

POWER VACUUM TUBES HANDBOOK

THIRD EDITION

Jerry C. Whitaker



CRC Press
Taylor & Francis Group

**POWER
VACUUM
TUBES
HANDBOOK**

Third Edition

ELECTRONICS HANDBOOK SERIES

Series Editor:

Jerry C. Whitaker

Technical Press

Morgan Hill, California

AC POWER SYSTEMS HANDBOOK, SECOND EDITION

Jerry C. Whitaker

THE COMMUNICATIONS FACILITY DESIGN HANDBOOK

Jerry C. Whitaker

DESIGN FOR RELIABILITY

Dana Crowe and Alec Feinberg

THE ELECTRONIC PACKAGING HANDBOOK

Glenn R. Blackwell

ELECTRONIC SYSTEMS MAINTENANCE HANDBOOK, SECOND EDITION

Jerry C. Whitaker

POWER VACUUM TUBES HANDBOOK, THIRD EDITION

Jerry C. Whitaker

THE RESOURCE HANDBOOK OF ELECTRONICS

Jerry C. Whitaker

THE RF TRANSMISSION SYSTEMS HANDBOOK

Jerry C. Whitaker

THERMAL DESIGN OF ELECTRONIC EQUIPMENT

Ralph Remsburg

POWER VACUUM TUBES HANDBOOK

Third Edition

Jerry C. Whitaker



CRC Press

Taylor & Francis Group

Boca Raton London New York

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

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2012 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20120119

International Standard Book Number-13: 978-1-4398-5065-7 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. 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, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

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

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

*This book is dedicated to my daughter, Alexis, as she begins
her adventure at UCLA. (Go Bruins!)*

Contents

Preface.....	xxiii
Acknowledgment.....	xxv
Author.....	xxvii
1. Power Vacuum Tube Applications	1
1.1 Introduction	1
1.2 Vacuum Tube Development	1
1.2.1 Pioneer Developers.....	2
1.2.1.1 Radio Central	4
1.2.1.2 WLW: The Nation's Station	4
1.2.1.3 UHF: A New Technical Challenge.....	5
1.2.1.4 Birth of the Klystron	6
1.2.1.5 Nuclear Magnetic Resonance	8
1.2.2 Transistor Is Born.....	8
1.2.3 Satellite Technology.....	9
1.2.4 Standardization.....	10
1.2.5 Transmission Systems	11
1.3 Vacuum Tube Applications.....	12
1.3.1 Market Overview	13
1.4 Electromagnetic Radiation Spectrum	14
1.4.1 Low Frequency: 30–300 kHz	15
1.4.2 Medium Frequency: 300 kHz–3 MHz	15
1.4.3 High Frequency: 3–30 MHz.....	16
1.4.4 Very High Frequency: 30–300 MHz	16
1.4.5 Ultrahigh Frequency: 300 MHz–3 GHz.....	16
1.4.6 Superhigh Frequency: 3–30 GHz	17
References	17
Bibliography	17
2. Modulation Systems and Characteristics	19
2.1 Introduction	19
2.2 Modulation Systems	19
2.2.1 Principles of Resonance	21
2.2.1.1 Series Resonant Circuits.....	21
2.2.1.2 Parallel Resonant Circuits.....	22
2.2.1.3 Cavity Resonators	23
2.2.2 Frequency Sources	25
2.2.2.1 Frequency Stabilization	27
2.2.2.2 Equivalent Circuit of a Quartz Resonator.....	27

2.2.2.3	Temperature Compensation	29
2.2.2.4	Stress Compensation	30
2.2.2.5	Aging Effects.....	30
2.3	Operating Class.....	31
2.3.1	Operating Efficiency.....	33
2.3.2	Broadband Amplifier Design.....	33
2.3.2.1	Stagger Tuning.....	33
2.3.2.2	Matching Circuits.....	34
2.3.2.3	Power Combining.....	34
2.4	Thermal and Circuit Noise.....	35
2.4.1	Thermal Noise.....	37
2.4.2	Noise in Systems of Cascaded Stages	37
2.5	Amplitude Modulation	37
2.5.1	High-Level AM Modulation.....	41
2.5.2	Vestigial-Sideband Amplitude Modulation.....	42
2.5.3	Single-Sideband Amplitude Modulation	42
2.5.4	Quadrature Amplitude Modulation	45
2.6	Frequency Modulation.....	46
2.6.1	Modulation Index	47
2.6.2	Phase Modulation.....	50
2.6.3	Modifying FM Waves.....	51
2.6.4	Preemphasis and De-Emphasis	51
2.6.5	Modulation Circuits	51
2.6.5.1	Direct-FM Modulator.....	52
2.6.5.2	VCO Direct-FM Modulator	54
2.6.6	FM Power Amplifiers	54
2.7	Pulse Modulation.....	56
2.7.1	Digital Modulation Systems.....	57
2.7.2	Pulse Amplitude Modulation	57
2.7.3	Pulse Time Modulation.....	59
2.7.4	Pulse Code Modulation	61
2.7.5	Delta Modulation.....	62
2.7.6	Digital Coding Systems	62
2.7.7	Baseband Digital Pulse Modulation	64
2.8	Spread Spectrum Systems	66
	References	67
	Bibliography.....	68
3.	Vacuum Tube Principles.....	69
3.1	Introduction.....	69
3.2	Characteristics of Electrons	69
3.2.1	Electron Optics.....	70
3.2.2	Magnetic Field Effects.....	71

3.2.3	Thermal Emission from Metals	72
3.2.4	Secondary Emission	73
3.3	Types of Vacuum Tubes.....	75
3.3.1	Diode.....	75
	3.3.1.1 Applications	76
3.3.2	Triode.....	77
3.3.3	Tetrode.....	82
	3.3.3.1 Application Example.....	83
3.3.4	Pentode	87
	3.3.4.1 Application Example.....	88
3.4	High-Frequency Operating Limits	91
3.4.1	Transit-Time Effects	93
3.5	Vacuum Tube Design.....	95
3.5.1	Device Cooling.....	95
	3.5.1.1 Air Cooling.....	96
	3.5.1.2 Water Cooling	96
	3.5.1.3 Vapor-Phase Cooling	98
	3.5.1.4 Special Applications.....	99
3.5.2	Cathode Assembly	99
	3.5.2.1 Oxide Cathode	100
	3.5.2.2 Thoriated-Tungsten Cathode	101
	3.5.2.3 Tungsten-Impregnated Cathode	102
	3.5.2.4 Cathode Construction	102
	3.5.2.5 Velocity of Emission.....	104
3.5.3	Grid Structures.....	104
	3.5.3.1 Wire Grids	104
	3.5.3.2 K-Grid	106
	3.5.3.3 Pyrolytic Grid	106
	3.5.3.4 Grid Physical Structure	107
	3.5.3.5 Secondary Emission Considerations	108
3.5.4	Plate Assembly	111
	3.5.4.1 Ceramic Elements.....	112
	3.5.4.2 Aluminum Oxide Ceramics.....	113
	3.5.4.3 Beryllium Oxide Ceramics.....	114
	3.5.4.4 Other Ceramics.....	114
3.6	Tube Construction.....	115
3.6.1	Connection Points.....	116
3.6.2	Tube Sockets	116
3.7	Neutralization.....	117
3.7.1	Circuit Analysis.....	118
3.7.2	Circuit Design	119
	3.7.2.1 Below VHF	120
	3.7.2.2 VHF and Above.....	124
3.7.3	Grounded-Grid Amplifier Neutralization	124

3.7.4	Grid impedance.....	126
3.7.4.1	Application Example.....	126
3.7.5	Self-Neutralizing Frequency	127
3.7.6	Neutralization Adjustment.....	130
3.8	Electron Tube Characteristics.....	131
3.8.1	Interpretation of Tube Data	134
	References	136
	Bibliography	136
4.	Designing Vacuum Tube Systems.....	139
4.1	Introduction	139
4.1.1	Class A Amplifier	140
4.1.2	Class B and AB Amplifiers	140
4.1.3	Class C Amplifier.....	140
4.2	Principles of RF Power Amplification.....	142
4.2.1	Drive Power Requirements.....	144
4.2.1.1	Operational Considerations for VHF and UHF	145
4.2.2	Mechanical and Electrical Considerations.....	146
4.2.2.1	Lead Lengths	147
4.2.2.2	Power Supply Considerations.....	147
4.2.3	Bypassing Tube Elements	148
4.2.3.1	Filament Bypassing.....	148
4.2.3.2	Screen and Suppressor Grid Bypassing.....	149
4.2.3.3	Application Examples	149
4.2.4	Parasitic Oscillations	150
4.2.4.1	Dynatron Oscillation	154
4.2.4.2	Harmonic Energy	154
4.2.5	Shielding	155
4.2.5.1	Pierced Shields.....	155
4.2.5.2	Metal Base Shells and Submounted Sockets.....	156
4.2.5.3	Compartments	156
4.2.6	Protection Measures	157
4.2.7	Cavity Amplifier Systems.....	158
4.2.8	Bandwidth and Efficiency	160
4.2.9	Current Paths.....	161
4.2.9.1	Input Circuit.....	161
4.2.9.2	Output Circuit	163
4.2.10	1/4-Wavelength Cavity.....	164
4.2.10.1	Tuning the Cavity.....	167
4.2.11	1/2-Wavelength Cavity.....	168
4.2.11.1	Lengthening the Plate Line.....	171
4.2.12	Folded 1/2-Wavelength Cavity	173
4.2.13	Wideband Cavity	174

4.2.14	Output Coupling	175
4.2.14.1	Inductive Coupling	175
4.2.14.2	Capacitive Coupling.....	177
4.2.14.3	PA Loading.....	179
4.2.15	Mechanical Design	181
4.3	High-Voltage Power Supplies	184
4.3.1	Silicon Rectifiers.....	184
4.3.1.1	Operating Rectifiers in Series.....	185
4.3.1.2	Operating Rectifiers in Parallel.....	186
4.3.1.3	Silicon Avalanche Rectifiers.....	187
4.3.2	Thyristor Servo Systems	188
4.3.2.1	Inductive Loads.....	189
4.3.2.2	Applications	191
4.3.2.3	Triggering Circuits.....	193
4.3.2.4	Fusing.....	196
4.3.2.5	Control Flexibility	196
4.3.3	Polyphase Rectifier Circuits	196
4.3.4	Power Supply Filter Circuits.....	197
4.3.4.1	Inductive Input Filter.....	200
4.3.4.2	Capacitive Input Filter	202
4.3.5	Parameter Sampling Circuits	204
4.3.6	Protection Systems.....	207
	References	208
	Bibliography	208
5.	Applying Vacuum Tube Devices	209
5.1	Introduction	209
5.2	AM Power Amplification Systems.....	209
5.2.1	Control Grid Modulation.....	210
5.2.2	Suppressor Grid Modulation	212
5.2.3	Cathode Modulation.....	212
5.2.4	High-Level AM Amplification	213
5.2.4.1	Class B Modulator	215
5.2.5	Pulse Width Modulation.....	216
5.3	Linear Amplification	219
5.3.1	Device Selection	219
5.3.2	Grid-Driven Linear Amplifier.....	220
5.3.3	Cathode-Driven Linear Amplifier	220
5.3.4	Intermodulation Distortion	221
5.4	High-Efficiency Linear Amplification.....	223
5.4.1	Chireix Outphasing Modulated Amplifier	223
5.4.2	Doherty Amplifier	225
5.4.3	Screen-Modulated Doherty-Type Amplifier	227
5.4.4	Terman-Woodyard Modulated Amplifier	228
5.4.5	Dome Modulated Amplifier.....	229

5.5	Television Power Amplifier Systems	231
5.5.1	System Considerations	231
5.5.2	Power Amplifier	233
5.5.2.1	Application Example	235
5.6	FM Power Amplifier Systems	236
5.6.1	Cathode-Driven Triode Amplifier	237
5.6.2	Grounded-Grid vs. Grid-Driven Tetrode	238
5.6.3	Grid-Driven Tetrode/Pentode Amplifiers	239
5.6.4	Impedance Matching into the Grid	241
5.6.4.1	Interstage Coupling	242
5.6.5	Neutralization	243
5.7	Special-Application Amplifiers	244
5.7.1	Distributed Amplification	244
5.7.2	Radar	245
5.7.2.1	Modulator	246
	References	249
	Bibliography	249
6.	Microwave Power Tubes	251
6.1	Introduction	251
6.1.1	Linear-Beam Tubes	252
6.1.2	Crossed-Field Tubes	253
6.1.2.1	Crossed-Field Amplifier	254
6.2	Grid Vacuum Tubes	255
6.2.1	Planar Triode	256
6.2.2	High-Power UHF Tetrode	258
6.2.3	Diacrode	259
6.3	Klystron	261
6.3.1	Reflex Klystron	261
6.3.2	Two-Cavity Klystron	264
6.3.2.1	Two-Cavity Klystron Oscillator	266
6.3.2.2	Two-Cavity Klystron Amplifier	266
6.3.3	Multicavity Klystron	266
6.3.3.1	Operating Principles	267
6.3.3.2	Emission	270
6.3.3.3	Modulating Anode	272
6.3.3.4	Magnetic Field	272
6.3.3.5	RF Structure	275
6.3.3.6	Phased Electron Operation	277
6.3.3.7	Types of Devices	279
6.3.4	Beam Pulsing	280
6.3.5	Integral vs. External Cavity	282
6.3.5.1	Number of Cavities	283
6.3.5.2	Efficiency	284
6.3.5.3	Performance Trade-Offs	284

6.3.6	MSDC Klystron	285
6.3.6.1	Theory of Operation.....	285
6.3.6.2	Electron Trajectories.....	286
6.3.6.3	Mechanical Construction	289
6.3.6.4	MSDC Power Supply	291
6.3.6.5	Device Performance.....	291
6.3.6.6	Applying the MSDC Klystron.....	294
6.4	Klystrode/Inductive Output Tube.....	296
6.4.1	Theory of Operation	296
6.4.2	Electron Gun.....	297
6.4.3	Grid Structure.....	299
6.4.4	Input Cavity.....	300
6.4.5	Output Cavity.....	301
6.4.6	Application Considerations.....	302
6.4.7	Continuing Research Efforts	304
6.5	Constant Efficiency Amplifier.....	304
6.5.1	Theory of Operation	304
6.6	Traveling Wave Tube	306
6.6.1	Theory of Operation	308
6.6.1.1	Interaction Circuit	308
6.6.1.2	Pulse Modulation	310
6.6.1.3	Electron Gun	310
6.6.1.4	Beam Focusing.....	311
6.6.1.5	Collector Assembly	313
6.6.2	Operating Efficiency.....	314
6.6.3	Operational Considerations	314
6.6.3.1	Intermodulation Distortion.....	315
6.6.3.2	Second-Harmonic Content.....	315
6.6.3.3	AM/PM Conversion.....	315
6.6.3.4	Phase Variation.....	315
6.7	Crossed-Field Tubes.....	316
6.7.1	Magnetron.....	316
6.7.1.1	Operating Principles.....	317
6.7.1.2	Coaxial Magnetron.....	319
6.7.1.3	Frequency-Agile Magnetron.....	319
6.7.1.4	Linear Magnetron	320
6.7.2	Backward Wave Oscillator.....	321
6.7.3	Strap-Fed Devices	322
6.7.4	Gyrotron.....	323
6.7.4.1	Theory of Operation.....	324
6.7.4.2	Gyrotron Design Variations.....	326
6.8	Other Microwave Devices.....	328
6.8.1	Quasiquantum Devices.....	329
6.8.2	Variations on the Klystron.....	330

6.9	Microwave Tube Life	331
6.9.1	Life-Support System	331
6.9.2	Protection Measures	332
	6.9.2.1 Heater Supply	333
	6.9.2.2 Beam Supply	333
	6.9.2.3 Magnet Supply	333
	6.9.2.4 RF Circuits	334
6.9.3	Filament Voltage Control	334
6.9.4	Cooling System	335
6.9.5	Reliability Statistics	335
	6.9.5.1 MTTR	335
	References	336
	Bibliography	337
7.	RF Interconnection and Switching	339
7.1	Introduction	339
	7.1.1 Skin Effect	339
7.2	Coaxial Transmission Line	341
	7.2.1 Electrical Parameters	342
	7.2.1.1 Transverse Electromagnetic Mode	342
	7.2.1.2 Dielectric	343
	7.2.1.3 Impedance	343
	7.2.1.4 Resonant Characteristics	345
7.2.2	Electrical Considerations	345
7.2.3	Coaxial Cable Ratings	347
	7.2.3.1 Power Rating	347
	7.2.3.2 Connector Effects	350
	7.2.3.3 Attenuation	350
	7.2.3.4 Phase Stability	354
7.2.4	Mechanical Parameters	354
7.3	Waveguide	355
	7.3.1 Propagation Modes	356
	7.3.1.1 Dual-Polarity Waveguide	358
	7.3.1.2 Efficiency	359
7.3.2	Ridged Waveguide	359
7.3.3	Circular Waveguide	359
	7.3.3.1 Parasitic Energy	360
7.3.4	Doubly Truncated Waveguide	361
7.3.5	Impedance Matching	362
	7.3.5.1 Waveguide Filters	363
7.3.6	Installation Considerations	364
	7.3.6.1 Tuning	364
	7.3.6.2 Waveguide Hardware	365
7.3.7	Cavity Resonators	365

7.4	RF Combiner and Diplexer Systems.....	367
7.4.1	Passive Filters	367
7.4.1.1	Filter Type.....	368
7.4.1.2	Filter Alignment	368
7.4.1.3	Filter Order	371
7.4.2	Four-Port Hybrid Combiner	372
7.4.3	Nonconstant-Impedance Diplexer.....	374
7.4.4	Constant-Impedance Diplexer	376
7.4.4.1	Band-Stop Diplexer	377
7.4.4.2	Bandpass Constant-Impedance Diplexer.....	378
7.4.4.3	Isolation of f_1 to f_2	380
7.4.4.4	Intermodulation Products.....	380
7.4.4.5	Group Delay	382
7.4.5	Microwave Combiners	383
7.4.6	Hot-Switching Combiners	383
7.4.6.1	Phase Relationships	386
7.4.6.2	Variable-Dielectric Vane.....	388
7.4.6.3	Dielectric Posts.....	388
7.4.6.4	Variable-Phase Hybrid.....	389
7.4.7	Phased-Array Antenna Systems.....	390
7.4.7.1	Phase-Shift Devices.....	393
7.4.7.2	Radar System Duplexer	394
7.5	High-Power Isolators	396
7.5.1	Theory of Operation	397
7.5.2	Applications.....	398
7.5.2.1	Hot Switch	400
7.5.2.2	Diplexer.....	400
7.5.2.3	Multiplexer	400
	References	401
	Bibliography	401
8.	Properties of Materials.....	403
8.1	Metals.....	403
8.1.1	Skin Depth	406
8.1.2	Heat Conduction	407
8.1.3	Thermal Expansion	407
8.2	Dielectrics.....	409
8.2.1	Frequency Dependence of the Properties	412
8.3	Tabular Data.....	413
8.3.1	Electrical Resistivity of Pure Metals	413
8.3.2	Electrical Resistivity of Selected Alloys	413
8.3.3	Resistance of Wires.....	413
8.3.4	Dielectric Constants of Common Materials.....	413
	References	432
	Bibliography	432

9. Cooling Considerations	435
9.1 Introduction.....	435
9.1.1 Thermal Properties.....	435
9.1.2 Heat Transfer Mechanisms.....	436
9.1.2.1 Conduction.....	436
9.1.2.2 Convection.....	438
9.1.2.3 Radiation.....	438
9.1.3 Physics of Boiling Water.....	439
9.2 Application of Cooling Principles.....	441
9.2.1 Forced-Air Cooling Systems.....	442
9.2.1.1 Cooling Airflow Data.....	443
9.2.1.2 Blower Selection.....	448
9.2.2 Water Cooling.....	448
9.2.2.1 Cooling System Design.....	449
9.2.2.2 Water Purity and Resistivity.....	451
9.2.2.3 Condensation.....	453
9.2.2.4 Preventive Maintenance.....	453
9.2.3 Vapor-Phase Cooling.....	454
9.2.3.1 Anode Design.....	456
9.2.3.2 Boiler.....	457
9.2.3.3 Insulating Tubing.....	459
9.2.3.4 Control Box.....	459
9.2.3.5 Equalizer Line.....	461
9.2.3.6 Condenser.....	461
9.2.3.7 Pressure Interlock.....	461
9.2.3.8 Piping.....	463
9.2.3.9 Automatic Refilling System.....	464
9.2.3.10 Alternative Vapor Cooling Systems.....	465
9.2.3.11 Maintenance.....	466
9.2.4 Temperature Measurements.....	466
9.2.4.1 “Crayon” Temperature Measurement.....	467
9.2.4.2 Phase-Change Fluid.....	468
9.2.4.3 Selection Process.....	469
9.2.5 Air-Handling System.....	469
9.3 Operating Environment.....	471
9.3.1 Air-Handling System.....	471
9.3.2 Air Cooling System Design.....	472
9.3.2.1 Case 1.....	473
9.3.2.2 Case 2.....	474
9.3.2.3 Case 3.....	475
9.3.3 Site Design Guidelines.....	475
9.3.3.1 Closed Site Design.....	477
9.3.3.2 Open Site Design.....	478
9.3.3.3 Hybrid Design.....	480

- 9.3.4 Water/Vapor Cooling System Maintenance 481
 - 9.3.4.1 Tests for Purity..... 482
 - 9.3.4.2 Foaming Test for Water Purity 483
- References 484
- Bibliography 485
- 10. Reliability Considerations..... 487**
 - 10.1 Introduction 487
 - 10.1.1 Terminology..... 487
 - 10.2 Quality Assurance 489
 - 10.2.1 Inspection Process..... 490
 - 10.2.2 Reliability Evaluation 490
 - 10.2.2.1 Parts-Count Method 491
 - 10.2.2.2 Stress-Analysis Method..... 491
 - 10.2.3 Failure Analysis..... 491
 - 10.2.4 Standardization..... 492
 - 10.3 Reliability Analysis..... 494
 - 10.3.1 Statistical Reliability 495
 - 10.3.1.1 Roller-Coaster Hazard Rate..... 497
 - 10.3.2 Environmental Stress Screening 498
 - 10.3.3 Latent Defects..... 501
 - 10.3.4 Operating Environment..... 503
 - 10.3.5 Failure Modes..... 503
 - 10.3.6 Maintenance Considerations..... 504
 - 10.3.6.1 Common-Mode Failure 504
 - 10.3.6.2 Spare Parts..... 505
 - 10.4 Vacuum Tube Reliability 505
 - 10.4.1 Thermal Cycling..... 506
 - 10.4.2 Tube-Changing Procedure..... 507
 - 10.4.3 Power Tube Conditioning..... 509
 - 10.4.3.1 Power Supply..... 510
 - 10.4.3.2 Conditioning Procedure 511
 - 10.4.3.3 Considerations for Very Large Tubes 511
 - 10.4.3.4 Safety..... 512
 - 10.4.4 Filament Voltage..... 512
 - 10.4.4.1 Black Heat..... 513
 - 10.4.5 Filament Voltage Management 514
 - 10.4.5.1 Klystron Devices 514
 - 10.4.6 PA Stage Tuning..... 516
 - 10.4.7 Fault Protection 517
 - 10.4.8 Vacuum Tube Life 519
 - 10.4.8.1 Catastrophic Failures..... 520
 - 10.4.9 Examining Tube Performance..... 521
 - 10.4.10 Shipping and Handling Vacuum Tubes 524

10.5	Klystron Reliability	525
10.5.1	Cleaning and Flushing the Cooling System	526
10.5.1.1	Transmitter Flushing	526
10.5.1.2	Flushing Klystron Water Lines	527
10.5.1.3	Cleaning Klystron Water Lines	528
10.5.1.4	Flushing and Cleaning Magnet Water Lines	529
10.5.1.5	General Cleaning	529
10.5.2	Cleaning Ceramic Elements	529
10.5.3	Reconditioning Klystron Gun Elements	530
10.5.4	Focusing Electromagnet Maintenance	531
10.5.5	Power Control Considerations	532
10.5.5.1	Primary Power Interruption	533
	References	534
	Bibliography	535
11.	Device Performance Criteria	539
11.1	Introduction	539
11.2	Measurement Parameters	540
11.2.1	Power Measurements	540
11.2.1.1	Root Mean Square	541
11.2.1.2	Average-Response Measurement	542
11.2.1.3	Peak-Response Measurement	543
11.2.1.4	Measurement Bandwidth	546
11.2.1.5	Meter Accuracy	546
11.2.1.6	Radio Frequency Power Measurement	547
11.2.2	Decibel Measurement	547
11.2.3	Noise Measurement	549
11.2.4	Phase Measurement	550
11.2.4.1	Relation to Frequency	551
11.2.5	Nonlinear Distortion	552
11.2.5.1	Harmonic Distortion	552
11.2.5.2	Intermodulation Distortion	557
11.2.5.3	Measurement Techniques	558
11.2.5.4	Addition and Cancellation of Distortion Components	559
11.2.5.5	Intermodulation Precorrection Techniques	561
11.3	Vacuum Tube Operating Parameters	564
11.3.1	Stage Tuning	564
11.3.1.1	Power Grid Tubes	564
11.3.1.2	Screen Voltage	565
11.3.1.3	Back-Heating by Electrons	566

11.3.2	Amplifier Balance	568
11.3.3	Parallel Tube Amplifiers	569
11.3.4	Harmonic Energy.....	569
11.3.4.1	Controlling Harmonics	570
11.3.5	Klystron Tuning Considerations.....	570
11.3.6	Intermodulation Distortion	573
11.3.6.1	Generation of Radio Frequency Intermodulation Products.....	574
11.3.6.2	Intermodulation as a Function of Turnaround Loss.....	574
11.4	Voltage Standing Wave Ratio	576
11.5	Radio Frequency System Performance	579
11.5.1	Key System Measurements.....	580
11.5.2	Synchronous AM in FM Systems	580
11.5.2.1	Bandwidth Limiting.....	581
11.5.2.2	Effects of Synchronous Amplitude Modulation.....	582
11.5.3	Incidental Phase Modulation	584
11.5.3.1	Bandwidth Considerations	584
11.5.3.2	Corrective Procedures	584
11.5.4	Carrier Amplitude Regulation.....	586
11.5.5	Site-Related Intermodulation Products	587
11.5.5.1	Receiver Intermod.....	588
11.5.5.2	Hardware Intermod.....	588
	References	588
	Bibliography	589
12.	RF System Maintenance and Troubleshooting	591
12.1	Introduction	591
12.2	Routine Maintenance.....	591
12.2.1	Maintenance Log	591
12.2.1.1	Case Study #1.....	592
12.2.1.2	Case Study #2.....	593
12.2.2	Preventive Maintenance Routine	594
12.2.2.1	Resistors and Capacitors	594
12.2.2.2	Power Supply Components.....	596
12.2.2.3	Coils and RF Transformers.....	596
12.2.2.4	Relay Mechanisms.....	598
12.2.2.5	Connection Points	599
12.2.3	Cleaning the System.....	599
12.3	Transmission Line and Antenna	602
12.3.1	Effects of Modulation	603
12.3.2	Maintenance Considerations.....	604

- 12.4 High-Voltage Power Supply 605
 - 12.4.1 Power Supply Maintenance 605
 - 12.4.2 Power Supply Metering 607
 - 12.4.3 Overload Sensor 607
 - 12.4.4 Transient Disturbances 608
 - 12.4.5 Single Phasing 608
- 12.5 Troubleshooting Procedure 610
 - 12.5.1 Factory Service Assistance 611
 - 12.5.2 Plate Overload Fault 612
 - 12.5.2.1 Troubleshooting Procedure 612
 - 12.5.2.2 Process of Elimination 614
 - 12.5.3 Radio Frequency System Faults 616
 - 12.5.3.1 Component Substitution 618
 - 12.5.3.2 VSWR Overload 619
 - 12.5.4 Power Control Faults 620
 - 12.5.4.1 Thyristor Control System 621
 - 12.5.4.2 Interlock Failures 621
 - 12.5.4.3 Step-Start Faults 623
 - 12.5.5 Protection Circuits 624
- 13. Safe Handling of Vacuum Tube Devices 627**
 - 13.1 Introduction 627
 - 13.2 Electric Shock 627
 - 13.2.1 Effects on the Human Body 628
 - 13.2.2 Circuit Protection Hardware 630
 - 13.2.3 Working with High Voltage 631
 - 13.2.3.1 Radio Frequency Considerations 632
 - 13.2.4 First Aid Procedures 632
 - 13.3 Operating Hazards 634
 - 13.3.1 OSHA Safety Considerations 634
 - 13.3.1.1 Protective Covers 636
 - 13.3.1.2 Identification and Marking 636
 - 13.3.1.3 Grounding 636
 - 13.3.2 Beryllium Oxide Ceramics 638
 - 13.3.3 Corrosive and Poisonous Compounds 638
 - 13.3.4 FC-75 Toxic Vapor 639
 - 13.4 Nonionizing Radiation 639
 - 13.4.1 NEPA Mandate 640
 - 13.4.2 Revised Guidelines 641
 - 13.4.3 Multiple-User Sites 642
 - 13.4.4 Operator Safety Considerations 644
 - 13.4.5 X-Ray Radiation Hazard 644
 - 13.4.6 Implosion Hazard 644
 - 13.4.7 Hot Coolant and Surfaces 644

13.4.8 Polychlorinated Biphenyls	645
13.4.8.1 Governmental Action	646
References	648
Bibliography	648
Appendix: Mathematics, Symbols, and Physical Constants	649

Preface

The term “high technology” is perhaps one of the more overused descriptions in our technical vocabulary. It is a term generally reserved for discussion of integrated circuits, fiber optics, satellite systems, and computers. Few people would associate high technology with vacuum tubes. The notion that vacuum tube construction is more art than science may have been true 20 or 30 years ago, but today it is a different story.

The demand on the part of industry for tubes capable of higher operating power and frequency and the economic necessity for tubes that provide greater efficiency and reliability have moved power tube manufacturers further into the high-tech arena. Advancements in tube design and construction have given end users new transmitters and RF generators that allow industry to grow and prosper.

If you bring up the subject of vacuum tubes to someone who has never worked on a transmitter or high-power RF generator, you are likely to get a blank stare and a question: “Do they make those anymore?” Although receiving tubes have more or less disappeared from the scene, power tubes are alive and well and are performing vital functions in thousands of divergent applications. Solid-state and tube technologies each have their place, each with its strengths and weaknesses. It should be noted that even receiving tubes are staging somewhat of a comeback in high-end audio applications.

Tube design and development, although accompanied by less fanfare, are advancing as are developments in solid-state power technology. Power tubes today are designed with an eye toward high operating efficiency and high gain/bandwidth properties. Above all, a tube must be reliable and provide long operating life. The design of a new power tube is a lengthy process that involves computer-aided calculations and advanced modeling.

Despite the inroads made by solid-state technology, the power vacuum tube occupies—and will continue to play—an important role in the generation of high-power radio frequency energy in the high-frequency regions and above. No other device can do the job as well. Certainly, solid-state technology cannot, especially if cost, size, and weight are important considerations.

The field of science encompassed by power vacuum tubes is broad and exciting. It is an area of enormous importance to military and industrial customers, and a discipline in which significant research is being conducted.

Power vacuum tubes include a wide range of devices, each for specific applications. Devices include power grid tubes (triodes, tetrodes, and pentodes) and microwave power tubes (klystrons, traveling wave tubes, gyrotrons, and numerous other high-frequency devices). Research is being conducted for both tube classes to extend output power and maximum frequency and to improve operating efficiency.

This book examines the underlying technology of each type of power vacuum tube device in common use today and provides examples of typical applications. This third edition of the *Power Vacuum Tubes Handbook* is directed toward engineering personnel involved in the design, specification, installation, and maintenance of high-power equipment utilizing vacuum tubes. Basic principles are discussed, with emphasis on how the underlying technology dictates the applications to which each device is dedicated. Supporting mathematics is included where appropriate to explain the material being discussed. Extensive use of technical illustrations and schematic diagrams aids the reader in understanding the fundamental principles of the subject.

Today's modern power tubes are unlike the power tubes in use decades ago. And with the trend in industry toward operation at higher power levels and higher frequencies, the vacuum tube is certain to remain on the scene for a long time to come.

Jerry C. Whitaker

Acknowledgment

I would like to express my appreciation to Varian Associates, whose support was invaluable in the preparation of this book.

Author

Jerry C. Whitaker is the vice president of Standards Development at the Advanced Television Systems Committee (ATSC). He supports the work of the various ATSC technology and planning committees and assists in the development of ATSC standards, recommended practices, and related documents. ATSC is an international, nonprofit organization developing voluntary standards for digital television.

Whitaker is a fellow of the Society of Broadcast Engineers and an SBE-certified professional broadcast engineer. He is also a fellow of the Society of Motion Picture and Television Engineers. He has been involved in various aspects of the electronics industry for over 30 years. Current CRC book titles include the following:

- *The Electronics Handbook*, second edition
- *Electronic System Maintenance Handbook*, second edition
- *AC Power Systems Handbook*, third edition
- *The RF Transmission Systems Handbook*
- *Microelectronics*, second edition

Whitaker has lectured extensively on the topic of electronic systems design, installation, and maintenance. He is the former editorial director and associate publisher of *Broadcast Engineering* and *Video Systems* magazines and a former radio station chief engineer and television news producer. He has twice received a Jesse H. Neal Award Certificate of Merit from the Association of Business Publishers for editorial excellence. He has also been recognized as Educator of the Year by the Society of Broadcast Engineers.

1

Power Vacuum Tube Applications

1.1 Introduction

The continuing demand for energy control devices capable of higher operating power, higher maximum frequency, greater efficiency, and extended reliability has pushed power vacuum tube manufacturers to break established performance barriers. Advancements in tube design and construction have given engineers new radio frequency (RF) generating systems that allow industry to grow and prosper. Power grid vacuum tubes have been the mainstay of high-power transmitters and other RF generation systems since the beginning of radio. Today, the need for new gridded and microwave power tubes is being met with new processes and materials.

Although low-power vacuum tubes have been largely replaced by solid-state devices, vacuum tubes continue to perform valuable service at high-power levels and, particularly, at high frequencies. The high-power capability of a vacuum device results from the ability of electron/vacuum systems to support high-power densities. Values run typically at several kilowatts per square centimeter, but may exceed $10\text{MW}/\text{cm}^2$. No known dielectric material can equal these values. For the foreseeable future, if high power is required, electron/vacuum devices will remain the best solution.

1.2 Vacuum Tube Development

Receiving tubes have more or less disappeared from the scene (high-end audio applications notwithstanding) because of the development of transistors and integrated circuits. Power grid and microwave tubes, however, continue to push the limits of technology. Power tubes are an important part of RF technology today.

From 1887—when Heinrich Hertz first sent and received radio waves—to the present, an amazing amount of progress has been made by engineers and scientists. The public takes for granted today what was considered science

fiction just a decade or two ago. The route from the primitive spark-gap transmitters to the present state of the art has been charted by the pioneering efforts of many. It is appropriate to review some of the milestones in electron tube development. Much of the fundamental work on power vacuum devices can be traced to early radio broadcasting, which—along with telephone technology—has brought the nations of the world closer together than the early pioneers of the art could have imagined. More than 100 years have passed since Charles D. (Doc) Herrold founded a *voice station* (as it then was known) at San Jose, CA. Developments since have been the result of many inspired breakthroughs and years of plain hard work.

1.2.1 Pioneer Developers

In 1895, 21-year-old Guglielmo Marconi and his brother Alfonso first transmitted radio signals across the hills behind their home in Bologna, Italy. Born in 1874 to an Italian merchant and a Scotch-Irish mother, young Marconi had learned of Hertzian waves from August Righi, a professor at the University of Bologna. Convinced that such waves could be used for wireless communication, Marconi conducted preliminary experiments using a spark-gap source and a coherer detector. Unable to interest the Italian government in his invention, Marconi took his crude transmitter and receiver to England, where he demonstrated his wireless system to officials of the British Post Office. Marconi received a patent for the device in July 1897. With the financial support of his mother's relatives, Marconi organized the Wireless Telegraph and Signal Company that same year to develop the system commercially. Regular transatlantic communications commenced in 1903 when a Marconi station at Cape Cod, MA, sent a short message from President Theodore Roosevelt to King Edward VII in England.

The invention of the vacuum tube diode by J. Ambrose Fleming in 1904 and the triode vacuum tube amplifier by Lee De Forest in 1906 launched the electronics industry as we know it. The De Forest invention was pivotal. It marked the transition of the vacuum tube from a passive to an active device. The new "control" electrode took the form of a perforated metal plate of the same size and shape as the existing anode, positioned between the filament and the anode. Encouraged by the early test results, De Forest worked to perfect his invention, trying various mechanical arrangements for the new grid. Additional refinements and innovative circuit applications were subsequently developed by Edwin Armstrong.

With the addition of the control grid, De Forest and Armstrong had set in motion a chain of events that led the vacuum tube to become the key element in the emerging discipline of electronics. Early experimenters and radio stations took this new technology and began developing their own tubes using in-house capabilities, including glassblowing. As the young electronics industry began to grow, vacuum tubes were produced in great quantity

and standardized (to a point), making it possible to share new developments and applications. A major impetus for standardization was the U.S. military, which required vacuum tubes in large numbers during World War I. Pushed by the navy, a standardized design, including base pins and operating parameters, was forged. The economic benefits to both the tube producers and tube consumers were quickly realized.

Most radio stations from 1910 through 1920 built their own gear. For example, at the University of Wisconsin, Madison, WI, special transmitting tubes were built by hand as needed to keep radio station 9XM, which later became WHA, on the air. The tubes were designed, constructed, and tested by Professor E. M. Terry and a group of his students in the university laboratories. Some of the tubes also were used in wireless telephonic experiments carried on with the Great Lakes Naval Training Station during 1918, when a wartime ban was imposed on wireless broadcasts.

It took many hours to make each tube. The air was extracted by means of a mercury vapor vacuum pump while the filaments were lighted and the plate voltage was on. As the vacuum increased, the plate current was raised until the plate became red hot. This out-gassing process was primitive, but it worked. The students frequently worked through the night to get a tube ready for the next day's broadcast. When completed, the device might last only a few hours before burning out.

Plate dissipation on Professor Terry's early tubes, designated #1, #2, and so on, was about 25W. Tube #5 had a power output of about 50W. Tubes #6-#8 were capable of approximately 75W. Tube #8 was one of the earliest handmade commercial products.

The addition of a "screen" grid marked the next major advancement for the vacuum tube. First patented in 1916 by Dr. Walter Schottky of the Siemens & Halske Company (Germany), the device garnered only limited interest until after World War I when the Dutch firm, Philips, produced a commercial product. The "double-grid" Philips type Q was introduced in May 1923. Later variations on the initial design were made by Philips and other manufacturers.

Early in the use of the tetrode it was determined that the tube was unsuited for use as an audio frequency power amplifier. Under certain operating conditions encountered in this class of service, the tetrode exhibited a negative-resistance characteristic caused by secondary emission from the anode being attracted to the positively charged screen. Although this peculiarity did not affect the performance of the tetrode as an RF amplifier, it did prevent use of the device for AF power amplification. The pentode tube, utilizing a third ("suppressor") grid was designed to overcome this problem. The suppressor grid provided a means to prevent the secondary emissions from reaching the screen grid. This allowed the full capabilities of the tube to be realized.

Philips researchers Drs. G. Holst and B. Tellegen are credited for the invention, receiving a patent for the new device in 1926.

1.2.1.1 Radio Central

The first major project that the young Radio Corporation of America (RCA) tackled was the construction of a huge radio transmitting station at Rocky Point, NY. The facility, completed in 1921, was hailed by President Harding as a milestone in wireless progress. The president, in fact, put the station into operation by throwing a switch that had been rigged up at the White House. Wireless stations around the globe had been alerted to tune in for a congratulatory statement by the president.

For a decade, this station—known as Radio Central—was the primary means of direct communications with Europe. It was also the “hopping off” point for messages transmitted by RCA to Central and South America.

The Rocky Point site was famous not only for its role in communications, but also for the pioneers of the radio age who regularly visited there. The guest book lists such pioneers as Guglielmo Marconi, Lee De Forest, Charles Steinmetz, and Nikola Tesla. Radio Central was a milestone in transatlantic communications.

Originally, two antenna structures stood at the Rocky Point site, each with six 410 ft towers. The towers stretched over a 3 mile area on the eastern end of Long Island.

The facility long outlived its usefulness. RCA demolished a group of six towers in the 1950s; five more were destroyed in early 1960. The last tower of the once mighty Radio Central was taken down on December 13, 1977.

1.2.1.2 WLW: The Nation's Station

Radio station WLW has a history as colorful and varied as any in the United States. It is unique in that it was the only station ever granted authority to broadcast with 500 kW. This accomplishment pushed further the limits of vacuum tube technology.

The station actually began with 20 W of power as a hobby of Powell Crosley, Jr. The first license for WLW was granted by the Department of Commerce in 1922. Crosley was authorized to broadcast on a wavelength of 360 m with a power of 50 W, three evenings a week. Growth of the station was continuous. WLW operated at various frequencies and power levels until, in 1927, it was assigned to 700 kHz at 50 kW and remained there. Operation at 50 kW began on October 4, 1928. The transmitter was located in Mason, OH. The station could be heard as far away as Jacksonville, FL, and Washington, DC.

The superpower era of WLW began in 1934. The contract for construction of the enormous transmitter was awarded to RCA in February 1933. Tests on the unit began on January 15, 1934. The cost of the transmitter and associated equipment was approximately \$400,000—a staggering sum in the middle of the Great Depression.

At 9:02 p.m. on May 2, 1934, programming was commenced with 500 kW of power. The superpower operation was designed to be experimental, but

Crosley managed to renew the license every 6 months until 1939. The call sign W8XO occasionally was used during test periods, but the regular call sign of WLW was used for programming.

“Immense” is the only way to describe the WLW facility. The antenna (including the flagpole at the top) reached a height of 831 ft. The antenna rested on a single ceramic insulator that supported the combined force of 135 ton of steel and 400 ton exerted by the guys. The tower was guyed with eight 1-7/8 in. cables anchored 375 ft from the base of the antenna.

The main antenna was augmented by a directional tower designed to protect CFRB, Toronto, when the station was using 500 kW at night. The directional system was unique in that it was the first designed to achieve both horizontal directivity and vertical-angle suppression.

A spray pond in front of the building provided cooling for the system, moving 512 gal of water per minute. Through a heat exchanger, the water then cooled 200 gal of distilled water in a closed system that cooled the transmitting tubes.

The transmitter consumed an entire building. Modulation transformers, weighing 37,000 lb each, were installed in the basement. Three plate transformers, a rectifier filter reactor, and a modulation reactor were installed outside the building. The exciter for the transmitter produced 50 kW of RF power. A motor-generator was used to provide 125 V dc for control circuits.

The station had its own power substation. While operating at 500 kW, the transmitter consumed 15,450,000 kWh per year. The facility was equipped with a complete machine shop because station personnel had to build much of the ancillary hardware needed. Equipment included gas, arc and spot welders, a metal lathe, milling machine, engraving machine, sander, drill press, metal brake, and a table saw. A wide variety of electric components were also on hand.

WLW operated at 500 kW until March 1, 1939, when the Federal Communications Commission (FCC) ordered the station to reduce power to 50 kW. The station returned to superpower operation a few times during World War II for government research. The days when WLW could boast of being “the nation’s station,” however, were in the past.

1.2.1.3 UHF: A New Technical Challenge

The early planners of the U.S. television system thought that 13 channels would more than suffice. The original channel 1 was from 44 to 50 MHz, but because of possible interference with other services, it was dropped before any active use. There remained 12 channels for broadcasting. Bowing to pressure from various groups, the FCC revised its allocation table in 1952 to permit ultrahigh frequency (UHF)-TV broadcasting for the first time. The new band was not, however, a bed of roses. Many people went bankrupt, building UHF stations only to find few receivers were available to the public. UHF converters soon became popular. The first converters

were so-called matchbox types that were good for one channel only. More expensive models mounted on top of the TV receiver and were tunable. Finally, the commission issued an edict that all TV set manufacturers had to include UHF tuning in their receivers. This move opened the doors for significant market penetration for UHF broadcasters.

The klystron has been the primary means of generating high-power UHF-TV signals since the introduction of UHF broadcasting.

1.2.1.4 Birth of the Klystron

Quietly developed in 1937, the klystron truly revolutionized the modern world. Indeed, the klystron may have helped save the world as we know it. Many decades after it was first operated in a Stanford University laboratory by Russell Varian and his brother Sigurd, the klystron and its offspring remain irreplaceable, even in the age of solid-state microelectronics.

The Varian brothers were unusually bright and extremely active. Mechanically minded, they produced one invention after another. Generally, Sigurd would think up an idea, Russell would devise a method for making it work, then Sigurd would build the device.

Through the influence of William Hansen, a former roommate of Russell and a physics professor at Stanford University, the Varians managed to get nonpaying jobs as research associates in the Stanford physics lab. They had the right to consult with members of the faculty and were given the use of a small room in the physics building.

Hansen's role, apparently, was to shoot down ideas as fast as the Varians could dream them up. As the story goes, the Varians came up with 36 inventions of varying impracticality. Then they came up with idea number 37. This time Hansen's eyes widened. On June 5, 1937, Russell proposed the concept that eventually became the klystron tube. The device was supposed to amplify microwave signals. With \$100 for supplies granted by Stanford, Sigurd built it.

The device was simple: A filament heated by an electric current in turn heated a cathode. A special coating on the cathode gave off electrons when it reached a sufficiently high temperature. Negatively charged electrons attracted by a positively charged anode passed through the first cavity of the klystron tube. Microwaves in the cavity interacted with the electrons and passed through a narrow passage called a *drift tube*. In the drift tube, the electrons tended to bunch up; some speeded up, some slowed down. At the place in the drift tube where the bunching was most pronounced, the electrons entered a second cavity, where the stronger microwaves were excited and amplified in the process.

The first klystron device was lit up on the evening of August 19, 1937. Performance was marginal, but confirmed the theory of the device. An improved klystron was completed and tested on August 30. The name for

the klystron came from the Greek verb *klyzo*, which refers to the breaking of a wave—a process much like the overtaking of slow electrons by fast ones.

The Varians published the results of their discovery in the *Journal of Applied Physics*. For reasons that have never been clear, their announcement immediately impressed British scientists working in the same field, but was almost entirely ignored by German scientists. The development of the klystron allowed British and American researchers to build smaller, more reliable radar systems. Klystron development paralleled work being done in England on the magnetron.

The successful deployment of microwave radar was accomplished by the invention of the cavity magnetron at Manchester University in the late 1930s. It was one of Britain's "Top Secrets" handed over to the Americans early in the war. The cavity magnetron was delivered to the Radiation Laboratory at MIT, where it was incorporated into later wartime radar systems. During the Battle of Britain in May 1940, British defenses depended upon longer wavelength radar (approximately 5 m), which worked but with marginal resolution. The magnetron provided high-power microwave energy at 10 cm wavelengths, which improved detection resolution enormously.

Armed with the magnetron and klystron, British and American scientists perfected radar, a key element in winning the Battle of Britain. So valuable were the secret devices that the British decided not to put radar in planes that flew over occupied Europe lest one of them crash and the details of the components be discovered.

After the war, the Varians—convinced of the potential for commercial value of the klystron and other devices they had conceived—established their own company. For Stanford University, the klystron represents one of its best investments: \$100 in seed money and use of a small laboratory room were turned into \$2.56 million in licensing fees before the patents expired in the 1970s, three major campus buildings and hundreds of thousands of dollars in research funding.

At about the same time the Varian brothers were working to perfect the klystron, Andrew Haeff of RCA Laboratories was developing what would come to be known as the inductive output tube (IOT). The crucial feature in both devices was that the electron beam, having given up a large part of its energy in a small, low-capacitance gap, tended to spread because of space-charge effects (the mutual repulsion of electrons). The characteristics of electron flow make it possible to use multiple collectors with differing potentials in tubes with large energy spread in the spent electron beam, such as the klystron and IOT. This multistage depressed-collector (MSDC) technique results in a considerable increase in operating efficiency. One of the earliest patents on the MSDC was issued in 1943 to Charles Litton, who later formed a company that would become Litton Industries. Figure 1.1 shows the basic MSDC structure, developed from the original Litton patent.

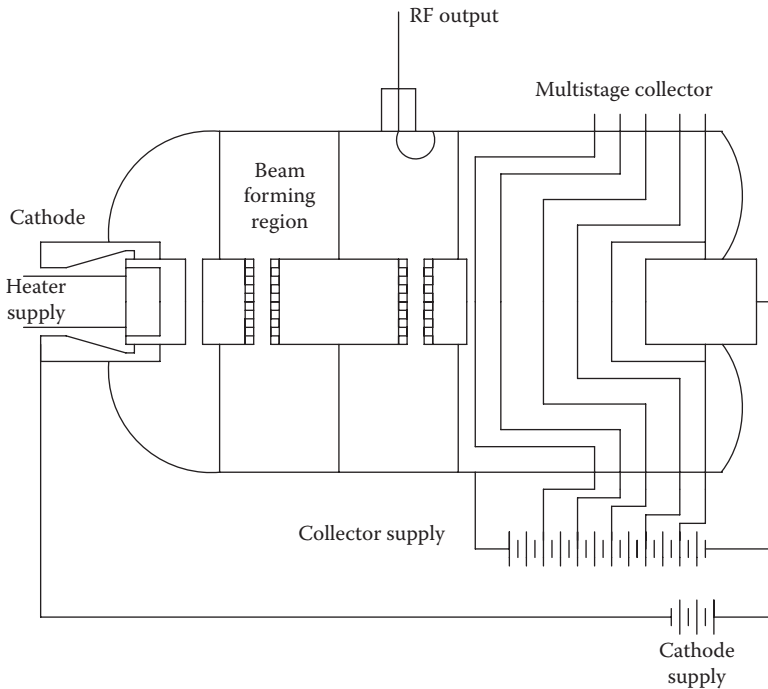


FIGURE 1.1
Simplified drawing of a multistage depressed-collector klystron using six collector elements.

1.2.1.5 Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR), a technique that has revolutionized chemistry throughout the world, had its origins in experiments carried out on opposite sides of the North American continent in the late 1940s. As is often the case with basic discoveries in science, the original intent of the work was far removed from the ultimate practical application. NMR was first demonstrated during the winter of 1945 by Professors Felix Bloch and William Hansen and their associates at Stanford University. At the same time another group was working independently at Harvard University, directed by Dr. E. M. Purcell. The Nobel Prize for physics was awarded to Bloch and Purcell in 1952 in recognition of their pioneering work in NMR, which allows chemists to make structural determinations of substances, and thereby see how a molecule is put together.

1.2.2 Transistor Is Born

In December 1947, Dr. William Shockley of Bell Laboratories changed the course of history by demonstrating to his colleagues a newly discovered device that exhibited what he called the *transistor effect*. From this

demonstration, and a later one at the Bell Labs in New York City on June 30, 1948, sprang one of the most important inventions of the twentieth century—the working transistor. (The phrase “transistor” came about from a contraction of “transfer resistor.”) For their development efforts, Bell Telephone scientists John Bardeen, Walter H. Brattain, and William Shockley received the Nobel Prize for physics in 1956.

The first experimental transistor made its debut as a 3/16 in. diameter, 5/8 in. long metal cylinder. The transistor was an economical and durable alternative to receiving vacuum tubes, which were big, fragile, and had a relatively short operating life. They consumed lots of power and, as a result, got very hot. Transistors offered a way to make a product that was compact, efficient, and reliable.

What made transistors distinctive was that they were fabricated from a single solid material that either insulated or conducted, depending on the purity of the base material. A solitary transistor could be scribed on nothing more than a chip of germanium (subsequently supplanted by silicon). Eager to exploit the riches of solid-state technology, engineers drew up complicated circuits, cramming boards of discrete devices into tight little islands.

1.2.3 Satellite Technology

The phrase “live via satellite” is commonplace today. Communications satellites have meant not only live coverage of world events, but also more service to more people at lower cost. The potential for use of satellites for communications was first demonstrated in 1960 with the launching by the United States of Echo 1 and Echo 2. These passive reflector satellites bounced radio signals across the Atlantic. This type of satellite, however, left a lot to be desired for communications purposes. Echo was superseded by Telstar 1, an active repeater satellite, launched 2 years later. Telstar demonstrated that color video signals could be reliably broadcast across the oceans and, in doing so, captured the interest and imagination of the public.

The first live transatlantic telecast was relayed by Telstar on July 10, 1962. The picture was of the American flag fluttering in front of the sending station at Andover, ME. More panoramic telecasts, showing life in widely distant places, were exchanged between the United States and Europe 13 days later.

With the potential of international communications becoming increasingly apparent, Congress passed the Communications Satellite Act in 1962. The legislation created, among other things, Comsat, which was to establish, in cooperation with organizations in other countries, a global commercial communications satellite system as quickly as possible.

U.S. initiative under the Communications Satellite Act, combined with growing international interest in the new technology, led to the formation of Intelsat in 1964. Acting as a technical manager of Intelsat during its initial growth period, Comsat developed Intelsat's first geosynchronous commercial communications satellite, Early Bird. The project brought to reality

the concept envisaged some 20 years earlier by Arthur C. Clarke, the noted British science fiction writer.

Early Bird (also known as Intelsat 1) was launched from Cape Kennedy on April 2, 1965, and placed into synchronous orbit 22,300 miles above the coast of Brazil. The launch marked the first step toward a worldwide network of satellites linking the peoples of many nations. Early Bird, the only mode of live transatlantic television, provided in July 1965 the first live telecast (via Intelsat satellite) to the United States, an American vs. Soviet track meet. All of these breakthroughs relied on new, compact, efficient, and lightweight microwave power tubes.

Although a dramatic improvement over the transatlantic telecommunications facilities at the time, Early Bird was nonetheless limited in capacity and capability. For example, in order for the only TV channel to be operative, all 240 voice channels had to be shut down. Furthermore, the cost of Early Bird time was quite high.

Following Early Bird's introduction to service in the Atlantic Ocean region, the challenge remained to develop a global network. The next step toward that goal was taken on July 11, 1967, with the successful launch of Intelsat 2, which established satellite communications between the U.S. mainland and Hawaii. Two years later, Intelsat 3 was launched for Indian Ocean region service, thereby completing the provisions of global coverage.

Fortunately, global coverage capacity was in place just in time for the TV audience—estimated to be the largest in world history—to see man set foot on the moon. The satellite system that had been a vision by Arthur C. Clarke two decades earlier, and a formidable legislative mandate more than a decade earlier, had emerged as a reality.

The demand for fixed satellite services (FSS) and capacity continues to grow. Rapid deployment of private networks using very small aperture terminals (VSATs)—which interconnect widely dispersed corporate offices, manufacturing, supply, and distribution centers—has contributed to this growth. The use of VSATs, which employ a low-cost terminal with a relatively small aperture antenna (1–2 m is typical), has steadily grown around the world. Internet and video distribution services are another major and rapidly increasing components of satellite services.

1.2.4 Standardization

Throughout the history of product development, design standardization has been critically important. To most engineers, the term “standards” connotes a means of promoting an atmosphere of interchangeability of basic hardware. The standardization process has been an important element in the advancement of power vacuum tube technology.

In 1836, the U.S. Congress authorized establishment of the Office of Weights and Measures (OWM) for the primary purpose of ensuring uniformity in customs house dealings. The Treasury Department was charged

with its operation. As advancements in science and technology fueled the industrial revolution, it was apparent that standardization of hardware and test methods was necessary to promote commercial development and to compete successfully in the world market. The industrial revolution in the 1830s introduced the need for interchangeable parts and hardware. Wide use of steam railways and the cotton gin, for example, was possible only with mechanical standardization.

By the late 1800s, professional organizations of mechanical, electrical, and chemical engineers were founded with this aim in mind. The American Institute of Electrical Engineers developed standards based on the practices of the major electrical manufacturers between 1890 and 1910. Because such activities were not within the purview of the OWM, there was no government involvement during this period. It took the pressures of war production in 1918 to cause the formation of the American Engineering Standards Committee (AESC) to coordinate the activities of various industry and engineering societies. This group became the American Standards Association (ASA) in 1928.

Parallel development occurred worldwide. The International Bureau of Weights and Measures was founded in 1875, the International Electrotechnical Commission (IEC) in 1904, and the International Federation of Standardizing Bodies in 1926. Following World War II, this group was reorganized as the International Standards Organization (ISO). The International Telecommunications Union (ITU) was founded in 1865 for the purpose of coordinating and interfacing telegraphic communications worldwide.

Many of the early standards relating to vacuum tubes in the United States were developed by equipment manufacturers, first under the banner of the Radio Manufacturers Association (RMA), then the Radio, Electronic, and Television Manufacturers Association (RETMA), and still later by the Electronic Industries Association (EIA). The Institute of Radio Engineers (the forerunner of the IEEE) was responsible for measurement standards and techniques.

Standards usually are changed only through natural obsolescence. Changes in basic quantities, such as units of length and volume, are extremely difficult for the general public to accept. In 1900, nearly all members of the scientific, commercial, and engineering communities supported the change to the metric system. But the general public finds the idea of changing ingrained yardsticks for weight and measure as unpalatable as learning to speak a new language in a native country. The effort to convert to the metric system is, in fact, still under way, with various degrees of success.

1.2.5 Transmission Systems

Radio communication is the grandfather of vacuum tube development. Amplitude modulation (AM) was the first modulation system that permitted voice communications to take place. This simple scheme was predominant

throughout the 1920s and 1930s. Frequency modulation (FM) came into regular broadcast service during the 1940s. TV broadcasting, which used AM for the visual portion of the signal and FM for the aural portion of the signal, became available to the public during the mid-1940s. These two basic approaches to modulating a carrier have served the communications industry well for many decades. Although the basic schemes still are used today, numerous enhancements have been made.

Technology also has changed the rules by which the communications systems developed. AM radio, as a technical system, offered limited audio fidelity but provided design engineers with a system that allowed uncomplicated transmitters and simple, inexpensive receivers. FM radio, on the other hand, offered excellent audio fidelity but required a complex and unstable transmitter (in the early days) and complex, expensive receivers. It is no wonder that AM radio flourished while FM remained relatively stagnant for at least 20 years after being introduced to the public. It was not until transistors and later integrated circuits became commonly available that FM receivers gained wide consumer acceptance.

TV broadcasting evolved slowly during the late 1940s and early 1950s. Color transmissions were authorized as early as 1952. Color receivers were not purchased in large numbers by consumers, however, until the mid-1960s. Early color sets were notorious for poor reliability and unstable performance. They also were expensive. As with FM, all that changed with the introduction of transistors and, later, integrated circuits.

1.3 Vacuum Tube Applications

The range of uses for power grid and microwave tubes is wide and varied. Some of the more common applications include the following:

- AM radio broadcasting, with power levels up to 50 kW
- Shortwave radio, with power levels of 50 kW to several megawatts
- FM radio broadcasting, with power levels up to 100 kW
- Television, with operating power levels up to several megawatts
- Radar, with widely varying applications ranging from airborne systems to over-the-horizon networks
- Satellite communications, incorporating the ground segment and space segment
- Industrial heating and other commercial processes

Figure 1.2 charts the major device types in relation to power and frequency.

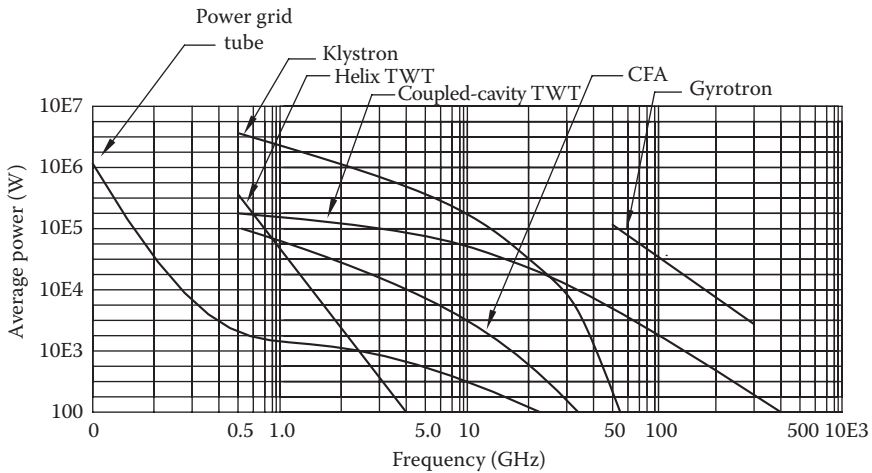


FIGURE 1.2

Generalized vacuum tube operating parameters based on power and frequency. (From Whitaker, J.C., *Microwave power tubes*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, 1996. With permission.)

1.3.1 Market Overview

There is a huge installed base of power vacuum tubes. Users cover a wide range of disciplines and facilities. The power tube market can be divided into the following general segments:

- *Medical.* Applications include diagnostics, such as x-ray and CAT scanners, and treatment, such as radio therapy. Devices used include x-ray tubes and klystrons.
- *Scientific.* Applications include various types of research, including particle accelerators for nuclear research. A wide variety of devices are used in this area. Specialized units can be found in “big science” projects, such as particle accelerators. The worldwide scientific market for microwave power tubes continues to grow as researchers reach for higher power levels and higher frequencies of operation.
- *Electronic warfare.* Applications include ship- and air-based radar systems, eavesdropping systems, and radio-signal jamming equipment. Electronic warfare applications are a driving force in power microwave tube development.
- *Communications.* Applications include space- and ground-based satellite systems, microwave relay, and specialized point-to-point communications systems. Power tubes play an important role in all types of high-power communications systems.
- *Radar.* Applications include ground-, ship-, and air-based radar for both commercial and military applications.

- *Industrial heating.* Applications center on material heating and chemical processing, and environmental uses such as smokestack scrubbing.
- *Broadcasting.* Applications are TV and radio broadcasting, including shortwave.

Despite the inroads made by solid-state technology, power vacuum tube applications continue to span a broad range of disciplines around the world.

1.4 Electromagnetic Radiation Spectrum

The usable spectrum of electromagnetic radiation frequencies extends over a range from below 100Hz for power distribution to 10^{20} Hz for the shortest x-rays (see Figure 1.3). The lower frequencies are used primarily for terrestrial broadcasting and communications. The higher frequencies include visible and near-visible infrared and ultraviolet light, and x-rays. The major frequency band designations are listed in Table 1.1.

Generally speaking, the fields of RF and microwave engineering together encompass the design and implementation of electronic systems utilizing frequencies in the electromagnetic spectrum from approximately 300kHz to over 100GHz [1]. The term “RF engineering” is typically used to refer to circuits and systems utilizing frequencies in a range from approximately 300 kHz at the low end to between 300MHz and 1GHz at the high end. The term “microwave engineering” meanwhile is used rather loosely to refer to design and implementation of electronic systems with operating frequencies in a range from 300MHz to 1GHz on the low end to upwards of 100GHz at the high end. The boundary between RF and microwave design is both somewhat indistinct and one that is continually shifting as device technologies and design methodologies advance.

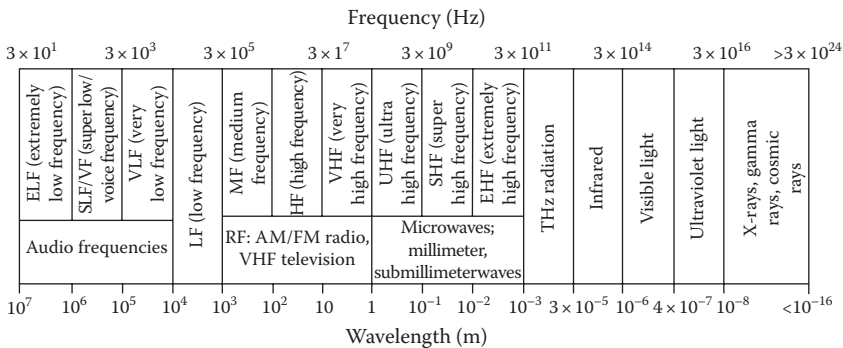


FIGURE 1.3 Electromagnetic frequency spectrum and associated wavelengths. (From Golio, M., *RF and Microwave Circuits, Measurements, and Modeling*, CRC Press, Boca Raton, FL, 2008. With permission.)

TABLE 1.1

Frequency Band Designations

Description	Designation	Frequency
Extremely low frequency	ELF (1) Band	3 Hz up to 30 Hz
Super low frequency	SLF (2) Band	30 Hz up to 300 Hz
Ultralow frequency	ULF (3) Band	300 Hz up to 3 kHz
Very low frequency	VLf (4) Band	3 kHz up to 30 kHz
Low frequency	LF (5) Band	30 kHz up to 300 kHz
Medium frequency	MF (6) Band	300 kHz up to 3 MHz
High frequency	HF (7) Band	3 MHz up to 30 MHz
Very high frequency	VHF (8) Band	30 MHz up to 300 MHz
Ultrahigh frequency	UHF (9) Band	300 MHz up to 3 GHz
Superhigh frequency	SHF (10) Band	3 GHz up to 30 GHz
Extremely high frequency	EHF (11) Band	30 GHz up to 300 GHz
—	—(12) Band	300 GHz up to 3 THz

1.4.1 Low Frequency: 30–300 kHz

The low frequency (LF) band is used for around-the-clock communications services over long distances and where adequate power is available to overcome high levels of atmospheric noise. Applications include

- Radionavigation
- Fixed/maritime communications and navigation
- Aeronautical radionavigation
- Low-frequency broadcasting
- Underwater submarine communications (10–30 kHz)

1.4.2 Medium Frequency: 300 kHz–3 MHz

The low-frequency portion of this band is used for around-the-clock communication services over moderately long distances. The upper portion of the medium frequency (MF) band is used principally for moderate-distance voice communications. Applications in this band include

- AM radio broadcasting
- Radionavigation
- Fixed/maritime communications
- Aeronautical radionavigation
- Fixed and mobile commercial communications
- Amateur radio
- Standard time and frequency services

1.4.3 High Frequency: 3–30 MHz

This band provides reliable medium-range coverage during daylight and, when the transmission path is in total darkness, worldwide long-distance service. The reliability and signal quality of long-distance service depends to a large degree upon ionospheric conditions and related long-term variations in sunspot activity affecting sky-wave propagation. Applications include

- Shortwave broadcasting
- Fixed and mobile service
- Telemetry
- Amateur radio
- Fixed/maritime mobile
- Standard time and frequency services
- Radio astronomy
- Aeronautical fixed and mobile

1.4.4 Very High Frequency: 30–300 MHz

The very high frequency (VHF) band is characterized by reliable transmission over medium distances. At the higher portion of the VHF band, communication is limited by the horizon. Applications include

- FM radio broadcasting
- Television broadcasting
- Commercial fixed and mobile radio
- Aeronautical radionavigation
- Space research
- Fixed/maritime mobile
- Amateur radio
- Radiolocation

1.4.5 Ultrahigh Frequency: 300 MHz–3 GHz

Transmissions in this band are typically line of sight. Short wavelengths at the upper end of the band permit the use of highly directional parabolic or multielement antennas. Applications include

- Television broadcasting
- Fixed and mobile communications
- Telemetry

- Meteorological aids
- Space operations
- Radio astronomy
- Radionavigation
- Satellite communications
- Point-to-point microwave relay

1.4.6 Superhigh Frequency: 3–30 GHz

Communication in this band is strictly line of sight. Very short wavelengths permit the use of parabolic transmit and receive antennas of exceptional gain. Applications include

- Satellite communications
- Point-to-point wideband relay
- Radar
- Specialized wideband communications
- Developmental research
- Military support systems
- Radiolocation
- Radionavigation
- Space research

References

1. Whitaker, J. C., Microwave power tubes, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 413–416, 1996.
2. Golio, M., *RF and Microwave Circuits, Measurements, and Modeling*, CRC Press, Boca Raton, FL, 2008.

Bibliography

- Battison, J., Making history, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, June 1986.
- Benson, K. B. and J. C. Whitaker, *Television and Audio Handbook for Technicians and Engineers*, McGraw-Hill, New York, 1989.

- Benson, K. B. and J. C. Whitaker (eds.), *Television Engineering Handbook*, revised edn., McGraw-Hill, New York, 1991.
- Brittain, J. E., Scanning the past: Guglielmo Marconi, *Proceedings of the IEEE*, 80(8), 335, August 1992.
- Burke, W., WLW: The nation's station, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, November 1967.
- Clerc, G. and William R. H., The case for tubes, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, pp. 67–82, May 1991.
- Dorschug, H., The good old days of radio, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1971.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineer's Handbook*, 2nd edn., McGraw-Hill, New York, 1982.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.
- Goldstein, I., Communications satellites: A revolution in international broadcasting, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1979.
- Golio, M., *RF and Microwave Applications and Systems*, CRC Press, Boca Raton, FL, 2008.
- Jordan, E. C. (ed.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 7th edn., Howard W. Sams Company, Indianapolis, IN, 1985.
- McCroskey, D., Setting standards for the future, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1989.
- Nelson, C., RCA demolishes last antenna tower at historic radio central, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, February 1978.
- Paulson, R., The house that radio built, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, April 1989.
- Peterson, B. B., Electronic navigation systems, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1710–1733, 1996.
- Pond, N. H. and C. G. Lob, Fifty years ago today or on choosing a microwave tube, *Microwave Journal*, 226–238, September 1988.
- Riggins, G., The real story on WLW's long history, *Radio World*, Falls Church, VA, March 8, 1989.
- Schow, E., A review of television systems and the systems for recording television, *Sound and Video Contractor*, Intertec Publishing, Overland Park, KS, May 1989.
- Schubin, M., From tiny tubes to giant screens, *Video Review*, April 1989.
- Stokes, J. W., *Seventy Years of Radio Tubes and Valves*, Vestal Press, New York, 1982.
- Symons, R. S., Tubes—Still vital after all these years, *IEEE Spectrum*, IEEE, New York, pp. 53–63, April 1998.
- Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.
- Varian Associates, *An Early History*, Varian, Palo Alto, CA.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1991.

