directivity of the antenna. However, too large a value of W might excite higher order modes. Since the operation of such an antenna depends on the resonator length L being approximately $\lambda_{geff}/2$, it is a narrow band antenna (being λ-dependent).

For an efficient transfer of power from the transmission line feeding input power to the patch antenna, the input impedance of the patch and the characteristic impedance of the transmission line have to be matched. It is observed that the impedance seen by the transmission line when attached

to the radiating edge is very high [as voltage is finite and current is zero, Figure 8.1b (ii)] and also the impedance decreases as one moves towards the center of the patch. Therefore, depending on the characteristic impedance of the transmission line, an appropriate point on the patch has to be chosen as the feed point.

A microstrip patch antenna can be fed in a number of ways. The coaxial probe is by far the most common feeding method (Figure 8.2a). The coaxial probe usually has a characteristic impedance of 50 Ω, thus the location of the probe should be at the 50 Ω point (which normally demands off-set feeding) of the patch to ensure impedance matching for optimum antenna performance. In practice, the center conductor of the coaxial line is soldered to the patch which needs a hole to be drilled through the substrate material and the connector protrudes outside the bottom ground plane, so the true planar nature of the antenna is not maintained. A patch excited by a microstrip line feed is shown in Figure 8.2b. This feed arrangement has the advantage that it can be etched on the same substrate, so the total structure remains planar. However, the microstrip line needs to be matched with the patch impedance with a quarter-wave impedance transformer. Further, radiation from the feed line leads to an increase of cross-polar level; and in the millimeter-wave range, the size of the feed line becomes comparable to the patch size causing increased undesired radiation. Other feeding methods like aperture-coupled feed and proximity feed are also used depending on the need of the application [3].

8.3 DESIGN OF CONVENTIONAL MICROSTRIP PATCH ANTENNA AND METAMATERIAL-INSPIRED MICROSTRIP PATCH ANTENNA

In this section, we will discuss two design problems: one is the conventional microstrip patch antenna and another one is metamaterial-inspired, i.e., CSRR-loaded microstrip patch antenna; both the design will be done with EDA simulation tool.

8.3.1 Conventional Microstrip Patch Antenna Design

Example 8.1 Design a microstrip patch antenna at 2.4 GHz with FR-4 as substrate material. The dielectric constant of the substrate material $\varepsilon_r = 4.4$ and the height h = 1.6 mm.

Solution

A schematic of the antenna layout for design is shown in Figure 8.3. To calculate the width (W) and length (L) of the patch, we use the following formulas [5]:

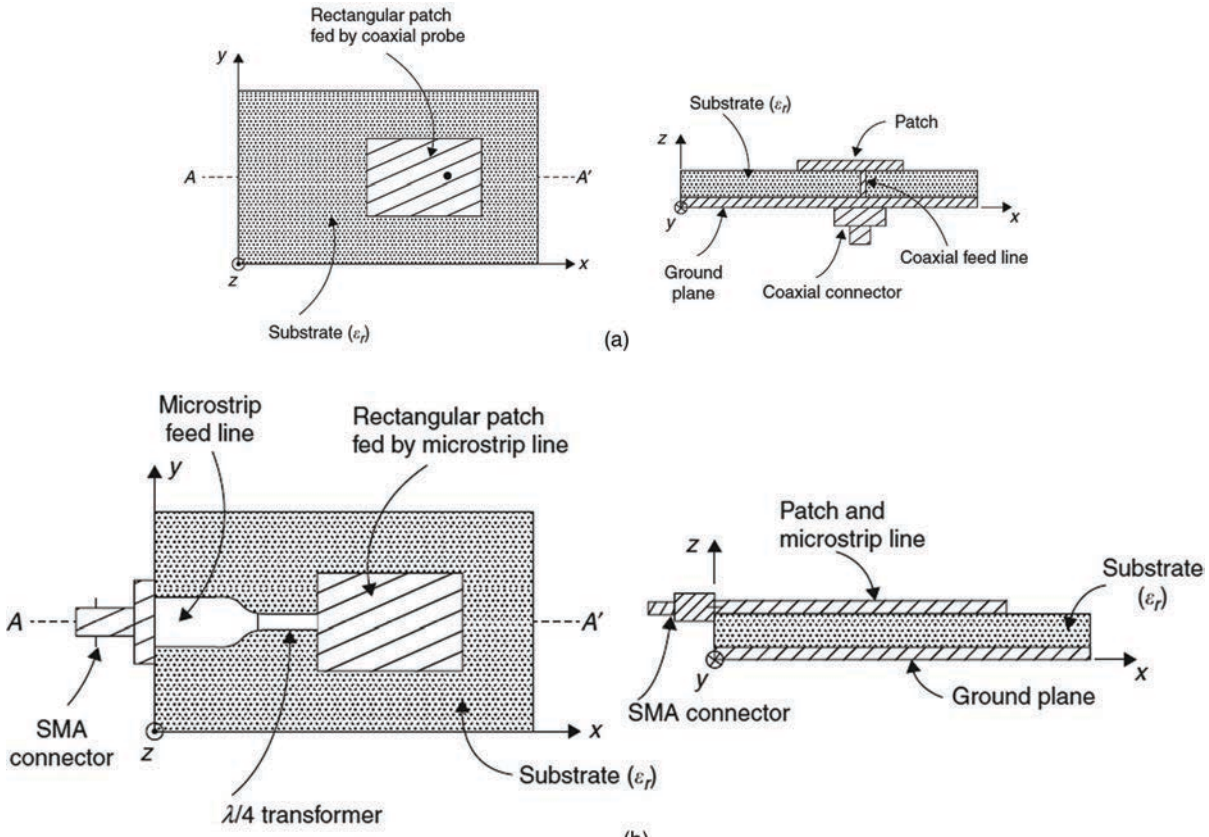


Figure 8.2 Feeding techniques for microstrip patch antenna: (a) Coaxial feed, and (b) Microstrip feed; in both cases shown top view on the left and side view on the right.

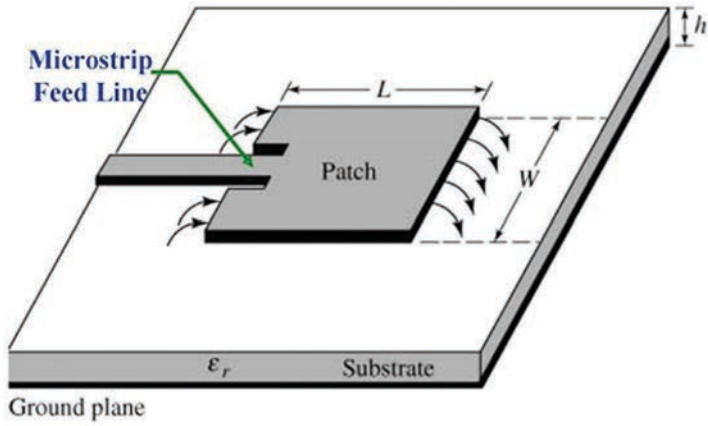


Figure 8.3 Schematic of the microstrip patch antenna for design.

$$W = \frac{c}{2f_r}\sqrt{\frac{2}{(\epsilon_r + 1)}} \tag{8.1}$$

$$L = \frac{c}{2f_r\sqrt{\epsilon_{eff}}} - 2\Delta L \tag{8.2}$$

$$\text{With: } \epsilon_{eff} = \frac{(\epsilon_r + 1)}{2} + \frac{(\epsilon_r - 1)}{2}\left\{1 + \left(\frac{12h}{W}\right)\right\}^{-1/2} \tag{8.3}$$

$$\Delta L = \frac{0.412h(\epsilon_{eff} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\epsilon_{eff} - 0.258)\left(\frac{W}{h} + 0.8\right)} \tag{8.4}$$

where f_r is the resonance frequency while c is the speed of light: 3×10^8. W, L, and h has been mentioned above as patch width, patch length, and height of patch from the ground plane (which is the height/thickness of the substrate material).

Using the above-mentioned formulas and using the given data, we get W = 39 mm and L = 29.4 mm. We have kept the substrate width and length same as the ground plane, i.e., 60 mm × 60 mm. We have used a microstrip line feeding technique. The central strip width is 3 mm, and slit width between the feed and patch is 1 mm on both

the sides see Figure 8.4a. The ground and patch material chosen is copper.

For simulating the antenna, we have defined a radiation box and assigned the box material to be air. A 3D view of the microstrip patch antenna as simulated in HFSS by Ansys is shown in Figure 8.4b. The plots for the E-field vector and the H-field vector obtained with the simulator is shown in Figure 8.5a and 8.5b, respectively.

Upon simulating, we have realized a 2.38 GHz patch antenna, and the gain of the antenna is approximately 3 dbi and the 3 dB beam width of the antenna is approximately 46 degrees, see Figure 8.6a for 2D gain plot and Figure 8.6b for 3D gain plot.

The return loss plot of the designed antenna has been shown in Figure 8.7a which shows a return loss of -26.05 dB at 2.39 GHz. Smith chart plot showing the impedance matching between the patch antenna and the microstrip feed is given in Figure 8.7b.

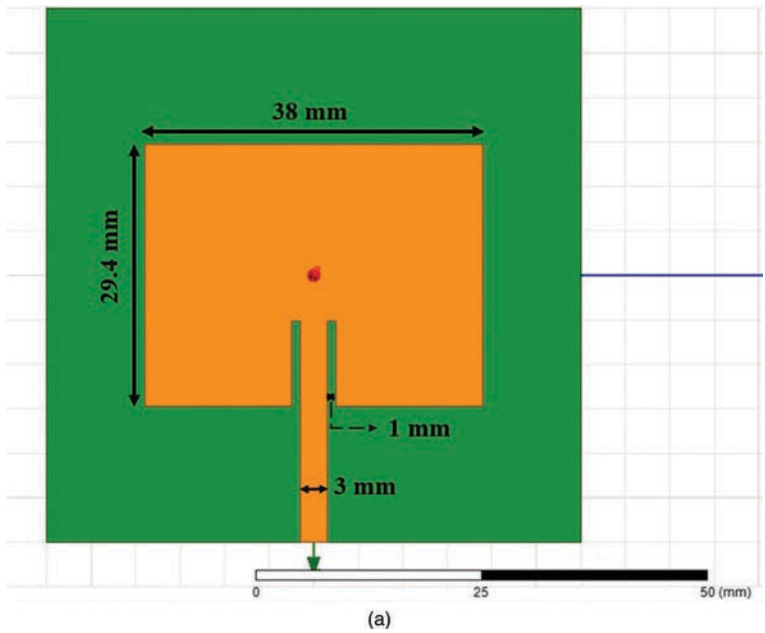
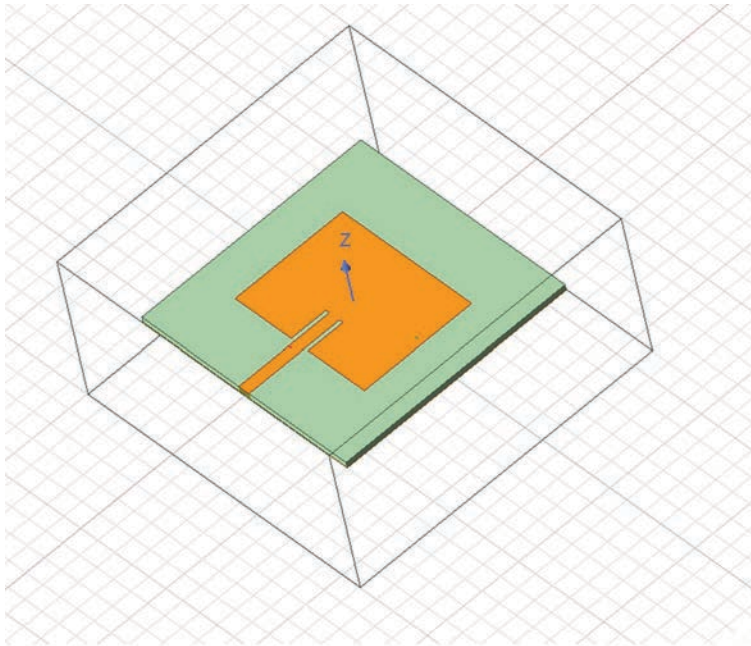
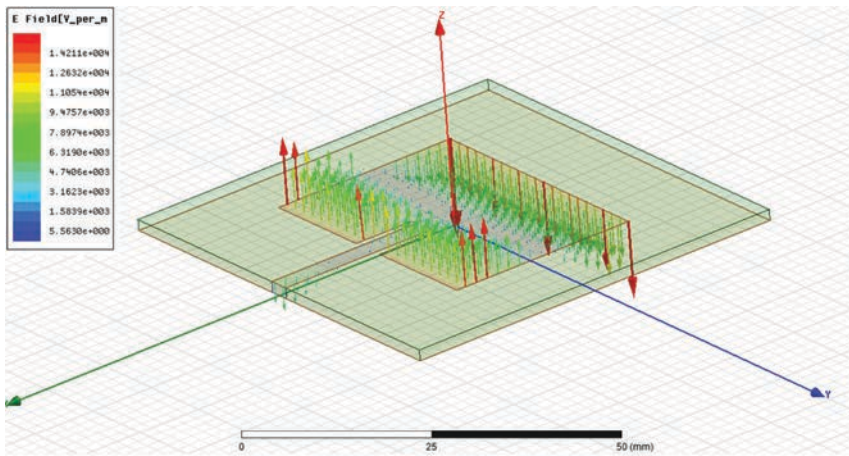


Figure 8.4 (a) Dimensions of the patch antenna to be designed, and (b) 3D view of the microstrip antenna as simulated in high frequency structure simulator (HFSS).



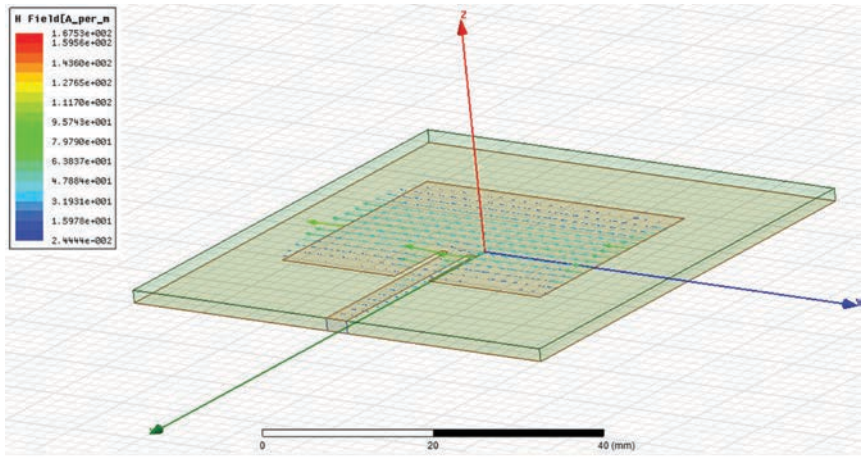
(b)

Figure 8.4 (Continued)



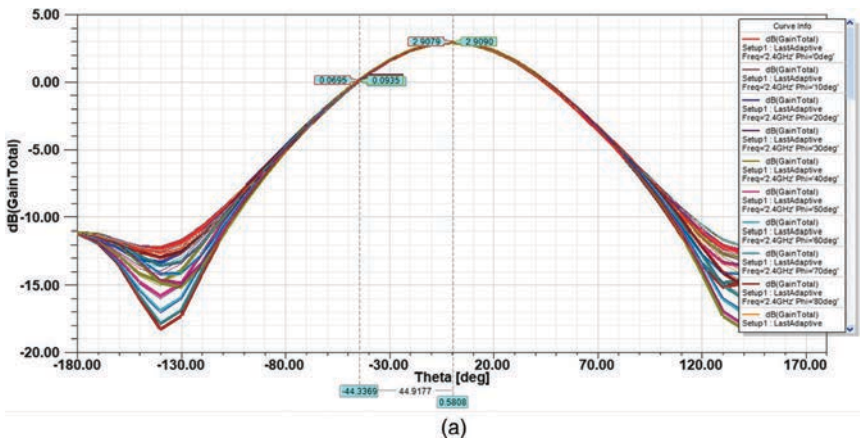
(a)

Figure 8.5 Field vector plots for the antenna: (a) E-field vector, and (b) H-field vector.



(b)

Figure 8.5 (Continued)



(a)

Figure 8.6 Gain plot of the designed antenna: (a) 2D plot, and (b) 3D plot.

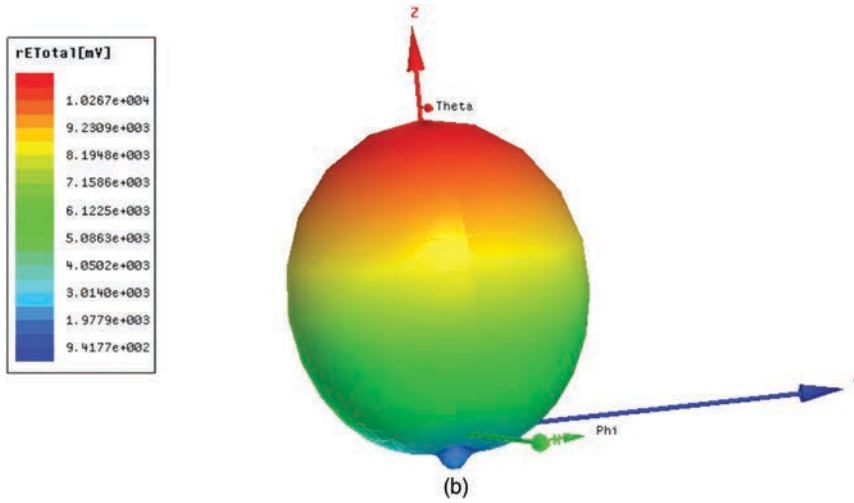


Figure 8.6 (Continued)

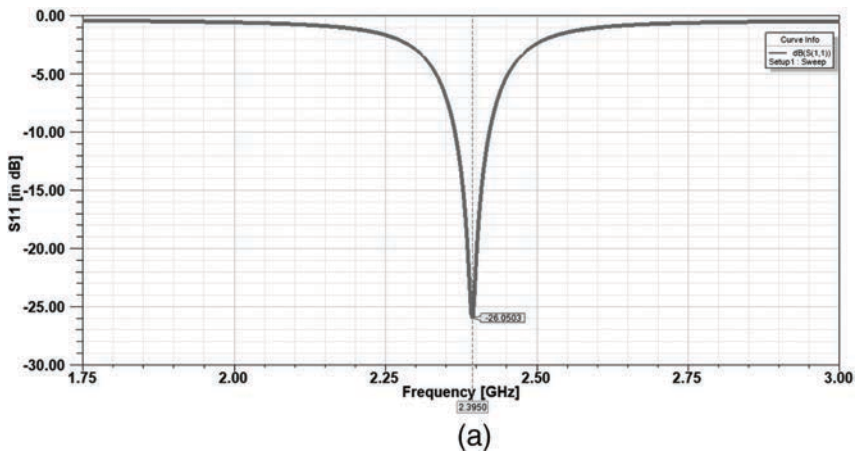


Figure 8.7 (a) Return loss plot of the designed antenna, and (b) Smith chart plot showing the impedance matching between the patch antenna and the microstrip feed.

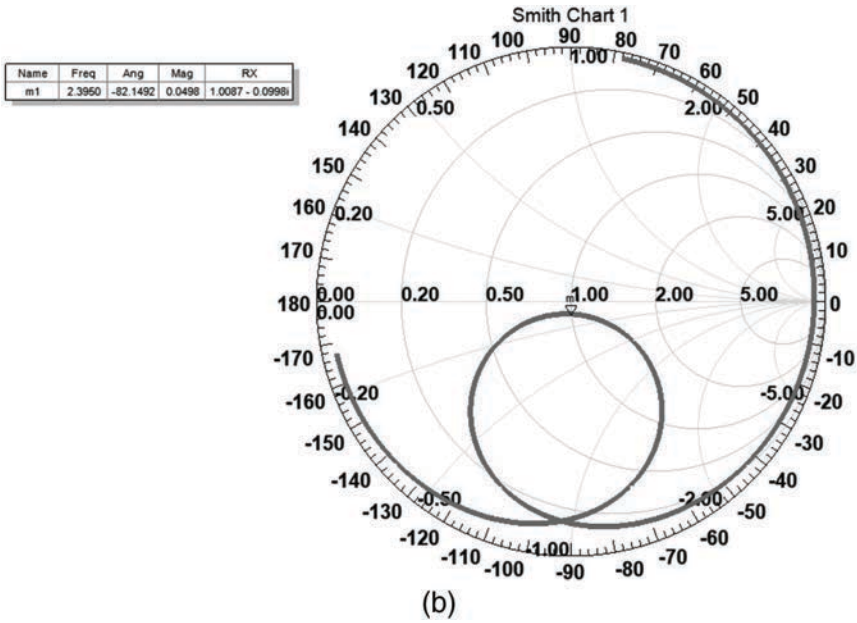


Figure 8.7 (Continued)

8.3.2 Metamaterial-Inspired Microstrip Patch Antenna Design

8.3.2.1 Introduction

Metamaterials are artificially structured (i.e., man-made) material, mimicking the natural material lattice structure constituted of the so-called "electric atom" and "magnetic atom" realized, respectively, with thin wire (TW) array and split-ring resonator (SRR) matrix, see Figure 8.8.

TW and SRR unit cells are, respectively, the electric and magnetic inclusion structures constituting the metamaterial. Anyway, the magnetic inclusion structure SRR of metamaterial and its complementary structure CSRR (complementary split-ring resonator) are useful resonant structures in designing different microwave passive circuits and antennas. Such metamaterial inclusion-based RF/microwave passive components and antennas are found to have better performance characteristics with significant size miniaturization. In the following, we will discuss metamaterial-inspired, i.e., CSRR-loaded microstrip patch antenna design and would find that such metamaterial-inspired antenna has 24% size reduction with improved performance when compared with conventional microstrip patch antenna with no metamaterial loading.

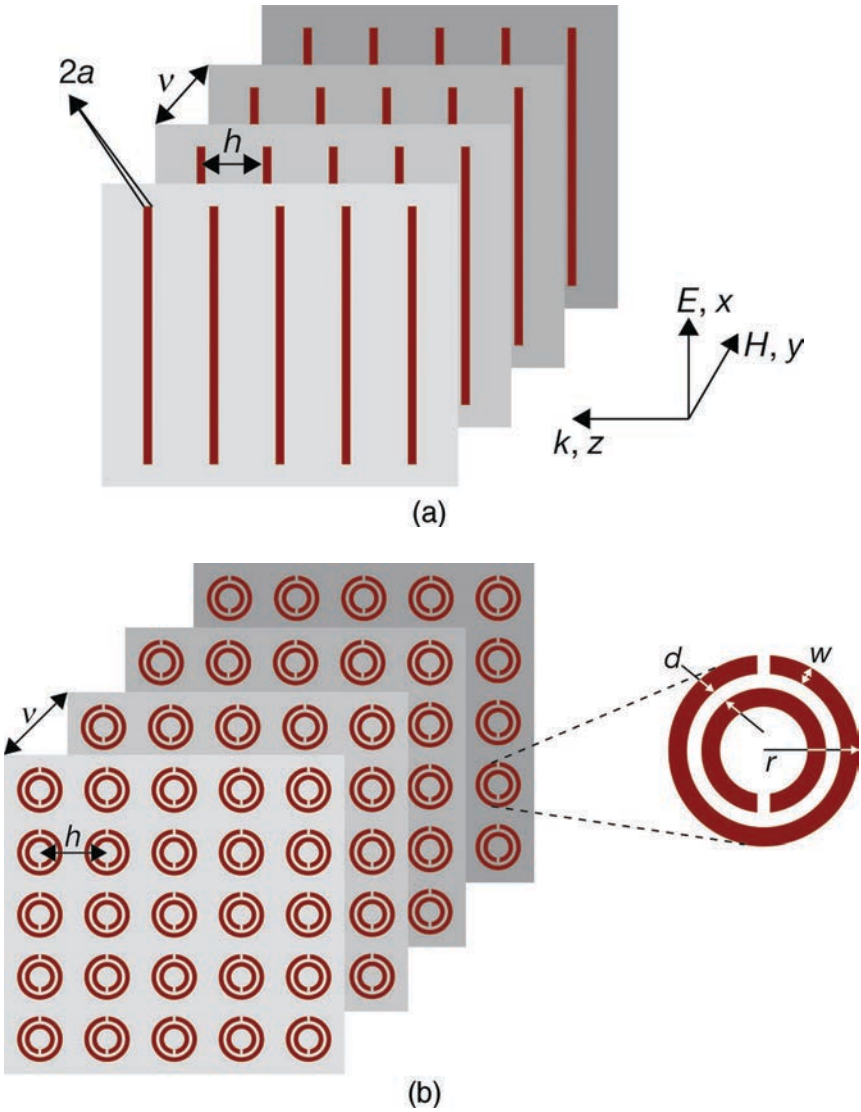


Figure 8.8 (a) An array of TW, and (b) A matrix of SRR.

CSRR represents the complementary version of the well-known SRR (see Figure 8.9a). In microstrip design, CSRR is the form of SRR which can be realized with planar fabrication technology. The gap portions of SRR are the metallic portions for CSRR and vice versa. Electrically, CSRR is the *dual* of SRR (as per Babinet's principle), the latter being a magnetic

inclusion while the former is an electric inclusion structure. The SRR can be considered as a resonant magnetic dipole that can be excited by an axial magnetic field, whereas the CSRR essentially behaves as an electric dipole with same frequency of resonance (as that of the equivalent SRR) and can be excited by an axial electric field. As CSRR is electrically dual of SRR, the equivalent-circuit inductance and capacitance of SRR will, respectively, correspond to the capacitance and inductance of CSRR. Since, dual parameters are only related by constants, it is evident that the resonant frequency which depends on the product of L and C will remain constant for a particular inclusion whether we treat it as an SRR or as a CSRR.

8.3.2.2 CSRR-Loaded Microstrip Patch Antenna Design

Example 8.2 Design a microstrip square-patch antenna with CSRR loading at 5.2 GHz and compare its performance with a conventional microwave square-patch antenna operating at the same frequency.

Solution

The schematic sketches of CSRR and CSRR-loaded microstrip square-patch antenna (MSPA) are shown in Figure 8.9a and 8.9b, respectively; in Figure 8.9b $\varepsilon_r = 2.2$ and $h = 0.508$ mm.

A conventional MSPA has been designed with metallic patch dimensions 18.45 mm × 18.45 mm printed on a substrate (Arlon Diclad 880 having $\varepsilon_r = 2.2$ and $\tan\delta = 0.0009$) with size 50 mm × 50 mm and height $h = 0.508$ mm that exhibit resonance at 5.2 GHz

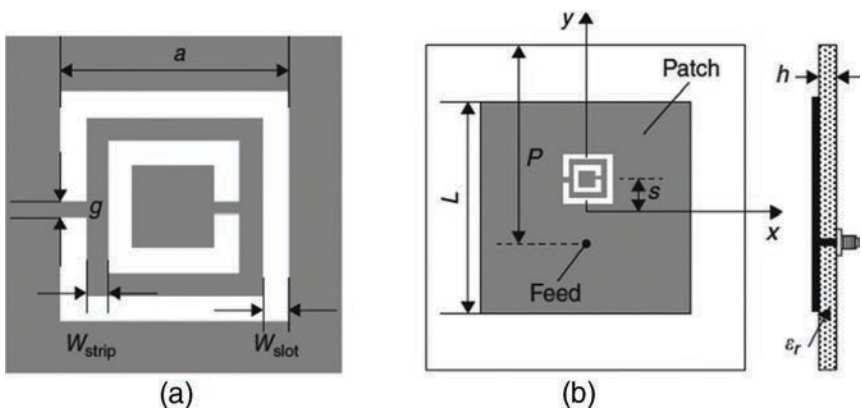


Figure 8.9 (a) Schematic of CSRR, and (b) CSRR-loaded MSPA.

[6]. It is fed with a standard 50 Ω coaxial cable having diameter of the inner and outer conductors 1.28 mm and 4.29 mm, respectively, with Teflon dielectric (having $\varepsilon_r = 2.1$). The impedance matching between the metal patch and the coaxial probe was done by optimizing the probe position *P* (see Figure 8.9b) using 3D electromagnetic field simulator. The MSPA schematic along with the Smith chart plot showing impedance matching is depicted in Figure 8.10a and 8.10b, respectively.

When CSRR resonator is loaded on the MSPA the performance optimization needs to be done by controlling the CSRR size and the parameter *s* (see Figure 8.9b). The schematic of microstrip square-patch loaded with square CSRR, having the line of cut in CSRR aligned orthogonal to the coaxial feed together with the surface current lines are shown in Figure 8.11a and 8.11b, respectively.

Other orientations of CSRR, one aligned with its line of cut along the line of the coaxial feed and facing the feed exhibits lossy nature while the other case when CSRR is aligned with its cut along the line of coaxial feed but rotated by 180° in clockwise direction exhibit no possibility of size reduction as it exhibits resonance at 5.8 GHz. The CSRR-loaded patch given in Figure 8.11a when optimized in terms of parametric study indicates that CSRR loading on patch brings down the resonance frequency below 5.2 GHz. To bring it back to 5.2 GHz with perfect impedance matching (see Figure 8.11c), the patch size was found to be reduced to 14 mm × 14 mm with other optimized dimensions as indicated in Table 8.1. It may be observed that a size reduction of more than 24% of the patch is realizable with CSRR-loading on MSPA. A fabricated CSRR-loaded MSPA antenna and a conventional MSPA antenna structure are shown in Figure 8.12.

A comparative performance of conventional MSPA antenna and CSRR-loaded MSPA antenna is given in Table 8.2.

The characteristics of the designed antenna are shown in Figure 8.13.

An analytical estimation of the resonant frequency of CSRR can be done as follows. Babinet's principle suggests that CSRR is the dual of SRR, so the resonant frequency of SRR and CSRR are expected to be the same [7]. The SRR is basically an LC resonator. The total inductance of SRR structure shown in Figure 8.14 can be calculated as

$$L_{eq} = \frac{L_{out} + L_{in}}{2} = 11.4nH \tag{8.5}$$

where L_{out} and L_{in} are calculated by considering each segment of the inner and outer rings as rectangular segments [8]. The evaluation of the equivalent inductance L_{eq} as the average of the L_{out} and L_{in} is justified.

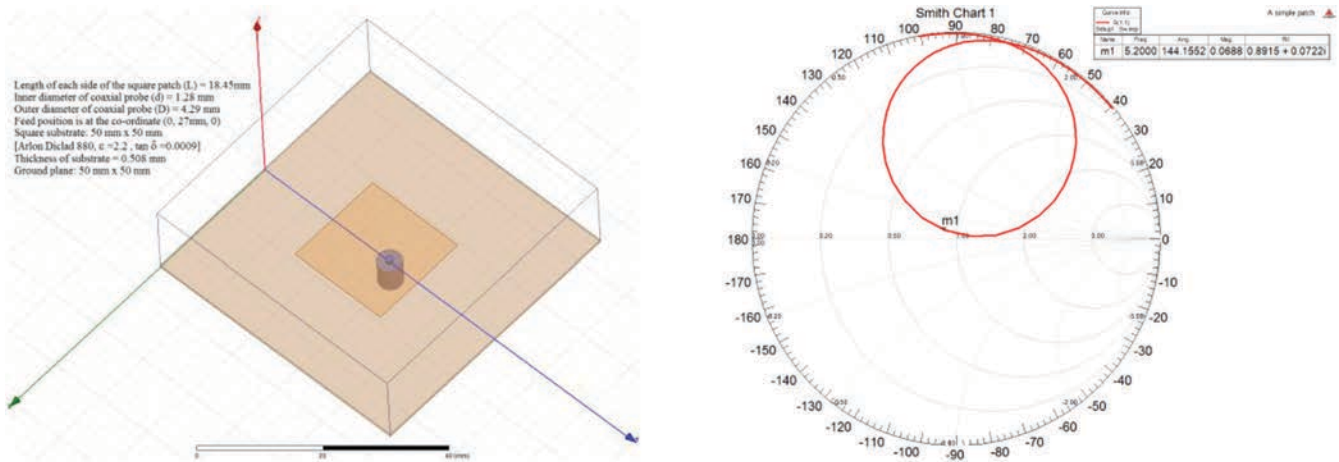
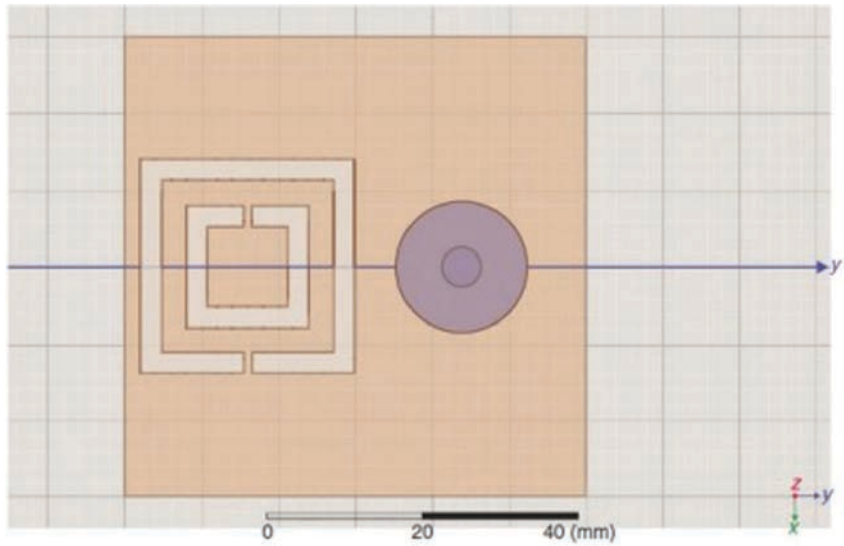


Figure 8.10 (a) Schematic of MSPA, and (b) Smith chart showing the impedance matching.