2

Modulation Systems and Characteristics

2.1 Introduction

Radio frequency (RF) power amplifiers are used in countless applications at tens of thousands of facilities around the world. The wide variety of applications, however, stem from a few basic concepts of conveying energy and information by means of an RF signal. Furthermore, the devices used to produce RF energy have many similarities, regardless of the final application. Although communications systems represent the most obvious use of high-power RF generators, numerous other common applications exist, including

- Induction heating and process control systems
- Radar (ground, air, and shipboard)
- Satellite communications
- Atomic science research
- Medical research, diagnosis, and treatment

The process of generating high-power RF signals has been refined over the years to an exact science. Advancements in devices and circuit design continue to be made each year, pushing ahead the barriers of efficiency and maximum operating frequency. Although different applications place unique demands on the RF design engineer, the fundamental concepts of RF amplification are applicable to virtually any system.

2.2 Modulation Systems

The primary purpose of most communications and signaling systems is to transfer information from one location to another. The message signals used in communication and control systems usually must be limited in frequency to provide for efficient transfer. This frequency may range from a

few hertz for control systems to a few megahertz for video signals to many megahertz for multiplexed data signals. To facilitate efficient and controlled distribution of these components, an *encoder* generally is required between the source and the transmission channel. The encoder acts to *modulate* the signal, producing at its output the *modulated waveform*. Modulation is a process whereby the characteristics of a wave (the *carrier*) are varied in accordance with a message signal, the modulating waveform. Frequency translation is usually a by-product of this process. Modulation may be continuous, where the modulated wave is always present, or pulsed, where no signal is present between pulses.

There are a number of reasons for producing modulated waves, including

- *Frequency translation.* The modulation process provides a vehicle to perform the necessary frequency translation required for distribution of information. An input signal may be translated to its assigned frequency band for transmission or radiation.
- *Signal processing.* It is often easier to amplify or process a signal in one frequency range as opposed to another.
- *Antenna efficiency.* Generally speaking, for an antenna to be efficient, it must be large compared with the signal wavelength. Frequency translation provided by modulation allows antenna gain and beamwidth to become part of the system design considerations. The use of higher frequencies permits antenna structures of reasonable size and cost.
- *Bandwidth modification.* The modulation process permits the bandwidth of the input signal to be increased or decreased as required by the application. Bandwidth reduction permits more efficient use of the spectrum at the cost of signal fidelity. Increased bandwidth, on the other hand, provides increased immunity to transmission channel disturbances.
- *Signal multiplexing.* In a given transmission system, it may be necessary or desirable to combine several different signals into one baseband waveform for distribution. Modulation provides the vehicle for such *multiplexing*. Various modulation schemes allow separate signals to be combined at the transmission end and separated (*demultiplexed*) at the receiving end. Multiplexing may be accomplished by using, among other systems, *frequency-domain multiplexing* (FDM) or *time-domain multiplexing* (TDM).

Modulation of a signal does not come without the possible introduction of undesirable attributes. Bandwidth restriction or the addition of noise or other disturbances are the two primary problems faced by the transmission system designer.

Spectral efficiency is of paramount importance when considering the design of virtually any commercial wireless communications system, whether that

system is intended for voice, video, control, or data. Spectral efficiency can be increased by using one or more of several methods, including

- Signal polarization, where, for example, uplink and downlink transmissions are on the same frequency but with different polarizations
- Access method, where a common resource (such as frequency) is shared among each of the users of the system
- Modulation method, where the characteristics of a signal can be changed, for example, amplitude, frequency, and phase
- Signal coding techniques, where the emission is given characteristics that allow it to be differentiated from other emissions

2.2.1 Principles of Resonance

All RF generations rely on the principles of resonance for operation. Three basic architectures exist:

- Series resonance circuits
- Parallel resonance circuits
- Cavity resonators

2.2.1.1 Series Resonant Circuits

When a constant voltage of varying frequency is applied to a circuit consisting of an inductance, capacitance, and resistance (all in series), the current that flows depends upon frequency in the manner shown in Figure 2.1. At low frequencies, the capacitive reactance of the circuit is large and the inductive reactance is small, so that most of the voltage drop is across the capacitor, while the current is small and leads the applied voltage by nearly 90° . At high frequencies, the inductive reactance is large and the capacitive reactance is

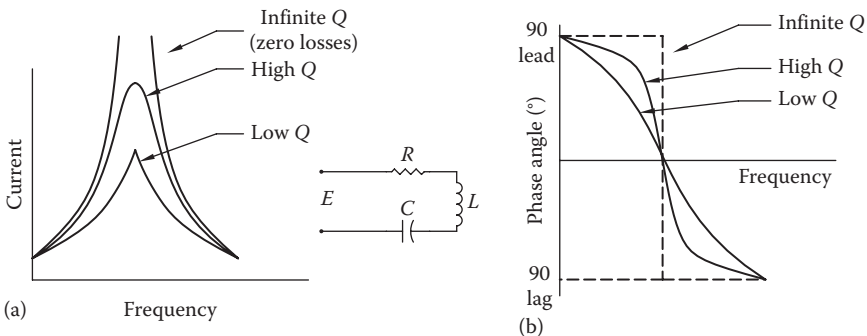


FIGURE 2.1

Characteristics of a series resonant circuit as a function of frequency for a constant applied voltage and different circuit Q s: (a) magnitude and (b) phase angle.

low, resulting in a small current that lags nearly 90° behind the applied voltage; most of the voltage drop is across the inductance. Between these two extremes is the *resonant frequency*, at which the capacitive and inductive reactances are equal and, consequently, neutralize each other, leaving only the resistance of the circuit to oppose the flow of current. The current at this resonant frequency is, accordingly, equal to the applied voltage divided by the circuit resistance, and it is very large if the resistance is low.

The characteristics of a series resonant circuit depend primarily upon the ratio of inductive reactance ωL to circuit resistance R , known as the circuit Q :

$$Q = \frac{\omega L}{R} \quad (2.1)$$

The circuit Q also may be defined by

$$Q = 2\pi \left(\frac{E_s}{E_d} \right) \quad (2.2)$$

where

E_s is the energy stored in the circuit

E_d is the energy dissipated in the circuit during one cycle

Most of the loss in a resonant circuit is the result of coil resistance; the losses in a properly constructed capacitor are usually small in comparison with those of the coil.

The general effect of different circuit resistances (different values of Q) is shown in Figure 2.1. As illustrated, when the frequency differs appreciably from the resonant frequency, the actual current is practically independent of circuit resistance and is nearly the current that would be obtained with no losses. On the other hand, the current at the resonant frequency is determined solely by the resistance. The effect of increasing the resistance of a series circuit is, accordingly, to flatten the resonance curve by reducing the current at resonance. This broadens the top of the curve, giving a more uniform current over a band of frequencies near the resonant point. This broadening is achieved, however, by reducing the selectivity of the tuned circuit.

2.2.1.2 Parallel Resonant Circuits

A parallel circuit consisting of an inductance branch in parallel with a capacitance branch offers an impedance of the character shown in Figure 2.2. At low frequencies, the inductive branch draws a large lagging current while the leading current of the capacitive branch is small, resulting in a large lagging line current and a low lagging circuit impedance. At high frequencies, the inductance has a high reactance compared with the capacitance, resulting in a large leading line current and a corresponding low circuit impedance that is leading in phase. Between these two extremes is a frequency at

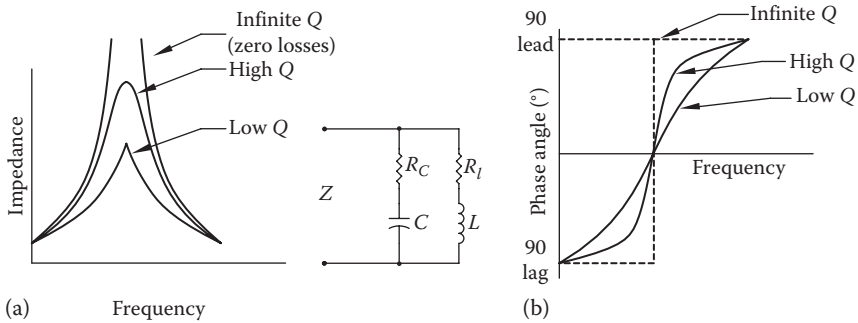


FIGURE 2.2 Characteristics of a parallel resonant circuit as a function of frequency for different circuit Q s: (a) magnitude and (b) phase angle.

which the lagging current taken by the inductive branch and the leading current entering the capacitive branch are equal. Being 180° out of phase, they neutralize, leaving only a small resultant in-phase current flowing in the line; the impedance of the parallel circuit is, therefore, high.

The effect of circuit resistance on the impedance of the parallel circuit is similar to the influence that resistance has on the current flowing in a series resonant circuit, as is evident when Figures 2.1 and 2.2 are compared. Increasing the resistance of a parallel circuit lowers and flattens the peak of the impedance curve without appreciably altering the sides, which are relatively independent of the circuit resistance.

The resonant frequency F_o of a parallel circuit can be taken as the same frequency at which the same circuit is in series resonance:

$$F_o = \frac{1}{2\pi\sqrt{LC}} \tag{2.3}$$

where

L is the inductance in the circuit

C is the capacitance in the circuit

When the circuit Q is large, the frequencies corresponding to the maximum impedance of the circuit and to unity power factor of this impedance coincide, for all practical purposes, with the resonant frequency defined in this way. When the circuit Q is low, however, this rule does not necessarily apply.

2.2.1.3 Cavity Resonators

Any space completely enclosed with conducting walls may contain oscillating electromagnetic fields. Such a space also exhibits certain resonant frequencies when excited by electrical oscillations. Resonators of this type, commonly termed *cavity resonators*, find extensive use as resonant circuits

at very high frequencies and above. For such applications, cavity resonators have a number of advantages, including

- Mechanical and electrical simplicity
- High Q
- Stable parameters over a wide range of operating conditions

If desired, a cavity resonator can be configured to develop high shunt impedance.

The simplest cavity resonators are sections of waveguides short-circuited at each end and $\lambda_g/2$ wavelengths long, where λ_g is the guide wavelength. This arrangement results in a resonance analogous to that of a $1/2$ -wavelength transmission line short-circuited at the receiving end.

Any particular cavity is resonant at a number of frequencies, corresponding to different possible field configurations that exist within the enclosure. The resonance having the longest wavelength (lowest frequency) is termed the *dominant* or *fundamental* resonance. In the case of cavities that are resonant sections of cylindrical or rectangular waveguides, most of the possible resonances correspond to various modes that exist in the corresponding waveguides.

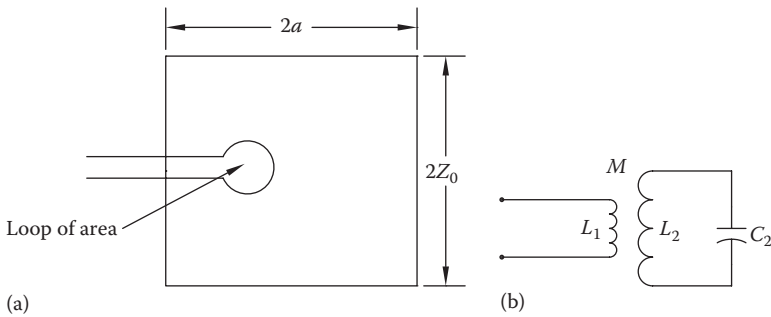
The resonant wavelength is proportional in all cases to the size of the resonator. If the dimensions are doubled, the wavelength corresponding to resonance will likewise be doubled. This fact simplifies the construction of resonators of unusual shapes whose proper dimensions cannot be calculated easily.

The resonant frequency of a cavity can be changed through one or more of the following actions:

- Altering the mechanical dimensions of the cavity
- Coupling reactance into the cavity
- Inserting a copper paddle inside the cavity and moving it to achieve the desired resonant frequency

Small changes in mechanical dimensions can be achieved by flexing walls, but large changes typically require some type of sliding member. Reactance can be coupled into the resonator through a coupling loop, thus affecting the resonant frequency. A copper paddle placed inside the resonator affects the normal distribution of flux and tends to raise the resonant frequency by an amount determined by the orientation of the paddle. This technique is similar to the way in which copper can be used to produce small variations in the inductance of a coil.

Coupling to a cavity resonator can be achieved by means of a coupling loop or a coupling electrode. Magnetic coupling is achieved by means of a

**FIGURE 2.3**

Cylindrical resonator incorporating a coupling loop: (a) orientation of loop with respect to cavity and (b) equivalent coupled circuit.

small coupling loop oriented so as to enclose magnetic flux lines existing in the desired mode of operation. A current passed through such a loop then excites oscillations of this mode. Conversely, oscillations existing in the resonator induce a voltage in such a coupling loop. The combination of a coupling loop and cavity resonator is equivalent to the ordinary coupled circuit shown in Figure 2.3. The magnitude of magnetic coupling can be readily controlled by rotating the loop. The coupling is reduced to zero when the plane of the loop is parallel to the magnetic flux.

2.2.2 Frequency Sources

Every RF amplifier requires a stable frequency reference. At the heart of most systems is a quartz crystal. Quartz acts as a stable high Q mechanical resonator. Crystal resonators are available for operation at frequencies ranging from 1 kHz to 300 MHz and beyond.

The operating characteristics of a crystal are determined by the *cut* of the device from a bulk “mother” crystal. The behavior of the device depends heavily on the size and shape of the crystal and the angle of the cut. To provide for operation at a wide range of frequencies, different cuts, vibrating in one or more selected modes, are used.

The properties of a piezoelectric quartz crystal can be expressed in terms of three sets of axes (see Figure 2.4). The axis joining the points at the ends of the crystal is known as the *optical axis*, and electrical stresses applied in this direction exhibit no piezoelectric properties. The three axes X' , X'' , and X''' passing through the corners of the hexagon that forms the section perpendicular to the optical axes, are known as the *electrical axes*. The three axes Y' , Y'' , and Y''' , which are perpendicular to the faces of the crystal, are the *mechanical axes*.

If a flat section is cut from a quartz crystal in such a way that the flat sides are perpendicular to an electrical axis, as shown in Figure 2.4b (X -cut),

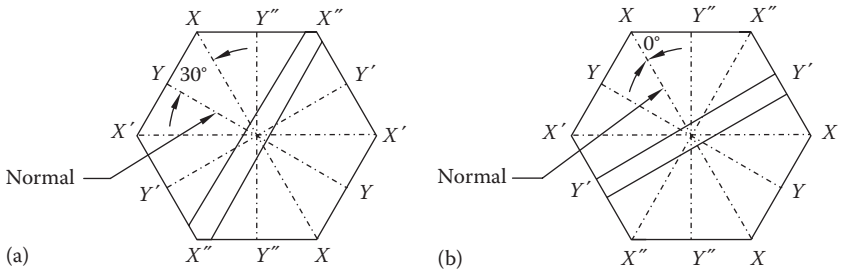


FIGURE 2.4

Cross section of a quartz crystal taken in the plane perpendicular to the optical axis: (a) Y-cut plate and (b) X-cut plate.

the mechanical stresses along the Y-axis of such a section produce electric charges on the flat sides of the crystal. If the direction of these stresses is changed from tension to compression, or vice versa, the polarity of the charges on the crystal surfaces is reversed. Conversely, if electric charges are placed on the flat sides of the crystal by applying a voltage across the faces, a mechanical stress is produced in the direction of the Y-axis. This property by which mechanical and electrical properties are interconnected in a crystal is known as the *piezoelectric effect* and is exhibited by all sections cut from a piezoelectric crystal. Thus, if mechanical forces are applied across the faces of a crystal section having its flat sides perpendicular to a Y-axis, as shown in Figure 2.4a, piezoelectric charges will be produced because forces and potentials developed in such a crystal have components across at least one of the Y- or X-axes, respectively.

An alternating voltage applied across a quartz crystal will cause the crystal to vibrate and, if the frequency of the applied alternating voltage approximates a frequency at which mechanical resonance can exist in the crystal, the amplitude vibrations will be large. Any crystal has a number of such resonant frequencies that depend upon the crystal dimensions, the type of mechanical oscillation involved, and the orientation of the plate cut from the natural crystal.

Crystals are temperature sensitive, as shown in Figure 2.5. The extent to which a device is affected by changes in temperature is determined by its cut and packaging. Crystals also exhibit changes in frequency with time. Such *aging* is caused by one or both of the following:

- Mass transfer to or from the resonator surface
- Stress relief within the device itself

Crystal aging is most pronounced when the device is new. As stress within the internal structure is relieved, the aging process slows.

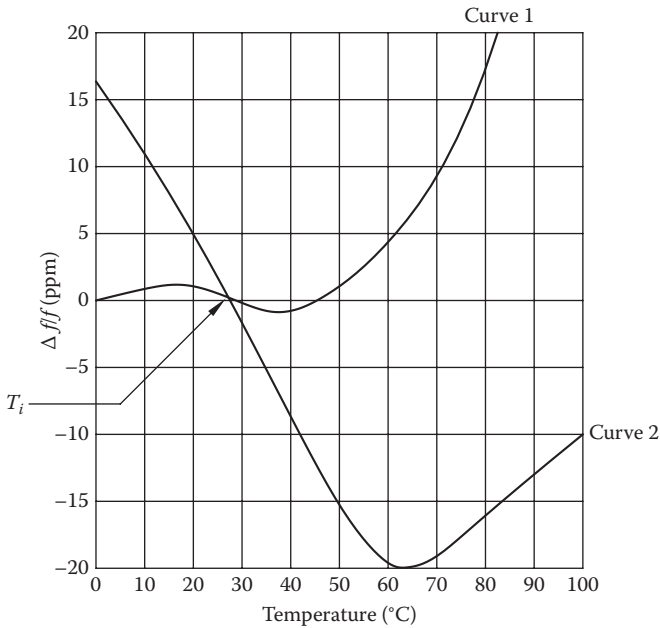


FIGURE 2.5
Effects of temperature on two types of AT-cut crystals.

2.2.2.1 Frequency Stabilization

The stability of a quartz crystal is inadequate for most commercial and industrial applications. Two common methods are used to provide the required long-term frequency stability:

- *Oven-controlled crystal oscillator.* The crystal is installed in a temperature-controlled box. Because the temperature in the box is constant, controlled by a thermostat, the crystal remains on-frequency. The temperature of the enclosure usually is set to the *turnover temperature* of the crystal (the turnover point is illustrated in Figure 2.5).
- *Temperature-compensated crystal oscillator (TCXO).* The frequency vs. temperature changes of the crystal are compensated by varying a load capacitor. A thermistor network typically is used to generate a correction voltage that feeds a voltage-controlled capacitor (*varactor*) to retune the crystal to the desired on-frequency value.

2.2.2.2 Equivalent Circuit of a Quartz Resonator

Because of the piezoelectric effect, the quartz resonator behaves physically as a vibrating object when driven by a periodic electric signal near a resonant

frequency of the cavity [1,2]. This resonant behavior may be characterized by equivalent electrical characteristics, which may then be used to accurately determine the response and performance of the electromechanical system. Dye and Butterworth (1926) [3] developed the equivalent lump electrical network that describes an electrostatic signal driving a mechanically vibrating system (an *electromechanical transducer*). The equivalent electric network that models the piezo resonator was later developed by Van Dyke (1928), Butterworth, and others, and is shown in Figure 2.6. In the equivalent network, the circuit parameters R_1 , L_1 , and C_1 represent parameters associated with the quartz crystal, while C_0 represents the shunt capacitance of the resonator electrodes in parallel with the container that packages the crystal.

The inductance L_1 is associated with the mass of the quartz slab, and the capacitance C_1 is associated with the stiffness of the slab. Finally, the resistance, R_1 is determined by the loss processes associated with the operation of the crystal. Each of the equivalent parameters can be measured using standard network measurement techniques. Typical parameters are shown in Table 2.1.

Application of an electric field across the electrodes of the crystal results in both face and thickness shear waves developing within the crystal resonator. For a Y-cut crystal, face shear modes occur in the X-Z plane, while thickness shear modes develop in the X-Y plane.

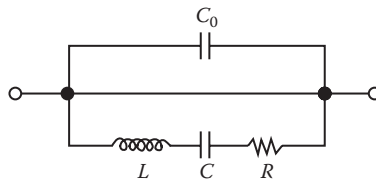


FIGURE 2.6

Equivalent network of a piezoelectric resonator. (After Tate, J.P. and Mead, P.F., Crystal oscillators, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 185–199, 1996.)

TABLE 2.1

Typical Equivalent Circuit Parameter Values for Quartz Crystals

Parameter	200 kHz Fundamental	2 MHz	30 MHz 3rd Overtone	90 MHz 5th Overtone
R_1	2 k Ω	100 Ω	20 Ω	40 Ω
L_1	27 H	520 mH	11 mH	6 mH
C_1	0.024 pF	0.012 pF	0.0026 pF	0.0005 pF
C_0	9 pF	4 pF	6 pF	4 pF
Q	18×10^3	18×10^3	18×10^3	18×10^3

Source: After Tate, J.P. and Mead, P.F., Crystal oscillators, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 185–199, 1996.

The vibration is enhanced for electric signals near a resonant frequency of the resonator. Low-frequency resonances tend to occur in the face shear modes, while high-frequency resonances occur in thickness shear modes. In either case, when the applied electric signal is near a resonant frequency, the energy required to maintain the vibrational motion is small.

2.2.2.3 Temperature Compensation

Compensation for temperature-related frequency shifts can be accomplished by connecting a reactance, such as a variable voltage capacitor (a *varactor*) in series with the resonator [1]. The effect of this reactance is to *pull* the frequency to cancel the effect of frequency drift caused by temperature changes. An example of a frequency vs. temperature characteristic both before and after compensation is shown in Figure 2.7.

Several techniques for temperature compensation are possible. They include

- Analog temperature compensation
- Hybrid analog–digital compensation
- Digital temperature compensation

For analog compensation, a varactor is controlled by a voltage divider network that contains both thermistors and resistors to manage the compensation. Stability values are in the range 0.5–10 ppm. The lowest value requires that individual components in the voltage divider network be adjusted under controlled-temperature conditions. This is difficult over large temperature ranges for analog networks because of the tight component value tolerances required and because of interactions between

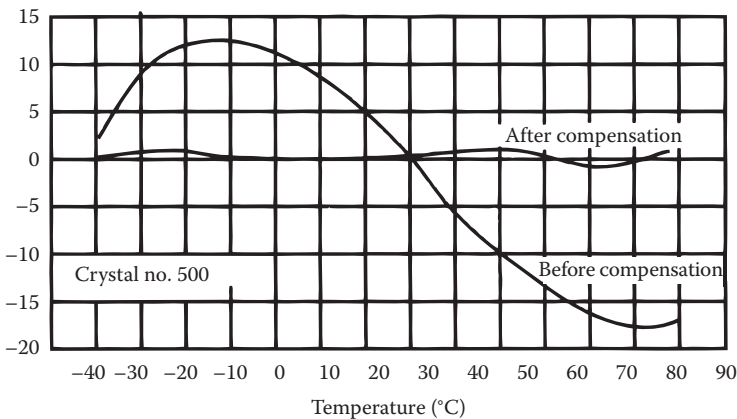


FIGURE 2.7

Frequency vs. temperature characteristics for a typical TCXO. (After Frerking, M.E., *Crystal Oscillator Design and Temperature Compensation*, Van Nostrand Reinhold, New York, 1978.)

different analog segments. Analog networks can be segmented in order to increase the independence of the component adjustments required to improve stability.

These problems can be alleviated through the use of some form of hybrid or digital compensation. In this scheme, the crystal oscillator is connected to a varactor in series, as in the analog case. This capacitor provides the coarse compensation (about ± 4 parts in 10^7). The fine compensation is provided by the digital network to further improve the stability. The hybrid analog–digital method is then capable of producing stability values of ± 5 parts in 10^8 over the range -40°C to $+80^\circ\text{C}$.

2.2.2.4 Stress Compensation

Stress compensation is employed in order to stabilize the crystal resonator frequency against shifts caused by mechanical stresses [1]. These result from temperature changes or from electrode changes, due to differences in the expansion coefficients, which can occur over time. The frequency variations are caused by alteration of the crystal elastic coefficients with stress bias. The method developed for stress compensation is a special crystal cut called the *stress-compensated* or SC-cut. This cut was developed to obtain a resonator whose frequency would be independent of specified stresses.

The SC-cut has some advantages over the more common AT- and BT-cuts. These include a flatter frequency-temperature response, especially at the turning points, and less susceptibility to aging because of changes in the mounting stresses. The lower sensitivity to frequency changes means that SC-cut resonators can be frequency-temperature tested more rapidly than their AT- and BT-cut counterparts. A disadvantage of the SC-cut is that it is sensitive to air damping and must be operated in a vacuum environment to obtain optimum Q factors. The frequency-temperature curve is also roughly 70° higher for the SC-cut over the more common AT-cut, which could also be a disadvantage in some applications.

2.2.2.5 Aging Effects

Aging in crystal oscillators generally refers to any change over time that will affect the frequency characteristics of the oscillator, or the physical parameters that describe the device, such as motional time constant and equivalent circuit parameters [1]. Some factors that influence aging include the following:

- Surface deterioration
- Surface contamination
- Electrode composition
- Environmental conditions

Deterioration on the surface contributes to strain in the crystal, and this type of strain can impact the operating frequency of the oscillator. More importantly, the surface strain is sensitive to humidity and temperature early in the life of the crystal. One method of combating this surface degradation mechanism is to etch away the strained surface layers. The initial slope of the frequency change vs. etch time curve is exponential. However, after the strained surface layers have been removed, the slope becomes linear.

Another surface issue is contamination from foreign materials. The presence of foreign substances will load the oscillator with an additional mass, and this can be reflected in the equivalent circuit parameters. The contaminants can be introduced during fabrication processes, from outgassing of other materials after the device has been packaged, or from contaminants in the environment during storage. The contamination process is often increased in higher temperature environments, and temperature cycling can lead to stability problems for the device.

Aging problems associated with the device electrodes includes delamination, leading to increased resistance or corrosion effects. Typical electrode materials include gold, silver, and aluminum. Of these, silver is the most popular material.

2.3 Operating Class

Amplifier stage operating efficiency is a key element in the design and application of an electron tube device or system. As the power level of an RF generator increases, the overall efficiency of the system becomes more important. Increased efficiency translates into lower operating costs and, usually, improved reliability of the system. Put another way, for a given device dissipation, greater operating efficiency translates into higher power output. The operating mode of the final stage, or stages, is the primary determining element in the maximum possible efficiency of the system.

An electron amplifying device is classified by its individual *class of operation*. Four primary class divisions apply to vacuum tube devices:

- *Class A*. A mode wherein the power amplifying device is operated over its linear transfer characteristic. This mode provides the lowest waveform distortion, but also the lowest efficiency. The basic operating efficiency of a class A stage is 50%. Class A amplifiers exhibit low intermodulation distortion, making them well suited to linear RF amplifier applications.
- *Class B*. A mode wherein the power amplifying device is operated just outside its linear transfer characteristic. This mode provides improved efficiency at the expense of some waveform distortion.

Class AB is a variation on class B operation. The transfer characteristic for an amplifying device operating in this mode is, predictably, between class A and class B.

- *Class C*. A mode wherein the power amplifying device is operated significantly outside its linear transfer characteristic, resulting in a pulsed output waveform. High efficiency (up to 90%) can be realized with class C operation, but significant distortion of the waveform will occur. Class C is used extensively as an efficient RF power generator.
- *Class D*. A mode that essentially results in a switched device state. The power amplifying device is either *on* or *off*. This is the most efficient mode of operation. It is also the mode that produces the greatest waveform distortion.

The angle of current flow determines the class of operation for a power amplifying device. Typically, the following generalizations regarding conduction angle apply:

- Class A: 360°
- Class AB: between 180° and 360°
- Class B: 180°
- Class C: less than 180°

Subscripts also may be used to denote grid current flow. The subscript "1" means that no grid current flows in the stage; the subscript "2" denotes grid current flow. Figure 2.8 charts operating efficiency as a function of conduction angle for an RF amplifier.

The class of operation is not directly related to the type of amplifying circuit. Vacuum tube stages of a given operating class may be grid or cathode driven.

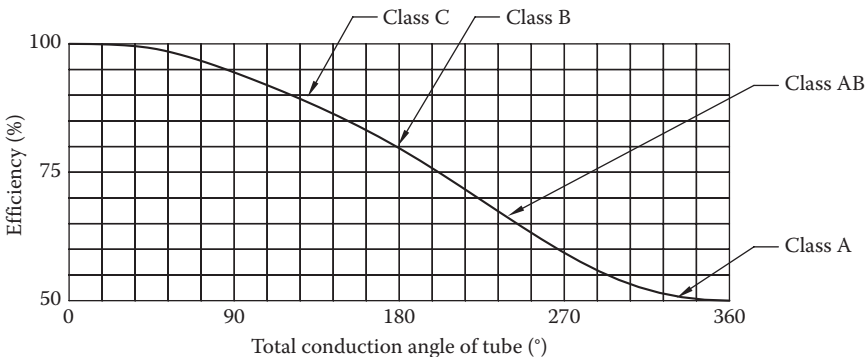


FIGURE 2.8

Plate efficiency as a function of conduction angle for an amplifier with a tuned load.

2.3.1 Operating Efficiency

The design goal of all RF amplifiers is to convert input power into an RF signal at the greatest possible efficiency. Direct current input power that is not converted into a useful output signal is, for the most part, converted to heat. This heat is wasted energy, which must be removed from the amplifying device. Removal of heat is a problem common to all high-power RF amplifiers. Cooling methods include

- Natural convection
- Radiation
- Forced convection
- Liquid
- Conduction
- Evaporation

The type of cooling method chosen is dictated in large part by the type of active device used and the power level involved. For example, liquid cooling is used almost exclusively for high-power (100 kW) vacuum tubes; conduction is used most often for low-power (10 kW) devices.

2.3.2 Broadband Amplifier Design

RF design engineers face a continuing challenge to provide adequate bandwidth for the signals to be transmitted while preserving as much efficiency from the overall system as possible. These two parameters, although not mutually exclusive, often involve trade-offs for both designers and operators.

An ideal RF amplifier will operate over a wide band of frequencies with minimum variations in output power, phase, distortion, and efficiency. The bandwidth of the amplifier depends to a great extent on the type of active device used, the frequency range required, and the operating power. As a general rule, a bandwidth of 20% or greater at frequencies above 100 MHz can be considered *broadband*. Below 100 MHz, broadband amplifiers typically have a bandwidth of one octave or more.

2.3.2.1 Stagger Tuning

Several stages with narrowband response (relative to the desired system bandwidth) can be cascaded and, through the use of *stagger tuning*, made broadband. Although there is an efficiency penalty for this approach, it has been used for years in all types of equipment. The concept is simple: Offset the center operating frequencies (and, therefore, peak amplitude response) of the cascaded amplifiers so that the resulting passband is flat and broad.

For example, the first stage in a three-stage amplifier is adjusted for peak response at the center operating frequency of the system. The second stage is adjusted above the center frequency, and the third stage is adjusted below center. The resulting composite response curve yields a broadband trace. The efficiency penalty for this scheme varies depending on the power level of each stage, the amount of stagger tuning required to achieve the desired bandwidth, and the number of individual stages.

2.3.2.2 Matching Circuits

The individual stages of an RF generator must be coupled together. Rarely do the output impedance and power level of one stage precisely match the input impedance and signal-handling level of the next stage. There is a requirement, therefore, for broadband matching circuits. Matching at radio frequencies can be accomplished with several different techniques, including

- *A ¼-wave transformer:* A matching technique using simply a length of transmission line ¼-wave long, with a characteristic impedance of $Z_{line} = \sqrt{Z_{in} \times Z_{out}}$, where Z_{in} and Z_{out} are the terminating impedances. Multiple ¼-wave transformers can be cascaded to achieve more favorable matching characteristics. Cascaded transformers permit small matching ratios for each individual section.
- *A balun transformer:* A transmission-line transformer in which the turns are arranged physically to include the interwinding capacitance as a component of the characteristic impedance of the transmission line. This technique permits wide bandwidths to be achieved in the device without unwanted resonances.

Balun transformers usually are made of twisted wire pairs or twisted coaxial lines. Ferrite toroids may be used as the core material.

2.3.2.3 Power Combining

The two most common methods of extending the operating power of active devices are *direct paralleling* of components and *hybrid splitting/combining*. Direct paralleling has been used for tube designs, but application of this simple approach is limited by variations in device-operating parameters. Two identical devices in parallel do not necessarily draw the same amount of current (supply the same amount of power to the load). Paralleling at UHF and above can be difficult because of the restrictions of operating wavelength.

The preferred approach involves the use of identical stages driven in parallel from a *hybrid coupler*. The coupler provides a constant-source impedance and directs any reflected energy from the driven stages to a *reject port* for dissipation.

A hybrid coupler offers a standing-wave-canceling effect that improves system performance. Hybrids also provide a high degree of isolation between active devices in a system.

2.4 Thermal and Circuit Noise

All electronic circuits are affected by any number of factors that cause their performance to be degraded from the ideal assumed in simple component models, in ways that can be controlled but not eliminated entirely [4]. One limitation is the failure of the model to account properly for the real behavior of components, either the result of an oversimplified model description or variations in manufacture. Usually, by careful design, the engineer can work around the limitations of the model and produce a device or circuit whose operation is very close to predictions. One source of performance degradation that cannot be easily overcome, however, is noise.

When vacuum tubes first came into use in the early part of the twentieth century, they were extensively used in radios as signal amplifiers and for other related signal conditioning functions. Thus, the measure of performance that was of greatest importance to circuit designers was the quality of the sound produced at the radio speaker. It was immediately noticed that the sound coming from speakers not only consisted of the transmitted signal but also of popping, crackling, and hissing sounds, which seemed to have no pattern and were distinctly different from the sounds that resulted from other interfering signal sources, such as other radio stations using neighboring frequencies. This patternless or random signal was labeled “noise” and has become the standard term for signals that are random and are combined with the circuit signal to affect the overall performance of the system.

As the study of noise has progressed, engineers have come to realize that there are many sources of noise in circuits. The following definitions are commonly used in discussions of circuit noise:

- *White noise.* Noise whose energy is evenly distributed over the entire frequency spectrum, within the frequency range of interest (typically below frequencies in the infrared range). Because noise is totally random it may seem inappropriate to refer to its frequency range, because it is not really periodic in the ordinary sense. Nevertheless, by examining an oscilloscope trace of white noise, one can verify that every trace is different—as the noise never repeats itself—and yet each trace looks the same. Similarly, an analog TV set tuned to an unused frequency displays never-ending “snow” that always looks the same, yet clearly is always changing. There is a strong theoretical foundation to represent the frequency content of such signals

as covering the frequency spectrum evenly. In this way the impact on other periodic signals can be analyzed. The term “white noise” arises from the fact that—similar to white light, which has equal amounts of all light frequencies—white noise has equal amounts of noise at all frequencies within circuit operating range.

- *Interference.* The name given to any predictable, periodic signal that occurs in an electronic circuit in addition to the signal the circuit is designed to process. This component is distinguished from a noise signal by the fact that it occupies a relatively small frequency range, and because it is predictable, it can often be filtered out. Interference usually results from another electronic system, such as an interfering radio source.
- *Thermal noise.* Any temperature-dependent noise generated within a circuit. This signal usually is the result of the influence of temperature directly on the operating characteristics of circuit elements, which—because of the random motion of molecules as a result of temperature—in turn creates a random fluctuation of the signal being processed.
- *Shot noise.* A type of circuit noise that is not temperature dependent and is not white noise in the sense that it tends to diminish at higher frequencies (see Figure 2.9). This noise usually occurs in components whose operation depends on a mean particle residence time for the active electrons within the device. The cutoff frequency above which noise disappears is closely related to the inverse of this characteristic particle residence time. It is called “shot noise” because in a radio it

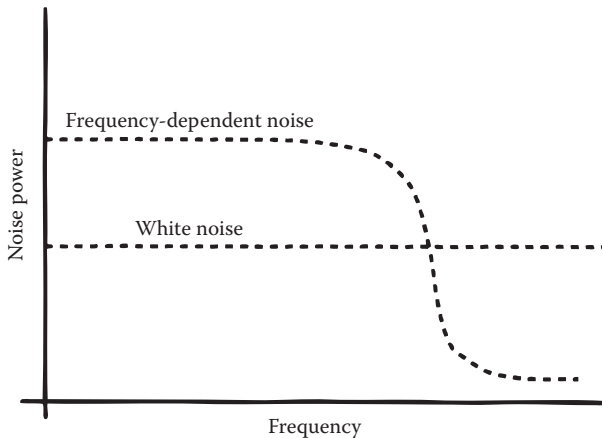


FIGURE 2.9

Plot of signal power vs. frequency for white noise and frequency limited noise. (After Douglass, B.G., Thermal noise and other circuit noise, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 30–36, 1996.)

can make a sound similar to buckshot hitting a drum, as opposed to white noise, which tends to sound more like hissing (because of the higher frequency content).

2.4.1 Thermal Noise

By definition, thermal noise is internally generated noise that is temperature dependent [4]. While first observed in vacuum tube devices (because their amplifying capabilities tend to bring out thermal noise), it is also observed in semiconductor devices. It is a phenomenon resulting from the ohmic resistance of devices that dissipate the energy lost in them as heat. Heat consists of random motion of molecules, which are more energetic as temperature increases. Because the motion is random, it is to be expected that as electrons pass through the ohmic device incurring resistance, there should be some random deviations in the rate of energy loss. This fluctuation has the effect of causing variation in the resulting current, which is noise. As the temperature of the device increases, the random motion of the molecules increases and so does the corresponding noise level.

2.4.2 Noise in Systems of Cascaded Stages

The noise level produced by thermal noise sources is not necessarily large; however, because source signal power may also be low, it is usually necessary to amplify the source signal. Because noise is combined with the source signal, and both are then amplified, with more noise added at each successive stage of amplification, noise can become a noticeable phenomenon [4].

2.5 Amplitude Modulation

In the simplest form of amplitude modulation, an analog carrier is controlled by an analog modulating signal. The desired result is an RF waveform whose amplitude is varied by the magnitude of the applied modulating signal and at a rate equal to the frequency of the applied signal. The resulting waveform consists of a carrier wave plus two additional signals:

- An upper-sideband signal, which is equal in frequency to the carrier *plus* the frequency of the modulating signal
- A lower-sideband signal, which is equal in frequency to the carrier *minus* the frequency of the modulating signal

This type of modulation system is referred to as *double-sideband amplitude modulation* (DSAM).

The radio carrier wave signal onto which the analog amplitude variations are to be impressed is expressed as

$$e(t) = AE_c \cos(\omega_c t) \quad (2.4)$$

where

$e(t)$ is the instantaneous amplitude of the carrier wave as a function of time (t)

A is a factor of amplitude modulation of the carrier wave

ω_c is the angular frequency of the carrier wave (rad/s)

E_c is the peak amplitude of the carrier wave

If A is a constant, the peak amplitude of the carrier wave is constant, and no modulation exists. Periodic modulation of the carrier wave results if the amplitude of A is caused to vary with respect to time, as in the case of a sinusoidal wave:

$$A = 1 + \left(\frac{E_m}{E_c} \right) \cos(\omega_m t) \quad (2.5)$$

where E_m/E_c is the ratio of modulation amplitude to carrier amplitude.

The foregoing relationship leads to

$$e(t) = E_c \left[1 + \left(\frac{E_m}{E_c} \right) \cos(\omega_m t) \right] \cos(\omega_c t) \quad (2.6)$$

This is the basic equation for periodic (sinusoidal) amplitude modulation. When all multiplications and a simple trigonometric identity are performed, the result is

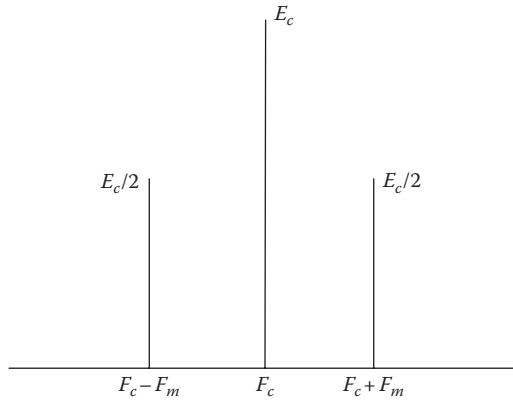
$$e(t) = E_c \cos(\omega_c t) + \left(\frac{M}{2} \right) \cos(\omega_c t + \omega_m t) + \left(\frac{M}{2} \right) \cos(\omega_c t - \omega_m t) \quad (2.7)$$

where M is the amplitude modulation factor E_m/E_c .

Amplitude modulation is, essentially, a multiplication process in which the time functions that describe the modulating signal and the carrier are multiplied to produce a modulated wave containing *intelligence* (information or data of some kind). The frequency components of the modulating signal are translated in this process to occupy a different position in the spectrum.

The bandwidth of an AM transmission is determined by the modulating frequency. The bandwidth required for full-fidelity reproduction in a receiver is equal to twice the applied modulating frequency.

The magnitude of the upper sideband and lower sideband will not normally exceed 50% of the carrier amplitude during modulation. This results in


FIGURE 2.10

Frequency-domain representation of an amplitude-modulated signal at 100% modulation. E_c , carrier power; F_c , frequency of the carrier; and F_m , frequency of the modulating signal.

an upper-sideband power of one-fourth the carrier power. The same power exists in the lower sideband. As a result, up to one-half of the actual carrier power appears additionally in the sum of the sidebands of the modulated signal. A representation of the AM carrier and its sidebands is shown in Figure 2.10. The actual occupied bandwidth, assuming pure sinusoidal modulating signals and no distortion during the modulation process, is equal to twice the frequency of the modulating signal.

The extent of the amplitude variations in a modulated wave is expressed in terms of the *degree of modulation or percentage of modulation*. For sinusoidal variation, the degree of modulation m is determined from the following:

$$m = \frac{E_{avg} - E_{min}}{E_{avg}} \quad (2.8)$$

where

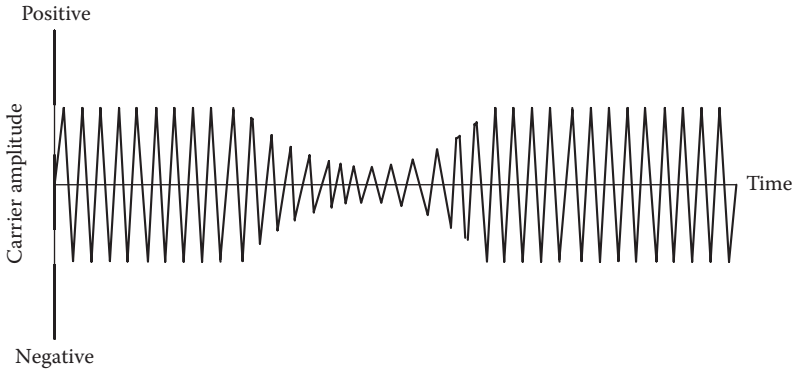
E_{avg} is the average envelope amplitude

E_{min} is the minimum envelope amplitude

Full (100%) modulation occurs when the peak value of the modulated envelope reaches twice the value of the unmodulated carrier, and the minimum value of the envelope is zero. The envelope of a modulated AM signal in the time domain is shown in Figure 2.11.

When the envelope variation is not sinusoidal, it is necessary to define the degree of modulation separately for the peaks and troughs of the envelope:

$$m_{pp} = \frac{E_{max} - E_{avg}}{E_{avg}} \times 100 \quad (2.9)$$

**FIGURE 2.11**

Time-domain representation of an amplitude-modulated signal. Modulation at 100% is defined as the point at which the peak of the waveform reaches twice the carrier level, and the minimum point of the waveform is zero.

$$m_{np} = \frac{E_{avg} - E_{min}}{E_{avg}} \times 100 \quad (2.10)$$

where

- m_{pp} is the positive peak modulation (%)
- E_{max} is the peak value of modulation envelope
- m_{np} is the negative peak modulation (%)
- E_{avg} is the average envelope amplitude
- E_{min} is the minimum envelope amplitude

When modulation exceeds 100% on the negative swing of the carrier, spurious signals are emitted. It is possible to modulate an AM carrier asymmetrically, that is, to restrict modulation in the negative direction to 100%, but to allow modulation in the positive direction to exceed 100% without a significant loss of fidelity. In fact, many modulating signals normally exhibit asymmetry, most notably human speech waveforms.

The carrier wave represents the average amplitude of the envelope and, because it is the same regardless of the presence or absence of modulation, the carrier transmits no information. The information is carried by the sideband frequencies. The amplitude of the modulated envelope may be expressed as [5]

$$E = E_0 + E_1 \sin(2\pi f_1 t + \Phi_1) + E_2 \sin(2\pi f_2 t + \Phi_2) \quad (2.11)$$

where

- E is the envelope amplitude
- E_0 is the carrier wave crest value (V)
- E_1 is $2 \times$ first sideband crest amplitude (V)

f_1 is the frequency difference between the carrier and the first upper/lower sidebands

E_2 is $2 \times$ second sideband crest amplitude (V)

f_2 is the frequency difference between the carrier and the second upper/lower sidebands

Φ_1 is the phase of the first sideband component

Φ_2 is the phase of the second sideband component

2.5.1 High-Level AM Modulation

High-level anode modulation is the oldest and simplest method of generating a high-power AM signal. In this system, the modulating signal is amplified and combined with the dc supply source to the anode of the final RF amplifier stage. The RF amplifier is normally operated class C. The final stage of the modulator usually consists of a pair of tubes operating class B in a push-pull configuration. A basic high-level modulator is shown in Figure 2.12.

The RF signal normally is generated in a low-level transistorized oscillator. It is then amplified by one or more solid-state or vacuum tube stages to provide final RF drive at the appropriate frequency to the grid of the final class C amplifier. The modulation input is applied to an intermediate power amplifier (usually solid state) and used to drive two class B (or class AB) push-pull output devices. The final amplifiers provide the necessary modulating power

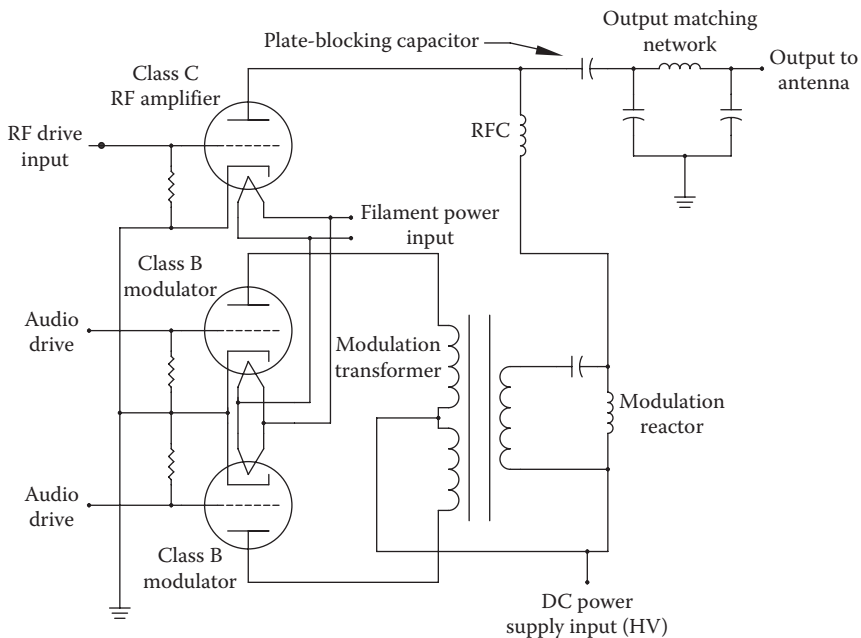


FIGURE 2.12
Simplified diagram of a high-level amplitude-modulated amplifier.

to drive the final RF stage. For 100% modulation, this modulating power is equal to 50% of the actual carrier power.

The modulation transformer shown in Figure 2.12 does not usually carry the dc supply current for the final RF amplifier. The modulation reactor and capacitor shown provide a means to combine the signal voltage from the modulator with the dc supply to the final RF amplifier. This arrangement eliminates the necessity of having direct current flow through the secondary of the modulation transformer, which would result in magnetic losses and saturation effects. In some transmitter designs, the modulation reactor is eliminated from the system, thanks to improvements in transformer technology.

The RF amplifier normally operates class C with grid current drawn during positive peaks of the cycle. Typical stage efficiency is 75%–83%. An RF tank following the amplifier resonates the output signal at the operating frequency and, with the assistance of a low-pass filter, eliminates harmonics of the amplifier caused by class C operation.

This type of system was popular in AM applications for many years, largely because of its simplicity. The primary drawback is low overall system efficiency. The class B modulator tubes cannot operate with greater than 50% efficiency. Still, with inexpensive electricity, this was not considered to be a significant problem. As energy costs increased, however, more efficient methods of generating high-power AM signals were developed. Increased efficiency normally came at the expense of added technical complexity.

2.5.2 Vestigial-Sideband Amplitude Modulation

Because the intelligence (modulating signal) of conventional AM transmission is identical in the upper *and* lower sidebands, it is possible to eliminate one sideband and still convey the required information. This scheme is implemented in *vestigial-sideband AM* (VSBAM). Complete elimination of one sideband (e.g., the lower sideband) requires an ideal high-pass filter with infinitely sharp cutoff. Such a filter is difficult to implement in any practical design. VSBAM is a compromise technique wherein one sideband (typically the lower sideband) is attenuated significantly. The result is a savings in occupied bandwidth and transmitter power.

VSBAM is used for analog television broadcast transmission and other applications. A typical bandwidth trace for a VSBAM analog TV transmitter is shown in Figure 2.13.

2.5.3 Single-Sideband Amplitude Modulation

The carrier in an AM signal does not convey any intelligence. All of the modulating information is in the sidebands. It is possible, therefore, to suppress the carrier upon transmission, radiating only one or both sidebands of the AM signal. The result is much greater efficiency at the transmitter (i.e., a reduction in the required transmitter power). *Single-sideband suppressed carrier AM* (SSB-SC) is

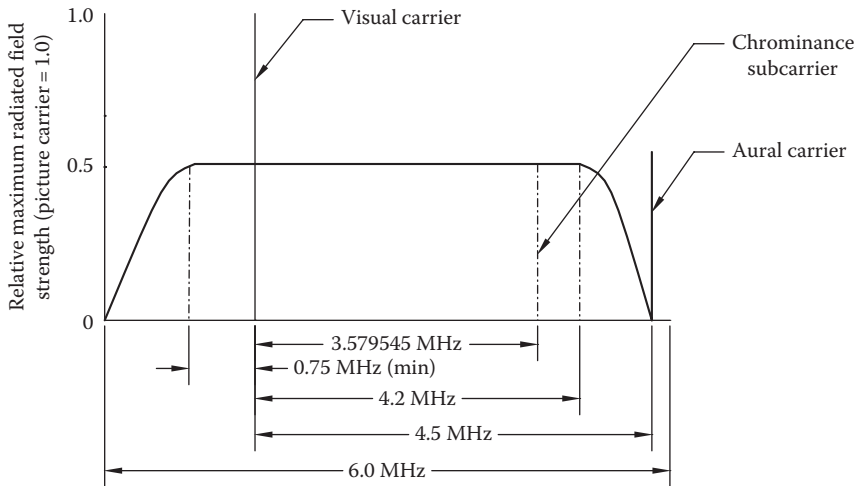


FIGURE 2.13

Idealized amplitude characteristics of the FCC standard waveform for monochrome and color analog TV transmission. (Adapted from FCC Rules, Sec. 73.699.)

the most spectrum- and energy-efficient mode of AM transmission. Figure 2.14 shows representative waveforms for suppressed carrier transmissions.

A waveform with carrier suppression differs from a modulated wave containing a carrier primarily in that the envelope varies at twice the modulating frequency. In addition, it will be noted that the SSB-SC wave has an apparent phase that reverses every time the modulating signal passes through zero. The wave representing a single sideband consists of a number of frequency components, one for each component in the original signal. Each of these components has an amplitude proportional to the amplitude of the corresponding modulating component and a frequency differing from that of the carrier by the modulating frequency. The result is that, in general, the envelope amplitude of the single-sideband signal increases with the degree of modulation, and the envelope varies in amplitude in accordance with the difference frequencies formed by the various frequency components of the single sideband interacting with each other.

An SSB-SC system is capable of transmitting a given intelligence within a frequency band only half as wide as that required by a DSAM waveform. Furthermore, the SSB system saves more than two-thirds of the transmission power because of the elimination of one sideband and the carrier.

The drawback to suppressed carrier systems is the requirement for a more complicated receiver. The carrier must be regenerated at the receiver to permit demodulation of the signal. Also, in the case of SSBAM transmitters, it is usually necessary to generate the SSB signal in a low-power stage and then amplify the signal with a linear power amplifier to drive the antenna. Linear amplifiers generally exhibit low efficiency.

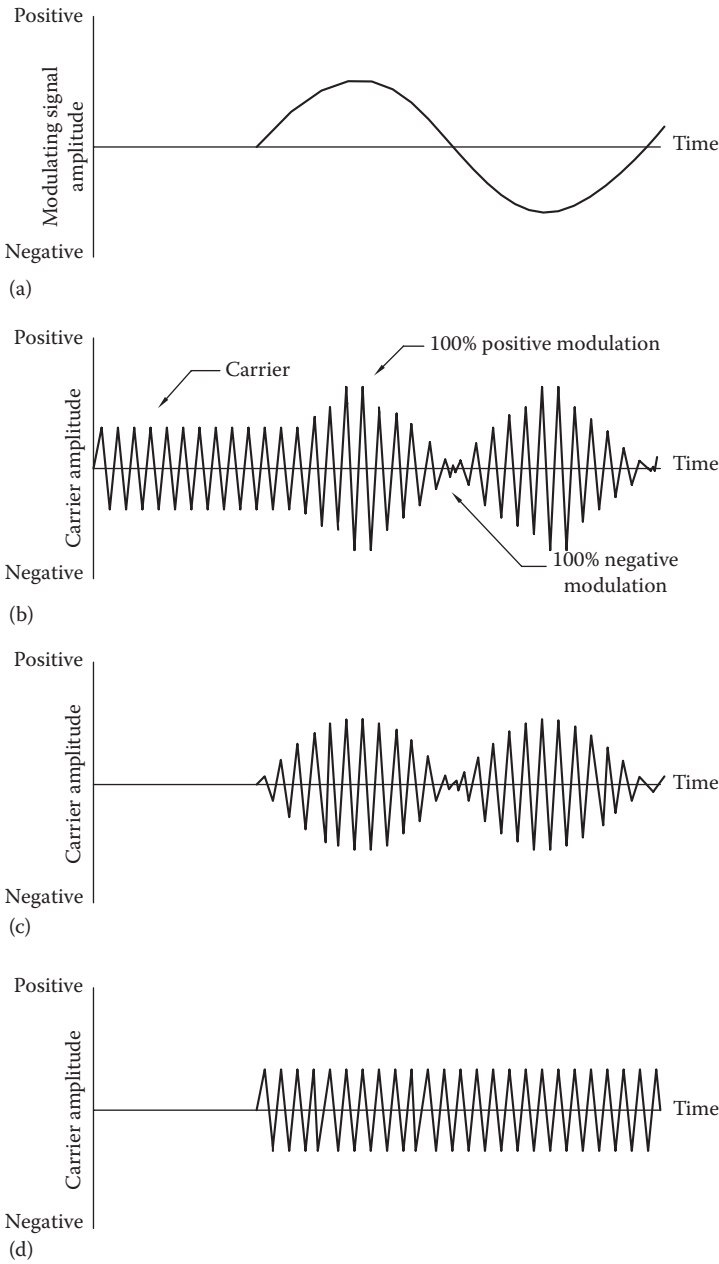


FIGURE 2.14 Types of suppressed carrier amplitude modulation: (a) the modulating signal, (b) double-sideband AM signal, (c) double-sideband suppressed carrier AM, and (d) single-sideband suppressed carrier AM.

2.5.4 Quadrature Amplitude Modulation

Single-sideband transmission makes very efficient use of the spectrum; for example, two SSB signals can be transmitted within the bandwidth normally required for a single DSB signal. However, DSB signals can achieve the same efficiency by means of *quadrature amplitude modulation* (QAM), which permits two DSB signals to be transmitted and received simultaneously using the same carrier frequency [6]. A basic QAM DSB modulator is shown schematically in Figure 2.15.

Two DSB signals coexist separately within the same bandwidth by virtue of the 90° phase shift between them. The signals are, thus, said to be in *quadrature*. Demodulation uses two local oscillator signals that are also in quadrature, that is, a sine and a cosine signal, as illustrated in Figure 2.16.

The chief disadvantage of QAM is the need for a coherent local oscillator at the receiver exactly in phase with the transmitter oscillator signal. Slight errors in phase or frequency can cause both loss of signal and interference between the two signals (co-channel interference or cross talk).

The relative merits of the various AM systems are summarized in Table 2.2.

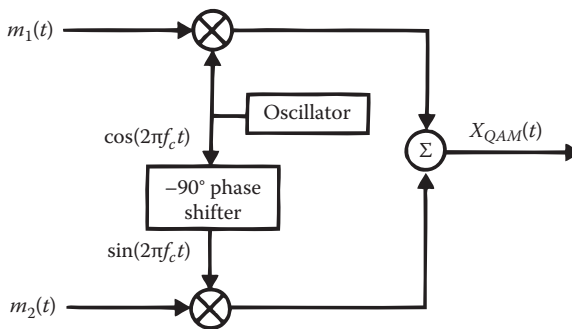


FIGURE 2.15 Simplified QAM modulator. (After Kubichek, R., Amplitude modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1175–1187, 1996.)

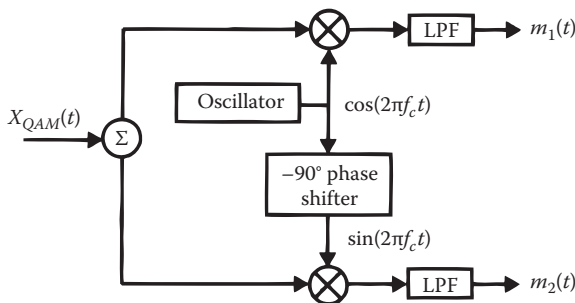


FIGURE 2.16 Simplified QAM demodulator. (After Kubichek, R., Amplitude modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1175–1187, 1996.)

TABLE 2.2

Comparison of Amplitude Modulation Techniques

Modulation Scheme	Advantages	Disadvantages	Comments
DSB SC	Good power efficiency; good low-frequency response	More difficult to generate than DSB + C; detection requires coherent local oscillator, pilot, or phase-locked loop; poor spectrum efficiency	
DSB + C	Easier to generate than DSB SC, especially at high-power levels; inexpensive receivers using envelope detection	Poor power efficiency; poor spectrum efficiency; poor low-frequency response; exhibits threshold effect in noise	Used in commercial AM broadcasting
SSB SC	Excellent spectrum efficiency	Complex transmitter design; complex receiver design (same as DSB SC); poor low-frequency response	Used in military communication systems, and to multiplex multiple phone calls onto long-haul microwave links
SSB + C	Good spectrum efficiency; low receiver complexity	Poor power efficiency; complex transmitters; poor low-frequency response; poor noise performance	
VSF SC	Good spectrum efficiency; excellent low-frequency response; transmitter easier to build than for SSB	Complex receivers (same as DSB SC)	
VSF + C	Good spectrum efficiency; good low-frequency response; inexpensive receivers using envelope detection	Poor power efficiency; poor performance in noise	Used in commercial TV broadcasting
QAM	Good low-frequency response; good spectrum efficiency	Complex receivers; sensitive to frequency and phase errors	Two SSB signals may be preferable

Source: After Kubichek, R., Amplitude modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1175–1187, 1996.

2.6 Frequency Modulation

Frequency modulation (FM) is a technique whereby the phase angle or phase shift of a carrier is varied by an applied modulating signal. The *magnitude* of frequency change of the carrier is a direct function of the *magnitude* of the modulating signal. The *rate* at which the frequency of the carrier is changed is a direct

function of the *frequency* of the modulating signal. In FM modulation, multiple pairs of sidebands are produced. The actual number of sidebands that make up the modulated wave is determined by the *modulation index* (MI) of the system.

2.6.1 Modulation Index

The MI is a function of the frequency deviation of the system and the applied modulating signal as follows:

$$MI = \frac{F_d}{M_f} \quad (2.12)$$

where

MI is the modulation index

F_d is the frequency deviation

M_f is the modulating frequency

The higher the MI , the more the sidebands produced. It follows that the higher the modulating frequency for a given deviation, the fewer the number of sidebands produced, but the greater their spacing.

To determine the frequency spectrum of a transmitted FM waveform, it is necessary to compute a Fourier series or Fourier expansion to show the actual signal components involved. This work is difficult for a waveform of this type, because the integrals that must be performed in the Fourier expansion or Fourier series are not easily solved. The result, however, is that the integral produces a particular class of solution that is identified as the *Bessel function*, illustrated in Figure 2.17.

The carrier amplitude and phase, plus the sidebands, can be expressed mathematically by making the MI the argument of a simplified Bessel function. The general expression is given from the following equations:

$$\text{RF output voltage} = E_t = E_c + S_{1u} - S_{1l} + S_{2u} - S_{2l} + S_{3u} - S_{3l} + S_{nu} - S_{nl}$$

$$\text{Carrier amplitude} = E_c = A[J_0(M)\sin\omega_c(t)]$$

$$\text{First-order upper sideband} = S_{1u} = J_1(M)\sin(\omega_c + \omega_m)t$$

$$\text{First-order lower sideband} = S_{1l} = J_1(M)\sin(\omega_c - \omega_m)t$$

$$\text{Second-order upper sideband} = S_{2u} = J_2(M)\sin(\omega_c + 2\omega_m)t$$

$$\text{Second-order lower sideband} = S_{2l} = J_2(M)\sin(\omega_c - 2\omega_m)t$$

$$\text{Third-order upper sideband} = S_{3u} = J_3(M)\sin(\omega_c + 3\omega_m)t$$

$$\text{Third-order lower sideband} = S_{3l} = J_3(M)\sin(\omega_c - 3\omega_m)t$$

$$\text{Nth-order upper sideband} = S_{nu} = J_n(M)\sin(\omega_c + n\omega_m)t$$

$$\text{Nth-order lower sideband} = S_{nl} = J_n(M)\sin(\omega_c - n\omega_m)t$$

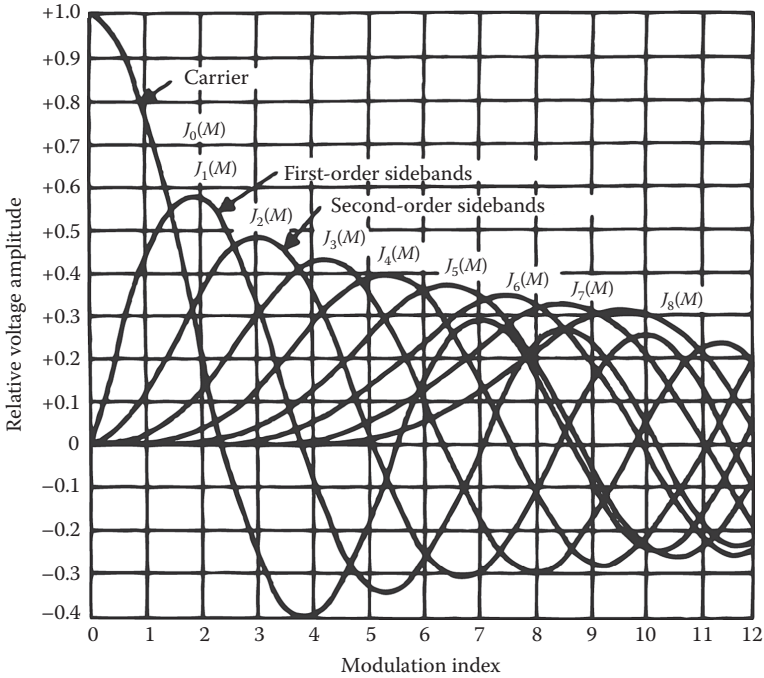


FIGURE 2.17
Plot of Bessel functions of the first kind as a function of MI.

where

- A is the unmodulated carrier amplitude constant
- J_0 is the modulated carrier amplitude
- $J_1, J_2, J_3, \dots, J_n$ are amplitudes of the n th-order sidebands
- M is the modulation index
- ω_c is $2\pi F_c$, the carrier frequency
- ω_m is $2\pi F_m$, the modulating frequency

Further supporting mathematics will show that an FM signal using the modulation indices that occur in a wideband system will have a multitude of sidebands. From the purist point of view, *all* sidebands would have to be transmitted, received, and demodulated to reconstruct the modulating signal with complete accuracy. In practice, however, the channel bandwidths permitted FM systems usually are sufficient to reconstruct the modulating signal with little discernible loss in fidelity, or at least an acceptable loss in fidelity.

Figure 2.18 illustrates the frequency components present for a MI of 5. Figure 2.19 shows the components for an index of 15. Note that the number of significant sideband components becomes quite large with a high MI . This simple representation of a single-tone frequency-modulated spectrum is useful for understanding the general nature of FM, and for making tests

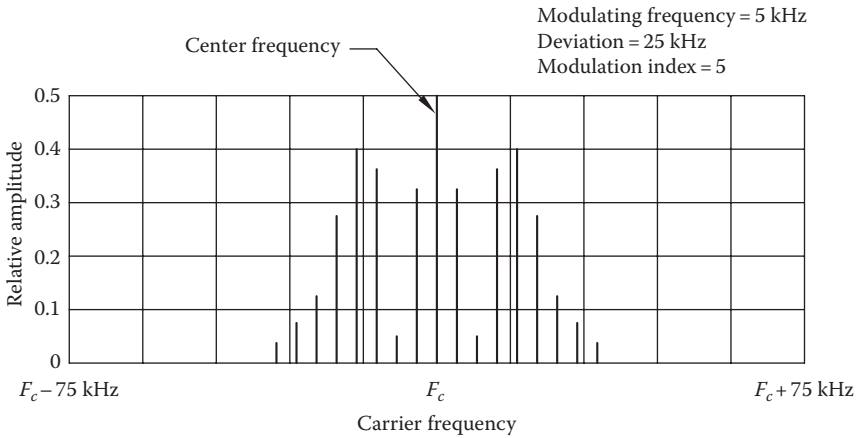


FIGURE 2.18 RF spectrum of a frequency-modulated signal with a *MI* of 5 and other operating parameters as shown.

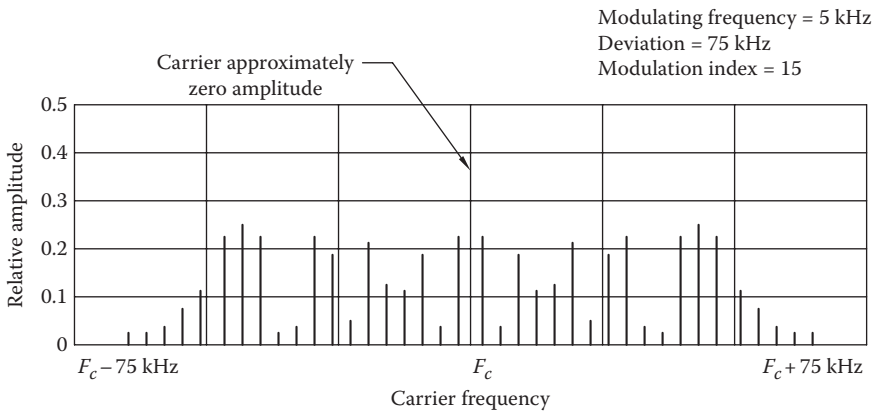


FIGURE 2.19 RF spectrum of a frequency-modulated signal with a *MI* of 15 and operating parameters as shown.

and measurements. When typical modulation signals are applied, however, many more sideband components are generated. These components vary to the extent that sideband energy becomes distributed over the entire occupied bandwidth, rather than appearing at discrete frequencies.

Although complex modulation of an FM carrier greatly increases the number of frequency components present in the frequency-modulated wave, it does not, in general, widen the frequency band occupied by the energy of the wave. To a first approximation, this band is still roughly twice the sum of the maximum frequency deviation at the peak of the modulation cycle plus the highest modulating frequency involved.

FM is not a simple frequency translation, as with AM, but involves the generation of entirely new frequency components. In general, the new spectrum is much wider than the original modulating signal. This greater bandwidth may be used to improve the *signal-to-noise ratio* (S/N) of the transmission system. FM thereby makes it possible to exchange bandwidth for S/N enhancement.

The power in an FM system is constant throughout the modulation process. The output power is increased in the amplitude modulation system by the modulation process, but the FM system simply distributes the power throughout the various frequency components that are produced by modulation. During modulation, a wideband FM system does not have a high amount of energy present in the carrier. Most of the energy will be found in the sum of the sidebands.

The constant-amplitude characteristic of FM greatly assists in capitalizing on the low noise advantage of FM reception. Upon being received and amplified, the FM signal normally is clipped to eliminate all amplitude variations above a certain threshold. This removes noise picked up by the receiver as a result of man-made or atmospheric signals. It is not possible (generally speaking) for these random noise sources to change the frequency of the desired signal; they can affect only its amplitude. The use of *hard limiting* in the receiver will strip off such interference.

2.6.2 Phase Modulation

In a phase modulation (PM) system, intelligence is conveyed by varying the phase of the RF wave. PM is similar in many respects to FM, except in the interpretation of the *MI*. In the case of PM, the *MI* depends only on the amplitude of the modulation; *MI* is independent of the frequency of the modulating signal. It is apparent, therefore, that the phase-modulated wave contains the same sideband components as the FM wave and, if the modulation indices in the two cases are the same, the relative amplitudes of these different components also will be the same.

The modulation parameters of a PM system relate as follows:

$$\Delta f = m_p \times f_m \quad (2.13)$$

where

Δf is the frequency deviation of the carrier

m_p is the phase shift of the carrier

f_m is the modulating frequency

In a phase-modulated wave, the phase shift m_p is independent of the modulating frequency; the frequency deviation Δf is proportional to the modulating frequency. In contrast, with a frequency-modulated wave, the frequency deviation is independent of modulating frequency. Therefore, a frequency-modulated wave can be obtained from a phase modulator by making the modulating voltage applied to the phase modulator inversely proportional to frequency. This can be readily achieved in hardware.

2.6.3 Modifying FM Waves

When a frequency-modulated wave is passed through a harmonic generator, the effect is to increase the MI by a factor equal to the frequency multiplication involved. Similarly, if the frequency-modulated wave is passed through a frequency divider, the effect is to reduce the MI by the factor of frequency division. Thus, the frequency components contained in the wave—and, consequently, the bandwidth of the wave—will be increased or decreased, respectively, by frequency multiplication or division. No distortion in the nature of the modulation is introduced by the frequency change.

When an FM wave is translated in the frequency spectrum by heterodyne action, the MI—hence, the relative positions of the sideband frequencies and the bandwidths occupied by them—remains unchanged.

2.6.4 Preemphasis and De-Emphasis

The FM transmission/reception system offers significantly better noise-rejection characteristics than AM. However, FM noise rejection is more favorable at low modulating frequencies than at high frequencies because of the reduction in the number of sidebands at higher frequencies. To offset this problem, the input signal to the FM transmitter may be *preemphasized* to increase the amplitude of higher-frequency signal components in normal program material. FM receivers utilize complementary *de-emphasis* to produce flat overall system frequency response.

FM broadcasting, for example, uses a $75\ \mu\text{s}$ preemphasis curve, meaning that the time constant of the resistance-inductance (RL) or resistance-capacitance (RC) circuit used to provide the boost of high frequencies is $75\ \mu\text{s}$. Other values of preemphasis are used in different types of FM communications systems. Figure 2.20 shows three common preemphasis curves.

2.6.5 Modulation Circuits

Early FM transmitters used *reactance modulators* that operated at a low frequency. The output of the modulator then was multiplied to reach the desired output frequency. This approach was acceptable for some applications and unsuitable for others. Modern FM systems utilize what is referred to as *direct modulation*; that is, the FM occurs in a modulated oscillator that operates on a center frequency equal to the desired transmitter output frequency.

Various techniques have been developed to generate the direct-FM signal. One of the most popular uses a variable-capacity diode as the reactive element in the oscillator. The modulating signal is applied to the diode, which causes the capacitance of the device to vary as a function of the magnitude of the modulating signal. Variations in the capacitance cause the frequency of the oscillator to change. The magnitude of the frequency shift is proportional to the amplitude of the modulating signal, and the rate of frequency shift is equal to the frequency of the modulating signal.

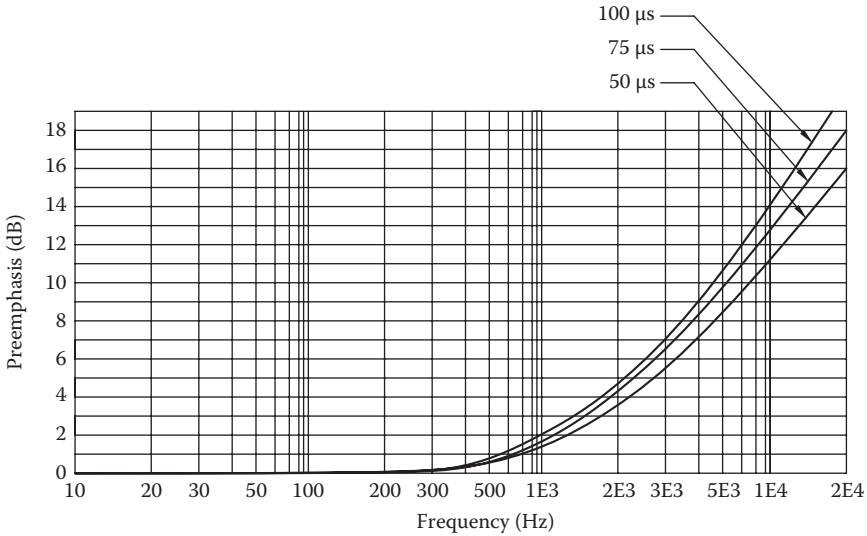


FIGURE 2.20
Preemphasis curves for time constants of 50, 75, and 100 μ s.

The direct-FM modulator is one element of an FM transmitter *exciter*, which generates the FM waveform. A block diagram of a complete FM exciter is shown in Figure 2.21. Input signals are buffered, filtered, and preemphasized before being summed to feed the modulated oscillator. Note that the oscillator is not normally coupled directly to a crystal, but to a free-running oscillator adjusted as closely as possible to the carrier frequency of the transmitter. The final operating frequency is maintained carefully by an automatic frequency control system employing a *phase locked loop* (PLL) tied to a reference crystal oscillator or frequency synthesizer.

2.6.5.1 Direct-FM Modulator

Many types of circuits have been used to produce a direct-FM signal [7]. In each case, a reactance device is used to shunt capacitive or inductive reactance

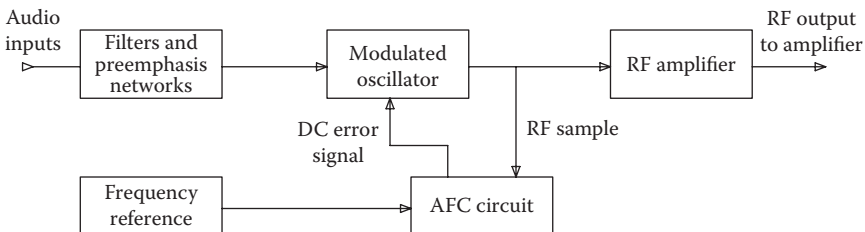


FIGURE 2.21
Block diagram of an FM exciter.

across an oscillator. The value of capacitive or inductive reactance is made to vary as the amplitude of the modulating signal varies. Because the reactive load is placed across an oscillator's tuned circuit, the frequency of the oscillator will therefore shift by a predetermined amount, thereby creating an FM signal.

A typical example of a reactance modulator is illustrated in Figure 2.22. The circuit uses a field-effect transistor (FET), where the modulating signal is applied to the modulator through C_1 . The actual components that affect the overall reactance consist of R_1 and C_2 . Typically, the value of C_2 is small as this is the input capacitance to the FET, which may only be a few picofarads. However, this capacitance will generally be much larger by a significant amount as a result of the *Miller effect*. Capacitor C_3 has no significant effect on the reactance of the modulator; it is strictly a blocking capacitor that keeps dc from changing the gate bias of the FET.

To further understand the performance of the reactance modulator, the equivalent circuit of Figure 2.22 is represented in Figure 2.23. The FET is represented as a current source, $g_m V_{g'}$ with the internal drain resistance, r_d . The impedances Z_1 and Z_2 are a combination of resistance and capacitive reactance, which are designed to provide a 90° phase shift.

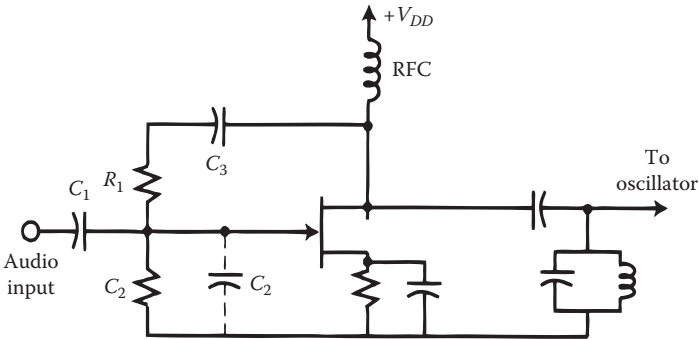


FIGURE 2.22 Simplified reactance modulator. (After Seymour, K., Frequency modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1188–1200, 1996.)

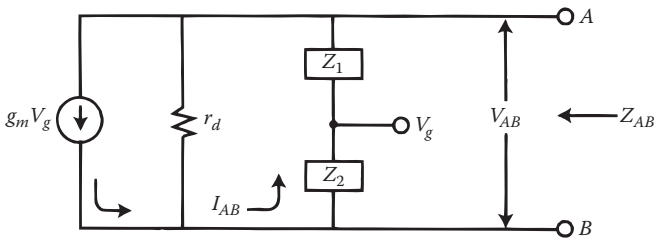
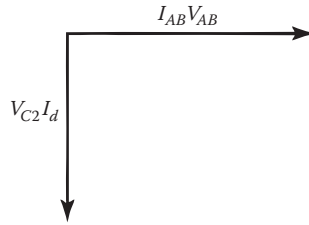


FIGURE 2.23 Equivalent circuit of the reactance modulator. (After Seymour, K., Frequency modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1188–1200, 1996.)

**FIGURE 2.24**

Vector diagram of a reactance modulator producing FM. Note: $V_s = V_{C2}$. (After Seymour, K., Frequency modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1188–1200, 1996.)

Using vector diagrams, we can analyze the phase relationship of the reactance modulator. Referring to Figure 2.22, the resistance of R_1 is typically high compared to the capacitive reactance of C_2 . The R_1C_2 circuit is then resistive. Because this circuit is resistive, the current, I_{AB} , that flows through it is in phase with the voltage V_{AB} . Voltage V_{AB} is also across R_1C_2 (or Z_1Z_2 in Figure 2.23). This is true because current and voltage tend to be in phase in a resistive network. However, voltage V_{C2} , which is across C_2 , is out of phase with I_{AB} . This is because the voltage that is across a capacitor lags behind its current by 90° (see Figure 2.24).

2.6.5.2 VCO Direct-FM Modulator

One of the more common direct-FM modulation techniques uses an analog *voltage-controlled oscillator* (VCO) in a PLL arrangement [7]. In this configuration, shown in Figure 2.25, a VCO produces the desired carrier frequency that is, in turn, modulated by applying the input signal to the VCO input via a varactor diode. The varactor is used to vary the capacitance of an oscillator tank circuit. Therefore, the varactor behaves as a variable capacitor whose capacitance changes as the signal voltage across it changes. As the input capacitance of the VCO is varied by the varactor, the output frequency of the VCO is shifted, which produces a direct-FM modulated signal.

2.6.6 FM Power Amplifiers

Nearly all high-power FM transmitters employ cavity designs. The $\frac{1}{4}$ -wavelength cavity is the most common. The design is simple and straightforward. A number of variations can be found in different transmitters, but the underlying theory of operation is the same. The goal of any cavity amplifier is to simulate a resonant tank circuit at the operating frequency and provide a means to couple the energy in the cavity to the transmission line.

A typical $\frac{1}{4}$ -wave cavity is shown in Figure 2.26. The plate of the tube connects directly to the inner section (tube) of the plate-blocking capacitor.

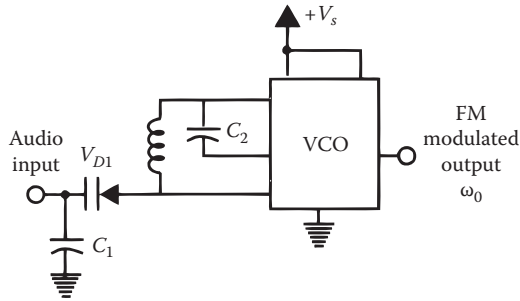


FIGURE 2.25 Voltage-controlled direct-FM modulator. (After Seymour, K., Frequency modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1188–1200, 1996.)

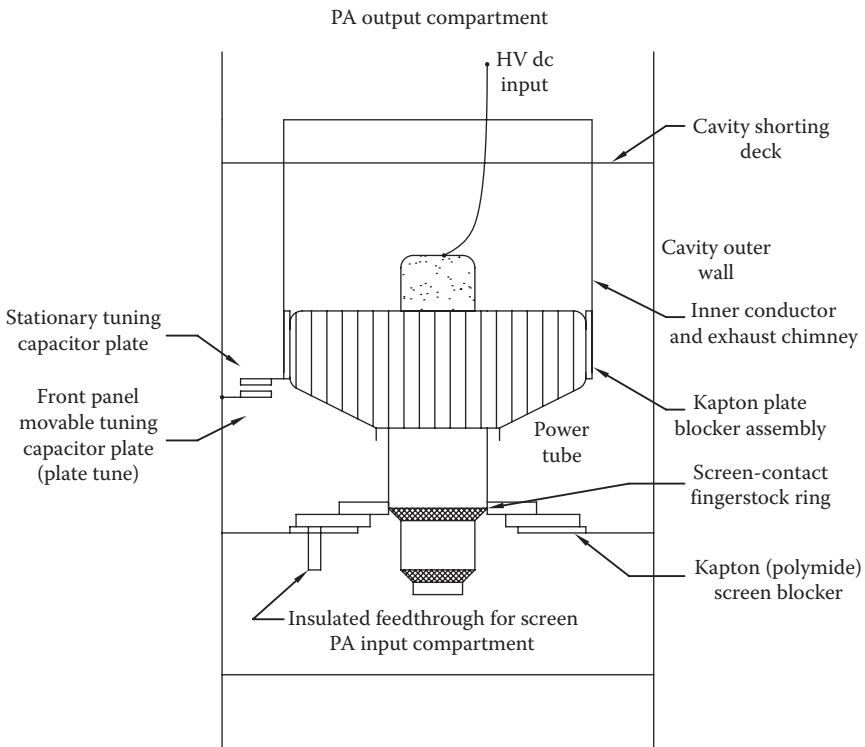


FIGURE 2.26 Mechanical layout of a common type of $\frac{1}{4}$ -wave PA cavity for FM service.

The blocking capacitor can be formed in one of several ways. In at least one design, it is made by wrapping the outside surface of the inner tube conductor with multiple layers of insulating film. The exhaust chimney/inner conductor forms the other element of the blocking capacitor. The cavity walls form the outer conductor of the $\frac{1}{4}$ -wave transmission line circuit. The dc plate voltage is applied to the PA tube by a cable routed inside the exhaust chimney and inner tube conductor. In this design, the screen-contact fingerstock ring mounts on a metal plate that is insulated from the grounded cavity deck by a blocking capacitor. This hardware makes up the screen-blocker assembly. The dc screen voltage feeds to the fingerstock ring from underneath the cavity deck using an insulated feedthrough.

Some transmitters that employ the $\frac{1}{4}$ -wave cavity design use a grounded-screen configuration in which the screen-contact fingerstock ring is connected directly to the grounded cavity deck. The PA cathode then operates at below ground potential (in other words, at a negative voltage), establishing the required screen voltage for the tube.

Coarse-tuning of the cavity is accomplished by adjusting the cavity length. The top of the cavity (the *cavity shorting deck*) is fastened by screws or clamps and can be raised or lowered to set the length of the assembly for the particular operating frequency. Fine-tuning is accomplished by a variable-capacity plate-tuning control built into the cavity. In the example shown in Figure 2.26, one plate of this capacitor—the stationary plate—is fastened to the inner conductor just above the plate-blocking capacitor. The movable tuning plate is fastened to the cavity box, the outer conductor, and is linked mechanically to the front-panel tuning control. This capacity shunts the inner conductor to the outer conductor and varies the electrical length and resonant frequency of the cavity.

2.7 Pulse Modulation

The growth of digital processing and communications has led to the development of modulation systems tailor-made for high-speed, spectrum-efficient transmission [8]. In a *pulse modulation* system, the unmodulated carrier usually consists of a series of recurrent pulses. Information is conveyed by modulating some parameter of the pulses, such as amplitude, duration, time of occurrence, or shape. Pulse modulation is based on the *sampling principle*, which states that a message waveform with a spectrum of finite width can be recovered from a set of discrete samples if the sampling rate is higher than twice the highest sampled frequency (the Nyquist criteria). The samples of the input signal are used to modulate some characteristic of the carrier pulses.

2.7.1 Digital Modulation Systems

Because of the nature of digital signals (on or off), it follows that the amplitude of the signal in a pulse modulation system should be one of two heights (present or absent/positive or negative) for maximum efficiency. Noise immunity is a significant advantage of such a system. It is necessary for the receiving system to detect only the presence or absence (or polarity) of each transmitted pulse to allow complete reconstruction of the original intelligence. The pulse shape and noise level have minimal effect (to a point). Furthermore, if the waveform is to be transmitted over long distances, it is possible to regenerate the original signal exactly for retransmission to the next relay point. This feature is in striking contrast to analog modulation systems in which each modulation step introduces some amount of noise and signal corruption.

In any practical digital data system, some corruption of the intelligence is likely to occur over a sufficiently large span of time. Data encoding and manipulation schemes have been developed to detect and correct or conceal such errors. The addition of error-correction features comes at the expense of increased system overhead and a resulting lower intelligence throughput.

2.7.2 Pulse Amplitude Modulation

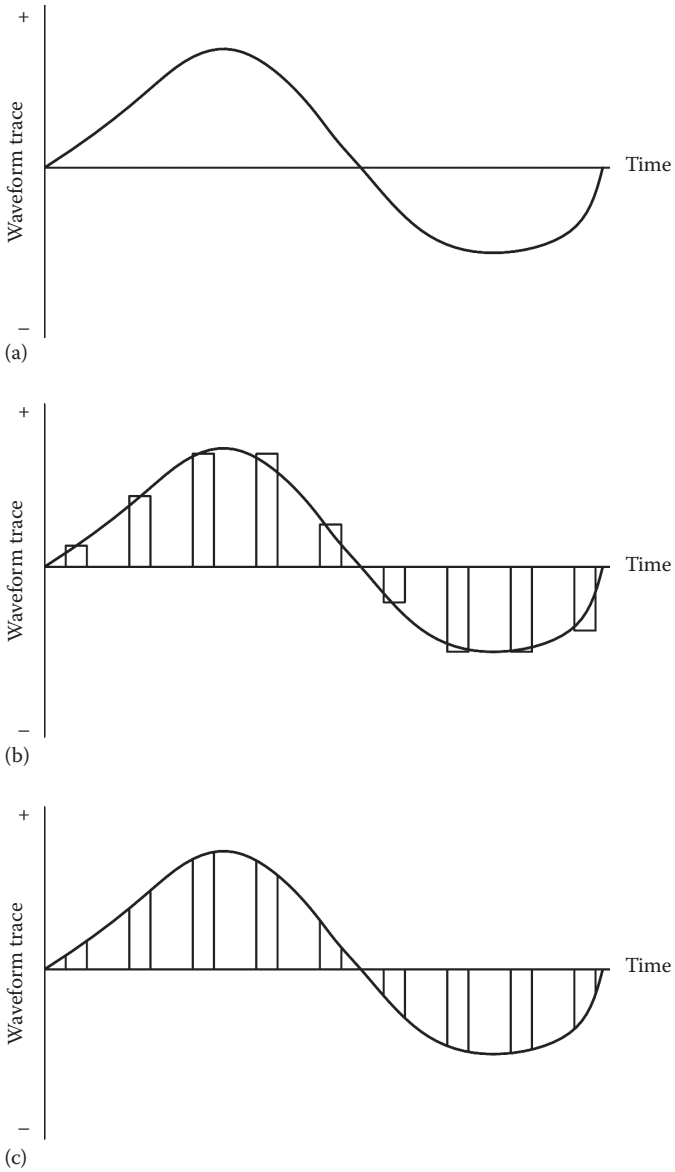
Pulse amplitude modulation (PAM) is one of the simplest forms of data modulation. PAM departs from conventional modulation systems in that the carrier exists as a series of pulses, rather than as a continuous waveform. The amplitude of the pulse train is modified in accordance with the applied modulating signal to convey intelligence, as illustrated in Figure 2.27.

There are two primary forms of PAM sampling:

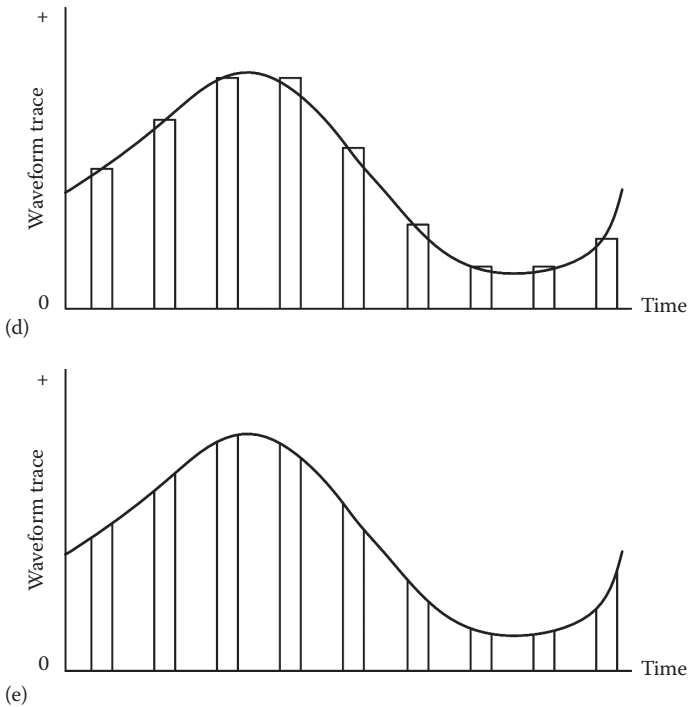
- *Natural sampling* (or *top sampling*), where the modulated pulses follow the amplitude variation of the sampled time function during the sampling interval.
- *Instantaneous sampling* (or *square-topped sampling*), where the amplitude of the pulses is determined by the instantaneous value of the sampled time function corresponding to a single instant of the sampling interval. This “single instant” may be the center or edge of the sampling interval.

There are two common methods of generating a PAM signal:

- Variation of the amplitude of a pulse sequence about a fixed nonzero value (or *pedestal*). This approach constitutes DSAM.
- Double-polarity modulated pulses with no pedestal. This approach constitutes double-sideband suppressed carrier modulation.

**FIGURE 2.27**

PAM waveforms: (a) modulating signal; (b) square-topped sampling, bipolar pulse train; (c) topped sampling, bipolar pulse train;

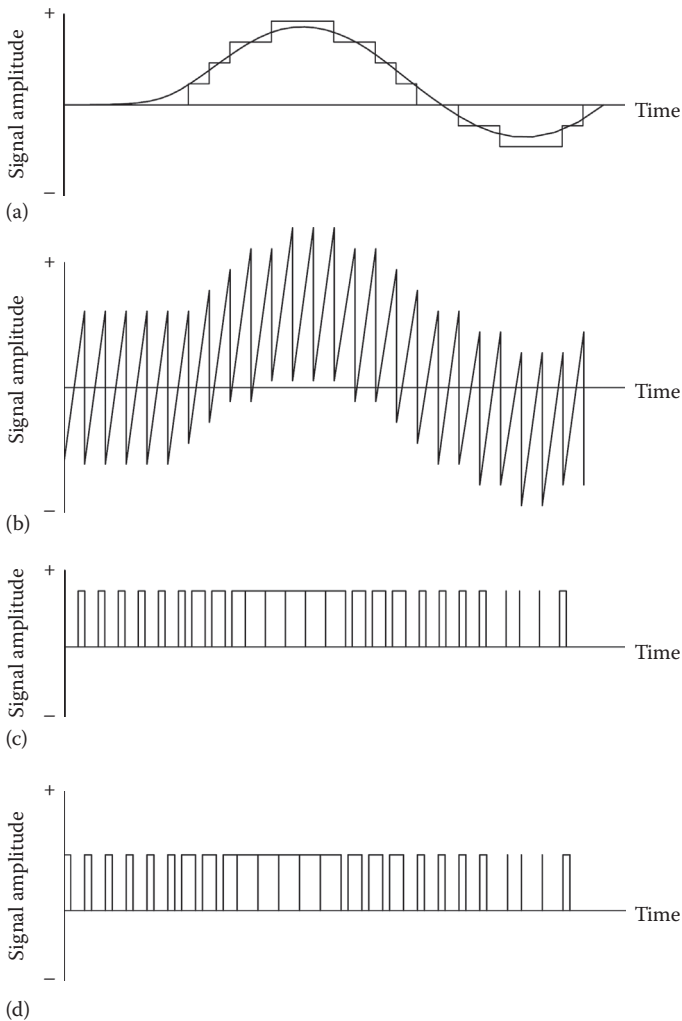
**FIGURE 2.27 (continued)**

(d) square-topped sampling, unipolar pulse train; and (e) top sampling, unipolar pulse train.

2.7.3 Pulse Time Modulation

A number of modulating schemes have been developed to take advantage of the noise immunity afforded by a constant amplitude modulating system. *Pulse time modulation* (PTM) is one of those systems. In a PTM system, instantaneous samples of the intelligence are used to vary the time of occurrence of some parameter of the pulsed carrier. Subsets of the PTM process include the following:

- *Pulse duration modulation* (PDM), where the time of occurrence of either the leading or trailing edge of each pulse (or both pulses) is varied from its unmodulated position by samples of the input modulating waveform. PDM also may be described as *pulse length* or *pulse width modulation* (PWM).
- *Pulse position modulation* (PPM), where samples of the modulating input signal are used to vary the position in time of pulses, relative to the unmodulated waveform. Several types of PTM waveforms are shown in Figure 2.28.

**FIGURE 2.28**

PTM waveforms: (a) modulating signal and sample-and-hold (S/H) waveforms, (b) sawtooth waveform added to S/H, (c) leading-edge PTM, and (d) trailing-edge PTM.

- *Pulse frequency modulation (PFM)*, where samples of the input signal are used to modulate the frequency of a series of carrier pulses. The PFM process is illustrated in Figure 2.29.

It should be emphasized that all of the pulse modulation systems discussed thus far may be used with both analog and digital input signals. Conversion is required for either signal into a form that can be accepted by the pulse modulator.

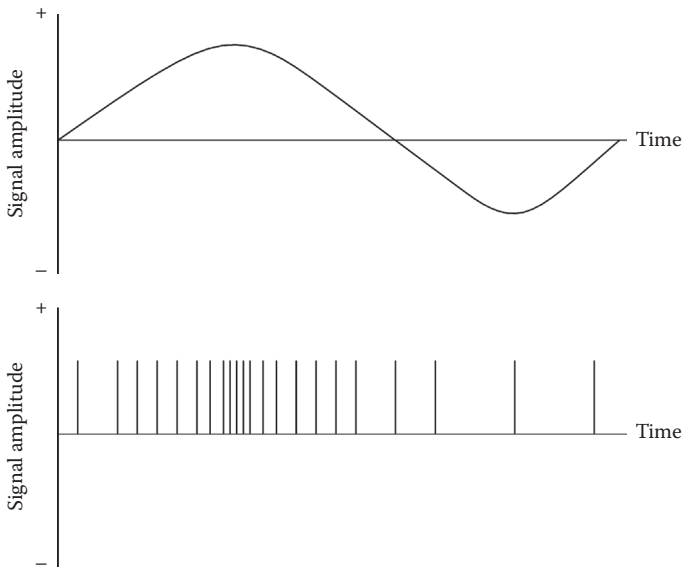


FIGURE 2.29
Pulse frequency modulation.

2.7.4 Pulse Code Modulation

The pulse modulation systems discussed previously are *unencoded* systems. *Pulse code modulation* (PCM) is a scheme wherein the input signal is *quantized* into discrete steps and then sampled at regular intervals (as in conventional pulse modulation). In the *quantization* process, the input signal is sampled to produce a code representing the instantaneous value of the input within a predetermined range of values. Figure 2.30 illustrates the concept. Only certain discrete levels are allowed in the quantization process. The code is then transmitted over the communications system as a pattern of pulses.

Quantization inherently introduces an initial error in the amplitude of the samples taken. This *quantization error* is reduced as the number of quantization steps is increased. In system design, trade-offs must be made regarding low quantization error, hardware complexity, and occupied bandwidth. The greater the number of quantization steps, the wider the bandwidth required to transmit the intelligence or, in the case of some signal sources, the slower the intelligence must be transmitted.

In the classic design of a PCM encoder, the quantization steps are equal. The quantization error (or *quantization noise*) usually can be reduced, however, through the use of nonuniform spacing of levels. Smaller quantization steps are provided for weaker signals, and larger steps are provided near the peak of large signals. Quantization noise is reduced by providing an encoder that is matched to the *level distribution* (*probability density*) of the input signal.

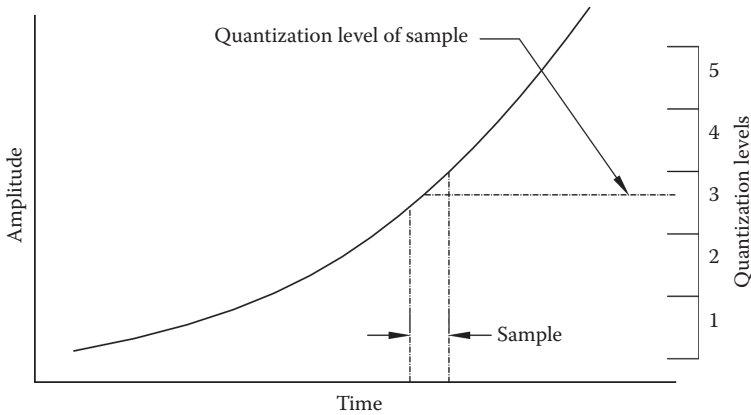


FIGURE 2.30
Quantization process.

Nonuniform quantization typically is realized in an encoder through processing of the input (analog) signal to compress it to match the desired nonuniformity. After compression, the signal is fed to a uniform quantization stage.

2.7.5 Delta Modulation

Delta modulation (DM) is a coding system that measures changes in the direction of the input waveform, rather than the instantaneous value of the wave itself. Figure 2.31 illustrates the concept. The clock rate is assumed to be constant. Transmitted pulses from the pulse generator are positive if the signal is changing in a positive direction; they are negative if the signal is changing in a negative direction.

As with the PCM encoding system, quantization noise is a parameter of concern for DM. Quantization noise can be reduced by increasing the sampling frequency (the pulse generator frequency). The DM system has no fixed maximum (or minimum) signal amplitude. The limiting factor is the slope of the sampled signal, which must not change by more than one level or step during each pulse interval.

2.7.6 Digital Coding Systems

A number of methods exist to transmit digital signals over long distances in analog transmission channels. Some of the more common systems include the following:

- *Binary on-off keying (BOOK)*, a method by which a high-frequency sinusoidal signal is switched on and off corresponding to 1 and 0 (on and off) periods in the input digital data stream. In practice, the transmitted sinusoidal waveform does not start or stop abruptly, but follows a predefined ramp up or down.

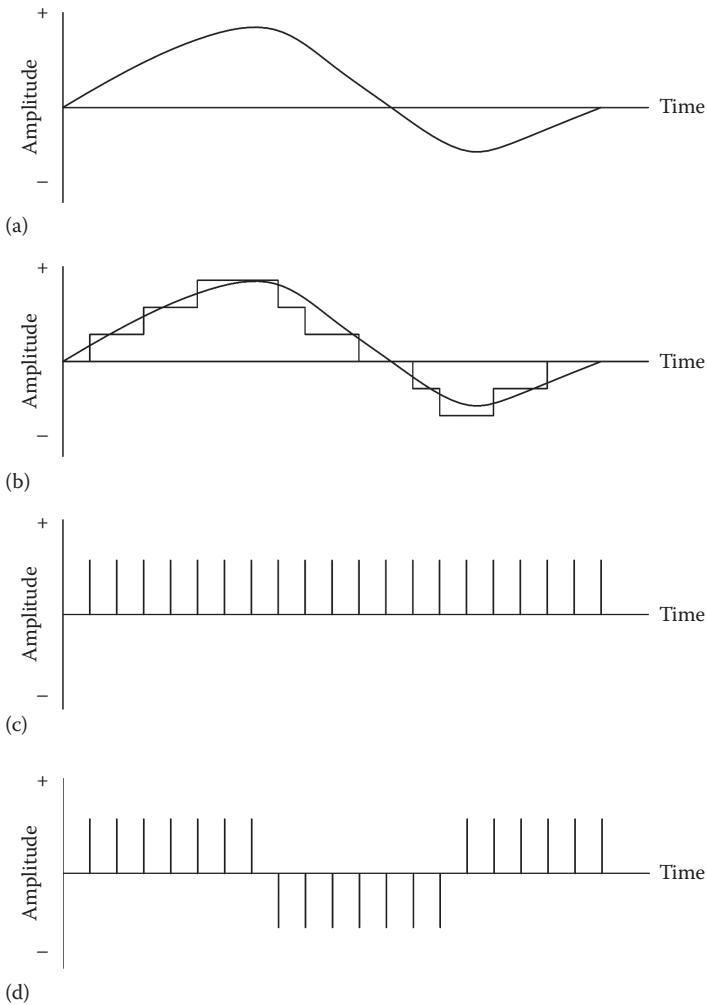


FIGURE 2.31 Delta modulation waveforms: (a) modulating signal, (b) quantized modulating signal, (c) pulse train, and (d) resulting delta modulation waveform.

- Binary frequency-shift keying (BFSK), a modulation method in which a continuous wave is transmitted that is shifted between two frequencies, representing 1s and 0s in the input data stream. The BFSK signal may be generated by switching between two oscillators (set to different operating frequencies) or by applying a binary baseband signal to the input of a VCO. The transmitted signals often are referred to as a mark (binary digit 1) or a space (binary digit 0). Figure 2.32 illustrates the transmitted waveform of a BFSK system.

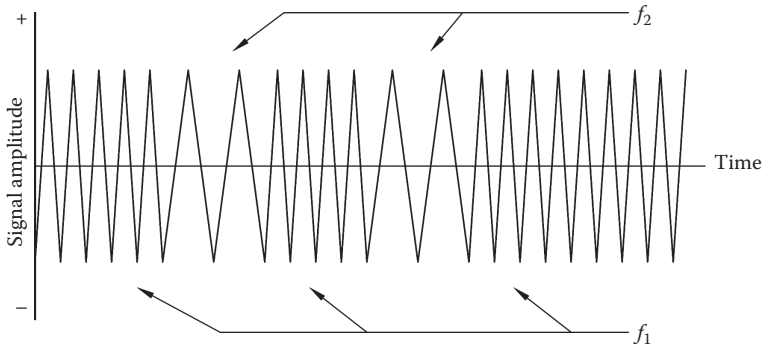


FIGURE 2.32
Binary FSK waveform.

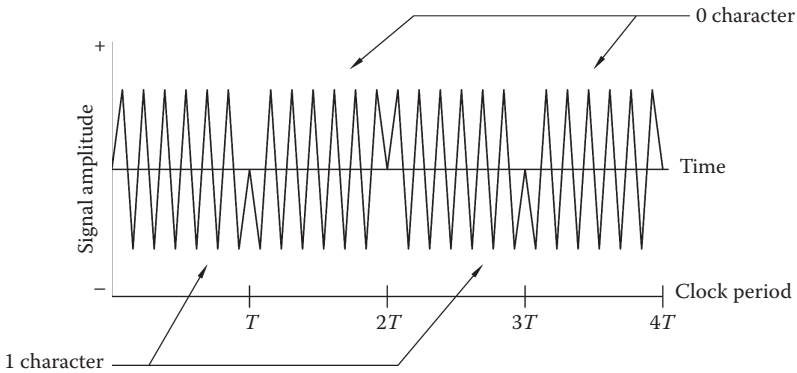


FIGURE 2.33
Binary PSK waveform.

- Binary phase-shift keying (BPSK), a modulating method in which the phase of the transmitted wave is shifted 180° in synchronism with the input digital signal. The phase of the RF carrier is shifted by $+\pi/2$ or $-\pi/2$ rad, depending upon whether the data bit is a 0 or a 1. Figure 2.33 shows the BPSK transmitted waveform.
- Quadrature phase-shift keying (QPSK), a modulation scheme similar to BPSK except that quaternary modulation is employed, rather than binary modulation. QPSK requires half the bandwidth of BPSK for the same transmitted data rate.

2.7.7 Baseband Digital Pulse Modulation

After the input samples have been quantized, they are transmitted through a channel, received, and converted back to their approximate original form [8]. The format (modulation scheme) applied to the quantized samples is determined by a number of factors, not the least of which

is the channel through which the signal passes. A number of different formats are possible and practical.

Several common digital modulation formats are shown in Figure 2.34. The first (Figure 2.34a) is referred to as *non-return-to-zero* (NRZ) polar because the waveform does not return to zero during each signaling interval, but switches from $+V$ to $-V$, or vice versa, at the end of each signaling interval (NRZ unipolar uses the levels V and 0). On the other hand, the

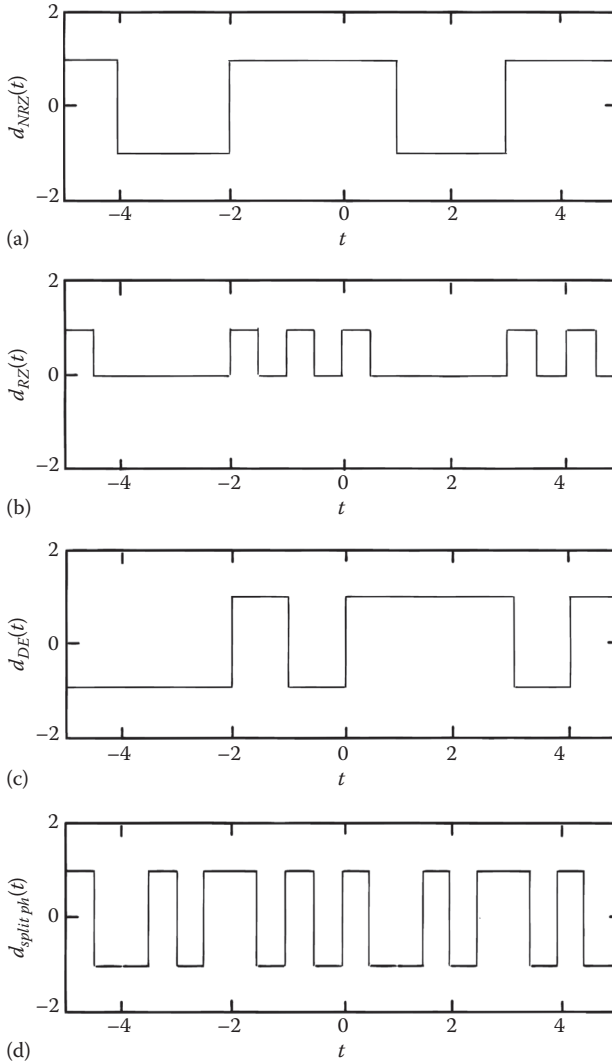


FIGURE 2.34 Various baseband modulation formats: (a) non-return-to zero, (b) unipolar return-to-zero, (c) differential encoded (NRZ-mark), and (d) split phase. (After Ziemer, R.E., Pulse modulation, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1201–1212, 1996.)

unipolar return-to-zero (RZ) format, shown in Figure 2.34b, returns to zero in each signaling interval. Because bandwidth is inversely proportional to pulse duration, it is apparent that RZ requires twice the bandwidth that NRZ does. Also, RZ has a nonzero dc component, whereas NRZ does not necessarily have a nonzero component (unless there are more 1s than 0s or vice versa). An advantage of RZ over NRZ is that a pulse transition is guaranteed in each signaling interval, whereas this is not the case for NRZ. Thus, in cases where there are long strings of 1s or 0s, it may be difficult to synchronize the receiver to the start and stop times of each pulse in NRZ-based systems.

A very important modulation format from the standpoint of synchronization considerations is *NRZ-mark*, also known as *differential encoding*, where an initial reference bit is chosen and a subsequent 1 is encoded as a change from the reference and a 0 is encoded as no change. After the initial reference bit, the current bit serves as a reference for the next bit, and so on. An example of this modulation format is shown in Figure 2.34c.

Manchester is another baseband data modulation format that guarantees a transition in each signaling interval and does not have a dc component. Also known as *biphase* or *split phase*, this scheme is illustrated in Figure 2.34d. The format is produced by *ORing* the data clock with an NRZ-formatted signal. The result is a + to – transition for a logic 1, and a – to + zero crossing for a logic 0.

A number of other data formats have been proposed and implemented, but further discussion is beyond the scope of this chapter.

2.8 Spread Spectrum Systems

As the name implies, a *spread spectrum* system requires a frequency range substantially greater than the basic information-bearing signal. Spread spectrum systems have some or all of the following properties:

- Low interference to other communications systems
- Ability to reject high levels of external interference
- Immunity to jamming by hostile forces
- Provision for secure communications paths
- Operability over multiple RF paths

Spread spectrum systems operate with an entirely different set of requirements than transmission systems discussed previously. Conventional modulation methods are designed to provide for the easiest possible reception and demodulation of the transmitted intelligence. The goals of spread spectrum

systems, on the other hand, are secure and reliable communications that cannot be intercepted by unauthorized persons. The most common modulating and encoding techniques used in spread spectrum communications include the following:

- *Frequency hopping*, where a random or *pseudorandom* number (PN) sequence is used to change the carrier frequency of the transmitter. This approach has two basic variations: *slow frequency hopping*, where the hopping rate is smaller than the data rate, and *fast frequency hopping*, where the hopping rate is larger than the data rate. In a fast frequency-hopping system, the transmission of a single piece of data occupies more than one frequency. Frequency-hopping systems permit multiple-access capability to a given band of frequencies because each transmitted signal occupies only a fraction of the total transmitted bandwidth.
- *Time hopping*, where a PN sequence is used to switch the position of a message-carrying pulse within a series of frames.
- *Message corruption*, where a PN sequence is added to the message before modulation.
- *Chirp spread spectrum*, where linear FM of the main carrier is used to spread the transmitted spectrum. This technique is commonly used in radar and also has been applied to communications systems.

In a spread spectrum system, the signal power is divided over a large bandwidth. The signal, therefore, has a small average power in any single narrowband slot. This means that a spread spectrum system can share a given frequency band with one or more narrowband systems. Furthermore, because of the low energy in any particular band, detection or interception of the transmission is difficult.

References

1. Tate, J. P. and P. F. Mead, Crystal oscillators, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 185–199, 1996.
2. Frerking, M. E., *Crystal Oscillator Design and Temperature Compensation*, Van Nostrand Reinhold, New York, 1978.
3. Dye, D. W., The piezo-electric quartz resonator and its equivalent electrical circuit, *Proceedings of the Physical Society*, 38, 399–457, 1926.
4. Douglass, B. G., Thermal noise and other circuit noise, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 30–36, 1996.
5. Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, p. 468, 1947.
6. Kubichek, R., Amplitude modulation, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1175–1187, 1996.

7. Seymour, K., Frequency modulation, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1188–1200, 1996.
8. Ziemer, R. E., Pulse modulation, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1201–1212, 1996.

Bibliography

- Benson, K. B. and J. C. Whitaker (eds.), *Television Engineering Handbook*, McGraw-Hill, New York, 1986.
- Benson, K. B. and J. C. Whitaker, *Television and Audio Handbook for Technicians and Engineers*, McGraw-Hill, New York, 1989.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, 1991.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 2nd edn., McGraw-Hill, New York, 1982.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.
- Golio, M., *RF and Microwave Applications and Systems*, CRC Press, Boca Raton, FL, 2008.
- Hulick, T. P., Using tetrodes for high power UHF, in *Proceedings of the Society of Broadcast Engineers*, Vol. 4, SBE, Indianapolis, IN, pp. 52–57, 1989.
- Jordan, E. C., *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 7th edn., Howard W. Sams, Indianapolis, IN, 1985.
- Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.
- Mendenhall, G. N., Fine tuning FM final stages, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1987.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1991.
- Whitaker, J. C., *Maintaining Electronic Systems*, CRC Press, Boca Raton, FL, 1992.

3

Vacuum Tube Principles

3.1 Introduction

A power grid tube is a device using the flow of free electrons in a vacuum to produce useful work [1]. It has an emitting surface (the cathode), one or more grids that control the flow of electrons, and an element that collects the electrons (the anode). Power tubes can be separated into groups according to the number of electrodes (grids) they contain. The physical shape and location of the grids relative to the plate and cathode are the main factors that determine the *amplification factor* (μ) and other parameters of the device. The physical size and types of material used to construct the individual elements determine the power capability of the tube. A wide variety of tube designs are available to commercial and industrial users. By far the most common are triodes and tetrodes.

3.2 Characteristics of Electrons

Electrons are minute, negatively charged particles that are constituents of all matter [1]. They have a mass of 9×10^{-28} g (1/1840 that of a hydrogen atom) and a charge of 1.59×10^{-19} C. Electrons are always identical, irrespective of their source. Atoms are composed of one or more such electrons associated with a much heavier nucleus, which has a positive charge equal to the number of the negatively charged electrons contained in the atom; an atom with a full quota of electrons is electrically neutral. The differences in chemical elements arise from differences in the nucleus and in the number of associated electrons.

Free electrons can be produced in a number of ways. *Thermonic emission* is the method normally employed in vacuum tubes. The principle of thermonic emission states that if a solid body is heated sufficiently, some of the electrons that it contains will escape from the surface into the surrounding space. Electrons also are ejected from solid materials as a result of the impact of rapidly moving electrons or ions. This phenomenon is referred to

as *secondary electron emission*, because it is necessary to have a primary source of electrons (or ions) before the secondary emission can be obtained. Finally, it is possible to pull electrons directly out of solid substances by an intense electrostatic field at the surface of the material.

Positive ions represent atoms or molecules that have lost one or more electrons and so have become charged bodies having the weight of the atom or molecule concerned, and a positive charge equal to the negative charge of the lost electrons. Unlike electrons, positive ions are not all alike and may differ in charge, weight, or both. They are much heavier than electrons and resemble the molecule or atom from which they are derived. Ions are designated according to their origin, such as mercury ions or hydrogen ions.

3.2.1 Electron Optics

Electrons and ions are charged particles and, as such, have forces exerted upon them by an electrostatic field in the same way as other charged bodies [1]. Electrons, being negatively charged, tend to travel toward the positive or anode electrode, while the positively charged ions travel in the opposite direction (toward the negative or cathode electrode). The force F exerted upon a charged particle by an electrostatic field is proportional to the product of the charge e of the particle and the voltage gradient G of the electrostatic field [1]:

$$F = G \times e \times 10^7 \quad (3.1)$$

where

F is the force in dyn

G is the voltage gradient in V/cm

e is the charge in C

This force upon the ion or electron is exerted in the direction of the electrostatic flux lines at the point where the charge is located. The force acts toward or away from the positive terminal, depending upon whether a negative or positive charge, respectively, is involved.

The force that the field exerts on the charged particle causes an acceleration in the direction of the field at a rate that can be calculated by the laws of mechanics where the velocity does not approach that of light:

$$A = \frac{F}{m} \quad (3.2)$$

where

A is the acceleration in cm/s

F is the force in dyn

m is the mass in g

The velocity an electron or ion acquires in being acted upon by an electrostatic field can be expressed in terms of the voltage through which the electron

(or ion) has fallen in acquiring the velocity. For velocities well below the speed of light, the relationship between velocity and the acceleration voltage is

$$v = \sqrt{\frac{2 \times V \times e \times 10^7}{m}} \quad (3.3)$$

where

v is the velocity in cm/s corresponding to V

V is the accelerating voltage

e is the charge in C

m is the mass in g

Electrons and ions move at great velocities in even moderate-strength fields. For example, an electron dropping through a potential difference of 2500 V will achieve a velocity of approximately one-tenth the speed of light.

Electron optics, as discussed in this section, relies on the principles of classical physics. While modern tube design uses computer simulation almost exclusively, the preceding information is nonetheless valid and provides a basis for the understanding of electron motion within a vacuum tube device.

3.2.2 Magnetic Field Effects

An electron in motion represents an electric current of magnitude ev , where e is the magnitude of the charge on the electron and v is its velocity [1]. A magnetic field accordingly exerts a force on a moving electron in a manner similar to the force it exerts on an electric current in a wire. The magnitude of the force is proportional to the product of the equivalent current ev represented by the moving electron and the strength of the component of the magnetic field in a direction at right angles to the motion of the electron. The resulting force is, then, in a direction at right angles both to the direction of motion of the electron and to the component of the magnetic field that is producing the force. As a result, an electron entering a magnetic field with a high velocity will follow a curved path. Because the acceleration of the electron that the force of the magnetic field produces is always at right angles to the direction in which the electron is traveling, an electron moving in a uniform magnetic field will follow a circular path. The radius of this circle is determined by the strength of the magnetic field and the speed of the electron moving through the field.

When an electron is subjected to the simultaneous action of both electric and magnetic fields, the resulting force acting on the electron is the vector sum of the force resulting from the electric field and the force resulting from the magnetic field, each considered separately.

Magnetic fields are not used for conventional power grid tubes. Microwave power tubes, on the other hand, use magnetic fields to confine and focus the electron stream.

3.2.3 Thermal Emission from Metals

Thermonic emission is the phenomenon of an electric current leaving the surface of a material as the result of thermal activation [1]. Electrons with sufficient thermal energy to overcome the surface-potential barrier escape from the surface of the material. This thermally emitted electron current increases with temperature because more electrons have sufficient energy to leave the material.

The number of electrons released per unit area of an emitting surface is related to the absolute temperature of the emitting material and a quantity b that is a measure of the work an electron must perform in escaping through the surface, according to the equation [1]

$$I = AT^2\epsilon^{-b/T} \quad (3.4)$$

where

T is the absolute temperature of the emitting material

b is the work an electron must perform in escaping the emitter surface

I is the electron current in A/cm²

A is a constant (value varies with the type of emitter)

The exponential term in the equation accounts for most of the variation in emission with temperature. The temperature at which the electron current becomes appreciable is accordingly determined almost solely by the quantity b . Figure 3.1 plots the emission resulting from a cathode operated at various temperatures.

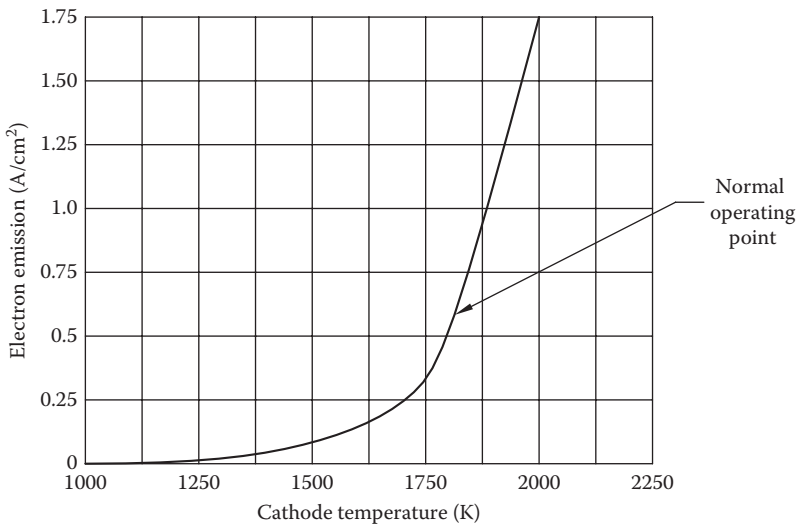
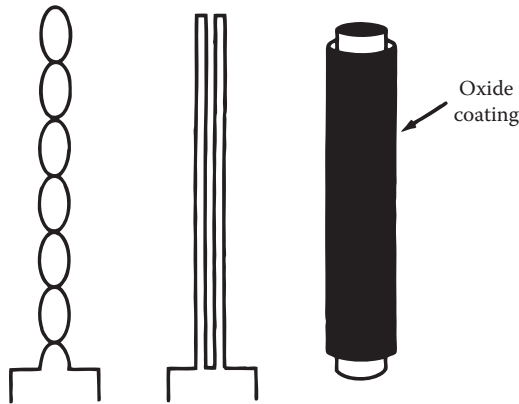


FIGURE 3.1

Variation of electron emission as a function of absolute temperature for a thoriated-tungsten emitter. (After Terman, F.E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.)

**FIGURE 3.2**

Common types of heater and cathode structures. (After Ferris, C.D., *Electron tube fundamentals*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 295–305, 1996.)

Thermal electron emission can be increased by applying an electric field to the cathode [2]. This field lowers the surface-potential barrier, enabling more electrons to escape. This field-assisted emission is known as the *Schottky effect*.

Figure 3.2 illustrates common heater-cathode structures for power tubes.

3.2.4 Secondary Emission

Almost all metals and some insulators will emit low-energy electrons (secondary electrons) when bombarded by other energetic electrons [1]. The number of secondary electrons emitted per primary electron is determined by the velocity of the primary bombarding electrons and the nature and condition of the material composing the surface being bombarded. Figure 3.3 illustrates a typical relationship for two types of surfaces. As shown in the figure, no secondary electrons are produced when the primary velocity is low. However, with increasing potential (and consequently higher velocity), the ratio of secondary to primary electrons increases, reaching a maximum and then decreasing. With pure metal surfaces, the maximum ratio of secondary to primary electrons ranges from less than 1 to approximately 3. Some complex surfaces based on alkali metal compounds yield ratios of secondary to primary electrons as high as 5–10.

The majority of secondary electrons emitted from a conductive surface have relatively low velocity. However, a few secondary electrons usually are emitted with a velocity nearly equal to the velocity of the bombarding primary electrons.

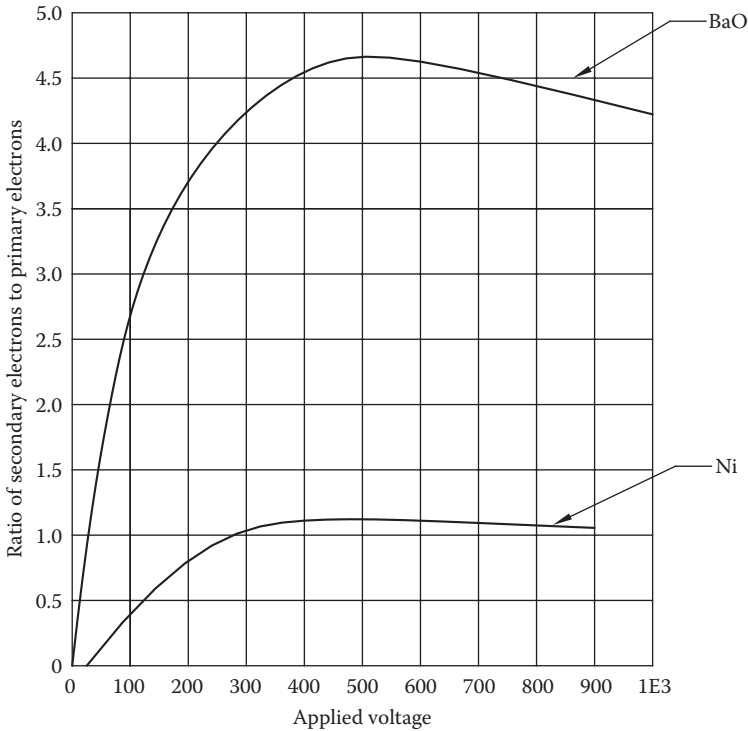


FIGURE 3.3

Ratio of secondary emission current to primary current as a function of primary electron velocity. (After Terman, F.E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.)

For insulators, the ratio of secondary to primary electrons as a function of primary electron potential follows along the same lines as for metals. The net potential of the insulating surface being bombarded is affected by the bombardment. If the ratio of secondary to primary current is less than unity, the insulator acquires a net negative charge because more electrons arrive than depart. This causes the insulator to become more negative and, finally, to repel most of the primary electrons, resulting in a blocking action. In the opposite extreme, when the ratio of secondary to primary electrons exceeds unity, the insulating surface loses electrons through secondary emission faster than they arrive; the surface becomes increasingly positive. This action continues until the surface is sufficiently positive that the ratio of secondary to primary electrons decreases to unity as a result of the increase in the velocity of the bombarding electrons, or until the surface is sufficiently positive that it attracts back into itself a significant number of secondary electrons. This process makes the number of electrons gained from all sources equal to the number of secondary electrons emitted.

3.3 Types of Vacuum Tubes

Electrons are of no value in a vacuum tube unless they can be put to work [3]. A tube is, therefore, designed with the elements necessary to utilize electrons as well as those required to produce them. These elements consist of a cathode and one or more supplementary electrodes. The electrodes are enclosed in an evacuated envelope having the necessary connections brought out through airtight seals. The air is removed from the envelope to allow free movement of the electrons and to prevent damage to the emitting surface of the cathode.

3.3.1 Diode

A diode is a two-electrode vacuum tube containing a cathode, which emits electrons by thermionic emission, surrounded by an anode (or plate) [1] (see Figure 3.4). Such a tube is inherently a rectifier because when the anode is positive, it attracts electrons; current, therefore, passes through the tube. When the anode is negative, it repels the electrons and no current flows.

The typical relationship between anode voltage and current flowing to the positive anode is shown in Figure 3.5. When the anode voltage is sufficiently high, electrons are drawn from the cathode as rapidly as they are emitted. The anode current is then limited by the electron emission of the cathode and, therefore, depends upon cathode temperature rather than anode voltage.

At low anode voltages, however, plate current is less than the emission of which the cathode is capable. This occurs because the number of electrons in transit between the cathode and plate at any instant cannot exceed the number that will produce a negative *space charge*, which completely neutralizes the attraction of the positive plate upon the electrons just leaving the cathode. All electrons in excess of the number necessary to neutralize the effects of the plate voltage are repelled into the cathode by the negative space

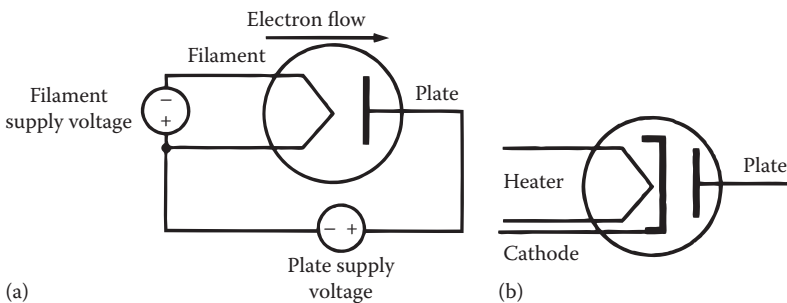


FIGURE 3.4

Vacuum diode: (a) directly heated cathode and (b) indirectly heated cathode. (After Ferris, C.D., *Electron tube fundamentals*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 295–305, 1996.)

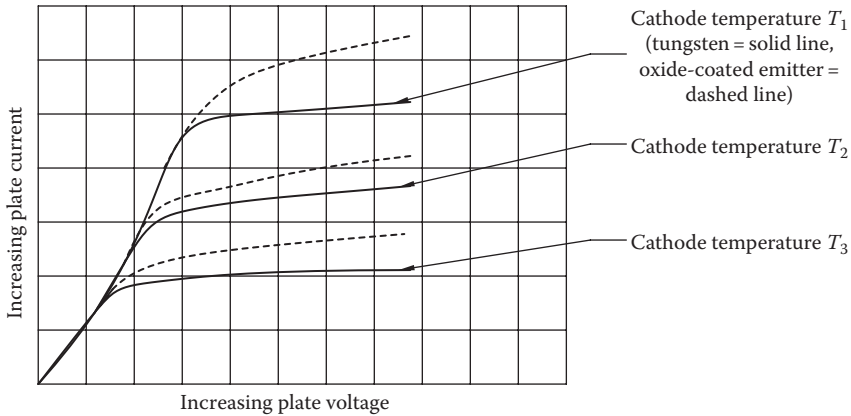


FIGURE 3.5

Anode current as a function of anode voltage in a two-electrode tube for three cathode temperatures. (After Terman, F.E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947)

charge of the electrons in transit; this situation applies irrespective of how many excess electrons the cathode emits. When the plate current is limited in this way by space charge, plate current is determined by plate potential and is substantially independent of the electron emission of the cathode.

Detailed examination of the space-charge situation will reveal that the negative charge of the electrons in transit between the cathode and the plate is sufficient to give the space in the immediate vicinity of the cathode a slight negative potential with respect to the cathode. The electrons emitted from the cathode are projected out into this field with varying emission velocities. The negative field next to the cathode causes the emitted electrons to slow as they move away from the cathode, and those having a low velocity of emission are driven back into the cathode. Only those electrons having the highest velocities of emission will penetrate the negative field near the cathode and reach the region where they are drawn toward the positive plate. The remainder (those electrons having low emission velocities) will be brought to a stop by the negative field adjacent to the cathode and will fall back into the cathode.

The energy that is delivered to the tube by the source of anode voltage is first expended in accelerating the electrons traveling from the cathode to the anode; it is converted into kinetic energy. When these swiftly moving electrons strike the anode, this kinetic energy is then transformed into heat as a result of the impact and appears at the anode in the form of heat that must be radiated to the walls of the tube.

3.3.1.1 Applications

Diodes are commonly used as rectifiers in ac-powered systems to convert the ac supply voltage to the dc voltage(s) necessary for other devices in the unit [3].

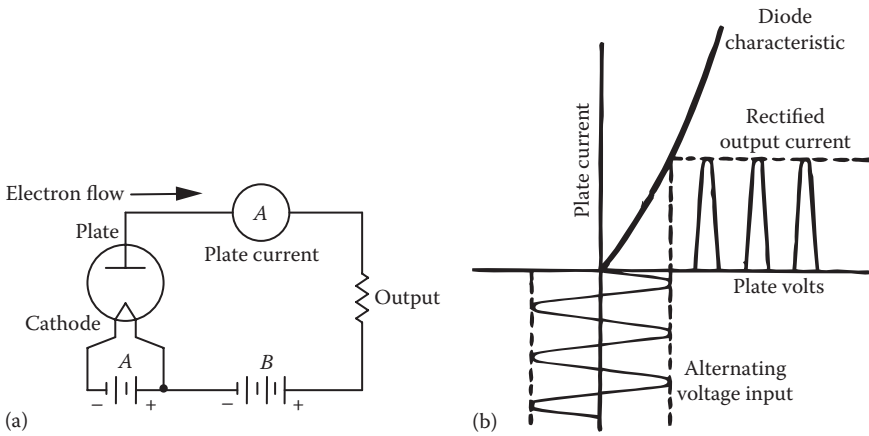


FIGURE 3.6 Basic diode circuit: (a) functioning circuit and (b) current characteristics of a rectifier circuit. (After *RCA Receiving Tube Manual*, RC-30, Radio Corporation of America, Camden, NJ, 1975.)

Figure 3.6a shows the basic diode circuit and Figure 3.6b illustrates the rectified output current produced by an alternating input voltage.

Rectifier tubes are commonly available in two primary configurations:

- One plate and one cathode for half-wave rectifier applications
- Two plates and one or two cathodes in the same device for full-wave rectifier applications

The basic function of a vacuum tube diode—to rectify an ac voltage—has been superseded by solid-state devices. An understanding of how the diode operates, however, is important in understanding the operation of triodes, tetrodes, and pentodes.

3.3.2 Triode

The power triode is a three-element device commonly used in a wide variety of RF generators. Triodes have three internal elements: the cathode, control grid, and plate. Most tubes are cylindrically symmetrical. The filament or cathode structure, the grid, and the anode are cylindrical in shape and are mounted with the axis of each cylinder along the center line of the tube, as illustrated in Figure 3.7.

The grid normally is operated at a negative potential with respect to the cathode and so attracts no electrons [1]. However, the extent to which it is negative affects the electrostatic field in the vicinity of the cathode and, therefore, controls the number of electrons that pass between the grid and the plate. The grid, in effect, functions as an imperfect electrostatic shield. It allows some, but not all, of the electrostatic flux from the anode

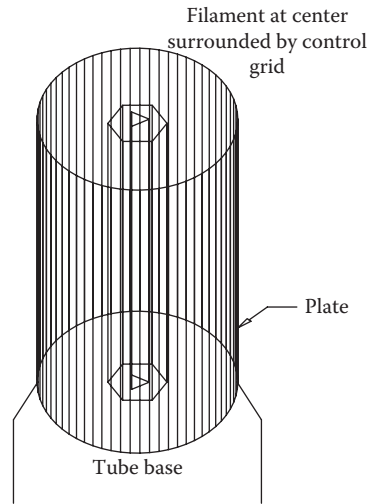


FIGURE 3.7
Mechanical configuration of a power triode.

to leak between its wires. The number of electrons that reach the anode in a triode tube under space-charge-limited conditions is determined almost solely by the electrostatic field near the cathode; the field in the rest of the interelectrode space has little effect. This phenomenon results because the electrons near the cathode are moving slowly compared with the electrons that have traveled some distance toward the plate. The result of this condition is that the volume density of electrons in proportion to the rate of flow is large near the cathode and low in the remainder of the interelectrode space. The total space charge of the electrons in transit toward the plate, therefore, consists almost solely of the electrons in the immediate vicinity of the cathode. After an electron has traveled beyond this region, it reaches the plate so quickly as to contribute to the space charge for only a brief additional time interval. The result is that the space current in a three-electrode vacuum tube is, for all practical purposes, determined by the electrostatic field that the combined action of the grid and plate potentials produces near the cathode.

When the grid structure is symmetrical, the field E at the surface of the cathode is proportional to the quantity

$$\frac{E_c + E_b}{\mu} \quad (3.5)$$

where

E_c is the control grid voltage (with respect to cathode)

E_b is the anode voltage (with respect to cathode)

μ is a constant determined by the geometry of the tube

The constant μ , the amplification factor, is independent of the grid and plate voltages. It is a measure of the relative effectiveness of grid and plate voltages in producing electrostatic fields at the surfaces of the cathode. Placement of the control grid relative to the cathode and plate determines the amplification factor. The μ , values of triodes generally range from 5 to 200. Key mathematical relationships include the following:

$$\mu = \frac{\Delta E_b}{\Delta E_{c1}} \quad (3.6)$$

$$R_p = \frac{\Delta E_b}{\Delta I_b} \quad (3.7)$$

$$S_m = \frac{\Delta I_b}{\Delta E_{c1}} \quad (3.8)$$

where

μ is the amplification factor (with plate current held constant)

R_p is the dynamic plate resistance

S_m is the transconductance (also may be denoted G_m)

E_b is the total instantaneous plate voltage

E_{c1} is the total instantaneous control grid voltage

I_b is the total instantaneous plate current

The total cathode current of an ideal triode can be determined from

$$I_k = \left\{ E_c + \frac{E_b}{\mu} \right\}^{3/2} \quad (3.9)$$

where

I_k is the cathode current

E_c is the grid voltage

E_b is the plate voltage

μ is the amplification factor

Figure 3.8a plots plate and grid current as a function of plate voltage at various grid voltages for a triode with a μ of 12 [4]. The tube, a 304TL, is a classic design and, while not used in new equipment, provides a common example of the relationship between the parameters plotted. Figure 3.8b plots the same parameters for a tube with a μ of 20. Observe how much more plate current at a given plate voltage can be obtained from the 304TL ($\mu = 12$) without driving the grid into the positive grid region. Note also how much more bias voltage is required for the 304TL to cut the plate current off at some given plate voltage. With this increased bias, there is a corresponding increase in grid voltage swing to drive up the zero grid voltage point on the curve.