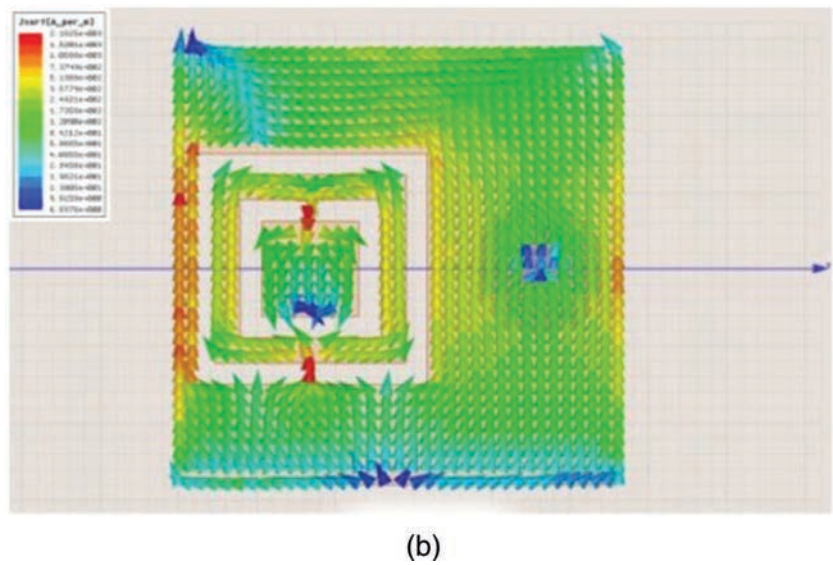


Figure 8.11 (a) Microstrip square-patch antenna loaded with CSRR aligned orthogonal to the coaxial feed, (b) Surface current lines, and (c) Smith chart plot showing matching between coaxial probe and the CSRR-loaded patch.

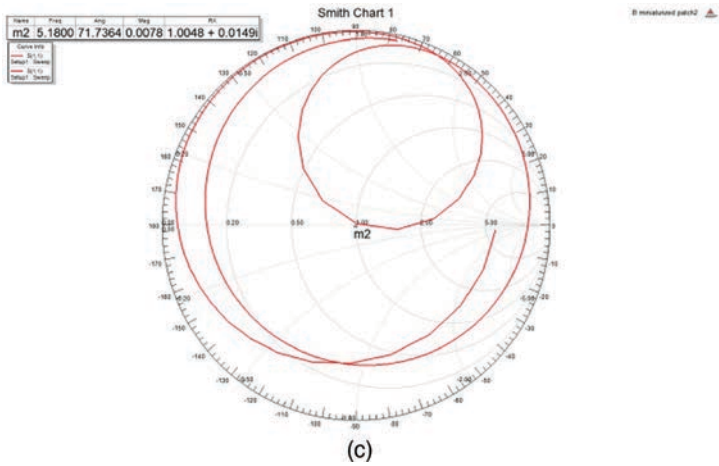


Figure 8.11 (Continued)

Table 8.1 Optimized dimensions of different parameters of CSRR-loaded patch antenna at 5.2 GHz

Dimensions of different parameters for CSRR-loaded MSPA	
Parameters	*Dimensions (in mm)*
L	14
a	7.5
Wstrip	0.8
Wslot	0.7
s	5.8
g	0.25
P	26.7

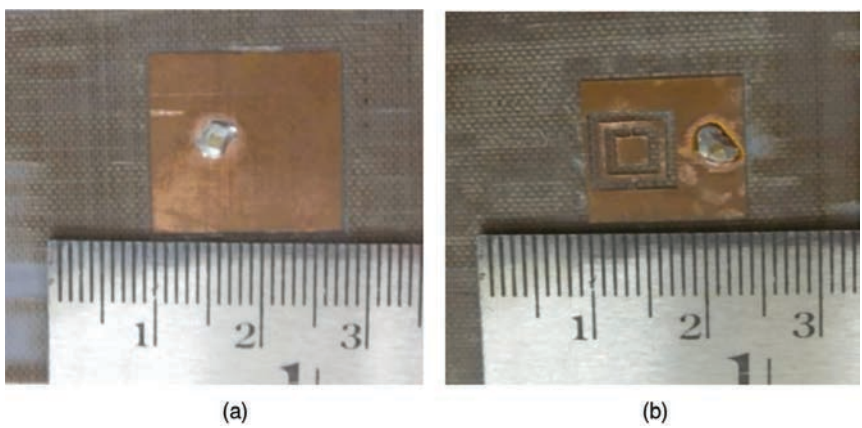


Figure 8.12 Fabricated antennas: (a) Conventional MSPA, and (b) CSRR-loaded MSPA (approximately 24% size reduction).

Table 8.2 A comparative performance of conventional MSPA and CSRR-Loaded MSPA antenna

Antenna type→		*Conventional MSPA*		*CSRR-loaded MSPA*	
	Obtained from→	*Simulation*	*Measurement*	*Simulation*	*Measurement*
E-Plane	Maximum realizable gain	6.86 dB	6.11 dB	4.2 dB	4.17 dB
	Maximum cross-polarization	–56.35 dB	–31.1 dB	–7.43 dB	–10.36 dB
H-Plane	Maximum realizable gain	6.86 dB	6.114 dB	4.2 dB	4.187 dB
	Maximum cross-polarization	–51.06 dB	–20.4 dB	–7.43 dB	– 6.39 dB
–10 dB Bandwidth		170.1 MHz	60 MHz	32.5 MHz	34MHz
Efficiency		81.89 %		50.73 %	

This is because the presence of small slits in both the ring results in circulation of current between the two rings. Thus, the closed current path formed by traversing through both the rings is exactly like a current in a single loop.

Method of Moments (MoM) has been used to find the capacitance per unit length between the two adjacent strips of SRR. Each strip is divided into N number of small rectangles such that surface charge distribution is nearly constant in a rectangle. The thickness of strips is very small (17 μm – 35 μm) compared to other dimensions and hence contribution due to this thickness in overall capacitance is neglected. Assuming unit potential on the strips, charge distribution on each rectangle is calculated from Poisson's equation which results in a $2N \times 2N$ matrix of $2N$ unknown variables. This matrix is solved using MATLAB to find the surface charge density on each rectangle. From surface charge density, total charge on a strip is calculated to find the capacitance for a given strip of width w, gap g, and length l. It is observed that capacitance between two adjacent plates is a linear function of the length of strips. So, capacitance per unit length can be determined and used for further calculations. For our case of SRR, the capacitance calculated using MoM between the split rings is 0.016 pF where the width of the strip is 0.7 mm. Gap between split rings is 0.8 mm. The half of SRR covers effective length of 10.5 mm and hence C_p (Figure 8.14) will be 0.16 pF. The total capacitance for the full SRR structure is a series combination of two capacitors of value C_p. So, the total equivalent capacitance for the SRR will be C_{eq} = 0.084 pF.

Using L_{eq} = 11.4nH and C_{eq} = 8.4 × 10^{-14}F, the resonant frequency of the SRR, f_0, is found to be 5.14 GHz. The CSRR-loaded patch antenna uses Arlon Diclad 880 substrate of dielectric constant 2.2 that has ε_{eff} = $(\varepsilon_r+1)/2$ =1.6. Thus, the resonant

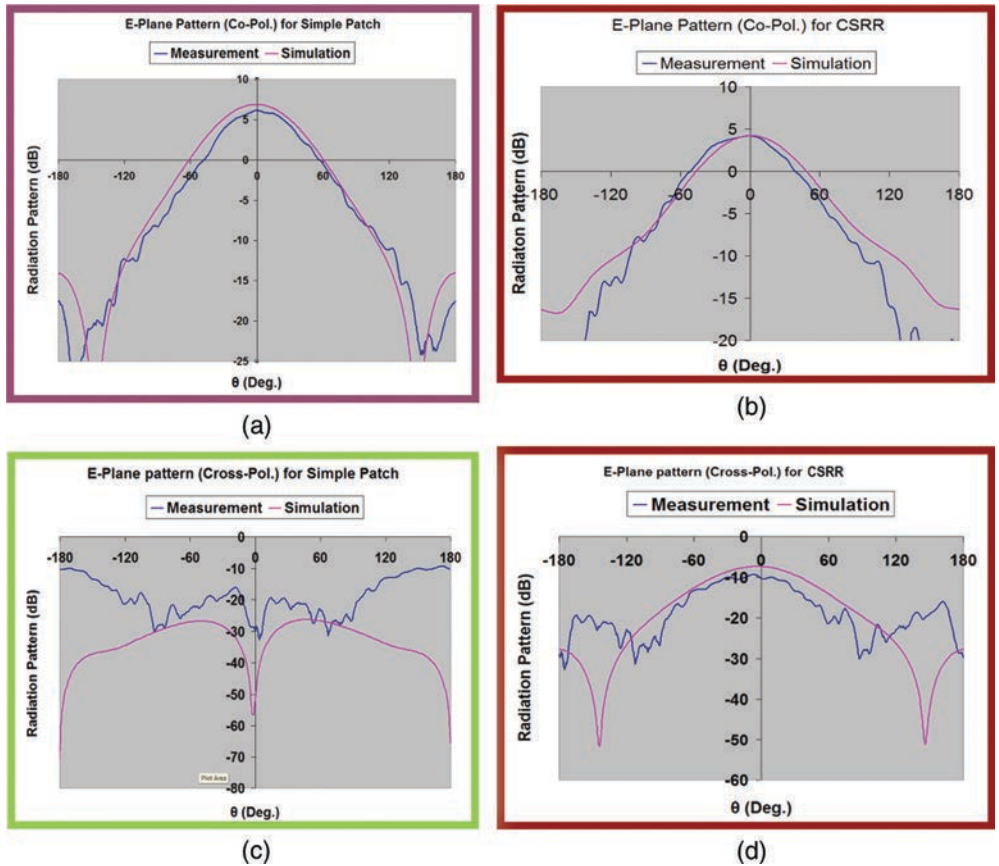


Figure 8.13 Characteristics of the designed CSRR-loaded patch antenna at 5.2 GHz: (a) E-plane pattern (co-pol) for simple patch, (b) E-plane pattern (co-pol) for CSRR-loaded patch, (c) E-plane pattern (cross-pol) for simple patch, and (d) E-plane pattern (cross-pol) for CSRR-loaded patch.

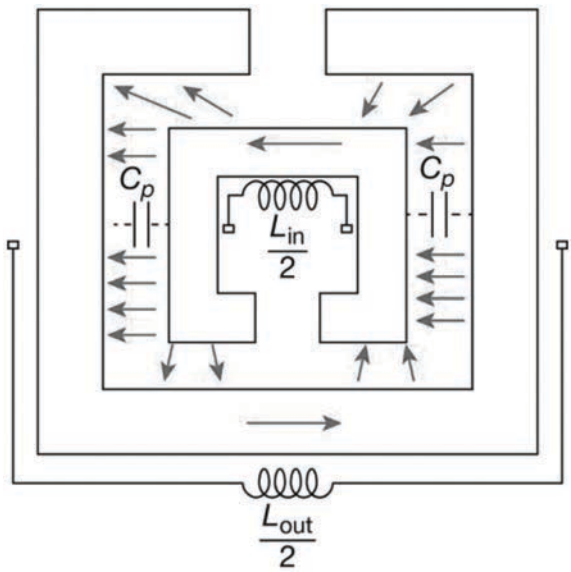


Figure 8.14 Equivalent circuit diagram for SRR to calculate resonant frequency.

frequency of the CSRR will be approximately $f_o/(\varepsilon_{eff})^{1/2}$ = 4.06 GHz. One thing is to be noted that f_o is the resonant frequency of the SRR whereas the metallic patch upon which the complementary form of SRR has been loaded is not an exact dual of SRR since the latter has been cut on a finite metal sheet (i.e., the patch). Due to this reason, some undue loading might have slightly changed the resonant frequency value of the CSRR. Possibly, this is the reason why the resonant frequency calculated analytically (4.06 GHz) and that obtained with simulation (3.03 GHz) of CSRR differs.

REFERENCES

[1] Fradin A Z, *Microwave Antennas*, Pergamon Press, Oxford, 1961.
[2] Kraus J D, Marhelfka R D, and Khan A S, *Antennas: For All Applications*, Third Edition, Tata Mc Graw-Hill Publishing Company Limited, New Delhi, 2006.
[3] Lee K F, Kwai K M, and LaiH W, *Microstrip Patch Antennas,* Second Edition, World Scientific Publishing Co. Pte. Ltd., Singapore,2017.
[4] Volakis J L, "Dielectric resonator antennas", Chap. 17 in *Antenna Engineering Handbook*, Fourth Edition, McGraw-Hill Education, New York, 2007.

[5] Huang Y, and Boyle K, *Antennas: From Theory to Practice*, Wiley, UK, 2008. Also www.everythingrf.com/rf-calculators/microstrip-patch-antenna-calculator

[6] Kar S, Ghosh M, Kumar A, and Majumder A, "Complementary split-ring resonator-loaded microstrip patch antenna useful for microwave communication", *International Journal of Electronics and Communication Engineering*, Vol. 10, pp. 1321–1325, 2016.

[7] Falcone F, Lopetegi T, Laso M A G, Baena J D, Bonache J, Beruete M, Marques R, Martin F, and Sorolla M, "Babinet principle applied to the design of metasurfaces and metamaterials", *Physical Review Letters*, Vol. 93, pp. 197401–197404, 2004.

[8] Wu R, Kuo C, and Chang K K, "Inductance and resistance computations for three-dimensional multiconductor interconnection structures", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 40, pp. 263–271, 1992.

Chapter 9

RF and Microwave System Design

We should aspire for a social system where everyone has mutual respect for each other

9.1 INTRODUCTION

Today, communications are the main driver of RF/microwave system development, leading to RF/microwave technology evolution at an unprecedented pace. Similar relationships exist for national security including radar and sensors used in detection and ranging. As relationships between RF/microwave communication, radar, and sensor design have become more intertwined, a broad appreciation of system design issues has become of prime importance to the RF and microwave design engineers. There are a number of books available on RF design techniques [1–3] which the readers might refer to.

Most RF/microwave system design is done by using modules such as amplifiers, integrated circuits (ICs), filters, frequency multipliers, and passive components putting them in cascade, and each module is separately designed and characterized. Modules are interconnected by transmission lines/waveguides with proper bias settings, and matching networks must be suitably designed. They often have 50Ω input and output impedances so that modules can be freely interconnected. In high volume applications, several modules could be monolithically integrated, but even then, design is based on the concept of modules. Many modules are available "off the shelf," and high-performance RF/microwave systems can be constructed using them. Much of RF/microwave design is about maximizing dynamic range, minimizing noise, and minimizing DC power consumption and so forth.

An RF/microwave signal is a signal that is coherently generated, radiated by a transmitting antenna, propagated through air or space, collected by a receiving antenna, and then amplified and finally the information is extracted. At microwave frequencies, compared to RF, the size of a circuit or structure is comparable to or greater than the wavelength of the electromagnetic signal. The division is arbitrary but if a circuit structure is greater

DOI: 10.1201/9781003729969-9

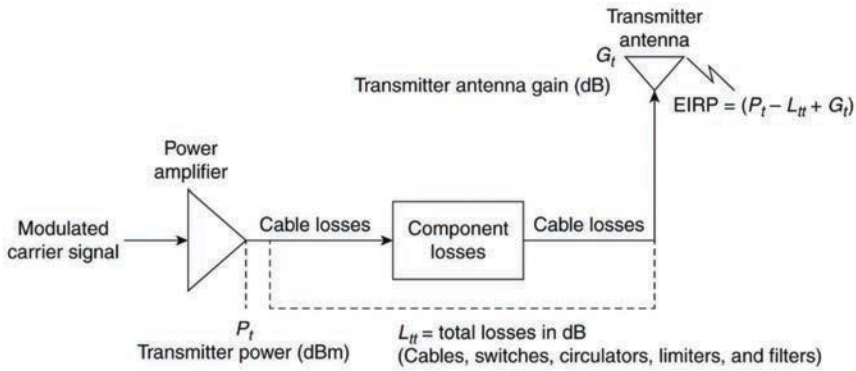


Figure 9.1 RF/microwave section of a communication transmitter.

than $(1/20)^{th}$ of a wavelength, then most engineers would regard the circuit as being a microwave circuit. For now, the microwave frequency range is generally taken as 300 MHz to 300 GHz. At these frequencies, distributed effects, sometimes called transmission line effects, must be considered.

A communication system is, in general, a transceiver of which the transmitter is the part of the transceiver that creates, modulates, and transmits the electromagnetic signal via antenna and through free-space to the receiver. The transmitter is responsible for providing the power required to transmit the signal through the link to the receiver. This includes the power amplifier (PA), the transmitter antenna, and the gains and losses associated with the process, such as cable and component losses, to provide the effective isotropic radiated power (EIRP) out of the antenna, see Figure 9.1.

Anyway, the receiver of a transceiver accepts/intercepts the electromagnetic signal, that was sent by the transmitter antenna, through the channel via the receiver antenna, see Figure 9.2. The losses from the antenna to the low-noise amplifier (LNA), which is the first amplifier in the receiver, should be kept as small as possible. This includes cable losses and component losses that are between the antenna and the LNA (see Figure 9.2). The distance between the receiver antenna and the LNA should also be kept as small as possible. Some systems actually include the LNA with the antenna system, separate from the rest of the receiver, to minimize this loss. The main job of the receiver is to receive the transmitted signal and detect the data that it carries in the most effective way without further degrading the signal.

RF/microwave design engineer should be conversant with some of the important issues like technical aspects related to transmitter and receiver, designer's considerations for transmitter and receiver, link budget, noise and distortion etc., related to the design of such systems.

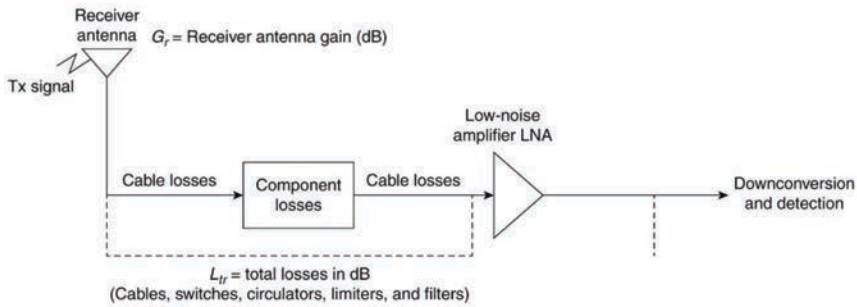


Figure 9.2 RF/microwave section of a communication receiver.

9.2 SOME TECHNICAL ASPECTS RELATED TO RF/MICROWAVE TRANSMITTER AND RECEIVER

We begin with the RF *power amplifier* (PA) that is used to provide the necessary power output from the transmitter to satisfy the link budget for the receiver. The general classifications of amplifiers in terms of the position of the quiescent point of the device was given earlier in Figure 6.16 of Chapter 6. In Section 6.1.3, a comparative discussion of power amplifier (PA) used in transmitters and LNA used in receiver front-end has been given; readers may refer to that and so it has not been repeated here for brevity of presentation.

Next, we consider the *effective isotropic radiated power* (EIRP) from the antenna of a transmitter. Most transceiver systems contain RF components such as a circulator or a transmit/receive (T/R) switch that enables the transceiver to use the same antenna for both transmitting and receiving. Also, if the antenna arrays use multiple antennas, some of their components will interconnect the individual antenna elements. Since these elements have a loss associated with them, they need to be taken into account in the overall output power of the transmitter. These losses directly reduce the signal level or the power output of the transmitter. EIRP provides a way to compare different transmitters and is given by: EIRP = $P_t - L_{tt} + G_t$; where P_t is the transmitter power in dBm, L_{tt} is the total losses from the PA to the antenna in dB (coaxial or waveguide line losses, switchers, circulators, antenna connections), G_t is the transmitter antenna gain in dB referenced to an isotropic antenna, see Figure 9.1. It may be noted that using larger diameter cables or higher quality cables can reduce the loss, which is a trade-off with cost. Another way to reduce the loss between the transmitter and the antenna is to locate the transmitter PA as close to the antenna as possible. This will reduce the length of the cable, which reduces the overall losses in the transmitter.

The amplifier that is used in the front-end of the receiver is a *low-noise amplifier* (LNA). The amplifier is referred to as a LNA because it is designed to have very low noise figure (NF). This is important for having overall NF of the system to be low, since it is determined mainly by the first amplifier stage in the receiver, as per Friss' formula for cascaded amplifier (mentioned in Section 6.1.3 of Chapter 6). *The receiver line losses* from the antenna to the LNA is given by L_{tr} = total losses in dB (due to cables, switches, circulators, limiters, and filters), see Figure 9.2, and thus *receiver gain before* LNA = $(G_r - L_{tr} - L_i) = R_{\text{gain}}$; where L_i is called the *implementation losses*. When implementing a receiver in hardware, there are several components that do not behave ideally. Losses associated with these devices degrade the signal-to-noise ratio (SNR), i.e., S/N. One of the contributors to this loss is the phase-noise or the time jitter of all of the oscillators and potential noise sources in the system, including the up-conversion and down-conversion oscillators, code and carrier tracking oscillators, phase-locked loops, matched filters, and analogue-to-digital converters. Another source of implementation loss is the detector process, including nonideal components and quantization errors. These losses directly affect the receiver's performance. Implementation losses (L_i) are included to account for the departure from the ideal design due to hardware implementation.

9.3 ISSUES RELATED TO RF AND MICROWAVE DESIGN

9.3.1 Designer's Considerations for RF/Microwave Transmitter and Receiver Design

There are a number of parameters that need to be taken care of by the designer with respect to the transmitter and receiver while designing an RF/microwave system. For transmitter, the parameters of concern are transmitter turn-on time, load-pull, spurious outputs, S/N, hum and noise, adjacent channel power, frequency stability, intermodulation distortion etc. While for receiver, designers need to take care of receiver sensitivity, receiver selectivity, spurious response rejection, intermodulation rejection, receiver self-quieting, and so forth. In the following, we will discuss in brief each of these parameters of RF/microwave transmitter and receiver.

9.3.1.1 Transmitter Parameters for Consideration

The function of a transmitter is to amplify the RF/microwae carrier modulated with the desired intelligence signal, adding a minimum of distortion to the encoded information. The specifications that quantify transmitter performance are described below.

A fundamental communication system parameter is the *power output*, whose definition depends on the modulation method used. PM and FM

systems use root mean square (rms) power, while AM system use peak envelop power. Continuous wave power or intermittent duty rating of pulse power will affect the type of thermal management required to control internal temperatures of the system.

Transmitter turn-on time, defined as the time for power output to reach 90% of its rated output, is important for digital communication systems and must be short in order not to limit system throughput, but not so short as to cause transient energy on adjacent channels. The amount of energy generated by rapid transmitter turn-on and turn-off of the transmitter power can be estimated by applying Fourier Transform theory to such signals.

Load-pull refers to the tendency of the transmitter frequency to change as the power output builds up. The said shift in frequency may be as a result of changing impedance either at antenna or internal to the transmitter. As the power output ramps up, the device impedances change; when the effect of these impedance changes propagates backward to the oscillator, its load impedance changes, causing its frequency shift. The cure for this problem is to provide enough buffering to isolate the oscillator from any load impedance variations. If significant amount of load-pull is present, then turn-on time of transmitter mentioned above refers to the time required for the frequency to settle to some fraction of the channel spacing used.

Spurious outputs are most frequently harmonically related to the main carrier, but other spurious emissions are also common in transmitters whose signal is generated by frequency synthesizers or by mixing process. Spurious outputs are usually 70 dB to 90 dB below the main carrier; some regulatory requirements specify an absolute maximum power allowed at any frequency other than carrier.

Signal-to-noise ratio of transmitter is described by *S/N, Hum*, and *Noise*. The *S/N* is usually limited by wideband noise far in excess of thermal noise produced in the transmitter circuits. Power supply hum, noise pickup, and nonlinearities in the modulation circuitry can often limit the ultimate S/N.

Adjacent channel power can result from high SSB phase-noise of the transmitter's oscillator. It can be produced by the modulation method itself or result from rapid transmitter turn-on and turn-off. In narrow band systems, the amount of adjacent channel power integrated over the corresponding receiver's IF bandwidth is usually required to be more than 50 dB below carrier power. In wideband system, 80 dB can be achieved.

Frequency stability is especially important in narrowband systems and is a closely regulated and monitored transmitter parameter.

Intermodulation distortion (IM) or *Spur* is produced when strong external signals are picked-up by the transmitter's antenna and the mixing products are then retransmitted, causing interference in other systems. One or more circulators can be placed between the last active stage and the antenna to reduce transmitter IM.

9.3.1.2 Receiver Parameters for Consideration

The purpose of an RF/microwave receiver is to process the incoming electromagnetic energy into useful information, adding a minimum distortion. How well a receiver performs its purpose is a function of the system design, its internal circuitry, and its working environment. The acceptable amount and type of introduced distortion vary with the application. Television system should maintain a signal-to-noise (S/N) ratio greater than 40dB and are much less tolerant of group delay distortion than analogue FM voice systems, for example. Receiver performance specifications in common usage are discussed below.

Receiver sensitivity, a fundamental property that intimately affects system performance, quantifies the ability of a receiver to respond to weak signal levels. The requirements of analogue receivers are maximum RF level to ensure a certain demodulated S/N ratio, while digital receivers use minimum possible bit error rate (BER) at certain RF level as measure of performance, typically 10^{-9} for telecommunication applications while 10^{-13} is a more appropriate minimum BER for digital transmission. The receiver sensitivity specification and its measurement rely on the assumption that thermal noise is the only limiting factor. However, this assumption may not be valid for systems operating at lower RF frequencies. In such systems, the input S/N ratio to achieve the required demodulated S/N output is a more useful parameter, as long as the ambient nonthermal noise level can be measured or estimated. Nonthermal and especially man-made noise (caused by automobile and aircraft ignition, electric motors, fluorescent lights etc., which is vertically polarized and falls between 1 – 600 MHz) do not have well-defined properties; the reliability of a receiver link in such systems is best described in statistical terms; for example, the output S/N will be better than 20 dB for 95% of the time at a certain location. The actual receiver sensitivity number is correlated to the square root of the receiver's bandwidth and range from 0.3 μV for narrowband systems to 0.5 mV in television receivers.

The ability of a receiver to reject unwanted signals on adjacent channel frequencies is known as *receiver selectivity*. Receiver selectivity is a parameter that quantifies the tendency of a receiver to respond to channels adjacent to the desired reception channel. Because international regulations are gradually moving to narrower channel spacings, receiver sensitivity assumes greater importance, since it frequently limits system performance and places restrictions on frequency allocation and system utilization. This specification, which ranges from 70 dB to 90 dB is sometimes so difficult to achieve that many systems do not allow simultaneous active adjacent channels in the same geographical area on the same cable system.

Receiver spurious responses are frequencies that are different from the desired receive frequency, yet that still produce demodulated output in the receiver. This situation is obviously undesirable and particularly troublesome

in receiver capable of tuning over a wide frequency range because the RF filters need to be wide to accommodate the wide frequency coverage. *Spurious response rejection* is necessary, especially in superheterodyne receivers that have the potential for responding to frequencies other than desired channel. This tendency needs to be minimized by the proper choice of the intermediate frequency (IF) and by use of the RF filters. Spurious response rejection of 70 dB to 100 dB can be achieved in practical receivers.

Intermodulation rejection measures the receiver's tendency to generate its own on-channel interference from two or more strong off-channel signals. Realization of 70 dB is easily achievable, while 90 dB is considered exceptional. Good receiver sensitivity and high intermodulation (IM) rejection are usually contradictory requirements, and some amount of compromise may be required between the two. Details of intercept points that measures a circuit or systems linearity and allows to calculate distortion or IM product has been given in section 9.3.3.3 of this chapter.

Receiver Self-quieting is the inability of a receiver to receive a weak external signal, because the detector has already been captured by some internally generated signal. In FM receivers, this phenomenon is manifested as reduced amplitude of squelch noise on a particular channel; the receiver noise is "quieter" there, hence the name. Thus, receiver self-quieting refers to reduced receiver sensitivity on some channels due to internally generated signals thus preventing or inhibiting the reception of desired, weak signal, and the RF designer should be aware of this.

Other receiver parameters, such as distortion of the demodulated signal, S/N ratio for strong signals (also referred to as the hum and noise), frequency stability, cochannel rejection, cross-modulation, radiated emissions, and susceptibility to high RF levels, are also prominent in many regulatory requirements because of their importance in dense carrier environments.

Environmental conditions, including operating temperature range, humidity, weatherproofing, and shock and vibration requirements can impose severe limitations on the component selection and mechanical packaging.

9.3.2 Link Budget

The link budget is a method used to determine the necessary parameters for successful transmission of a signal from a transmitter to a receiver. The term "link" refers to linking or connecting the transmitter to the receiver, which is done by sending out RF waves through free-space. The term "budget" refers to the allocation of RF power, gains, and losses and tracks both the signal and the noise levels throughout the entire system, including the link between the transmitter and the receiver. The main items that are included in the budget are the required power output level from the transmitter power amplifier (PA), the gains and losses throughout the system and link, and the

signal-to-noise ratio (SNR) for reliable detection; the ratio of bit energy to noise spectral density (E_b/N_o) to produce the desired bit error rate (BER); or the probability of detection and probability of false alarm at the receiver for pulse radar systems. Therefore, when certain parameters are known or selected, the link budget allows the system designer to calculate unknown parameters.

Several of the link budget parameters are given or chosen, as there are many variables and tradeoffs in the design of a transceiver, and each one needs to be evaluated for each system design. For example, there are tradeoffs between the power output required from the PA and the size of the antenna. The larger the antenna (producing more gain), the less power is required from the PA. However, the cost and size of the antenna may be too great for the given application. On the other hand, the cost and size of the PA increase as the power output increases, which may be the limiting factor. If the power output requirement is large enough, a solid-state amplifier may not be adequate, and therefore, a traveling-wave tube (TWT) amplifier may be needed. The TWT amplifier requires a special high-voltage power supply, which generally increases the size and cost. By making these kinds of tradeoff studies, an optimum data link solution can be designed for a specific application.

Before starting the link budget, all fixed or specified information concerning the transceiver needs to be examined to determine which parameters to calculate in the link budget. These concessions need to be evaluated before the link budget is performed and then must be re-evaluated to ensure that the right decisions have been made. Proper transceiver design is critical in the sense that the cost and performance of a data link must be optimized. To provide the optimal design for the transceiver, a link budget is used to allocate the gains and losses in the link and to perform trade-offs of various parts of the system. The link budget also uses the required SNR or the (E_b/N_o) for a given probability of error. These required levels are derived by using probability of error curves for a certain type of modulation. The parameters for a link budget and the probability of error curves are available in published literature. Generally, since there are both known and unknown variances in the link budget, a link budget will provide an additional SNR or E_b/N_o, which is referred to as the link margin.

Link margin for analogue communication systems =
$\text{SNR (calculated)}_{\text{dB}} - \text{SNR (required)}_{\text{dB}}$
Link margin for digital communication systems =
$(E_b/N_o)_{\text{(calculated)dB}} - (E_b/N_o)_{\text{(required)dB}}$

This link margin is used to provide a margin of error in the analysis, hardware implementation, and other factors that can affect the desired performance.

9.3.3 Noise and Distortion

9.3.3.1 Introduction

Measures of noise, distortion, and dynamic range are system metrics describing the sensitivity of active systems and modules. Noise and nonlinear distortion set the bounds on the range of signals that can be processed by an RF circuit. Noise (i.e., random fluctuations of voltage and current) establishes the minimum detectable signal, while nonlinear distortion sets the level of the largest signal from which information can be reliably extracted or which can be transmitted without affecting other systems.

9.3.3.2 Noise

Various types of noise are important in electronic circuits, and these types range from noise, such as thermal noise, that is very well understood, to noise that has been observed and seriously affect the performance of RF/microwave circuits but not well understood. An example of the latter is phase-noise on oscillators which manifests itself as random fluctuations of the phase or frequency of an oscillation signal. The impact of noise on an RF/microwave system is described in the frequency domain whereas the physical origins of noise in the real world is in the time domain. The impact of noise in digital systems, in general, is in the time domain resulting in bit errors of a received, demodulated, and processed digitally modulated signal.