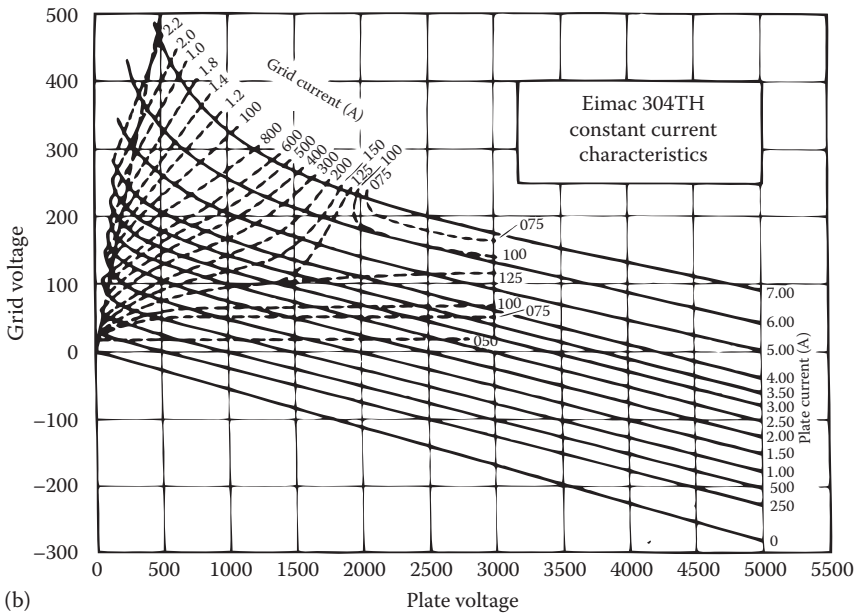
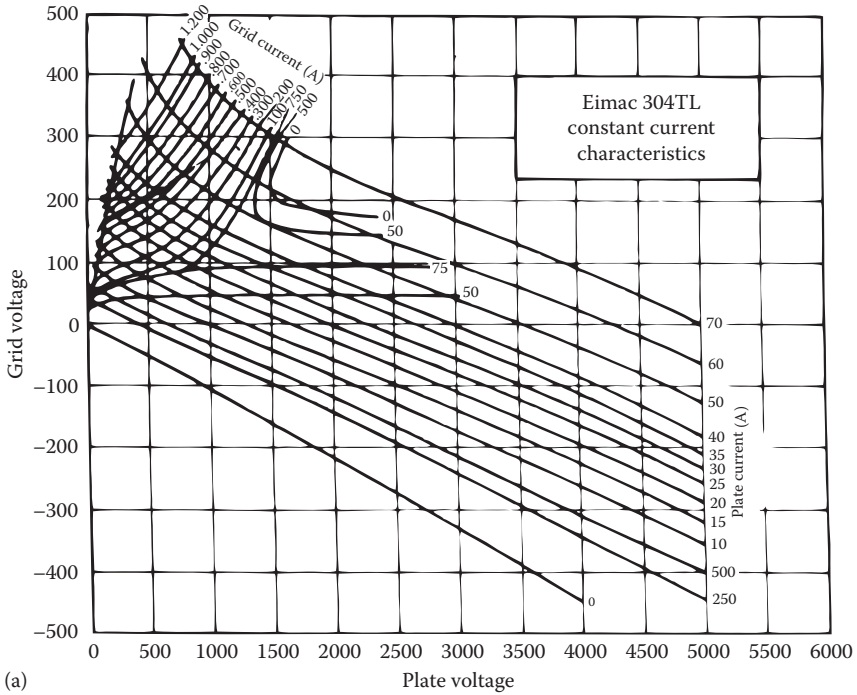


FIGURE 3.8 Constant-current characteristics for a triode: (a) with a μ of 12 and (b) with a μ of 20. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

Low- μ tubes have lower voltage gain by definition. This fact can be seen by comparing Figure 3.8a and b.

Triodes with a μ of 20–50 generally are used in conventional RF amplifiers and oscillators. High- μ triodes (200 or more) may be designed so that the operating bias is zero, as depicted in Figure 3.9. These *zero-bias triodes* are available with plate dissipation ratings of 400W–10kW or more. The zero-bias triode commonly is used in grounded-grid amplification. The tube offers good power gain and circuit simplicity. No bias power source is required. Furthermore, no protection circuits for loss of bias or drive are needed. Despite these attributes, present-day use of the zero-bias triode is limited.

Low- and medium- μ devices usually are preferred for induction heating applications because of the wide variations in load that an induction or dielectric heating oscillator normally works into. Such tubes exhibit lower grid-current variation with a changing load. The grid current of a triode with a μ of 20 will rise substantially less under a light- or no-load condition than a triode with a μ of 40. High- μ triode oscillators can be designed for heating applications, but extra considerations must be given to current rise under no-load conditions.

Vacuum tubes specifically designed for induction heating are available, intended for operation under adverse conditions. The grid structure is

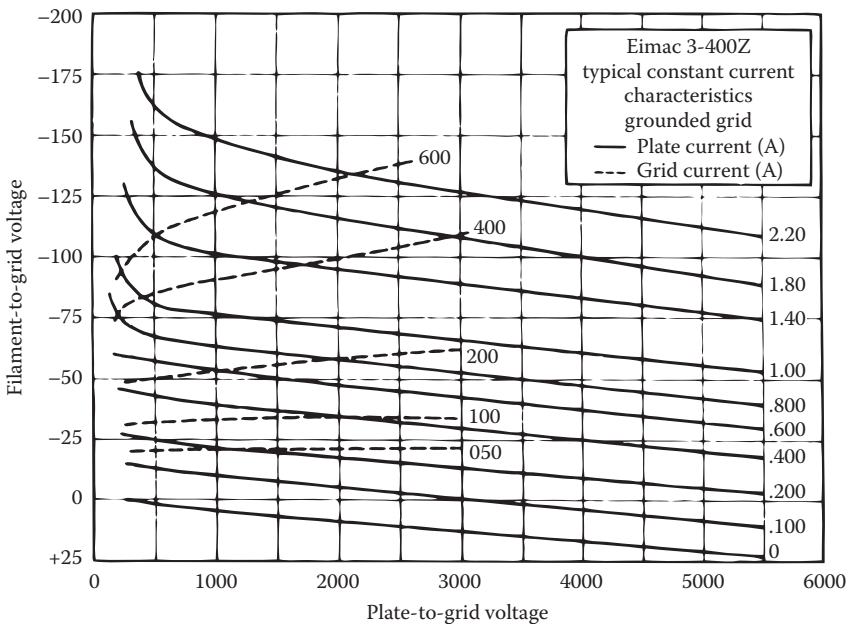


FIGURE 3.9 Grounded-grid constant-current characteristics for a zero-bias triode with a μ of 200. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

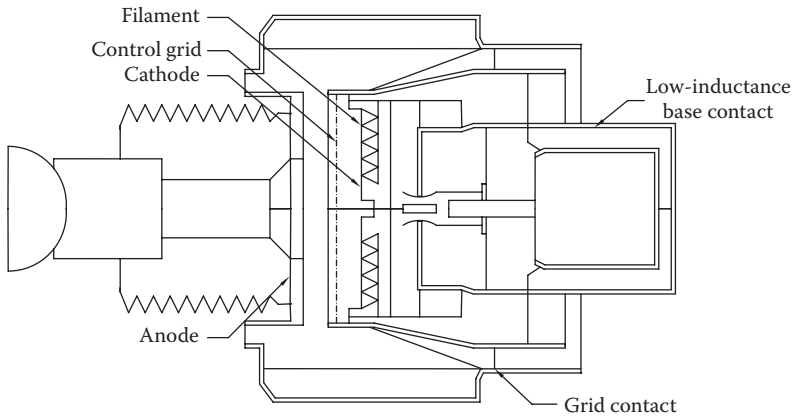


FIGURE 3.10

Internal configuration of a planar triode. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

ruggedized with ample dissipation capability to deal with wide variations in load. As the load decreases, grid dissipation increases.

Triodes also are manufactured with the cathode, grid, and anode in the shape of a flat surface, as shown in Figure 3.10. Tubes so constructed are called *planar triodes*. This construction technique permits operation at high frequencies. The close spacing reduces electron *transit time*, allowing the tube to be used at high frequencies (up to 3GHz or so). The physical construction of planar triodes results in short lead lengths, which reduces lead inductance. Planar triodes are used in both *continuous wave* (CW) and pulsed modes. The contacting surfaces of the planar triode are arranged for easy integration into coaxial and waveguide resonators.

3.3.3 Tetrode

The tetrode is a four-element tube with two grids [4]. The control grid serves the same purpose as the grid in a triode, while a second (*screen*) grid with the same number of vertical elements (bars) as the control grid is mounted between the control grid and the anode. The grid bars of the screen grid are mounted directly behind the control-grid bars, as observed from the cathode surface, and serve as a shield or screen between the input circuit and the output circuit of the tetrode. The principal advantages of a tetrode over a triode include the following:

- Lower internal plate-to-grid feedback.
- Lower drive power requirements. In most cases, the driving circuit need supply only 1% of the output power.
- More efficient operation. Tetrodes allow the design of compact, simple, flexible equipment with little spurious radiation and low intermodulation distortion.

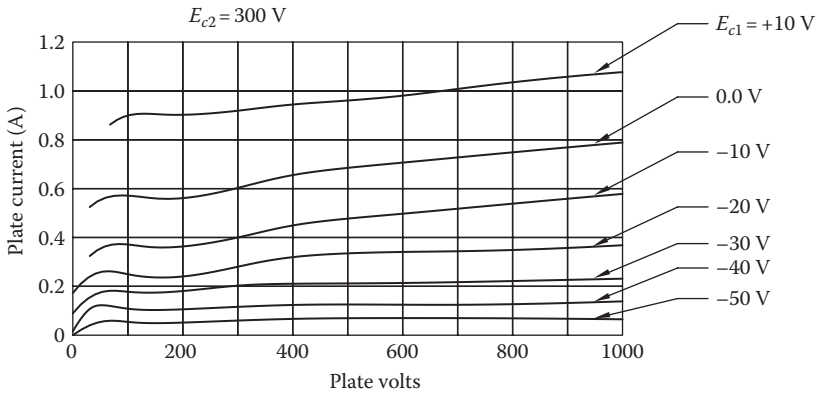


FIGURE 3.11 Tetrode plate current characteristics. Plate current is plotted as a function of plate voltage, with grid voltages as shown. (After Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.)

Plate current is almost independent of plate voltage in a tetrode. Figure 3.11 plots plate current as a function of plate voltage at a fixed screen voltage and various grid voltages. In an ideal tetrode, a change in plate current does not cause a change in plate voltage. The tetrode, therefore, can be considered a *constant-current device*. The voltages on the screen and control grids determine the amount of plate current.

The total cathode current of an ideal tetrode is determined by

$$I_k = K \left\{ E_{c1} + \frac{E_{c2}}{\mu_s} + \frac{E_b}{\mu_p} \right\}^{3/2} \tag{3.10}$$

where

- I_k is the cathode current
- K is a constant determined by tube dimensions
- E_{c1} is the control grid voltage
- E_{c2} is the screen grid voltage
- μ_s is the screen amplification factor
- μ_p is the plate amplification factor
- E_b is the plate voltage

The arithmetic value of the screen μ generally is not used in the design of an RF amplifier. In most tetrode applications, the screen amplification factor is useful to categorize roughly the performance to be expected.

3.3.3.1 Application Example

Figure 3.12 shows a radial beam power tetrode (4CX15000A) designed for class AB_1 or class C power amplification. The device is particularly well



FIGURE 3.12
Radial beam power tetrode (4CX15000A).
(Courtesy of Svetlana Electron Devices,
Huntsville, AL.)

suiting for RF linear power amplifier service. The tube has a directly heated thoriated-tungsten mesh filament for mechanical ruggedness and high efficiency. The maximum rated plate dissipation of the tube is 15 kW using air cooling.

The tube must be protected from damage that may result from an internal arc occurring at high plate voltage. A protective resistance typically is inserted in series with the tube anode to help absorb stored power supply energy in case an internal arc occurs.

The maximum control grid dissipation is 200 W, determined (approximately) by the product of the dc grid current and the peak positive grid voltage.

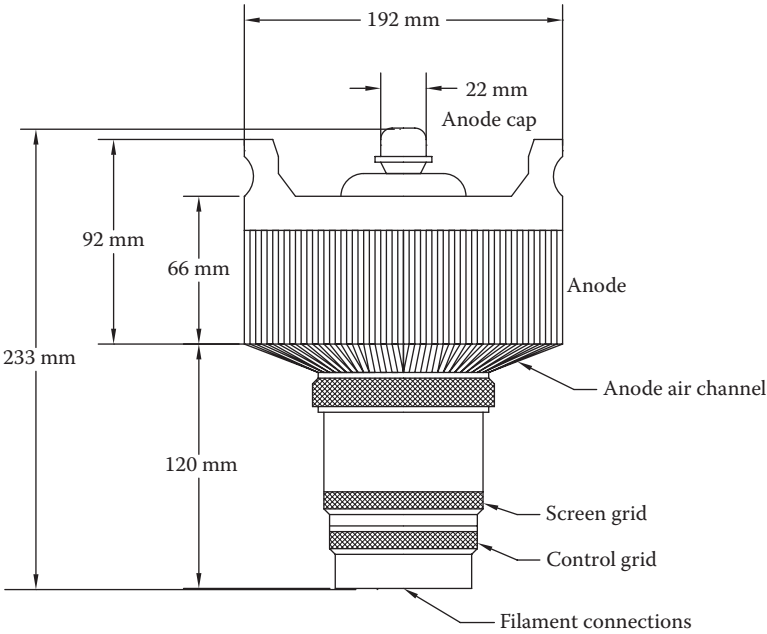
Screen grid maximum dissipation is 450 W. With no ac applied to the screen grid, dissipation is the product of dc screen voltage and dc screen current. Plate voltage, plate loading, and/or bias voltage must never be removed while filament and screen voltages are present.

The 4CX15000A must be mounted vertically, base up or down. The tube requires forced-air cooling in all applications. The tube socket is mounted in a pressurized compartment where cooling air passes through the socket and is guided to the anode cooling fins by an air chimney. Adequate movement of cooling air around the base of the tube keeps the base and socket contact fingers at a safe operating temperature. Although the maximum temperature rating for seals and the anode is 250°C, good engineering practice dictates that a safety factor be provided. Table 3.1 lists cooling parameters for the tube with the cooling air at 50°C and a maximum anode temperature of 225°C. The figures given in the table apply to designs in which air passes in the base-to-anode direction. Pressure drop values shown are approximate and apply to the tube/socket/chimney combination.

TABLE 3.1
Minimum Cooling Airflow Requirements for the
4CX15000A Power Tetrode at Sea Level

Plate Dissipation (W)	Airflow (CFM)	Pressure Drop (Inches of Water)
7,500	230	0.7
12,500	490	2.7
15,000	645	4.6

Source: Courtesy of Svetlana Electron Devices, Huntsville, AL.



Note: Dimensions shown are approximate.

FIGURE 3.13
Principal dimensions of 4CX15000A tetrode. (Courtesy of Svetlana Electron Devices, Huntsville, AL.)

At altitudes significantly above sea level, the flow rate must be increased for equivalent cooling. At 5,000 ft above sea level, both the flow rate and the pressure drop are increased by a factor of 1.20; at 10,000 ft, both flow rate and pressure drop are increased by 1.46.

Anode and base cooling is applied before or simultaneously with filament voltage turn-on, and normally should continue for a brief period (5 min) after shutdown to allow the tube to cool properly.

An outline of the principal tube dimensions is given in Figure 3.13. General specifications are listed in Table 3.2.

TABLE 3.2

General Characteristics of the 4CX15000A Power Tetrode

<i>Electrical characteristics</i>		
Filament type		Thoriated-tungsten mesh
Filament voltage		6.3 ± 0.3 V
Filament current		164 A (at 6.3V)
Amplification factor (average), grid to screen		4.5
Direct interelectrode capacitance (grounded cathode)	C_{in}	158 pF
	C_{out}	25.8 pF
	C_{pk}	1.3 pF
Direct interelectrode capacitance (grounded grid)	C_{in}	67 pF
	C_{out}	25.6 pF
	C_{gk}	0.21 pF
Maximum frequency for full ratings (CW)		110 MHz
<i>Mechanical characteristics</i>		
Length		238 mm (9.38 in.)
Diameter		193 mm (7.58 in.)
Net weight		5.8 kg (12.8 lb)
Operating position		Axis vertical, base up or down
Maximum operating temperature (seals/envelope)		250°C
Cooling method		Forced air
Base type		Coaxial
<i>Radio frequency power amplifier (class C FM) (absolute maximum ratings)</i>		
DC plate voltage		10,000 V
DC screen voltage		2,000 V
DC grid voltage		-750 V
DC plate current		5.0 A
Plate dissipation		15 kW
Screen dissipation		450 W
Grid dissipation		200 W
<i>Typical operation (frequencies up to 110 MHz)</i>		
DC plate voltage	7.5 kV dc	10.0 kV dc
DC screen voltage	750 V dc	750 V dc
DC grid voltage	-510 V dc	-550 V dc
DC plate current	4.65 A dc	4.55 A dc
DC screen current	0.59 A dc	0.54 A dc
DC grid current	0.30 A dc	0.27 A dc
Peak RF grid voltage	730 V	790 V
Calculated driving power	220 W	220 W
Plate dissipation	8.1	9.0 kW
Plate output power	26.7	36.5 kW

Source: Courtesy of Svetlana Electron Devices, Huntsville, AL.

3.3.4 Pentode

In all electron tubes, electrons striking the plate may, if moving at sufficient speed, dislodge other electrons [3]. In two- and three-electrode types, these dislodged electrons usually do not cause problems because no positive electrode other than the plate itself is present to attract them. These electrons, therefore, are drawn back to the plate. Emission caused by bombardment of an electrode by electrons from the cathode is called secondary emission.

In the case of screen-grid tubes, the proximity of the positive screen grid to the plate offers a strong attraction to these secondary electrons, and particularly so if the plate voltage swings lower than the screen-grid voltage. This effect reduces the plate current and limits the useful plate voltage swing for a tetrode.

The pentode is a five-electrode tube incorporating three grids [4]. The control and screen grids perform the same function as in a tetrode. The third grid, the *suppressor grid*, is mounted in the region between the screen grid and the anode. The suppressor grid produces a *potential minimum*, which prevents secondary electrons from being interchanged between the screen and plate. The pentode's main advantages over the tetrode include the following:

- Reduced secondary emission effects.
- Good linearity.
- Ability to let plate voltage swing below the screen voltage without excessive screen dissipation. This allows slightly higher power output for a given operating plate voltage.

Because of the design of the pentode, plate voltage has even less effect on plate current than in the tetrode. The same total space-current equation applies to the pentode as with the tetrode:

$$I_k = K \left\{ E_{c1} + \frac{E_{c2}}{\mu_s} + \frac{E_b}{\mu_p} \right\}^{3/2} \quad (3.11)$$

where

I_k is the cathode current

K is a constant determined by tube dimensions

E_{c1} is the control grid voltage

E_{c2} is the screen grid voltage

μ_s is the screen amplification factor

μ_p is the plate amplification factor

E_b is the plate voltage

The suppressor grid may be operated negative or positive with respect to the cathode. It also may be operated at cathode potential. It is possible to control plate current by varying the potential on the suppressor grid. Because of this ability, a modulating voltage can be applied to the suppressor to achieve amplitude modulation. The required modulating power is low because of the low electron interception of the suppressor.

3.3.4.1 Application Example

The 5CX1500B is a ceramic/metal power pentode designed for use as a class AB₁ linear amplifier or FM VHF amplifying device. The exterior of the device is shown in Figure 3.14. Basic specifications for the tube are given in Table 3.3.

The filament of the 5CX1500B is rated for 5.0 V. Filament voltage, measured at the socket, is maintained to $\pm 5\%$ of the rated value for maximum tube life.

The rated dissipation of the control grid is 25 W. This value is approximately the product of dc grid current and peak positive grid voltage. Operation at bias and drive levels near those listed in Table 3.3 will ensure safe operation.

The power dissipated by the screen grid must not exceed 75 W. Screen dissipation, in cases where there is no ac applied to the screen, is the simple product of screen voltage and screen current. If the screen voltage is modulated,



FIGURE 3.14
5CX1500B power pentode device. (Courtesy of Svetlana Electron Devices, Huntsville, AL.)

TABLE 3.3

General Characteristics of the 5CX1500B Power Pentode

<i>Electrical characteristics</i>		
Filament type		Thoriated-tungsten mesh
Filament voltage		5.0 ± 0.25 V
Filament current		38.5 A (at 5.0 V)
Transconductance (avg.), $I_b = 1.0$ A dc, $E_{c2} = 500$ V dc		24,000 μmhos
Amplification factor (average), grid to screen		5.5
Direct interelectrode capacitance (grounded cathode)	Input	75 pF
	Output	17.8 pF
	Feedback	0.20 pF
Frequency (maximum) CW		110 MHz
<i>Mechanical characteristics</i>		
Cooling method		Forced air
Base		Ring and breechblock
Recommended air system socket		SK-840 series
Recommended (air) chimney		SK-806
Operating position		Axis vertical, base down or up
Maximum operating temperature		250°C
Maximum dimensions	Length	130 mm (5.2 in.)
	Diameter	85.6 mm (3.37 in.)
Net weight		850 gm (30 oz)
<i>Radio frequency linear amplifier (class C, CW conditions) maximum ratings</i>		
Plate voltage		5000 V
Screen voltage		750 V
Plate dissipation		1500 W
Suppressor dissipation		25 W
Screen dissipation		75 W
Grid dissipation		25 W
<i>Typical operation (frequencies to 30 MHz)</i>		
Plate voltage	3000 V dc	4500 V dc
Suppressor voltage	0	0
Screen voltage	500 V dc	500 V dc
Grid voltage	-200 V dc	-200 V dc
Plate current	900 mA dc	900 mA dc
Screen current	94 mA dc	88 mA dc
Grid current	35 mA dc	34 mA dc
Peak RF grid voltage	255 V	255 V

(continued)

TABLE 3.3 (continued)

General Characteristics of the 5CX1500B Power Pentode

Calculated driving power	9.0 W	9.0 W
Plate input power	2700 W	4050 W
Plate dissipation	720 W	870 W
Plate output power	1980 W	3180 W

Source: Courtesy of Svetlana Electron Devices, Huntsville, AL.

screen dissipation will depend on the rms screen current and voltage. Screen dissipation is likely to rise to excessive levels if the plate voltage, bias voltage, or plate load are removed with filament and screen voltages applied. Suitable protective means must be provided to limit screen dissipation to 75 W in the event of a circuit failure.

The rated dissipation of the suppressor grid is 25 W. Suppressor current is zero or nearly zero for all typical operating conditions specified in Table 3.3. The 5CX1500B has been designed for zero voltage operation of the suppressor grid for most applications.

The plate dissipation rating of the 5CX1500B is 1500 W. The tube and associated circuitry should be protected against surge current in the event of an arc inside the tube. A current-limiting resistance of 10–25 Ω between the power supply and plate is sufficient. The resistor should be capable of withstanding the surge currents that may develop; it should not be used as a fuse.

Cooling requirements for the 5CX1500B follow along the same lines as for the 4CX15000A discussed in the previous section. Table 3.4 lists the airflow requirements for the pentode. Note that the power dissipated by the device is equal to the power dissipated by the plate, heater, and grids (control, screen, and suppressor) combined. In the case of the 5CX1500B, the heater and three grids account for approximately 350 W dissipation.

Mechanical dimensions for the 5CX1500B are given in Figure 3.15.

TABLE 3.4

Minimum Cooling Airflow Requirements for the 5CX1500B Power Tetrode

Plate Dissipation (W)		Airflow (CFM)	Pressure Drop (Inches of Water)
At sea level	1000	27	0.33
	1550	47	0.76
At 6000 ft	1000	33	0.40
	1550	58	0.95

Source: Courtesy of Svetlana Electron Devices, Huntsville, AL.

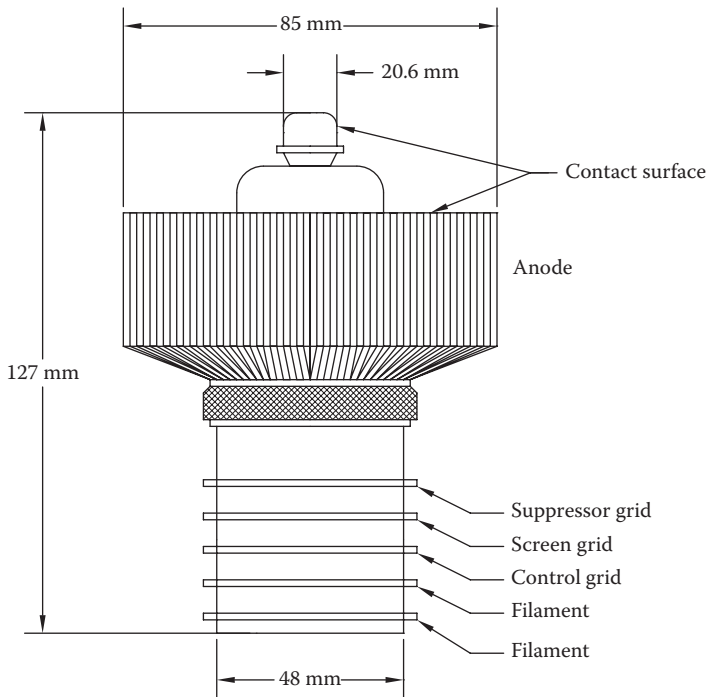


FIGURE 3.15

Principal dimensions of 5CX1500B pentode. (Courtesy of Svetlana Electron Devices, Huntsville, AL.)

3.4 High-Frequency Operating Limits

As with most active devices, performance of a given vacuum tube deteriorates as the operating frequency is increased beyond its designed limit [5]. Electron *transit time* is a significant factor in the upper-frequency limitation of electron tubes. A finite time is taken by electrons to traverse the space from the cathode, through the grid, and travel on to the plate. As the operating frequency increases, a point is reached at which the electron transit-time effects become significant. This point is a function of the accelerating voltages at the grid and anode and their respective spacing. Tubes with reduced spacing in the grid-to-cathode region exhibit reduced transit-time effects.

A power limitation also is interrelated with the high-frequency limit of a device. As the operating frequency is increased, closer spacing and smaller-sized electrodes must be used. This reduces the power-handling capability of the tube. Figure 3.16 illustrates the relationship.

Gridded tubes at all power levels for frequencies up to about 1 GHz are invariably cylindrical in form. At higher frequencies, planar construction is almost universal. As the operating frequency is increased beyond

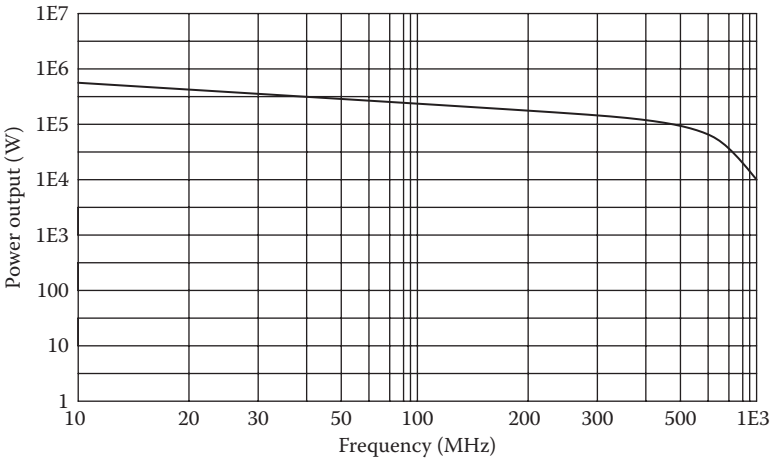


FIGURE 3.16 Continuous wave output power capability of a gridded vacuum tube. (After Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989)

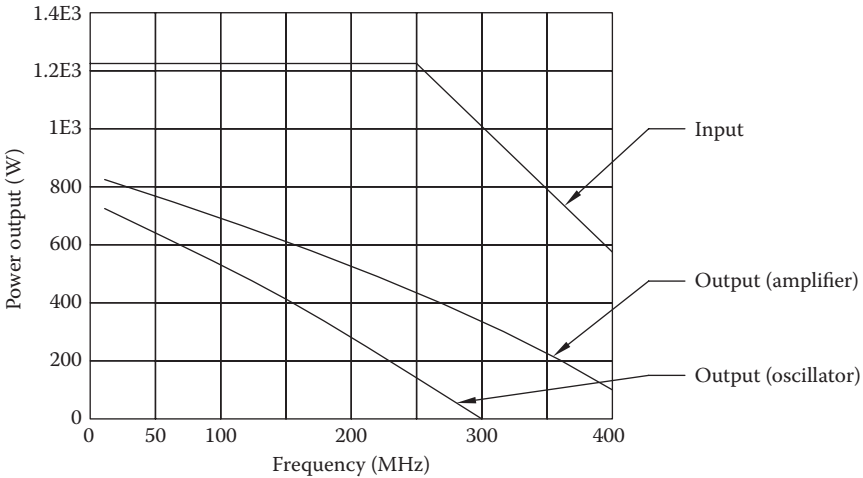


FIGURE 3.17 Performance of a class C amplifier as the operating frequency is increased beyond the design limits of the vacuum tube.

design limits, output power and efficiency both decrease. Figure 3.17 illustrates the relationship.

Transit time typically is not a problem for power grid tubes operating below 30MHz. Depending on the application, power grid tubes can be used at 100MHz and above without serious consideration of transit-time effects. Klystrons and other microwave tubes actually take advantage of transit time in their basic design.

3.4.1 Transit-Time Effects

When class C, class B, or similar amplifier operations are carried out at frequencies sufficiently high that the transit time of the electrons is not a negligible fraction of the waveform cycle, the following complications are observed in grid-based vacuum tubes:

- *Back heating* of the cathode
- Loading of the control grid circuit as a result of energy transferred to electrons that do not necessarily reach the grid to produce a dc grid current
- Debunching of plate current pulses
- Phase differences between the plate current and the exciting voltage applied to the control grid

Back heating of the cathode occurs when the transit time in the grid-cathode space is sufficiently great to cause an appreciable number of electrons to be in transit at the instant the plate current pulse would be cut off in the case of low-frequency operation. A considerable fraction of the electrons thus trapped in the interelectrode space are returned to the cathode by the negative field existing in the grid-cathode space during the cutoff portion of the cycle. These returning electrons act to heat the cathode. At very high frequencies, this back-heating is sufficient to supply a considerable fraction of the total cathode heating required for normal operation. Back-heating may reduce the life of the cathode as a result of electron bombardment of the emitting surface. It also causes the required filament current to depend upon the conditions of operation within the tube.

Energy absorbed by the control grid as a result of input loading is transferred directly to the electron stream in the tube. Part of this stream acts to produce back-heating of the cathode. The remainder affects the velocity of the electrons as they arrive at the anode of the tube. This portion of the energy is not necessarily all wasted. In fact, a considerable percentage of it may, under favorable conditions, appear as useful output in the tube. To the extent that this is the case, the energy supplied by the exciting voltage to the electron stream is simply transferred directly from the control grid to the output circuits of the tube without amplification.

An examination of the total time required by electrons to travel from the cathode to the anode in a triode, tetrode, or pentode operated as a class C amplifier reveals that the resulting transit times for electrons at the beginning, middle, and end of the current pulse will differ as the operating frequency is increased. In general, electrons traversing the distance during the first segment of the pulse will have the shortest transit time, while those near the middle and end of the pulse will have the longest transit times, as illustrated in Figure 3.18. The first electrons in the pulse have a short transit time

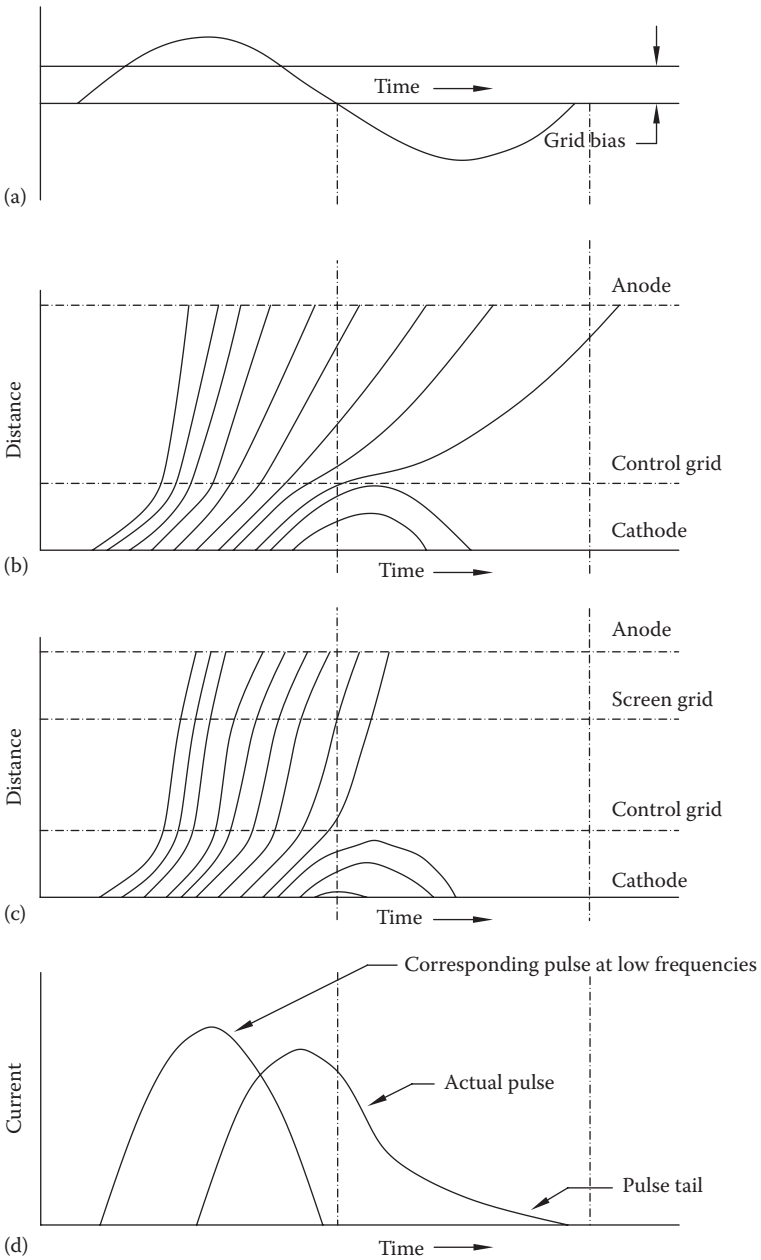


FIGURE 3.18 Transit-time effects in a class C amplifier: (a) control grid voltage, (b) electron position as a function of time (triode case), (c) electron position as a function of time (tetrode case), and (d) plate current (triode case).

because they approach the plate before the plate potential is at its minimum value. Electrons near the middle of the pulse approach the plate with the instantaneous plate potential at or near minimum and, consequently, travel less rapidly in the grid-plate space. Finally, those electrons that leave the cathode late in the current pulse (those just able to escape being trapped in the control grid-cathode space and returned toward the cathode) will be slowed as they approach the grid, and so have a large transit time. The net effect is to cause the pulse of plate current to be longer than it would be in operation at a low frequency. This causes the efficiency of the amplifier to drop at high frequencies, because a longer plate current pulse increases plate losses.

3.5 Vacuum Tube Design

Any particular power vacuum tube may be designed to meet a number of operating parameters, the most important usually being high operating efficiency and high gain/bandwidth properties. Above all, the tube must be reliable and provide long operating life. The design of a new power tube is a lengthy process that involves computer-aided calculations and modeling. The design engineers must examine a laundry list of items, including

- *Cooling*: how the tube will dissipate heat generated during normal operation. A high-performance tube is of little value if it will not provide long life in typical applications. Design questions include whether the tube will be air cooled or water cooled, the number of fins the device will have, and the thickness and spacing of the fins.
- *Electro-optics*: how the internal elements line up to achieve the desired performance. A careful analysis must be made of what happens to the electrons in their paths from the cathode to the anode, including the expected power gain of the tube.
- *Operational parameters*: what the typical interelectrode capacitances will be, and the manufacturing tolerances that can be expected. This analysis includes spacing variations among elements within the tube, the types of materials used in construction, the long-term stability of the internal elements, and the effects of thermal cycling.

3.5.1 Device Cooling

The main factor that separates tube types is the method of cooling used: air, water, or vapor. Air-cooled tubes are common at power levels below 50kW. A water cooling system, although more complicated, is more effective than air

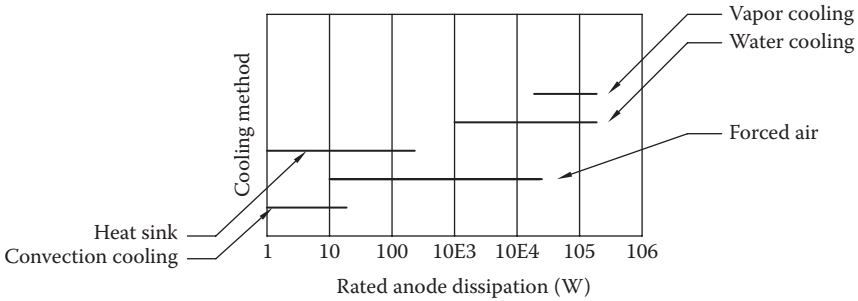


FIGURE 3.19 Relationship between anode dissipation and cooling method. (After Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.)

cooling—by a factor of 5–10 or more—in transferring heat from the device. Air cooling at the 100kW level is virtually impossible because it is difficult to physically move enough air through the device (if the tube is to be of reasonable size) to keep the anode sufficiently cool. Vapor cooling provides an even more efficient method of cooling a power amplifier (PA) tube than water cooling, for a given water flow and a given power dissipation. Naturally, the complexity of the external blowers, fans, ducts, plumbing, heat exchangers, and other hardware must be taken into consideration in the selection of a cooling method. Figure 3.19 shows how the choice of cooling method is related to anode dissipation.

3.5.1.1 Air Cooling

A typical air cooling system for a transmitter is shown in Figure 3.20. Cooling system performance for an air-cooled device is not necessarily related to air-flow volume. The cooling capability of air is a function of its mass, not its volume. An appropriate airflow rate within the equipment is established by the manufacturer, resulting in a given resistance to air movement.

The altitude of operation is also a consideration in cooling system design. As altitude increases, the density (and cooling capability) of air decreases. To maintain the same cooling effectiveness, increased airflow must be provided.

3.5.1.2 Water Cooling

Water cooling usually is preferred over air cooling for power outputs above about 50kW. Multiple grooves on the outside of the anode, in conjunction with a cylindrical jacket, force the cooling water to flow over the surface of the anode, as illustrated in Figure 3.21.

Because the water is in contact with the outer surface of the anode, a high degree of purity must be maintained. A resistivity of 1mΩ cm (at 25°C) typically is specified by tube manufacturers. Circulating water can remove about 1kW/cm² of effective internal anode area. In practice, the temperature of water leaving the tube must be limited to 70°C to prevent the possibility of spot boiling.

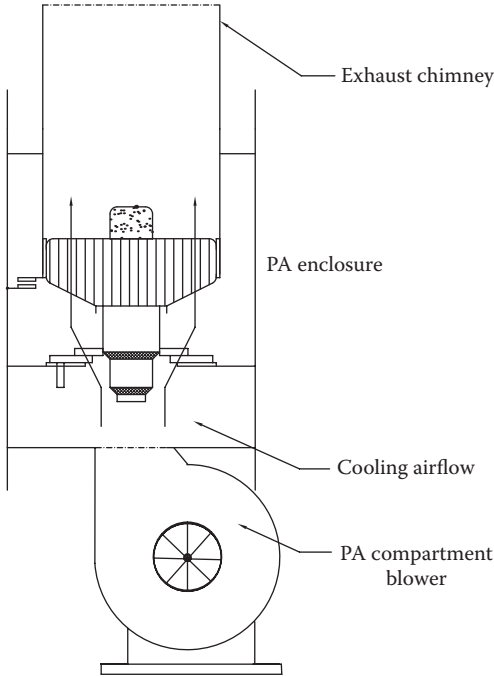


FIGURE 3.20
Typical transmitter PA stage cooling system.

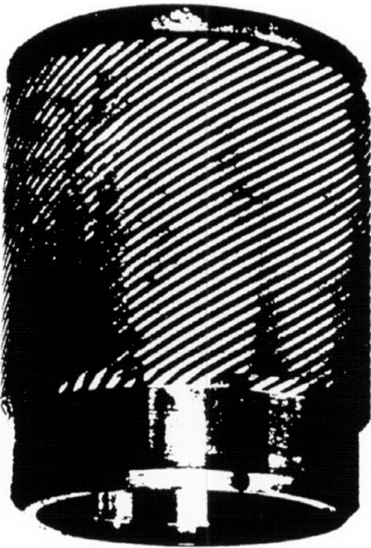


FIGURE 3.21
Water-cooled anode with grooves for controlled water flow. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

After leaving the anode, the heated water is passed through a heat exchanger where it is cooled to 30°C–40°C before being pumped back to the tube.

In typical modern water-cooled tubes, the dissipation ranges from 500 W to 1 kW/cm². Water-cooled anodes have been designed that can operate at 10 kW/cm² and above. However, with the exception of “big science” applications, dissipations this high are seldom required.

3.5.1.3 Vapor-Phase Cooling

Vapor cooling allows the permissible output temperature of the water to rise to the boiling point, giving higher cooling efficiency compared with water cooling [4]. The benefits of vapor-phase cooling are the result of the physics of boiling water. Increasing the temperature of 1 g of water from 40° to 70°C requires 30 cal of energy. However, transforming 1 g of water at 100°C into steam vapor requires 540 cal. Thus, a vapor-phase cooling system permits essentially the same cooling capacity as water cooling, but with greatly reduced water flow. Viewed from another perspective, for the same water flow, the dissipation of the tube may be increased significantly (all other considerations being the same).

A typical vapor-phase cooling system is shown in Figure 3.22. A tube incorporating a specially designed anode is immersed in a boiler filled with distilled water. When power is applied to the tube, anode dissipation heats the water to the boiling point, converting the water to steam vapor. The vapor passes to a condenser, where it gives up its energy and reverts to a liquid state. The condensate then is returned to the boiler, completing the cycle. Electric valves and interlocks are included in the system to provide for operating safety and maintenance. A vapor-phase cooling system for a transmitter with multiple PA tubes is shown in Figure 3.23.

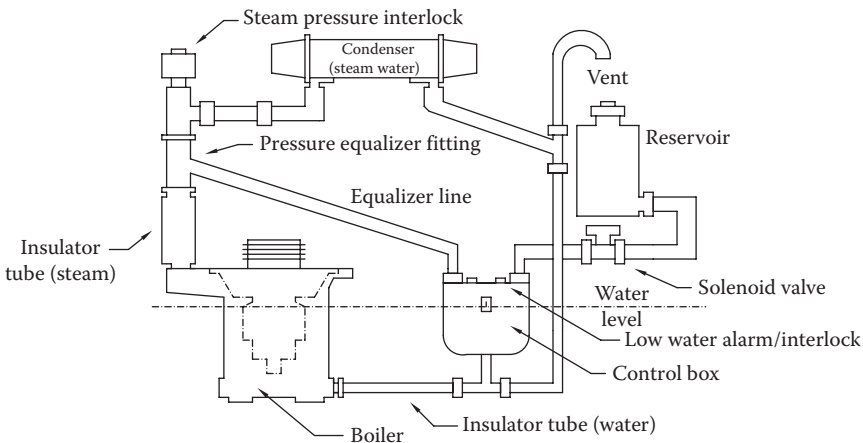


FIGURE 3.22

Typical vapor-phase cooling system. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

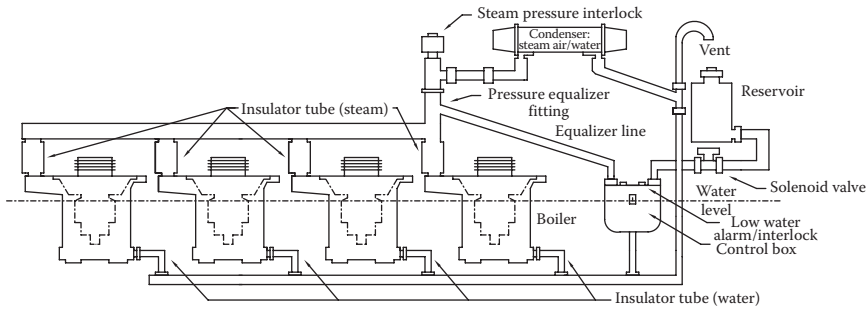


FIGURE 3.23

Vapor-phase cooling system for a 4-tube transmitter using a common water supply. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

3.5.1.4 Special Applications

Power devices used for research applications must be designed for transient overloading, requiring special considerations with regard to cooling. Oil, heat pipes, refrigerants, and, where high-voltage hold-off is a problem, gases (such as sulfurhexafluoride) are sometimes used to cool the anode of a power tube.

3.5.2 Cathode Assembly

The ultimate performance of any vacuum tube is determined by the accuracy of design and construction of the internal elements. The requirements for a successful tube include the ability to operate at high temperatures and withstand physical shock. Each element is critical to this objective.

The cathode used in a power tube obtains the energy required for electron emission from heat. The cathode may be directly heated or indirectly heated. A directly heated cathode, or filament cathode, is a wire heated by the passage of an electric current. An indirectly heated cathode, or heater cathode, consists of a filament (heater) enclosed in a metal sleeve. The sleeve carries the electron-emitting material on its outside surface and is heated by radiation and conduction from the heater element.

The three types of emitting surfaces most commonly used are

- Thoriated tungsten
- Alkaline-earth oxides
- Tungsten-barium-aluminate-impregnated emitters

The thoriated-tungsten and tungsten-impregnated cathodes are preferred in power tube applications because they are more tolerant to *ion bombardment*. The characteristics of the three emitting surfaces are summarized in Table 3.5.

A variety of materials may be used as a source of electrons in a vacuum tube. Certain combinations of materials are preferred however, for reasons of performance and economics.

TABLE 3.5
Characteristics of Common Thermionic Emitters

Emitter	Heating Method ^a	Operating Temp. (°C)	Emission Density	
			(A/cm ²) Average	Peak
Oxide	Direct and indirect	700–820	0.10–0.50	0.10–20
Thoriated tungsten	Direct	1600–1800	0.04–0.43	0.04–10
Impregnated tungsten	Direct and indirect	900–1175	0.5–8.0	0.5–12

Source: After Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.

^a Directly heated refers to a filament-type cathode.

3.5.2.1 Oxide Cathode

The conventional production-type oxide cathode consists of a coating of barium and strontium oxides on a base metal such as nickel [4]. Nickel alloys, in general, are stronger, tougher, and harder than most nonferrous alloys and many steels. The most important property of nickel alloys is their ability to retain strength and toughness at elevated temperatures. The oxide layer is formed by first coating a nickel structure (a can or disk) with a mixture of barium and strontium carbonates, suspended in a binder material. The mixture is approximately 60% barium carbonate and 40% strontium carbonate.

During vacuum processing of the tube, these elements are *baked* at high temperatures. As the binder is burned away, the carbonates subsequently are reduced to oxides. The cathode is then said to be *activated* and will emit electrons.

An oxide cathode operates CW at 700°C–820°C and is capable of an average emission density of 100–500 mA/cm². High emission current capability is one of the main advantages of the oxide cathode. Other advantages include high peak emission for short pulses and low operating temperature. As shown in Table 3.5, peak emission of up to 20 A/cm² is possible from an oxide cathode. A typical device is shown in Figure 3.24.

Although oxide-coated emitters provide more peak emission per watt of heating power than any other type, they are not without their drawbacks. Oxide emitters are more easily damaged or “poisoned” than other emitters and also deteriorate more rapidly when subjected to bombardment from high-energy particles.

The oxide cathode material will evaporate during the life of the tube, causing free barium to migrate to other areas within the device. This evaporation can be minimized in the design stage by means of a high-efficiency cathode that runs as cool as possible but still is not emission limited at the desired heater voltage. In the field, the heater voltage must not exceed the design value. An oxide cathode that is overheated produces little, if any, additional emission. The life of the tube under such operation, however, is shortened significantly.

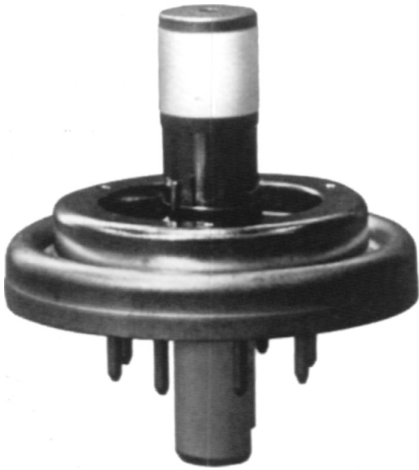
**FIGURE 3.24**

Photo of a typical oxide cathode. (Courtesy of Varian, Palo Alto, CA.)

3.5.2.2 Thoriated-Tungsten Cathode

The thoriated-tungsten filament is another form of atomic-film emitter commonly used in power grid tubes [4]. Tungsten is stronger than any other common metal at temperatures of more than 3500°F. The melting point of tungsten is 6170°F, higher than that of any other metal. The electrical conductivity of tungsten is approximately one-third that of copper, but much better than the conductivity of nickel, platinum, or iron-based alloys. The resistivity of tungsten in wire form is exploited in filament applications.

The thoriated-tungsten filament (or cathode) is created in a high-temperature gaseous atmosphere to produce a layer of *ditungsten carbide* on the surface of the cathode element(s). Thorium is added to tungsten in the process of making tungsten wire. The thorium concentration is typically about 1.5%, in the form of thoria. By proper processing during vacuum pumping of the tube envelope, the metallic thorium is brought to the surface of the filament wire. The result is an increase in emission of approximately 1000 times over a conventional cathode.

At a typical operating temperature of 1600°C–1800°C, a thoriated-tungsten filament will produce an average emission of 40–430 mA/cm². Peak current ranges up to 10 A/cm² or more.

One of the advantages of a thoriated-tungsten cathode over an oxide cathode is the ability to operate the plate at higher voltages. Because oxide cathodes are susceptible to deterioration caused by ion bombardment, plate voltage must be limited. A thoriated-tungsten cathode is more tolerant of ion bombardment, so higher plate voltages can be safely applied.

The end of useful life for a thoriated-tungsten tube occurs when most of the carbon has evaporated or has combined with residual gas, depleting

the carbide surface layer. Theoretically, a 3% increase in filament voltage will result in a 20 Kelvin (K) increase in cathode temperature, a 20% increase in peak emission, and a 50% decrease in tube life because of carbon loss. This cycle works in reverse, too. For a small decrease in temperature and peak emission, the life of the carbide layer—hence, the tube—may be increased.

3.5.2.3 Tungsten-Impregnated Cathode

The tungsten-impregnated cathode typically operates at 900°C–1175°C and provides the highest average emission density of the three types of cathodes discussed in this section (500 mA/cm²–8 A/cm²) [4]. Peak power performance ranges up to 12 A/cm².

Tungsten, as an element, is better able than other emitters to withstand bombardment by high-energy positive ions without suffering emission impairment. These positive ions are always present in small numbers in vacuum tubes as a result of ionization by collision with the residual gas.

3.5.2.4 Cathode Construction

Power tube filaments can be assembled in several different configurations [4]. Figure 3.25 shows a spiral-type filament, and Figure 3.26 shows a bar-type design. The spiral filament is used extensively in low-power tubes. As the size of the tube increases, mechanical considerations dictate a bar-type filament with spring loading to compensate for thermal expansion. A mesh

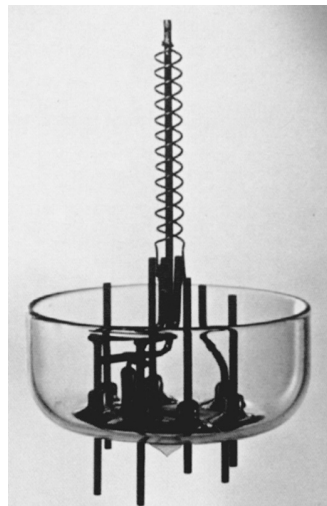


FIGURE 3.25
Spiral-type tungsten filament. (Courtesy of Varian, Palo Alto, CA.)

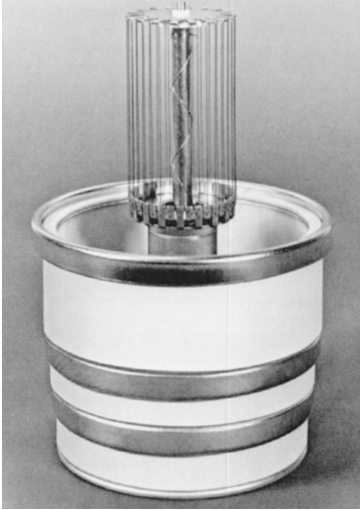


FIGURE 3.26
Bar-type tungsten filament. (Courtesy of Varian, Palo Alto, CA.)

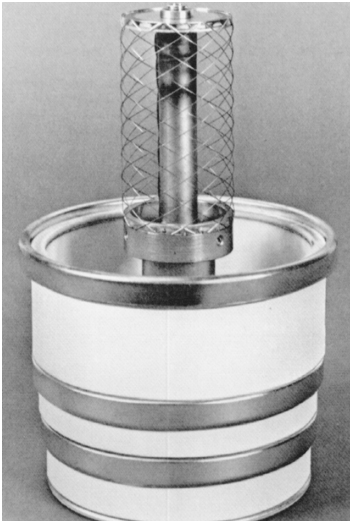


FIGURE 3.27
Mesh tungsten filament. (Courtesy of Varian, Palo Alto, CA.)

filament can be used for both small and large tubes. It is more rugged than other designs and less subject to damage from shock and vibration.

The rigidity of a cylindrical mesh cathode is determined by the following parameters:

- The diameter of the cathode
- The number, thickness, and length of the wires forming the cathode
- The ratio of welded to total wire crossings

A mesh cathode is shown in Figure 3.27.

Most power grid tubes are designed as a series of electron gun structures arranged in a cylinder around a centerline. This construction allows large amounts of plate current to flow and to be controlled with a minimum of grid interception. With reduced grid interception, less power is dissipated in the grid structures. In the case of the control grid, less driving power is required for the tube.

In certain applications, the construction of the filament assembly may affect the performance of the tube and that of the RF system as a whole. For example, filaments built in a basket-weave mesh arrangement usually offer lower distortion in critical high-level AM modulation circuits.

3.5.2.5 Velocity of Emission

The electrons emitted from a hot cathode depart with a velocity that represents the difference between the kinetic energy possessed by the electron just before emission and the energy that must be given up to escape. Because the energy of different electrons within the emitter is not the same, the velocity of emission will vary as well, ranging from zero to a maximum value determined by the type and style of emitter.

3.5.3 Grid Structures

The type of grid used for a power tube is determined principally by the power level and operating frequency required. For most medium-power tubes (5–25 kW dissipation) welded-wire construction is common. At higher power levels, laser-cut *pyrolytic graphite* grids may be found. The grid structures of a power tube must maintain their shape and spacing at elevated temperatures. They also must withstand shock and vibration.

3.5.3.1 Wire Grids

Conventional wire grids are prepared by operators that wind the assemblies using *mandrels* (forms) that include the required outline of the finished grid [4]. The operators spot weld the wires at intersecting points, as shown in Figure 3.28. Most grids of this type are made with tungsten or *molybdenum*, which exhibit stable physical properties at elevated temperatures. On a strength basis, pure molybdenum generally is considered the most suitable of all refractory metals at temperatures of 1600°F–3000°F. The thermal conductivity of molybdenum is more than three times that of iron and almost half that of copper.

Grids for higher-power tubes typically are built using a bar-cage type of construction. A number of vertical supports are fastened to a metal ring at the top and to a base cone at the bottom. The lower end of the assembly is bonded to a contact ring. The construction of the ring, metal

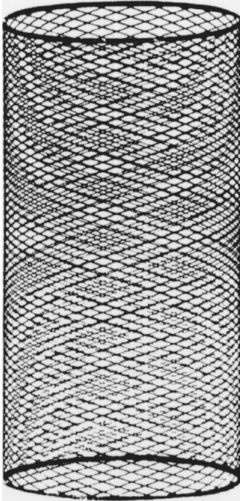


FIGURE 3.28
Mesh grid structure. (Courtesy of Varian, Palo Alto, CA.)

base cone, and cylindrical metal base give the assembly low lead inductance and low RF resistance.

The external loading of a grid during operation and the proximity of the grid to the hot cathode impose severe demands on both the mechanical stability of the structure and the physical characteristics of its surface. The grid absorbs a high proportion of the heat radiated by the cathode. It also intercepts the electron beam, converting part of its kinetic energy into heat. Furthermore, high-frequency capacitive currents flowing in the grid create additional heat.

The result is that grids are forced to work at temperatures as high as 1500°C. Their primary and secondary emissions, however, must be low. To prevent grid emission, high electron affinity must be ensured throughout the life of the tube, even though it is impossible to prevent material that evaporates from the cathode from contaminating the grid surface.

In tubes with oxide cathodes, grids made of tungsten or molybdenum wire are coated with gold to reduce primary emission caused by deposition. The maximum safe operating temperature for gold plating, however, is limited (about 550°C). Special coatings developed for high-temperature applications are effective in reducing grid emission. In tubes with thoriated-tungsten cathodes, grids made of tungsten or molybdenum are coated with proprietary compounds to reduce primary emission.

Primary grid emission is usually low in a thoriated-tungsten cathode device. In the case of an oxide cathode, however, free barium may evaporate from the cathode coating material and find its way to the control and screen grids. The rate of evaporation is a function of cathode temperature and time. A grid contaminated with barium will become another emitting surface. The hotter the grid, the greater the emissions will be.

3.5.3.2 K-Grid

To permit operation at higher powers (and, therefore, higher temperatures) the *K-grid* has been developed (Philips). The K-grid is a spot-welded structure of molybdenum wire doped to prevent brittleness and recrystallization. To eliminate mechanical stresses, the grid is annealed at 1800 K. It is then baked under vacuum at 2300 K to degas the molybdenum wire. The grid structure next is coated with zirconium carbide (sintered at 2300 K) and, finally, with a thin layer of platinum.

Despite high grid temperatures, the platinum coating keeps primary emissions to a minimum. The surface structure maintains low secondary emissions.

3.5.3.3 Pyrolytic Grid

Pyrolytic grids are a high-performance alternative to wire or bar grid assemblies. Used primarily for high-power devices, pyrolytic grids are formed by laser-cutting a graphite cup of the proper dimensions. The computer-controlled laser cuts numerous holes in the cup to simulate a conventional-style grid. Figure 3.29 shows a typical pyrolytic-type grid before and after laser processing.

Pyrolytic (or oriented) graphite is a form of crystallized carbon produced by the decomposition of a hydrocarbon gas at high temperatures in a controlled environment. A layer of pyrolytic graphite is deposited on a grid form. The thickness of the layer is proportional to the amount of time

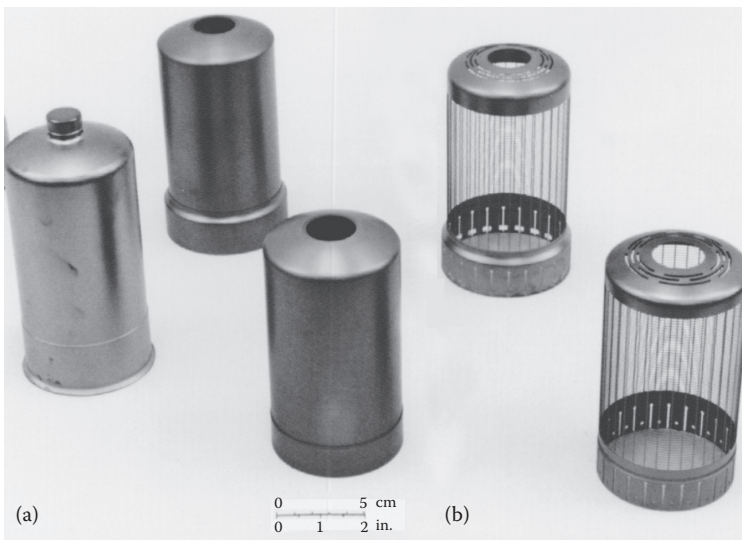


FIGURE 3.29

Pyrolytic graphite grid: (a) before laser processing and (b) completed assembly. (Courtesy of Varian, Palo Alto, CA.)

deposition is allowed to continue. The structural and mechanical properties of the deposited graphite depend upon the imposed conditions.

Pyrolytic grids are ideal vacuum tube elements because they do not expand like metal. Their small coefficient of expansion prevents movement of the grids inside the tube at elevated temperatures. This preserves the desired electrical characteristics of the device. Because tighter tolerances can be maintained, pyrolytic grids can be spaced more closely than conventional wire grids. Additional benefits are that the grid

- Is a single structure having no weld points
- Has a thermal conductivity in two of the three planes nearly that of copper
- Can operate at high temperatures with low vapor pressure

3.5.3.4 Grid Physical Structure

The control, screen, and suppressor grids are cylindrical and concentric. Each is slightly larger than the previous grid, as viewed from the cathode. Each is fastened to a metal base cone, the lower end of which is bonded to a contact ring. Figure 3.30 shows the construction of a typical screen grid assembly. Figure 3.31 provides a cutaway view of a tetrode power tube.

The shape of the control grid and its spacing from the cathode define, in large part, the operating characteristics of the tube. For best performance, the grid must be essentially transparent to the electron path from the cathode to the plate. In a tetrode, the control and screen grids must be precisely aligned to minimize grid current. For pentode tubes, these two conventions apply,

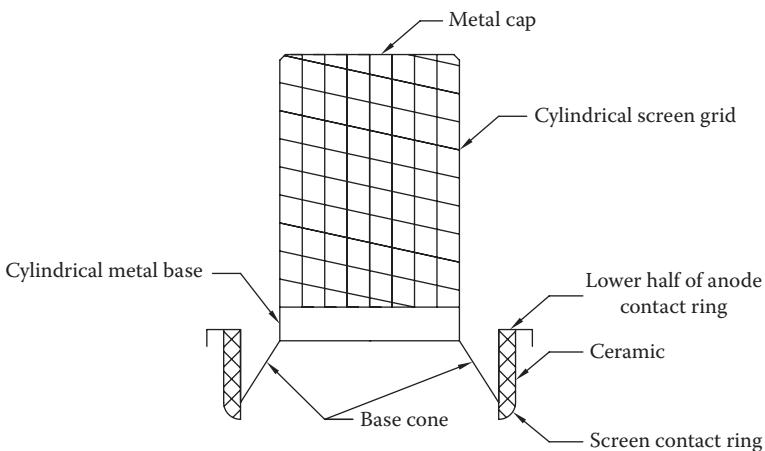


FIGURE 3.30

Screen grid assembly of a typical tetrode PA tube.

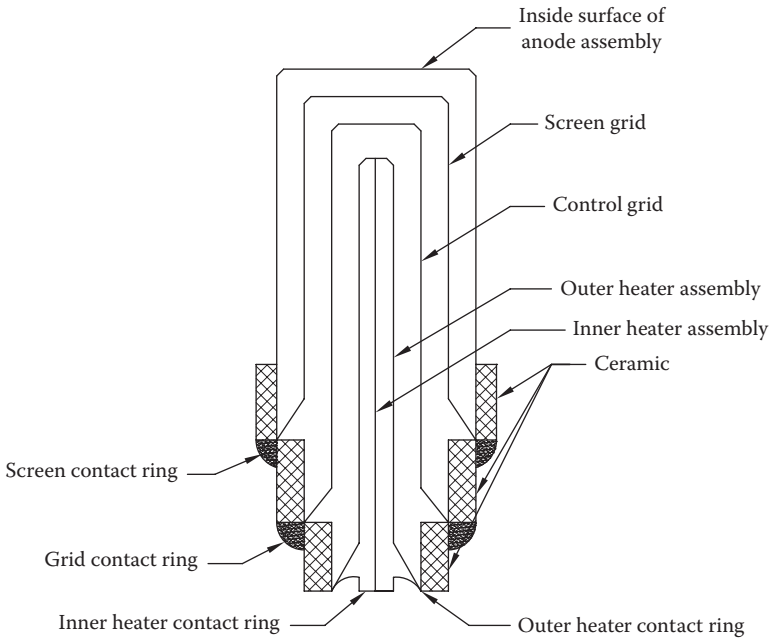


FIGURE 3.31

Internal arrangement of the anode, screen, control grid, and cathode assemblies of a tetrode power tube.

in addition to the requirement for precise alignment and minimum beam interception for the suppressor grid.

3.5.3.5 Secondary Emission Considerations

The relationship of the properties of secondary electrons to the grid structures and other elements of a vacuum tube must be considered carefully in any design. As the power capability of a tube increases, the physical size of the elements also increases. This raises the potential for secondary emission from the control, screen, and suppressor grids [4]. Secondary emission may occur regardless of the type of cathode used. The yield of secondary electrons may be reduced through the application of surface treatments.

In a tetrode, the screen is operated at a relatively low potential, necessary to accelerate the electrons emitted from the cathode. Not all electrons pass through the screen on their way to the plate. Some are intercepted by the screen grid. As the electrons strike the screen, other low-energy electrons are emitted. If these electrons have a stronger attraction to the screen, they will fall back to that element. If, however, they pass into the region between the screen and the plate, the much higher anode potential will attract them. The result is electron flow from screen to plate.

Because of the physical construction of a tetrode, the control grid will have virtually no control over screen-to-plate current flow as a result of secondary electrons. During a portion of the operating cycle of the device, it is possible that more electrons will leave the screen grid than will arrive. The result will be a reverse electron flow on the screen element, a condition common to high-power tetrodes. A low-impedance path for reverse electron flow must be provided.

Tube manufacturers typically specify the recommended value of bleeding current from the screen power supply to counteract the emission current. Two common approaches are illustrated in Figures 3.32 and 3.33. If the screen power supply impedance is excessively high in the direction of reverse electron flow, the screen voltage will attempt to rise to the plate voltage. Note the

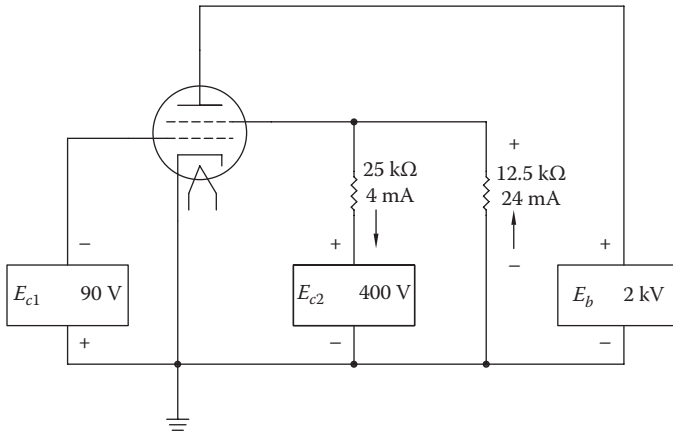


FIGURE 3.32 Circuit for providing a low-impedance screen supply for a tetrode. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

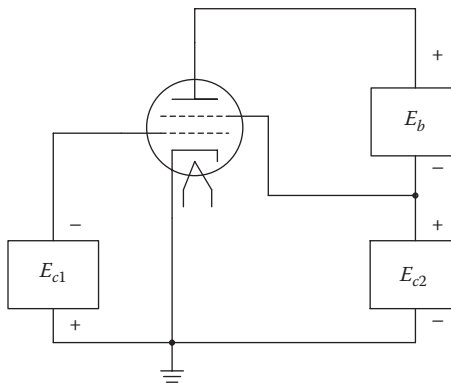


FIGURE 3.33 Series power supply configuration for swamping the screen circuit of a tetrode. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

emphasis on low impedance in the reverse electron flow direction. Most regulated power supplies are low impedance in the forward electron flow direction only. If the supply is not well bled, the reverse electrons will try to flow from anode to cathode in the regulator series pass element. As the screen voltage rises, the secondary and plate currents will increase, and the tube will enter a runaway condition.

As shown in Figure 3.32, the addition of a 12.5k Ω resistor from screen to ground provides a path for screen grid emission (20mA for the circuit shown). In the circuit of Figure 3.33, plate current flows through the screen power supply, swamping the screen supply. The screen power supply must, obviously, carry the normal screen and plate currents. This scheme is used extensively in circuits where the screen is operated at dc ground potential. The plate-to-cathode voltage is then the sum of the E_b and E_{c2} power supplies.

The suppressor grid of a pentode lessens the effects of secondary emission in the device, reducing the requirement to provide a reverse electron flow path for the screen grid power supply. The screen current requirement for a pentode, however, may be somewhat higher than for a tetrode of the same general characteristics.

The designer also must consider the impedance of the control grid circuit. Primary grid emission can result in operational problems if the grid circuit impedance is excessively high. Primary grid emission, in the case of an oxide cathode tube, will increase with tube life.

The size and power of gridded tubes dictate certain characteristics of electrical potential. As this geometry increases in electrical terms, significant secondary emission from the control grid can occur. Control grid secondary emission can exist whether the cathode is a thoriated-tungsten or oxide emitter, and it can occur in a triode, tetrode, or pentode. A typical curve of grid current as a function of grid voltage for a high-power thoriated-tungsten filament tetrode is shown in Figure 3.34. As shown in the figure, grid current decreases and eventually takes a reverse direction as the grid voltage increases. This reduction and reversal of grid current can be explained by the normal secondary emission characteristics of the metals used in the grid structure.

The secondary emission characteristics of common metals are presented in a curve form in Figure 3.35. The ratio of secondary-to-primary electron current is given as a function of primary electron potential. An examination of the chart will show the region between 200 and 600 V to be rather critical as far as secondary emission is concerned. Any power grid tube that normally operates with 200–600 V on the grid can exhibit the negative resistance characteristic of decreasing grid current with increasing grid voltage when another electrode—such as the anode in a triode or the screen grid in a tetrode—is at a sufficiently high potential to attract the emitted electrons from the control grid. A driver stage that works into such a nonlinear load normally must be designed in such a manner as to tolerate this condition. One technique involves swamping the driver so that changes in the load resulting from secondary grid emission represent only a small percentage of the total load the driver works into.

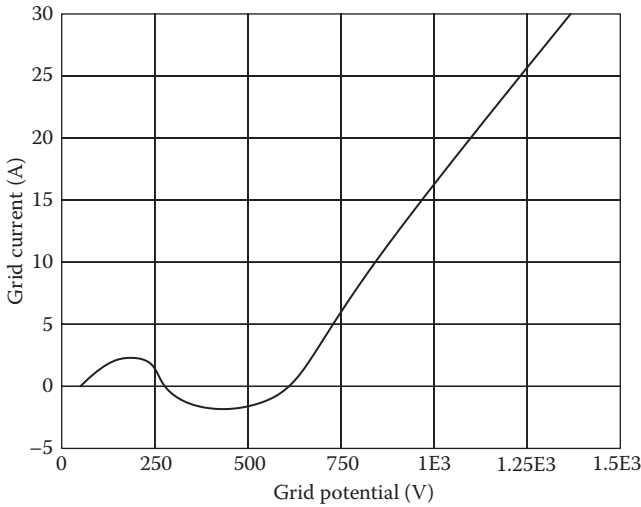


FIGURE 3.34 Typical curve of grid current as a function of control grid voltage for a high-power thoriated-tungsten filament tetrode. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

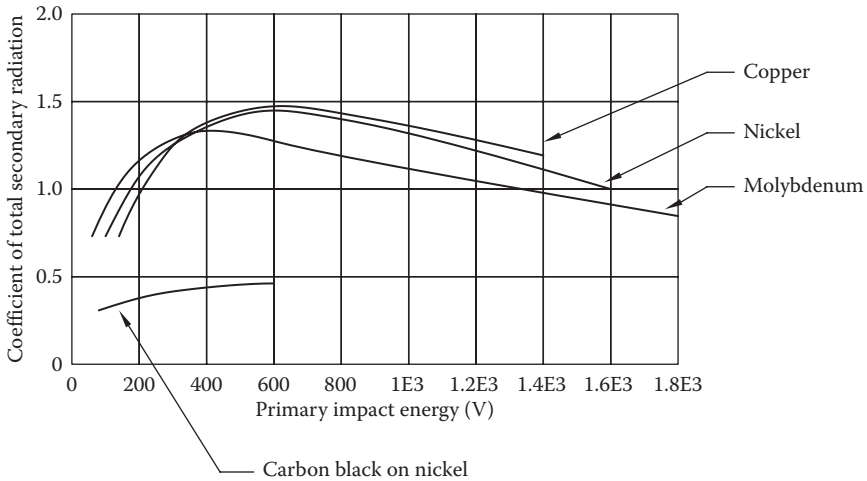


FIGURE 3.35 Secondary emission characteristics of various metals under ordinary conditions. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

3.5.4 Plate Assembly

The plate assembly of a power tube is typically a collection of many smaller parts that are machined and assembled to tight specifications [4]. Copper generally is used to construct the anode. It is an excellent material for this purpose because it has the highest electrical conductivity of any metal

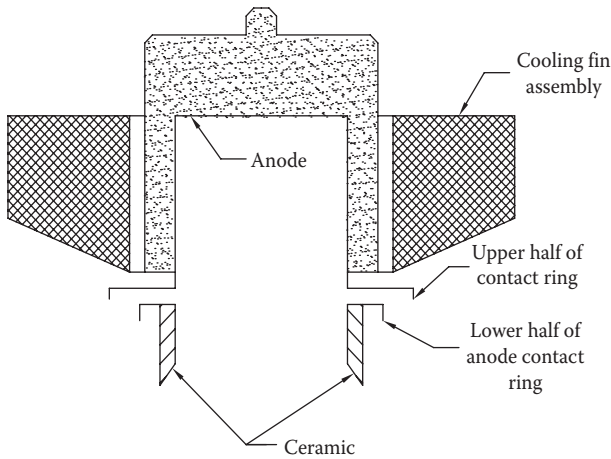


FIGURE 3.36
Cutaway view of the anode structure of an RF power amplifier tube.

except pure silver. Furthermore, copper is easily fabricated and ideally suited to cold-forming operations such as deep drawing, bending, and stamping.

The anode and cooling fins (in the case of an air-cooled device) begin as flat sheets of copper. They are stamped by the tube manufacturer into the necessary sizes and shapes. After all of the parts have been machined, the anode and cooling fins are stacked in their proper positions, clamped, and brazed into one piece in a brazing furnace.

The plate of a power tube resembles a copper cup with the upper half of a plate contact ring welded to the mouth and cooling fins silver-soldered or welded to the outside of the assembly. The lower half of the anode contact ring is bonded to a base ceramic spacer. At the time of assembly, the two halves of the ring are welded together to form a complete unit, as shown in Figure 3.36.

In most power tubes, the anode is a part of the envelope and, because the outer surface is external to the vacuum, it can be cooled directly. Glass envelopes were used in older power tubes. Most have been replaced, however, with devices that use ceramic as the envelope material.

3.5.4.1 Ceramic Elements

Ceramics are an integral part of modern power vacuum tubes. Three types of ceramics are in common usage in the production of vacuum devices:

- *Aluminum oxide*-based ceramics
- *Beryllium oxide* (BeO)-based ceramics
- *Aluminum nitride* (AlN)-based ceramics

3.5.4.2 Aluminum Oxide Ceramics

Aluminum oxide-based ceramic insulators are a common construction material for a wide variety of electric components, including vacuum tubes. Aluminum oxide is 20 times higher in thermal conductivity than most oxides. The flexure strength of commercial high-alumina ceramics is two to four times greater than that of most oxide ceramics. There are drawbacks, however, to the use of alumina ceramics, including

- Relatively high thermal expansion (approximately 7 ppm/°C), compared to other ceramic materials, such as BeO
- Moderately high dielectric constant (approximately 10)

Aluminas are fabricated from aluminum oxide powders with various percentages of *sintering promoters*, which produce the *glassy phase*. The later additives reduce the densification temperatures to between 1500°C and 1600°C. Based on the final application, the powders may be pressed, extruded, or prepared in slurries for slip casting or tape casting. The surface finish of aluminas is typically 3–25 µm/in. as a result of normal processing. For very smooth finishes (2 µm/in.), the surfaces may be lapped or polished.

Alumina ceramics for vacuum tube applications are rarely used apart from being bonded to metals. The means by which this is accomplished frequently dictate the processing technique. Metallization of aluminas usually is accomplished by either high-temperature firing or low-temperature thick-film processing. Fired, shaped aluminas usually are postmetallized by refiring the formed article after it has been coated with a slurry of molybdenum and manganese powder or tungsten metal powder. Based on the purity of the alumina, glass powder may be added to the metal powder. The mechanism of metallization requires an internal glassy phase or the added glassy phase for proper bonding. This is accomplished by firing in slightly reducing and moist atmospheres at temperatures above 1500°C. The resulting metallization usually is plated electrochemically with nickel, copper, or both.

Thick-film processing of alumina ceramics traditionally has been performed in oxidizing atmospheres at moderate temperatures (800°C–1000°C). Precious metals are used, such as gold, silver, platinum, palladium, and their combinations. A glassy phase usually is incorporated in the thick-film paste.

Each approach to metallization has advantages and disadvantages. The advantage of the high-temperature metallization schemes with molybdenum or tungsten is their moderate cost. The disadvantage is the high resistivity of the resulting films. These attributes are reversed for thick-film materials. Another advantage of thick-film metallization is that the process often is applied by the device fabricator, not the ceramic vendor. This can allow for greater efficiency in design modification.

In tubes designed for operation at 100 MHz and below, alumina typically is the ceramic of choice.

3.5.4.3 Beryllium Oxide Ceramics

BeO-based ceramics are in many ways superior to alumina-based ceramics. The major drawback is the toxicity of BeO. Beryllium and its compounds are a group of materials that are potentially hazardous and must be handled properly. With the necessary safeguards, BeO has been used successfully in many tube designs.

BeO materials are particularly attractive for use in power vacuum tubes because of their electrical, physical, and chemical properties. The thermal conductivity of BeO is approximately 10 times higher than that of alumina-based materials. Figure 3.37 compares the thermal conductivity of BeO with that of alumina and some alternative materials. As the chart illustrates, BeO has a lower dielectric constant and coefficient of thermal expansion than alumina. It is, however, also slightly lower in strength.

Beryllia materials are fabricated in much the same way as alumina compounds, although the toxic properties of the powders mandate that they be processed in laboratories equipped to handle them safely. Simple cutting, drilling, and postmetallization also are handled by vendors with special equipment. Reliable thick-film systems may be applied to BeO substrates. Such coatings usually necessitate less elaborate safety precautions and are often applied by the device fabricator.

3.5.4.4 Other Ceramics

AlNi-based ceramics have been developed as an alternative to the toxicity concerns of BeO-based materials. As shown in Figure 3.37, the thermal conductivity of AlNi is comparable to that of BeO but deteriorates less with temperature.

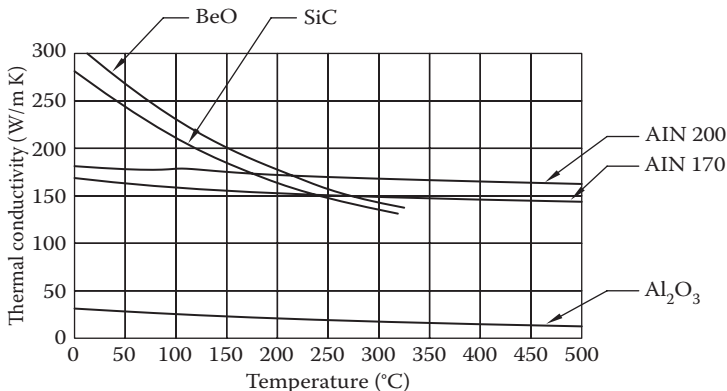


FIGURE 3.37

Thermal conductivity of common ceramics.

The dielectric constant of AlNi is comparable to that of alumina (a drawback), but its thermal expansion is low (4 ppm/°C).

Other ceramic materials that find some use in vacuum devices include silicon carbide (SiC)-based substances and boron nitride ceramics.

3.6 Tube Construction

Each type of power grid tube is unique insofar as its operating characteristics are concerned. The basic physical construction, however, is common to most devices. A vacuum tube is built in essentially two parts:

- The base, which includes the filament and supporting stem, control grid, screen grid, and suppressor grid (if used)
- The anode, which includes the heat-dissipating fins made in various machining steps

The base subassembly is welded using a *tungsten-inert gas* (TIG) process in an oxygen-free atmosphere (a process sometimes referred to as *Heliarc welding*) to produce a finished base unit.

The ceramic elements used in a vacuum tube are critical parts of the device. Assembled in sections, each element builds upon the previous one to form the base of the tube. The ceramic-to-metal seals are created using a material that is *painted* onto the ceramic and then heated in a brazing oven. After preparation in a high-temperature oven, the painted area provides a metallic structure that is molecularly bonded to the ceramic and provides a surface suitable for brazing.

This process requires temperature sequences that dictate completion of the highest-temperature stages first. As the assembly takes form, lower oven temperatures are used so that completed bonds will not be compromised.

Despite all the advantages that make ceramics one of the best insulators for a tube envelope, their brittleness is a potential cause of failure. The smallest cracks in the ceramic, not in themselves damaging, can cause the ceramic to break when mechanically stressed by temperature changes.

After the base assembly has been matched with the anode, the completed tube is brazed into a single unit. The device then goes through a *bake-out* procedure. Baking stations are used to evacuate the tube and bake out any oxygen or other gases from the copper parts of the assembly. Although oxygen-free copper is used in tube construction, some residual oxygen exists in the metal and must be driven out for long component life. A typical vacuum reading of 10^{-8} Torr is specified for most power tubes (formerly expressed as 10^{-8} mm of mercury). For comparison, this is the degree of vacuum in outer space about 200 miles above the earth.

A vacuum offers excellent electrical insulation characteristics. This property is essential for reliable operation of a vacuum tube, the elements of which typically operate at high potentials with respect to each other and to the surrounding environment. An electrode containing absorbed gases, however, will exhibit reduced breakdown voltage because the gas will form on the electrode surface, increasing the surface gas pressure and lowering the breakdown voltage in the vicinity of the gas pocket.

To maintain a high vacuum during the life of the component, power tubes contain a *getter* device. The name comes from the function of the element: to “get” or trap and hold gases that may exist inside the tube. Materials used for getters include zirconium, cerium, barium, and titanium.

The operation of a vacuum tube is an evolving chemical process. End of life in a vacuum tube generally is caused by loss of emission.

3.6.1 Connection Points

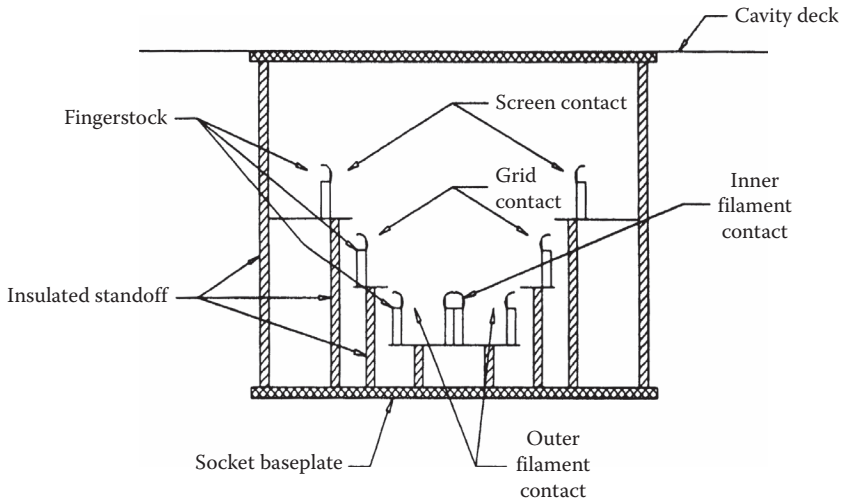
The high power levels and frequencies at which vacuum tubes operate place stringent demands on the connectors used to tie the outside world to the inside elements. Tubes are designed to be mounted vertically on their electrical connectors. The connectors provide a broad contact surface and mechanical support for the device.

The cathode and grids typically are mounted on ring-shaped *Kovar* bases, which also serve as contact rings for the external connections. Kovar is an iron-nickel-cobalt alloy whose coefficient of thermal expansion is comparable to that of aluminum oxide ceramic. The different diameters of the various contact rings allow them to be grouped coaxially. The concentric tube/connector design provides for operation at high frequencies. Conductivity is improved by silver plating.

3.6.2 Tube Sockets

Any one tube design may have several possible socket configurations, depending upon the frequency of operation. If the tube terminals are large cylindrical surfaces, the contacting portions of the socket consist of either spring *collets* or an assembly of spring fingerstock. Usually, these multiple-contacting surfaces are made of beryllium copper to preserve spring tension at the high temperatures present at the tube terminals. The fingers are silver plated to reduce RF resistance. Figure 3.38 shows a cutaway view of the base of a tetrode socket.

If the connecting fingers of a power tube socket fail to provide adequate contact with the tube element rings, a concentration of RF currents will result. Depending on the extent of this concentration, damage may result to the socket. After a connecting finger loses its spring action, the heating effect is aggravated, and tube damage is possible.

**FIGURE 3.38**

Cross section of the base of a tetrode socket showing the connection points.

A series of specialized power tubes is available with no sockets at all. Intended for cathode-driven service, the grid assembly is formed into a flange that is bolted to the chassis. The filament leads are connected via studs on the bottom of the tube. Such a configuration eliminates the requirement for a socket. This type of device is useful for low-frequency applications, such as induction heating.

3.7 Neutralization

An RF power amplifier must be properly neutralized to provide acceptable performance in most applications. The means to accomplish this end vary considerably from one design to another. An RF amplifier is neutralized when two operating conditions are met:

- The interelectrode capacitance between the input and output circuits is canceled.
- The inductance of the screen grid and cathode assemblies (in a tetrode) is canceled.

Cancellation of these common forms of coupling between the input and output circuits of vacuum tube amplifiers prevents self-oscillation and the generation of spurious products.

3.7.1 Circuit Analysis

Figure 3.39 illustrates the primary elements that affect neutralization of a vacuum tube RF amplifier operating in the VHF band [4]. (Many of the following principles also apply to lower frequencies.) The feedback elements include the residual grid-to-plate capacitance (C_{gp}), plate-to-screen capacitance (C_{ps}), and screen grid lead inductance (L_s). The RF energy developed in the plate circuit (E_p) causes a current (I) to flow through the plate-to-screen capacitance and the screen lead inductance. The current through the screen inductance develops a voltage ($-E$) with a polarity opposite that of the plate voltage (E_p). The $-E$ potential often is used as a method of neutralizing tetrode and pentode tubes operating in the VHF band.

Figure 3.40 graphically illustrates the electrical properties at work. The circuit elements of the previous figure have been arranged so that the height above or below the zero potential line represents the magnitude and polarity of the RF voltage for each part of the circuit with respect to ground (zero). For the purposes of this illustration, assume that all of the circuit elements involved are pure reactances. The voltages represented by each, therefore, are either in phase or out of phase and can be treated as positive or negative with respect to each other.

The voltages plotted in the figure represent those generated as a result of the RF output circuit voltage (E_p). No attempt is made to illustrate the typical driving current on the grid of the tube. The plate (P) has a high positive potential above the zero line, established at the ground point. Keep in mind that the distance above the baseline represents increasing positive potential.

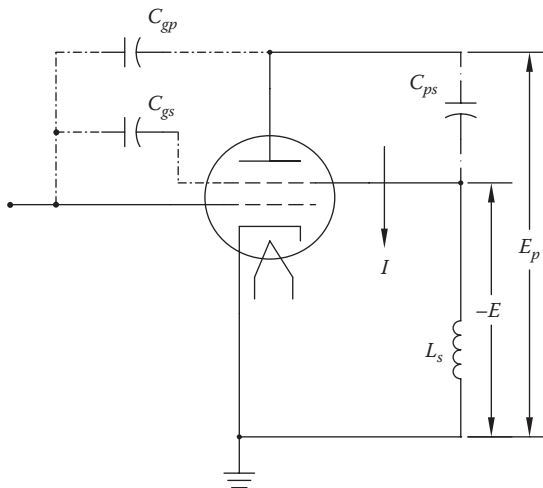


FIGURE 3.39

Elements involved in the neutralization of a tetrode PA stage. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

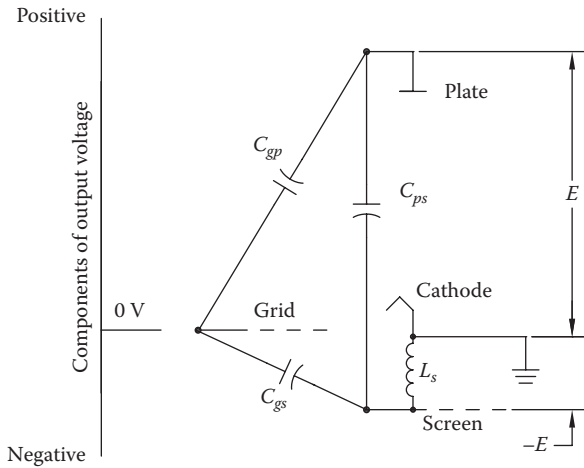


FIGURE 3.40 Graphical representation of the elements involved in the self-neutralization of a tetrode RF stage. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

The effect of the out-of-phase screen potential developed as a result of inductance L_s is shown, resulting in the generation of $-E$.

As depicted, the figure constitutes a perfectly neutralized circuit. The grid potential rests at the zero baseline. The grid operates at filament potential insofar as any action of the output circuit on the input circuit is concerned.

The total RF voltage between plate and screen is made up of the plate potential and screen lead inductance voltage, $-E$. This total voltage is applied across a divider circuit that consists of the plate-to-grid capacitance and grid-to-screen capacitance (C_{gp} and C_{gs}). When this potential divider is properly matched for the values of plate RF voltage (E_p) and screen lead inductance voltage ($-E$), the control grid will exhibit zero voltage difference with respect to the filament as a result of E_p .

3.7.2 Circuit Design

A variety of methods may be used to neutralize a vacuum tube amplifier [4]. Generally speaking, a grounded-grid, cathode-driven triode can be operated into the VHF band without external neutralization components. The grounded-grid element is sufficient to prevent spurious oscillations. Tetrode amplifiers generally will operate through the MF band without neutralization. However, as the gain of the stage increases, the need to cancel feedback voltages caused by tube interelectrode capacitances and external connection inductances becomes more important. At VHF and above, it is generally necessary to provide some form of stage neutralization.

3.7.2.1 Below VHF

For operation at frequencies below the VHF band, neutralization for a tetrode typically employs a capacitance bridge circuit to balance out the RF feedback caused by residual plate-to-grid capacitance [4]. This method assumes that the screen is well bypassed to ground, providing the expected screening action inside the tube.

Neutralization of low-power push-pull tetrode or pentode tubes can be accomplished with cross-neutralization of the devices, as shown in Figure 3.41. Small-value neutralization capacitors are used. In some cases, neutralization can be accomplished with a simple wire connected to each side of the grid circuit and brought through the chassis deck. Each wire is positioned to “look at” the plate of the tube on the opposite half of the circuit. Typically, the wire (or a short rod) is spaced a short distance from the plate of each tube. Fine adjustment is accomplished by moving the conductor in or out from its respective tube.

A similar method of neutralization can be used for a cathode-driven symmetrical stage, as shown in Figure 3.42. Note that the neutralization capacitors (C_n) are connected from the cathode of one tube to the plate of the opposite tube. The neutralizing capacitors have a value equal to the internal cathode-to-plate capacitance of the PA tubes.

In the case of a single-ended amplifier, neutralization can be accomplished using either a push-pull output or push-pull input circuit. Figure 3.43 shows a basic push-pull grid neutralization scheme that provides the out-of-phase voltage necessary for proper neutralization. It is usually simpler to create a push-pull network in the grid circuit than in the plate because of the lower voltages present. The neutralizing capacitor, C_n , is small and may consist of

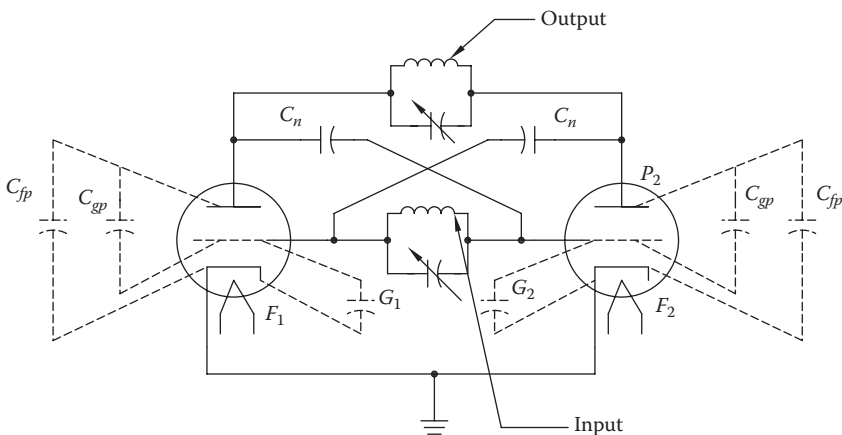


FIGURE 3.41

Push-pull grid neutralization. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

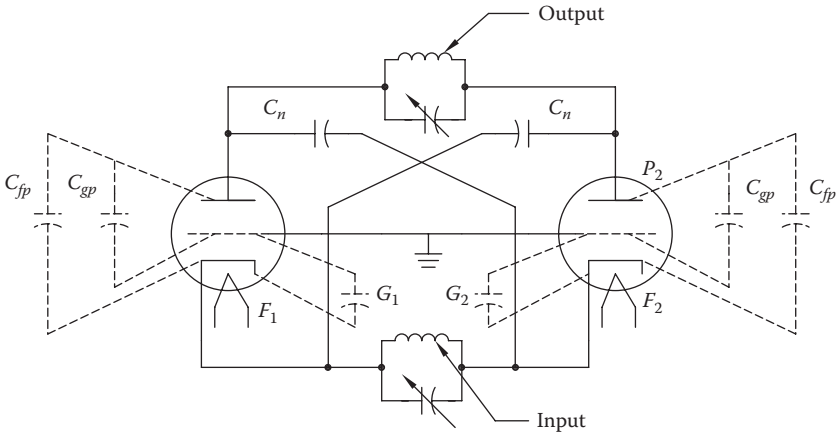


FIGURE 3.42 Symmetrical stage neutralization for a grounded-grid circuit. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

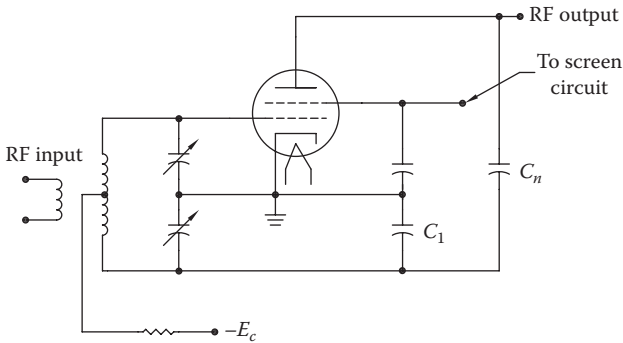


FIGURE 3.43 Push-pull grid neutralization in a single-ended tetrode stage. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

a simple feedthrough wire (described previously). A padding capacitor in parallel with C_v often is added to maintain the balance of the input circuit while tuning. The padding capacitor generally is equal in size to the input capacitance of the tube.

Single-ended tetrode and pentode stages also can be neutralized using the method shown in Figure 3.44. The input resonant circuit is placed above ground by a small amount because of the addition of capacitor C . The voltage to ground that develops across C upon the application of RF drive is out of phase with the grid voltage, and is fed back to the plate through C_n to provide neutralization. In such a design, C_n is considerably larger in value than the grid-to-plate interelectrode capacitance.

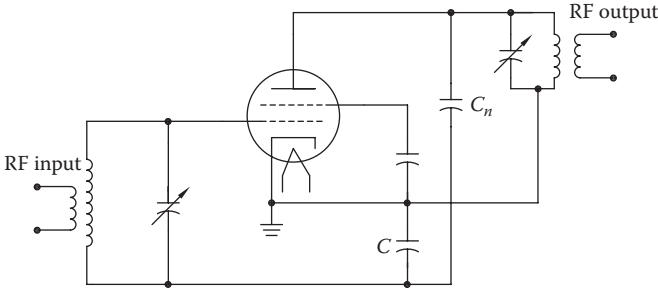


FIGURE 3.44
Single-ended grid neutralization for a tetrode. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

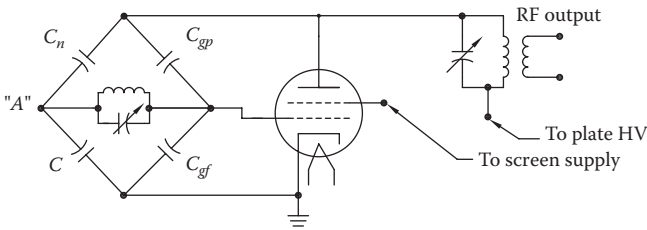


FIGURE 3.45
Previous figure redrawn to show the elements involved in neutralization. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

The single-ended grid neutralization circuit is redrawn in Figure 3.45 to show the capacitance bridge that makes the design work. Balance is obtained when the following condition is met:

$$\frac{C_n}{C} = \frac{C_{gp}}{C_{gf}} \tag{3.12}$$

where

- C_n is the neutralization capacitance
- C is the input circuit bypass capacitor
- C_{gp} is the grid-to-plate interelectrode capacitance
- C_{gf} is the total input capacitance, including tube and stray capacitance

A single-ended amplifier also can be neutralized by taking the plate circuit slightly above ground and using the tube capacitances as part of the neutralizing bridge. This circuit differs from the usual RF amplifier design in that the plate bypass capacitor is returned to the screen side of the screen bypass capacitor, as shown in Figure 3.46. The size of screen bypass capacitor C_s and the amount of stray capacitances in C are chosen to balance the voltages induced in the input by the internal tube capacitances, grid-to-plate (C_{gp}) and

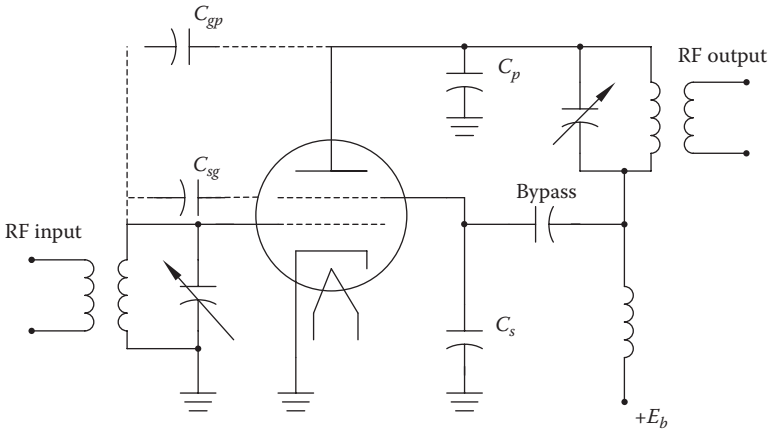


FIGURE 3.46 Single-ended plate neutralization. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

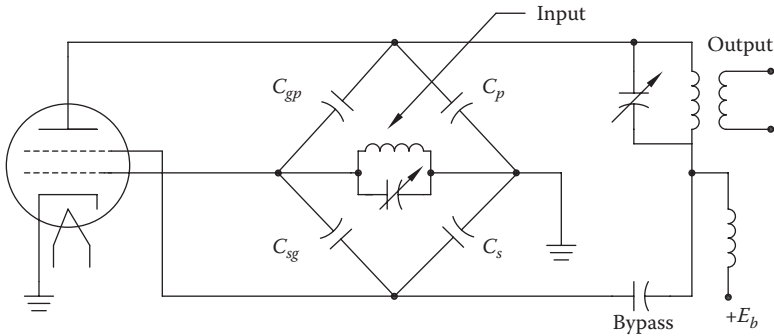


FIGURE 3.47 Single-ended plate neutralization showing the capacitance bridge present. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

screen-to-grid (C_{sg}). This circuit is redrawn in Figure 3.47 in the usual bridge form. Balance is obtained when the following condition is met:

$$\frac{C_p}{C_s} = \frac{C_{gp}}{C_{sg}} \tag{3.13}$$

In usual tetrode and pentode structures, the screen-to-grid capacitance is approximately half the published tube input capacitance. The tube input capacitance is primarily the sum of the grid-to-screen capacitance and the grid-to-cathode capacitance.

Note that in the examples given, it is assumed that the frequency of operation is low enough that inductances in the socket and connecting leads can

be ignored. This is basically true in MF applications and below. At higher bands, however, the effects of stray inductances must be considered, especially in single-ended tetrode and pentode stages.

3.7.2.2 VHF and Above

Neutralization of power grid tubes operating at very high frequencies provides special challenges and opportunities to the design engineer. At VHF and above, significant RF voltages may develop in the residual inductance of the screen, grid, and cathode elements. If managed properly, these inductances can be used to accomplish neutralization in a simple, straightforward manner.

At VHF and above, neutralization is required to make the tube input and output circuits independent of each other with respect to reactive currents. Isolation is necessary to ensure independent tuning of the input and output. If variations in the output voltage of the stage produce variations of phase angle of the input impedance, phase modulation will result.

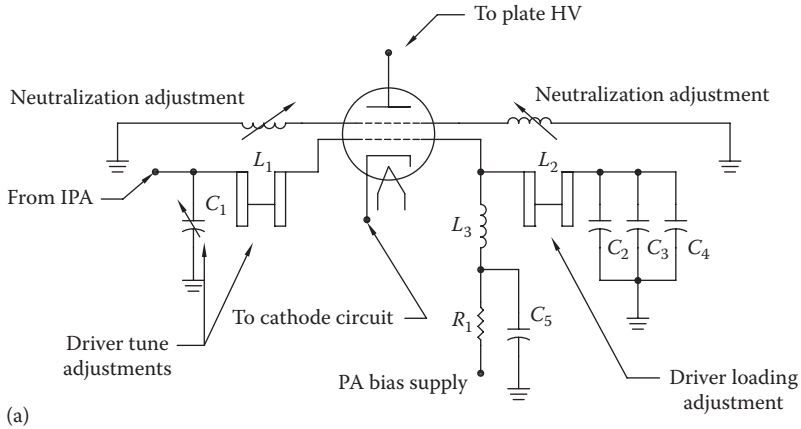
As noted previously, a circuit exhibiting independence between the input and output circuits is only half of the equation required for proper operation at radio frequencies. The effects of incidental inductance of the control grid also must be canceled for complete stability. This condition is required because the suppression of coupling by capacitive currents between the input and output circuits is not, by itself, sufficient to negate the effects of the output signal on the cathode-to-grid circuit. Both conditions—input and output circuit independence and compensation for control grid lead inductance—must be met for complete stage stability at VHF and above.

Figure 3.48 shows a PA stage employing stray inductance of the screen grid to achieve neutralization. In this grounded-screen application, the screen is bonded to the cavity deck using six short connecting straps. Two additional adjustable ground straps are set to achieve neutralization.

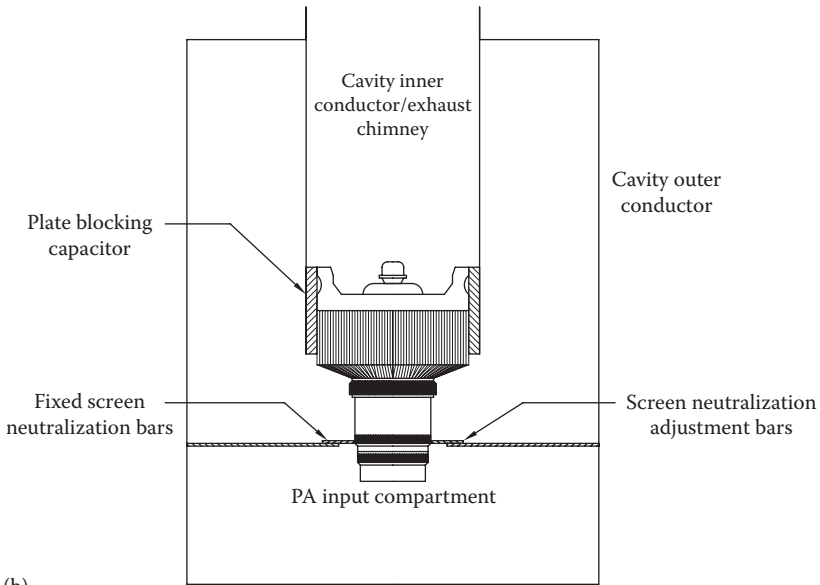
3.7.3 Grounded-Grid Amplifier Neutralization

Grounded-grid amplifiers offer an attractive alternative to the more common grid-driven circuit [4]. The control grid is operated at RF ground and serves as a shield to capacitive currents from the output to the input circuit. Generally speaking, neutralization is not required until the control grid lead inductive reactance becomes significant. The feedback from the output to the input circuit is no longer the result of plate-to-filament capacitance. The physical size of the tube and the operating frequency determine when neutralization is required.

Two methods of neutralization commonly are used with grounded-grid amplifiers. In the first technique, the grids of a push-pull amplifier are connected to a point having zero impedance to ground, and a bridge of neutralizing capacitances is used that is equal to the plate-filament capacitances of the tubes.



(a)



(b)

FIGURE 3.48

Grounded-screen PA stage neutralized through the use of stray inductance between the screen connector and the cavity deck: (a) schematic diagram and (b) mechanical layout of cavity.

The second method of neutralization requires an inductance between the grid and ground, or between the grids of a push-pull amplifier, of a value that will compensate for the coupling of a push-pull input and output circuits resulting from the internal capacitances of the tubes.

Behavior of these two circuits is quite different. They may be considered as special forms of the more general case in which the neutralizing capacitors have values differing from the internal capacitances of the tubes, and in which an appropriate reactance is connected between the grids. Under these

conditions, the value of neutralizing capacitance permits continuous variation of power amplification, stability, and negative feedback.

3.7.4 Grid impedance

In the special case of a grounded-grid amplifier having a grid impedance and the reactive currents neutralized, the following equations apply (see Figure 3.49):

$$C_n = C_{fp} - \frac{C_{fg}}{\mu} \quad (3.14)$$

$$Z_g = -\frac{1}{j\omega C_{fg} + C_{gp}(1 + \mu)} \quad (3.15)$$

If in solving the equation for C_n the sign is negative, this indicates that in-phase neutralization is required [4]. Conversely, if the sign of C_n is positive, then out-of-phase neutralization is needed. A negative value of Z_g indicates capacitive reactance required, and a positive value indicates that inductive reactance is to be used.

3.7.4.1 Application Example

If the grids of a push-pull cathode-driven amplifier are not at ground potential because the inductance of the leads is not negligible, coupling may exist between the input and output circuits through the plate-grid capacitances,

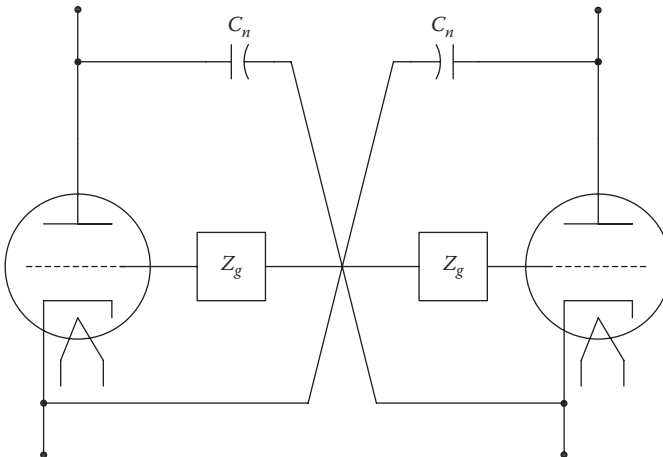
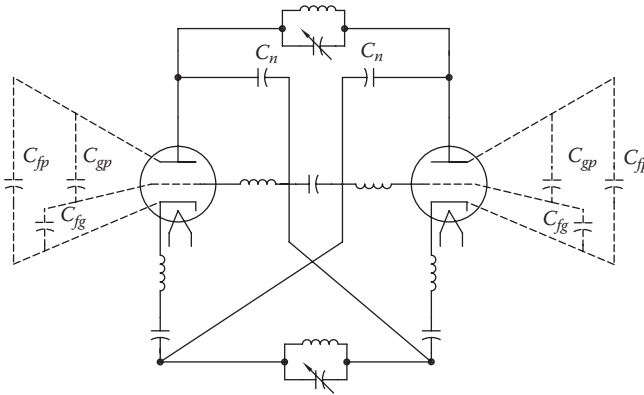


FIGURE 3.49

Circuit of grounded-grid amplifier having grid impedance and neutralized for reactive currents. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

**FIGURE 3.50**

Neutralization by cross-connected capacitors of a symmetrical cathode-excited amplifier with compensation of lead inductance. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

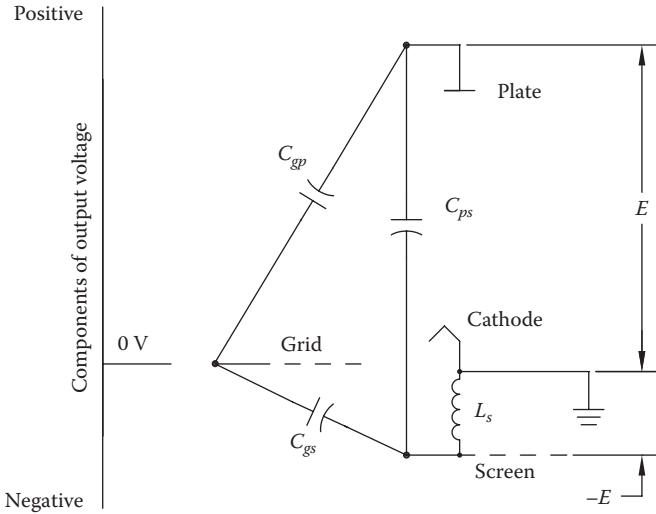
cathode-grid capacitances, and grid-to-grid inductance [4]. One method of reducing this coupling is to insert between the grids a series-tuned circuit that has zero reactance at the operating frequency. This technique is illustrated in Figure 3.50. This neutralization scheme is useful only for the case in which no grid current flows. If grid current flows, a grid resistance will appear in parallel with the grid-to-filament capacitance. If the resistance is small in comparison with the reactance of this grid-to-filament capacitance, phase modulation will result.

Another important property of the preceding neutralization scheme is that power amplification is a function of the neutralizing capacitance, while the independence of cathode and plate circuits from the viewpoint of reactive currents may be obtained with any value of neutralizing capacitance. If the neutralizing capacitance is less than the plate-to-filament capacitance of the tube, the stage will operate with low excitation power and high power amplification.

If the neutralizing capacitance is greater than the plate-to-filament capacitance, the power amplification will be quite low, but the total output power possible will be increased.

3.7.5 Self-Neutralizing Frequency

The voltage-dividing action between the plate-to-grid capacitance (C_{pg}) and the grid-to-screen capacitance (C_{gs}) will not change with variations in operating frequency [4]. The voltage division between the plate and screen, and screen and ground caused by the charging current (I) will, however, vary significantly with frequency. There will be a particular frequency, therefore, at which this potential-dividing circuit will effectively place the grid at filament potential insofar as the plate is concerned. This point is known as the *self-neutralizing frequency*, illustrated in Figure 3.51.

**FIGURE 3.51**

Graphical representation of the elements of a tetrode when self-neutralized. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

At the self-neutralizing frequency, the tetrode or pentode is inherently neutralized by the circuit elements within the device itself, and external screen inductance to ground. When a device is operated below its self-neutralizing frequency, the normal cross-neutralization circuits apply. When the operating frequency is above the self-neutralizing frequency, the voltage developed in the screen lead inductance is too large to give the proper voltage division between the internal capacitances of the device. One approach to neutralization in this case involves adjusting the inductive reactance of the screen lead to ground to lower the total reactance. In the example shown in Figure 3.52, this is accomplished with a series variable capacitor.

Another approach is shown in Figure 3.53, in which the potential divider network made up of the tube capacitance is changed. In the example, additional plate-to-grid capacitance is added external to the tube. The external capacitance (C_{ext}) can take the form of a small wire or rod positioned adjacent to the plate of the tube. This approach is similar to the one described previously, except that in this case the neutralizing probe is connected to the grid of the tube, rather than to an opposite polarity in the circuit.

If the RF power amplifier is operating above the self-neutralizing frequency of the tube and must be tuned over a range of frequencies, it is probably easier to use the screen series-tuning capacitor method and make this control available to the operator. If operation is desired over a range of frequencies including the self-neutralizing frequency of the tube, this circuit is also desirable because the incidental lead inductance in the variable capacitor lowers the self-neutralizing frequency of the circuit so that the neutralizing series capacitor can be made to operate over the total desired frequency range. If this range is

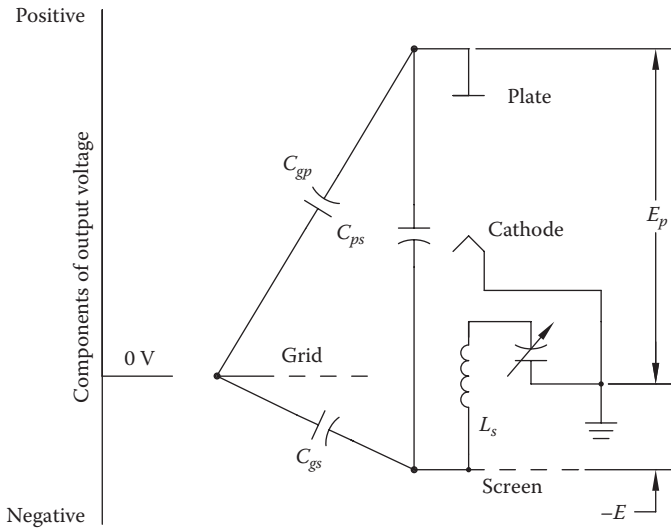


FIGURE 3.52 Components of the output voltage of a tetrode when neutralized by adding series screen-lead capacitance. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

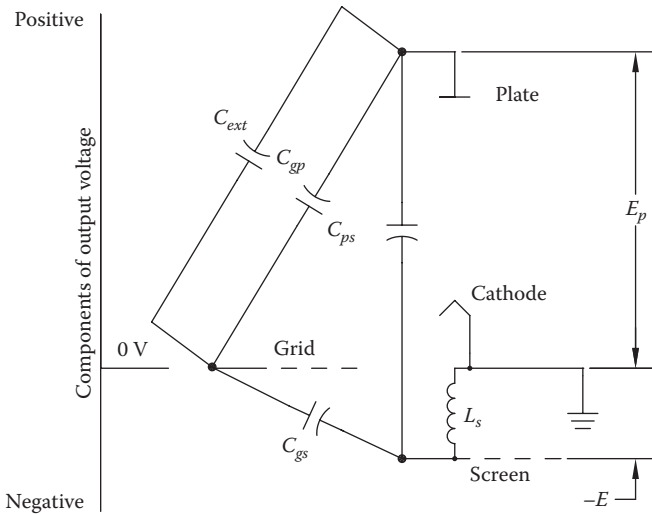
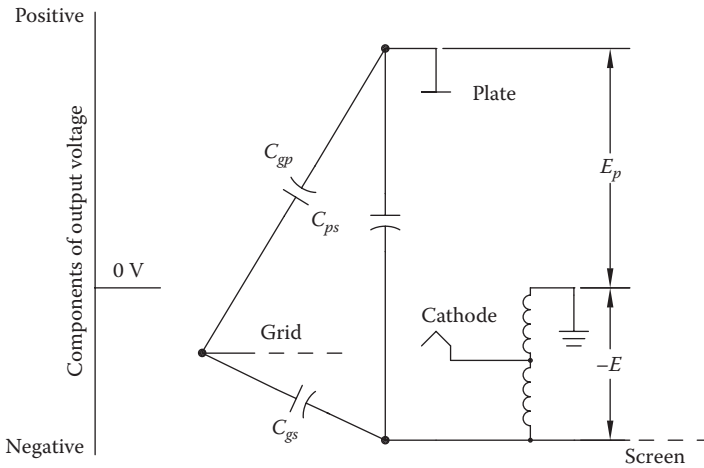


FIGURE 3.53 Components of the output voltage of a tetrode when neutralized by adding external grid-to-plate capacitance. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

**FIGURE 3.54**

Components of the output voltage of a tetrode neutralized by adding inductance common to the screen and cathode return. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

too great, switching of neutralizing circuits will be required. A small 50–100 pF variable capacitor in the screen lead often has been found to be satisfactory.

Another method of changing the self-neutralizing frequency of a tetrode or pentode can be fashioned from the general bypassing arrangement of the screen and filament shown in Figure 3.54. The screen lead is bypassed with minimum inductance to the filament terminal of the tube. Some inductance is introduced in the common filament and screen ground leads. The grid is shown below the zero voltage or chassis potential, indicating that the voltage developed in the screen lead inductance to chassis is excessive. If the filament is tapped on this inductance, a point can be found where the voltage difference between the grid and filament is zero, as far as the components of plate voltage are concerned. This arrangement will be found to self-neutralize at a higher frequency than if the filament and screen were bypassed separately to the chassis. Thus, increasing the self-neutralizing frequency of the tube and screen bypass arrangement reduces the tendency of the VHF parasitic to occur.

If the frequency of the VHF parasitic is reduced by increasing the inductance of the plate lead (presuming this is the principal frequency-defining circuit), the circuit can be made to approach the self-neutralizing frequency of the tube and, therefore, suppress the parasitic.

3.7.6 Neutralization Adjustment

Most neutralization circuits must be adjusted for operation at a given frequency [4]. The exact procedure followed to make these adjustments varies from one circuit to the next. The following generalizations, however, apply to most systems.

The first step in the process of neutralization is to break the dc connections to the plate voltage and screen voltage supplies, leaving the RF circuits intact. If the direct current path is not broken, some current can flow in either of these circuits even though the voltages are zero. The presence of this current causes the amplifier to work in the normal manner, generating RF power in the plate circuit. It would then be incorrect to adjust for zero power in the plate circuit. Sufficient RF grid drive must be applied to provide some grid current or to cause a sensitive RF meter coupled to the plate to give an indication of feedthrough power. When the plate circuit is tuned through resonance, the grid current will dip when the circuit is out of neutralization, or the RF meter will peak. The neutralization adjustments are made until the indication is minimum.

Another powerful tool for roughly neutralizing an RF amplifier is to feed the power output from a signal generator into the grid circuit. A sensitive RF detector is inserted between the output connector and the load. Neutralization can then be adjusted for minimum feedthrough. This technique is useful in working with prototype equipment. Actual qualitative measurements can be made. If the insertion loss of the amplifier is less than the expected gain, oscillation will occur. Circuit modifications can be made until the isolation is sufficient to warrant a test with high voltages applied. The advantages of this “cold” system test are that

- No components are subjected to unusual stress if the amplifier is unstable.
- Circuit adjustments can be made safely because no high voltages are present.

For final trimming of the neutralization adjustment, the stage should be returned to operating condition at reduced power (similar to that used when testing for parasitic oscillations), or under the final loaded operating conditions. At higher frequencies, particularly in the VHF region, it will be found that a small additional trimming adjustment of the neutralization circuit is usually required. When the plate circuit is tuned through resonance, minimum plate current and maximum control grid current should occur simultaneously. In the case of the tetrode and pentode, the dc screen current should be maximum at the same time.

These neutralizing procedures apply not only to the HF radio frequencies, but also to the VHF and UHF regions.

3.8 Electron Tube Characteristics

The term “characteristics” is used to identify the distinguishing electrical features and values of an electron tube [3]. These values may be shown in curve form or they may be tabulated (or both). When the characteristic values

are given in curve form, the curves may be used to determine tube performance and to calculate additional factors. Tube characteristics are obtained from electrical measurements of a device in various circuits under defined conditions.

Static characteristics may be shown by plate characteristics curves and transfer (mutual) characteristics curves. These curves present the same information but in two different forms to increase their usefulness:

- Plate characteristics curves are obtained by varying the plate voltage and measuring plate current for different grid-bias voltages.
- Transfer characteristics curves are obtained by varying the grid-bias voltage and measuring plate current for different plate voltages.

A plate characteristics family of curves is shown in Figure 3.55, while Figure 3.56 illustrates the transfer characteristics family of curves for the same tube.

Dynamic characteristics include amplification factor, plate resistance, control-grid plate transconductance, plate efficiency, power sensitivity, and other parameters. These characteristics may be shown in curve form for variations in specified operating conditions. Specifically,

- *Amplification factor* (μ): the ratio of the change in plate voltage to a change in control-electrode voltage in the opposite direction, under the conditions that the plate current remains unchanged and that all other electrode voltages are maintained constant. The μ of a device is useful for calculating stage gain.

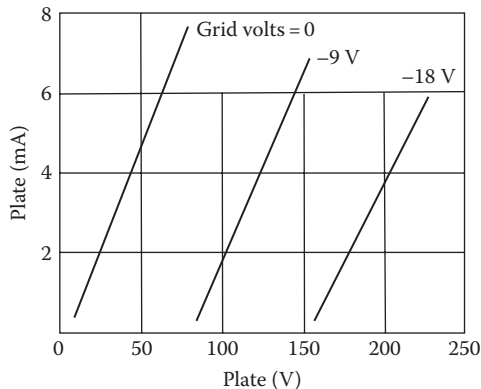
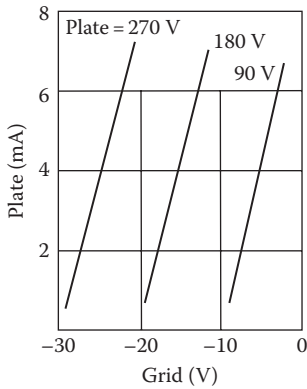


FIGURE 3.55

Family of plate characteristics curves. (After *RCA Receiving Tube Manual*, RC-30, Radio Corporation of America, Camden, NJ, 1975.)

**FIGURE 3.56**

Family of transfer characteristics curves. (After *RCA Receiving Tube Manual*, RC-30, Radio Corporation of America, Camden, NJ, 1975.)

Example: if, when the plate voltage is made 1 V more positive, the control-electrode (grid #1) voltage must be made 0.1 V more negative to hold plate current unchanged, the amplification factor is 1 divided by 0.1 or 10. In other words, a small voltage variation in the grid current of the tube has the same effect on plate current as a large plate voltage change—the latter equal to the product of the grid voltage change and the amplification factor.

- *Plate resistance* (r_p): the resistance of the path between the cathode and plate to the flow of alternating current. This parameter is the quotient of a small change in plate voltage divided by the corresponding change in plate current and is expressed in ohms.

Example: if a change of 0.1 mA (0.0001 A) is produced by a plate voltage variation of 1 V, the plate resistance is 1 divided by 0.0001 or 100,000 Ω .

- *Transconductance* (g_m): a factor that combines in one term the amplification factor and the plate resistance, and is the quotient of the first divided by the second. More correctly described as control grid-to-plate transconductance, this term has also been known as *mutual conductance*. Transconductance may be more strictly defined as the quotient of a small change in plate current (amperes) divided by the small change in control-grid voltage producing it, under the condition that all other voltages remain unchanged.

Example: if a grid voltage change of 0.5 V causes a plate current change of 1 mA (0.001 A), with all other voltages held constant, the transconductance is 0.001 divided by 0.5 or 0.005 mho (“mho” is the unit of conductance). For convenience, a millionth of a mho, or a micromho (μmho) is commonly used to express transconductance. Thus, in this example 0.002 mho is equal to 2000 μmho .

- *Plate efficiency*: the ratio of the ac power output (P_o) of an amplifier tube to the product of the average dc plate voltage (E_b) and the dc plate current (I_b) at full signal level:

$$\text{Plate efficiency (\%)} = \frac{P_o}{E_b \times I_b} \times 100 \quad (3.16)$$

where

P_o is the ac power output in W
 E_b is the average dc plate voltage in V
 I_b is the dc plate current in A

- *Power sensitivity*: the ratio of the power output (P_o) of an amplifier tube to the square of the input signal voltage (E_{in}), expressed in mhos:

$$\text{Power sensitivity (mhos)} = \frac{P_o}{E_{in}^2} \quad (3.17)$$

where

P_o is the ac power output in W
 E_{in} is the input signal voltage (rms)

3.8.1 Interpretation of Tube Data

The data provided by electron-tube manufacturers include typical operating values, characteristics, and characteristic curves [3]. Ratings are established on electron-tube types to help equipment designers utilize the performance and service capabilities of each tube type to the best advantage. Ratings are given for those characteristics that careful study and experience indicate must be kept within certain limits to insure satisfactory performance.

Various rating systems have been used by the electron-tube industry over the years, including the following:

- *Absolute maximum*: limiting values that should not be exceeded with any tube of the specified type under any condition of operation. These ratings are not used too often for receiving tubes, but are generally used for transmitting and industrial tubes.
- *Design center*: limiting values that should not be exceeded with a tube of the specified type having characteristics equal to the published values under normal operating conditions. These ratings, which include

allowances for normal variations in both tube characteristics and operating conditions, were used for most receiving tubes prior to 1957.

- *Design maximum*: limiting values that should not be exceeded with a tube of the specified type having limiting characteristics equal to the published values under any conditions of operation. These ratings include allowances for normal variations in tube characteristics but do not provide for variations in operating conditions. Design maximum ratings were adopted for receiving tubes in 1957.

Many ratings, such as voltage and current ratings, are in general self-explanatory, but additional detail is useful:

- *Plate dissipation*: the power dissipated in the form of heat by the plate as a result of electron bombardment. This value is the difference between the power supplied to the plate of the tube and the power delivered by the tube to the load.
- *Peak heater-cathode voltage*: the highest instantaneous value of voltage that a tube can safely stand between its heater and cathode. This rating is applied to tubes having a separate cathode terminal and used in applications where excessive voltage may be introduced between the heater and the cathode.
- *Maximum dc output current*: the highest average plate current that can be handled continuously by a rectifier tube. This value for any rectifier is based on the permissible plate dissipation of that tube type. Under operating conditions involving a rapidly repeating duty cycle (steady load), the average plate current may be measured with a dc ammeter.
- *Maximum peak plate current*: the highest instantaneous plate current that a tube can safely carry recurrently in the direction of normal current flow.
- *Maximum peak inverse plate voltage*: the highest instantaneous plate voltage that the tube can withstand recurrently in the direction opposite to that in which it is designed to pass current.
- *Typical operation values*: values for typical operation given to serve as guiding information for the use of each tube type. These values should not be confused with *ratings*, because a tube can be used under any suitable conditions within its maximum ratings, according to the application. The power output value for any operating condition is an approximate tube output; that is, plate input minus plate loss. Circuit losses must be subtracted from the tube output in order to determine the useful output.
- *Interelectrode capacitance*: discrete capacitances measured between specified elements or groups of elements in electron tubes. Unless otherwise

indicated, all capacitances are measured with filament or heater cold, with no direct voltages present, and with no external shields installed.

- *Grid #2 (screen grid) input*: the power applied to the grid #2 electrode and consists essentially of the power dissipated in the form of heat by the grid as a result of electron bombardment.

References

1. Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.
2. Ferris, C. D., Electron tube fundamentals, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 295–305, 1996.
3. Radio Corporation of America, *RCA Receiving Tube Manual*, RC-30, Camden, NJ, 1975.
4. Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.
5. Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.

Bibliography

- Birdsall, C. K., *Plasma Physics via Computer Simulation*, Adam Hilger, New York, 1991.
- Block, R., CPS microengineers new breed of materials, *Ceramic Industry*, 51–53, April 1988.
- Buchanan, R. C., *Ceramic Materials for Electronics*, Marcel Dekker, New York, 1986.
- Ceramic Products, Brush Wellman, Cleveland, OH.
- Chaffee, E. L., *Theory of Thermionic Vacuum Tubes*, McGraw-Hill, New York, 1939.
- Combat boron nitride, solids, powders, coatings, Carborundum Product Literature, form A-14, 011, September 1984.
- Coors Ceramics—Materials for Tough Jobs, Coors Data Sheet K.P.G.-2500-2/87 6429.
- Cote, R. E. and R. J. Bouchard, Thick film technology, in *Electronic Ceramics*, L. M. Levinson (ed.), Marcel Dekker, New York, pp. 307–370, 1988.
- Dettmer, E. S. and H. K. Charles, Jr., AlNi and SiC substrate properties and processing characteristics, *Advances in Ceramics*, 31, 1989.
- Dettmer, E. S., H. K. Charles, Jr., S. J. Mobley, and B. M. Romensko, Hybrid design and processing using aluminum nitride substrates, in *ISHM 88 Proceedings*, Seattle, WA, pp. 545–553, 1988.
- Eastman, A. V., *Fundamentals of Vacuum Tubes*, McGraw-Hill, New York, 1941.
- Floyd, J. R., How to tailor high-alumina ceramics for electrical applications, *Ceramic Industry*, 44–47, February 1969; 46–49, March 1969.
- Gray, T. S., *Applied Electronics*, John Wiley & Sons, New York, 1954.
- Harper, C. A., *Electronic Packaging and Interconnection Handbook*, McGraw-Hill, New York, 1991.

- High power transmitting tubes for broadcasting and research, Philips Technical Publication, Eindhoven, the Netherlands, 1988.
- Iwase, N. and K. Shinozaki, Aluminum nitride substrates having high thermal conductivity, *Solid State Technology*, 29, 135–138, October 1986.
- Jordan, E. C. (ed.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 7th edn., Howard W. Sams, Indianapolis, IN, 1985.
- Kingery, W. D., H. K. Bowen, and D. R. Uhlmann, *Introduction to Ceramics*, John Wiley & Sons, New York, pp. 637, 1976.
- Kohl, W., *Materials Technology for Electron Tubes*, Reinhold, New York, 1951.
- Lafferty, J. M., *Vacuum Arcs*, John Wiley & Sons, New York, 1980.
- Mattox, D. M. and H. D. Smith, The role of manganese in the metallization of high alumina ceramics, *Journal of the American Ceramic Society*, 64, 1363–1369, 1985.
- Mistler, R. E., D. J. Shanefield, and R. B. Runk, Tape casting of ceramics, in *Ceramic Processing Before Firing*, G. Y. Onoda, Jr. and L. L. Hench (eds.), John Wiley & Sons, New York, pp. 411–448, 1978.
- Muller, J. J., Cathode excited linear amplifiers, *Electrical Communications*, 23, 297–305, 1946.
- Powers, M. B., *Potential Beryllium Exposure While Processing Beryllia Ceramics for Electronics Applications*, Brush Wellman, Cleveland, OH.
- Reich, H. J., *Theory and Application of Electronic Tubes*, McGraw-Hill, New York, 1939.
- Roth, A., *Vacuum Technology*, 3rd edn., Elsevier Science Publishers B. V., Amsterdam, the Netherlands, 1990.
- Sawhill, H. T., A. L. Eustice, S. J. Horowitz, J. Gar-El, and A. R. Travis, Low temperature co-fireable ceramics with co-fired resistors, *Proceeding of the International Symposium on Microelectronics*, Atlanta, GA, pp. 173–180, 1986.
- Schwartz, B., Ceramic packaging of integrated circuits, in *Electronic Ceramics*, L. M. Levinson (ed.), Marcel Dekker, New York, pp. 1–44, 1988.
- Strong, C. E., The inverted amplifier, *Electrical Communications*, 19(3), 32–36, 1941.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1991.

4

Designing Vacuum Tube Systems

4.1 Introduction

Any number of configurations may be used to generate radio frequency (RF) signals using vacuum tubes. Circuit design is dictated primarily by the operating frequency, output power, type of modulation, duty cycle, and available power supply. Tube circuits can be divided generally by their operating class and type of modulation employed. The angle of plate current flow determines the class of operation:

- Class A = 360° conduction angle
- Class B = 180° conduction angle
- Class C = conduction angle less than 180°
- Class AB = conduction angle between 180° and 360°

The class of operation has nothing to do with whether the amplifier is grid driven or cathode driven [1]. A cathode-driven amplifier, for example, can be operated in any desired class. The class of operation is only a function of the plate current conduction angle. The efficiency of an amplifier is also a function of the plate current conduction angle.

The efficiency of conversion of dc to RF power is one of the more important characteristics of a vacuum tube amplifier circuit. The dc power that is not converted into useful output energy is, for the most part, converted to heat. This heat represents wasted power; the result of low efficiency is increased operating cost for energy. Low efficiency also compounds itself. This wasted power must be dissipated, requiring increased cooling capacity. The efficiency of the amplifier must, therefore, be carefully considered, consistent with the other requirements of the system. Figure 4.1 shows the theoretical efficiency attainable with a tuned or resistive load assuming that the peak ac plate voltage is equal to the plate supply voltage.

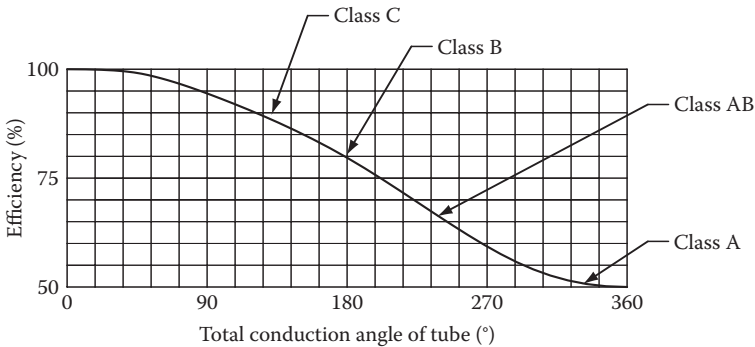


FIGURE 4.1

Plate efficiency as a function of conduction angle for an amplifier with a tuned load. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

4.1.1 Class A Amplifier

A class A amplifier is used in applications requiring low harmonic distortion in the output signal [1]. A class A amplifier can be operated with low intermodulation distortion in linear RF amplifier service. Typical plate efficiency for a class A amplifier is about 30%. Power gain is high because of the low drive power required. Gains as high as 30 dB are typical.

4.1.2 Class B and AB Amplifiers

A class AB power amplifier (PA) is capable of generating more power—using the same tube—than the class A amplifier, but more intermodulation distortion also will be generated [1]. A class B RF linear amplifier will generate still more intermodulation distortion, but is acceptable in certain applications. The plate efficiency is typically 66%, and stage gain is about 20–25 dB.

4.1.3 Class C Amplifier

A class C PA is used where large amounts of RF energy need to be generated with high efficiency [1]. Class C RF amplifiers must be used in conjunction with tuned circuits or cavities, which restore the amplified waveform through the *flywheel effect*.

The grounded-cathode class C amplifier is the building block of RF technology. It is the simplest method of amplifying continuous wave (CW), pulsed, and FM signals. The basic configuration is shown in Figure 4.2. Tuned input and output circuits are used for impedance matching and to resonate the stage at the desired operating frequency. The cathode is bypassed to ground using low-value capacitors. Bias is applied to the grid as shown. The bias power supply may be eliminated if a self-bias configuration is used. The typical operating efficiency of a class C stage ranges from 65% to 85%.

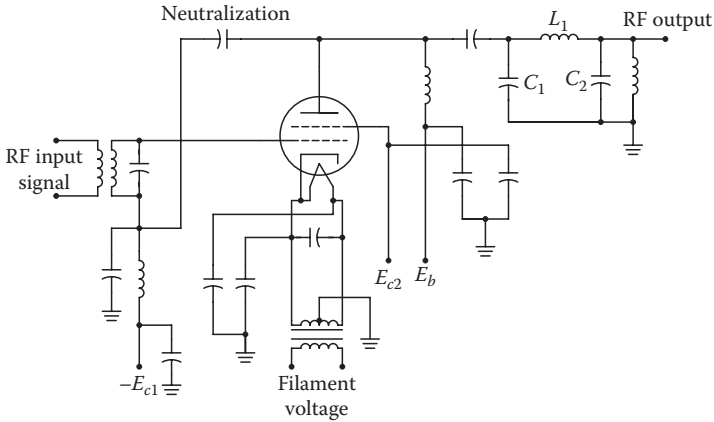


FIGURE 4.2 Basic grounded-cathode RF amplifier circuit. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

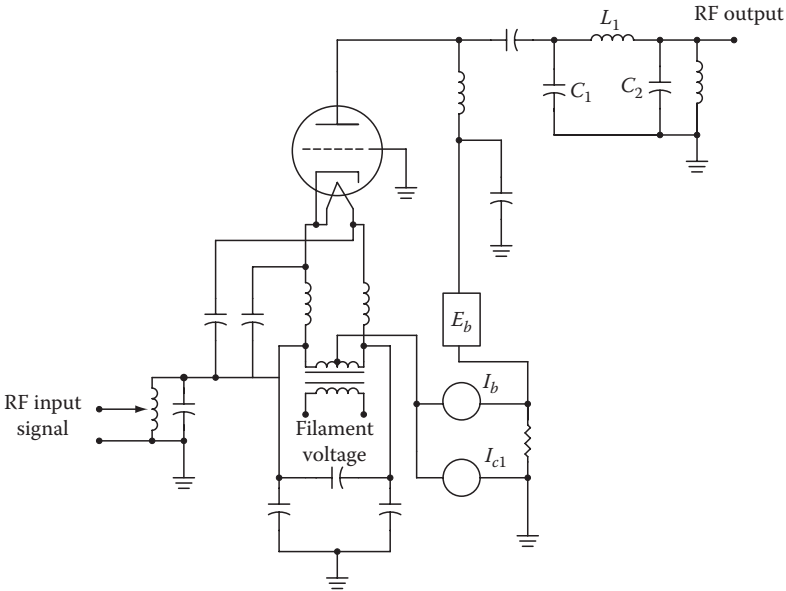


FIGURE 4.3 Typical amplifier circuit using a zero-bias triode. Grid current is measured in the return lead from ground to the filament. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

Figure 4.3 illustrates the application of a zero-bias triode in a grounded-grid arrangement. Because the grid operates at RF ground potential, this circuit offers stable performance without the need for neutralization (at MF and below). The input signal is coupled to the cathode through a matching network. The output of the triode feeds a π network through a blocking capacitor.

4.2 Principles of RF Power Amplification

In an RF PA, a varying voltage is applied to the control grid (or cathode, in the case of a grounded-grid circuit) from a driver stage whose output is usually one of the following [1]:

- Carrier-frequency signal only
- Modulation (intelligence) signal only
- Modulated carrier signal

Simultaneous with the varying control grid signal, the plate voltage will vary in a similar manner, resulting from the action of the amplified current flowing in the plate circuit. In RF applications with resonant circuits, these voltage changes are smooth sine wave variations, 180° out of phase with the input. The relationship is illustrated in Figure 4.4. Note how these variations center about the dc plate voltage and the dc control grid bias. In Figure 4.5, the variations have been indicated next to the plate voltage and grid voltage scales of a typical constant-current curve. At some instant in time, shown as t on the timescales, the grid voltage has a value denoted E_g on the grid voltage sine wave.