Noise in RF and microwave systems includes noise from the environment as well as noise generated within the circuitry itself. Noise from the environment can have galactic origins, when it is known as cosmic background noise, from black-body radiation, or can be artificially generated noise. In cellular communication systems, the major source of interference is from other phones and base stations in the cellular system. Provided this is uniformly random over the communication band, it can be treated as random noise. Uniformly random noise (i.e., white noise) can be modelled by a resistor held at what is called the noise temperature.

A few words about phase-noise that we come across in RF/microwave design will be worthwhile at this point.

An ideal oscillator is supposed to generate a pure sine wave given by

$$V(t) = V_0 \sin(2\pi f_0 t) \tag{9.1a}$$

where V_0 is the nominal amplitude of the signal and $2\pi f_0 t$ is the linearly growing phase component.

However, due to the presence of noise caused by inevitable reasons, the signal from a practical oscillator is better modelled as

$$V(t) = [V_0 + a(t)]\sin[2\pi f_0 t + \varphi(t)] \tag{9.1b}$$

where $a(t)$ relates to the amplitude fluctuation while $\varphi(t)$ is the fluctuating phase term caused by the presence of noise.

It may be noted that amplitude fluctuation caused by noise do not affect the zero-crossing, while the phase fluctuation does not affect the amplitude of the signal peaks. Well-designed oscillators have small amplitude fluctuations, if any, can be removed by using automatic level control (ALC) systems or by passing the signal through a limiting amplifier (as its output is determined only by the zero-crossing of the signal). Amplitude noise is also rejected to some extent by mixers used in radio systems.

There are two types of fluctuating phase term: deterministic and random. The deterministic ones originate from power line frequency, vibration frequency, and mixer products etc., and appear as distinct components in the power spectral density plot and are commonly known as spurious signals (see Figure 9.3). The random phase fluctuations originate from thermal noise, shot noise, and flicker noise. The random phase fluctuation (and hence eventual frequency fluctuation, since frequency and phase are related to each other) may be long-term or short-term in nature. The long-term frequency fluctuation is specified in terms of *frequency drift* (expressed in parts per million per hour, day, month, or year), while the short-term (may be over a span of few seconds or minutes) frequency fluctuation is characterized with *phase-noise*.

Phase-noise basically causes phase modulation of carrier. In time domain, these results in phase off-set of the sinusoidal signal of the carrier according to the modulation caused by the noise signal giving rise to a "time jitter" in the zero-crossing of the carrier sinusoid (see Figure 9.4a). The timing jitter can become significantly deleterious in some applications, for example, if the signal with time jitter is used as data clock in digital transmission systems

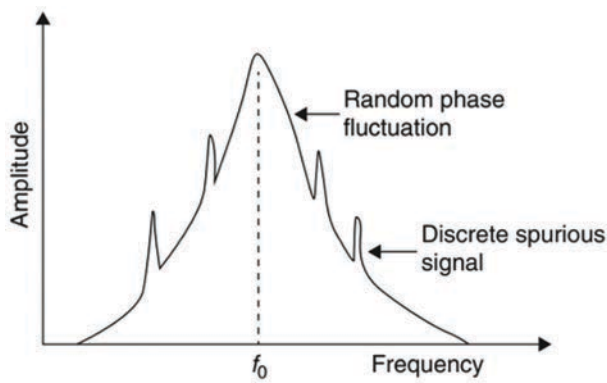


Figure 9.3 Power density spectrum showing the deterministic and random phase fluctuation giving rise to spurious signals and phase-noise.

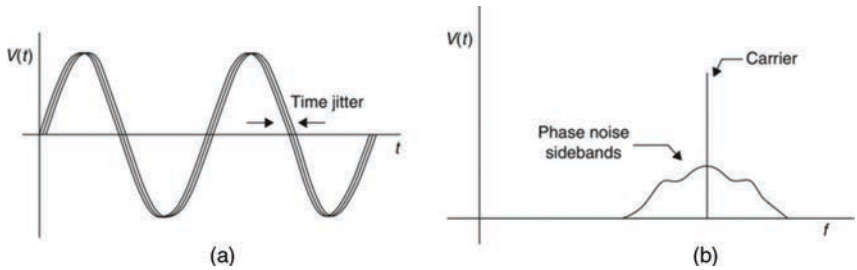


Figure 9.4 (a) Time jitter of a sinusoidal signal due to phase-noise, and (b) Phase-noise sideband spread around the carrier.

the timing jitter may cause erroneous data sampling. Normally, phase-noise results in low-level phase modulation, and thus, the equivalent time jitter is very small to be detected in time domain (i.e., in an oscilloscope). Phase-noise characterization is normally best suited in frequency domain (i.e., with a spectrum analyzer) where the phase-noise shows up as noise sidebands on the carrier (see Figure 9.4b). Phase-noise can have a major impact on radio system performance. Minute amount of phase-noise in a transmitter signal might cause significant interference to other services, while the phase-noise in a receiver's local oscillator can severely reduce the receiver selectivity or cause other undesirable effects.

The spectral purity of high-frequency RF oscillators at microwave and millimeter-wave frequency thus needs to be characterized in terms of phase-noise. The National Bureau of Standards defines Single Side Band (SSB) phase-noise ($L_{\varphi}(f)$) as the ratio of power (P_{ssb}) in one phase modulation sideband per Hertz bandwidth, at an off-set of f Hertz from the carrier, to the total signal power in the carrier (P_{sc}) (see Figure 9.5a). Mathematically,

$$L_{\phi}(f) = \frac{P_{sc}}{P_{ssb}} \tag{9.1c}$$

The SSB phase-noise is usually expressed logarithmically, that is:

$$L_{\phi}(f)_{dB} = 10\log_{10}[L_{\phi}(f)] \tag{9.1d}$$

The unit of phase-noise is dBc/Hz. In practice, phase-noise is depicted in spectrum analyzer (in frequency domain) as how many dB down from the carrier signal peak is the selected sideband phase modulated signal at the specified frequency off-set f from the carrier frequency f_{c} (see Figure 9.5b); in

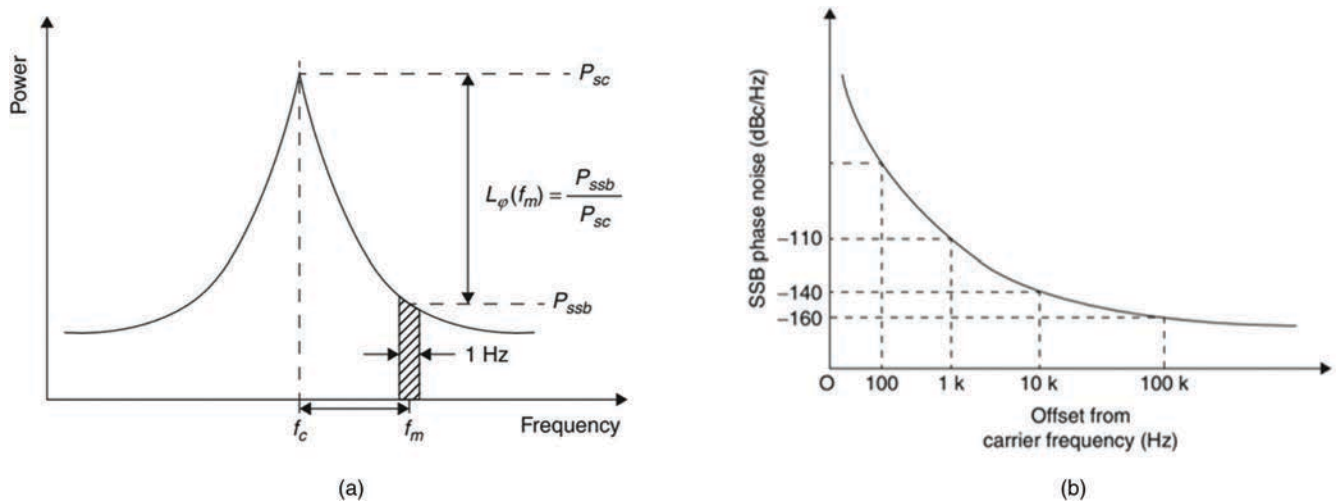


Figure 9.5 (a) SSB phase-noise with respect to carrier, and (b) SSB phase-noise representation in logarithmic scale.



Figure 9.6 Typical phase-noise plot of an oscillator in spectrum analyzer. (Courtesy: SAMEER Kolkata Centre, India).

which frequency is shown in the logarithmic scale. It may be observed that for a given oscillator (say) better phase-noise is realizable with larger frequency off-set, however, when close to the carrier the phase-noise degrades.

Phase-noise of an oscillator (i.e., the phase instability of an oscillator caused by noise) can, in general, be reduced by increasing the quality factor Q of the oscillator system. This is the reason why dielectric resonator oscillator (DRO) using dielectric resonator (DR) with comparatively large Q value as resonant circuit for the oscillator has better phase-noise (of the order of –110 dBc/Hz to –120 dBc/Hz) compared to normal waveguide cavity type microwave oscillators (of the order of –80 dBc/Hz to –90 dBc/Hz) at a frequency off-set of 10 kHz from carrier at an operating frequency of 10GHz. Phase-noise of an oscillator can be characterized with spectrum analyzer (fundamentals of spectrum analyzer is given in Appendix VII); a typical phase-noise plot of oscillator in spectrum analyzer is shown in Figure 9.6.

9.3.3.3 Nonlinearity and Distortion

Active devices like diodes and transistors have operational characteristics that exhibit both linear and nonlinear regions of operation. Depending upon the application, the active device may be biased in the linear or nonlinear region of the characteristics—amplification, in general, is in linear region while detection/mixing (frequency conversion) takes place in the nonlinear region. The nonlinearity in the device characteristics of diode and transistors, however, would inevitably introduce distortions like gain compression and spurious frequency components due to intermodulation/cross-modulation etc., in amplifier and mixer performance.

The output voltage response of a nonlinear device may be modelled with a Taylor series:

$$v_0 = a_0 + a_1 v_1 + a_2 v_i^2 + a_3 v_i^3 + \ldots\ldots \tag{9.2a}$$

where v_i is the input voltage signal to the device. The Taylor coefficients in Eq. (9.2a) are given by

$$a_0 = v_0(0),\ \ a_1 = \left.\frac{dv_0}{dv_i}\right|_{v_i=0},\ a_2 = \left.\frac{d^2 v_0}{dv_i^2}\right|_{v_i=0};$$
$$\text{and other higher order terms} \tag{9.2b}$$

The constant term (with coefficient a_0) in Equation (9.2a) leads to rectification (AC to DC). The linear term (with coefficient a_1) is useful for amplification while the second-order term (with coefficient a_2) can be used for frequency conversion (mixing).

<u>*Gain compression*</u>: Let v_i be a single frequency sinusoid given by: $v_i = V_0 \cos \omega_0 t$; substituting this in Equation (9.2a), we get

$$\begin{aligned} v_0 &= a_0 + a_1 V_0 \cos\omega_0 t + a_2 V_0^2 \cos^2 \omega_0 t + a_3 V_0^3 \cos^3 \omega_0 t + \ldots\ldots \\ &= \left(a_0 + 0.5 a_2 V_0^2\right) + \left(a_1 V_0 + 0.75 a_3 V_0^3\right) \\ &\quad \cos\omega_0 t + 0.5 a_2 V_0^2 \cos 2\omega_0 t + 0.25 a_3 V_0^3 \cos 3\omega_0 t + \end{aligned} \tag{9.2c}$$

Thus, the voltage gain of the signal component at frequency ω_0 is given by

$$G_v = \frac{v_i\big|_{\omega_0}}{v_0\big|_{\omega_0}} = \frac{a_1 V_0 + 0.75 a_3 V_0^3}{V_0} = a_1 + 0.75 a_3 V_0^2 \tag{9.2d}$$

where terms only upto third-order has been retained.

It may be observed from Equation (9.2d) that the voltage gain is equal to the coefficient of the linear term, a_1; to which an additional term is also added. For most practical amplifiers, a_3 has opposite sign to that of a_1 and hence for large values of V_0 the output of the amplifier reduces and departs from linear behavior as shown in Figure 9.7a. This nonlinear behavior of output with respect to the input of an amplifier showing saturation of the output is known as *gain compression*.

Figure 9.7a plots the power of the output sinewave against the power of the input sinewave. The plot is called the AM-to-AM (AM-AM) response (the amplitude-dependent amplitude response) of the amplifier. The AM–AM characteristic is linear at low input powers, but eventually the gain

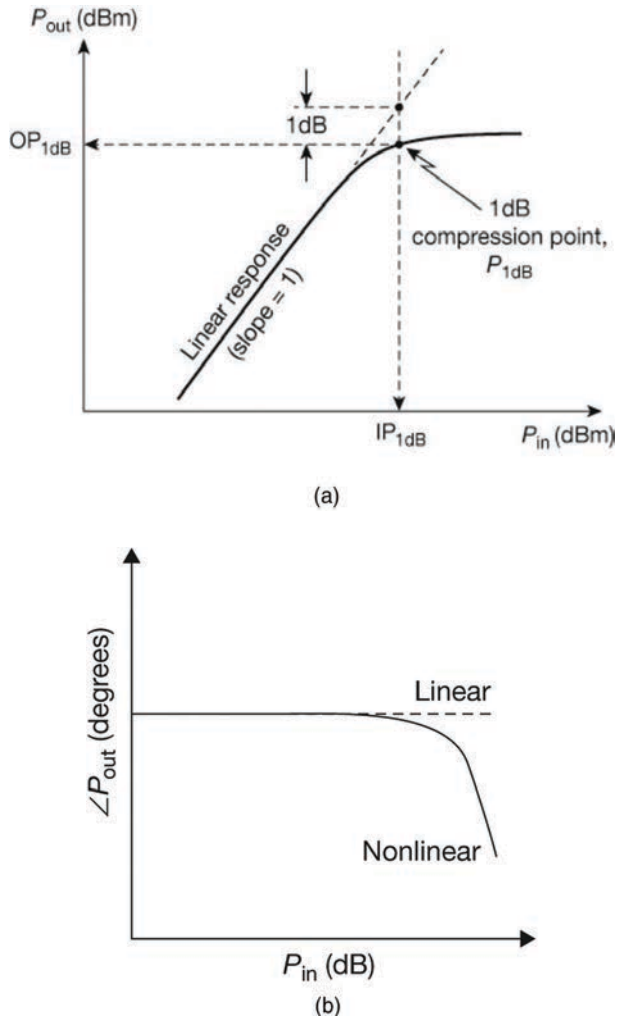


Figure 9.7 Nonlinear effects introduced by RF hardware: (a) Amplitude-dependent amplitude (AM–AM) distortion showing 1-dB compression point of an amplifier, and (b) Amplitude-dependent phase (AM–PM) distortion.

reduces (compresses) and the output power drops below the linear extension of the small-signal response. In Figure 9.7a, the 1-*dB gain compression point* or just 1-*dB compression point* is the point where the difference between the extrapolated linear response exceeds the actual gain by 1 dB. P_{1dB} is the output power at the 1-dB compression point, and it can be stated either in terms of input power (IP_{1dB}), as in case of mixer, or output power (OP_{1dB}), as in case of amplifier. A gain compression of 1dB for an amplifier

corresponds to a signal voltage amplitude reduction of 11% for amplifier, so P_{1dB} indicates a small but significant reduction in the linear gain of an amplifier. For mixer, 1-dB gain compression point is about 15dB below IP_{1dB}.

The small-signal operation of an amplifier takes place in the linear region while for large-signal operation it enters into the 1 dB compression point region. Amplifiers are normally operated with input power limited to 1-dB compression point (i.e., with P_{1dB} as in case of LNA of receivers and also gain amplifiers of transmitters). However, driver stage amplifiers and especially the final stage power amplifiers of transmitters are operated with significant amount of nonlinearity, where the amplifier operation is with P_{2dB} or even with P_{3dB}.

At large powers, the parasitic capacitances of the transistors in the amplifier vary the phase of the output signal, and hence phase distortion occurs. Figure 9.7(b) shows what is called the AM-to-PM (AM-PM) characteristic which is a measure of the amplitude-dependent phase distortion. For a digitally modulated signal, phase distortion would be significant if it caused the wrong constellation point to be selected in a receiver.

Intermodulation distortion: Intermodulation distortion is a property of all systems that exhibit a nonlinear transfer function, such as amplitude compression at sufficiently high level. This is the primary cause of distortion introduced in amplifier and mixer performance due to nonlinearity of active devices. All practical amplifiers and mixers are operated with multi-tone (more than one frequency) input signal, which results in intermodulation distortion. For single tone input (with $v_i = V_0 \cos \omega_0 t$), the output will consist of harmonics of the input frequency given by: $n\omega_0$, for n = 0, 1, 2, …. These harmonics would often lie outside the pass-band of the amplifier and has no chance to interfere with the desired signal at ω_0. But when the input signal has two closely spaced frequencies, i.e., for a two-tone input signal of the form: $v_1 = V_0\left(\cos \omega_1 t + \cos \omega_2 t\right)$ the output is given from Eq. (9.2a) as

$$\begin{aligned} v_0 &= a_0 + a_1 V_0\left(\cos \omega_1 t + \cos \omega_2 t\right) + a_2 V_0^2\left(\cos \omega_1 t + \cos \omega_2 t\right)^2 \\ &\quad + a_3 V_0^3\left(\cos \omega_1 t + \cos \omega_2 t\right)^3 \\ &= a_0 + a_1 V_0 \cos \omega_1 t + a_1 V_0 \cos \omega_2 t + 0.5 a_2 V_0^2\left(1 + \cos 2\omega_1 t\right) \\ &\quad + 0.5 a_2 V_0^2\left(1 + \cos 2\omega_2 t\right) \\ &\quad + a_2 V_0^2 \cos\left(\omega_1 - \omega_2\right)t + a_2 V_0^2 \cos\left(\omega_1 + \omega_2\right)t \\ &\quad + a_3 V_0^3\left(0.75 \cos \omega_1 t + 0.25 \cos 3\omega_2 t\right) + a_3 V_0^3 \\ &\quad \left(0.75 \cos \omega_2 t + 0.25 \cos 3\omega_2 t\right) \\ &\quad + a_3 V_0^3\left[1.5 \cos \omega_2 t + 0.75 \cos\left(2\omega_1 - \omega_2\right)t + 0.75 \cos\left(2\omega_1 + \omega_2\right)t\right] \\ &\quad + a_3 V_0^3\left[1.5 \cos \omega_1 t + 0.75 \cos\left(2\omega_2 - \omega_1\right)t + 0.75 \cos\left(2\omega_2 + \omega_1\right)t\right] + \ldots\ldots \end{aligned} \tag{9.2e}$$

Where we have used standard trigonometric identities to expand the initial expressions in order to obtain the final expression in a logical form.

It may be observed that the output spectrum consists of harmonics of the form: $m\omega_1 + n\omega_2$, with $m, n = 0 \pm 1, \pm 2, \pm 3, \ldots..$ These combinations of two input frequencies is known as *intermodulation products* (IMP); the order for a given IMP is given by $|m| + |n|$. On investigating Equation (9.2e) we find that there are four second order IMPs ($2\omega_1$, $2\omega_2$, $\omega_1 \pm \omega_2$) and six third order IMPs ($3\omega_1$, $3\omega_2$, $2\omega_1 \pm \omega_2$, $2\omega_2 \pm \omega_1$). If ω_1 and ω_2 are close, all of the second order IMPs will be far from ω_1 or ω_2 and can easily be prevented from reaching the output of the amplifier/mixer with properly designed filter; which is also true for four of the third order IMPs: $3\omega_1$, $3\omega_2$, $2\omega_1 + \omega_2$, $2\omega_2 + \omega_1$. However, two third-order IMPs with difference terms viz. $(2\omega_1 - \omega_2)$ and $(2\omega_2 - \omega_1)$, generated due to the mixing of the second-order products $2\omega_1$ and $2\omega_2$ with the first-order products ω_1 and ω_2, are the most problematic as they will appear near the input frequencies ω_1 and ω_2, (shown in the bell-shaped grey part in Figure 9.8) and are impossible to filter out from the pass-band of an amplifier or would be very close to the desired IF output of mixers; causing the harmful *third-order intermodulation distortion*.

Thus, if the frequencies ω_1 and ω_2 corresponded to communication channels, then the channels occupied by the third order products will be unavailable for normal communication (they would interfere with actual signals on those channels). For example, in the diagram above, if CH01 and

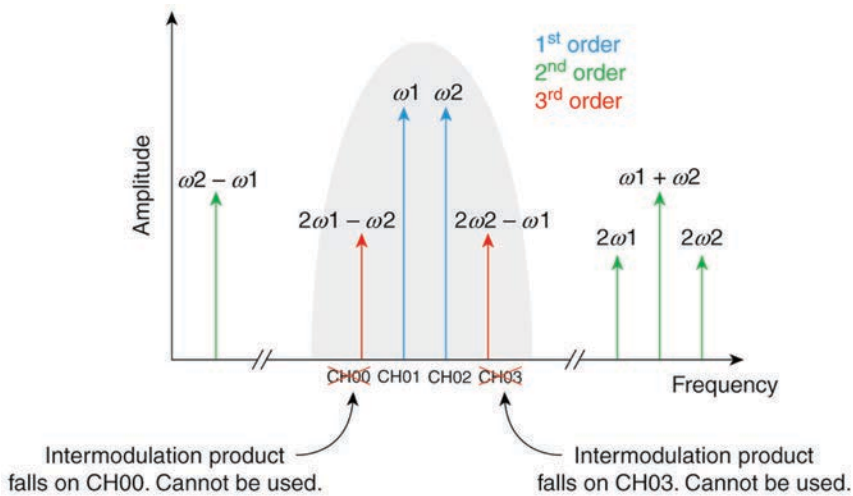


Figure 9.8 Intermodulation product diagram showing the first- and second-order intermodulation products by blue and green while the two problematic third-order intermodulation product is shown by red.

CH02 are the input frequencies, the channels CH00 and CH03 cannot be used as those are corrupted by the problematic third-order intermodulation products. Anyway, "Circuit Design" has a tool at www.cdt21.com/technical_tools/channel-planning/ that calculates when the third-order intermodulation will occur and can also tell the users which channels are safe to use and which are not; in fact channels settings are to be properly chosen to avoid to be corrupted by third-order intermodulation distortion.

Few more points need to be discussed while talking about the IMD. Two-tone *third order intercept point* (IP_3) is a well-established metric for characterizing intermodulation distortion (IMD) in amplifiers and mixers operating under large-signal condition. As was mentioned earlier that the "intercept point" is a measure of a circuit or systems linearity that allows us to calculate distortion or IM product levels from the input signal amplitude. The order of intercept point refers to how fast the amplitude of the distortion products increase with an increase in input power level, see Figure 9.9a. For example, for third-order intercept point (IP_3) the IM product will increase in amplitude by 3dB when the input signal is raised by 1dB. It is seen from Equation (9.2e) that as the input voltage V_0 increases, the voltage associated with the third order IMPs increases as V_0^3 and hence its power ($P \propto V^2$) with respect to the input power. Thus, for lower input power third order IMPs will be very small, but will increase rapidly as the input power increases. Graphical representation of the effect of third-order intercept is shown in Figure 9.9b, where the output power for the first- and third-order products have been plotted in dB (i.e., in logarithmic terms, $\log x^n = n$) scale.

The first-order (linear response) product will have a slope 1 while the third-order (cube response) product will have a slope 3; with both the curves showing compression at higher input power level. Thus, when the curves are extrapolated (shown by dotted lines) they will intersect at a point above the onset of compression. This intersection point, where the power of the first- and third-order product is equal is known as *third-order intercept point*; which may be expressed either in terms of output power (OIP_3), as for an amplifier, or in terms of input power (IIP_3), as in case of a mixer; higher of the two is taken for consideration (since for amplifiers $OIP_3 > IIP_3$ while for mixers $IIP_3 > OIP_3$; $OIP_3 = G.\ IIP_3$, where G is the gain of amplifier or conversion loss of mixer). The third-order intercept point is used as a rule-of-thumb measure to estimate the nonlinearity of an amplifier/mixer; a higher intercept point ensures better linearity in performance. Usually, IP_3 is taken to be 10–15 dB greater than P_{1dB}.

A little more elaboration about the "intercept points (IPs)" would be worthwhile for the understanding of the readers regarding IP and its relation to IMD. The output power (in dB) plotted against the input power is a linear function up to the so-called 1-dB compression point, see Figure 9.7a. The slope of this line is constant with gain; i.e., $P_{out}(\text{dBm}) = P_{in}(\text{dBm}) + G(\text{dB})$.

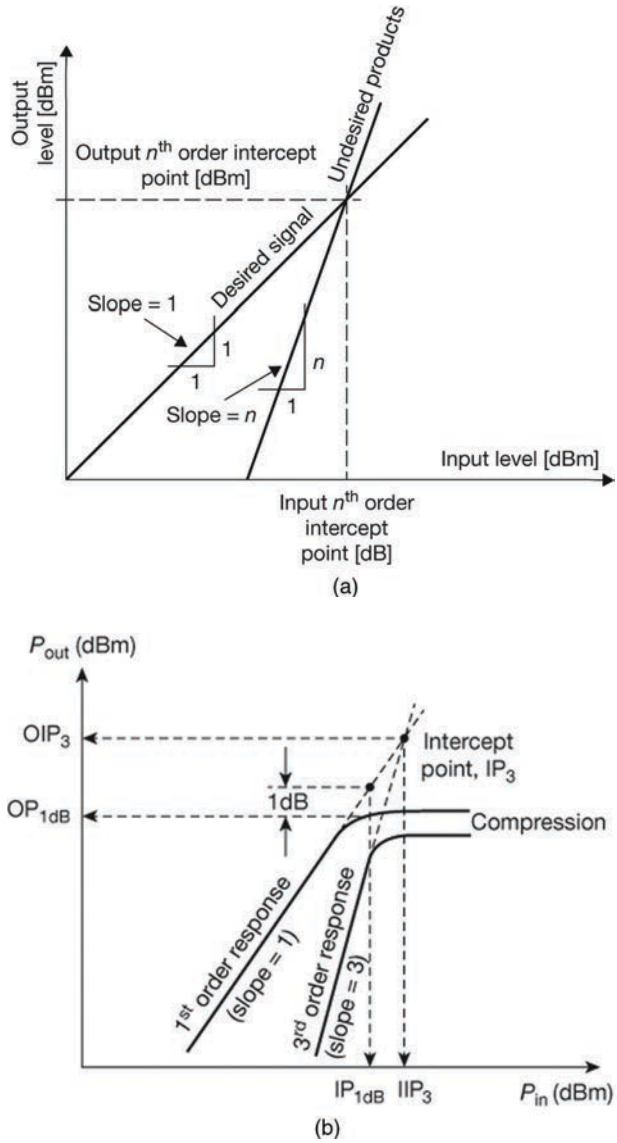


Figure 9.9 (a) Intercept point definition, and (b) Third-order intercept point determination of a nonlinear device.

The slope of the second-order response ($n = 2$), see Figure 9.8a, is equal to twice that of the fundamental; that is, for every 1 dB of increase in output fundamental power, the second-order output signal power increases by 2 dB. The slope of the third-order response ($n = 3$) is equal to three times that of the fundamental; that is, for every 1 dB of increase in output fundamental

power, the third-order output signal power increases by 3 dB. If the fundamental plot is extended linearly beyond the 1-dB compression point, then this line will intersect the second- and third-order lines. The point where the second-order curve crosses the fundamental is called the second-order intercept point. The point where the third-order curve crosses the fundamental is called the third-order intercept point, see Figure 9.8b, and is generally not at the same location as the second-order intercept point. The intercept points are given in dBm. The intermodulation levels can then be calculated for a given signal level.

The second-order intermodulation signal will be two times farther down the power scale from the intercept point as the fundamental. Therefore, the actual power level (in dB) of the second-order intermodulation is: $IM_2 = IP_2 - 2\,(IP_2 - P_{out})$; where IM_2 is the second-order intermodulation power level (in dBm), IP_2 is the second-order intercept point, and P_{out} is the output power of the fundamental signal.

The third-order intermodulation signal will be three times farther down the power scale from the intercept point as the fundamental. Therefore, the actual power level (in dB) of the third-order intermodulation is $IM_3 = IP_3 - 3\,(IP_3 - P_o)$; where IM_3 is the third-order intermodulation power level (in dBm), IP_3 is the third-order intercept point, and P_{out} is the output power of the fundamental signal.

9.4 DESIGN OF RF/MICROWAVE COMMUNICATION TRANSMITTER AND RECEIVER INCLUDING RADAR SYSTEM USING EDA SIMULATION TOOL

We would use here SystemVue EDA simulation tool, a brief note on which is given in Appendix V, that helps in understanding the workflow and signal characterization at every node of the system in the design of RF/microwave systems. It would provide an idea of physical layer design of complex RF system which can be ultimately translated to circuit-level layout design using Keysight's another EDA tool, advanced design system (ADS), and ultimately to translate that to hardware implementation. Anyway, in this book, in the purview of system design, we would restrict our approach for system design with SystemVue EDA simulation tool only to get a flavor of the design approach for complex RF systems that will provide a rudimentary insight of the RF/microwave system design.

Example 9.1 Design the RF block-diagram scheme for a satellite communication transmitter-receiver system with step-by-step signal characterization. The specifications for the desired design of the system are given below.

(I) The RF transmitter has an input signal of 0dBm at 1GHz and is expected to deliver 30 dBm (1 Watt) power at 37 GHz at the output of the transmitter.
(II) The RF receiver will receive the 37 GHz at -70 dBm and will have power output around 35 dBm with NF less than 5 dB.

In both the cases, purity of the transmitted signal and that of the retrieved received signal need to be ensured. Use SystemVue EDA tool for the purpose of this design.

Solution

(I) Design and Characterization of Transmitter

The transmitter block diagram that relates to the design is shown in Figure 9.10 while the signal characterization at every node of the block-diagram design has been shown in Figure 9.11. The signal source provides the 1 GHz signal at 0 dBm power (shown in Figure 9.11a) to the *first* mixer at the IF port (node 1) of Figure 9.10. The *first* local oscillator (LO) generates a power of 18 dBm at 5 GHz, which is feed to a band-pass filter (BPF), refer node 2, for reducing the spurs of the LO. The BPF used has f_L (low) and f_H (high) set to 4.8 GHz and 5.3 GHz, respectively. The insertion loss (IL) of the filter is 0.6 dB. The output of the BPF is fed to the *first* mixer (node 3). Figure 9.11b shows the output spectrum of this BPF and the output spectrum of the *first* mixer is shown in Figure 9.11c. The mixer is translating or up-converting the low frequency signal to a higher frequency with the help of the LO. At the output of this mixer (node 4), another BPF (f_L and f_H are 5.75 GHz and 6.25 GHz, respectively) is being used to reduce the spurs and reject the unwanted signals due to intermodulation. Figure 9.11d shows the spectrum after this BPF at node 5. The *first* amplifier's output (at node 6) following the preceding up-converting is shown in Figure 9.11e. Gain of the amplifier is 7.75 dB and the noise figure (NF) is 3.4 dB. The other parameters like, OP1 dB, OPSAT, OIP3, OIP2 are 18.48 dBm, 20 dBm, 28.97 dBm, 32 dBm, respectively. For increasing the power to a certain level, another, i.e., *second* amplifier is used along with a BPF to further reduce the spurs at the second amplifier's out. The output spectrum of BPF, at node 9, just before further up-converting to the desired frequency, is shown in Figure 9.11f. This signal of 6 GHz is then up-converted to the desired frequency at Ka Band (37 GHz) using LO of 31 GHz. To reduce the spurs from the *second* LO signal a BPF is used at node 10, whose output spectrum at node 11 is shown in Figure 9.11g. The filter's lower and higher cut-off frequency is at 30.1 GHz and 31.8 GHz, respectively. The IL is 1.1 dB.

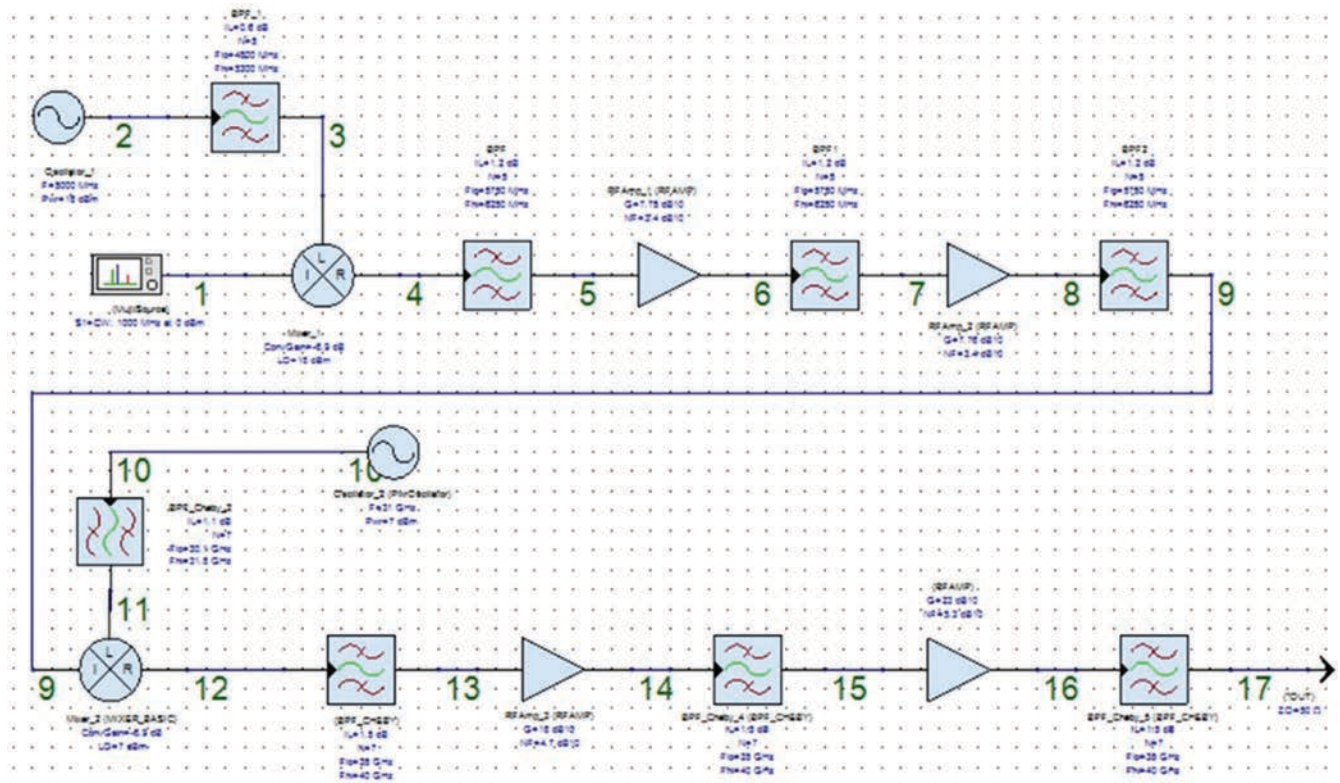
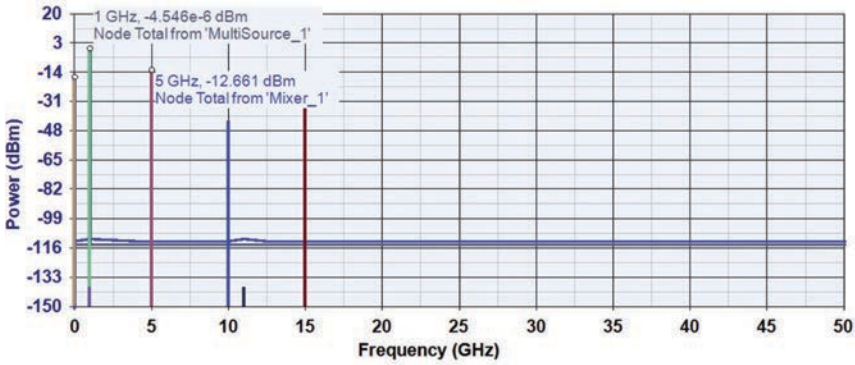
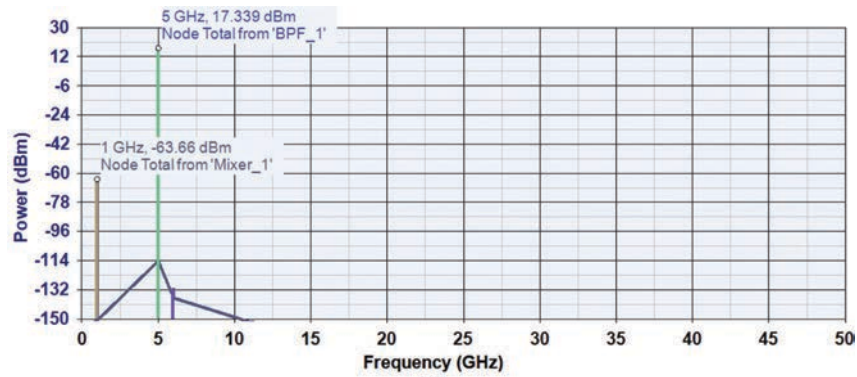


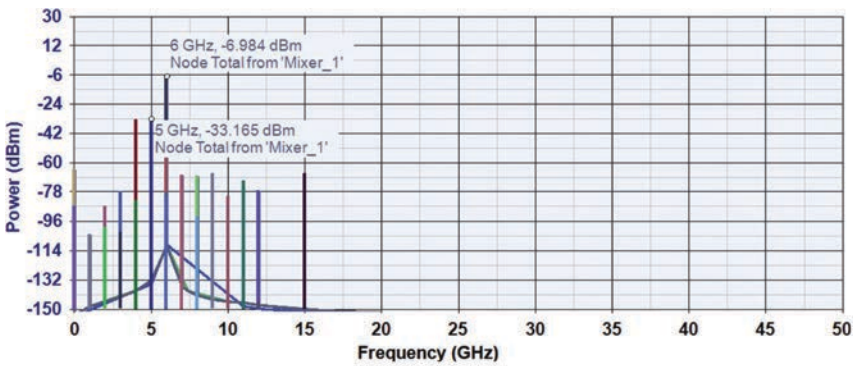
Figure 9.10 Transmitter block diagram scheme for design and characterization.



(a)

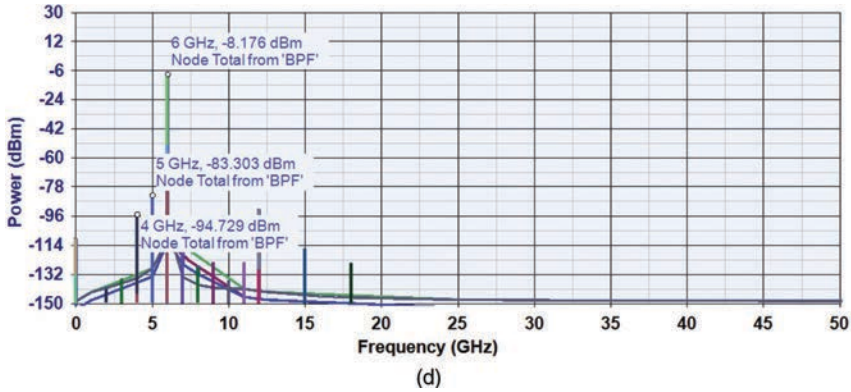


(b)

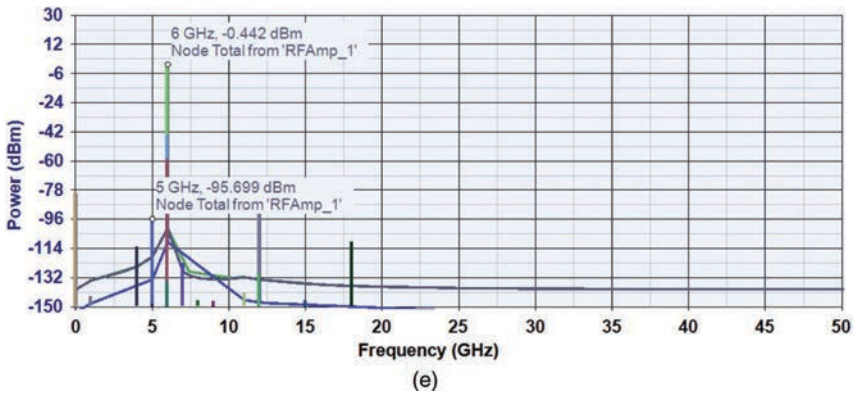


(c)

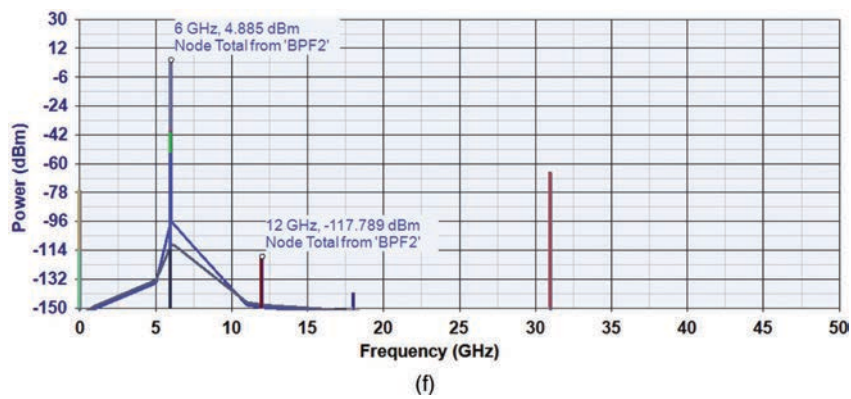
Figure 9.11 Signal output of the transmitter at different nodes as mentioned in the text.



(d)

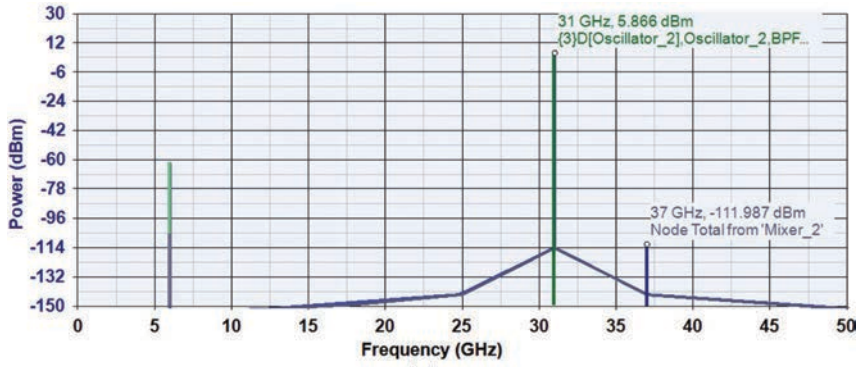


(e)

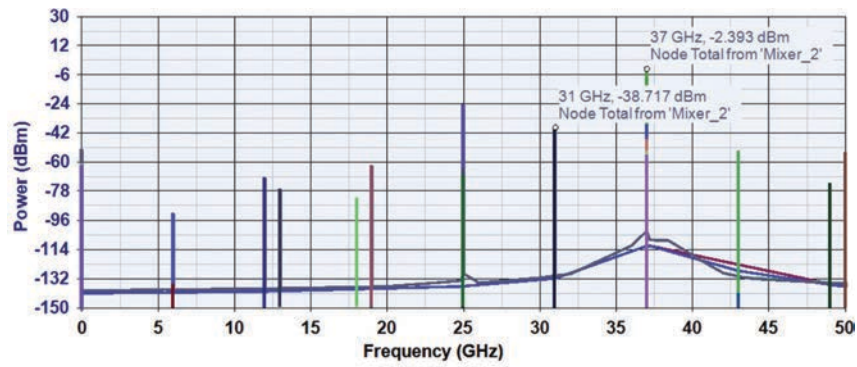


(f)

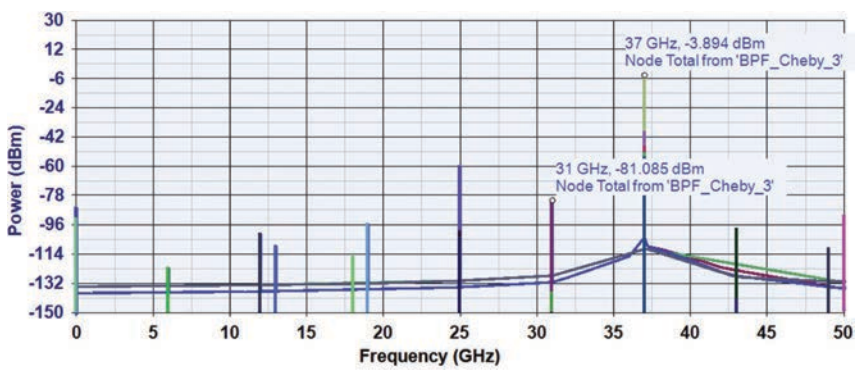
Figure 9.11 (Continued)



(g)

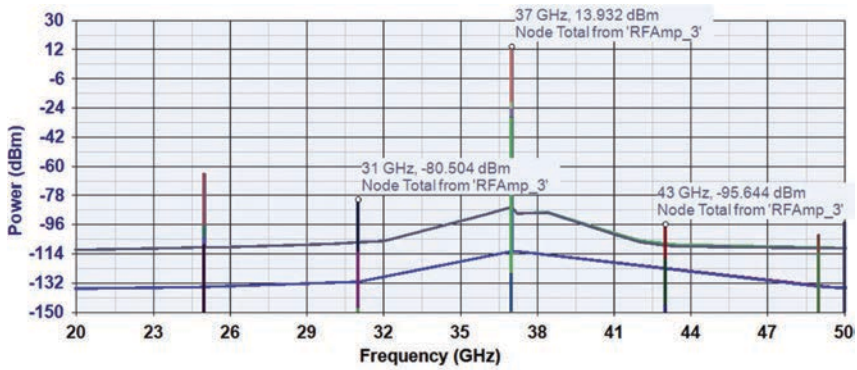


(h)

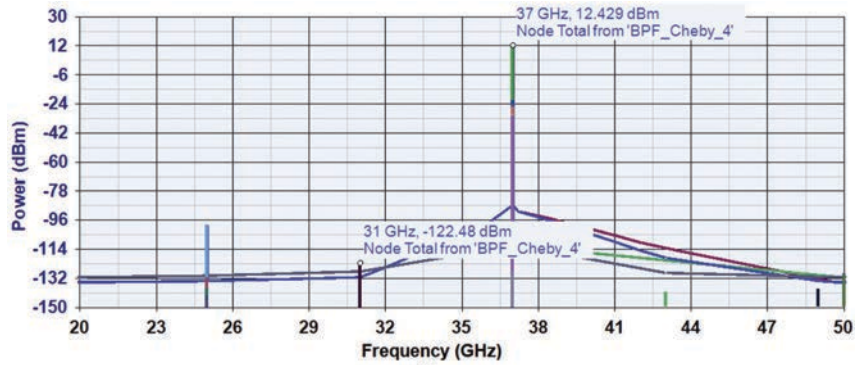


(i)

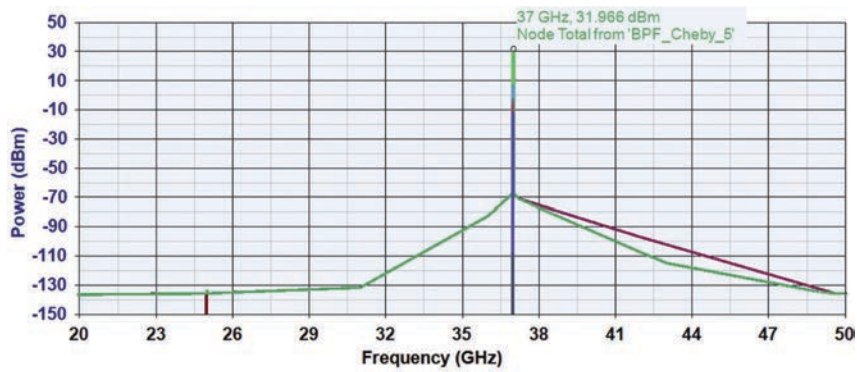
Figure 9.11 (Continued)



(j)



(k)



(l)

Figure 9.11 (Continued)

The output spectrum of the *second* mixer is shown in Figure 9.11h. After up-converting, we need to reduce and reject all the unwanted signals. Just at the output of this *second* mixer at node 12, another BPF is used. Its IL is 1.5 dB and bandwidth is 5 GHz (f_L is 35 GHz and f_H 40 GHz). The output spectrum of this BPF at node 13 is shown in Figure 9.11i. After filtering, another amplifier is used to boost the strength of the signal and the output spectrum (seen at node 14) is shown in Figure 9.11j. The gain of the amplifier is 18 dB and NF is 4.7 dB. OP1 dB, OPSAT, OIP3, OIP2 are 25 dBm, 27 dBm, 34 dBm, and 41 dBm, respectively. To reduce the further spurs from this amplifier's output another BPF is used whose IL is 1.5 dB and f_L is 35 GHz and f_H 40 GHz (spectrum at node 15 is shown in Figure 9.11k). Finally, to increase the signal strength to the desired power level, yet another amplifier is used whose gain is 23 dB and NF is 5.3 dB. OP1 dB, OPSAT, OIP3, OIP2 are 35 dBm, 39 dBm, 48 dBm, and 47 dBm, respectively. Another BPF is used, as usual, at the output of this final amplifier to suppress the unwanted signals and pass only the desired frequency band. The final output spectrum (at node 17) is shown in Figure 9.11l.

Finally, we have got the 37 GHz output signal with approximately 30 dBm (or 1W) output power from the designed transmitter.

II Design and Characterization of the Receiver

The receiver block-diagram that relates to the design is shown in Figure 9.12 while the signal characterization at every node of the block-diagram design has been shown in Figure 9.13. The microwave signal from the transmitter at 37 GHz with –70 dBm power level is received at node 1 (shown by a signal generator in Figure 9.12). At the output spectrum of the *first* BPF, noted at node 2, is shown in Figure 9.13a. This filter has f_L and f_H set to 34.7 GHz and 38.2 GHz respectively with insertion loss (IL) set to 1.5 dB. It has been used to reduce the spur and reject the out-of-band signals. This BPF is known as the pre-selector filter for the received signal. The output of this BPF is fed to a LNA to increase the power of the received signal. Gain of this *first* LNA is set to 18.5 dB and NF is 3 dB. OP1dB, OPSAT, OIP3, OIP 2 are set to 18 dBm, 21 dBm, 29 dBm, and 39 dBm, respectively. The output spectrum of this LNA is shown in Figure 9.13b. To reduce the spur from this LNA output, at node 3, another BPF is used that has f_L and f_H set to 34.7 GHz and 38.2 GHz with IL being 1.5 dB. A *second* LNA is also used, for the signal output, at node 4, from preceding BPF to increase the signal strength which is further filtered out using the next BPF with f_L and f_H set to 34.7 GHz and 38.2 GHz, respectively. A 32 GHz frequency signal is generated by a local

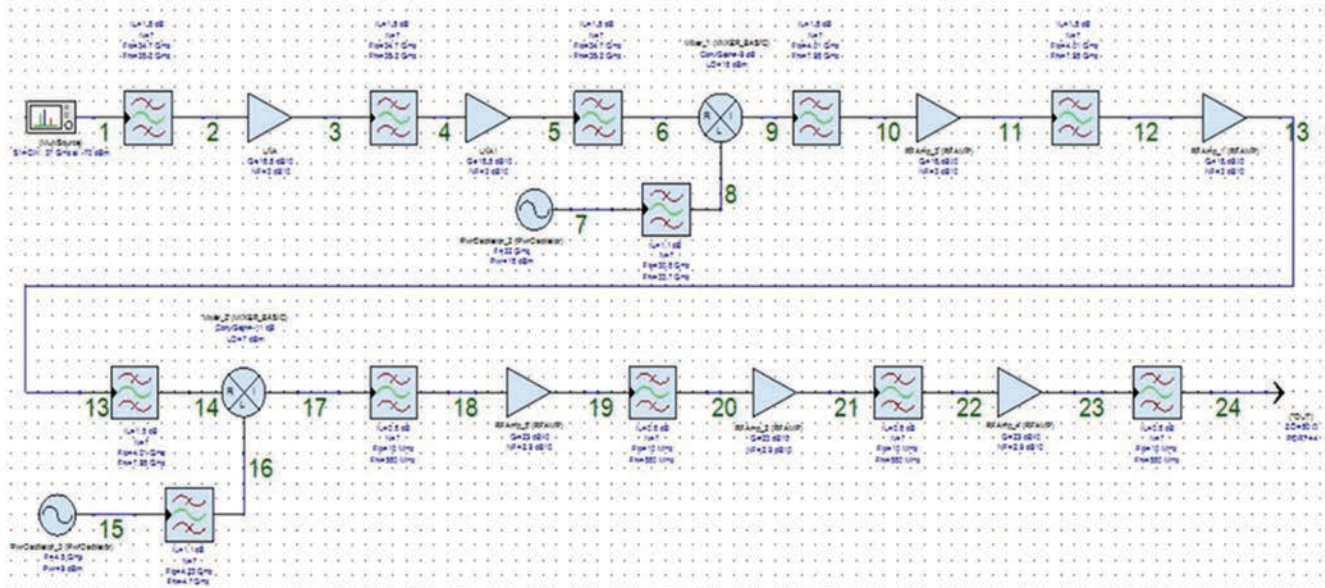


Figure 9.12 Receiver block diagram scheme for design and characterization.

oscillator and power is set to 18 dBm (node 7). To reduce the spur from the LO output, we use a BPF whose output spectrum is shown in Figure 9.13c (refer to node 8). Here the f_L and f_H set to 30.8 GHz and 33.7 GHz with IL set to 1.1 dB. The output of the BPF is fed to the *first* mixer at node 8. To reduce the spurs and intermodulation components of the signal from this mixer, another BPF is used at node 9 whose spectrum is shown in Figure 9.13d. Here the f_L and f_H of the BPF is set to 4.01 GHz and 7.95 GHz and the IL is set to 1.5 dB. Next, a *third* LNA is used to increase the power level. Gain of the amplifier is set to 18 dB and NF is 3 dB. OP1dB, OPSAT, OIP3, OIP 2 are set to 22 dBm, 28 dBm, 32 dBm, and 35 dBm respectively. The output spectrum of this LNA, at node 11, is shown in Figure 9.13e. Again, a BPF is used to reduce the spur of the signal from the amplifier. Here the f_L and f_H are respectively 4.01 GHz and 7.95 GHz and the IL of the BPF is 1.5 dB. For further increasing the signal strength, a *fourth* LNA along with a BPF is used at node 12 and node 13, respectively. The output of this BPF, at node 14, is fed to a *second* mixer which is driven by a *second* local oscillator to perform the final down-conversion to 500 MHz. This LO generates 4.5 GHz frequency with 9 dBm power out (referred to node 15). A BPF follows the LO to reduce the spur from the LO signal. For this BPF the f_L and f_H are 4.23 GHz and 4.7 GHz, respectively, with IL set to 1.1dB. The output spectrum of this BPF, at node 16, is shown in Figure 9.13f. Then the signal is down-converted with the above mentioned *second* mixer. The output spectrum of the mixer, at node 17, is shown in Figure 9.13g. Again, to reduce the spur from the output signal of this mixer another BPF is used at node 17. Here the f_L and f_H are 10 MHz and 580 MHz, respectively, with IL being set to 0.8dB. The output spectrum of this BPF (refer to node 18) is shown in Figure 9.13h. The filter is then followed by another LNA, the *third* LNA. The Gain of this LNA is set to 23 dB and NF is 2.9 dB. OP 1dB, OPSAT, OIP3, and OIP2 are set to 29 dBm, 35 dBm, 49 dBm, and 72 dBm, respectively. The output spectrum of the *third* LNA, at node 19, is shown in Figure 9.13i. The output of this LNA is then fed again to another BPF to reduce the spur level. Here the f_L and f_H of the BPF are 10 MHz and 580 MHz, respectively, with IL as 0.8dB. The spectrum of this BPF output, at node 20, is shown in Figure 9.13j. To increase the power of the signal further, two amplifier is cascaded with BPFs have been used. The output spectrum of the final BPF, at node 24, is shown in Figure 9.13k. The output power of the receiver is approximately 35 dBm, and the NF of the receiver chain is less than 5 dB as shown in Figure 9.13l.

A NOTE: All the components that have been used in the transmitter-receiver design has been realized using commercial products. Every

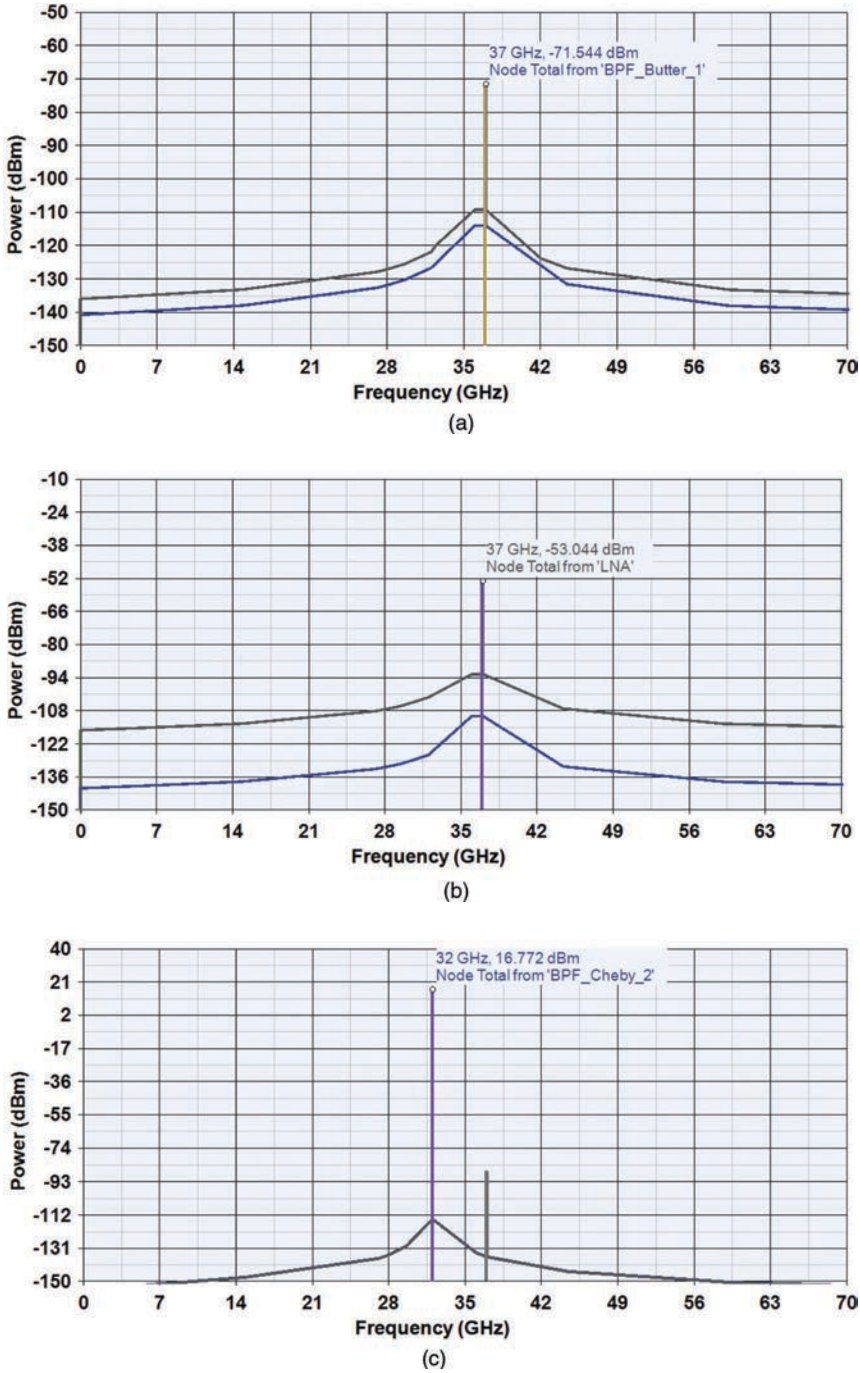
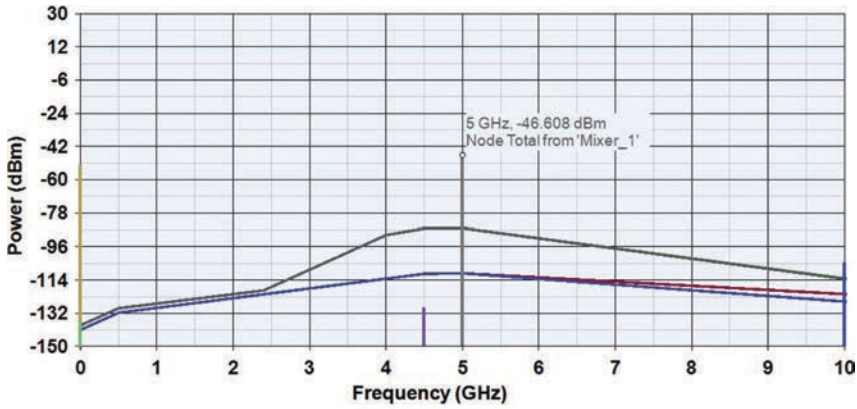
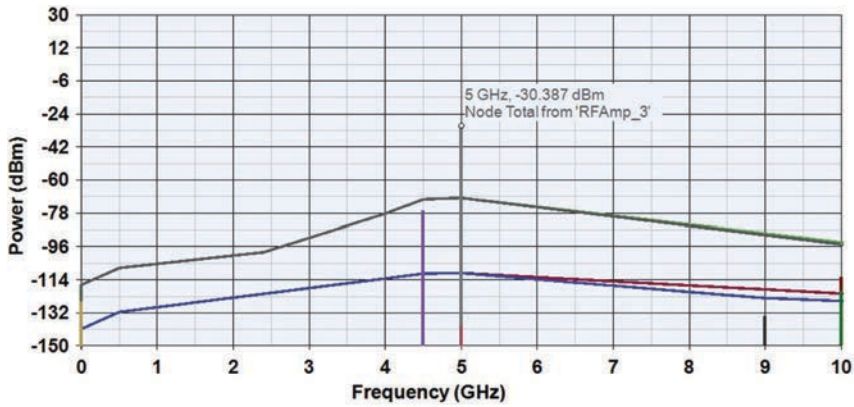


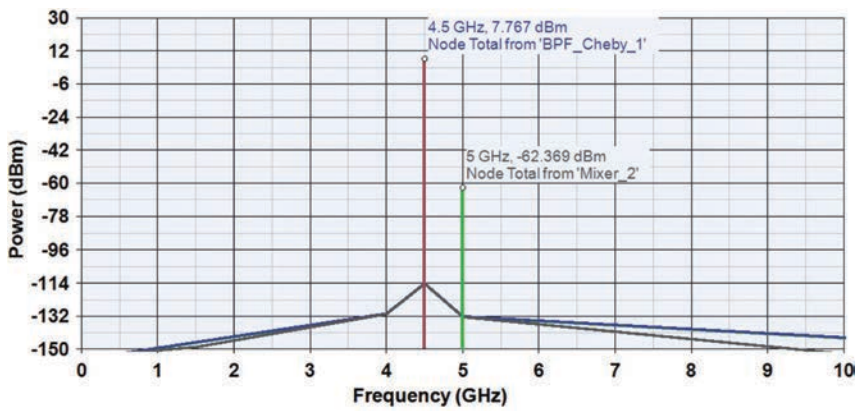
Figure 9.13 Signal output of the receiver at different nodes as mentioned in the text.