Any point on the operating line (when drawn on constant-current curves as illustrated in Figure 4.5) tells the instantaneous values of plate current,

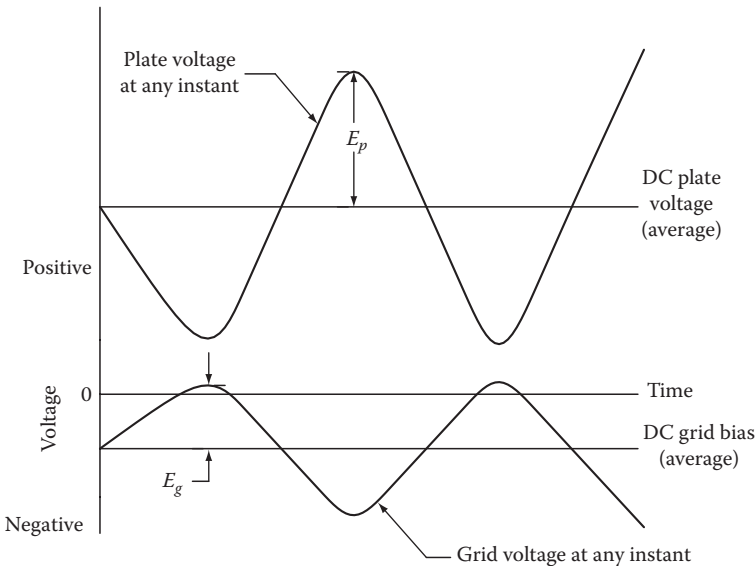


FIGURE 4.4

Variation of plate voltage as a function of grid voltage. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

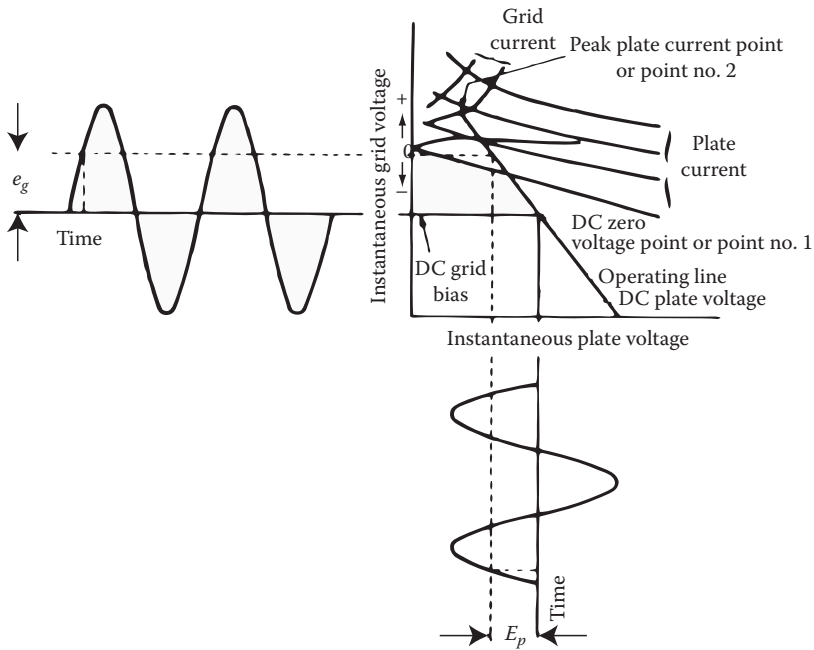


FIGURE 4.5

Relationship between grid voltage and plate voltage plotted on a constant-current curve. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

screen current, and grid current that must flow when these particular values of grid and plate voltage are applied to the tube. Thus, by plotting the values of plate and grid current as a function of time t , it is possible to produce a curve of instantaneous values of plate and grid current. Such plots are shown in Figure 4.6.

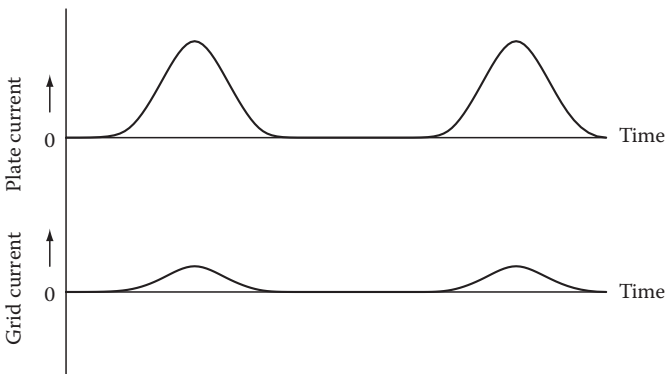


FIGURE 4.6

Instantaneous values of plate and grid current as a function of time. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

By analyzing the plate and grid current values, it is possible to predict with accuracy the effect on the plate circuit of a change at the grid. It follows that if a properly loaded resonant circuit is connected to the plate, a certain amount of RF power will be delivered to that circuit. If the resonant circuit is tuned to the fundamental frequency (the same frequency as the RF grid voltage), the power delivered will be that of the fundamental or principal RF component of plate current. If the circuit is tuned to a harmonic of the grid voltage frequency, the power delivered will be the result of a harmonic component of the plate current.

4.2.1 Drive Power Requirements

The technical data sheet for a given tube type lists the approximate drive power required for various operating modes [1]. As the frequency of operation increases and the physical size of the tube structure becomes large with respect to this frequency, the drive power requirements also will increase.

The drive power requirements of a typical grounded-cathode amplifier consist of six major elements:

- The power consumed by the bias source, given by

$$P_1 = I_{c1} \times E_{c1} \quad (4.1)$$

- The power dissipated in the grid as a result of rectified grid current

$$P_2 = I_{c1} \times e_{cmp} \quad (4.2)$$

- The power consumed in the tuned-grid circuit

$$P_3 = I_{c,rms}^2 \times R_{rf} \quad (4.3)$$

- The power lost as a result of transit-time effects

$$P_4 = \left\{ \frac{e_{c,rms}}{R_t} \right\}^2 \quad (4.4)$$

R_t is that part of the resistive component of the tube input impedance resulting from transit-time effects and is given by

$$R_t = \frac{1}{Kg_m f^2 T^2} \quad (4.5)$$

- The power consumed in that part of the resistive component of the input impedance resulting from cathode lead inductance

$$P_5 = \frac{e_g^2}{R_s} \quad (4.6)$$

Input resistance resulting from the inductance of the cathode leads is found from the following:

$$R_s = \frac{1}{\omega^2 g_m L_k C_{gk}} \quad (4.7)$$

- The power dissipated in the tube envelope because of dielectric loss

$$P_6 = 1.41 f E_1^2 \epsilon \quad (4.8)$$

where

I_{c1} is the dc grid current

E_{c1} is the dc grid voltage

e_{cmp} is the maximum positive grid voltage

$I_{c,rms}$ is the rms value of RF grid current

R_{rf} is the RF resistance of grid circuit

$e_{c,rms}$ is the rms value of RF grid voltage

R_t is the resistance resulting from transit-time loading

K is a constant (function of the tube geometry)

g_m is the transconductance

f is the frequency (Hz)

T is the transit-time, cathode to grid

R_s is the cathode lead inductance input resistance loading

$\omega = 2\pi f$

L_k is the cathode lead inductance (H)

C_{gk} is the grid-to-cathode capacitance (F)

E_1 is the voltage gradient (kV/in., rms)

ϵ is the loss factor of dielectric materials

The total drive power P_t is, then, equal to

$$P_t = P_1 + P_2 + P_3 + P_4 + P_5 + P_6 \quad (4.9)$$

Particular attention must be given to grid dissipation when a tube is operated in the very high frequency (VHF) and ultrahigh frequency (UHF) regions. The total driving power required for a given output may be greater than the grid dissipation capability of the device.

4.2.1.1 Operational Considerations for VHF and UHF

For operation of a tube in the VHF and UHF regions, several techniques may be applied to minimize the driving power without appreciably affecting plate conversion efficiency [1]. The most common techniques are as follows:

- Use the minimum dc control bias. Frequently, it is advisable to bring the bias down to approximately cutoff.
- Maintain a high value of dc screen voltage, even though it appears to increase the fraction of the cycle during which plate current flows.

- Use the minimum RF excitation voltage necessary to obtain the desired plate circuit performance, even though the dc grid current is considerably lower than would be expected at lower frequencies.
- Keep the cathode lead inductance to the output and input circuits as low as possible. This can be accomplished by (1) using short and wide straps, (2) using two separate return paths for the input and output circuits, or (3) properly choosing a cathode bypass capacitor(s).

These techniques do not necessarily decrease the plate efficiency significantly when the circuit is operated at VHF and UHF. The steps should be tried experimentally to determine whether the plate circuit efficiency is appreciably affected. It is usually acceptable, and even preferable, to sacrifice some plate efficiency for improved tube life when operating at VHF and UHF.

Optimum power output at these frequencies is obtained when the loading is greater than would be used at lower frequencies. Fortunately, the same condition reduces driving power and screen current (for the tetrode and pentode) and improves tube life expectancy in the process.

4.2.2 Mechanical and Electrical Considerations

To maintain proper isolation of the output and input circuits, careful consideration must be given to the location of the component parts of the amplifier [1]. All elements of the grid or input circuit and any earlier stages must be kept out of the plate circuit compartment. Similarly, plate circuit elements must be kept out of the input compartment. Note, however, that in the case of the tetrode and pentode, the screen lead of the tube and connections via the socket are common to both the output and input resonant circuits. Because of the plate-to-screen capacitance of a tetrode or pentode, the RF plate voltage (developed in the output circuit) causes an RF current to flow out of the screen lead to the chassis. In the case of a push-pull stage, this current may flow from the screen terminal of one tube to the screen terminal of the other tube. Similarly, because of the grid-to-screen capacitance of the tube, the RF voltage in the input circuit will cause an RF current to flow in this same screen lead to the chassis, or to the opposite tube of the push-pull circuit.

The inductance of the lead common to both the output and input circuits has the desirable feature of providing voltage of opposite polarity to neutralize the feedback voltage of the residual plate to control grid capacitance of the tube. Note, however, that the mutual coupling from the screen lead to the input resonant circuit may be a possible source of trouble, depending on the design.

4.2.2.1 Lead Lengths

The interconnecting lead wires close to the tube should be designed with low inductance to minimize the generation of VHF parasitic oscillations [1]. If two or more tubes are used in a given circuit, they should be placed reasonably close together to facilitate short interconnecting leads. The lead lengths of HF circuits usually can be much longer; the length depends, to a large extent, upon the frequency of the fundamental. All of the dc, keying, modulation, and control circuit wires can be relatively long if properly filtered and arranged physically apart from the active RF circuits. The following interconnecting lead wires in a tetrode or pentode PA should have low inductance:

- Filament and screen bypass leads
- Suppressor bypass leads
- Leads from the grid and the plate to the tuning capacitor of the RF circuit (and return)
- Interconnections from tube to tube in push-pull or parallel arrangements (except for parasitic suppressors in the plate)

For a lead to have low inductance, it must have a large surface and be short in length, such as a strap or a ribbon. This consideration also applies to that portion of a lead inside a bypass capacitor.

4.2.2.2 Power Supply Considerations

The power supply requirements for a triode are straightforward [1]. The degree of regulation and ripple depends upon the requirements of the system. In the case of a linear RF amplifier, it is important to have good plate power supply regulation. Without tight regulation, the plate voltage will drop during the time the plate is conducting current heavily. This drop will cause *flat topping* and will appear as distortion in the output. In push-pull applications where grid current flows, it is important to keep the grid circuit resistance to a minimum. If this is not done, positive peak clipping will occur.

In the case of the tetrode and pentode, the need for screen voltage introduces some new considerations and provides some new operating possibilities. Voltage for the screen grid of a low-power tetrode or pentode can readily be taken from the power supply used for the plate of the tube. In this case, a series resistor or potential-dividing resistor is chosen so that, with the intended screen current flowing, the voltage drop through the resistor is adequate to give the desired screen voltage. The potential-dividing resistor is the preferred technique for those tubes with significant secondary screen emission.

For high-power tubes, screen voltage commonly is taken from a separate power supply. A combination scheme also may be employed, where a dropping resistor is used in conjunction with a low-voltage or intermediate-voltage supply. Often, a combination of series resistor and voltage source can be chosen so that the rated screen dissipation will not be exceeded, regardless of variations in screen current. With a fixed screen supply, there are advantages in using an appreciable amount of fixed grid bias so as to provide protection against the loss of excitation, or for cases where the driver stage is being keyed.

If the screen voltage is taken through a dropping resistor from the plate supply, there is usually little point in using a fixed grid bias because an unreasonable amount of bias would be required to protect the tube if the excitation failed. Under operating conditions with normal screen voltage, the cutoff bias is low (screen voltage divided by the screen μ). When a stage loses excitation and runs statically, the screen current falls to nearly zero. (See the static curves of the tube in question.) If the screen voltage is obtained through a simple dropping resistor from the plate supply, the screen voltage will then rise close to full plate voltage. Because the cutoff bias required is proportional to screen voltage, the grid bias required will be much greater than the amount of bias desired under normal operating conditions. When a screen dropping resistor is used, most of the bias normally is supplied through a grid resistor, and other means are used for tube protection.

The power output from a tetrode or pentode is sensitive to screen voltage. For this reason, any application requiring a high degree of linearity through the amplifier requires a well-regulated screen power supply. A screen dropping resistor from the plate supply is not recommended in such cases.

The suppressor grid power supply requirements are similar to those of the control grid power supply. The suppressor grid intercepts little current, so a low-power supply may be used. Any variation in suppressor voltage as a result of ripple or lack of regulation will appear at the output of the amplifier because of suppressor grid modulation of the plate current.

4.2.3 Bypassing Tube Elements

Operation at high frequencies requires attention to bypassing of the tube elements. Areas of concern include

- Filament circuit
- Screen grid circuit
- Suppressor grid circuit

4.2.3.1 Filament Bypassing

Low-inductance bypass capacitors should be used at the filament in an RF amplifier [1]. Good engineering practice calls for placement of the capacitor directly between the filament socket terminals. If the circuit design allows it,

strap one filament directly to the chassis; if not, use a second bypass capacitor from one terminal to the chassis.

If two or more tubes are used in a push-pull or parallel circuit, a short strap interconnecting one of the filament terminals of each socket can be used. The midpoint of the interconnecting strap is then bypassed or grounded directly.

For tubes equipped with an isolating screen cone terminal, the general circuit arrangement is usually different. The filament or cathode should go directly, or through bypass capacitors, to the cavity wall or chassis to which the screen terminal is bypassed.

4.2.3.2 Screen and Suppressor Grid Bypassing

Low-inductance leads are generally advisable for screen and suppressor grid terminal connections [1]. For all frequencies, good engineering practice calls for routing of the screen bypass capacitors directly from the screen to one filament terminal. The suppressor grid is bypassed in the same manner when the suppressor is operated with a potential other than cathode potential. With the suppressor operating at cathode potential, the suppressor should be grounded to the chassis directly in a circuit where the cathode is at chassis potential. This applies to tubes in push-pull as well as tubes in single-ended stages.

In the VHF region, the connection to the screen terminals—for those tubes with two screen pins—should be made to the midpoint of a strap placed between the two screen terminals of the socket. This arrangement provides for equal division of the RF currents in the screen leads and minimizes heating effects.

At operating frequencies above the self-neutralizing frequency of the tetrode or pentode being used, the screen bypass capacitors are sometimes variable. By proper adjustment of this variable capacitor, the amount and phase of the screen RF voltage can be made to cancel the effects of the feedback capacitance within the tube. Thus, neutralization is accomplished. The screen lead inductance and the variable capacitor are not series resonant. The variable capacitor is adjusted so that a net inductive reactance remains to provide the proper voltage and phase for neutralization.

The preceding guidelines apply directly to tubes having the screen and suppressor grids mounted on internal supporting lead rods. Tube types having isolating screen cone terminals tend to work best when the screen or suppressor bypass capacitor is a flat sandwich type built directly onto the peripheral screen-contacting collet of the socket. The size of the bypass capacitor is a function of the operating frequency. The dielectric material may be of Teflon, mica, isomica, or a similar material. (Teflon is a trademark of DuPont.)

4.2.3.3 Application Examples

The use of bypassing capacitors for a grounded-cathode RF amplifier is illustrated in Figure 4.7 [1]. Bypassing for a grounded-grid amplifier is shown

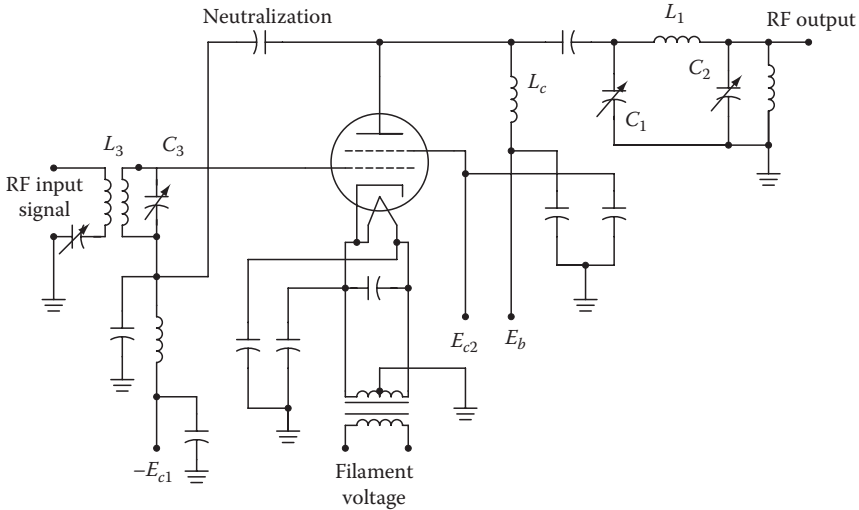


FIGURE 4.7

Schematic of a tetrode RF power amplifier showing tube element bypass capacitors. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

in Figure 4.8. The control grid may be operated at dc ground potential, shown in Figure 4.8a, or bypassed directly at the socket, shown in Figure 4.8b. For the sake of simplicity, operating the grid at dc ground potential is preferred. Functionally, these two circuits are identical. They vary only in the method used to measure grid current. In Figure 4.8a, grid current is measured in the return lead from ground to filament. In Figure 4.8b, the grid is raised 1 Ω above dc ground to allow grid current to be measured.

4.2.4 Parasitic Oscillations

Most self-oscillations in RF PA using gridded tubes have been found to fall within the following three classes [1]:

- Oscillation at VHF from about 40–200 MHz, regardless of the operating frequency of the amplifier
- Self-oscillation on the fundamental frequency of the amplifier
- Oscillation at a low RF below the normal frequency of the amplifier

Low-frequency oscillation in an amplifier usually involves the RF chokes, especially when chokes are used in both the output and input circuits. Oscillation near the fundamental frequency involves the normal resonant circuits and brings up the question of neutralization. When a parasitic self-oscillation is found on a VHF, the interconnecting leads of the tube, tuning capacitor, and bypass capacitors typically are involved.

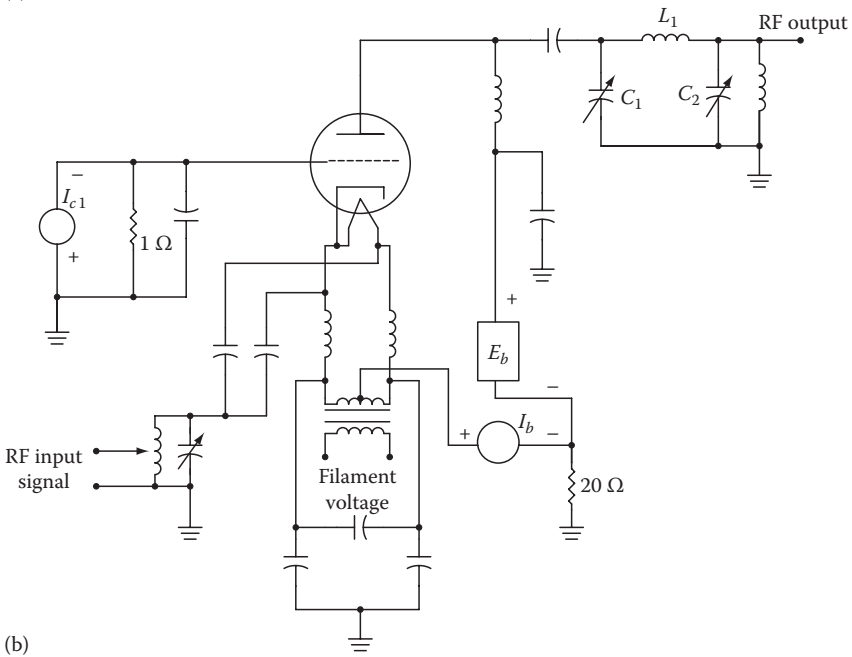
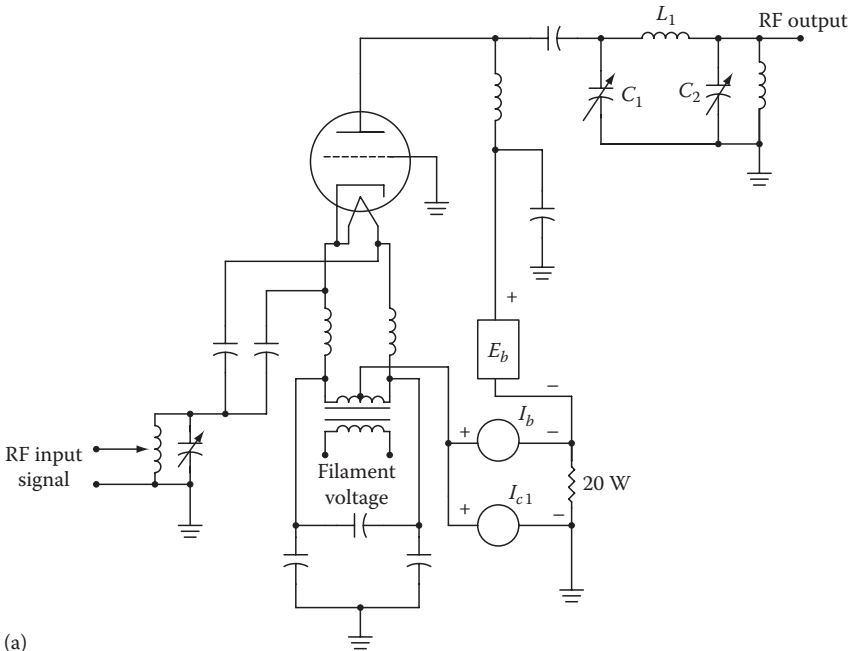
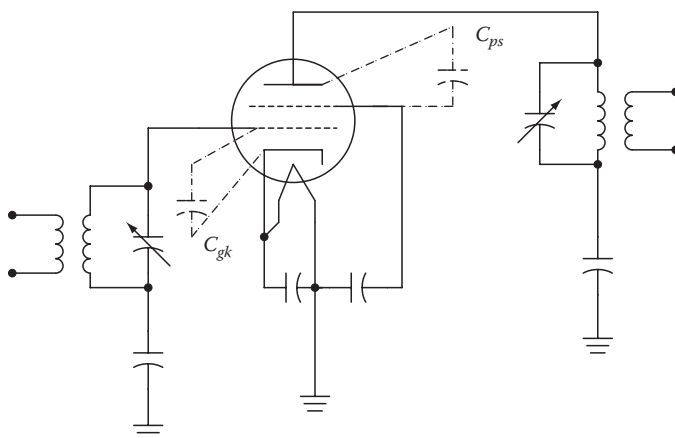


FIGURE 4.8

Schematic of a triode RF power amplifier showing tube element bypass capacitors: (a) control grid at dc ground potential and (b) control grid bypassed to ground. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

**FIGURE 4.9**

Interelectrode capacitances supporting VHF parasitic oscillation in a high-frequency RF amplifier. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

VHF oscillation occurs commonly in amplifiers where the RF circuits consist of coils and capacitors, as opposed to cavity designs. As illustrated in Figure 4.9, the tube capacitances effectively form a tuned-plate, tuned-grid oscillator.

The frequency of a VHF parasitic typically is well above the self-neutralizing frequency of the tube. However, if the self-neutralizing frequency of the device can be increased and the frequency of the parasitic lowered, complete suppression of the parasitic may result, or its suppression by resistor-inductor parasitic suppressors may be made easier.

It is possible to predict, with the use of a grid dip wavemeter, the parasitic frequency to be expected in a given circuit. The circuit should be complete and have no voltages applied to the tube. Couple the meter to the plate or screen lead, and determine the resonant frequency.

Elimination of the VHF parasitic oscillation may be accomplished using one of the following techniques:

- Place a small coil and resistor combination in the plate lead between the plate of the tube and the tank circuit, as shown in Figure 4.10a. The resistor-coil combination usually is made up of a noninductive resistor of about 25–100 Ω , shunted by three or four turns of approximately 1/2 in. diameter and frequently wound around the resistor. In some cases, it may be necessary to use such a suppressor in both the plate and grid leads. The resistor-coil combination operates on the principle that the resistor loads the VHF circuit but is shunted by the coil for the lower fundamental frequency. In the process of adjusting the resistor-coil combination, it is often found that the

resistor runs hot. This heat usually is caused by the dissipation of fundamental power in the resistor, and it is an indication of too many turns in the suppressor coil. Use only enough turns to suppress the parasitic, and no more. Once the parasitic has been suppressed, there will be no parasitic voltage or current present. Therefore, there is no parasitic power to be dissipated.

- Use small parasitic chokes in the plate lead, as shown in Figure 4.10b. The size of the coil will vary considerably depending upon the tube and circuit layout. A coil of 4–10 turns with a diameter of approximately 1/2 in. is typical. The presence of the choke in the frequency-determining part of the circuit lowers the frequency of a possible VHF parasitic so that it falls near the self-neutralizing frequency of

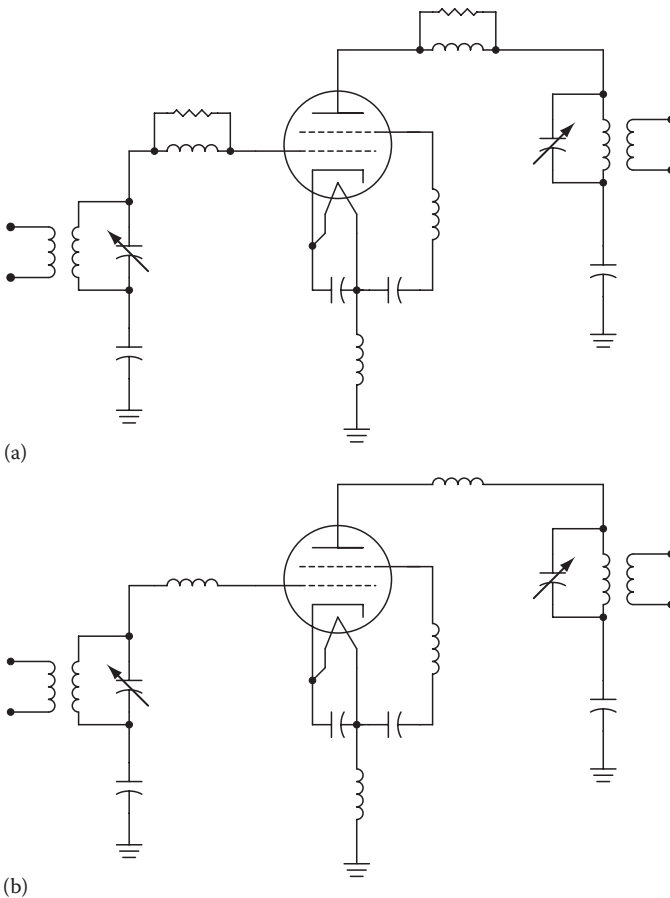


FIGURE 4.10 Placement of parasitic suppressors to eliminate VHF parasitic oscillations in a high-frequency RF amplifier: (a) resistor–coil combination and (b) parasitic choke. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

the tube and bypass leads. In addition to variation in the size of the suppressor choke, the amount of inductance common to the screen and filament in the filament grounding strap may be a factor. This parameter can be varied simultaneously with the suppressor choke.

Of the two methods outlined for suppressing VHF parasitic oscillations, the first is probably the simpler and is widely employed.

4.2.4.1 Dynatron Oscillation

Another form of commonly encountered self-oscillation is known as *dynatron oscillation*, caused when an electrode in a vacuum tube has *negative resistance* [1]. At times, there may be more electrons leaving the screen grid than arriving. If the screen voltage is allowed to increase under these conditions, even more electrons will leave the grid. This phenomenon implies a negative resistance characteristic. If there is high alternating current impedance in the circuit from the screen grid through the screen grid power supply, and from the plate power supply to the plate, dynatron oscillation may be sustained.

Dynatron oscillation typically occurs in the region of 1–20 Hz. This low-frequency oscillation usually is accompanied by another oscillation in the 1–2 kHz region. Suppression of these oscillations can be accomplished by placing a large bypass capacitor (1000 μF) across the output of the screen grid power supply. The circuit supporting the oscillation also can be detuned by a large inductor. Increasing the circuit losses at the frequency of oscillation is also effective.

4.2.4.2 Harmonic Energy

It is generally not appreciated that the pulse of grid current contains energy on harmonic frequencies and that control of these harmonic energies may be important. The ability of the tetrode and pentode to isolate the output circuit from the input over a wide range of frequencies is significant in avoiding feedthrough of harmonic voltages from the grid circuit. Properly designed tetrode and pentode amplifiers provide for complete shielding in the amplifier layout so that coupling external to the tube is prevented.

In RF amplifiers operating either on the fundamental or on a desired harmonic frequency, the control of unwanted harmonics is important. The following steps permit reduction of the unwanted harmonic energies present in the output circuit:

- Keep the circuit impedance between the plate and cathode low for the high harmonic frequencies. This requirement may be achieved by having some or all of the tuning capacitance of the resonant circuit close to the tube.

- Completely shield the input and output compartments.
- Use inductive output coupling from the resonant plate circuit and possibly a capacitive or Faraday shield between the coupling coil and the tank coil, or a high-frequency attenuating circuit such as a π or π -L network.
- Use low-pass filters on all supply leads and wires coming into the output and input compartments.
- Use resonant traps for particular frequencies.
- Use a low-pass filter in series with the output transmission line.

4.2.5 Shielding

In an RF amplifier, shielding between the input and output circuits must be considered [1]. Triode amplifiers are more tolerant of poor shielding because power gain is relatively low. If the circuit layout is reasonable and no inductive coupling is allowed to exist, a triode amplifier usually can be built without extensive shielding. Even if shielding is not necessary to prevent fundamental-frequency oscillation, it will aid in eliminating any tendency toward parasitic oscillation. The higher the gain of an amplifier, the more important the shielding.

4.2.5.1 Pierced Shields

Tetrode and pentode PA require comprehensive shielding to prevent input-to-output circuit coupling [1]. It is advisable to use nonmagnetic materials such as copper, aluminum, or brass in the RF fields to provide the shielding. Quite often, a shield must have holes through it to allow the passage of cooling air. In the LF and part of the HF range, the presence of small holes will not impair shield effectiveness. As the frequency is increased, however, the RF currents flowing around the hole in one compartment cause fields to pass through the hole into another compartment. Currents, therefore, are induced on the shield in the second compartment. This problem can be eliminated by using holes that have significant length. A piece of pipe with a favorable length-to-diameter ratio as compared to the frequency of operation will act as a *waveguide-beyond-cutoff* attenuator.

If more than one hole is required to pass air, a material resembling a honeycomb may be used. This material is available commercially and provides excellent isolation with a minimum of air pressure drop. A section of honeycomb shielding is shown in Figure 4.11. Some tube sockets incorporate a waveguide-beyond-cutoff air path. These sockets allow the tube in the amplifier to operate at high gain and up through VHF.

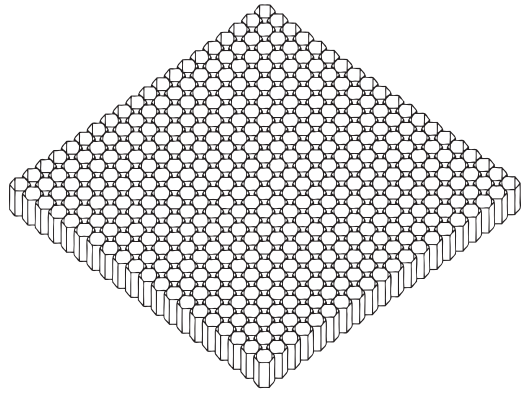


FIGURE 4.11
Section of honeycomb shielding material used in an RF amplifier.

4.2.5.2 Metal Base Shells and Submounted Sockets

Some tetrodes and pentodes are supplied with metal base shells [1]. The shell is typically grounded by the clips provided with the socket. This completes the shielding between the output and input circuits because the base shell of the tube comes up opposite the screen shield within the tube itself.

Some pentodes use this metal base shell as the terminal for the suppressor grid. If the suppressor is to be at some potential other than ground, then the base shell must not be dc grounded. The base shell is bypassed to ground for RF and insulated from ground for dc.

There is a family of tetrodes and pentodes without the metal base shell. For this type of tube structure, it is good practice to submount the socket so that the internal screen shield is at the same level as the chassis deck. This technique will improve the input-to-output circuit shielding. In submounting a tube, it is important that adequate clearance be provided around the base of the device for the passage of cooling air.

4.2.5.3 Compartments

By placing the tube and related circuits in completely enclosed compartments and properly filtering incoming supply wires, it is possible to prevent coupling of RF energy out of the circuit by means other than the desired output coupling [1]. Such filtering prevents the coupling of energy that may be radiated or fed back to the input section or to earlier stages in the amplifier chain. Energy fed back to the input circuit causes undesirable interaction in tuning and/or self-oscillation. If energy is fed back to earlier stages, significant operational problems may result because of the large power gain over several stages.

In the design of an RF amplifier, doors or removable panels typically must be used. The requirement for making a good, low-resistance joint at the discontinuity must be met. There are several materials available commercially

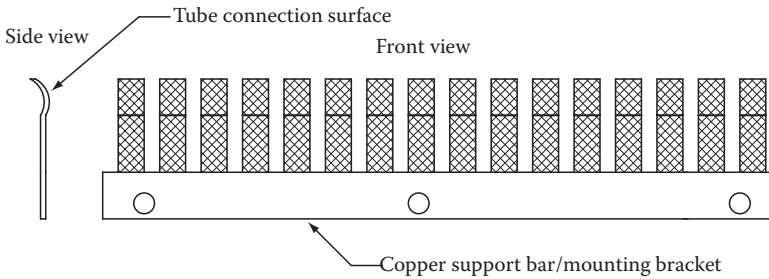


FIGURE 4.12
Selection of common fingerstock.

for this purpose. Fingerstock, shown in Figure 4.12, has been used for many years. Sometimes, after the wiring has been completed, further shielding of a particular conductor is required. Various types of shielding tapes can be wound on a conductor as a temporary or permanent solution.

4.2.6 Protection Measures

Power grid tubes are designed to withstand considerable abuse [1]. The maximum ratings for most devices are conservative. For example, the excess anode dissipation resulting from detuning the plate circuit will have no ill effects on most tubes if it is not applied long enough to overheat the envelope and the seal structure.

Similarly, the control, screen, and suppressor grids will stand some excess dissipation. Typically, the maximum dissipation for each grid indicated on the data sheet should not be exceeded except for time intervals of less than 1 s. The maximum dissipation rating for each grid structure is usually considerably above typical values used for maximum output so that ample operating reserve is provided. The time of duration of overload on a grid structure is necessarily short because of the small heat-storage capacity of the grid wires. Furthermore, grid temperatures cannot be measured or seen, so no warning of accidental overload is apparent.

The type and degree of protection required in an RF amplifier against circuit failure varies with the type of screen and grid voltage supply. Table 4.1 lists protection criteria for tetrode and pentode devices. The table provides guidelines on the location of a suitable relay that should act to remove the principal supply voltage from the stage or transmitter to prevent damage to the tube.

For designs where screen voltage is taken through a dropping resistor from the plate supply, a plate relay provides almost universal protection. For a fixed screen supply, a relay provides protection in most cases. For protection against excessive antenna loading and subsequent high plate dissipation, a screen undercurrent relay also may be used in some services.

TABLE 4.1

Protection Guidelines for Tetrode and Pentode Devices

Circuit Failure Type	Fixed Screen Supply		Screen Voltage through Dropping Resistor	
	Fixed Grid Bias	Resistor Grid Bias	Fixed Grid Bias	Resistor Grid Bias
Loss of excitation	No protection required	Plate current relay	Plate current relay	Plate current relay or screen control circuit
Loss of antenna loading	Screen current relay	Screen current relay	Grid current relay	No protection required
Excess antenna loading	Screen undercurrent relay	Screen undercurrent relay	Plate current relay	Plate current relay
Failure of plate supply	Screen current relay	Screen current relay	Grid current relay	No protection required
Failure of screen supply	Grid current relay	No protection required	Does not apply	Does not apply
Failure of grid bias supply	Plate current relay or screen current relay	Does not apply	Plate current relay or grid current relay	Does not apply

Source: After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.

The plate, screen, and bias voltages may be applied simultaneously to a tetrode. The same holds true for a pentode, plus the application of the suppressor voltage. In a grid-driven amplifier, grid bias and excitation usually can be applied alone to the tube, especially if a grid leak resistor is used. Plate voltage can be applied to the tetrode and pentode before the screen voltage with or without excitation to the control grid. Never apply screen voltage before the plate voltage. The only exceptions are when the tube is cut off so that no space current (screen or plate current) will flow, or when the excitation and screen voltage are low. If screen voltage is applied before the plate voltage and screen current can flow, the maximum allowable screen dissipation will almost always be exceeded, and tube damage will result.

Table 4.2 lists protection guidelines for a triode. The table covers the grid-driven triode amplifier and the high- μ (zero-bias) cathode-driven triode amplifier. Drive voltage must never be applied to a zero-bias triode amplifier without plate voltage. The table indicates the recommended location of a suitable relay that should act to remove the principal supply voltage from the stage or transmitter to prevent damage to the tube or transmitter.

4.2.7 Cavity Amplifier Systems

Power grid tubes are ideally suited for use as the power generating element in a cavity amplifier. Because of the physical dimensions involved, cavity

TABLE 4.2

Protection Guidelines for a Triode

Circuit Failure Type	Triode		
	Fixed Grid Bias	Resistor Grid Bias	Zero-Bias Triode
Loss of excitation	No protection required	Plate overcurrent relay	No protection required
Loss of antenna loading	RF output detector and relay	RF output detector and relay	Grid overcurrent relay
Excess antenna loading	RF output detector and relay	RF output detector and relay	RF output detector and relay
Failure of plate supply	No protection required	No protection required	Grid overcurrent relay
Failure of grid bias supply	Plate overcurrent relay	Does not apply	Does not apply

Source: After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.

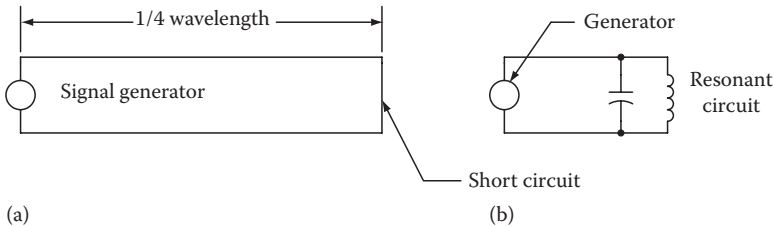


FIGURE 4.13

Shorted 1/4-wavelength line: (a) physical circuit and (b) electrical equivalent.

designs typically are limited to VHF and above. Lower-frequency RF generators utilize discrete *L* and *C* devices to create the needed tank circuit. Two types of cavity amplifiers commonly are used: 1/4-wavelength and 1/2-wavelength systems.

In a cavity amplifier, the tube becomes part of a resonant transmission line. The stray interelectrode and distributed capacity and inductance of the tube are used to advantage as part of the resonant line. This resonant line is physically larger than the equivalent lumped constant *LRC* resonant circuit operating at the same frequency, and this larger physical size aids in solving the challenges of high-power operation, skin effect losses, and high-voltage-standoff concerns.

A shorted 1/4-wavelength transmission line has a high, purely resistive input impedance. Electrically, it appears as a parallel resonant circuit, as shown in Figure 4.13. When the physical length of the line is less than 1/4 wavelength, the impedance will be lower and the line will appear inductive, as illustrated in Figure 4.14. This inductance is used to resonate with the capacitive reactance in the tube and the surrounding circuit.

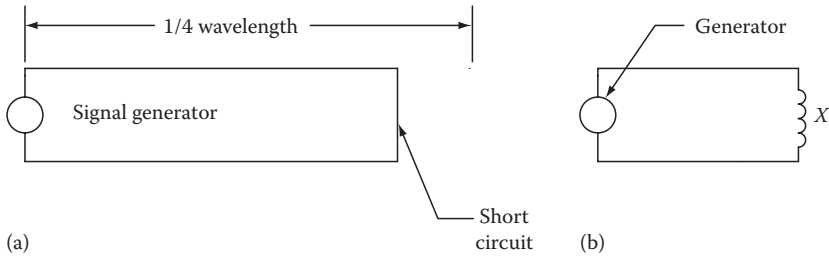


FIGURE 4.14 Shorted line less than 1/4 wavelength: (a) physical circuit and (b) electrical equivalent.

4.2.8 Bandwidth and Efficiency

PA bandwidth has a significant effect on modulation performance. Available bandwidth determines the amplitude response, phase response, and group-delay response. Performance trade-offs often must be considered in the design of a cavity amplifier, including bandwidth, gain, and efficiency.

PA bandwidth is restricted by the equivalent load resistance across the parallel tuned circuits in the stage. Tuned circuits are necessary to cancel low reactive impedance presented by the relatively high input and output capacitances of the amplifying device. The bandwidth for a single tuned circuit is proportional to the ratio of capacitive reactance X_c to load resistance R_l appearing across the tuned circuit:

$$BW \approx \frac{K}{2\pi f_c C R_l} \approx \frac{K(X_c)}{R_l} \tag{4.10}$$

where

- BW is the bandwidth between half-power (-3 dB) points
- K is the proportionality constant
- R_l is the load resistance appearing across tuned circuit
- C is the total capacitance of tuned circuit (includes stray capacitances and output or input capacitances of the tube)
- X_c is the capacitive reactance of C
- f_c is the carrier frequency

The RF voltage swing across the tuned circuit also depends on the load resistance. For the same power and efficiency, the bandwidth can be increased if the capacitance is reduced.

The efficiency of a cavity power amplifier (PA) depends primarily on the RF plate voltage swing, the plate current conduction angle, and the cavity efficiency. Cavity efficiency is related to the ratio of loaded to unloaded Q as follows:

$$N = 1 - \frac{Q_l}{Q_u} \times 100 \tag{4.11}$$

where

N is the cavity efficiency in percent

Q_l is the cavity loaded Q

Q_u is the cavity unloaded Q

The loaded Q is determined by the plate load impedance and output circuit capacitance. Unloaded Q is determined by the cavity volume and the RF resistivity of the conductors resulting from the skin effect. For best cavity efficiency, the following conditions are desirable:

- High unloaded Q
- Low loaded Q

As the loaded Q increases, the bandwidth decreases. For a given tube output capacitance and power level, loaded Q decreases with decreasing plate voltage or with increasing plate current. The increase in bandwidth at reduced plate voltage occurs because the load resistance is directly related to the RF voltage swing on the tube anode. For the same power and efficiency, the bandwidth also can be increased if the output capacitance is reduced. Power tube selection and minimization of stray capacitance are important considerations when designing for maximum bandwidth.

4.2.9 Current Paths

The operation of a cavity amplifier is an extension of the current paths inside the tube. Two elements must be examined in this discussion:

- The input circulating currents
- The output circulating currents

4.2.9.1 Input Circuit

The grid/cathode assembly resembles a transmission line whose termination is the RF resistance of the electron stream within the tube. Figure 4.15 shows the current path of an RF generator (the RF driver stage output) feeding a signal into the grid/cathode circuit.

The outer contact ring of the cathode heater assembly makes up the inner conductor of a transmission line formed by the cathode and control grid assemblies. The filament wires are returned down the center of the cathode. For the input circuit to work correctly, the cathode must have a low RF impedance to ground. This cathode bypassing may be accomplished in one of several ways.

Below 30 MHz, the cathode can be grounded to RF voltages by simply bypassing the filament connections with capacitors, as shown in Figure 4.16a. Above 30 MHz, this technique does not work well because of the stray inductance of the filament leads. Notice that in Figure 4.16b, the filament leads

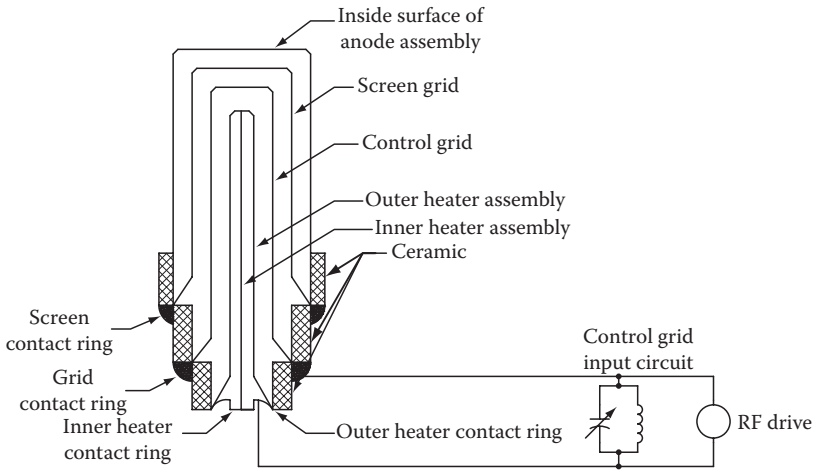


FIGURE 4.15
Simplified representation of the grid input circuit of a PA tube.

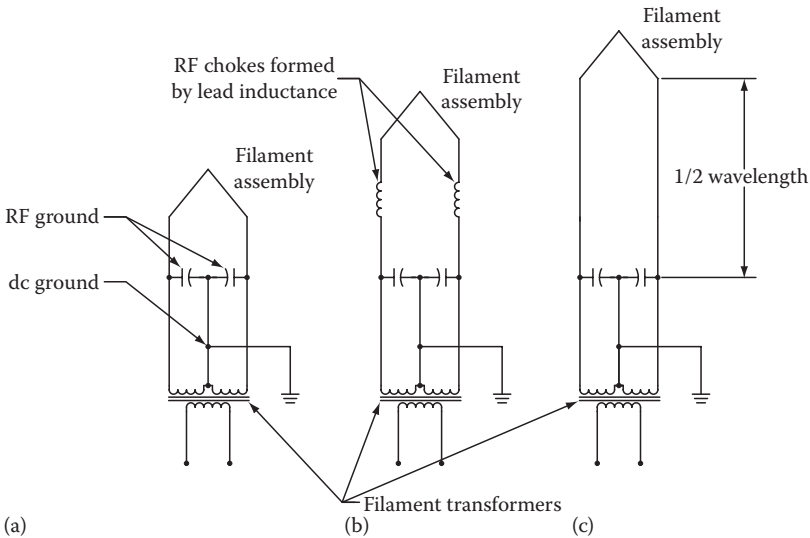


FIGURE 4.16
Three common methods for RF bypassing of the cathode of a tetrode PA tube: (a) grounding the cathode below 30 MHz, (b) grounding the cathode above 30 MHz, and (c) grounding the cathode via a 1/2-wave transmission line.

appear as RF chokes, preventing the cathode from being placed at RF ground potential. This causes negative feedback and reduces the efficiency of the input and output circuits.

In Figure 4.16c, the cathode circuit is configured to simulate a 1/2-wave transmission line. The line is bypassed to ground with large-value capacitors 1/2 wavelength from the center of the filament (at the filament voltage feed

point). This transmission line RF short circuit is repeated $1/2$ wavelength away at the cathode (heater assembly), effectively placing it at ground potential.

Because $1/2$ -wavelength bypassing is usually bulky at VHF (and may be expensive), RF generators often are designed using certain values of inductance and capacitance in the filament/cathode circuit to create an artificial transmission line that will simulate a $1/2$ -wavelength shorted transmission line. As illustrated in the figure, the inductance and capacitance of the filament circuit can resemble an artificial transmission line of $1/2$ wavelength, if the values of L and C are properly selected.

4.2.9.2 Output Circuit

The plate-to-screen circulating current of the tetrode is shown in Figure 4.17. For the purposes of example, consider that the output RF current is generated by an imaginary current generator located between the plate and screen grid. The RF current travels along the inside surface of the plate structure (because of the skin effect), through the ceramic at the lower half of the anode contact ring, across the bottom of the fins, and to the band around the outside of the fins. The RF current then flows through the plate bypass capacitor to the RF tuned circuit and load, and returns to the screen grid.

The return current travels through the screen bypass capacitor (if used) and screen contact ring, up the screen base cone to the screen grid, and back to the imaginary generator.

The screen grid has RF current returning to it, but because of the assembly's low impedance, the screen grid is effectively at RF ground potential. The RF current generator, therefore, appears to be feeding an open-ended

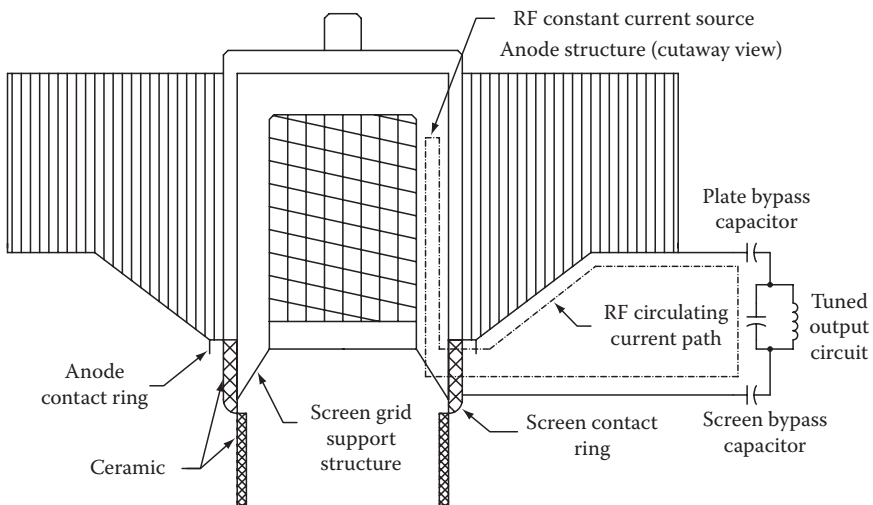


FIGURE 4.17 RF circulating current path between the plate and screen in a tetrode PA tube.

transmission line consisting of the anode (plate) assembly and the screen assembly. The RF voltage developed by the anode is determined by the plate impedance (Z_p) presented to the anode by the resonant circuit and its load.

When current flows on one conductor of a transmission line cavity circuit, an equal-magnitude current flows in the opposite direction on the other conductor. This means that a large value of RF circulating current is flowing in the cavity amplifier outer conductor (the cavity box). All of the outer conductor circulating currents start at and return to the screen grid (in a tetrode-based system). The front or back access panel (door) of the cavity is part of the outer conductor, and large values of circulating current flow into it, through it, and out of it. A mesh contact strap generally is used to electrically connect the access panel to the rest of the cavity.

4.2.10 1/4-Wavelength Cavity

The 1/4-wavelength PA cavity is common in transmitting equipment. The design is simple and straightforward. A number of variations can be found in different RF generators, but the underlying theory of operation is the same.

A typical 1/4-wave cavity is shown in Figure 4.18. The plate of the tube connects directly to the inner section (tube) of the plate-blocking capacitor. The exhaust chimney/inner conductor forms the other plate of the blocking capacitor. The cavity walls form the outer conductor of the 1/4-wave transmission line circuit. The dc plate voltage is applied to the PA tube by a cable

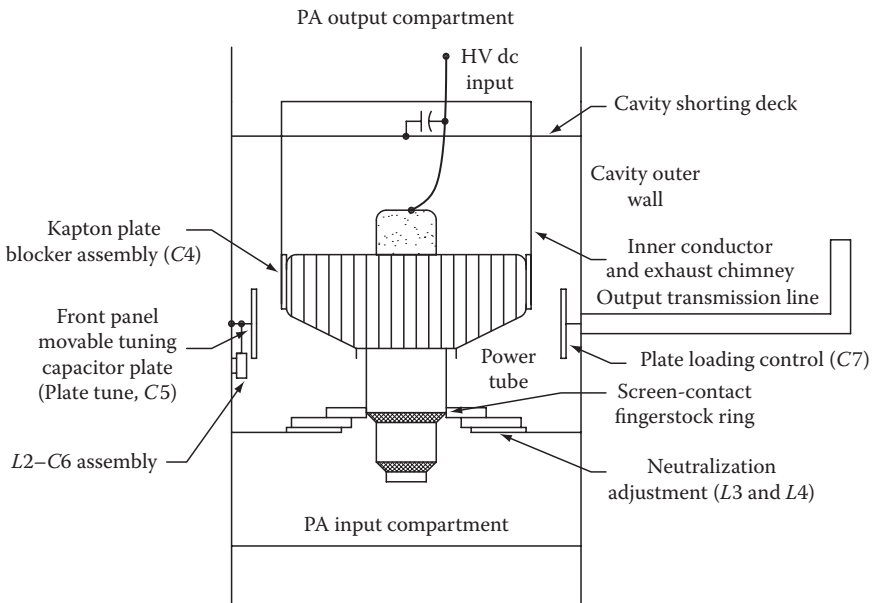


FIGURE 4.18

Physical layout of a common type of 1/4-wavelength PA cavity designed for VHF operation.

routed inside the exhaust chimney and inner tube conductor. In the design shown in the figure, the screen-contact fingerstock ring mounts on a metal plate that is insulated from the grounded cavity deck by an insulating material. This hardware makes up the screen-blocking capacitor assembly. The dc screen voltage feeds to the fingerstock ring from underneath the cavity deck through an insulated feedthrough assembly.

A grounded screen configuration also may be used in this design in which the screen-contact fingerstock ring is connected directly to the grounded cavity deck. The PA cathode then operates at below ground potential (in other words, at a negative voltage), establishing the required screen voltage to the tube.

The cavity design shown in the figure is set up to be slightly shorter than a full 1/4 wavelength at the operating frequency. This makes the load inductive and resonates the tube's output capacity. Thus, the physically foreshortened shorted transmission line is resonated and electrically lengthened to 1/4 wavelength.

Figure 4.19 illustrates the paths taken by the RF circulating currents in the circuit. RF energy flows from the plate, through the plate-blocking capacitor, along the inside surface of the chimney/inner conductor (because of the skin effect), across the top of the cavity, down the inside surface of the cavity box, across the cavity deck, through the screen-blocking capacitor, over the screen-contact fingerstock, and into the screen grid.

Figure 4.20 shows a graph of RF current, voltage, and impedance for a 1/4-wavelength coaxial transmission line. The plot shows that infinite impedance, zero RF current, and maximum RF voltage occur at the feed point. This would not be suitable for a practical PA circuit because arcing could result from the high RF voltage, and poor efficiency would be caused by the mismatch between the tube and the load.

Notice, however, the point on the graph marked at slightly less than 1/4 wavelength. This length yields an impedance of 600–800Ω and is ideal for

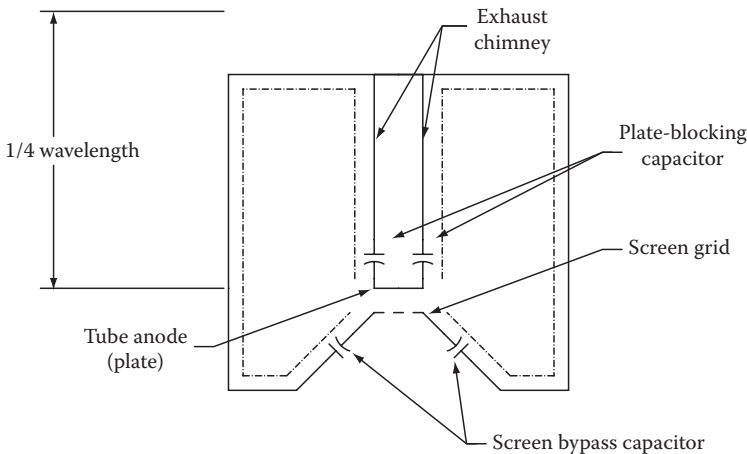


FIGURE 4.19

RF circulating current paths for the 1/4-wavelength cavity shown in Figure 4.18.

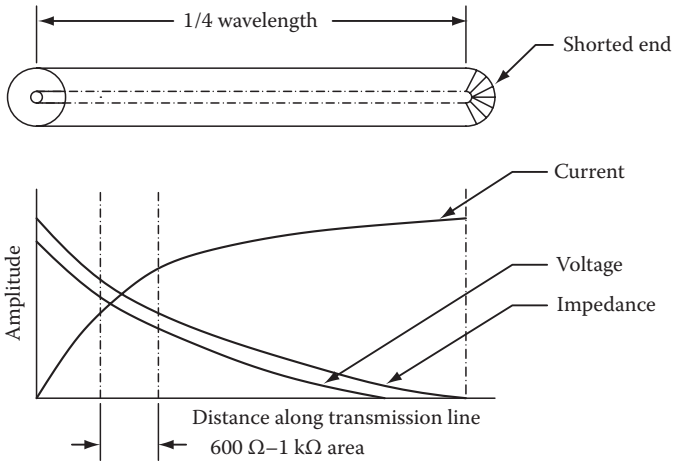


FIGURE 4.20
Graph of the RF current, RF voltage, and RF impedance for a $1/4$ -wavelength shorted transmission line. At the feed point, RF current is zero, RF voltage is maximum, and RF impedance is infinite.

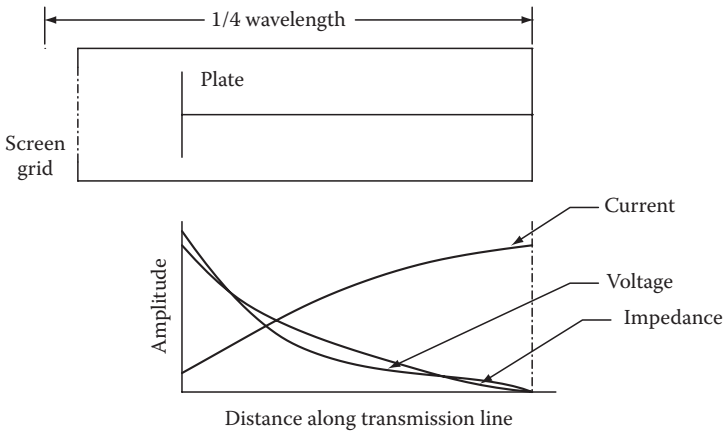


FIGURE 4.21
Graph of the RF current, RF voltage, and RF impedance produced by the physically foreshortened coaxial transmission line cavity.

the PA plate circuit. It is necessary, therefore, to physically foreshorten the shorted coaxial transmission line cavity to provide the correct plate impedance. Shortening the line also is a requirement for resonating the tube’s output capacity, because the capacity shunts the transmission line and electrically lengthens it.

Figure 4.21 shows a graph of the RF current, voltage, and impedance presented to the plate of the tube as a result of the physically foreshortened line. This plate impedance is now closer to the ideal $600\text{--}800\Omega$ value required by the tube anode.

4.2.10.1 Tuning the Cavity

Coarse tuning of the cavity is accomplished by adjusting the cavity length. The top of the cavity (the cavity shorting deck) is fastened by screws and clamps and can be raised or lowered to set the length of the cavity for the particular operating frequency. Fine-tuning is accomplished by a variable-capacity plate-tuning control built into the cavity. In a typical design, one plate of the tuning capacitor—the stationary plate—is fastened to the inner conductor just above the plate-blocking capacitor. The movable tuning plate is fastened to the cavity box—the outer conductor—and mechanically linked to the front-panel tuning control. This capacity shunts the inner conductor to the outer conductor and is used to vary the electrical length and resonant frequency of the cavity.

The 1/4-wavelength cavity is inductively coupled to the output port. This coupling is usually on the side opposite the cavity access door. The inductive pickup loop can take one of several forms. In one design it consists of a half-loop of flat copper bar stock that terminates in the loading capacitor at one end and feeds the output transmission line inner conductor at the other end. This configuration is shown in Figure 4.22. The inductive pickup ideally would be placed at the maximum current point in the 1/4-wavelength cavity. However, this point is located at the cavity shorting deck, and, when the deck is moved for coarse tuning, the magnetic coupling will be changed. A compromise in positioning, therefore, must be made. The use of a broad,

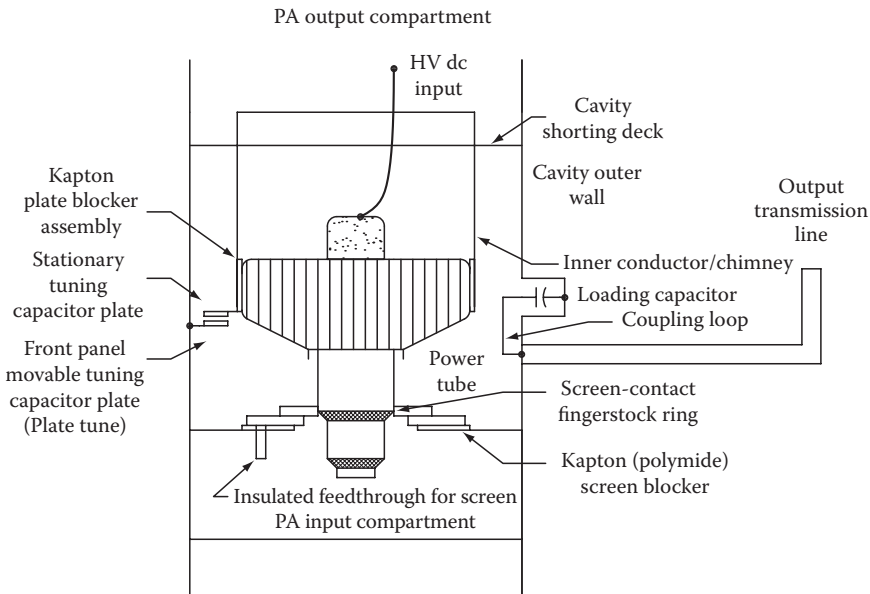


FIGURE 4.22
Basic loading mechanism for a VHF 1/4-wave cavity.

flat copper bar for the coupling loop adds some capacitive coupling to augment the reduced magnetic coupling.

Adjustment of the loading capacitor couples the $50\ \Omega$ transmission line impedance to the impedance of the cavity. Heavy loading lowers the plate impedance presented to the tube by the cavity. Light loading reflects a higher load impedance to the amplifier plate.

Methods commonly used to increase the operating bandwidth of the cavity include minimizing added tuning capacitance. The ideal case would be to resonate the plate capacitance alone with a “perfect” inductor. Practical cavities, however, require either the addition of a variable capacitor or a variable inductor using sliding contacts for tuning. Other inherent mechanical and electrical compromises include

- The requirement for a plate dc-blocking capacitor.
- The presence of maximum RF current at the grounded end of the simulated transmission line where the conductor may be nonhomogeneous. This can result in accompanying losses in the contact resistance.

4.2.11 1/2-Wavelength Cavity

The operation of a 1/2-wavelength cavity follows the same basic idea as the 1/4-wavelength design outlined previously. The actual construction of the system depends upon the power level used and the required bandwidth. An example will help to illustrate how a 1/2-wavelength cavity operates.

The design of a basic 1/2-wavelength PA cavity for operation in the region of 100 MHz is shown in Figure 4.23. The tube anode and a silver-plated brass pipe serve as the inner conductor of the 1/2-wave transmission line. The cavity box serves as the outer conductor. The transmission line is open at the far end and repeats this condition at the plate of the tube. The line is, in effect, a parallel resonant circuit for the PA tube.

The physical height of the circuit shown (67 in.) was calculated for operation down to 88 MHz. To allow adequate clearance at the top of the transmission line and space for input circuitry at the bottom of the assembly, the complete cavity box would have to be almost 8 ft tall. This is too large for any practical transmitter.

Figure 4.24 shows RF voltage, current, and impedance for the inner conductor of the transmission line and the anode of the tube. The load impedance at the plate is thousands of ohms. Therefore, the RF current is extremely small, and the RF voltage is extremely large. In the application of such a circuit, arcing would become a problem. The high plate impedance also would make amplifier operation inefficient.

The figure also shows an area between the anode and the 1/4-wavelength location where the impedance of the circuit is 600–800 Ω . As noted previously,

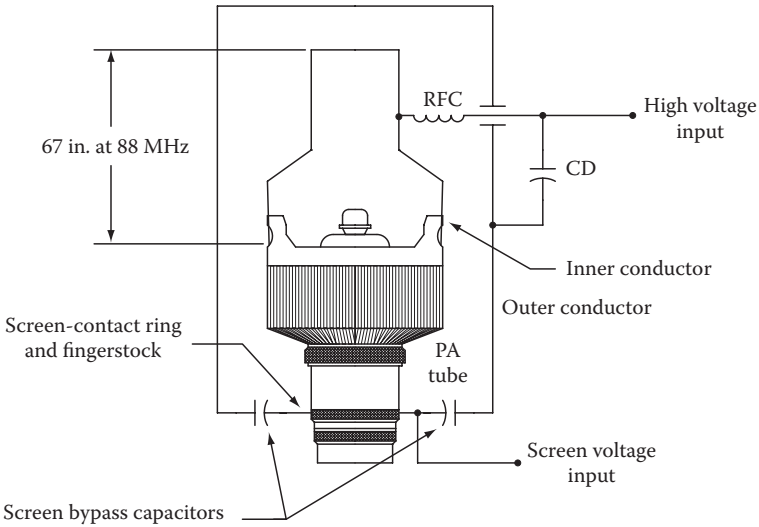


FIGURE 4.23
1/2-wavelength PA cavity in its basic form.

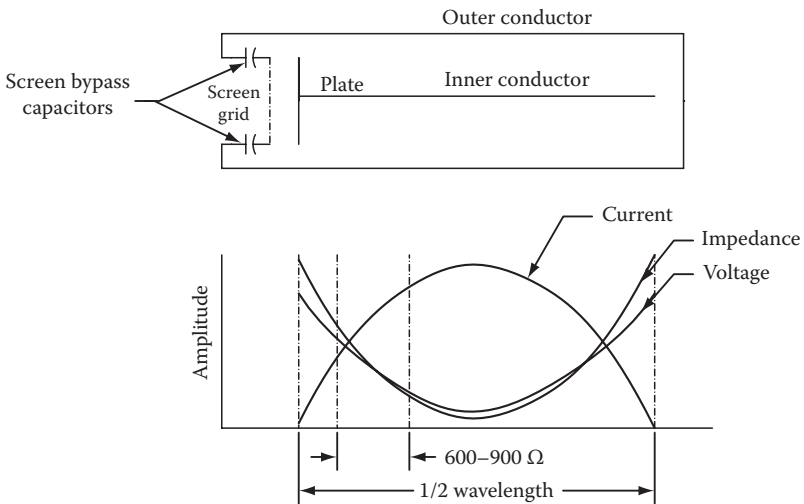


FIGURE 4.24
Distribution of RF current, voltage, and impedance along the inner conductor of a 1/2-wavelength cavity.

this value is ideal for the anode of the PA tube. To achieve this plate impedance, the inner conductor must be less than a full 1/2 wavelength. The physically foreshortened transmission-line circuit must, however, be electrically resonated (lengthened) to 1/2 wavelength for proper operation.

If the line length were changed to operate at a different frequency, the plate impedance also would change because of the new distribution of RF voltage and current on the new length of line. The problem of frequency change, therefore, is twofold: The length of the line must be adjusted for resonance, and the plate impedance of the tube must be kept constant for good efficiency.

To accommodate operation of this system at different frequencies, while keeping the plate impedance constant, two forms of coarse-tuning and one form of fine-tuning are built into the 1/2-wave PA cavity. Figure 4.25 shows the tube and its plate line (inner conductor). The inner conductor is U-shaped to reduce the cavity height. With the movable section (the plate tune control) fully extended, the inner conductor measures 38 in., and the anode strap measures 7 in. The RF path from the anode strap to the inside of the tube plate (along the surface because of the skin effect) is estimated to be about 8 in. This makes the inner conductor maximum length about 53 in. This is too short to be a 1/2 wavelength at 88 MHz, the target low-end operating frequency in this example. The full length of a 1/2-wave line is 67.1 in. at 88 MHz.

The coarse-tuning and fine-tuning provisions of the cavity, coupled with the PA tube's output capacity, resonate the plate line to the exact operating frequency. In effect, they electrically lengthen the physically foreshortened line. This process, along with proper loading, determines the plate impedance and, therefore, the efficiency.

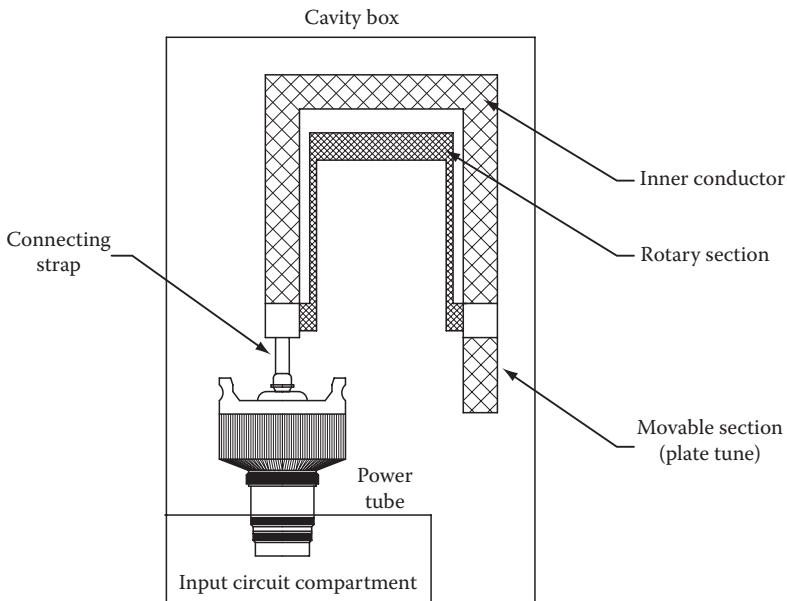


FIGURE 4.25

Configuration of a practical 1/2-wavelength PA cavity.