(d)

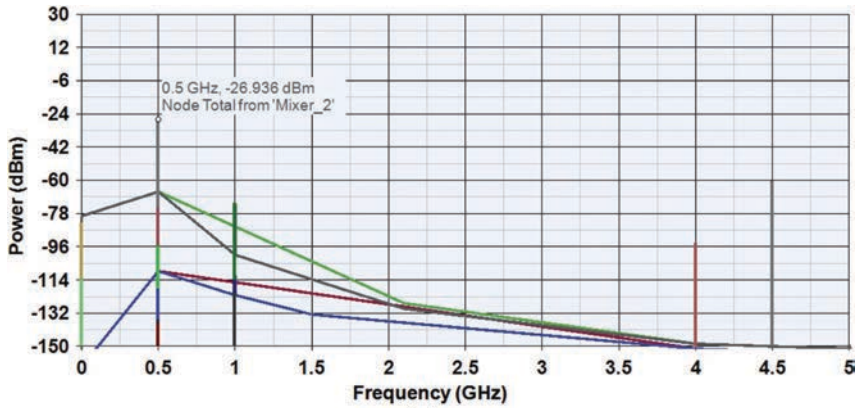


(e)

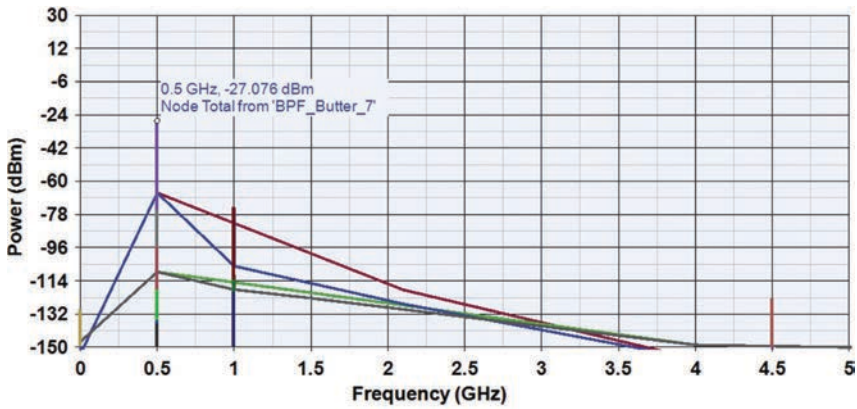


(f)

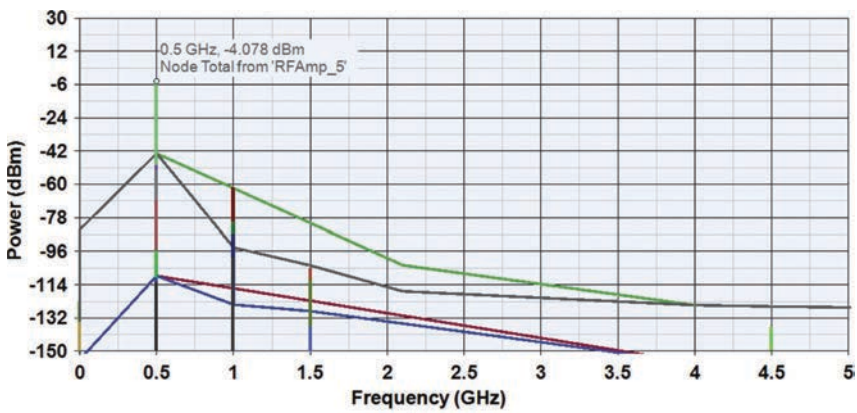
Figure 9.13 (Continued)



(g)

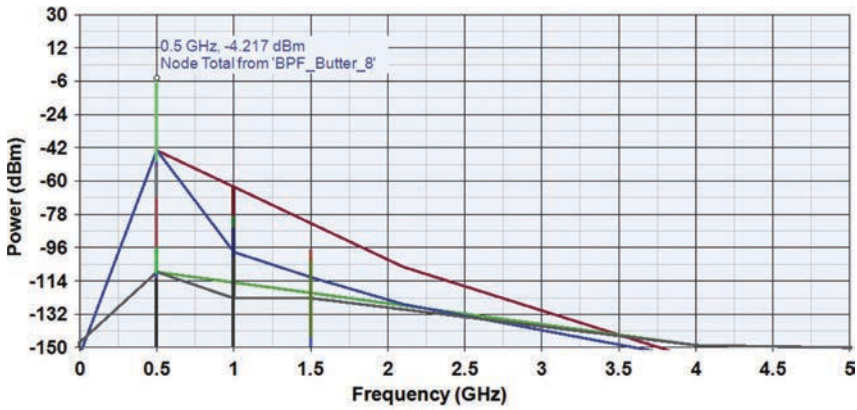


(h)

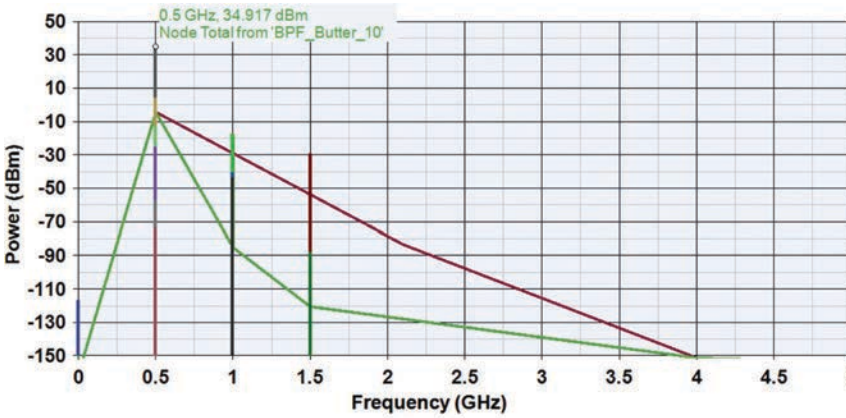


(i)

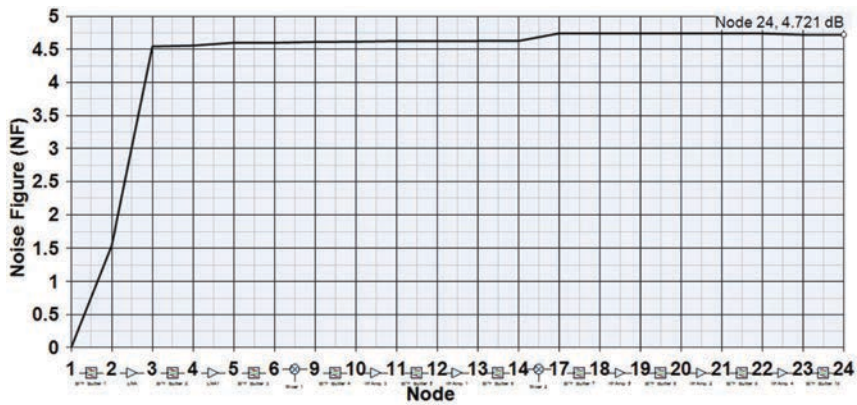
Figure 9.13 (Continued)



(j)



(k)



(l)

Figure 9.13 (Continued)

component's *S*-parameter, non-linear parameters like OP1 dBm, conversion gain/loss, NF, OIP3 dBm, etc., have been incorporated while realizing the transmitter and receiver architecture.

Example 9.2 Design a radar transmitter-receiver system operating at X-band beginning with an input signal at 70 MHz using two approaches:

(A) Frequency up/down conversion with transmitter output at 10.07 GHz.
(B) Frequency multiplier with transmitter output at 10.05 GHz having spur-rejection more than 50dB.

Use SystemVue EDA tool for the purpose of this design.

Solution

(A) (i) **Transmitter Design**: The signal frequency 70 MHz is first up-converted using a mixer that is being fed with a local oscillator (LO) signal at 10 GHz to get 10.07 GHz signal at the output of the mixer, see Figure 9.14a. The output spectrum of this single-stage upconverted signal (at node 3) is shown in Figure 9.14b. After upconversion, LO leakage (@ 10 GHz) and lower side-band leakage (@9.93 GHz) are of the same level. A filter may be incorporated at the output of the mixer of this single-stage up-conversion but the percentage bandwidth requirement of the filter is of the order of 0.3 % that is difficult to realize in practical terms.

However, if double-stage up-conversion is used, see Figure 9.15a, the problem with the practical filter design of single-stage up-conversion, mentioned above, can be circumvented. Here, the LO frequency for first stage up-conversion is set at 1800 MHz thus the up-converted frequency of the 1st mixer is 1870 MHz. The LO frequency (1800 MHz) and lower side-band frequency (1730 MHz) can now be rejected using narrow band-pass filter; as the percentage bandwidth requirement is now around 2%. The second LO frequency is chosen as 8200 MHz so that the up-converted output of the second mixer becomes the desired transmitter frequency of 10.07 GHz. Now the LO leakage (@ 8.2 GHz) and lower side-band leakage (@6.330 GHz) are widely separated that can be easily rejected using band-pass filter. Finally, a power amplifier is incorporated as the output stage of the transmitter to get the desired output power that is required to be transmitted. Output spectrum of the transmitter (referred to node 10) at 10070 MHz has been shown in Figure 9.15b.

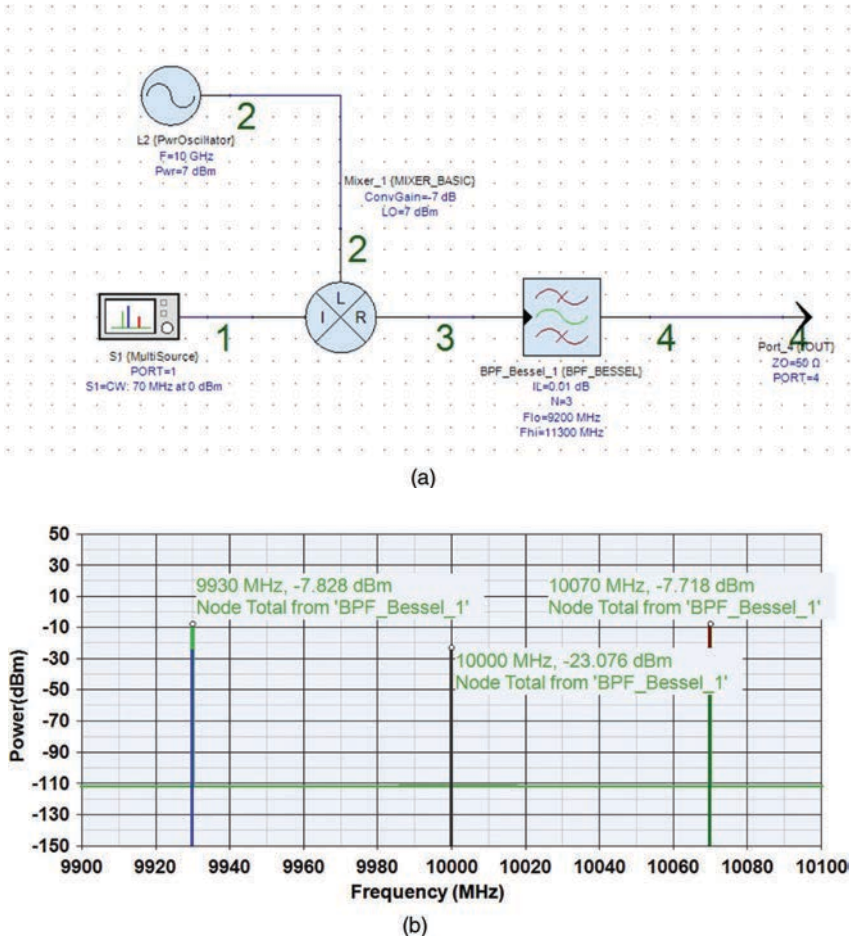


Figure 9.14 (a) Single-stage up-conversion in radar transmitter, and (b) Output spectrum.

(A) (ii) **Receiver Design:** If we use single-stage mixing of the received signal then we again face some practical problem for filter design in order to reject the so-called image frequency. The received signal from transmitter at 10,070 MHz is required to be down-converted to 70 MHz with a local oscillator signal of 10,000 MHz. The image frequency will be at ($f_{RF} - 2f_{IF}$), i.e., at 9930 MHz. So, for an input received signal at 9930 MHz, the IF will be at 70 MHz. Thus, to reject the image at 9930 MHz, we need to use image rejection filter at the front of the receiver. However, the image signal is very close to the desired received signal, and it is very difficult to reject it using practically realizable filter.

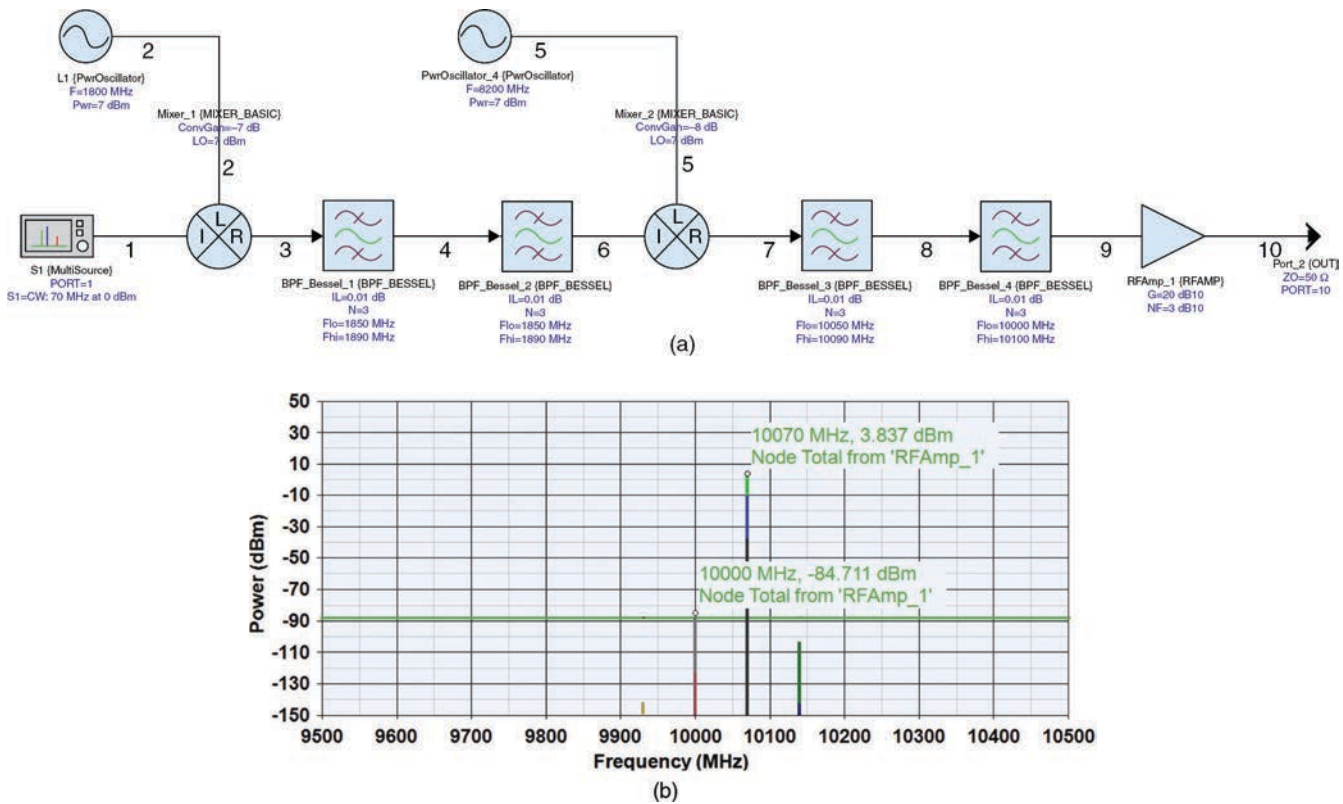


Figure 9.15 (a) Two-stage radar transmitter architecture, and (b) Output spectrum of the transmitter at 10070 MHz.

However, if we make two stages of down conversion, see Figure 9.16a, with 1st IF at 1500 MHz when the image frequency will be at ($f_{RF} - 2f_{IF}$), i.e., at 7070 MHz. Thus, the separation of the image signal is now far away from the desired spectrum at 10070 MHz that can be easily filtered out using band-pass filter at the front of the receiver.

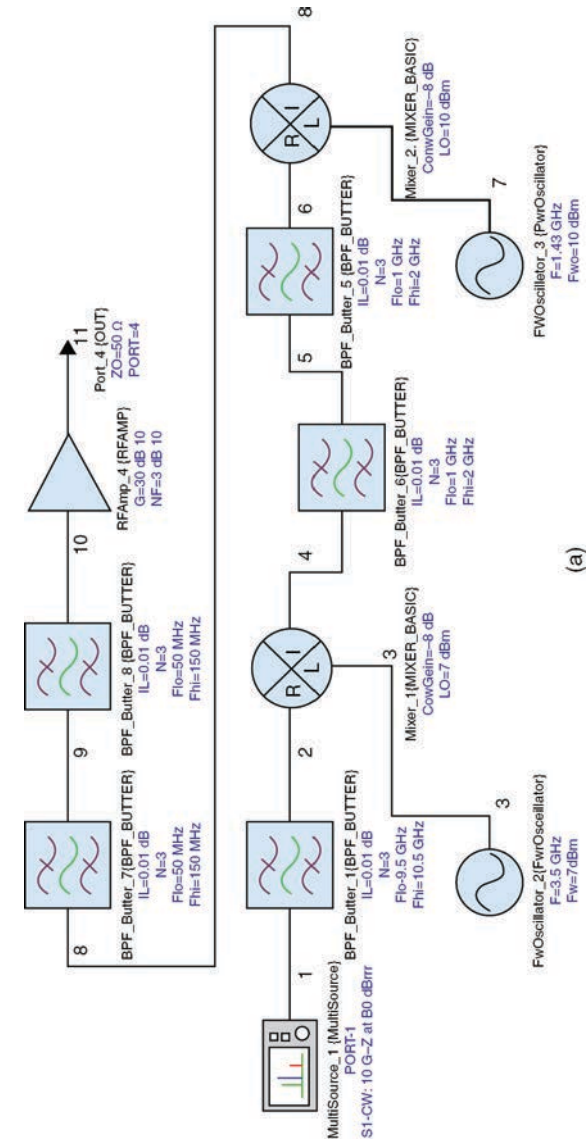


Figure 9.16 (a) Two-stage radar receiver architecture, and (b) Output spectrum of the receiver at 70 MHz.

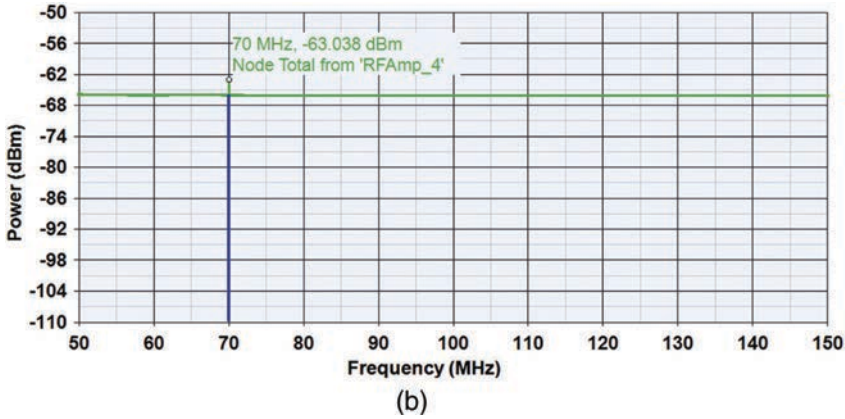


Figure 9.16 (Continued)

In the second stage, IF frequency of 1500 MHz is down-converted to 70 MHz using LO at 1430 MHz. At the final output stage, IF amplifiers are incorporated to compensate for the down-conversion losses. The output spectrum of the receiver is shown at 70 MHz (refer to node 11), see Figure 9.16b. For low-noise receiver, one low-noise amplification stage is added at the front section of the receiver (after the image rejection filter). To protect the receiver from high power jamming, signal limiter is placed just after the antenna at the front section of the receiver.

(B) In radar transmitter where a narrow pulse is transmitted (with low, medium or high pulse repetition frequency) at microwave or millimeter-wave frequency, the bandwidth requirement is very less. This type of transmitter uses multiplier for up-conversion (other than mixer for final up-conversion of frequency) of the signal to the desired frequency range for transmission. Use of multipliers reduces the overall spur in the output spectrum and the total number of filter requirement in the chain also reduces. A specific example is simulated to compare the overall spur level at the transmitter spectrum. In the example, an X-band transmitter is simulated, and the spectrum at different stages are plotted.

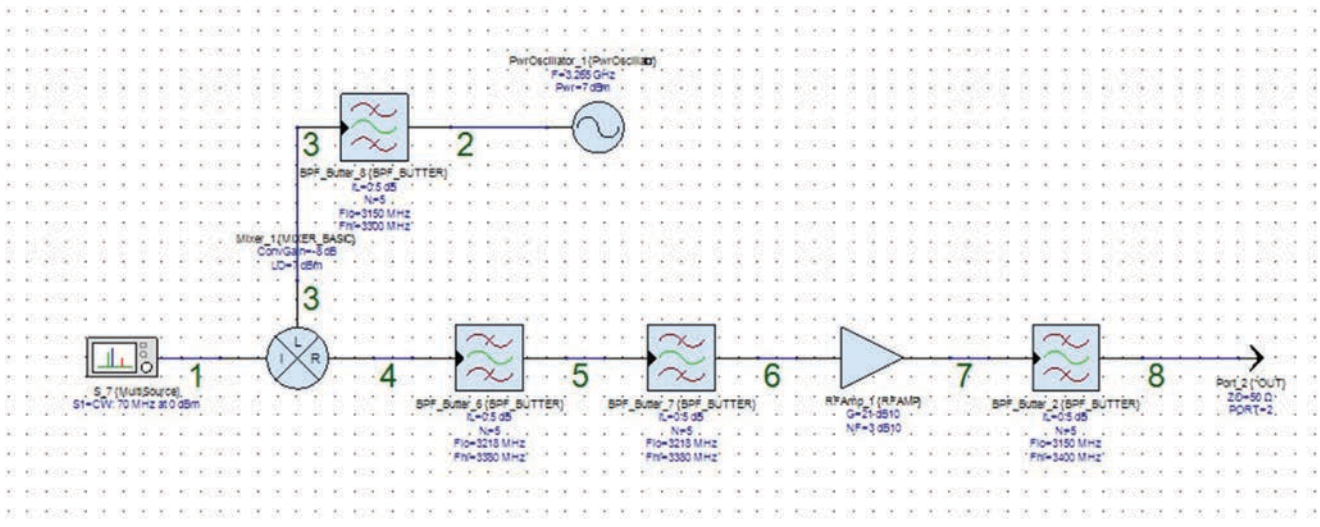


Figure 9.17 (a) Initial scheme of transmitter without multiplier, and (b) Output spectrum showing the power level and spur rejection.

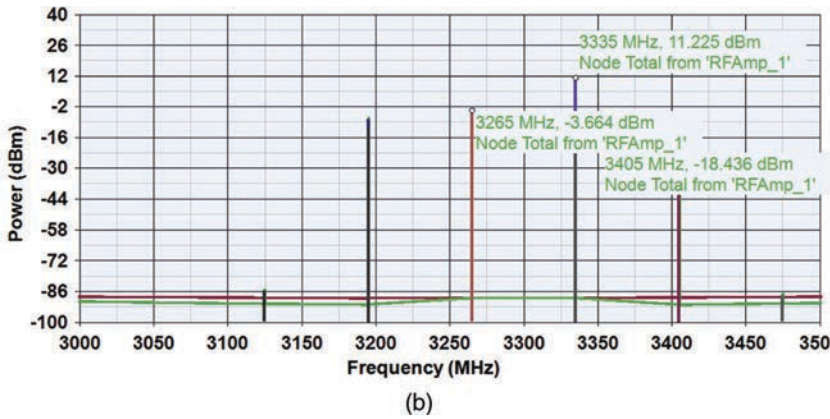


Figure 9.17 (Continued)

The stable signal generated from crystal oscillator (at 70 MHz) is up-converted by the mixer using LO signal at 3.265 GHz, see Figure 9.17a. Two band-pass filters are connected in cascade to reduce the unwanted sideband level and LO leakage. Amplifier is used to boost the power level of the signal. Another filter is used to reject the unwanted signals and the output spectrum, at node 8, of the desired up-converted frequency 3335 MHz is shown in Figure 9.17b, with power level of 11.225 dBm and the spur rejection is around 15 dB.

Next, x3 multiplier is used to multiply the center frequency 3 times so as to get the desired frequency of 10.005 GHz, see Figure 9.18a. The spectrum of the multiplier output, at node 9, is shown in Figure 9.18b.

To reduce the spur level, we may use multiple band-pass filters and by using two-stage amplification, with final output power of 1W is realized. For final transmitter scheme and the output spectrum, see Figure 9.19a and. 9.19b, respectively. It may be observed, at node 15, that the output spectrum of the transmitter at 10.005 GHz with power level 1W has the spur rejection more than 50 dB.

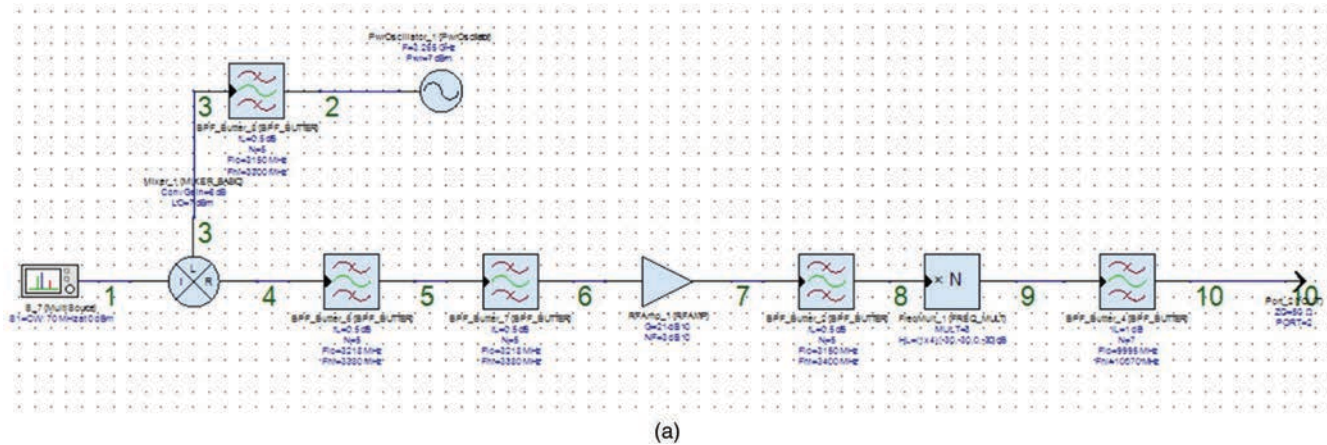


Figure 9.18 (a) Transmitter scheme with multiplier, and (b) Spectrum of multiplier output.

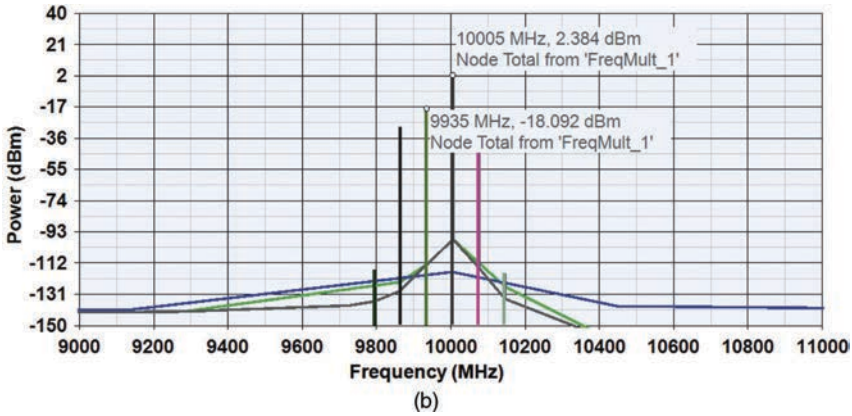
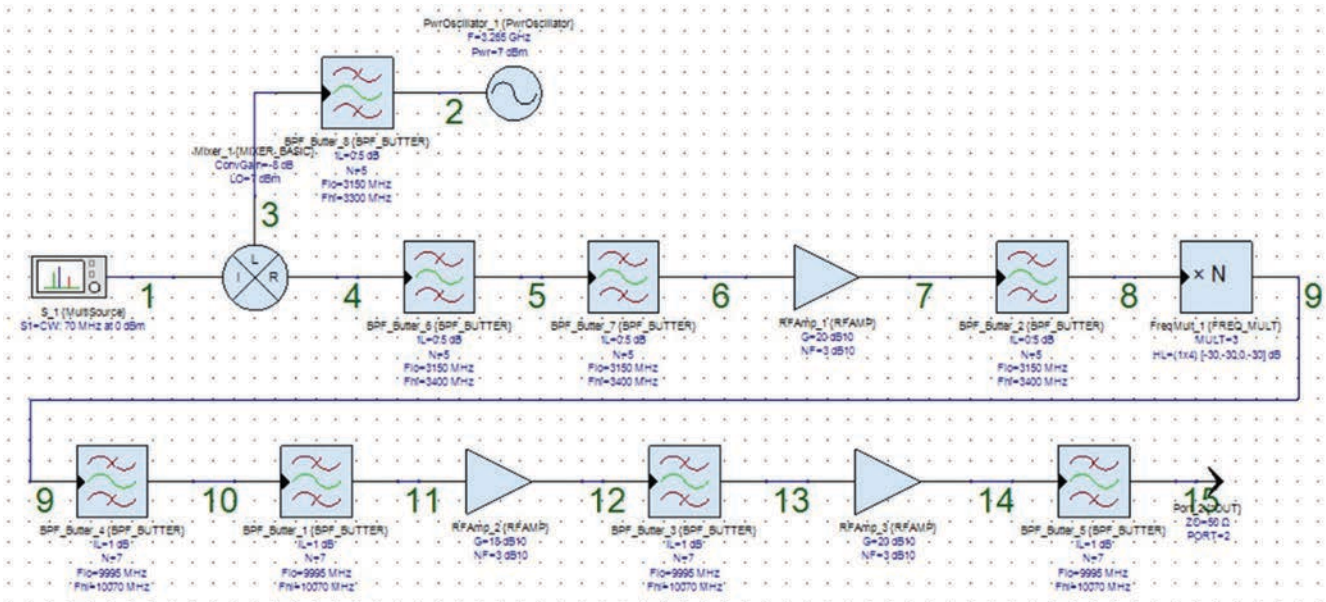


Figure 9.18 (Continued)

A NOTE: A question might come to the mind of the readers, why we have chosen the starting frequency as 70 MHz? The answer for this is given below.

The IF frequency in most radar systems is typically selected around 70–100 MHz because it provides an optimal balance between bandwidth, filtering performance, and ease of system implementation. Radar signal bandwidths are generally below 100 MHz, and in this frequency range high-Q crystal, ceramic, and SAW filters are readily available, offering excellent selectivity, strong rejection, and suitability for image rejection, channelization, and pulse compression. An IF in this range also increases the separation between the desired signal and the image frequency, making image-frequency filtering simpler and more effective. Additionally, modern digital radars can directly digitize IF signals in this band using commercially available high-speed ADCs, enabling straightforward integration into digital processing chains. Overall, using an IF around 70 MHz has become a mature and well-optimized choice that meets the performance, filtering, and digitization needs of contemporary radar systems.



(a)

Figure 9.19 (a) Scheme for reducing spur level, and (b) Output spectrum showing spur rejection by more than 50 dB.

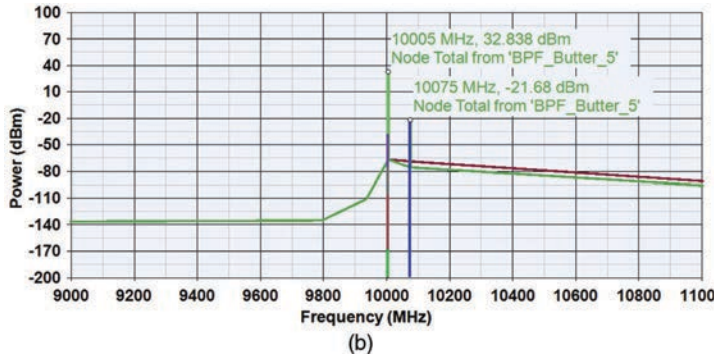


Figure 9.19 (Continued)

REFERENCES

[1] Steer M, *Microwave and RF Design Modules* (Vol. 4), Third Edition, The University of North Carolina Press, North Carolina, 30 September 2019.
[2] Vizmuller P, *RF Design Guide: Systems Circuits and Equations,* Artech House Incorporated, Inc, Norwood, MA, 1995.
[3] Bullock S R, *Transceiver and System Design for Digital Communications*, Fifth Edition, SciTech Publishing, Inc. Raleigh, NC, 2018, IET Telecommunication Series 80.

Appendix I

RF and Microwave Band Designation

ITU* Radio Wave Bands

Band Designation	*Frequency Range*
Very Low Frequency (VLF)	3–30 kHz
Low Frequency (LF)	30–300 kHz
Medium Frequency (MF)	300–3,000 kHz
High Frequency (HF)	3–30 MHz
Very High Frequency (VHF)	30–300 MHz
Ultra-High Frequency (UHF)	300–3,000 MHz
Super High Frequency (SHF)	3–30 GHz
Extra High Frequency (EHF)	30–300 GHz
Terahertz (THz) or Tremendously High Frequency (THF)	300–3,000 GHz [NOTE: 300 GHz to 10 THz; in the present designation]

* International Union of Telecommunication.

IEEE Microwave Bands**

Band Designation	*Frequency Range*
L	1–2 GHz
S	2–4 GHz
C	4–8 GHz
X	8–12 GHz
K	12–18 GHz
Ku	18–27 GHz
Ka	27–40 GHz
V	40–75 GHz
W	75–110 GHz
mm-wave	110–300 GHz

** Institute of Electrical and Electronics Engineers.

Commercial and Military Microwave Bands

Commercial Microwave Bands		Military Microwave Bands	
Frequency Range (GHz)	*Band Designation**	*Frequency Range (GHz)*	*Band Designation***
1.12–1.70	L	0.100–0.250	A
1.70–2.60	LS	0.250–0.500	B
2.60–3.95	S	0.500–1.000	C
3.94–5.99	C	1.000–2.000	D
5.38–8.17	J	2.000–3.000	E
8.20–12.40	X	3.000–4.000	F
12.40–18.00	Ku	4.000–6.000	G
18.00–26.40	K	6.000–8.000	H
26.40–40.00	Ka, R	8.000–10.000	I
33.00–50.00	Q	10.000–20.000	J
40.00–60.00	U	20.000–40.000	K
50.00–75.00	V	40.000–60.000	L
60.00–90.0	E	60.000–100.000	M
75.00–110.00	W		
90.00–140	F		
110.00–170	D		
170.00–260	Y		

* Used by Hewlett Packard, Narda, etc., companies.
** New U.S. Military designation.

Appendix II

MATLAB Programs

AII.1 MATLAB Program to Draw the Curves in Figure 4.12 (Example 4.5)

```
Z0=58;
ZL=222;
pL=(log(ZL/Z0));
p0=0.5*pL;
pm=0.025;
A=acosh(p0/pm);
t1=0.5*log(222*58);
t2=pm*(A^2);
a=22.86;%                                (waveguide shorter dimen-
                                          sion in mm)
lambda0=3;%                              (free-space wavelength in cm)
lambdag=3.975;%                          (guide wavelength in cm)
Z01=377;%                                (free-space characteristic
                                          impedance)
znorm=0:(1/99):1;%                       (normalized length)
p=znorm*log(222/58);
Ze=58*exp(p);%                           (exponential)
be=(Ze*lambda0*a)/(Z01*lambdag);%        (exponential)
figure(1)
   plot(znorm,Ze)
   hold on

   [x,y]=size(znorm);
Zt=zeros(1,y);
bt=zeros(1,y);
for m=1:1:y
if znorm(m)<=0.5
Zt(m)=Zt(m)+58*exp(2*((znorm(m)).^2)*log(222/58));% (triangular)
end
if znorm(m)>0.5
Zt(m)=Zt(m)+58*exp(((4*znorm(m))-(2*((znorm(m)).^2))-1)*log(222/58));
```

```
end
bt(m)=bt(m)+(Zt(m)*lambda0*a)/(Z01*lambdag);
end
figure(1)
   plot(znorm,Zt,'r')

   hold on
%klopfenstein%
zL4=0:(1/99):1;
   [x1,y1]=size(zL4);
   Z3=zeros(1,y1);
   bk=zeros(1,y1);
for m=1:1:y1

   syms y;
phi(m)=int(besseli(0,A*sqrt(1-(y^2)))/(A*sqrt(1-(y^2))),0,((2*zL4(m))-1));
if A==0
phi(m)=0.5*(((2*zL4(m))-1));
end
end
[x2,y2]=size(phi);
phi(1)=0;
phi(y2)=((cosh(A))-1)/(A^2);
for m=1:1:y1
   Z3(m)=Z3(m)+exp(t1+(t2*phi(m)));
   bk(m)=bk(m)+(Z3(m)*lambda0*a)/(Z01*lambdag);
end
figure(1);
   plot(zL4,Z3,'k');
   % Plotting reflection coefficient%
   hold on
   for BL=0:0.01:5*pi
      Tei=(pL/2)*((sin(BL))/BL);
      Teia = abs(Tei);
      figure (2)
      set(gca, 'xtick', [0, pi,2*pi,3*pi,4*pi,5*pi]);
      set(gca, 'xtickLabel', {'0','p','2p','3p','4p','5p'})
      set(gca, 'Fontname', 'symbol')
      xlim([0 5*pi])

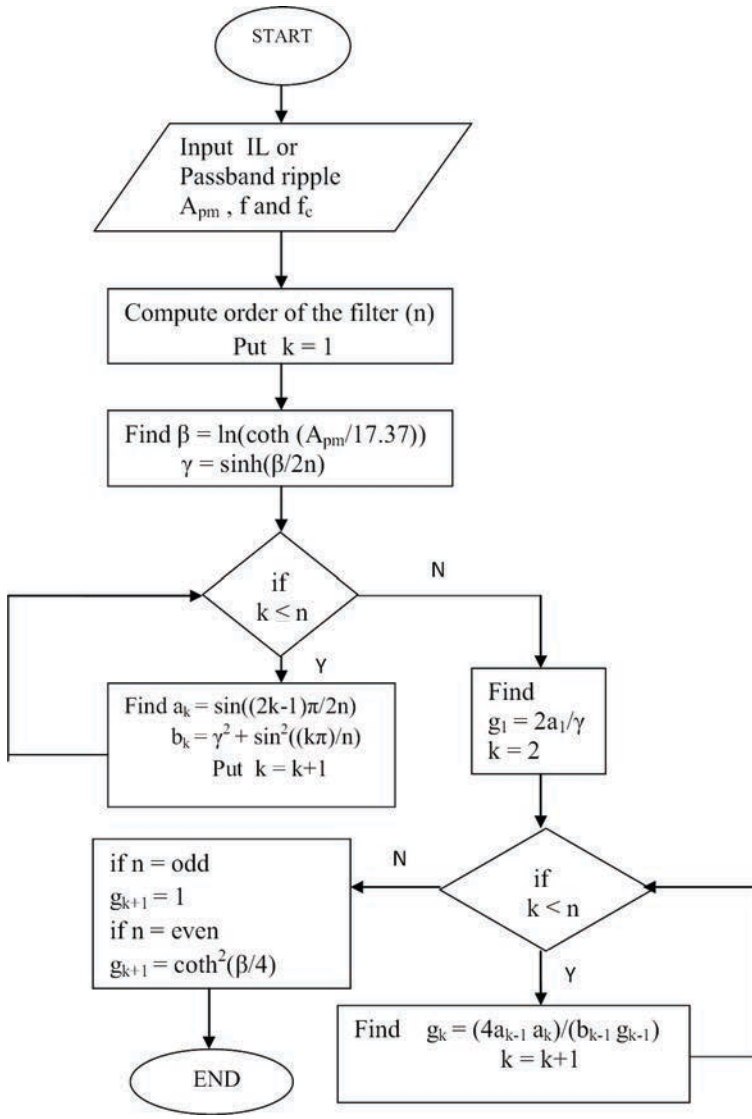
      plot(BL,Teia)
      hold on
 end
```

```
  for BL2=0:0.01:5*pi
     Tti=(pL/2)*(((sin(BL2/2))/(BL2/2))^2);
     Ttia = abs(Tti);
     figure (2)
     set(gca, 'xtick', [0, pi,2*pi,3*pi,4*pi,5*pi]);
     set(gca, 'xtickLabel', {'0','p','2p','3p','4p','5p'})
     set(gca, 'Fontname', 'symbol')
     xlim([0 5*pi])
%       ylim([0 0.5])
     plot(BL2,Ttia,'r')
     hold on
   end
%
%
  for BL3=0:0.01:5*pi
     T0=0.5*pL;
     Tm=0.025;
     A=acosh(T0/Tm);

     Tki=T0*exp(-1i*BL3)*(cos(sqrt(((BL3)^2)-(A^2)))/cosh(A));
     Tkia = abs(Tki);
     figure (2)
     set(gca, 'xtick', [0, pi,2*pi,3*pi,4*pi,5*pi]);
     set(gca, 'xtickLabel', {'0','pi ','2p','3p','4p','5p'})
     set(gca, 'Fontname', 'symbol')
     xlim([0 5*pi])
%       ylim([0 0.5])
     plot(BL3,Tkia,'k')
     hold on
 end
```

AII.2 Flow Chart and MATALB Programs to Generate Tables in Fig. 5.4(a): (i) to (v), and Fig. 5.4(b)

i. **For Table 5.1:** Flowchart to write program for the generation of prototype element (g_k) values for different order (n) of the filter with Tchebyschev (i.e., equal-ripple) response.



Program for Equal-Ripple Case

```
Apm=input('give the value of Apm');
Ap2=(10^(Apm/10))-1;
disp(Ap2);
il=input('give the value of il');
plr=10^(il/10);
disp(plr);
f=input('give the value of f');
w=2*3.14*f;
disp(f);
fc=input('give the value of fc');
wc=2*3.14*fc;
disp(fc);
h=2*(w/wc)
n1=(0.5)*log((4*(plr-1)/Ap2)-h);
disp(n1);

n2=round(n1)
if n2>n1
   n=n2-1;
else
   n=n2;
end
b=log(coth((Apm)/17.37));
disp(b);
y=sinh(b/(2*n));
disp(y);

for   i=1:n
   ak(1,i)= sin((((2*(i))-1)*3.14)/(2*n));
   disp(ak(1,i));
   a1=ak(1,1);
   disp(a1(1,1));
   bk(1,i)=y^2+(sin((i*3.14)/n))^2;
   disp(bk(1,i));

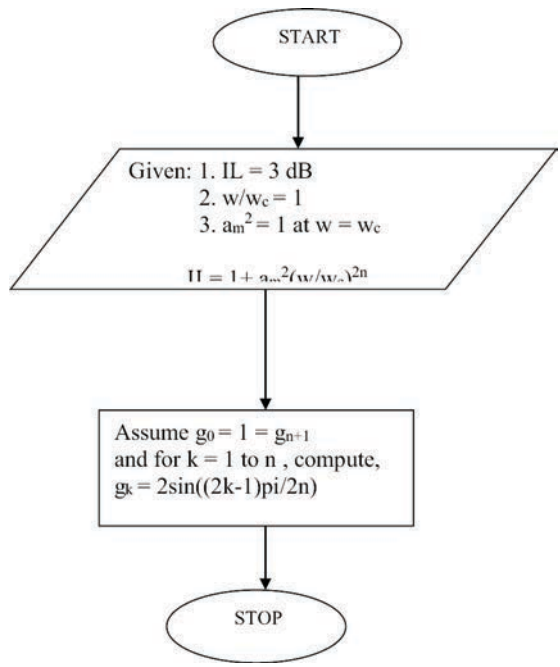
end
g(1,1)=2*a1/y;
   u=g(1,1);

   for i=1:n
      if i<=1
      gk(1,1)=2*a1/y;
```

```
    else
  gk(1,i)=4*(ak(i-1)*ak(i))/(bk(i-1))/(gk(1,i-1));
end

if mod(i,2) == 0
   %number is even
   gev(1,i)=(coth(b/4))^2;
   gk(1,i+1)=gev(1,i);
else
   %number is odd
   god(1,i)=1.0000;
   gk(1,i+1)=god(1,i);
end
   end
```

ii. **For Table 5.2:** Flowchart to write program for the generation of prototype element (g_k) values for Butterworth filter having a given insertion loss IL = 3 dB.



Program for Butterworth Case

[***To generate the component values only]***

```
for i=1:n

   gk(1,i)= 2*sin((((2*(i))-1)*3.14)/(2*n));
end

god(1,i)=1.0000;
gk(1,i+1)=god(1,i);
```

[To generate values along with P_{LR} and other quantities]

```
k2=input('give the value of k²');
disp(k2);
il=input('give the value of il');
plr=10^(il/10);
disp(plr);
f=input('give the value of f');
w=2*3.14*f;
disp(f);
fc=input('give the value of fc');
wc=2*3.14*fc;
disp(fc);
h=(w/wc)
n1=(0.5)*log(((plr-1)/k2)-h);
disp(n1);

n2=round(n1)
if n2>n1
   n=n2-1;
else
   n=n2;
end

for i=1:n

   gk(1,i)= 2*sin((((2*(i))-1)*3.14)/(2*n));
end

god(1,i)=1.0000;
gk(1,i+1)=god(1,i);
```

AII.3 MATLAB Program to Solve Example 5.2

```
clc;
clear all;
format long;
Apm=input('give the value of Apm');
Ap2=(10^(Apm/10))-1
il=input('give the value of il');
plr=10^(il/10)
f=input('give the value of f');
w=2*3.14*f*10^9
fc=input('give the value of fc');
wc=2*3.14*fc*10^9
h=(w/wc)
n1=(acosh(sqrt((plr-1)/Ap2)))/acosh(h)
n=round(n1)

% Find parameter values from table %%

g1=input('give the value of g1');
g2=input('give the value of g2');
g3=input('give the value of g3');
g4=input('give the value of g4');
g5=input('give the value of g5');
g6=input('give the value of g6');

Rl = 50;
C1=g1/(Rl*wc)
L2=(Rl*g2)/wc
C3=g3/(Rl*wc)
L4=(Rl*g4)/wc
C5=g5/(Rl*wc)
Zoc=25;
Zol=120;
er=10;

A=((Zol/60)*sqrt((er+1)/2))+(((er-1)/(er+1))*(0.23+(0.11/er)));
B=(377*3.14)/(2*Zoc*sqrt(er));
A0=((Rl/60)*sqrt((er+1)/2))+(((er-1)/(er+1))*(0.23+(0.11/er)));

Wh1=(8*exp(A))/(exp(2*A)-2)
Wh2=(2/3.14)*(B-1-log((2*B)-1)+(((er-1)/(2*er))*(log(B-1)+0.39-(0.61/
er))))
Wh0=(8*exp(A0))/(exp(2*A0)-2)
```

```
W1=Wh2*0.65
W2=Wh1*0.65
W3=Wh2*0.65
W4=Wh1*0.65
W5=Wh2*0.65
W6=Wh0*0.65

erff1=((er+1)/2)+((er-1)/2)*((1+(12/Wh1))^(-1/2)+(0.04*(1-Wh1)^2))
erff2=((er+1)/2)+((er-1)/2)*((1+(12/Wh2))^(-1/2))

lambda0=(3*10^8)/(fc*10^9)
lambdagl=lambda0/sqrt(erff1)
lambdagc=lambda0/sqrt(erff2)

lC1=fc*10^9*lambdagl*Zoc*C1
lL2=fc*10^9*lambdagl*L2/Zol
lC3=fc*10^9*lambdagl*Zoc*C3
lL4=fc*10^9*lambdagl*L4/Zol
lC5=fc*10^9*lambdagl*Zoc*C5
```

AII.4 MATLAB Program to Solve Example 5.3

```
fc1=2.27;
fc2=2.53;
f0=sqrt(fc1*fc2);
Delta=(fc2-fc1)/2.4;
w0=2.4;
wc2=2.53;
wc1=2.27;
Wp=1;
ws=2.65;

Ap=sqrt((10^0.02)-1);
W1=((Wp*w0)/(wc2-wc1))*((ws/w0)-(w0/ws));
Plr=10^(3);
Tn1=(Plr-1)/((Ap^2)*W1);
Tn=sqrt(Tn1);
x=acosh(Tn);
n=x/(acosh(W1));
g1=1.3394;
g5=g1;
g2=1.337;
g4=g2;
```

```
g3=2.166;
g0=1.0;
g6=g0;

J10z0=sqrt((3.14*Delta)/(2*g0*g1));
J65z0=J10z0;
J21z0=(3.14*Delta)/(2*sqrt(g1*g2));
J54z0=J21z0;
J32z0=(3.14*Delta)/(2*sqrt(g3*g2));
J43z0=J32z0;
z0=50;
% k=0
Z0e10=z0*(1+J10z0+J10z0^2);
Z0es10=Z0e10/2;
Z0o10=z0*(1-J10z0+J10z0^2);
Z0os10=Z0o10/2;
k=1;
Z0e21=z0*(1+J21z0+J21z0^2);
Z0es21=Z0e21/2;
Z0o21=z0*(1-J21z0+J21z0^2);
Z0os21=Z0o21/2;
k=2;
Z0e32=z0*(1+J32z0+J32z0^2);
Z0es32=Z0e32/2;
Z0o32=z0*(1-J32z0+J32z0^2);
Z0os32=Z0o32/2;
syms res10

vpasolve(((60*log(8/res10)+(res10/4))/((11/2)+(9/2)*((1+(12/res10))^-0.5)+
0.04*((1-res10)^2))^0.5)==Z0es10,res10)
pes10=ans;
syms res21
vpasolve(((60*log(8/res21)+(res21/4))/((11/2)+(9/2)*((1+(12/res21))^-0.5)+
0.04*((1-res21)^2))^0.5)==Z0es21,res21)
pes21=ans;
syms res32
vpasolve(((60*log(8/res32)+(res32/4))/((11/2)+(9/2)*((1+(12/res32))^-0.5)+
0.04*((1-res32)^2))^0.5)==Z0es32,res32)
pes32=ans;
syms ros10
vpasolve((((11/2)+((9/2)*((1+(12/ros10))^-0.5)))^0.5)*(ros10+1.4+
0.667*(log(ros10+1.444)))==(377/Z0os10),ros10)
pos10=ans;
syms ros21
```

```
vpasolve((((11/2)+((9/2)*((1+(12/ros21))^-0.5)))^0.5)*(ros21+1.4+
0.667*(log(ros21+1.444)))==(377/Z0os21),ros21)
pos21=ans;
syms ros32
vpasolve((((11/2)+((9/2)*((1+(12/ros32))^-0.5)))^0.5)*(ros32+1.4+
0.667*(log(ros32+1.444)))==(377/Z0os32),ros32)
pos32=ans;

t10=(cosh((3.14/2)*(pes10)));
t110=(cosh((3.14/2)*(pos10)));
t102=((t10+t110-2)/(t110-t10))
t103=acosh(t102);
t104=(2/3.14)*t103
sh10=t104
g10=cosh((3.14/2)*sh10);
des10=((g10+1)*(cosh((3.14/2)*(pes10))))+g10-1
Res10=(acosh(des10)-(3.14/2)*(sh10))/3.14
rres10=Res10
t21=(cosh((3.14/2)*(pes21)));

t121=(cosh((3.14/2)*(pos21)));
t212=((t21+t121-2)/(t121-t21))
t213=acosh(t212);
t214=(2/3.14)*t213
sh21=t214
g21=cosh((3.14/2)*sh21);
des21=((g21+1)*(cosh((3.14/2)*(pes21))))+g21-1
Res21=(acosh(des21)-(3.14/2)*(sh21))/3.14
rres21=Res21

t32=(cosh((3.14/2)*(pes32)));
t132=(cosh((3.14/2)*(pos32)));
t322=((t32+t132-2)/(t132-t32))
t323=acosh(t322);
t324=(2/3.14)*t323
sh32=t324
g32=cosh((3.14/2)*sh32);
des32=((g32+1)*(cosh((3.14/2)*(pes32))))+g32-1
Res32=(acosh(des32)-(3.14/2)*(sh32))/3.14
rres32=Res32

z0esair10=((60*log(8/pes10)+(pes10/4))/((2/2)^0.5))
z0eair10=2*z0esair10
Age10=(300/2.4)*(50/z0eair10)
```

```
z0osair10=377/(((2/2)^0.5)*(pos10+1.4+0.667*(log(pos10+1.444))))
z0oair10=2*z0osair10
Ago10=(300/2.4)*(50/z0oair10)

Agm10=(Age10+Ago10)/2
l10=Agm10/4

z0esair21=((60*log(8/pes21)+(pes21/4))/((2/2)^0.5))
z0eair21=2*z0esair21
Age21=(300/2.4)*(50/z0eair21)

z0osair21=377/(((2/2)^0.5)*(pos21+1.4+0.667*(log(pos21+1.444))))
z0oair21=2*z0osair21
Ago21=(300/2.4)*(50/z0oair21)

Agm21=(Age21+Ago21)/2
l21=Agm21/4

z0esair32=((60*log(8/pes32)+(pes32/4))/((2/2)^0.5))
z0eair32=2*z0esair32
Age32=(300/2.4)*(50/z0eair32)

z0osair32=377/(((2/2)^0.5)*(pos32+1.4+0.667*(log(pos32+1.444))))
z0oair32=2*z0osair32
Ago32=(300/2.4)*(50/z0oair32)

Agm32=(Age32+Ago32)/2
l32=Agm32/4
```

Appendix III

Oscillator Configurations with IMPATT/Gunn Diodes

Commonly used waveguide oscillator circuits of an IMPact Avalanche Transit Time (IMPATT)/Gunn diode are shown in Figure AIII.1(b–d), while the simplest waveguide oscillator circuit may be a bias-post mounted full-height waveguide system (Figure AIII.1a). Here, the diode is mounted along with a bias-post inside the full-height waveguide and the cavity is closed at one end by a sliding short and on the other end is an iris through which the oscillator drives the generated microwave signal to the load. In such an oscillator, the predominantly capacitive device impedance resonates with the inductive impedance of the bias-post generating oscillations which emanate out of the bias-post as a vertically polarized signal resembling radiation from a quarter-wave vertical antenna. The oscillator frequency can be controlled either by changing the bias-post diameter as this will change the inductive impedance presented by it or by changing the sliding short position (which is very effective for mechanical tuning of the oscillator). The simple post-mounted oscillator has the problem of impedance matching via the bias-post between the IMPATT/Gunn device (for X-band diode the magnitude of R_D is of the order of 1 Ω) and the waveguide circuit impedance which is more than 200 Ω.

A common technique at microwave frequencies for realizing better impedance matching between the device and the circuit is to use reduced-height waveguide in which the height (b) of the device may be reduced (thereby reducing the characteristic impedance of the guide, since $Z_0 \; \alpha$ b), and a post of a smaller length may be placed above the diode. In such a configuration of the oscillator, an impedance transformer (stepped impedance transformer or tapered impedance transformer, sections 4.2.2 and 4.2.3 of Chapter 4) is used to match the impedance of the reduced-height waveguide to the full-height waveguide to connect it to the load (see Figure AIII.1b). Another technique to realize device–circuit impedance matching in a full-height waveguide for narrowband operation at microwave/millimeter-wave frequency is a resonant-cap structure (see Figure AIII.1c). The resonant-cap is a circular metal disc attached to the end of a bias-post which is inserted

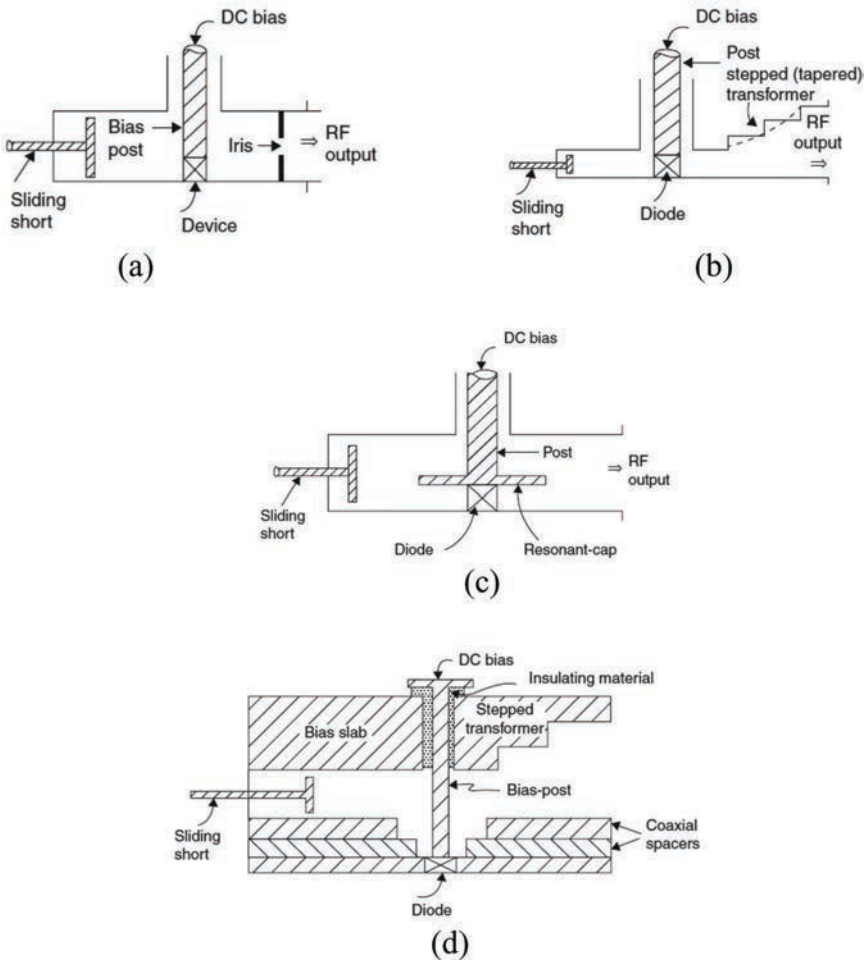


Figure AIII.1 Oscillator configurations with IMPATT/Gunn diodes using: (a) Bias-post-type cavity, (b) Reduced-height waveguide-type cavity, (c) Resonant-cap-type cavity, and (d) Coaxial-waveguide-type cavity.

through the top broad wall of the waveguide. The cap along with the lower broad wall of the waveguide forms a localized radial-line cavity around the diode, with the radius of the cap approximately a quarter wavelength. The frequency of the oscillator is not easily tuned in this type of configuration, since the diameter of the radial-line cavity (which has a very high Q value) set the frequency of oscillation quite precisely and the realizable bandwidth of the oscillator is typically 1%–2%. Schematic cross-section of

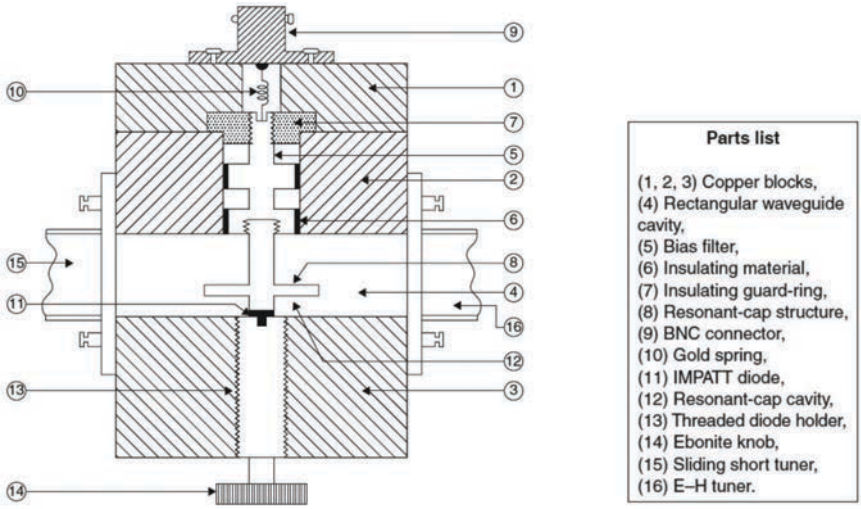


Figure AIII.2 Schematic cross-section of a resonant-cap-type IMPATT Oscillator/amplifier with integral heat-sink.

a resonant-cap-type IMPATT oscillator/amplifier with an integral heat-sink is shown in Figure AIII.2. Yet another technique useful at millimeter-wave frequencies for broad-band operation is coaxial-waveguide configuration in reduced-height waveguide (see Figure AIII.1d). In a coaxial-waveguide cavity, the impedance of the millimeter-wave guide is first reduced with a stepped impedance transformer and then the final impedance matching is realized using coaxial spacers as shown. The power and bandwidth of the coaxial-waveguide type oscillator are controllable by changing the hole-diameter of the coaxial spacers, that is, its coupling with the IMPATT device.

Appendix IV

Short-Slot Hybrid and Branch-Line Quadrature Hybrid

The waveguide version of quadrature hybrid is known as ***short-slot hybrid***, the schematic of which is shown in Figure AIV.1. The principle of operation of this hybrid may be understood as follows. A signal fed at the input port terminal 1_i of the main guide proceeds along that guide until it encounters the coupling slot in the common narrow wall of the main and the auxiliary guides. The presence of the slot in the common narrow wall of the two guides produces two signals at the output port terminals 1_o and 2_o those are equal in magnitude (3 dB power division) but the signal at terminal 2_o leads by 90^0 corresponding to the signal at terminal 1_o, with all ports perfectly matched. The reason for this 3 dB quadrature coupling at the two output terminals may be understood in terms of even- and odd-mode coupling. In fact, every time the traveling wave propagating down either of the guide faces the slot it advances in phase by 90^0.

The ***branch-line quadrature hybrid*** (see Figure AIV.2) is another type of four-port microwave network in which the phase difference of the signals arriving at the two output ports is 90^0. In this coupler, a signal fed at port 1 is evenly divided into ports 2 and 3, with the signal at the coupled port 3 advanced in phase by 90^0 with respect to the signal at the through port 2 (as shown). But the signals arriving at port 4 are in phase opposition and thus cancel out, making port 4 as the isolated port. The coupler behaves identically when power is applied at any other port, for instance, if a signal is applied at port 2, it is split evenly between ports 1 and 4 with no power being coupled to port 3. For 3 dB power coupling between the two output ports of the coupler with the phase relationship of the signals as mentioned above, the characteristic impedance of the different line sections should be as follows: $Z_{01} = \sqrt{Z_0 \times \frac{Z_0}{2}} = Z_0 / \sqrt{2}$ (as it has to transform the Z_0 impedance of the line at input port 1 to the parallel combination of two output lines of impedance Z_0, each) and $Z_{02} = Z_0$ as shown in Figure AIV.2.

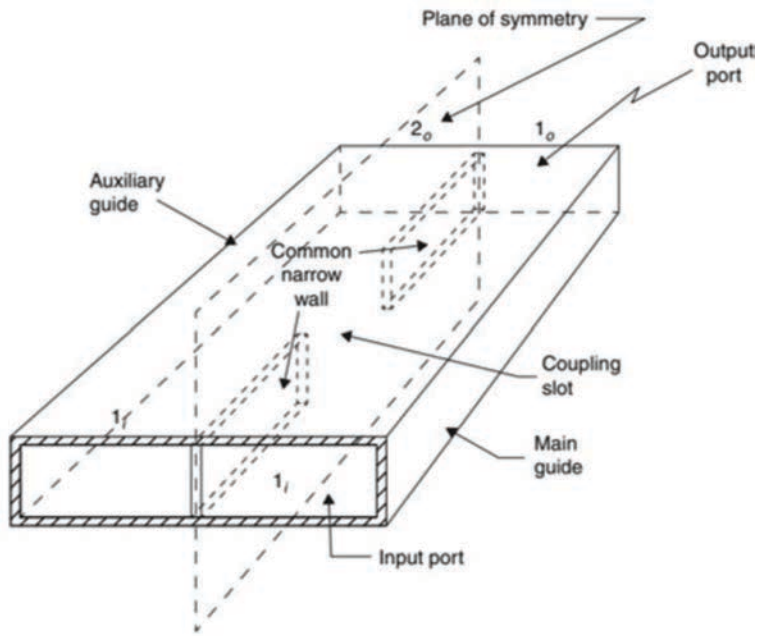


Figure AIV.1 Waveguide version of quadrature hybrid or short-slot hybrid.

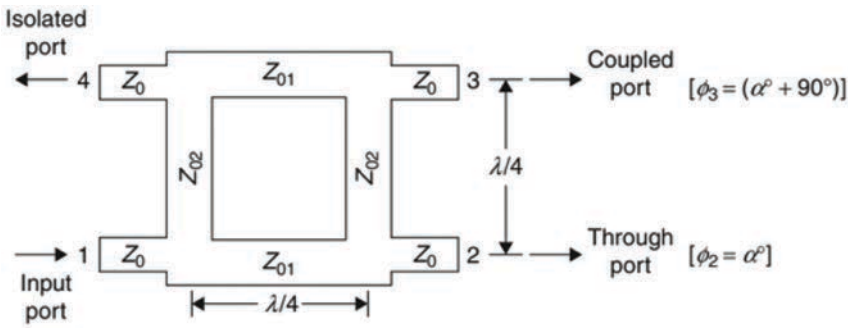


Figure AIV.2 Branch-line quadrature hybrid.

Appendix V

Computational Electromagnetics (CEM), Numerical Techniques, and Simulators

AV.1 Computational Electromagnetics (CEM)

CEM deals with numerical approximation of Maxwell's equations. CEM techniques have practically developed and matured during the last four decades and the electromagnetic (EM) simulators based on them are now an invaluable part of RF and microwave engineering design. CEM is a multidisciplinary field with its core being electromagnetics and numerical methods; however, for useful implementation geometrical modeling and visualization, computer science and algorithm, all have an important role to play.

The "big three" techniques in CEM are *Method of Moments* (MoM), *Finite Element Method* (FEM), and *Finite Difference Time Domain* (FDTD) method. These three methods have been implemented in successful commercial codes (simulators) and are used by RF and microwave engineers quite frequently for design of antennas, filters, microwave integrated circuits (MICs), etc. These three techniques are also known as *Full-Wave CEM* that approximates Maxwell's equations numerically, without using any physical approximations. There is another class of numerical methods for solving Maxwell's equations, known as *Asymptotic Techniques*, that require fundamental approximations in Maxwell's equations—the validity of which increases asymptotically with frequency, normally used in Optics.

Central to all full-wave technique is to use the idea of "*discretizing*" some unknown electromagnetic property, typically the surface current for the MoM, and the electric field for the EFM and FDTD method; for the latter magnetic field may also be discretized. The process of discretization is also known as "*meshing*." It entails subdividing the geometry into a large number of small "*elements*." These may be one-dimensional segments, two-dimensional surface patches (often triangle), or three-dimensional tetrahedral elements, depending on the problem at hand and the method used. With each element, a simple functional dependence is assumed for the spatial variation of the unknown—for instance, a linear approximation—but the amplitude (and possibly phase) of the unknown is determined by application of the method to the patchwork of elements which approximates the original geometry. This functional dependence is also known as the "*basis*

(or expansion) function." Generally, the accuracy of the method is related to the discretization (i.e., mesh size)—the finer the mesh size the better is the accuracy of the method (exceptions are also there as in the case of MoM). Largest mesh size (i.e., the finest geometrical resolution) is limited by the available computational resources.

AV.2 Various Numerical Techniques Those Are the Basis of Electronic Design Automation (EDA) Software Tools

AV.2.1 Introduction

As is the case in many fields of engineering, besides analytical approach (which will always remain essential for developing creative solutions), the role of numerical methods is becoming increasingly important in the characterization and design of RF and microwave devices, circuits, and systems in some form or the other. Indeed, they are useful to rapidly validate new concepts and also to efficiently optimize initial designs.