4.2.11.1 Lengthening the Plate Line

The output capacity of the tube is the first element that electrically lengthens the plate line. A 1/2-wave transmission line that is too short offers a high impedance that is both resistive and inductive. The output capacity of the tube resonates this inductance. The detrimental effects of the tube's output capacity, therefore, are eliminated. The anode strap and the cavity inner conductor rotary section provide two methods of coarse frequency adjustment for resonance.

The anode strap, shown in Figure 4.26, has a smaller cross-sectional area than the inner conductor of the transmission line. Therefore, it has more inductance than an equal length of the inner conductor. The anode-coupling strap acts as a series inductance and electrically lengthens the plate circuit.

At low frequencies, one narrow anode strap is used. At midband frequencies, one wide strap is used. The wide strap exhibits less inductance than the narrow strap and does not electrically lengthen the plate circuit as much. At the upper end of the operating band, two anode straps are used. The parallel arrangement lowers the total inductance of the strap connection and adds still less electrical length to the plate circuit.

The main section of the plate resonant line, together with the rotary section, functions as a parallel inductance. RF current flows in the same direction through the transmission line and the rotary section. Therefore, the magnetic fields of the two paths add. When the rotary section is at maximum height, the magnetic coupling between the main section of the transmission

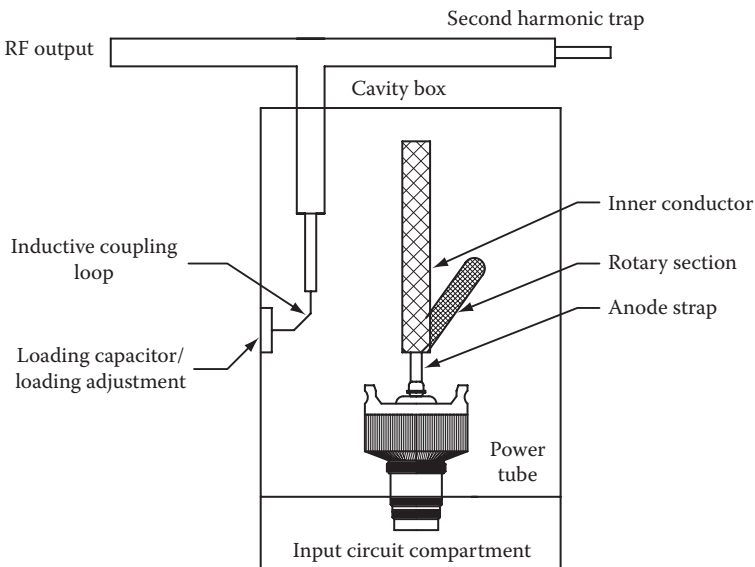


FIGURE 4.26
Coarse-tuning mechanisms for a 1/2-wave cavity.

line and the rotary assembly is maximum. Because of the relatively large mutual inductance provided by this close coupling, the total inductance of these parallel inductors increases. This electrically lengthens the transmission line and lowers the resonant frequency. The concept is illustrated in Figure 4.27a.

When the rotary section is at minimum height, the magnetic coupling between the two parts of the inner conductor is minimum. This reduced coupling lowers the mutual inductance, which lowers the total inductance of the parallel combination. The reduced inductance allows operation at a higher resonant frequency. This condition is illustrated in Figure 4.27b. The rotary section provides an infinite number of coarse settings for various operating frequencies.

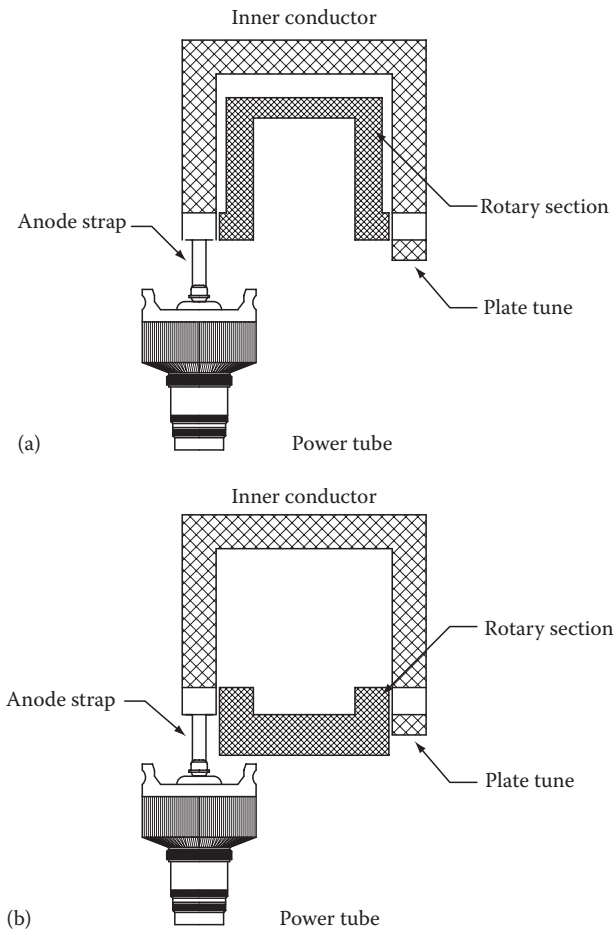


FIGURE 4.27

Using the cavity's movable section to adjust for resonance: (a) the rotary section at maximum height and (b) the rotary section at minimum height.

The movable plate-tune assembly is located at the end of the inner-plate transmission line. It can be moved up and down to change the physical length of the inner conductor by about 4-3/4in. This assembly is linked to the front-panel plate-tuning knob, providing a fine adjustment for cavity resonance.

4.2.12 Folded 1/2-Wavelength Cavity

A special case of the 1/2-wavelength PA cavity is shown in Figure 4.28. The design employs a folded 1/2-wave resonator constructed with coaxial aluminum and copper tubing. This cavity arrangement eliminates the high-voltage blocking capacitor and high-current shorting contacts of conventional designs by connecting the main transmission-line resonant circuit conductor directly to the anode of the power tube. A grounded, concentric transmission-line center conductor tunes the cavity with a variable reentrant length inserted into the end of the main conductor opposite the tube.

The main conductor (the fixed portion of the plate line) is insulated from ground and carries the anode dc potential. High-voltage power is fed at the fundamental-frequency RF voltage null point, approximately 1/4 wavelength from the anode, for easy RF decoupling. A large surface area without sliding contacts results in minimal loss.

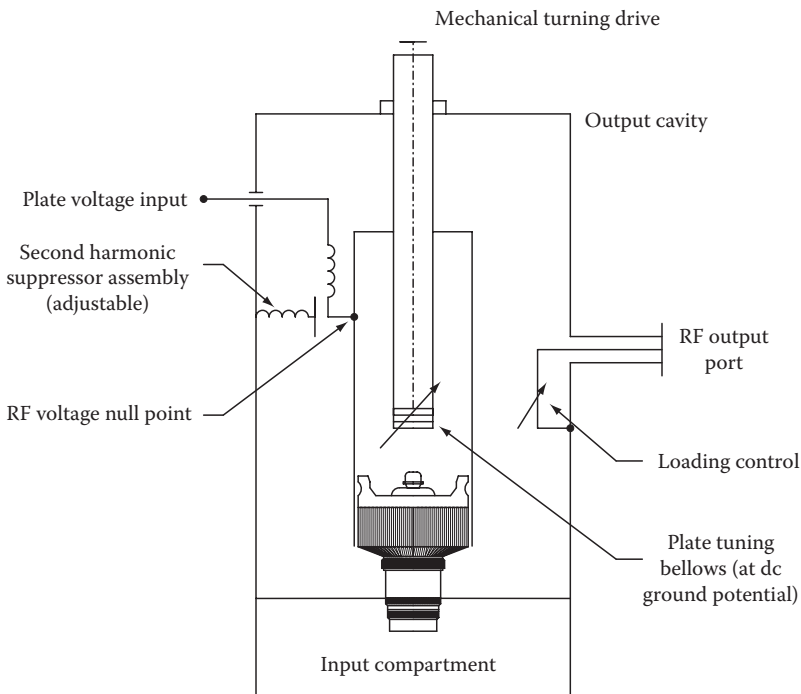


FIGURE 4.28
Basic design of a folded 1/2-wavelength cavity.

Incorporated into the tank design is a second-harmonic suppressor. Rather than attenuating the second harmonic after the signal has been generated and amplified, this design essentially prevents formation of second-harmonic energy by series-*LC* trapping the second-harmonic waveform at the point where the wave exhibits a high impedance (approximately 1/4 wavelength from the anode). The second harmonic will peak in voltage at the same point that the dc-plate potential is applied.

Plate tuning is accomplished by an adjustable bellows on the center portion of the plate line, which is maintained at chassis ground potential. Output coupling is accomplished with an untuned loop intercepting the magnetic field concentration at the voltage null (maximum RF current) point of the main line. The PA loading control varies the angular position of the plane of the loop with respect to the plate line, changing the amount of magnetic field that it intercepts. Multiple phosphor-bronze leaves connect one side of the output loop to ground and the other side to the center conductor of the output transmission line. This allows for mechanical movement of the loop by the PA loading control without using sliding contacts. The grounded loop improves immunity to lightning and to static buildup on the antenna.

4.2.13 Wideband Cavity

The cavity systems discussed so far in this section provide adequate bandwidth for many applications. Some uses, however, dictate an operating bandwidth beyond what a conventional cavity amplifier can provide. One method of achieving wider operating bandwidth involves the use of a double-tuned overcoupled amplifier, shown in Figure 4.29. The system includes four controls to accomplish tuning:

- *Primary tune*, which resonates the plate circuit and tends to tilt the response and slide it up and down the bandpass.
- *Coupling*, which sets the bandwidth of the PA. Increased coupling increases the operating bandwidth and lowers the PA efficiency. When the coupling is adjusted, it can tilt the response and change the center of the bandpass, necessitating readjustment of the plate-tune control.
- *Secondary tune*, which resonates the secondary cavity and will tilt the response if so adjusted. The secondary tune control typically can be used to slide the response up and down the bandpass, similar to the primary tune control.
- *Loading*, which determines the value of ripple in the bandpass response. Adjustment of the loading control usually tilts the response and changes the bandwidth, necessitating readjustment of the secondary tune and coupling controls.

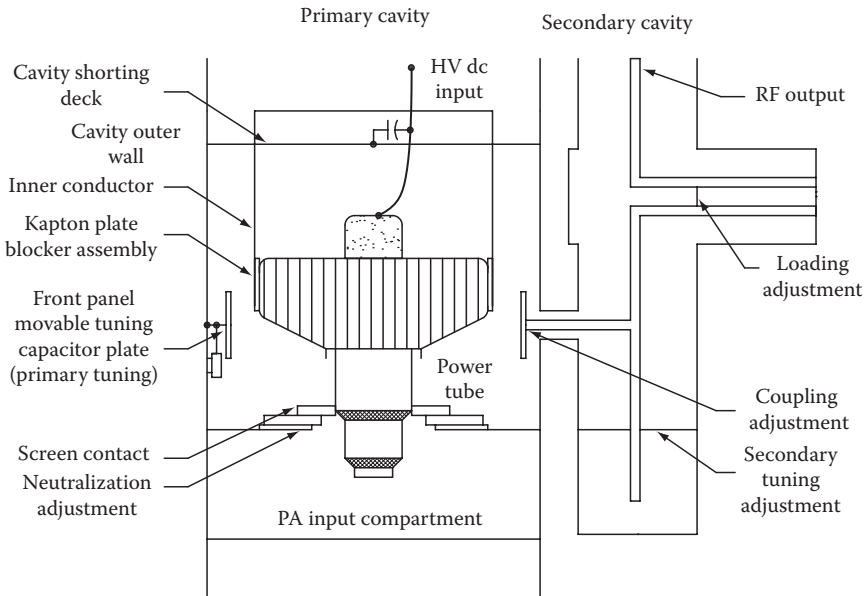


FIGURE 4.29
Double-tuned overcoupled broadband cavity amplifier.

4.2.14 Output Coupling

Coupling is the process by which RF energy is transferred from the amplifier cavity to the output transmission line. Wideband cavity systems use coupling to transfer energy from the primary cavity to the secondary cavity. Coupling in tube-type PA usually transforms a high (plate or cavity) impedance to a lower output (transmission line) impedance. Both capacitive (electrostatic) and inductive (magnetic) coupling methods are used in cavity RF amplifiers. In some designs, combinations of the two methods are used

4.2.14.1 Inductive Coupling

Inductive (magnetic) coupling employs the principles of transformer action. The efficiency of the coupling depends upon three conditions:

- The cross-sectional area under the coupling loop, compared to the cross-sectional area of the cavity (see Figure 4.30). This effect can be compared to the turns ratio principle of a transformer.
- The orientation of the coupling loop to the axis of the magnetic field. Coupling from the cavity is proportional to the cosine of the angle at which the coupling loop is rotated away from the axis of the magnetic field, as illustrated in Figure 4.31.

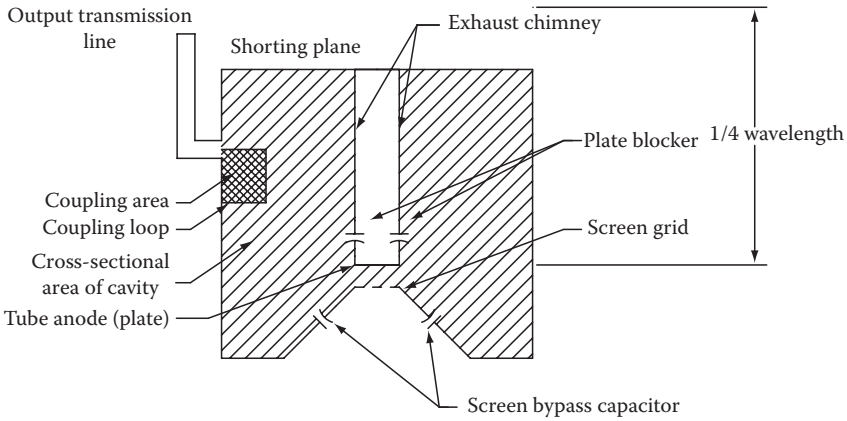


FIGURE 4.30
Use of inductive coupling in a 1/4-wavelength PA stage.

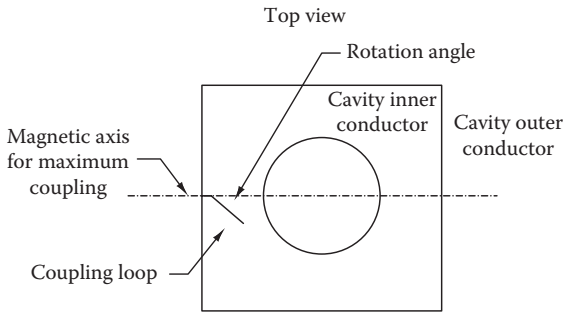


FIGURE 4.31
Top view of a cavity showing the inductive coupling loop rotated away from the axis of the magnetic field of the cavity.

- The amount of magnetic field that the coupling loop intercepts (see Figure 4.32). The strongest magnetic field will be found at the point of maximum RF current in the cavity. This is the area where maximum inductive coupling is obtained. Greater magnetic field strength also is found closer to the center conductor of the cavity. Coupling, therefore, is inversely proportional to the distance of the coupling loop from the center conductor.

In both 1/4- and 1/2-wavelength cavities, the coupling loop generally feeds a 50Ω transmission line (the load). The loop is in series with the load and has considerable inductance at very high frequencies. This inductance will reduce the RF current that flows into the load, thus reducing power output. This effect can be overcome by placing a variable capacitor in series with the output coupling loop. The load is connected to one end of the coupling loop,

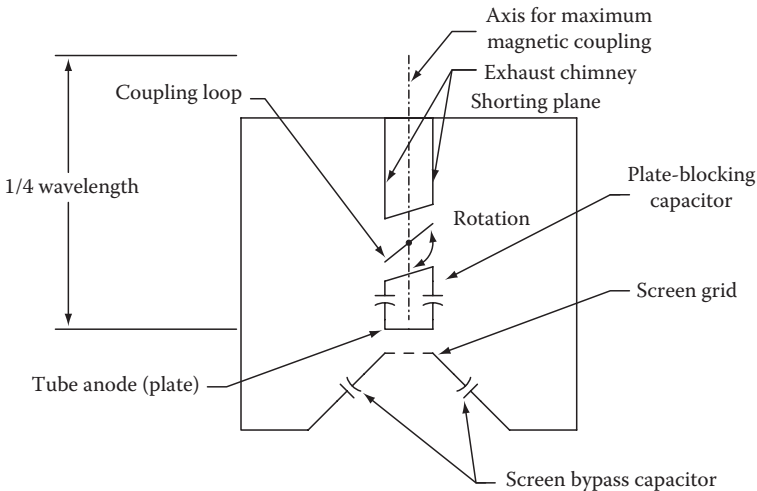


FIGURE 4.32 Cutaway view of the cavity showing the coupling loop rotated away from the axis of the magnetic field.

and the variable capacitor ties the other end of the loop to ground. The variable capacitor cancels some or all of the loop inductance. It functions as the PA-stage loading control.

Maximum loop current and output power occurs when the loading capacitor cancels all of the inductance of the loading loop. This lowers the plate impedance and results in heavier loading.

Light loading results if the loading capacitance does not cancel all of the loop inductance. The loop inductance that is not canceled causes a decrease in load current and power output, and an increase in plate impedance.

4.2.14.2 Capacitive Coupling

Capacitive (electrostatic) coupling, which physically appears to be straightforward, often baffles the applications engineer because of its unique characteristics. Figure 4.33 shows a cavity amplifier with a capacitive coupling plate positioned near its center conductor. This coupling plate is connected to the output load, which may be a transmission line or a secondary cavity (for wideband operation). The parameters that control the amount of capacitive coupling include

- The area of the coupling capacitor plate (the larger the area, the greater the coupling)
- The distance from the coupling plate to the center conductor (the greater the distance, the lighter the coupling)

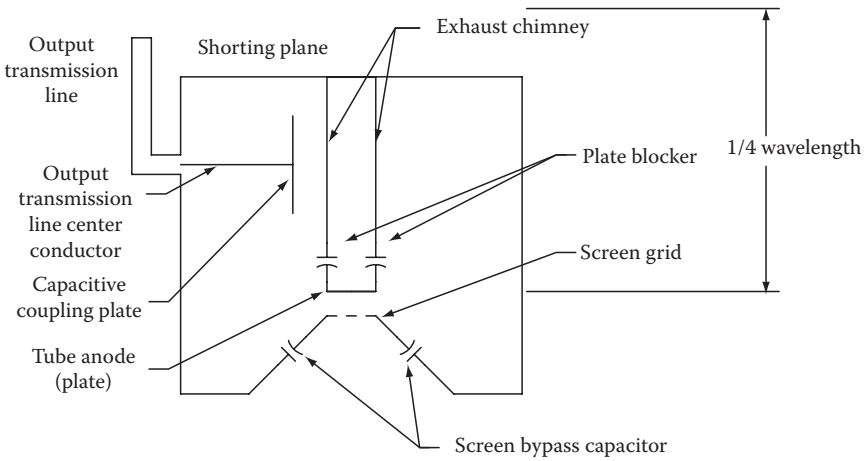


FIGURE 4.33
 $1/4$ -wavelength cavity with capacitive coupling to the output load.

Maximum capacitive coupling occurs when the coupling plate is at the maximum voltage point on the cavity center conductor.

To understand the effects of capacitive coupling, observe the equivalent circuit of the cavity. Figure 4.34 shows the PA tube, cavity (functioning as a parallel resonant circuit), and output section. The plate-blocking capacitor isolates the tube's dc voltage from the cavity. The coupling capacitor and output load are physically in series, but electrically they appear to be in parallel, as shown in Figure 4.35. The resistive component of the equivalent parallel circuit is increased by the coupling reactance. The equivalent parallel coupling reactance is absorbed into the parallel resonant circuit. This explains the need to retune the plate after changing the PA stage coupling (loading). The coupling reactance may be a series capacitor or inductor.

The series-to-parallel transformations are explained by the following formula:

$$R_p = \frac{R_s^2 + X_s^2}{R_s} \quad (4.12)$$

$$X_p = \frac{R_s^2 + X_s^2}{R_s} \quad (4.13)$$

where

R_p is the effective parallel resistance

R_s is the actual series resistance

X_s is the actual series reactance

X_p is the effective parallel reactance

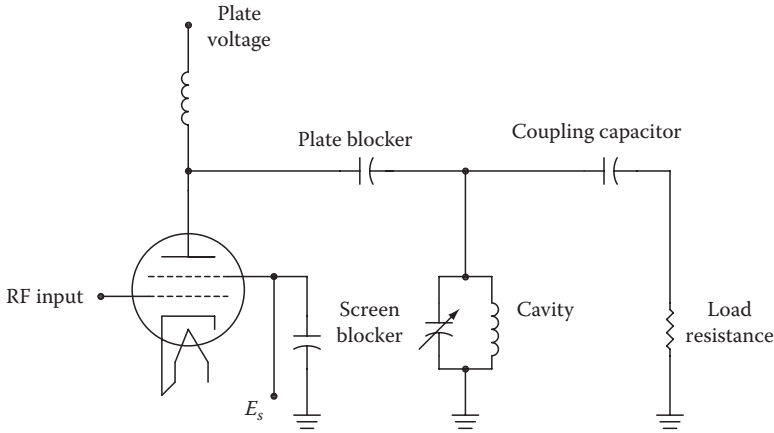


FIGURE 4.34 Equivalent circuit of a 1/4-wavelength cavity amplifier with capacitive coupling.

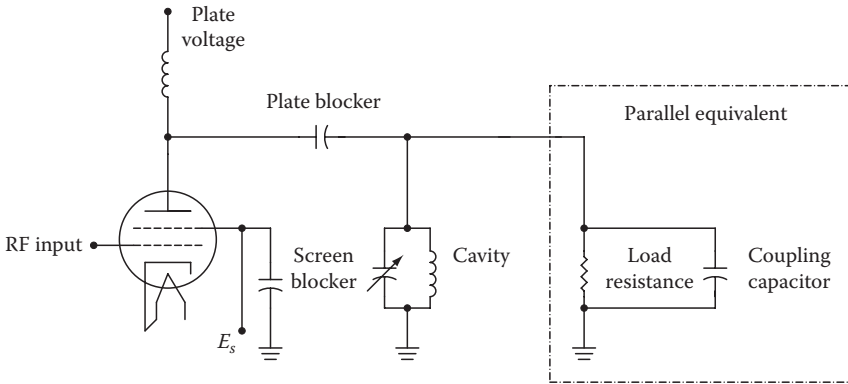


FIGURE 4.35 Equivalent circuit of a 1/4-wavelength cavity amplifier showing how series capacitive coupling appears electrically as a parallel circuit to the PA tube.

4.2.14.3 PA Loading

Proper loading of a cavity PA stage to the output transmission line is critical to dependable operation and optimum efficiency of the overall system. Light coupling produces light loading and results in a high plate impedance; conversely, heavy coupling results in heavier loading and a lower plate impedance. Maximum output power, coinciding with maximum efficiency and acceptable dissipation, dictates a specific plate impedance for a cavity of given design. This plate impedance is also dependent upon the operating values of dc plate voltage (E_p) and plate current (I_p).

Plate impedance dictates the cavity parameters of loaded Q , RF circulating current, and bandwidth. The relationships can be expressed as follows:

- Loaded Q is directly proportional to the plate impedance and controls the other two cavity parameters. Loaded $Q = X_p/X_i$, where Z_p is the cavity plate impedance and X_i is the cavity inductive reactance.
- Circulating current in the cavity is much greater (by a factor of the loaded Q) than the RF current supplied by the tube. Circulating current = $Q \times I_p$, where I_p is the RF current supplied to the cavity by the tube.
- The cavity bandwidth is dependent on the loaded Q and operating frequency. Bandwidth = F_r/Q , where F_r is the cavity resonant frequency.

Heavy loading lowers the PA plate impedance and cavity Q . A lower Q reduces the cavity RF circulating currents. In some cavities, high circulating currents can cause cavity heating and premature failure of the plate- or screen-blocking capacitors. The effects of lower plate impedance—a by-product of heavy loading—are higher RF and dc plate currents and reduced RF plate voltage. The instantaneous plate voltage is the result of the RF plate voltage added to the dc plate voltage. The reduced swing of plate voltage causes less positive dc screen current to flow. Positive screen current flows only when the plate voltage swings close to or below the value of the positive screen grid.

Light loading raises the plate impedance and cavity Q . A higher Q will increase the cavity circulating currents, raising the possibility of component overheating and failure. The effects of higher plate impedance are reduced RF and dc plate current and increased RF and dc plate voltage excursions. The higher cavity RF or peak dc voltages may cause arcing in the cavity.

There is one value of plate impedance that will yield optimum output power, efficiency, dissipation, and dependable operation. It is dictated by cavity design and the values of the various dc and RF voltages and currents supplied to the stage.

Depending on the cavity design, light loading may seem deceptively attractive. The dc plate voltage is constant (set by the power supply), and the lower dc plate current resulting from light loading reduces the tube's overall dc input power. The RF output power may change with light loading, depending on the plate impedance and cavity design, while efficiency will probably increase or, at worst, remain constant. Caution must be exercised with light loading, however, because of the increased RF voltages and circulating currents that such operation creates. Possible adverse effects include cavity arcing and overheating of cavity components, such as capacitors. The manufacturer's recommendations on PA tube loading should, therefore, be carefully observed.

Despite the many similarities among various cavity designs, each imposes its own set of operational requirements and limitations. No two cavity systems will tune up in exactly the same fashion. Given proper maintenance, a cavity amplifier will provide years of reliable service.

4.2.15 Mechanical Design

The operation of a cavity amplifier can be puzzling because of the nature of the major component elements. It is often difficult to relate the electrical schematic diagram to the mechanical assembly that exists within the transmitter. Consider the PA cavity schematic diagram shown in Figure 4.36. The grounded screen stage is of conventional design. Decoupling of the high-voltage power supply is accomplished by C1, C2, C3, and L1. Capacitor C3 is located inside the PA chimney (cavity inner conductor). The RF sample lines provide two low-power RF outputs for a modulation monitor or other test instrument. Neutralization inductors L3 and L4 consist of adjustable grounding bars on the screen grid ring assembly.

Figure 4.37 shows the electrical equivalent of the PA cavity schematic diagram. The 1/4-wavelength cavity acts as the resonant tank for the PA.

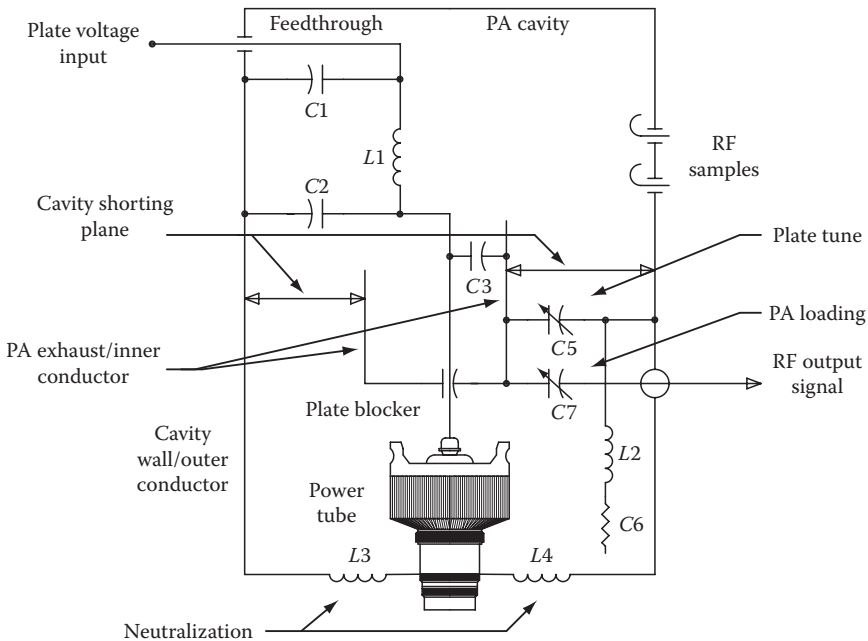


FIGURE 4.36
Typical VHF cavity amplifier.

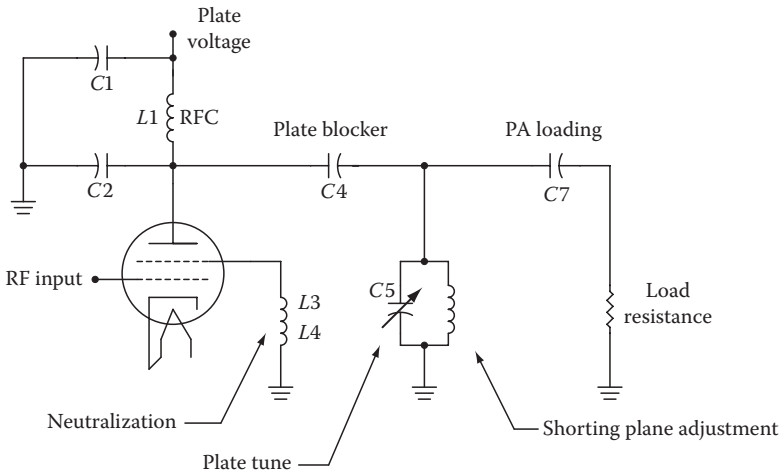
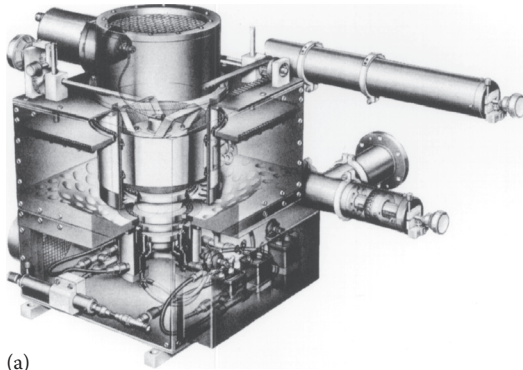


FIGURE 4.37
Electrical equivalent of the cavity amplifier shown in Figure 4.36.

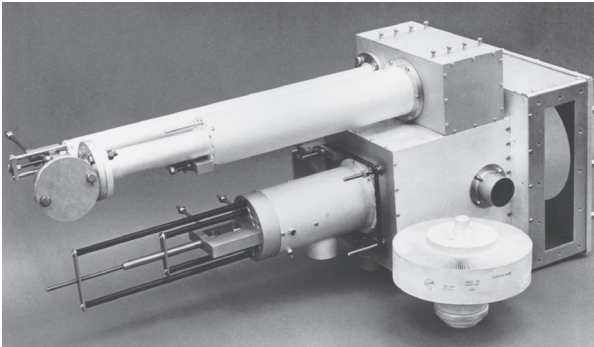
Coarse-tuning of the cavity is accomplished by adjustment of the shorting plane. Fine-tuning is performed by the PA tuning control, which acts as a variable capacitor to bring the cavity into resonance. The PA loading control consists of a variable capacitor that matches the cavity to the load. The assembly made up of $L2$ and $C6$ prevents spurious oscillations within the cavity.

Blocking capacitor $C4$ is constructed of a roll of *Kapton* insulating material sandwiched between two circular sections of aluminum. (*Kapton* is a registered trademark of Du Pont.) PA plate-tuning control $C5$ consists of an aluminum plate of large surface area that can be moved in or out of the cavity to reach resonance. PA loading control $C7$ is constructed much the same as the PA tuning assembly, with a large-area paddle feeding the harmonic filter, located external to the cavity. The loading paddle may be moved toward the PA tube or away from it to achieve the required loading. The $L2$ – $C6$ damper assembly actually consists of a 50Ω noninductive resistor mounted on the side of the cavity wall. Component $L2$ is formed by the inductance of the connecting strap between the plate-tuning paddle and the resistor. Component $C6$ is the equivalent stray capacitance between the resistor and the surrounding cavity box.

It can be seen that cavity amplifiers involve as much mechanical engineering as they do electrical engineering. The photographs of Figure 4.38 show graphically the level of complexity that a cavity amplifier may involve. Figure 4.38a depicts a VHF PA (Philips) with broadband input circuitry. Figure 4.38b shows a wideband VHF amplifier intended for television service (Varian, Palo Alto, CA). Figure 4.38c illustrates a VHF cavity amplifier designed for FM broadcast service (Varian, Palo Alto, CA).



(a)



(b)



(c)

FIGURE 4.38
VHF cavity amplifiers: (a) cross-sectional view of a broadband design, (b) cavity designed for television service, and (c) FM broadcast cavity amplifier.

4.3 High-Voltage Power Supplies

Virtually all power grid and microwave vacuum tubes require one or more sources of high voltage to operate. The usual source of operating power is a single-phase or multiphase supply operating from ac line current. Most supplies use silicon rectifiers as the primary ac-to-dc converting device.

Obviously, a power supply does not produce power for its load. Rather, it takes power from a source and reconditions the power so that the load can operate without damage of interference.

4.3.1 Silicon Rectifiers

Rectifier parameters generally are expressed in terms of reverse-voltage ratings and mean forward-current ratings in a 1/2-wave rectifier circuit operating from a 60 Hz supply and feeding a purely resistive load. The three principle reverse-voltage ratings are

- *Peak transient reverse voltage* (V_{rm}). The maximum value of any nonre-current surge voltage. This value must never be exceeded.
- *Maximum repetitive reverse voltage* ($V_{RM(rep)}$). The maximum value of reverse voltage that may be applied recurrently (in every cycle of 60 Hz power). This includes oscillatory voltages that may appear on the sinusoidal supply.
- *Working peak reverse voltage* ($V_{RM(wkg)}$). The crest value of the sinusoidal voltage of the ac supply at its maximum limit. Rectifier manufacturers generally recommend a value that has a significant safety margin, relative to the peak transient reverse voltage (V_{rm}), to allow for transient overvoltages on the supply lines.

There are three forward-current ratings of similar importance in the application of silicon rectifiers:

- *Nonrecurrent surge current* ($I_{FM,surge}$). The maximum device transient current that must not be exceeded at any time. $I_{FM,surge}$ is sometimes given as a single value, but more often is presented in the form of a graph of permissible surge-current values versus time. Because silicon diodes have a relatively small thermal mass, the potential for short-term current overloads must be given careful consideration.
- *Repetitive peak forward current* ($I_{FM,rep}$). The maximum value of forward current reached in each cycle of the 60 Hz waveform. This value does not include random peaks caused by transient disturbances.

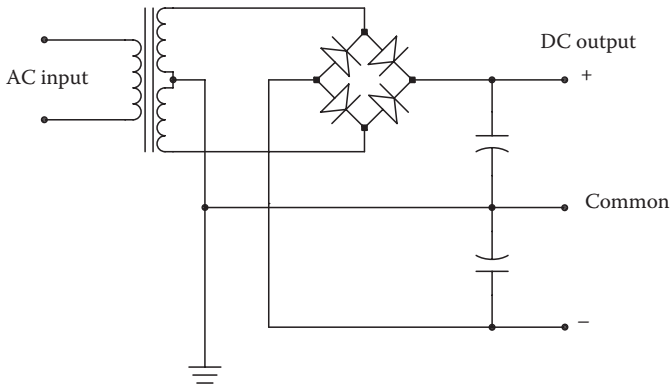


FIGURE 4.39
Conventional capacitor input filter full-wave bridge.

- *Average forward current* ($I_{FM,av}$). The upper limit for average load current through the device. This limit is always well below the repetitive peak forward-current rating to ensure an adequate margin of safety.

Rectifier manufacturers generally supply curves of the instantaneous forward voltage versus instantaneous forward current at one or more specific operating temperatures. These curves establish the forward-mode upper operating parameters of the device.

Figure 4.39 shows a typical rectifier application in a bridge circuit.

4.3.1.1 Operating Rectifiers in Series

High-voltage power supplies (5 kV and greater) often require rectifier voltage ratings well beyond those typically available from the semiconductor industry. To meet the requirements of the application, manufacturers commonly use silicon diodes in a series configuration to give the required working peak reverse voltage. For such a configuration to work properly, the voltage across any one diode must not exceed the rated peak transient reverse voltage (V_{rm}) at any time. The dissimilarity commonly found between the reverse leakage current characteristics of different diodes of the same type number makes this objective difficult to achieve. The problem normally is overcome by connecting shunt resistors across each rectifier in the chain, as shown in Figure 4.40. The resistors are chosen so that the current through the shunt elements (when the diodes are reverse biased) will be several times greater than the leakage current of the diodes themselves.

The *carrier storage* effect also must be considered in the use of a series-connected rectifier stack. If precautions are not taken, different diode recovery times (caused by the carrier storage phenomenon) will effectively force the full applied reverse voltage across a small number of diodes, or even a

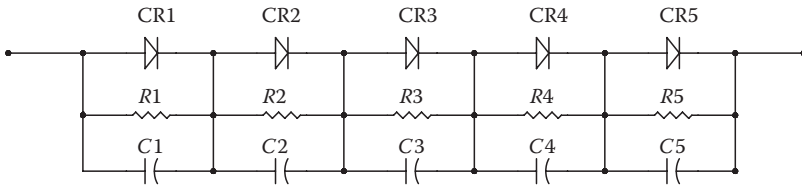


FIGURE 4.40
Portion of a high-voltage series-connected rectifier stack.

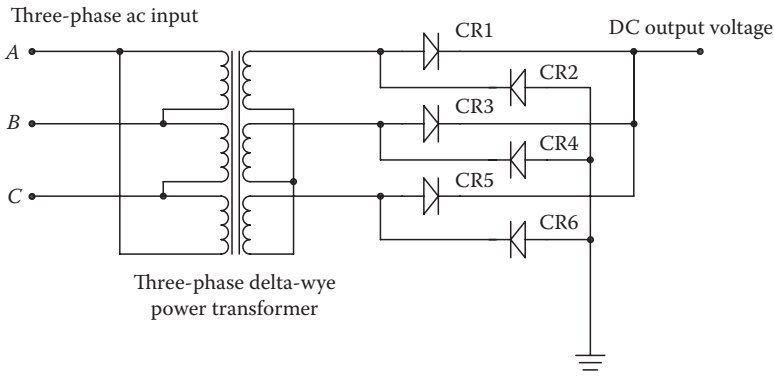


FIGURE 4.41
Three-phase delta-connected high-voltage rectifier.

single diode. This problem can be prevented by connecting small-value capacitors across each diode in the rectifier stack. The capacitors equalize the transient reverse voltages during the carrier storage recovery periods of the individual diodes.

Figure 4.41 illustrates a common circuit configuration for a high-voltage three-phase rectifier bank. A photograph of a high-voltage series-connected three-phase rectifier assembly is shown in Figure 4.42.

4.3.1.2 Operating Rectifiers in Parallel

Silicon rectifiers are used in a parallel configuration when a large amount of current is required from the power supply. *Current sharing* is the major design problem with a parallel rectifier assembly because diodes of the same type number do not necessarily exhibit the same forward characteristics.

Semiconductor manufacturers often divide production runs of rectifiers into tolerance groups, matching forward characteristics of the various devices. When parallel diodes are used, devices from the same tolerance group must be used to avoid unequal sharing of the load current. As a margin of safety, designers allow a substantial derating factor for devices in a parallel assembly to ensure that the maximum operating limits of any one component are not exceeded.

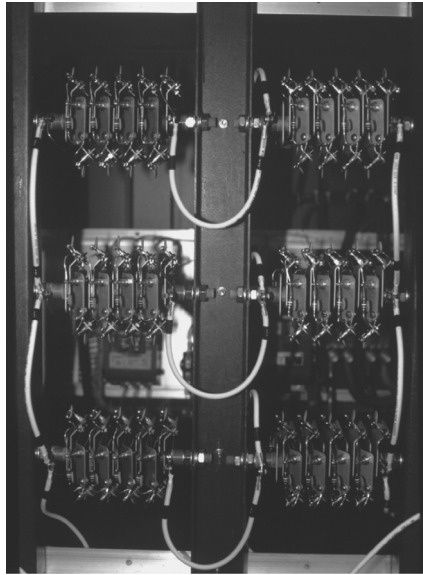


FIGURE 4.42
High-voltage rectifier assembly for a three-phase delta-connected circuit.

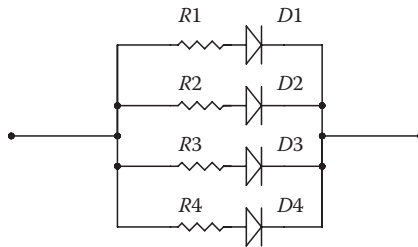


FIGURE 4.43
Using buildout resistances to force current sharing in a parallel rectifier assembly.

The problems inherent in a parallel rectifier assembly can be reduced through the use of a resistance or reactance in series with each component, as shown in Figure 4.43. The buildout resistances ($R1-R4$) force the diodes to share the load current equally. Such assemblies can, however, be difficult to construct and may be more expensive than simply adding diodes or going to higher-rated components.

4.3.1.3 Silicon Avalanche Rectifiers

The *silicon avalanche diode* is a special type of rectifier that can withstand high reverse power dissipation. For example, an avalanche diode with a normal forward rating of 10 A can dissipate a reverse transient of 8kW for 10ms without damage. This characteristic of the device allows elimination of the

surge-absorption capacitor and voltage-dividing resistor networks needed when conventional silicon diodes are used in a series rectifier assembly. Because fewer diodes are needed for a given applied reverse voltage, significant under-rating of the device (to allow for reverse-voltage transient peaks) is not required.

When an extra-high-voltage rectifier stack is used, it is still advisable to install shunt capacitors—but not resistors—in an avalanche diode assembly. The capacitors are designed to compensate for the effects of carrier storage and stray capacitance in a long series assembly.

4.3.2 Thyristor Servo Systems

Thyristor control of ac power has become a common method of regulating high-voltage power supplies. The type of servo system employed depends on the application. Figure 4.44 shows a basic single-phase ac control circuit

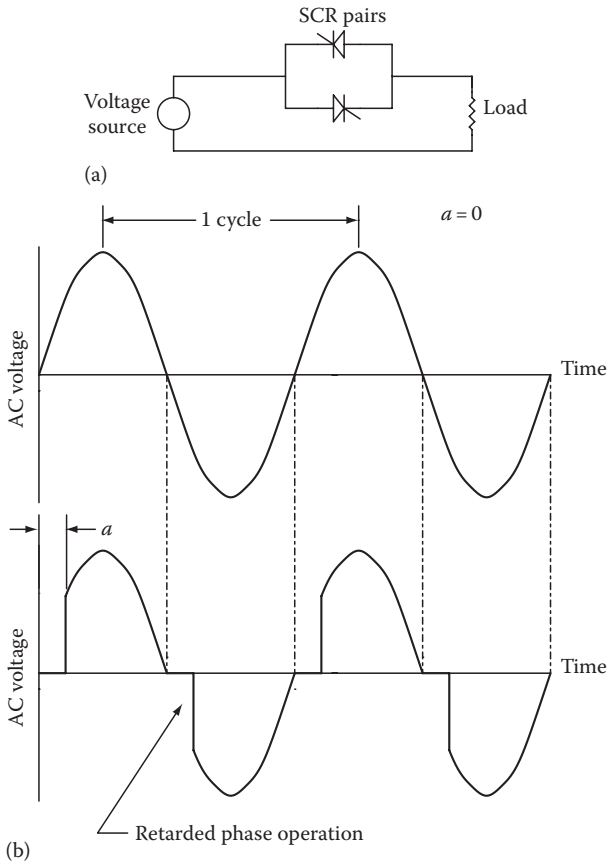


FIGURE 4.44

Inverse-parallel thyristor ac power control: (a) circuit diagram and (b) voltage and current waveforms for full and reduced thyristor control angles. The waveforms apply for a purely resistive load.

using discrete thyristors. The rms load current (I_{rms}) at any specific phase delay angle (α) is given in terms of the normal full-load rms current at a phase delay of zero (I_{rms-0}):

$$I_{rms} = I_{rms-0} \left\{ 1 - \frac{\alpha}{\pi} + (2\pi) - 1 \sin 2\alpha \right\}^{\frac{1}{2}} \quad (4.14)$$

The load rms voltage at any particular phase-delay angle bears the same relationship to the full-load rms voltage at zero phase delay as the previous equation illustrates for load current. An analysis of the mathematics shows that although the theoretical delay range for complete control of a resistive load is 0° – 180° , a practical span of 20° – 160° gives a power-control range of approximately 1%–99% of maximum output to the load. Figure 4.45 illustrates typical phase-control waveforms.

The circuit shown in Figure 4.44 requires a source of gate trigger pulses that must be isolated from each other by at least the peak value of the applied ac voltage. The two gate pulse trains must also be phased 180° with respect to each other. Also, the gate pulse trains must shift together with respect to the ac supply voltage phase when power throughput is adjusted.

Some power-control systems use two identical, but isolated, gate pulse trains operating at a frequency of twice the applied supply voltage (120 Hz for a 60 Hz system). Under such an arrangement, the forward-biased thyristor will fire when the gate pulses are applied to the silicon controlled rectifier (SCR) pair. The reverse-biased thyristor will not fire. Normally, it is considered unsafe to drive a thyristor gate positive while its anode–cathode is reverse biased. In this case, however, it may be permissible because the thyristor that is fired immediately conducts and removes the reverse voltage from the other device. The gate of the reverse-biased device is then being triggered on a thyristor that essentially has no applied voltage.

4.3.2.1 Inductive Loads

The waveforms shown in Figure 4.46 illustrate effects of phase control on an inductive load. When inductive loads are driven at a reduced conduction angle, a sharp transient change of load voltage occurs at the end of each current pulse (or loop). The transients generally have no effect on the load, but they can be dangerous to proper operation of the thyristors. When the conducting thyristor turns off, thereby disconnecting the load from the ac line supply, the voltage at the load rapidly drops to zero. This rapid voltage change, in effect, applies a sharply rising positive anode voltage to the thyristor opposing the device that has been conducting. If the thyristor dv/dt (change in voltage as a function of a change in time) rating is exceeded, the opposing device will turn on and conduction will take place, independent of any gate drive pulse.

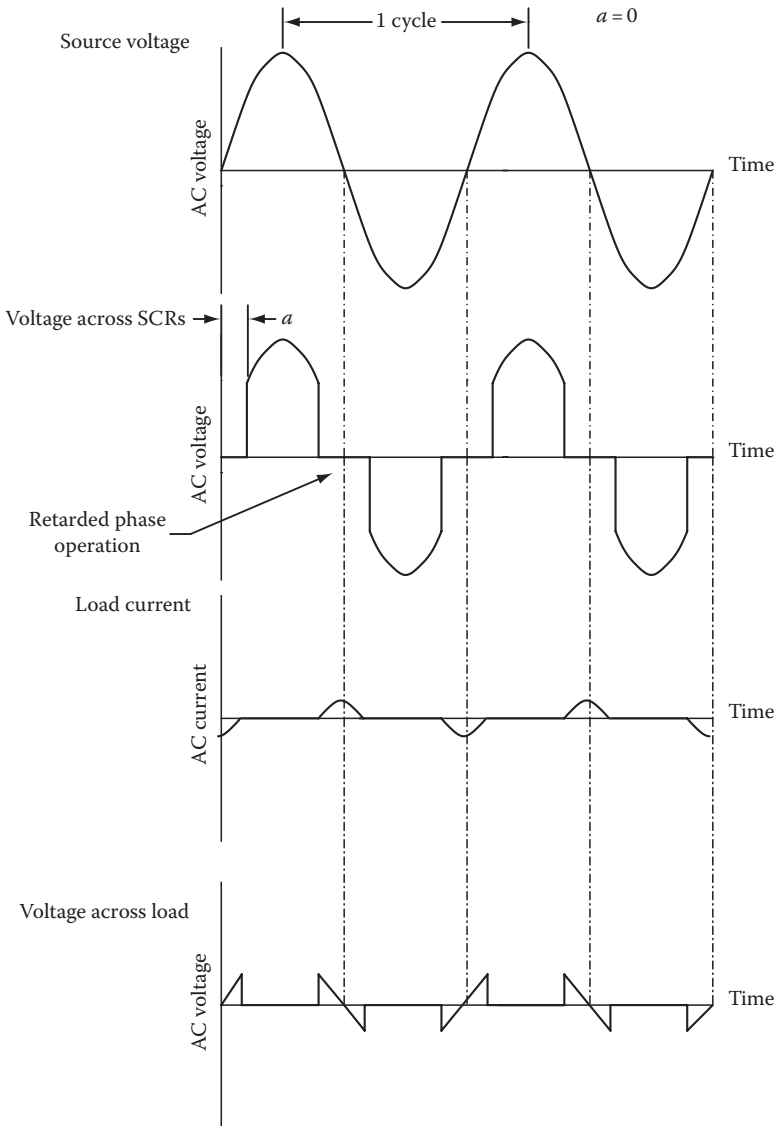


FIGURE 4.45
Waveforms in an ac circuit using thyristor power control.

A common protective approach involves the addition of a resistor–capacitor (RC) snubber circuit to control the rate of voltage change seen across the terminals of the thyristor pair. Whenever a thyristor pair is used to drive an inductive load, such as a power transformer, it is critically important that each device fires at a point in the applied waveform exactly 180° relative to the other. If proper timing is not achieved, the positive and negative current

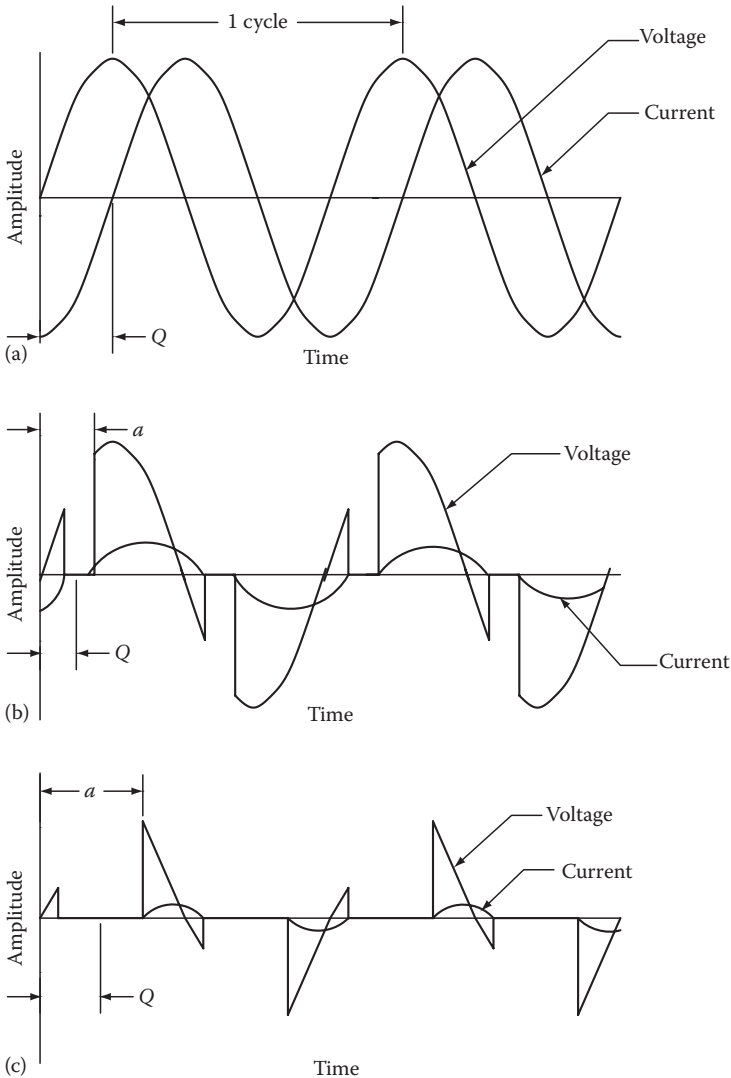


FIGURE 4.46 Voltage and current waveforms for inverse-parallel thyristor power control with an inductive load: (a) full conduction, (b) small-angle phase reduction, and (c) large-angle phase reduction.

loops will differ in magnitude, causing a direct current to flow through the primary side of the transformer. A common trigger control circuit, therefore, should be used to determine gate timing for thyristor pairs.

4.3.2.2 Applications

Several approaches are possible for thyristor power control in a three-phase ac system. The circuit shown in Figure 4.47 consists of essentially three

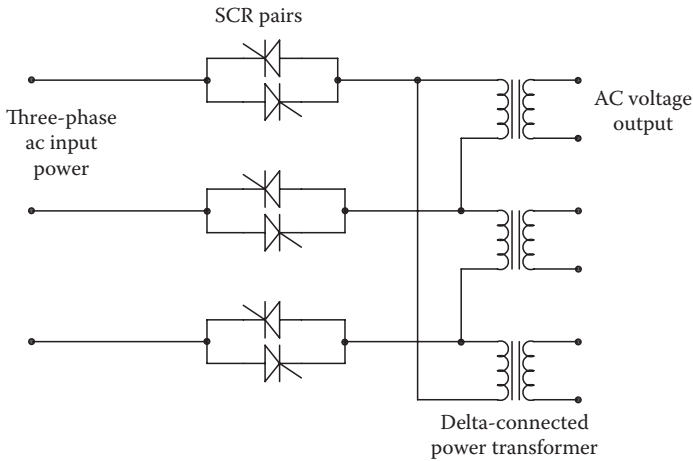


FIGURE 4.47

Modified full-thyristor three-phase ac control of an inductive delta load.

independent, but interlocked, single-phase thyristor controllers. This circuit is probably the most common configuration found in commercial and industrial equipment.

In a typical application, the thyristor pairs feed a power transformer with multitap primary windings, thereby giving the user an adjustment range to compensate for variations in utility company line voltages from one location to another. A common procedure specifies selection of transformer tap positions that yield a power output of 105% when nominal utility company line voltages are present. The thyristor power-control system then is used to reduce the angle of conduction of the SCR pairs as necessary to cause a reduction in line voltage to the power transformer to yield 100% rated power output from the power supply. A servo loop from a sample point at the load may be used to compensate automatically for line-voltage variations. With such an arrangement, the thyristors are kept within a reasonable degree of retarded phase operation. Line voltages will be allowed to sag 5% or so without affecting the dc supply output. Utility supply voltage excursions above the nominal value simply will result in further delayed triggering of the SCR pairs.

Thyristor control of high-power loads (200 kW and above) typically uses transformers that provide 6- or 12-phase outputs. Although they are more complicated and expensive, such designs allow additional operational control, and filtering requirements are reduced significantly. Figure 4.48 shows a six-phase *boost rectifier* circuit. The configuration consists basically of a full-wave three-phase SCR bridge connected to a wye-configured transformer secondary. A second bridge, consisting of six diodes, is connected to low-voltage taps on the same transformer. When the SCRs are fully on, their output is at a higher voltage than the diode bridge. As a result, the diodes are

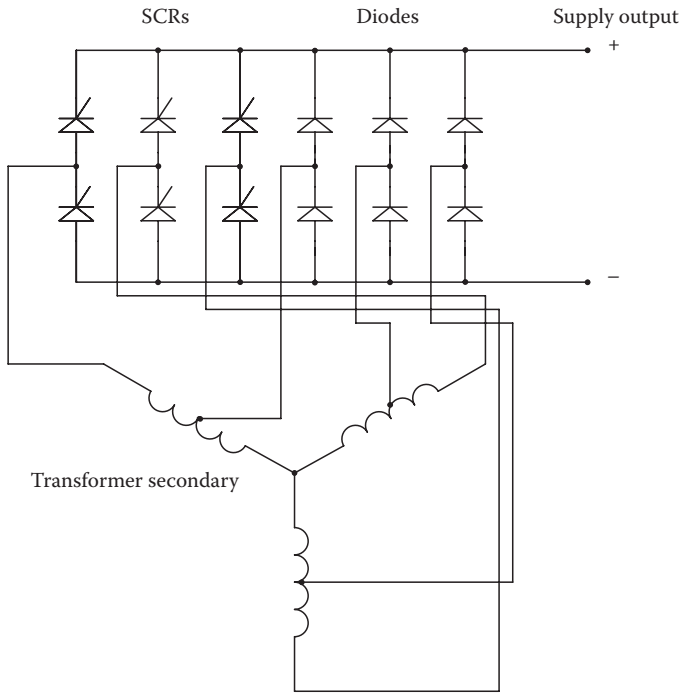


FIGURE 4.48
Six-phase boost rectifier circuit.

reverse biased and turned off. When the SCRs are partially on, the diodes are free to conduct. The diodes improve the quality of the output waveform during low-voltage (reduced conduction angle) conditions. The minimum output level of the supply is determined by the transformer taps to which the diodes are connected.

A thyristor-driven three-phase power-control circuit is shown in Figure 4.49. A single-phase power-control circuit is shown in Figure 4.50.

4.3.2.3 Triggering Circuits

Accurate, synchronized triggering of the gate pulses is a critical element in thyristor control of a three-phase power supply. The gate signal must be synchronized properly with the phase of the ac line that it is controlling. The pulse also must properly match the phase-delay angle of the gates of other thyristors in the power-control system. Lack of proper synchronization of gate pulse signals between thyristor pairs can result in improper current sharing (*current hogging*) among individual legs of the three-phase supply.

The gate circuit must be protected against electrical disturbances that could make proper operation of the power-control system difficult or unreliable. Electrical isolation of the gate is a common approach.

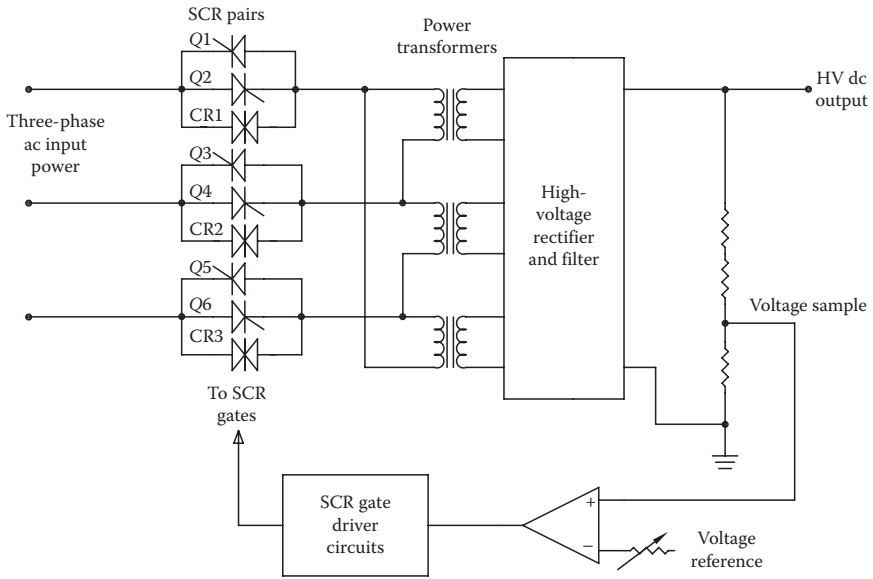


FIGURE 4.49
Thyristor-controlled high-voltage servo power supply.

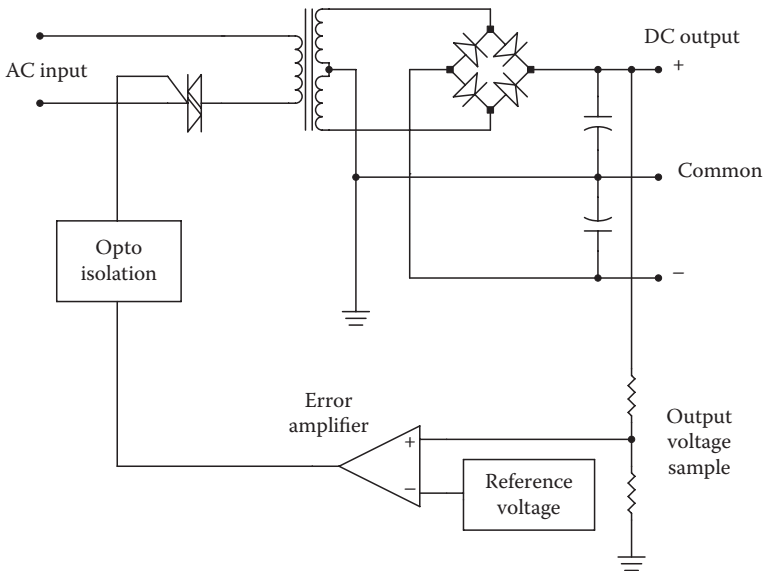


FIGURE 4.50
Phase-controlled power supply with primary regulation.

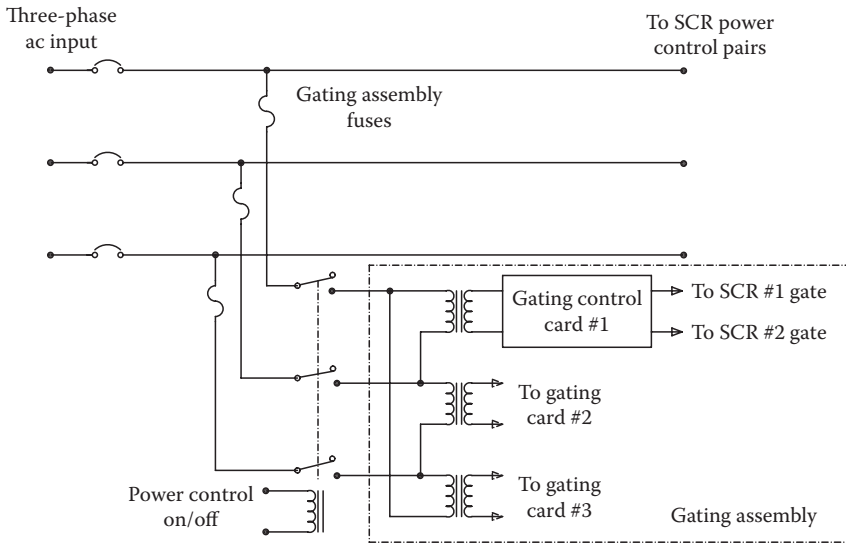


FIGURE 4.51
Simplified block diagram of the gating circuit for a phase-control system using back-to-back SCRs.

Standard practice calls for the use of gate pulse transformers in thyristor servo system gating cards. Pulse transformers are ferrite-cored devices with a single primary winding and (usually) multiple secondary windings that feed, or at least control, the individual gates of a back-to-back thyristor pair. This concept is illustrated in Figure 4.51. Newer thyristor designs may use optocouplers to achieve the necessary electrical isolation between the trigger circuit and the gate.

It is common practice to tightly twist together the control leads from the gate and cathode of a thyristor to the gating card assembly. This practice provides a degree of immunity to high-energy pulses that might inadvertently trigger the thyristor gate. The gate circuit must be designed and configured carefully to reduce inductive and capacitive coupling that might occur between power and control circuits. Because of the high di/dt conditions commonly found in thyristor-controlled power circuits, power wiring and control (gate) wiring must be separated physically as much as possible. Shielding of gating cards in a metal card cage is advisable.

Equipment manufacturers use various means to decrease gate sensitivity to transient sources, including placement of a series resistor in the gate circuit and/or a shunting capacitor between the gate and cathode. A series resistor has the effect of decreasing gate sensitivity, increasing the allowable dv/dt of the thyristor and reducing the turn-off time, which simultaneously increases the required holding and latching currents. The use of a shunt capacitor between the gate and cathode leads reduces high-frequency noise

components that might be present on the gate lead and increases the dv/dt withstand capability of the thyristor.

4.3.2.4 Fusing

Current limiting is a basic method of protection for any piece of equipment operated from the utility ac line. The device typically used for breaking fault currents is either a fuse or a circuit breaker. Some designs incorporate both components. *Semiconductor fuses* often are used in conjunction with a circuit breaker to provide added protection. Semiconductor fuses operate more rapidly (typically within 8.3 ms) and more predictably than common fuses or circuit breakers. Surge currents caused by a fault can destroy a semiconductor device, such as a power thyristor, before the ac line circuit breaker has time to act. Manufacturers of semiconductor fuses and thyristors usually specify in their data sheets the I^2t ratings of each device. Because the thyristor rating normally assumes that the device is operating at maximum rated current and maximum junction temperature (conditions that do not represent normal operation), a safety factor is ensured.

4.3.2.5 Control Flexibility

Thyristor servo control of a high-voltage power supply is beneficial to the user for a number of reasons, including

- Wide control over ac input voltages
- Capability to compensate automatically for line-voltage variations
- Capability to soft-start the dc supply

Thyristor control circuits typically include a ramp generator that increases the ac line voltage to the power transformer from zero to full value within 2–5 s. This prevents high surge currents through rectifier stacks and filter capacitors during system start-up.

Although thyristor servo systems are preferred over other power-control approaches from an operational standpoint, they are not without their drawbacks. The control system is complex and can be damaged by transient activity on the ac power line. Conventional power contactors are simple and straightforward. They either make contact or they do not. For reliable operation of the thyristor servo system, attention must be given to transient suppression at the incoming power lines.

4.3.3 Polyphase Rectifier Circuits

High-voltage power supplies typically used in vacuum tube circuits incorporate multiphase rectification of the ac line voltage. Common configurations

include 3-, 6-, and 12-phase, with 3-phase rectification being the most common. Figure 4.52 illustrates four approaches to three-phase rectification:

- *Three-phase half-wave wye* (Figure 4.52a). Three half-wave rectifiers are used in each leg of the secondary *wye* forming one phase. In such an arrangement, each diode carries current one-third of each cycle, and the output wave pulses at three times the frequency of the ac supply. In order to avoid direct-current saturation in the transformer, it is necessary to employ a three-phase transformer rather than three single-phase transformers.
- *Three-phase full-wave bridge* (Figure 4.52b). Six diodes are used in this circuit to produce a low ripple output with a frequency of six times the input ac waveform. It is permissible in this configuration to use three single-phase transformers, if desired.
- *Six-phase star* (Figure 4.52c). This circuit, also known as a three-phase diametric configuration, uses six diodes with a transformer secondary configured as a star, as illustrated in the figure. The output ripple frequency is six times the input ac waveform.
- *Three-phase double-wye* (Figure 4.52d). This circuit uses six diodes and a complicated configuration of transformer windings. Note the balance coil (*interphase transformer*) in the circuit.

The relative merits of these rectifier configurations are listed in Table 4.3.

Polyphase rectifiers are used when the dc power required is on the order of 2kW or more. The main advantages of a polyphase power supply over a single-phase supply include the following:

- Division of the load current between three or more lines to reduce line losses.
- Significantly reduced filtering requirements after rectification because of the low ripple output of a polyphase rectifier.
- Improved voltage regulation when using an inductive-input filter. Output voltage *soaring* is typically 6% or less from full-load to no-load conditions.
- Greater choice of output voltages from a given transformer by selection of either a *delta* or *wye* configuration.

The main disadvantage of a polyphase system is its susceptibility to phase imbalance. Resulting operational problems include increased ripple at the output of the supply and uneven sharing of the load current by the transformer windings.