The starting point of all the numerical techniques is of course the source-free Maxwell's curl equations in isotropic medium (see Appendix VI), which are given below for the time-dependent case and also for time-harmonic case.

Time-dependent case:

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} \quad \nabla \times \vec{H} = \sigma E + \varepsilon \frac{\partial \vec{E}}{\partial t} \tag{AV.1a}$$

Time-harmonic case:

$$\nabla \times \vec{E} = -j\omega\mu\vec{H} \quad \nabla \times \vec{H} = \sigma\vec{E} + j\omega\varepsilon\vec{E} \tag{AV.1b}$$

There are several techniques, such as plane-wave expansion and spherical-wave expansion, FDTD method, transfer-matrix method (TMM), FEM, method of moment (MoM) and so on, to solve Maxwell's equations numerically. Finite differences and finite elements, as well as finite integration (which presents some similarities with finite differences), are generally used for solving equations in differential form. Regarding the integral-equation formulation, MoM is clearly the most popular technique. It is generally confined to fields in the harmonic regime, since time-domain solutions remain a challenge from stability and memory points of view. For the characterization of RF/microwave devices and circuits and also for various design problems we use simulation tools that use numerical techniques like FEM, FDTD, TMM, MoM, and so on. Thus, in this appendix we will discuss all these techniques in brief. For thorough discussions of individual techniques one might refer to standard texts [1, 2].

AV.2.2 MoM, FEM, and FDTD—An Overview

MoM or synonymously the "method of weighted residuals" owes its name to the process of taking moments by multiplying with appropriate weighting functions and integrating. The method has been applied successfully to a wide variety of electromagnetic problems of practical interest, such as radiation due to thin-wire elements and arrays, scattering problems, analysis of microstrip and lossy structures, etc. In MoM the structure under study is represented by equivalent surface current. The surface current is discretized into wire segments and/or surface patches. A matrix equation is then derived, representing the effect of every segment/patch on every other segment/patch (hence the taking of moment). This interaction is computed using Green's function for the problem. The relevant boundary condition is then applied to all the interactions, yielding a set of linear equations. The solution of this linear system yields the (approximate) current on each segment/patch.

FEM divides or discretizes the solution region into subregions or subdomains, i.e., to "finite elements" with the help of two-dimensional (typically triangles for surface mesh) or three-dimensional (typically tetrahedron for volume mesh) elements. Although MoM and FDTD are conceptually simpler and easier to program—FEM is a more powerful and versatile numerical technique for handling problems involving complex geometries and inhomogeneous media. FEM can be applied to two different types of problems, viz., eigen analysis (source-free) and deterministic (driven) one. Problems without any internal/external field source fall into the category of eigen analysis problems. A classic example is a cavity resonator. What emerges from the analysis is a set of eigenvalues and associated eigen modes; these represent the resonant frequencies and associated field distribution within the cavity. In the design of microwave oven FEM can be used to find the design data for the feed location and optimum load position. Traditionally, the FEM has been formulated in the frequency domain, although time domain formulations have also been used for specialized applications. Several companies market commercial FEM products for RF and microwave design. Ansys's HFSS package is widely used and regarded as market leader. FEMLAB has also attracted users.

FDTD is partial differential equation based like FEM, however, this method does not use weighted residuals like FEM but are based upon approximations which permit replacing differential equations by finite difference equations. In other words, the finite difference method relates the value of the dependent variable at a point in the solution region to the value at some neighboring points, i.e., it directly approximates the differential operators in Maxwell's curl equations on "grid" staggered in time and space. This is the only widely used CEM scheme to operate in time domain (MoM and FEM may also operate in time domain, which are otherwise

exclusively formulated in frequency domain, but for specialized applications just as frequency domain formulation of finite difference method is not at all popular for general applications). The apparent simplicity of basic implementation of FDTD and the availability of large memory desktop PCs (in fact FDTD has a voracious appetite for computer memory space) made it very popular in recent times, both in industry and academia, and "do-it-yourself" FDTD codes are commonly developed by graduate students and researchers. FDTD sprang to prominence in 1980s both with the availability of large memory computers and increasing interest in modelling of inhomogeneous materials, in particular for the assessment of human exposure to RF fields.

Time domain formulation of the FDTD method is both an attractive feature and a drawback. Using wideband sources, the FDTD method can compute a wideband response in one run as well as the transient response, whereas frequency domain methods cannot predict the transient response and must obviously re-compute the system response for each frequency point. However, the majority of RF devices operate over quite a narrow frequency band, and this may be less of an advantage than one might expect. In particular, for high Q devices, a very large number of time steps may be required to obtain sufficient frequency resolution and the determination of the smallest time step is subject to Courant's stability condition. Many systems exhibit dispersive properties; examples are waveguides and most real dielectrics and magnetic materials. In frequency domain, this is simple to handle by obvious expedient of simply changing the material/device properties with frequency, but in the time domain this is more challenging since a convolution is implied.

FDTD code-based commercial simulator packages include Computer Simulation Technology (CST) MICROWAVE STUDIO (though it uses finite integration technique but very closely related to FDTD), and REMCON's XFDTD. The CST microwave studio suite (2021) is considered to be the most powerful multilayer 3D full-wave electromagnetic solver, which accurately solves Maxwell equations using the FEM technique.

AV.3 Simulators: Circuit Simulators and EM Simulators

The simulators used for the design of impedance matching networks, filters, oscillators, etc., in this book are either EDA tool or 3D electromagnetic field simulators based on one of the three full-wave CEM techniques discussed above. There are, in general, different types of commercially available simulators for RF and microwave design those are EM simulators (with or without circuit simulators) having different trade names with different capabilities and limitations. A comparative table of some of those simulators is given in Tables AV.1 and AV.2. A few points may be noted with respect to those tables.

Table AV.1 EM simulators

Name of simulator	*Vendor*	*Simulation engines available*	*Applications*
HFSS	Ansys	Frequency domain (finite element method); Transient (discontinuous Galerkin); Integral equation (3D MOM); Physical optics	Antenna and antenna array; Planar and nonplanar passive microwave circuits; Periodic structures; Cavity; Large electromagnetic structures; EMI
CST Microwave Studio	CST	Finite integration; Frequency domain	Antenna and antenna array; Planar and nonplanar passive microwave circuits; Periodic structures; EMI

A circuit simulator can simulate an arbitrary network of active and passive components in either time or frequency domain depending on the type of simulation engine being used. EM simulator can simulate a passive arbitrary structure in time or frequency domain and predict the response or spatial field for an arbitrary (though mostly limited to time harmonic) stimulus.

By linear circuit simulator we understand that the input–output relation of the circuit being simulated is linear in nature. For active devices it implicitly means small-signal analysis of the circuit. One common example of linear analysis is S-parameter analysis.

In nonlinear circuit simulator, the nonlinear behavior of circuit elements is simulated either in time or in frequency domain. One example is harmonic balance analysis, which is a frequency domain tool used to solve the circuit at several harmonics of the applied signal or for two-tone analysis, whereas SPICE is a time domain nonlinear circuit simulator.

The electromagnetic (EM) simulators may be 2D, 2.5D, and 3D planer simulators or full 3D arbitrary geometry simulators. In 2D planer EM simulators, strip-type transmission lines may be analyzed with propagation modes TEM in strip-line and quasi-TEM in microstrip line. The electric and magnetic field vectors are in the cross-sectionnal plane of the line and the variation of the fields being along the length of the line and is a function of time. Electric currents are allowed across the width (y-direction) and length (x-direction) of the metal strip but not vertically (z-direction). 2.5D planer EM simulator is similar to the 2D one; however, extra half dimension means that the solver can calculate currents in the z-direction (usually for through-substrate vias,) in addition to that in x- and y-directions (e.g., Momentum). 3D plane EM simulators are similar, but they are also able to calculate the

Table AV.2 Linear and nonlinear circuit simulators

Name of simulator	*Vendor*	*Capabilities*	*Simulation engines available*	*Applications*	*Coupled EM simulation availability*
Advanced Design System	Keysight Technologies	Schematic/layout design entry; Linear and nonlinear circuit simulation; EM analysis, Synthesis, optimization, and yield analysis; Design rule checking/layout vs schematic	DC., Linear circuit; Harmonic balance; Transient; 2.5D MOM; Electrothermal	Microwave Integrated circuits (MICs); Monolithic MICs (MMICs); RF printed circuit boards (PCBs); Integrated microwave assemblies	Available with momentum (2.5 D Method of Moments simulator); 3D EM with EMPro
Microwave Office	NI (National Instruments)	Schematic/layout design entry; Linear and nonlinear circuit simulation; EM analysis, Synthesis, optimization, and yield analysis; Design rule checking/layout vs schematic	DC., Linear circuit; Multirate Harmonic Balance; Transient; 2.5D MOM	Microwave integrated circuits (MICs); Monolithic MICs (MMICs); RF printed circuit boards (PCBs); Integrated microwave assemblies	Available with AXIEM (2.5 D Method of moments simulator)
Ansys Designer	Ansys	Schematic/layout design entry; Linear and nonlinear circuit simulation; EM analysis, Synthesis, optimization, and yield analysis; Design rule checking, layout vs schematic	DC; Linear circuit; Harmonic balance; High Frequency Spice; 2.5D MOM	Microwave integrated circuits (MICs); Monolithic MICs (MMICs); RF printed circuit boards (PCBs)	Available with Designer (2.5 D Method of moments simulator) and HFSS (3D EM simulator)

[NOTE: Almost all circuit simulators provide industry standard interface to import data from other 2.5D or 3D electromagnetic (EM) simulators for incorporating in circuit design.]

effects of conductor thickness and resistivity (e.g., Sonnet Lite and EM3DS). 3D arbitrary geometry EM simulators define the system to be analyzed volumetrically with no restriction being imposed on the directions of the fields or currents and are also able to handle the effects of metal thickness and resistivity (e.g., HFSS and CST Microwave Studio).

AV.4 SystemVue EDA Simulation Tool

A brief note on SystemVue EDA simulation tool, now known as PathWave system design, by Keysight Technologies, would be worthwhile as this is a system-level design used by RF and microwave engineers. It goes beyond math-based modeling capability with a complete RF/microwave design work flow that helps the engineers to innovate the physical layer of the RF/microwave communication and radar systems. It provides capabilities like system frequency planning, root cause analysis for spurious signals, and integration with RF design tools. It also allows for co-design and co-verification by connecting dataflow simulators (like SystemVue) with circuit simulators (like ADS) that helps to translate the final simulator-based design for hardware implementation.

It is not used for direct hardware design at circuit level; it is essential for architectural design and system-level simulation of RF and microwave systems providing signal characteristics at different intermediate ports of the system. It is primarily used to analyze and verify system performance before committing to a final hardware implementation. The software can simulate complex RF/microwave system effects that are difficult to model with circuit simulators alone for example, intermodulation, and so forth. SystemVue is often used alongside Keysight's other simulation tools like ADS for circuit-level physical layout design.

REFERENCES

[1] Davidson D B, *Computational Electromagnetics for RF and Microwave Engineering*, Second Edition, Cambridge University Press, Cambridge, New York, 2010.

[2] Sadiku M N O, *Numerical Techniques in Electromagnetics with MATLAB*, Third Edition., CRC Press, Taylor and Francis Group, New York, 2011.

Appendix VI

Maxwell's Equations and Electromagnetic Wave Equations—A Brief Review

AVI.1 Introduction

Maxwell's equations and electromagnetic wave equations are the basis of RF and microwave engineering and therefore a brief review of that will help the RF/microwave engineers to refresh their ideas regarding the insight and fallout of these equations. In addition to being the basis of waveguide, cavity resonators, and antenna analysis, Maxwell's equations are also the basis of the electronic design automation (EDA) software tools (discussed in brief in Appendix V) which the modern RF/microwave designers use every now and then.

In 1873, James Clerk Maxwell wrote his now famous *Treatise on Electricity and Magnetism* where he made it amply clear that his treatise is a sort of summary or the unification of the knowledge in electric and magnetic fields as put forward by others, including those who preceded Faraday (Ampère, Gauss, Coulomb, and others). It is, however, more than summarizing what others done prior to him—in fact his main contribution is in proposing the inclusion of "displacement current" in Ampère's circuital law. This seemingly minor change in the field equations, as known before his time, was, in fact, a fundamental change in the theory of electromagnetics. Maxwell himself had no experimental proof for the existence of displacement currents, but it is obvious from reading his book that he considered both displacement currents and the implications of their existence as facts. It was found that pre-Maxwell equations were not consistent with continuity equation and the conservation of charge is not satisfied. However, with the inclusion of displacement current in Ampère's circuital law (by Maxwell) made that consistence possible with charge conservation. Experimental proof of the existence of electromagnetic waves at frequencies well below those of light came only in 1888 with the spark-gap experiment of Heinrich Hertz—10 years after Maxwell died and 15 years after he wrote the Treatise. Maxwell's equations established for the first time the interdependence of electric and magnetic fields and the existence of electromagnetic waves. He established that in free-space, the speed of light is finite (3×10^8m/s) and

light is an electromagnetic wave similar to our well-known radio waves, microwaves, X-rays, and Gamma rays, only with a different frequency band of operation (see Figure 1.4 of Chapter 1 for electromagnetic spectrum).

It was J C Maxwell who first established in 1865 that all the basic principles of electromagnetism can be formulated in terms of four fundamental equations—now known as Maxwell's equations of electromagnetism. These include the Gauss's law of electrostatics and magnetostatics (that basically comes from the non-existent magnetic monopole), Faraday's law of electromagnetic induction and Ampère's circuital law corrected by him with the introduction of displacement current, giving rise to modified Ampère's law or more fittingly which can be mentioned to be the "Ampère–Maxwell law." These are a self-consistent set of equations that are used to solve all the problems of electromagnetism as boundary-value problems with suitable boundary conditions imposed on the electric and magnetic fields as the physical situation demands. Maxwell's equations can be expressed either in the differential or in the integral forms. Maxwell's equations in the differential form are valid at all points in space, that is, they permit us to study directly the interaction between the field vectors and their relationships with source densities at individual points. However, in explaining electromagnetic phenomena in a physical environment, we must deal with finite objects of specified shapes and boundaries. Thus, to deal with a practical physical situation, we need to convert Maxwell's equations in differential forms into their integral form equivalents. Further, the actual type of time functions for the field quantities depends on the source functions ρ_v and $\vec{J}$. In engineering applications, sinusoidal time functions are commonly used when we need to use the phasor form of Maxwell's equations; details of all these are available in standard text books [1,2].

AVI.2 Differential Form of Maxwell's Equations

For time-varying fields, *Maxwell's equations in the differential or point form* are written as [2]:

$$\nabla.\vec{D} = \rho \,(\text{Gauss' law}) \tag{AVI.1a}$$

$$\nabla \cdot \vec{B} = 0 \,(\text{non-existent magnetic monopole}) \tag{AVI.1b}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} (\text{rotational electric field, Faraday's law}) \tag{AVI.1c}$$

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \text{ (rotational magnetic field, Ampère's law modified by Maxwell)} \tag{AVI.1d}$$

With continuity equation as:

$$\nabla \cdot \vec{J} = -\frac{\partial \rho}{\partial t} \tag{AVI.1e}$$

The *constitutive relations* for homogeneous, isotropic medium being:

$$\vec{D} = \varepsilon \vec{E}, \quad \vec{B} = \mu \vec{H}, \quad \vec{J} = \sigma \vec{E} \tag{AVI.1f}$$

NOTE: For the static case, i.e., for time-independent case, Maxwell's equations have the following form:

$$\nabla \cdot \vec{D} = \rho \quad \text{and} \quad \nabla \cdot \vec{B} = 0$$

$$\nabla \times \vec{E} = 0 \quad \text{and} \quad \nabla \times \vec{H} = \vec{J} \tag{AVI.1g}$$

In the above equations, $\vec{D}\left(= \varepsilon \vec{E}\right)$ is the electric displacement, and $\vec{B}\left(= \mu \vec{H}\right)$ is the magnetic flux density, where ε and μ, respectively, represent the permittivity and permeability of the medium. (In simple terms, the permittivity of a medium/material is a measure of how easily the electric field lines can pass through the medium/material, and permeability is the similar term applicable to magnetic field.) E and H are, respectively, electric and magnetic field intensities, and J is the current density, while ρ is the electric charge density in the medium/material. The symbol ∇ is called the *del* (or *nabla*) operator. The *del* operator is a vector operator whose dot $(\cdot)$ product is the divergence of the physical vector quantity (E or H), while the cross $(\times)$ product is the rotation/curl of the physical vector quantity.

It may be seen that even for the time-varying case the magnetic field remains *solenoidal* (i.e., the lines of magnetic flux close upon themselves having the appearance of tube or pipe, meaning the Greek word *sol-een*), while both the electric and magnetic fields become *rotational*, i.e., they have non-zero curl and hence become the *vortex field*.

These four equations together with the current continuity equation $[\nabla.\vec{J} = -(\partial\rho / \partial t)]$ and Lorentz's force equation $[\vec{F} = q\left(\vec{E} + \vec{v} \times \vec{B}\right)]$ form the foundation of electromagnetic theory. These equations can be used to explain and predict all macroscopic electromagnetic phenomena.

If we consider the source-free case when $\rho = 0$ and $\vec{J} = 0$, the set of equations given by AVI.1(a)–AVI.1(d) becomes perfectly symmetrical! Indeed, when Maxwell first introduced the new term of displacement current in the "generalized Ampère's law" or Ampère–Maxwell law, this was his major motivation: he felt that there must be an underlying symmetry between electric and magnetic fields. It is a mark of his remarkable insight that this feeling of him was entirely correct.

Even more significantly, these equations point to an extremely interesting and important coupling between the magnetic and electric fields. If we imagine we have a time-varying electric field, this shows that it will "source" some kind of magnetic field. That magnetic field will be time varying, and so it will source an electric field. That electric field will then "source" a magnetic field, which will source an electric field, which will end up with an infinite chain of electric to magnetic to electric to … This is the basis of the concept of electromagnetic radiation which takes place from an antenna [2].

AVI.3 Maxwell's Wave Equations

The waves of electric and magnetic fields are immediately possible to be derived from Maxwell's two curl equations leading to the so-called *wave equation* of electromagnetics. The solution of these wave equations gives the propagating electromagnetic waves. The general solution of the wave equation is of little practical value; particular solutions are much more important and contribute a great deal to our understanding of certain physical phenomena. One simple solution of wave equation corresponds to a *plane wave*; its importance stems from the fact that complicated fields can be constructed from plane wave solutions by the application of Fourier's theorem. Further, a plane wave can be produced experimentally to a fair degree of accuracy in the shape of a parallel beam of an electromagnetic wave.

The general forms of the so-called *wave equations* obeyed by $\vec{E}$ and $\vec{H}$ fields in source-free unbounded media are given by [2]:

$$\nabla^2\tilde{\vec{E}} - \mu\sigma\frac{\partial\tilde{\vec{E}}}{\partial t} - \mu\varepsilon\frac{\partial^2\tilde{\vec{E}}}{\partial t^2} = 0 \qquad \text{(AVI.2a)}$$

$$\nabla^2\tilde{\vec{E}} - \mu\sigma\frac{\partial\tilde{\vec{H}}}{\partial t} - \mu\varepsilon\frac{\partial^1\tilde{\vec{H}}}{\partial t^2} = 0 \qquad \text{(AVI.2b)}$$

These wave equations govern the electromagnetic fields in a linear, homogeneous, and isotropic medium (conducting or non-conducting) in which the charge density is zero (that is, source-free). The mathematical symmetry of Eqs. (AVI.2a) and (AVI.2b) implies that the solutions of both will have the same mathematical form. Hence, we can solve either one of them and infer the solution for the other.

The wave equations for the time-harmonic fields may be immediately written (by replacing $\tilde{\vec{E}}(r,t)$ by $\vec{E}_s(r)$ and each time derivative in Eq. (AVI.2) by a $j\omega$ factor) as:

$$\nabla^2\vec{E}_s + \left(\omega^2\mu\varepsilon - j\omega\mu\sigma\right)\vec{E}_s = 0$$

$$\text{or } \nabla^2\vec{E}_s - \gamma^2\vec{E}_s = 0 \qquad \text{(AVI.3)}$$

where $\gamma^2 = j\omega\mu(\sigma + j\omega\varepsilon)$.
Again, the complex propagation factor: $\gamma = \alpha + j\beta$, with:

$$\alpha = \omega\sqrt{\frac{\mu\varepsilon}{2}}\left[\sqrt{1+\left(\frac{\sigma}{\omega\varepsilon}\right)^2} - 1\right]^{1/2} \tag{AVI.4a}$$

$$\beta = \omega\sqrt{\frac{\mu\varepsilon}{2}}\left[\sqrt{1+\left(\frac{\sigma}{\omega\varepsilon}\right)^2} + 1\right]^{1/2} \tag{AVI.4b}$$

Equation (AVI.4) is known as the *Vector Helmholtz equation*. A similar equation also holds for the $\vec{H}$ field. It may be noted from Eqs. (AVI.4a and AVI.4b) that both the attenuation constant (α) and the propagation constant (β) of a general lossy medium are nonlinear functions of frequency (ω).

For free-space, which is an unbounded dielectric medium with zero conductivity (i.e., a lossless medium), the wave equation represented by Eq. (AVI.3) simplifies to:

$$\nabla^2\vec{E}_s + k^2\vec{E}_s = 0 \tag{AVI.5}$$

$$\text{Or} \quad \nabla^2\vec{E}_s + \frac{\omega^2}{c^2}\vec{E}_s = 0$$

$$\text{where} \quad k = \omega\sqrt{\mu\varepsilon} = \beta \tag{AVI.6}$$

In Eq. (AVI.6), k is called the *wavenumber* and β is known as the *propagation constant* which are used interchangeably in the literature representing the same physical quantity. It may be noted from Eq. (AVI.6) that for a lossless medium β, k is a linear function of frequency.

As mentioned above, the most easily treated solutions of the wave equation are known as the plane wave solutions. A *plane wave* is defined as a wave in which the phase of the electric field and the phase of the magnetic field at any instant of time are constants over infinite planes (xy-plane) orthogonal to the direction of propagation (z-direction). If the equiphase surface is also the equiamplitude surface, it is a uniform plane wave (nonuniform plane waves are encountered in waveguides).

It is known from electromagnetic theory that the phase velocity of electromagnetic wave is given by:

$$v_p = \frac{\omega}{k} = \frac{1}{\sqrt{\mu\varepsilon}} = k \tag{AVI.7}$$

which in lossless free-space (with $\mu = \mu_0$ and $\varepsilon = \varepsilon_0$) was found to be equal to the velocity of light—this was a major conclusion from Maxwell's contribution in electromagnetism. However, for waves propagating in a lossy medium or those that propagate in guided structures this does not hold true.

The propagation constant β bears a definite relation to the wavelength:

$$\beta = \omega\sqrt{\mu\varepsilon} = \frac{\omega}{v_p} = \frac{2\pi}{\lambda} = k \tag{AVI.8}$$

Thus, k is a measure of the number of wavelengths in a span of 2π, hence the name *wavenumber*. It may be noted that physicists prefer to use k while microwave engineers prefer to use β for the wavenumber/propagation constant.

The total electric and magnetic field strengths of an electromagnetic field bear a constant ratio given by:

$$\frac{E}{H} = \eta = \sqrt{\frac{j\omega\mu}{\sigma + j\omega\varepsilon}} \tag{AVI.9}$$

where η is known as the *wave impedance or the intrinsic* (*i.e., the characteristic*) *impedance* of the medium. For purely dielectric medium with zero conductivity ($\sigma = 0$), $\eta = \sqrt{\mu / \varepsilon}$ and thus for free-space $\eta_0 = 120\ \pi = 377\ \Omega$. It is interesting to note that for DC and power frequency (50/60 Hz) case where voltage (V) and current (I) are measurable quantities, impedance is defined as the ratio of V to I, while for RF and microwave frequency where V is derived from E field and I is derived from H field we talk in terms of *wave impedance*.

The mutual orientation of the electric and magnetic field vectors and the direction of propagation of uniform plane wave are such that $\vec{E}$ and $\vec{H}$ fields are at right angles to each other and the direction of propagation (z-direction) is perpendicular (i.e., transverse) to the plane (wavefront) containing the E and H field vectors in the x-y plane. Thus, a uniform plane wave propagates through the free-space as a *Transverse Electro-Magnetic* (i.e., TEM) wave as shown in Figure 1.9(a) of Chapter 1.

One of the direct applications of Maxwell's wave equation, as described above, is to solve the waveguide (rectangular/circular) problems using appropriate boundary conditions; the waveguide represents a metal–dielectric interfacing problem (the body of the waveguide is metallic while inside of it is a dielectric, commonly air). Waveguides are transmission structure at microwave and higher frequencies and most of the passive components at these frequencies are designed with waveguides. Cavity resonators, which are the electromagnetic equivalent of LC circuits at microwave and higher

frequencies, are terminated waveguides (terminated at both the open ends of the waveguide with short circuit having waveguide length: $l = n\lambda/2$, where n = 1, 2, 3, …) and also demands solution of Maxwell's wave equations with proper boundary conditions to understand their electromagnetic properties and performance. The radiation phenomenon from antenna is another example which may be understandable in terms of wave propagation concept that is based on Maxwell's two curl equations. There are a number of good texts [3–5, 1] that deals with the detailed analysis of waveguide and cavity resonators and waveguide-based components.

[*NOTE*: While discussing antenna radiation problems it is more convenient to use Maxwell's wave equation not in terms of E and H field intensities (as discussed above) but the wave equations need to be expressed in terms of *magnetic vector potential* $\vec{A}$ and the field solutions are obtained in terms of the so-called *retarded vector potential*. By manipulating the two curl equations given by Eq. (AVI.1c) and (AVI.1d) we may derive the following two equations [1]:

$$\nabla^2 \vec{A} - \mu\epsilon \left(\frac{\partial^2 \vec{A}}{\partial t^2} \right) = -\mu \vec{J} + \mu\epsilon \nabla \left(\frac{\partial V}{\partial t} \right) + \nabla \left(\nabla . \vec{A} \right) \tag{AVI.10a}$$

$$\nabla^2 V + \nabla . \frac{\partial \vec{A}}{\partial t} = -\frac{\rho}{\epsilon} \tag{AVI.10b}$$

where $\vec{A}$ is the magnetic vector potential and V is the scalar potential. Eqs. (AVI.10a) and (AVI.10b) are complicated in the sense that they are coupled, i.e., the unknowns $\vec{A}$ and V appear in both equations. A further complication is the fact that Eqs. (AVI.10a) and (AVI.10b) do not yield a unique solution for the potentials. The truth of the later statement may be inferred from the *Helmholtz theorem: Any vector field due to finite source is specified uniquely if both the curl and divergence of the field are specified*. The curl of $\vec{A}$ is known to be: $\mu\vec{H} = \nabla \times \vec{A}$, but its divergence remains unspecified. An examination of Eqs. (AVI.10a) and (AVI.10b) shows that if the divergence of $\vec{A}$ is set to $-\mu\varepsilon(\partial V/\partial t)$, these equations become uncoupled and reduce to standard wave equations with source terms included, viz:

$$\nabla^2 \vec{A} - \mu\epsilon \left(\frac{\partial^2 \vec{A}}{\partial t^2} \right) = -\mu \vec{J} \tag{AVI.11a}$$

$$\nabla^2 V - \mu\epsilon \left(\frac{\partial^2 V}{\partial t^2} \right) = -\frac{\rho}{\epsilon} \tag{AVI.11b}$$

The particular choice used above, i.e.,

$$\nabla.\vec{A} = -\mu\varepsilon\left(\frac{\partial V}{\partial t}\right) \tag{AVI.12}$$

is known as the *Lorentz Gauge condition*.

AVI.4 Retarded Vector Potential for Antenna Analysis

It may be recalled that in the study of electrostatic field and the steady magnetic field it was found possible, and often simpler, first to set up potentials (scalar electric potential, V, in terms of charges and magnetic vector potential, $\vec{A}$, in terms of currents) and then to obtain the electric or magnetic fields from these potentials. Similarly, for the radiated electromagnetic field it appears to be much simpler first to set up potentials that are related to charges or currents of the source of excitation, and then to obtain the $\vec{E}$ and $\vec{B}$ of the radiated field from these potentials. But for radiated field (which is caused by time varying current flowing through the antenna element) one cannot use magnetic vector potential used for static field (which will be applicable only if the antenna size is assumed to be extremely small so that propagation effects do not jeopardize the physical situation). The problem may be solved using wave equation but that approach is very complicated even with simple antenna geometry. However, a logical approach that supports the physical situation yet allows us to obtain the correct solution without solving the wave equation directly is the use of *retarded magnetic vector potential*. The idea may be elucidated as follows: Suppose that the (time varying) current carrying antenna element generates the magnetic vector potential in the space surrounding it. The solution must follow wave equation; thus, the waves generated at the source (the antenna element) propagate in all directions with a constant and finite phase velocity $\left(v_p = 1/\sqrt{\mu\varepsilon}\right)$. An observer located at a distance r form the source would receive the signal after a time $t = r/v_p$. Thus, the wave that is received at the observer's location is a retarded wave and we have to modify the magnetic vector potential expression of static field with the inclusion of this time delay and obtain the expression for retarded vector potential as [1]:

$$\vec{A}(\vec{r},t) = \frac{\mu}{4\pi}\int \frac{\vec{J}\left(r', t - r/v_p\right)}{r} dv'. \tag{AIV.13}$$

Thus, the vector potential at the observation point at a time t due to the time varying current flowing through the antenna element (dv' being the

elemental volume of the conductor the antenna element) has been generated by a current at time $(t - r/v_p)$ at the source point (i.e., at the antenna). In other words, the current itself gets retarded by factor a $(t - r/v_p)$ as this current for radiation to take place is a time varying one.

REFERENCES

[1] Jordan E, and Balmain K G, *Electromagnetic Waves and Radiating Systems*, Second Edition, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1968.
[2] Kar S, *Wireless Technology—Kilohertz to Megahertz*, CRC Press, Taylor and Francis Group, London, 2025.
[3] Rizzi P A, *Microwave Engineering Passive Circuits*, Prentice-Hall of India Pvt. Ltd., New Delhi, 1999.
[4] Liao S Y, *Microwave Devices and Circuits*, Third Edition, Pearson Education Inc., India, 2003.
[5] Pozar D M, *Microwave Engineering*, Fourth Edition, Wiley, New York, 2012.

Appendix VII

Fundamentals of Spectrum Analyzer and Network Analyzer

AVII.1 Introduction

For radio frequency (RF) and microwave engineers, the spectrum analyzer and network analyzer are the two most important equipment with which every RF/microwave engineer should be conversant. Both spectrum analyzer and network analyzer, in general, make frequency domain measurements; although with "zero-span" of spectrum analyzer and the "time domain reflectometry" option of network analyzer, time domain measurements can also be done. A spectrum analyzer is used to study the signal characteristics of signal sources, while a network analyzer is used to characterize components, devices, circuits, and subassemblies. To be more explicit, a spectrum analyzer is intended for analyzing the nature of (unknown) signals that are fed into them, that is, the signature of the energy source. The network analyzer, on the other hand, contains a reference signal source (test signal) within it and uses this to analyze a network, that is, to characterize the network signature. Although there are many similarities between spectrum analyzer and network analyzer (both contain tuned receivers operating over similar frequency ranges), they are optimized for very different measurement applications. A spectrum analyzer is used to characterize the signal level and its harmonics including phase-noise of oscillators and signal or carrier level, side bands, and intermodulation products of a communication or radar signal (just to mention a few). On the other hand, a network analyzer, on the basis of the S-parameter (for linear network) or X-parameter (for nonlinear network) measurements, characterizes the network under investigation. With a network analyzer we get both the (ratioed) amplitude and phase information, but spectrum analyzer does not, in general, provide the phase information of the signal.

From a system engineer's point of view, a spectrum analyzer is a swept-tuned super-heterodyne receiver that provides a cathode ray tube (CRT)/liquid crystal display (LCD) display of amplitude versus frequency. It is essentially a frequency selective, peak responding voltmeter calibrated to display the root mean squared (RMS) value of a sine wave. The spectrum analyzer can show the individual frequency components and the relative

power each component contributes to make up a composite signal (which otherwise appears as a composite voltage waveform in the time domain as displayed in an oscilloscope). A network analyzer, on the other hand, contains both source and multiple receivers (in contrast to the spectrum analyzer which is configured as a single-channel receiver, without a source) and displays (in CRT/LCD) the amplitude and phase information in various formats including one on a Smith chart. With vector-error correction, the network analyzer provides a much higher measurement accuracy than the spectrum analyzer.

Why do both the spectrum analyzer and network analyzer, in general, make measurements in the frequency domain? This is because frequency domain measurements have several distinct advantages. First, narrowband frequency domain measurements provide a greater sensitivity and dynamic range than time domain measurements. Since the measurement bandwidth can be narrowed almost arbitrarily, frequency domain analyzers can greatly reduce the amount of noise present in the measurement. Also, narrowband measurements can remove large interfering signals at undesired frequencies. Second, some systems are inherently frequency domain oriented. For example, many telecommunication systems use what is called the Frequency Division Multiple Access (FDMA) or Frequency Division Multiplexing (FDM). In these systems, different users are assigned different frequencies for transmitting and receiving, such as with a cellular phone. These types of systems must be analyzed in the frequency domain in order to minimize the possibilities of mutual interference. Even when systems are not inherently in the frequency domain, they might require frequency domain measurements. For example, the bandwidth and the speed of a digital pulse may be affected by stray capacitance and resistive losses in a high-speed digital circuit. With the help of a network analyzer, the bandwidth of the circuit may be determined by measuring its transfer function in the frequency domain. Again, suppose a switching power supply operating at a desired switching frequency is corrupted with the harmonics of the line frequency (50/60 Hz), then it is difficult to characterize the signal with multiple corrupting frequencies with a time domain measurement in an oscilloscope. A spectrum analyzer, on the other hand, can separate these frequency components and can measure them accurately.

AVII.2 Spectrum Analyzer Basics

AVII.2.1 Introduction

An RF/microwave spectrum analyzer works typically in the frequency range from 10 MHz to 40 GHz, which covers the vast majority of communication, navigation, radar systems, and industrial applications. For operation at higher millimeter-wave frequency, external waveguide mixers can be used to extend the frequency range of operation.

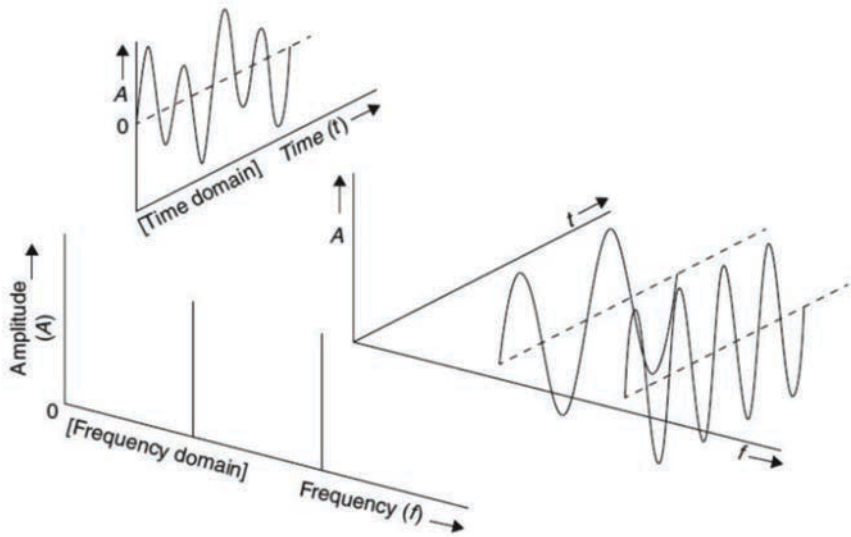


Figure AVII.1 Representation of signals in the time and frequency domains.

Oscilloscopes make measurements in the *time domain* while in spectrum analyzers measurement is usually done in the *frequency domain*. The two axes in the plot in an oscilloscope are time and amplitude, whereas in a spectrum analyzer they are frequency and amplitude. Electrical signals can be represented and observed in both the time domain and frequency domain (see Figure AVII.1). These two methods of representation are linked together via the *Fourier transform*, which means that there is a characteristic frequency spectrum for every signal that can be represented in the time domain.

A spectrum analyzer displays each frequency component of a signal as an individual vertical line (deflection) in the CRT. For example, a pure sine wave would be a single vertical line displayed at the frequency corresponding to the frequency of the sine wave. For nonsinusoidal waveforms, vertical lines representing the fundamental and harmonic (odd and/or even) components would be displayed. For example, a square wave displayed in an oscilloscope (in time domain) will appear in a spectrum analyzer (in the frequency domain) with a vertical line corresponding to the fundamental frequency (f) and multiple vertical lines with diminishing amplitude at $3f$, $5f$, $7f$… corresponding to the odd harmonics. Because of this capability of the spectrum analyzer, it is possible to study the power density spectrum and phase-noise of oscillators, carrier modulation (including pulsed RF signal) and unexpected modulation, harmonics, and intermodulation levels. In addition to swept-tuned frequency mode, a spectrum analyzer can be used in

fixed-tuned mode (zero-span) to provide time-domain measurement capability much like that of an oscilloscope.

AVII.2.2 Basic Principle and Block Diagram of Spectrum Analyzer

The job of a spectrum analyzer is to extract the frequency content of a general time domain signal, say, $x(t)$ and to plot amplitude versus frequency of all the frequency components contained in the signal. A phenomenological solution may be to pass $x(t)$ through a long series/bank of band-pass filters (BPF) with overlapping adjacent passbands, that is, $f_1 + BW/2 = f_2 - BW/2$ and so on, where f_1 and f_2 are, respectively, the center frequency of filters 1 and 2, and BW is the bandwidth of the BPFs (see Figure AVII.2a). If the BPFs are of very narrow bandwidth, then the extracted filter outputs give us the frequency components $X(f_1)$, $X(f_2)$, …. and so on. However, this may not be a practical approach and a simple and better solution may be to use a property of Fourier transform which states that, if we multiply (in the time domain) a signal $x(t)$ by a sinusoid, the spectrum of the signal is shifted in the frequency domain by an amount equal to the frequency (f_0) of the sinusoid. Mathematically,

$$x(t).\cos(2\pi f_0 t) \rightarrow 0.5X(f - f_0) + 0.5X(f + f_0)$$

Thus, instead of using a bank of narrowband filters, we can now use only one narrowband filter centered at a fixed frequency, say, f_{IF}, and scan the signal spectrum across this filter by multiplying $x(t)$ with a sinusoid of varying frequency f_0. In other words, we can use heterodyning ("heterodyne" in Greek means mixing different frequencies) principle (see Figure AVII.2b) which we use in the ordinary amplitude modulation (AM) and frequency modulation (FM) radio receivers. While tuning the AM/FM radio receiver, we select f_0 (the frequency of the local oscillator, LO) so that the desired signal will pass through the IF filter; while in a spectrum analyzer, the frequency f_0 is swept over a desired range so that the frequency component $X(f)$ is shifted to f_{IF} and passed by the filter. For example, if we want to view the frequency content of $x(t)$ from f_1 to f_2, then we must select f_0 to scan from $(f_1 + f_{IF})$ to $(f_2 + f_{IF})$.

Thus, the most common type of spectrum analyzer is basically swept-tuned super-heterodyne receiver (also popularly known as swept/scanning spectrum analyzer), of course, having a much more complicated signal conditioning going on inside the analyzer than the one shown in Figure AVII.2. A simplified block diagram of a swept spectrum analyzer is shown in Figure AVII.3.

A ramp (sweep) generator is used to produce a linearly increasing voltage which drives the tuning port of a voltage-controlled oscillator (VCO) acting as LO. The same ramp voltage is also applied to the horizontal axis of the

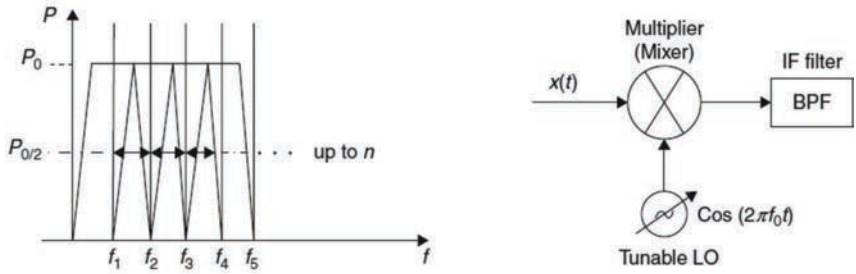


Figure AVII.2 (a) A long series/bank of band-pass filters with overlapping adjacent pass bands and (b) Frequency mixing (heterodyning).

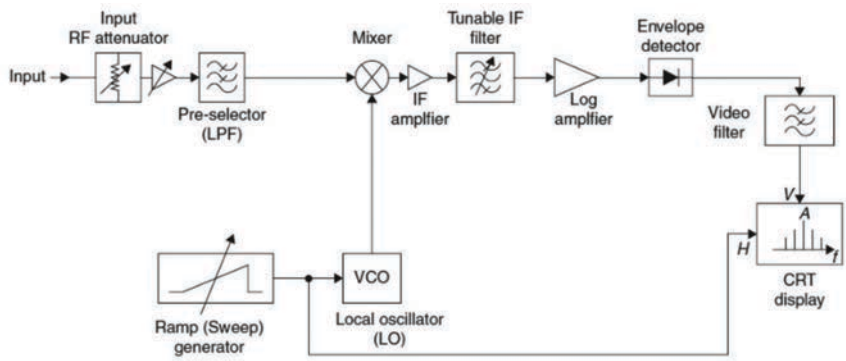


Figure AVII.3 Block diagram of a swept-tuned spectrum analyzer.

display (say, a CRT) to facilitate synchronization. The RF signal under test is heterodyned (mixed) with the LO signal to generate the desired intermediate frequency (IF). The LO in Figure AVII.3 is tuned by a sweep (ramp) generator in order to convert the entire input frequency range to a constant IF. The IF signal is amplified and arrives at the IF filter with a definable bandwidth. The input signal is essentially "swept past" this filter with a fixed center frequency. The IF filter determines the resolution bandwidth (RBW) of the analyzer. In order to be able to display signals with a wide variety of amplitude levels on the screen simultaneously, the IF signal is compressed with the aid of a logarithmic amplifier. After that, the envelope detector and the video filter, respectively, acquire the signal's envelope, and the noise is reduced with the aid of an averaging process (by the video filter), which smooths out the displayed video signal. As the LO is swept in frequency, the spectrum of the input signal is automatically plotted on the display.

The input side of the spectrum analyzer block diagram has a variable RF attenuator, often followed by a variable gain amplifier. The purpose is to control the signal level applied to the rest of the instrument. If the signal level is too large, the analyzer circuit will distort the signal, causing distortion products to appear along with the desired signal. If the signal level is too low, the signal may be masked by noise present in the analyzer. Either problem would otherwise reduce the dynamic range of the measurement and hence needs to be addressed properly by setting the input signal level to the instrument.

The conceptual basis for understanding the operation of the spectrum analyzer discussed above uses all analogue implementation. However, in this digital age, the use of microprocessor and digital signal processing (DSP) has become very common with all modern spectrum analyzers. The swept LO may be a digital synthesized source with microprocessor control. Also, the output of the detector may be sampled and converted to a digital signal with analogue-to-digital (ADC) converter which may be read by the microprocessor. Further, the display in a modern spectrum analyzer is always a digital graphic display using LCD. The photograph of a modern spectrum analyzer depicting phase-noise of a microwave oscillator was shown in Figure 9.6 of Chapter 9.

In addition to the above-mentioned method of frequency-domain measurement using the swept-tuned spectrum analyzer technique, with the availability of advanced ADC and DSP techniques, the *Fourier analyzer* is becoming useful for sophisticated measurements at lower RF frequencies. The Fourier analyzer basically takes a time-domain signal, digitizes it using digital sampling, and then performs the mathematics required to convert it to the frequency domain, and display the resulting spectrum. With its real-time signal analysis capability, the Fourier analyzer is able to capture periodic as well as random and transient events. It also can provide significant speed improvement over the more traditional swept analyzer and can measure magnitude as well as phase. However, it has its limitations, particularly in the areas of frequency range, sensitivity, and dynamic range. These analyzers can offer significant performance improvement over conventional spectrum analyzers, but often with a price premium.

State-of-the-art spectrum analyzers of today combine the benefits of both the swept-tuned and Fourier analyzer leading to *real-time spectrum analyzer*, which captures and processes the spectrum of the signal so fast that no part of the signal is missed.

AVII.2.3 A Few Passing Comments

A few passing comments may be made regarding some of the blocks used in the diagram above (Figure AVII.3) and the use of a spectrum analyzer for best performance on the basis of application requirements.

The block diagram shown in Figure AVII.3 uses a single mixer and a single IF stage; however, all modern spectrum analyzers use multiple conversion stages. Each conversion section contains a mixer, an LO and IF filter. LOs may all be derived from a master oscillator.

The function of the pre-selector filter is signal conditioning—to select or allow only a narrowband of frequencies to pass into the analyzer. It is a sweeping filter that tracks the frequency the analyzer is tuned to at any particular point of time. The function it performs is to inhibit harmonic mixing within the first mixer/converter. By eliminating the harmonic conversions, unwanted mixing products do not appear as spurious signals in the spectrum being displayed.

The choice of resolution bandwidth (RBW) and span/division (span/div) is an important issue for observing the spectrum of the signal under investigation in a spectrum analyzer. The RBW determines the finest frequency line that may be visible/obscured in the spectrum. The smaller the RBW, the finer is the resolution and hence the fine details of the spectrum will be visible. However, with larger RBW, the finer details are obscured/lost.

The span/div determines how the spread/width of the frequency spectrum being displayed on the screen can be varied/controlled. The span/div determines the sweeping speed across the spectrum. With the so-called zero span, the analyzer no longer sweeps across a frequency spectrum, but behaves like a super-heterodyne receiver. The analyzer under this situation behaves just like a typical oscilloscope where the display indicates the modulation of any signal at the frequency selected by the frequency control.

The choice of an extremely small RBW with an extremely wide span/div would make the signal spectrum appear exceptionally narrow and may be overlooked easily (practically invisible!) and the measurement process very slow or time consuming. Again, if a very wide RBW is chosen with respect to span/div, then low-level signals nearby will be masked in the filter skirts of the RBW filter. Unless special requirement dictates a specific RBW, the RBW selected should be somewhat greater than 1/50 the span/div. Most modern spectrum analyzers through the use of microprocessors have the capability to select the optimum RBW depending on span/div. At times, however, manual control of RBW may be desired.

RBW is a functional control that selects one of several band-pass filters physically located in the IF signal processing chain of the instrument. It is defined in terms of Hz and is a measure of the width of the filter characteristics either 3 dB or 6 dB down from its peak value, depending on the manufacturer of the analyzer. The shape of the signal being traced on the screen is a combination of the RBW filter and the signal, and not just the shape of the signal being analyzed. The limitations imposed on an analyzer by the RBW filter are significant. The sweep speed (the rate at which the analyzer sweeps the frequency being present) must be slow enough to allow the RBW filter

to reach peak amplitude, otherwise an inaccurate signal amplitude would result.

In general, narrow RBW (optimum with respect to span/div) is desired from the viewpoint of many applications. When displaying a pulsed RF signal (commonly used in radar systems), the RBW should be selected narrow enough so that each spectral line (which is close by) of the pulsed RF spectrum becomes distinctly visible. Further, with the choice of narrow RBW, the noise floor (sometimes known as "grass" because of its appearance) of the analyzer goes down, in addition to making the spectrum of the signal being displayed narrow, and thus helps to identify low-level signal and sidebands close by, if any, which otherwise get obscured in higher noise floor levels with the spreading out of the signal spectrum over a wide frequency range.

AVII.3 Network Analyzer Basics

AVII.3.1 Introduction

The network analyzer provides a complete description of linear network behavior (in terms of S-parameters) and of nonlinear network (in terms of X-parameters) in the frequency domain. This is demanded by RF/microwave design engineers (who need to predict with some certainty the final network performance of a multi-component RF/microwave network from the knowledge of the individual components) and production manager (who needs to know the allowable tolerance on the products manufactured and if the final product meets the specified tolerance). The job of the network analysis is to create a data model (of active/passive network) of component parameters as impedance and transfer function through stimulus response testing (with test signal source built into the system) over the frequency range of interest. Generally, a sine wave test is used for linear networks. The swept nature of analysis is required to have both phase and amplitude of the signal characteristics of the component under test. A computer-controlled network analyzer can be programmed to set up and make many measurements automatically. The measurement process is further accelerated by the computer's ability to store, transform, summarize, and give the output of the data in a variety of formats to a number of peripherals. This capability makes the computer-controlled network analyzer ideal for both computer-aided design (CAD) and automatic product testing (APT).

The scalar network analyzer (SNA) is a form of RF/microwave network analyzer that only measures the amplitude properties of the device under test. The vector network analyzer (VNA) is a more useful form of RF/microwave network analyzer than the SNA as it is able to measure the amplitude response and the phase response as well. As a result, VNA may also be called a gain-phase meter or an automatic network analyzer. The large-signal network analyzer (LSNA) is a highly specialized form of RF network

analyzer that is able to investigate the characteristics of devices under large-signal conditions. It is able to look at the harmonics and nonlinearity of a network; the previous version of LSNA was known as the microwave transition analyzer (MTA).

Vector network analysis is a method of accurately characterizing the electrical performance of components and circuits by measuring their effect on the amplitude and phase of swept-frequency/swept-power test signals. Measuring both magnitude and phase of components is important for several reasons. Though in some applications, for example, when we are only concerned about the gain of an amplifier or the stop-band rejection of a filter, magnitude-only data is sufficient for our needs. However, complete characterization of devices and networks involves measurement of phase as well as magnitude. This is necessary for developing circuit models for simulation and to design matching circuits based on conjugate matching techniques. Time-domain characterization requires magnitude and phase information to perform the inverse-Fourier transform. Further, for best measurement accuracy, phase data is required to perform vector error correction which is common and an indispensable part of VNA measurement.

AVII.3.2 Block Diagram to Understand VNA Operation

The block diagram of a VNA is shown in Figure AVII.4 with reference to which its operation can be understood.

The VNA primarily consists of four sections—(i) source for stimulus, (ii) signal-separation devices, (iii) receivers that down-convert and detect the signals, and (iv) processor and display for depicting the results in the desired format and for reviewing.

The variable frequency CW *signal source* supplies the stimulus for the stimulus–response test system which is a sweep oscillator. Traditionally, network analyzers used a separate source. These sources were either based on open-loop VCOs which were cheaper (but inaccurate for narrowband measurements due to high phase-noise), or more expensive synthesized sweepers which provide better performance, especially for measuring narrowband devices. Today, network analyzers integrated with synthesized sources are used, which can provide excellent frequency resolution and stability.

The *signal separation* devices are directional couplers having low insertion loss and good isolation and directivity. There are two functions to be served by the signal separation hardware—to measure a portion of the incident signal to provide a reference for ratioing and to separate the incident (forward) and reflected (reverse) traveling waves—the latter coming from the device under test (DUT).

The next section of the network analyzer is the *signal detection* block. There are two basic ways of providing signal detection—diode detection

technique and tuned receiver technique. Diode detection is inherently scalar as phase information of the RF carrier is lost; however, they are of low cost and can provide a broad frequency range coverage at the cost of sensitivity and dynamic range. In the tuned receiver technique, an LO (a reference oscillator) is used to heterodyne the RF signal down to a lower IF. The LO is either locked to the RF or the IF signal so that the receivers in the network analyzer are always tuned to the RF signal present at the input. The IF signal is bandpass filtered, which narrows the receiver bandwidth and greatly improves sensitivity and dynamic range. Modern analyzers use an ADC and DSP to extract magnitude and phase information from the IF signal. Both vector network analyzers and spectrum analyzers use tuned-receiver approach.

The last major block in the hardware of a network analyzer is the *processor and display* section. This is where the reflection and transmission data are formatted in ways that make it easy to interpret the measurement results. Most network analyzers have features like linear and logarithmic sweeps, linear and log formats, polar plots, Smith chart, and so on.

In a generalized two-port VNA (Figure AVII.4) the two ports P_1 and P_2 of DUT are connected to the two test ports of the VNA via suitable adapter connectors A_1 and A_2 and precision connecting cables PC_1 and PC_2, respectively. The input reflection coefficient (S_{11}) and forward transmission

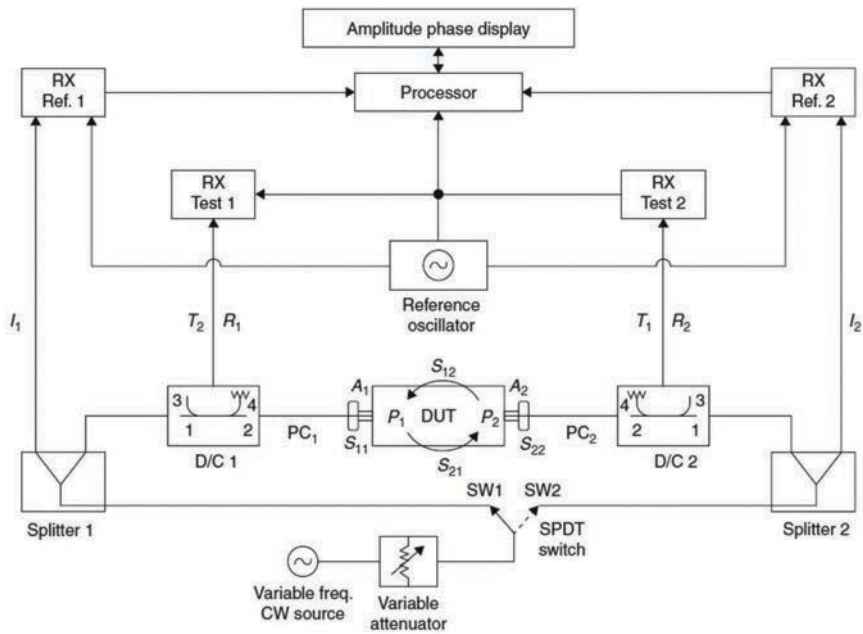


Figure AVII.4 Block diagram of a vector network analyzer.

coefficient (S_{21}) are measured at port P_1 in the forward path while the output reflection coefficient (S_{22}) and reverse transmission coefficient (S_{12}) are measured at port P_2 in the reverse path. The test frequency is generated by a variable frequency CW source and its power level is set with the help of a variable attenuator. The SPDT switch can be connected either to position SW1 or position SW2. When it is in position SW1, the test signal is incident on the DUT at P_1, which is appropriate for measuring S_{11} and S_{21}. The test signal is fed by SW1 to the common port of splitter 1, one arm of which (the reference channel) feeds the reference incident signal I_1 to a reference receiver for P_1 (RX Ref. 1) and the other (the test channel) connects to P_1 via the directional coupler D/C 1, PC_1, and A_1. The third port of D/C 1 couples off the power reflected from P_1 via A_1 and PC_1 as R_1 and finally feeds it to test receiver 1 (RX Test 1). Similarly, transmitted signals leaving P_2 (caused by incident signal at P_1) pass via A_2, PC_2, and D/C 2 as T_1 to RX Test 2. RX Ref. 1, RX Test 1, RX Ref. 2, and RX Test 2 are coherent receivers as they share the same reference oscillator and are capable of measuring the test signal's amplitude and phase at the test frequency being processed in the processor. When the SPDT switch is set to position SW2, the test signals are applied to P_2, the reference is measured by RX Ref. 2; reflections from P_2 are coupled off by D/C 2 as R_2 and measured by RX Test 2, and the signal leaving P_1 as transmitted signal T_2 is coupled off by D/C 1 and measured by RX Test 1. This position of SPDT switch is appropriate for measuring S_{22} and S_{12}.

A typical RF network analyzer can make measurements from 45 MHz to 40 GHz and with external mixer up to 110 GHz. The photograph of a modern VNA depicting the impedance locus of a microwave circuit on Smith chart was shown in Figure 3.1 of Chapter 3.

AVII.3.3 VNA Calibration

VNA calibration is an important issue that makes *vector error correction* for accurate measurement with such an analyzer. Calibration by the user (over and above the factory-made calibration) before making any measurement is critical to making good S-parameter measurements with VNA. User calibration is crucial to ensure the accuracy and repeatability of VNA measurements.

The VNA calibration can be understood on the basis of its error model. The VNA error model includes all possible areas in which systematic inaccuracies can arise—for example, the *directivity* of the couplers used to configure the reflectometer that separates the incident and reflected signals from DUT, the *match* at the input and output test ports of the DUT, the *frequency response* of the transmission line cable used between the DUT and the instrument ports, and *cross-talk* or *lack of perfect isolation* (leakage signal coupling) among the couplers.

While the VNA is a highly linear receiver and has sufficient spectral purity in its sources to make good measurements, there are a number of imperfections that limit measurements done without calibrations. These imperfections are related to *match* (VNA being broadband, a match more than the typical value of 20 dB generally used in practice, is desired to reduce potential errors), *directivity and isolation* (the directional couplers used in the heart of VNA even though are of very high quality, there may be a certain amount of coupled signal even with perfect termination, thereby deteriorating directivity, and can influence measurements of very small reflection coefficients; also the forward and reverse couplers might have leakage coupling or cross-talk with imperfect isolation between the two paths), and *frequency response* (any cables connected externally will have some frequency dependence that must be calibrated out for high-quality measurements even though the internal frequency response of the VNA could have been calibrated at the factory). These calibrations carried out before making every measurement with VNA are known as *vector error correction* of VNA.

The user calibration accounts for 12 sources of systematic errors in a VNA. For a forward measurement path, when the signal source (the test signal or stimulus) is connected to the DUT at port 1, the first source of systematic error is due to possible *impedance mismatch* at the input (source side) and output (load side) of the DUT. The second source of error is due to *non-ideal directivity* (the directivity not being infinite) which is true even with a very high quality practical directional coupler and little *imperfect isolation* caused by small amount of leakage radiation being coupled from port 1 to port 2 of DUT, that is, between the two directional couplers. The third source of error is caused by the *frequency dependence* of the signal paths (that of the cables) for the transmission and reflection measurement. Thus, for the forward measurement path there are in total six error terms and similarly six more error terms will arise when the signal source is connected to port 2 of the DUT for the reverse path, making the *total error terms* 12. In the error correction process, the measured magnitude and phase of each error signal appearing as a vector are applied mathematically via a proper algorithm to the measurement signal, thereby realizing *vector error correction*.

The most comprehensive VNA calibration is full two-port calibration method in which the parameters calibrated are S_{11}, S_{12}, S_{21}, and S_{22}. One may use other approaches, such as one-port and (frequency) response calibrations, but they are usually subsets of a full two-port calibration. The wide assortment of VNA user calibrations is divided into three commonly used categories or families— Short, Open, Load, Thru (SOLT), Thru, Reflect, Line (TRL), and automatic calibration. SOLT calibration is perhaps the most familiar of all VNA calibrations. It uses a well-defined short, open, and load (of characteristic impedance, usually 50 Ω or 75 Ω). Each standard is connected one by one at the reference plane so that the VNA can measure

them. When these steps are completed, one needs to connect the two reference planes together to form a through (thru) connection for the final measurement. The point where the calibration standards are connected defines the measurement plane. Also called the reference plane, this is the point at which the correction factors are applied and the measurement begins. For example, if coaxial cables are connected to the VNA ports, and the calibration standards are connected to the end of the cables during calibration, then the measurement plane is at the cable ends.

In recent years, automatic calibration modules have become a popular alternative to manual techniques such as SOLT and TRL. Because of internal switching, heating, and other uncertainties, early automatic calibration modules lacked the accuracy and repeatability of manual calibration. However, the current generation of modules offers accurate results and other benefits that make them an attractive choice.

Index

A
Accelerators/colliders, 1
Applegate diagram, 9
ABCD-parameters, 18
Available power gain, 40, 186
Amplifier with three-terminal microwave device, 182
Amplifier gain, 185
Automatic Design System (ADS) software, 228, 240, 353

B
Binomial stepped impedance transformer, 72

C
Cavity-stabilized Gunn oscillator, 172
Class A amplifier, 195
Class B and AB amplifier, 195
Class C amplifier, 196
Coaxial transmission line, 12
Composite right-/left-handed (CRLH) line, 157
Compressed Smith chart, 52
Computational electromagnetics (CEM), 348
Conjugate matching, 184
Coupled-line band-pass filter, 139
Coupled-resonator band-stop filter, 145
CRLH-type band-pass filter, 158
CSRR (complementary split-ring resonator), 274, 275
CSRR-loaded microstrip patch antenna design, 276

D
Decomposition rules, 35
Design of LNA and PA using EDA simulation tool, 196
Design of microwave power combiners, 228
Design of RF/microwave communication transmitter and receiver including radar system using EDA simulation tool, 304
Designer's considerations for RF/microwave transmitter and receiver design, 288
Distributed components, 18
Distributed parameters, 6
Double-stub matching, 57

E
Effective isotropic radiated power (EIRP), 287
Electromagnetic spectrum, 8
Electromagnetic wave equations, 16
Electronic design automation (EDA) tool, 104, 134, 145, 154, 197, 208, 255, 349
Elliptic function filter, 109
Equal-ripple/Tchebyscheff filter, 108
Equal-ripple/Tchebyscheff transformer, 72
Evanescent and propagation regions in waveguide, 15
Exponential taper, 89

F
Feeding techniques for microstrip patch antenna 268
Filter transformations from prototype, 115
Four-way power combiner using short-slot hybrid 225
Friss' formula, 195
Fundamental theoretical background of RF/microwave diode oscillator/amplifier, 170

G
Gain compression, 298

H
HEMT device, 10
Heterojunction transistor, 10
High Frequency Structure Simulator (HFSS), 228, 270, 352
Homojunction transistor, 10

I
Image frequency, 246
Image-rejection mixer, 253
IMPATT oscillator and VCO design with 3D electromagnetic field simulator, 174
Impedance (*K*) or admittance (*J*) inverters, 126
Impedance matching with lumped-component/planar distributed-line reactive elements, 97
Impedance transformers, 67
Injection-locked oscillator/amplifier, 172
Insertion loss method, 105
Intercept point, 302
Intermodulation distortion, 289, 300
Intermodulation products, 247, 301
Intermodulation rejection, 291

K
Klopfenstein taper, 91
Kuroda's identities, 122

L
Link budget, 291
Load-pull, 289
Load-pull test, 209
Lorentz gauge condition, 362
Low-noise amplifier (LNA), 195, 288
Low-noise amplifier (LNA) *vs.* power amplifier (PA), 195
Low-pass prototype filter design technique, 110
Lumped circuit analysis, 5
Lumped components, 18

M
Mason's rule, 35
MATLAB programs, 331
Maximally flat/binomial transformer, 72
Maximally flat/Butterworth filter, 106
Maxwell's equations and electromagnetic wave equations—a brief review, 355
Metamaterial-inspired filter design, 157
Metamaterial-inspired microstrip patch antenna design, 274
Metamaterials or left-handed materials, 157
Microstrip line, 16
Microstrip patch antenna (MPA), 263
Microstrip patch antenna design, 267
Microwave filter design with CRLH transmission line, 164
Microwave frequency bands, 6, 229
Microwave mixers, 246
Mixer design using EDA simulation tool, 255
MoM, FEM, and FDTD—an overview, 350

N
Network analyzer basics, 371
Noise and distortion, 293
Nonlinearity and distortion, 297
N-way in single-stage power combiner, 223

O
1-dB compression point, 299
Operating power gain or simply power gain, 42, 186
Oscillator configurations with IMPATT/Gunn diodes, 343

P
Planar transmission lines, 16
Properties of S-parameters, 23
Power amplifier, 195, 208
Phase-noise, 294, 295
Phase-noise plot of an oscillator in spectrum analyzer, 297

Q
Quarter-wave impedance transformer, 68
Quasi-TEM propagation, 16

R
Receiver sensitivity, 290
Reflection type amplifier, 172
Retarded magnetic vector potential, 362
RF and microwave band designation, 229
RF/microwave solid-state devices, 7
RF/microwave tube devices, 6
RF transmission line, 12
Richard's transformation, 119

S
Short-slot hybrid and branch-line quadrature hybrid, 346
Signal flow graphs, 34
Simulators:
  circuit simulators and EM simulators, 351
Single-balanced and double-balanced mixers, 249
Single-device/single-ended diode mixer, 249

Single-stub matching, 54
S-matrix of directional coupler, 29
S-matrix of magic tee, 32
Smith and Jones chart, 50, 99, 100
Smith chart, 43, 49, 59, 62, 65
Some technical aspects related to RF/microwave transmitter and receiver, 287
S-parameters, 4, 18, 19, 23
Spatial power combining, 227
Spectrum analyzer, 297, 365
Split-ring resonator (SRR), 157
SRR-loaded waveguide filter, 157
Stability considerations for amplifier, 186
Stepped-impedance (high-low) low-pass filter, 129
Stub matching, 53
Sub-harmonically pumped mixer, 252
SystemVue EDA simulation tool, 305, 318, 354

T
Taper with triangular distribution, 89
Tapered impedance matching transformers, 88
Tchebyscheff stepped impedance transformer, 80
TE mode or H-wave, 13
TEM mode, 14
Theory of small reflections, 69
Third order intercept point (IP3), 302
3D electromagnetic field simulator, 170, 174, 180
T-matrix, 19
TM mode or E-wave, 13
Transducer power gain, 39, 187
Transmission line sections and their LC equivalents, 11
Twin-cap combiner, 223
Twin-wire line, 11

U
Unity property of S-matrix, 24

V
Vector network analyzer, 44
VNA calibration, 374

W
Waveguide, 13
Waveguide and planar transmission line-based RF/microwave passive components, 5
Wilkinson power combiner/divider, 225

X
X-parameters, 19

Z
Z-Y Smith chart, 97



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com