4.3.4 Power Supply Filter Circuits

The filter network for a high-voltage power supply typically consists of a series inductance and one or more shunt capacitances. Bleeder resistors and

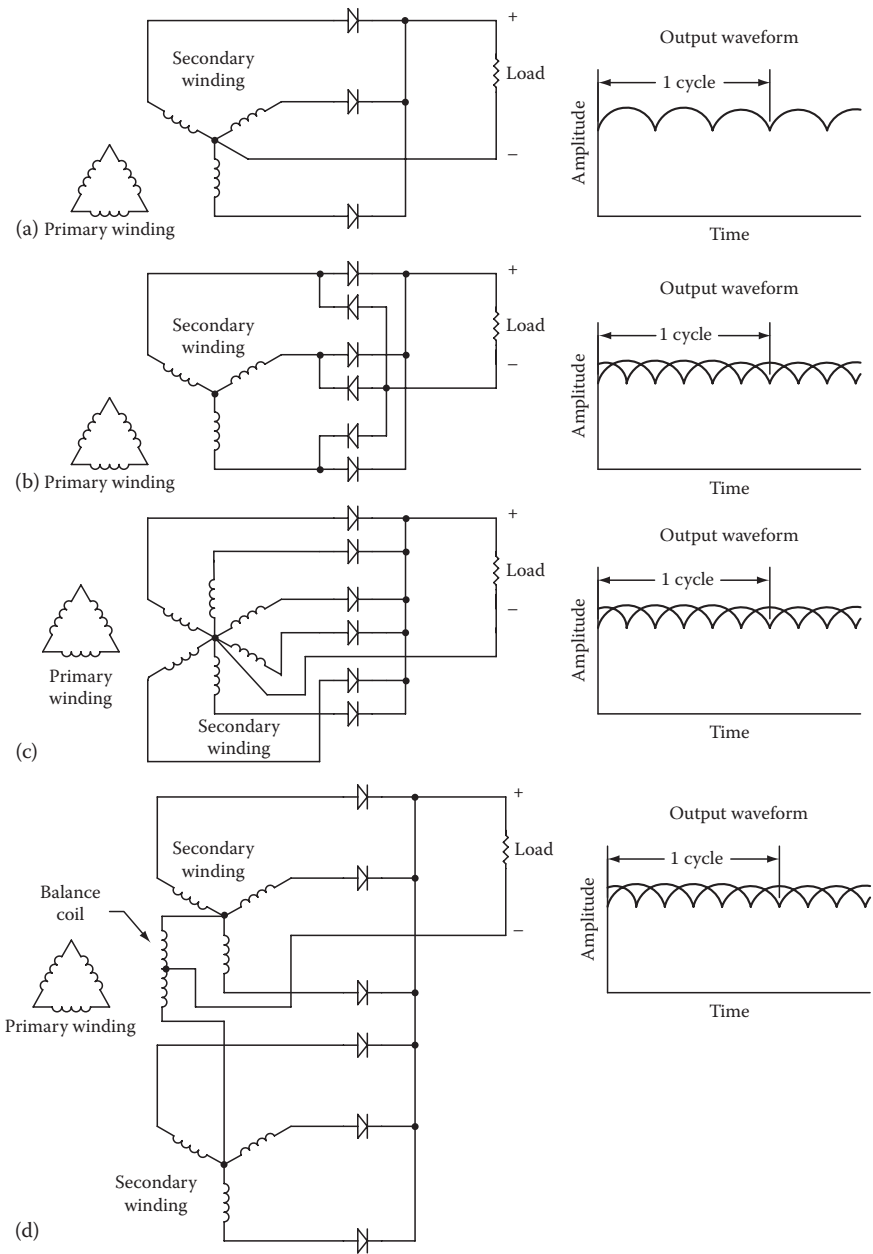


FIGURE 4.52 Basic three-phase rectifier circuits: (a) half-wave wye, (b) full-wave bridge, (c) six-phase star, and (d) three-phase double-wye.

TABLE 4.3
Operating Parameters of Three-Phase Rectifier Configurations

Parameter	Three-Phase Star	Three-Phase Bridge	Six-Phase Star	Three-Phase Double-Y	Multiplier ^a
Rectifier elements	3	6	6	6	
rms dc voltage output	1.02	1.00	1.00	1.00	Average dc voltage output
Peak dc voltage output	1.21	1.05	1.05	1.05	Average dc voltage output
Peak reverse volts per rectifier	2.09	1.05	2.09	2.42	Average dc voltage output
	2.45	2.45	2.83	2.83	rms secondary volts per transformer leg
	1.41	1.41	1.41	1.41	rms secondary volts line-to-line
Average dc output current per rectifier	0.333	0.333	0.167	0.167	Average dc output current
rms current per rectifier, resistive load	0.587	0.579	0.409	0.293	Average dc output current
rms current per rectifier, inductive load	0.578	0.578	0.408	0.289	Average dc output current
Percent ripple	18.3	4.2	4.2	4.2	
Ripple frequency	3	6	6	6	Line frequency
AC line power factor	0.826	0.955	0.955	0.955	
Transformer secondary rms volts per leg ^b	0.855	0.428	0.740	0.855	Average dc voltage output
Transformer secondary rms volts line-to-line	1.48	0.740	1.48 (max)	1.71 (max, no load)	Average dc voltage output
Secondary line current	0.578	0.816	0.408	0.289	Average dc output current
Transformer secondary volt-amperes	1.48	1.05	1.81	1.48	DC watts output
Primary line current	0.817	1.41	0.817	0.707	(Average load 1 × secondary leg V)/ primary line V

^a To determine the value of a parameter in any column, multiply the factor shown by the value given in this column.

^b For inductive load or large choke input filter.

various circuit protection devices are usually incorporated as well. Filter systems can be divided into two basic types:

- *Inductive input:* filter circuits that present a series inductance to the rectifier output
- *Capacitive input:* filter circuits that present a shunt capacitance to the rectifier output

4.3.4.1 Inductive Input Filter

The inductive input filter is the most common configuration found in high-power RF equipment. A common circuit is shown in Figure 4.53, along with typical current waveforms. When the input inductance is infinite, current through the inductance is constant and is carried at any moment by the rectifier anode that has the most positive voltage applied to it at that instant. As the alternating voltage being rectified passes through zero, the current suddenly transfers from one anode to another, producing square current waves through the individual rectifier devices.

When the input inductance is finite (but not too small), the situation changes to that shown by the solid lines of Figure 4.53. The current through the input inductance tends to increase when the output voltage of the rectifier exceeds the average or dc value, and to decrease when the rectifier output voltage is less than the dc value. This causes the current through the individual anodes to be modified as shown. If the input inductance is too small, the current decreases to zero during a portion of the time between the peaks of the rectifier output voltage. The conditions then correspond to a capacitor input filter system.

The output wave of the rectifier can be considered as consisting of a dc component upon which are superimposed ac voltages (ripple voltages). To a first approximation, the fluctuation in output current resulting from a finite input inductance can be considered as the current resulting from the lowest-frequency component of the ripple voltage acting against the impedance of the input inductance. This assumption is permissible because the higher-frequency components in the ripple voltage are smaller, and at the same time encounter higher impedance. Furthermore, in practical filters the shunting capacitor following the input inductance has a small impedance at the ripple frequency compared with the reactance of the input inductance. The peak current resulting from a finite input inductance, therefore, is given approximately by the relation [2]

$$\frac{I_f}{I_i} = 1 + \frac{E_1}{E_0} \frac{R_{eff}}{\omega L_1} \quad (4.15)$$

where

I_f is the peak current with finite input inductance

I_i is the peak current with infinite input inductance

E_1/E_0 is the ratio of lowest-frequency ripple component to the dc voltage in the rectifier output

R_{eff} is the effective load resistance

ωL_1 is the reactance of the incremental value of the input inductance at the lowest ripple frequency

This equation is derived as follows:

- The peak alternating current through the input inductance is approximately $E_1/\omega L_1$.
- The average or dc current is E_0/R_{eff} .

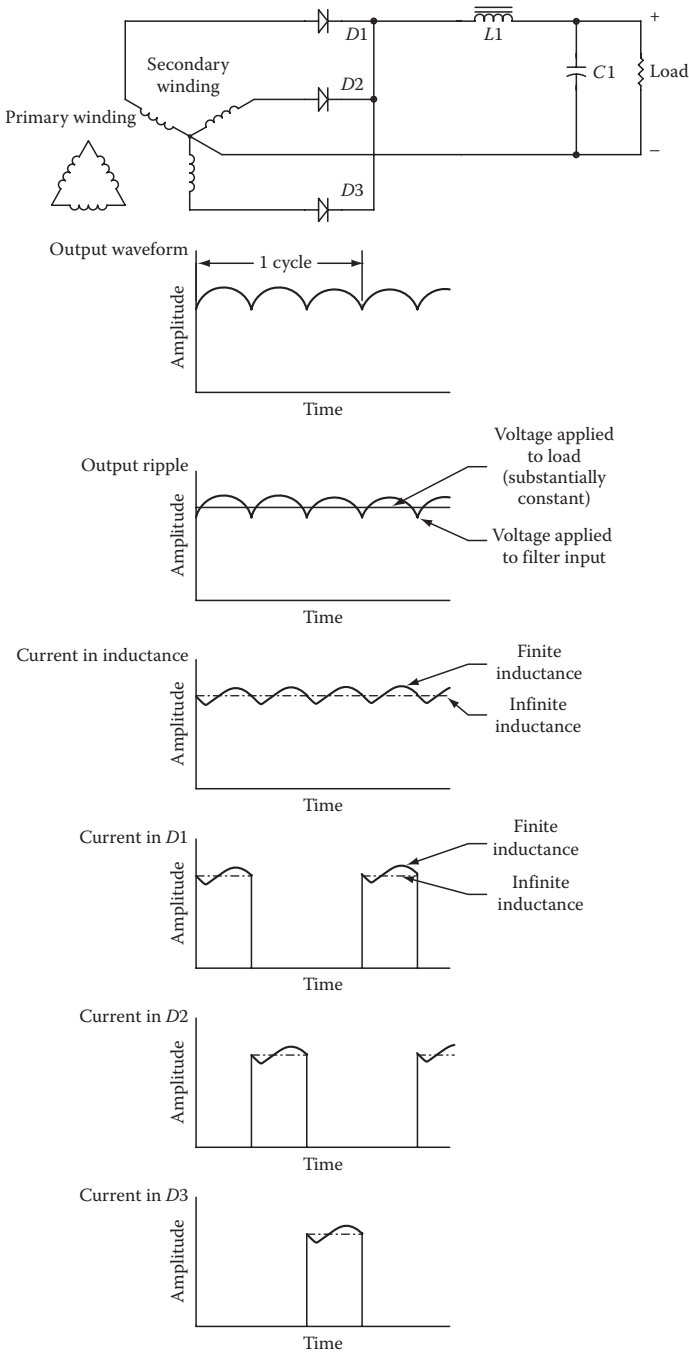


FIGURE 4.53 Voltage and current waveshapes for an inductive-input filter driven by a three-phase supply.

- The peak current with finite inductance is, therefore, $(E_1/\omega L_1) + (E_0/R_{eff})$.
- The current with infinite inductance is E_0/R_{eff} .

The effective load resistance value consists of the actual load resistance plus filter resistance plus equivalent diode and transformer resistances.

The normal operation of an inductive input filter requires that there be a continuous flow of current through the input inductance. The peak alternating current flowing through the input inductance, therefore, must be less than the dc output current of the rectifier. This condition is realized by satisfying the approximate relation

$$\omega L_1 = R_{eff} \frac{E_1}{E_0} \quad (4.16)$$

In the practical case of a 60Hz single-phase full-wave rectifier circuit, the foregoing equation becomes

$$L_1 = \frac{R_{eff}}{1130} \quad (4.17)$$

In a polyphase system, the required value of L_1 is significantly less. The higher the load resistance (the lower the dc load current), the more difficult it is to maintain a continuous flow of current and, with a given L_1 , the previous equation will not be satisfied when the load resistance exceeds a critical value.

The minimum allowable input inductance (ωL_1) is termed the *critical inductance*. When the inductance is less than the critical value, the filter acts as a capacitor input circuit. When the dc drawn from the rectifiers varies, it is still necessary to satisfy the ωL_1 equation at all times, particularly if good voltage regulation is to be maintained. To fulfill this requirement at small load currents without excessive inductance, it is necessary to place a bleeder resistance across the output of the filter system in order to limit R_{eff} to a value corresponding to a reasonable value of L_1 .

4.3.4.2 Capacitive Input Filter

When a shunt capacitance rather than a series inductance is presented to the output of a rectifier, the behavior of the circuit is greatly modified. Each time the positive crest alternating voltage of the transformer is applied to one of the rectifier anodes, the input capacitor charges up to just slightly less than this peak voltage. The rectifier then ceases to deliver current to the filter until another anode approaches its peak positive potential, when the capacitor is charged again. During the interval when the voltage across the input capacitor is greater than the potential of any of the anodes, the voltage across the input capacitor drops off nearly linearly with time, because the first filter inductance draws a substantially constant current from the input capacitor. A typical set of voltage and current waves is illustrated in Figure 4.54.

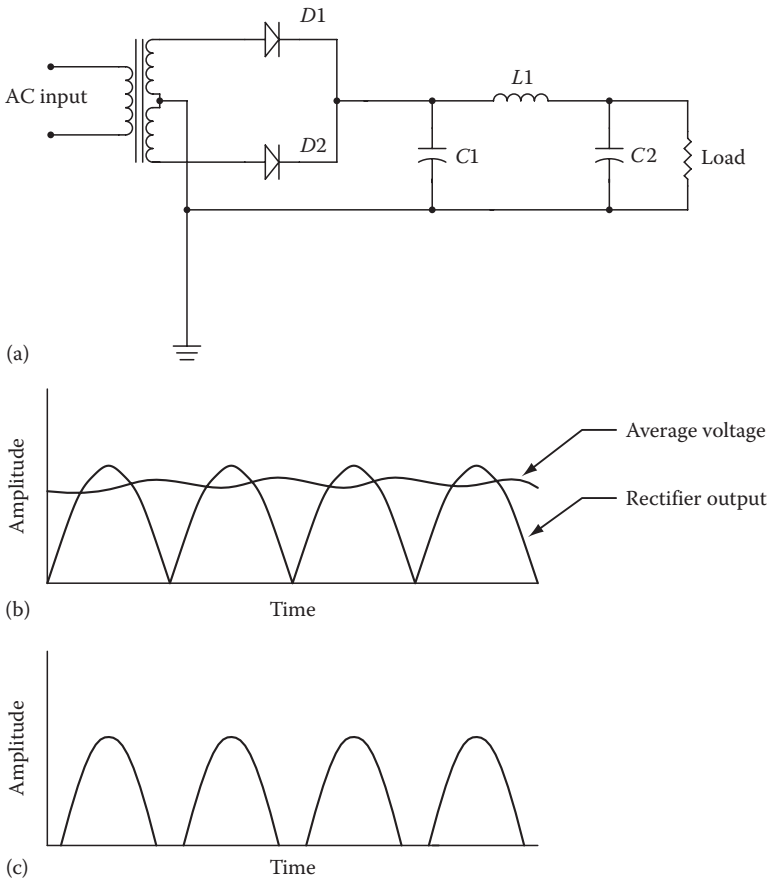


FIGURE 4.54

Characteristics of a capacitive input filter circuit: (a) schematic diagram, (b) voltage waveshape across the input capacitor and (c) waveshape of current flowing through the diodes.

The addition of a shunt capacitor to the input of a filter produces fundamental changes in behavior, including the following:

- The output voltage is appreciably higher than with an inductance input.
- The ripple voltage is less with a capacitance input filter than with an inductance input filter.
- The dc voltage across the filter input drops as the load current increases for the capacitive input case, instead of being substantially constant as for the inductive input case.
- The ratio of peak-to-average anode current at the rectifiers is higher in the capacitive case.
- The utilization factor of the transformer is less with a capacitive input configuration.

Filters incorporating shunt capacitor inputs generally are employed when the amount of dc power required is small. Inductance input filters are used when the amount of power involved is large; the higher utilization factor and lower peak current result in important savings in rectifier and transformer costs under these conditions. Inductance input systems are almost universally employed in polyphase rectifiers.

4.3.5 Parameter Sampling Circuits

Most RF system parameters can be reduced to a *sample voltage* by using voltage dividers or current-sense resistors in the circuit to be measured [3]. Remote metering is accomplished by sending a signal representing the sample voltage to the remote metering point, and then displaying a value representing the original parameter. In practice, the actual sampling ratio (sample voltage/parameter value) is not particularly important as long as the sampling ratio is stable and the sample voltage is reasonable (i.e., low enough to be easily handled, yet high enough to be immune to noise). Figure 4.55 shows some typical voltage sampling circuits for high-power RF systems.

Figure 4.55a shows how a voltage divider can be used to sample plate voltage. The circuit is independent of the transmitter front panel plate voltage meter. A zener clamping diode is placed across the sample line for safety; if

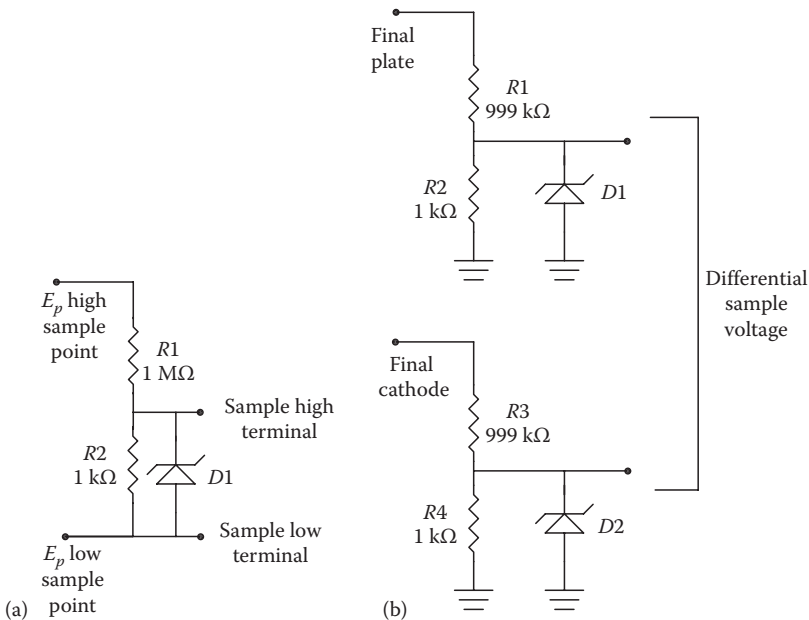


FIGURE 4.55 Plate voltage sampling circuits: (a) single ended and (b) differential.

the shunt resistor opened, the full plate voltage would be present at the sampling output terminals (assuming minimal loading by the remote indicator). The zener will typically fail in a short-circuit mode, protecting the external sampling equipment.

The sampling circuit of Figure 4.55a is adequate for systems that measure plate voltage from a ground reference. In the event that the cathode operates at a potential other than ground, the sampling circuit of Figure 4.55b may be used. The sampling line is balanced, with both terminals floating above ground by $1\text{ k}\Omega$. The differential output voltage is independent of the common mode input voltage if all the resistors are precisely matched.

Several techniques are available for dealing with differential sampling voltages (where neither side of the sample is grounded). Typical approaches include the following:

- *Isolation amplifier.* Isolation amplifiers provide isolation between the input and output by using magnetic, capacitive, optical, or thermal coupling between the input and output. Some units include feedback techniques to reduce nonlinearities induced by the coupling method.
- *Differential amplifier.* In most cases, electrical isolation between the sampling circuit and the remote metering equipment is not required. Instead, a differential voltage must be changed to a *single-ended* voltage (one side grounded). This can be accomplished using the differential amplifier shown in Figure 4.56a or the instrumentation amplifier shown in Figure 4.56b.

Measurement of plate and/or screen current is typically done by inserting a sampling resistor into the ground return lead of the respective supplies. This approach is shown in Figure 4.57. Notice that the remote plate current

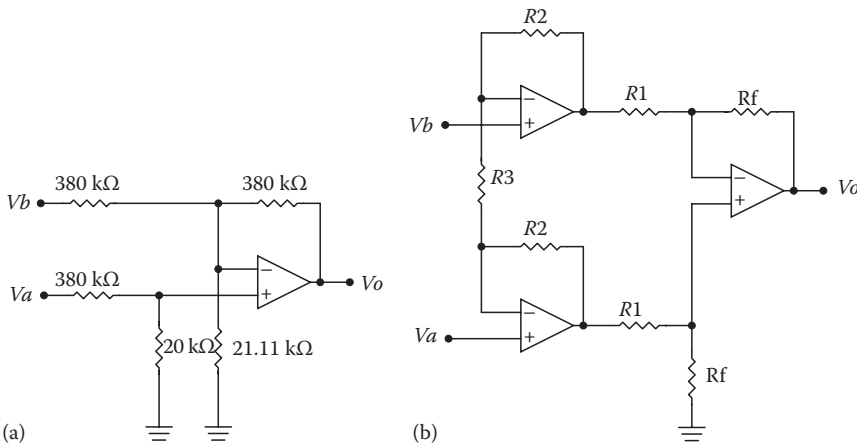


FIGURE 4.56 Differential amplifier circuits: (a) basic circuit and (b) instrumentation amplifier.

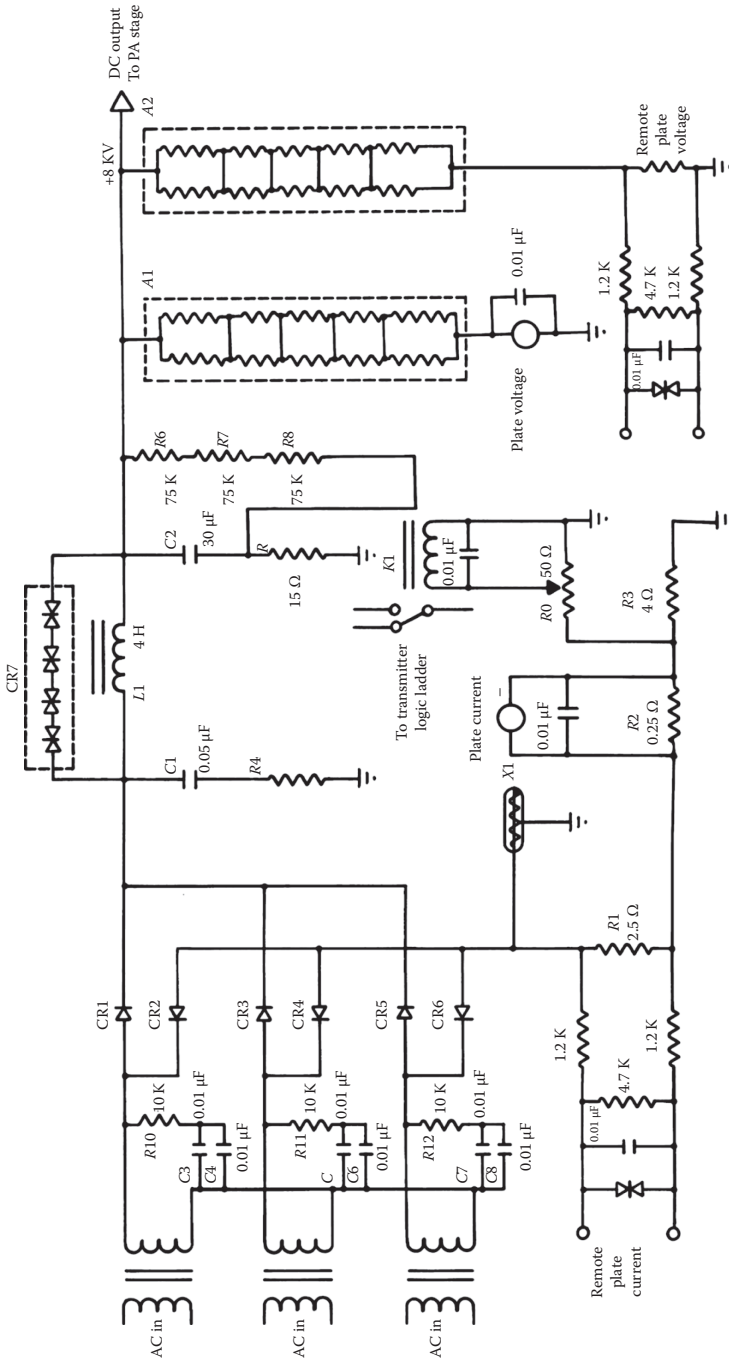


FIGURE 4.57 Transmitter high-voltage power supply showing remote plate current and voltage sensing circuits.

metering is taken across $R1$, with balancing resistors of $1.2\text{ k}\Omega$ used to lift the sample above the ground reference. A bidirectional zener diode is placed across the output of the sample terminals for protection. The plate current meter on the transmitter is sampled across $R2$, and the overcurrent relay is sampled across $R3$. Device $X1$ is a crowbar that will shunt the plate current to ground in the event of a failure in the $R1$ – $R3$ chain.

The initial tolerance of resistors used as current sense elements or in a voltage divider circuit is usually not particularly critical. An inexact resistance will cause the sample voltage to be something other than that expected. However, whatever sample voltage appears, the remote meter can be calibrated to agree with the local meter. The more important attribute is resistance changes with time and temperature.

A carbon composition resistor can commonly have a temperature coefficient of about $\pm 600\text{ ppm}/^\circ\text{C}$ for resistors up to $1\text{ k}\Omega$. This increases to $\pm 1875\text{ ppm}/^\circ\text{C}$ for resistors up to $1\text{ M}\Omega$. If these resistors were used in a common plate voltage sampling circuit, such as that shown in Figure 4.57, the sample voltage could change as much as 20% over a 0°C – 50°C temperature range. Metal film resistors are available with temperature coefficients as low as $15\text{ ppm}/^\circ\text{C}$. Bulk metal resistors are available with temperature coefficients of $5\text{ ppm}/^\circ\text{C}$. If the resistors in a voltage divider are at the same temperature, the matching between the *temperature coefficient of resistance* (TCR) is more important than the actual TCR. Because the TCR matching is generally better than the TCR of an individual resistor, sampling systems should be designed to keep the resistors in the network at the same operating temperature.

In differential amplifiers, the matching of resistor values is critical. To insure that the resistance values track with temperature, it is common practice to use a resistor network made using hybrid film techniques or integrated circuit techniques. Because the resistors are made at the same time and of the same materials, the TCRs track closely.

4.3.6 Protection Systems

Monitoring circuits are often built into a power supply to check for elevated temperature, output overvoltage, undervoltage, and overcurrent. Overvoltage and overcurrent protection are necessary to limit the propagation of damage in the system. In order to accomplish this protection, something needs to happen to clear the fault. Options include a series relay that will remove power from the supply, which is either latched open or reset after a specified length of time. The overcurrent situation can be addressed by a *brick wall* current limit or current *foldback*. Limiting the maximum current that the supply can deliver establishes the brick wall limit. Undervoltage conditions do not usually cause damage in a system, but they can lead to unreliable operation or operation outside of established parameters.

References

1. Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.
2. Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, p. 560, 1947.
3. Hallikainen, H., Transmitter control systems, in *NAB Engineering Handbook*, 9th edn., J. C. Whitaker (ed.), National Association of Broadcasters, Washington, DC, 1998.

Bibliography

- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, 1992.
- Golio, M., *RF and Microwave Applications and Systems*, CRC Press, Boca Raton, FL, 2008.
- Gray, T. S., *Applied Electronics*, Massachusetts Institute of Technology, Cambridge, MA, 1954.
- Heising, R. A., Modulation in radio telephony, *Proceedings of the IRE*, 9(3), 305, 1921.
- Heising, R. A., Transmission system, U.S. Patent no. 1,655,543, January 1928.
- Jordan, E. C. (ed.), *Reference Data for Engineers: Radio, Electronics, Computers, and Communications*, 7th edn., Howard W. Sams, Indianapolis, IN, 1985.
- Martin, T. L., Jr., *Electronic Circuits*, Prentice Hall, Englewood Cliffs, NJ, 1955.
- Mendenhall, G. N., M. B. Shrestha, and E. J. Anthony, FM broadcast transmitters, in *NAB Engineering Handbook*, 8th edn., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, pp. 429–474, 1992.
- Shrestha, M. B., Design of tube power amplifiers for optimum FM transmitter performance, *IEEE Transactions on Broadcasting*, 36(1), 46–64, 1990.
- Terman, F. E. and J. R. Woodyard, A high efficiency grid-modulated amplifier, *Proceedings of the IRE*, 26(8), 929, 1938.
- Weirather, R. R., G. L. Best, and R. Plonka, TV transmitters, in *NAB Engineering Handbook*, 8th edn., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, pp. 499–551, 1992.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1991.
- Whitaker, J. C., *AC Power Systems*, 2nd edn., CRC Press, Boca Raton, FL, 1998.
- Woodard, G. W., AM transmitters, in *NAB Engineering Handbook*, 8th edn., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, pp. 353–381, 1992.

5

Applying Vacuum Tube Devices

5.1 Introduction

Numerous circuits have been developed to meet the varied needs of high-power radio frequency (RF) amplification. Designers have adapted standard vacuum tube devices to specific applications that require unique performance criteria. Most of these circuits are built upon a few basic concepts.

5.2 AM Power Amplification Systems

Class C plate modulation is the classic method of producing an amplitude modulation (AM) waveform [1]. Figure 5.1 shows the basic circuit. Triodes and tetrodes may be used as the modulator or carrier tube. Triodes offer the simplest and most common configuration.

Numerous variations on the basic design exist, including a combination of plate and screen grid modulation. The carrier signal is applied to the control grid, and the modulating signal is applied to the screen and plate. The plate is fully modulated, and the screen is modulated 70%–100% to achieve 100% carrier modulation. Modulation of the screen can be accomplished using one of the following methods:

- Screen voltage supplied through a dropping resistor connected to the unmodulated dc plate supply, shown in Figure 5.2a
- An additional (third) winding on the modulation transformer, illustrated in Figure 5.2b
- A modulation choke placed in series with a low-voltage fixed screen supply, shown in Figure 5.2c

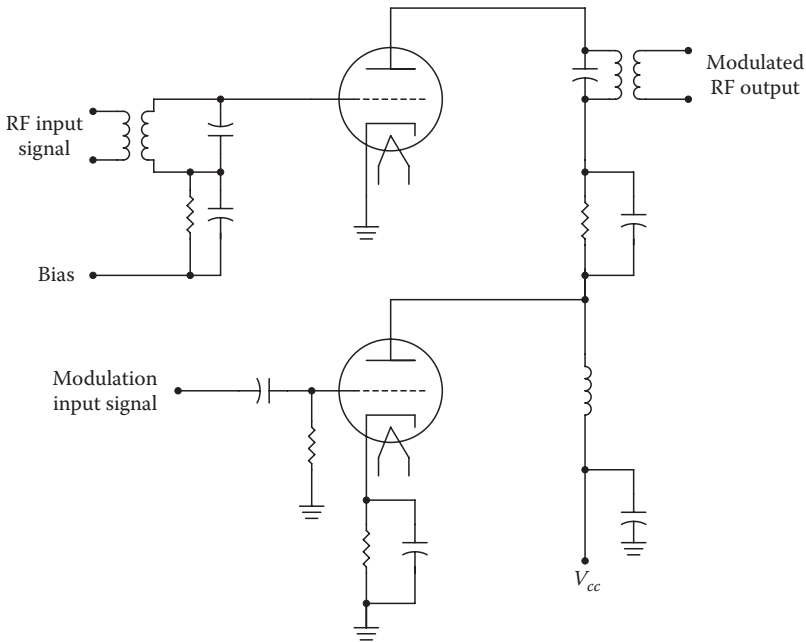


FIGURE 5.1

Classic plate-modulated PA circuit for AM applications. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

Depending on the design, screen/plate modulation also may require partial modulation of the control grid to achieve the desired performance characteristics.

During the portion of the modulation cycle when the plate voltage is increased, the screen current decreases. If the screen is supplied through an impedance, such as the screen dropping resistance of a modulation choke, the voltage drop in this series impedance becomes less and the screen voltage rises in the desired manner. During the part of the modulation cycle when the plate voltage is decreased, the screen current increases, causing a greater voltage drop in the screen series impedance, and lowering the voltage on the screen of the tube. The value of the screen bypass capacitor in the class C stage is a compromise between good RF bypassing and the shunting effect of this capacitance on the screen modulation circuit.

5.2.1 Control Grid Modulation

Although not as common as class C plate modulation, a class C RF amplifier may be modulated by varying the voltage on the control grid of a triode [1]. This approach, shown in Figure 5.3, produces a change in the magnitude of the plate current pulses and, therefore, variations in the output waveform taken from the plate tank. Both the carrier and modulation signals are applied to the grid.

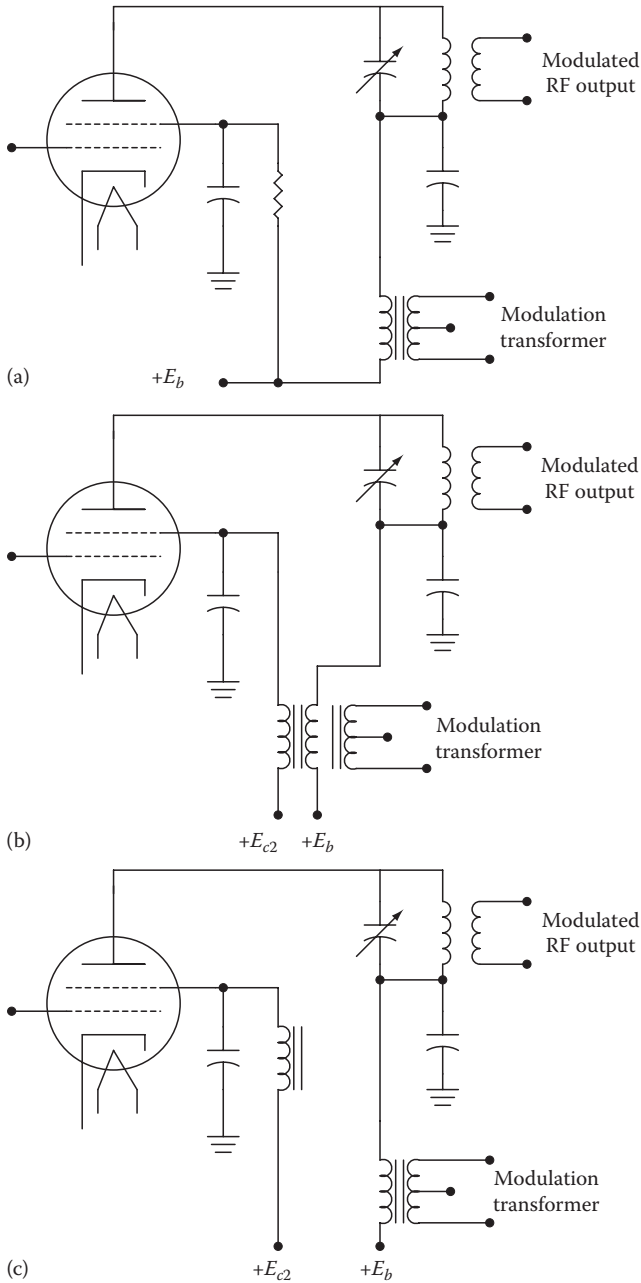


FIGURE 5.2 Basic methods of screen and plate modulation: (a) plate modulation, (b) plate and screen modulation, and (c) plate modulation using separate plate/screen supplies. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

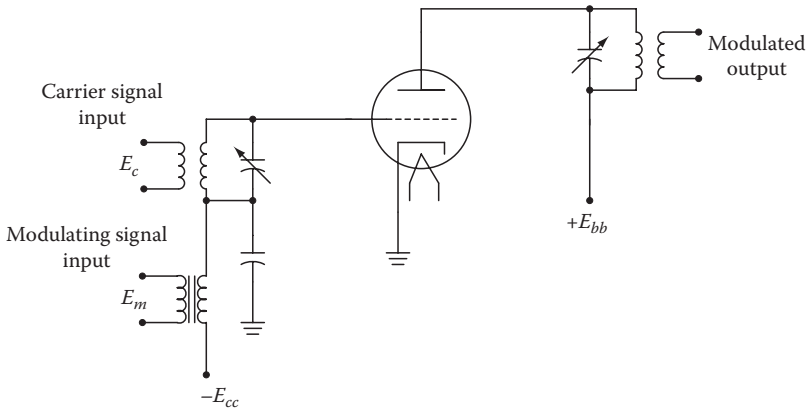


FIGURE 5.3

Grid modulation of a class C RF amplifier. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

The primary benefit of grid modulation is that a modulator is not required. Grid modulation, however, requires a fixed plate supply voltage that is twice the peak RF voltage without modulation. The result is higher plate dissipation at lower modulation levels, including carrier level. The typical plate efficiency of a grid-modulated stage may range from only 35% to 45% at carrier. Grid modulation typically is used only in systems where plate modulation transformers cannot provide adequate bandwidth for the intended application.

When grid modulation is used, the screen voltage and grid bias must be taken from sources with good regulation. This usually means separate low-voltage power supplies.

5.2.2 Suppressor Grid Modulation

The output of a class C pentode amplifier may be controlled by applying a modulating voltage, superimposed on a suitable bias, to the suppressor grid to produce an AM waveform [1]. As the suppressor grid becomes more negative, the minimum instantaneous potential at which current can be drawn to the plate is increased. Thus, as modulation varies the suppressor grid potential, the output voltage changes.

This method of modulating an RF amplifier provides about the same plate efficiency as a grid-modulated stage. The overall operating efficiency, however, is slightly lower because of increased screen grid losses associated with the design. Linearity of the circuit is not usually high.

5.2.3 Cathode Modulation

Cathode modulation incorporates the principles of both control grid and plate modulation [1]. As shown in Figure 5.4, a modulation transformer in the

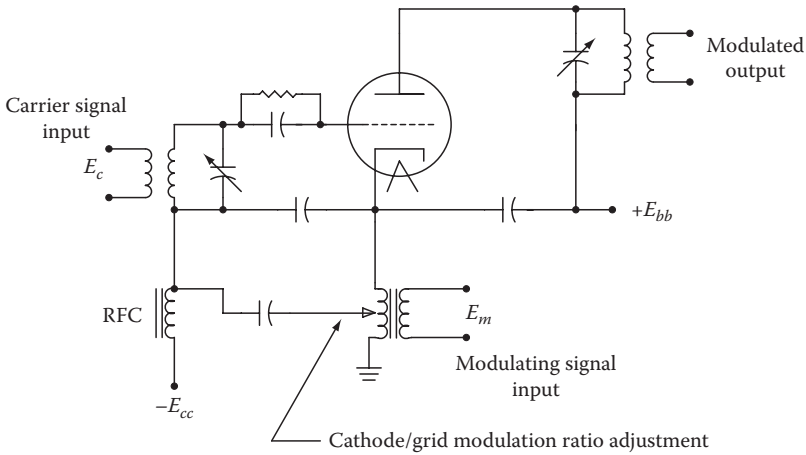


FIGURE 5.4 Cathode modulation of a class C RF amplifier. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

cathode circuit varies the grid-cathode potential, as well as the plate-cathode potential. The ratio of grid modulation to plate modulation is set by adjustment of the tap shown in the figure. Grid leak bias typically is employed to improve linearity of the stage.

5.2.4 High-Level AM Amplification

The basic tuned-anode vacuum tube power amplifier is described in graphical form in Figure 5.5 [2]. The tube may be a triode, tetrode, or pentode. Tetrode final amplifiers are the most common. The RF excitation voltage is applied to the grid of the PA tube, and the ratio of dc grid bias voltage to peak RF excitation voltage (shown sinusoidal in Figure 5.5) determines the conduction angle θ_c of anode current:

$$\theta_c = 2 \arccos \left\{ \frac{E_{cc}}{E_g - E_{cc}} \right\} \tag{5.1}$$

where

- E_g is the exciting grid signal
- E_{cc} is the dc grid bias voltage

The shape of the anode current pulse is determined by the vacuum tube transfer characteristics and the input waveshape. The pulse of current thus generated is supplied by the dc power supply E_{bb} and passes through the resonant anode tank circuit. The tank is assumed to have sufficient operating Q to force anode voltage E_p to be essentially sinusoidal and of the same periodic frequency as the RF excitation voltage and resultant anode

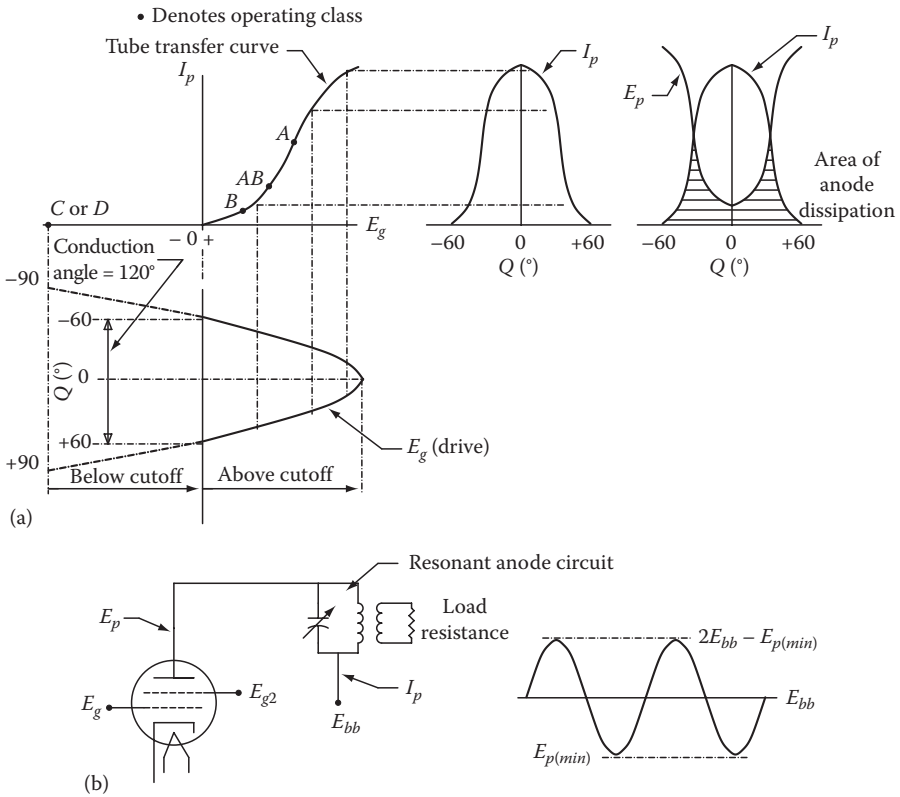


FIGURE 5.5

Classic vacuum tube class C amplifier: (a) representative waveforms and (b) schematic diagram. Conditions are as follows: sinusoidal grid drive, 120° anode current, resonant anode load. (After Woodard, G.W., AM transmitters, in *NAB Engineering Handbook*, 8th edn., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, p. 356, 1992.)

current pulse. The instantaneous anode dissipation is the product of instantaneous tube anode voltage drop and anode current. The tube transfer characteristic is a variable, dependent upon device geometry and other factors, including the maximum drive signal E_g . The exact shape and magnitude of the current waveform is normally obtained from a load-line plot on the constant-current characteristic tube curves supplied by the device manufacturer. The resonant anode load impedance is chosen and adjusted to allow $E_{p(min)}$ to be as low as possible without causing excessive screen grid (tetrode case) or control grid dissipation.

The overall anode efficiency of this circuit can be extended beyond the normal limits for typical class C amplification by employing a third-harmonic resonator between the output anode connection and the fundamental resonant circuit. In some cases, fifth-harmonic resonators are also employed. This technique has the effect of squaring up the anode voltage waveform E_p ,

thus causing the integral of the $E_p \times i_p$ product (anode dissipation) to be smaller, resulting in lower dissipation for a given RF power output. An amplifier employing the third-harmonic anode trap commonly is referred to as a class C–D stage, suggesting efficiency ranging between conventional class C operation and class D operation. The use of the third-harmonic technique permits efficiencies of 85% or more for transmitters of 10 kW power output and above.

5.2.4.1 Class B Modulator

The most common method of applying low-frequency (typically audio) intelligence to a high-level amplitude-modulated amplifier is the class B push–pull system illustrated in Figure 5.6 [1]. The vacuum tubes used in such a circuit may be triodes, tetrodes, or pentodes. The output circuit includes a modulation transformer, audio coupling capacitor, and dc shunt feed inductor. The capacitor and shunt inductor network is used to prevent unbalanced dc from magnetizing the modulation transformer core, which would result in poor low-frequency performance. Advanced core materials and improved transformer design have permitted elimination of the coupling capacitor and the shunt reactor in modern systems. The direct current to the modulated RF amplifier anode, therefore, flows directly through the secondary of the modulation transformer.

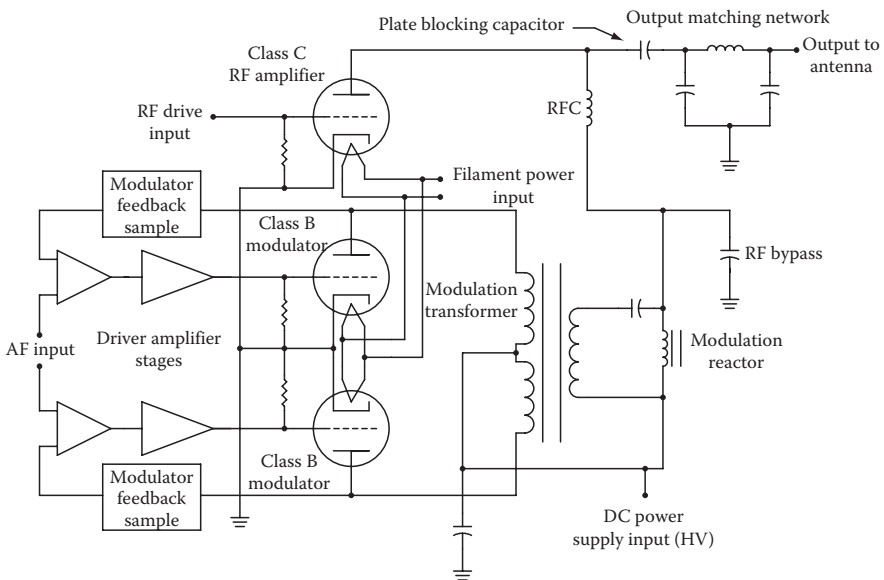


FIGURE 5.6

Class B push–pull modulator stage for a high-level AM amplifier. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

Elimination of the coupling capacitor and shunt reactor results in improved performance from the modulator circuit. The L and C components effectively form a three-pole high-pass filter that causes low-frequency transient distortion to be generated when driven with complex audio waveforms. The simplified circuit constitutes a single-pole high-pass filter, greatly reducing low-frequency transients.

Even in the improved design, the transient distortion performance of the class B push-pull modulator can be insufficient for some applications. Complex modulating waveforms with sharp transitions present special problems. Stray internal winding capacitances and leakage inductances effectively form a multipole low-pass filter at the high-frequency end of the audio spectrum. This filter can produce overshoots when driven with complex waveforms. Such distortion can be reduced by filtering the input signal, at the cost of somewhat lower high-frequency response.

Negative feedback is used to reduce nonlinear distortion and noise in the push-pull circuit. Negative feedback, however, usually increases high-frequency transient distortion.

5.2.5 Pulse Width Modulation

Pulse width modulation (PWM), also known as *pulse duration modulation* (PDM), is a common transmission system for modern vacuum tube transmitters. Figure 5.7 shows the PDM scheme (Harris) applied to an amplitude-modulated PA.

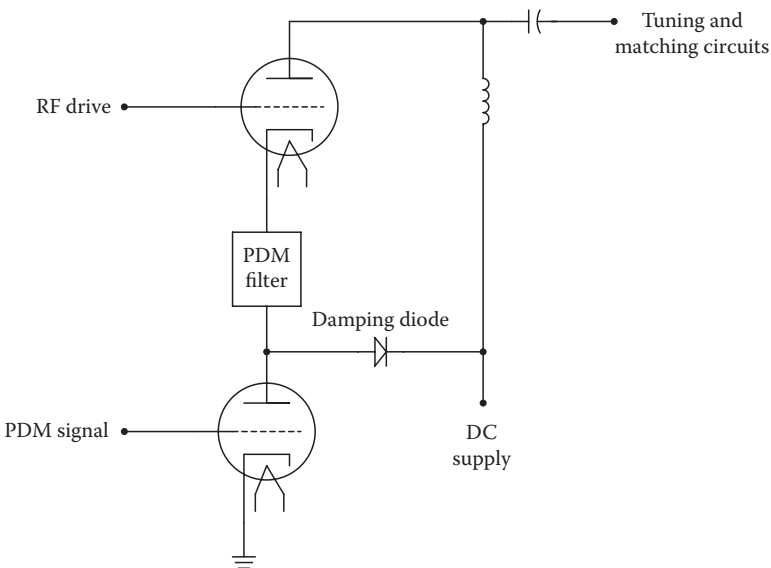


FIGURE 5.7
PDM method of PWM.

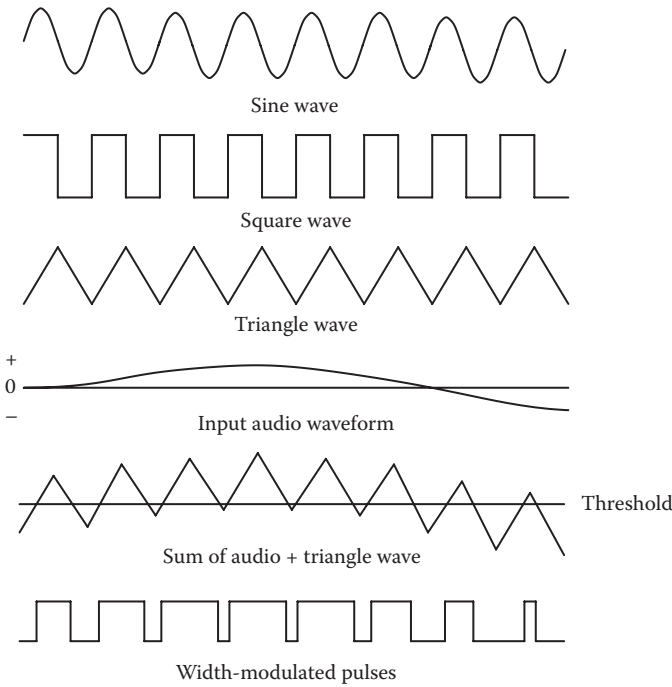


FIGURE 5.8
Principal waveforms of the PDM system.

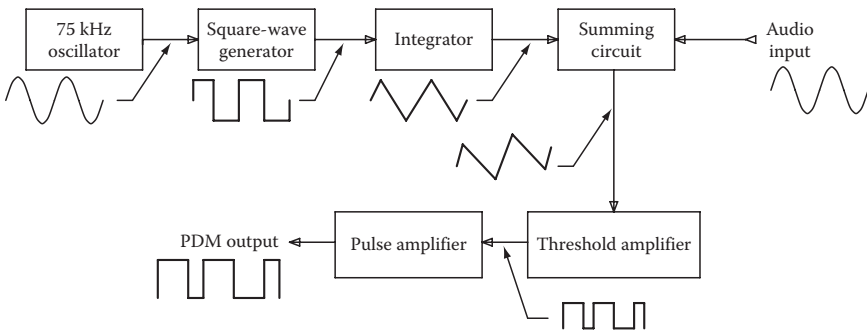


FIGURE 5.9
Block diagram of a PDM waveform generator.

The PDM system works by utilizing a square wave switching system, illustrated in Figure 5.8.

The PDM process begins with a signal generator (Figure 5.9). A 75 kHz sine wave is produced by an oscillator and used to drive a square wave generator, resulting in a simple 75 kHz square wave. The square wave is then integrated, resulting in a triangular waveform that is mixed with the input

audio in a summing circuit. The resulting signal is a triangular waveform that rides on the incoming signal, typically audio. This composite signal is then applied to a threshold amplifier, which functions as a switch that is turned on whenever the value of the input signal exceeds a certain limit. The result is a string of pulses in which the width of the pulse is proportional to the period of time the triangular waveform exceeds the threshold. The pulse output is applied to an amplifier to obtain the necessary power to drive subsequent stages. A filter eliminates whatever transients may exist after the switching process is complete.

The PDM scheme is, in effect, a digital modulation system with the input information being sampled at a 75 kHz rate. The width of the pulses contains all the intelligence. The pulse-width-modulated signal is applied to a *switch* or *modulator tube*. The tube is simply turned *on*, to a fully saturated state, or *off* in accordance with the instantaneous value of the pulse. When the pulse goes positive, the modulator tube is turned on and the voltage across the tube drops to a minimum. When the pulse returns to its minimum value, the modulator tube turns off.

This PDM signal becomes the power supply to the final RF amplifier tube. When the modulator is switched on, the final amplifier will experience current flow and RF will be generated. When the switch or modulator tube goes off, the final amplifier current will cease. This system causes the final amplifier to operate in a highly efficient class D switching mode. A dc offset voltage to the summing amplifier is used to set the carrier (no modulation) level of the transmitter.

A high degree of third-harmonic energy will exist at the output of the final amplifier because of the switching-mode operation. This energy is eliminated by a third-harmonic trap. The result is a stable amplifier that normally operates in excess of 90% efficiency. The power consumed by the modulator and its driver is usually a fraction of a full class B amplifier stage.

The *damping diode* shown in Figure 5.7 is included to prevent potentially damaging transient overvoltages during the switching process. When the switching tube turns off the supply current during a period when the final amplifier is conducting, the high current through the inductors contained in the PDM filters could cause a large transient voltage to be generated. The energy in the PDM filter is returned to the power supply by the damping diode. If no alternative route were established, the energy would return by arcing through the modulator tube itself.

The PWM system makes it possible to eliminate audio frequency transformers in the transmitter. The result is wide frequency response and low distortion. Note that variations on this amplifier and modulation scheme have been used by other manufacturers for standard broadcast, shortwave service, and other applications.

5.3 Linear Amplification

The following features are desirable for vacuum tubes used in RF linear amplifier service [1]:

- High power gain
- Low plate-to-grid capacitance
- Good efficiency
- Linear characteristics that are maintained without degradation across the desired operating band

For linear service, RF amplifiers may be operated in class A, AB₁, AB₂, or B modes. Triode, tetrode, or pentode devices may be used, either grid or cathode driven. The choice of mode, tube, and driving method depends upon the operational specifications of each application.

5.3.1 Device Selection

The triode tube, having a large plate-to-grid interelectrode capacitance, requires neutralization in grid-driven linear service to prevent oscillation [1]. A triode having a low amplification factor is suitable for class AB and AB₂ grid-driven operation. The RF grid excitation voltage for this type of service will be high; grid excursions into the positive region are normal for class AB₂. A swamping resistor is typically used across the input tuned circuit to maintain a constant input impedance to the stage and to provide for stability. With a low value of swamping resistance, the grid current drawn is only a small part of the total grid load, and the driver load impedance is relatively constant. The swamping resistor improves RF stability by providing a low impedance to ground for regenerative feedback through the plate-to-grid capacitance.

High-amplification-factor triodes perform exceptionally well in circuits where the grid is grounded and the cathode is driven. Under these conditions, the control grid acts as a shield between the input and output circuits. Neutralization, therefore, is not normally required. Zero-bias triodes operate in the class AB₂ mode and require only filament, plate, and drive power. For optimum linear operation, a tuned circuit is placed in the cathode RF return path to maintain a sinusoidal waveshape over the drive cycle. The tuned circuit will reduce the intermodulation distortion produced by the amplifier and also will reduce drive power requirements.

If the driver and PA stages are located close to each other, the tuned cathode circuit can be a part of the output circuit of the driver. If, however, the amplifiers are far removed and coupled by a length of coaxial

cable, it is recommended that a tuned cathode circuit with a Q of between 2 and 4 be used.

Most tetrode and pentode amplifiers are designed to be grid driven to take advantage of the high power gain of these devices.

In all linear amplifier systems, the driver output impedance should be kept low because of the nonlinear input loading characteristics of the amplifier tube as it approaches maximum power. The lower the driver amplifier impedance, the smaller the effect of the nonlinear input loading.

5.3.2 Grid-Driven Linear Amplifier

A linear amplifier utilizing a tetrode or pentode is usually grid driven to take advantage of the inherent high gain of the tube [1]. Such an amplifier can be driven into the grid-current region under the proper circumstances. In any case, the input circuit will be loaded by the tube grid. For the no-grid-current case, the driver will work into the input conductance loading; for the grid-current case, the driver will work into the input conductance loading plus grid-current loading. It is, therefore, desirable (and necessary in the grid-current case) to swamp the input circuit with an appropriate noninductive resistor. The resistor will maintain an almost constant load on the driver and minimize the effects of any nonlinearity in grid loading.

5.3.3 Cathode-Driven Linear Amplifier

The cathode-driven amplifier may use triode, tetrode, or pentode tubes [1]. The drive signal is applied to the cathode in this class of operation. The cathode-driven amplifier is particularly suitable for high-power stages using high- μ triodes in the HF and VHF region. This class of operation normally eliminates the need for neutralization because the control grid screens the plate from the input circuit. The power gain for suitable triode class AB cathode-driven amplifiers is on the order of 7–20. The actual tube power gain is approximately equal to the ratio of RF plate voltage to RF cathode voltage. This relationship is true because the fundamental component of the plate current is common to the input and output circuits.

Tetrode tubes also can be used in cathode-driven operation. Power gain is considerably higher than that of the triode, on the order of 20–50. It is important to note that screen grid current loads the input circuit, just as control grid current does.

For an amplifier located some distance from the driver, an improvement in intermodulation distortion can be realized by tuning the cathode circuit. When the driver is located close to the amplifier (0.1 wavelength or so) other means may be used to minimize the nonlinear loading of the cathode-driven stage.

5.3.4 Intermodulation Distortion

When an RF signal with varying amplitude is passed through a nonlinear device, many new products are produced [1]. The frequency and amplitude of each component can be determined mathematically because the nonlinear device can be represented by a power series expanded about the zero-signal operating point. The example of a typical two-tone signal serves to summarize this mathematical relationship. Assume that two equal-amplitude test signals ($f_1 = 2.001$ MHz, and $f_2 = 2.003$ MHz) are applied to a linear amplifier. The output spectrum of the amplifier is shown in Figure 5.10. Many of the distortion products fall outside the passband of the amplifier tuned circuits. If no impedance exists at the frequencies of the distortion components, then no voltage can be developed. Further study of this spectrum discloses that no even-order products fall near the two desired signals. Some odd-order products, however, fall near the desired frequencies and possibly within the passband of the tuned circuits.

The distortion products usually given in tube data sheets are the third- and fifth-order intermodulation distortion products that can fall within the amplifier passband. Using the same f_1 and f_2 frequencies of the previous example, the frequencies of the third-order products are

$$2f_1 - f_2 = 1.999 \text{ MHz}$$

$$2f_2 - f_1 = 2.005 \text{ MHz}$$

The frequencies of the fifth-order products are

$$3f_1 - 2f_2 = 1.997 \text{ MHz}$$

$$3f_2 - 2f_1 = 2.007 \text{ MHz}$$

These frequencies are well within the passband of a tuned circuit intended to pass voice frequencies and, therefore, power will be delivered to the antenna at these frequencies. All intermodulation distortion power is wasted and serves no purpose other than to cause interference to adjacent channels.

Intermodulation distortion in a power amplifier tube is the result of variations of the transfer characteristics of the device from the ideal case. An ideal transfer characteristic curve is shown in Figure 5.11. Even-order products do not contribute to the intermodulation distortion problem because they fall outside the amplifier passband. Therefore, if the transfer characteristic produces an even-order curvature at the small-signal end of the curve (from point *A* to point *B*) and the remaining portion of the curve (point *B* to point *C*) is linear, the tube is considered to have an ideal transfer characteristic. If the operating point of the amplifier is set at point *O* (located midway horizontally between point *A* and point *B*), there will be no distortion in a class AB amplifier. However, no tube has this idealized transfer characteristic. It is possible, by manipulation of the electron ballistics within a given tube structure, to alter the transfer characteristic and minimize the distortion products.

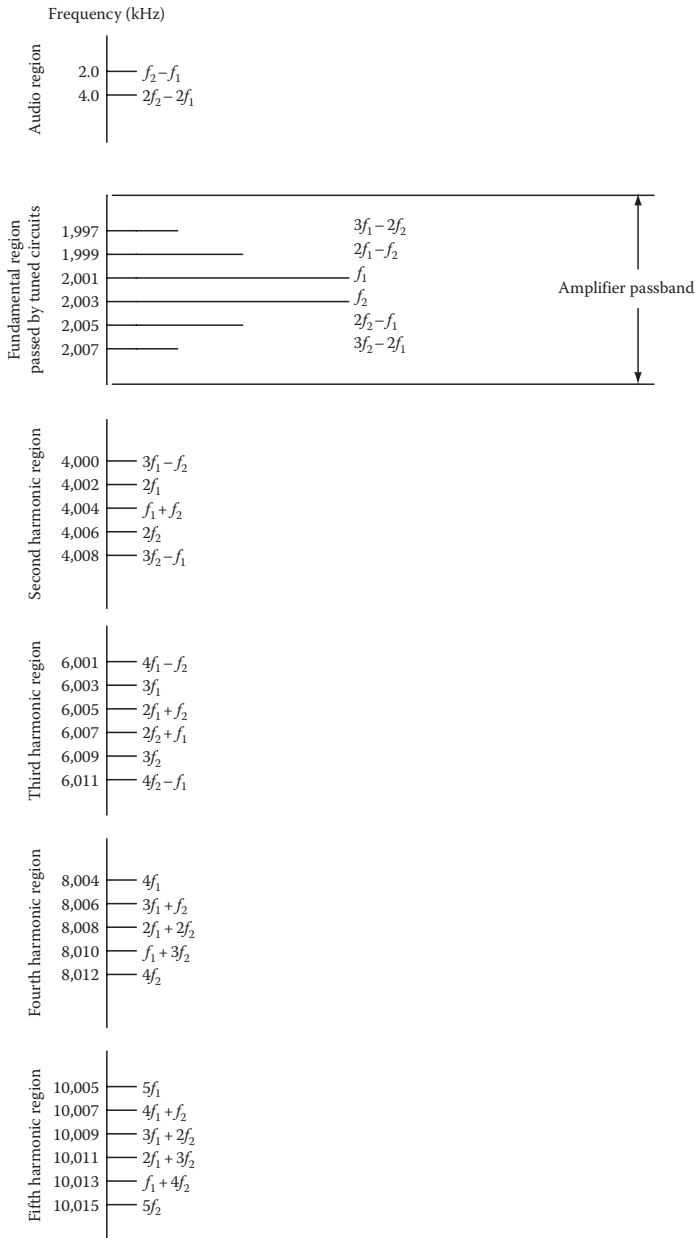


FIGURE 5.10

Spectrum at the output of a nonlinear device with an input of two equal-amplitude sine waves of $f_1 = 2.001$ MHz and $f_2 = 2.003$ MHz. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.)

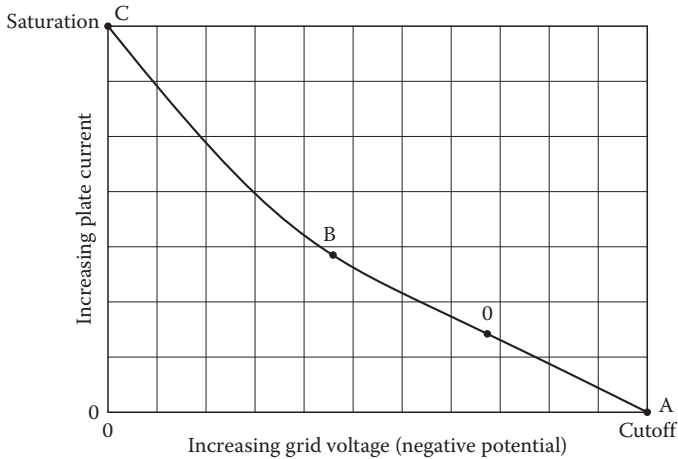


FIGURE 5.11
Ideal grid-plate transfer curve for class AB operation.

5.4 High-Efficiency Linear Amplification

The class B RF amplifier is the classic means of achieving linear amplification of an AM signal. A significant efficiency penalty, however, is paid for linear operation. The plate efficiency at carrier level in a practical linear amplifier is about 33%, while a high-level (plate)-modulated stage can deliver a basic efficiency of 65%–80%. In a real-world transmitter, power consumed by the modulator stage must also be taken into account. However, because the output power of the modulator is, at most, one-third of the total power output, the net efficiency of the plate-modulated amplifier is still higher than a conventional linear amplifier of comparable power.

Still, linear amplification is attractive to the users of high-power transmitters because low-level modulation can be employed. This eliminates the need for modulation transformers and reactors, which—at certain power levels and for particular applications—may be difficult to design and expensive to construct. To overcome the efficiency penalty of linear amplification, special systems have been designed that take advantage of the benefits of a linear mode, while not imposing an excessive efficiency burden.

5.4.1 Chireix Outphasing Modulated Amplifier

The *Chireix* amplifier employs two output stage tubes. The grids of the tubes are driven with signals whose relative phase varies with the applied modulating signal. The output waveforms of the tubes are then applied to a summing network, and finally to the load. Figure 5.12 shows a simplified schematic diagram.

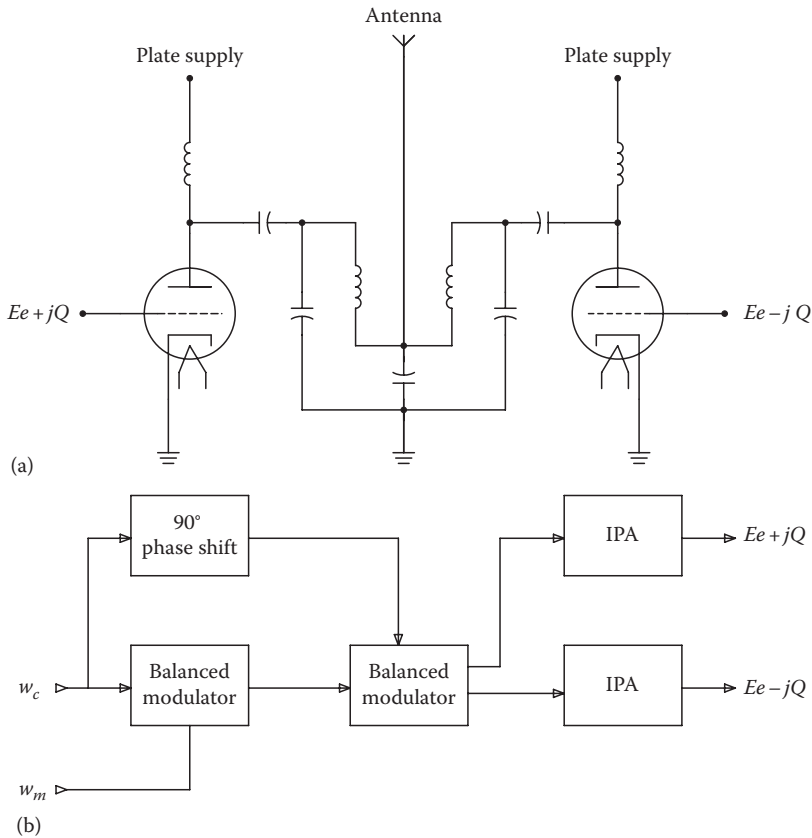


FIGURE 5.12

Simplified schematic diagram of the Chireix outphasing modulated amplifier: (a) output circuit and (b) drive signal generation circuit.

Operation of the circuit is straightforward. When the grids of the tubes are driven with signals that are 180° out of phase, power output will be zero. This corresponds to a negative AM peak of 100%. Full (100%) positive modulation occurs when both grids are driven with signals that are in phase. The AM signal is produced by varying the relative phase of the drive signals.

The phase-modulated carrier-frequency grid waveforms are generated in a low-level balanced modulator and amplified to the level required by the PA tubes. Stability of the exciter in the Chireix system is critical. Any shift in the relative phase of the exciter outputs will translate to AM of the transmitted waveform.

Because of the outphasing method employed to produce the AM signal, the power factor associated with the output stage is of special significance. The basic modulating scheme would produce a unity power factor at zero output. The power factor would then decrease with increasing RF output. To avoid this undesirable characteristic, the output plate circuits are detuned, one

above resonance and one below, to produce a specified offset. This *dephasing*, coupled with adjustment of the relative phase of the driving signals, maintains the power factor within a reasonable range.

There are, however, two major disadvantages to this system. First, the efficiency of the amplifier is compromised because at all instantaneous levels of modulation except one, the anode circuits must work into a reactive load. Typical overall efficiency is on the order of 60%. Second, output carrier power adjustment is sensitive to tuning of any stage in the chain. This problem can be reduced by employing broadband amplifying circuits at lower-level stages.

5.4.2 Doherty Amplifier

The *Doherty*-modulated amplifier has seen considerable use in high-efficiency linear AM transmitters. The system is a true high-efficiency linear amplifier, rather than a hybrid amplifier/modulator. The Doherty amplifier employs two tubes and a 90° network as an *impedance inverter* to achieve load-line modification as a function of power level. The basic circuit is shown in Figure 5.13.

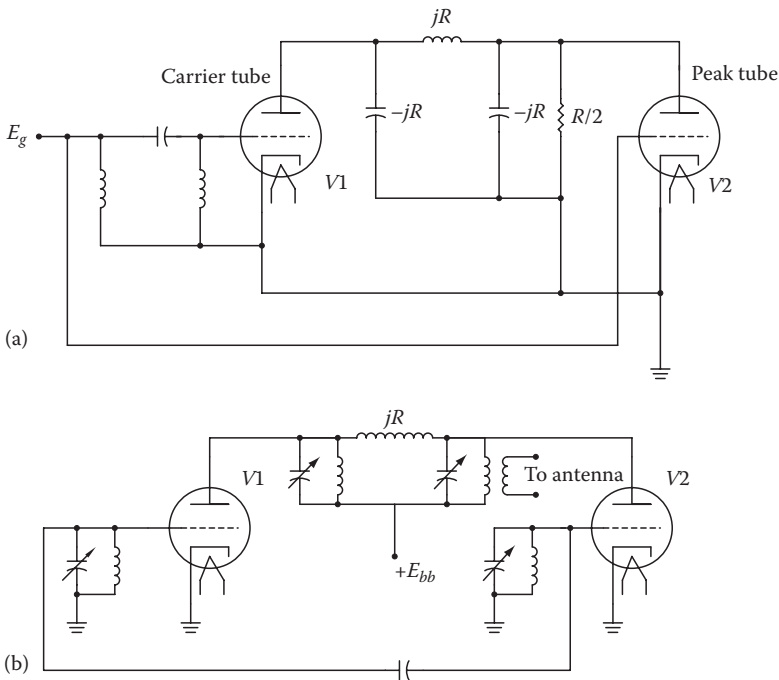


FIGURE 5.13 Doherty high-efficiency AM amplifier: (a) operating theory, and (b) schematic diagram of a typical system.

Two tubes are used in the Doherty amplifier, a *carrier tube* ($V1$) and a *peak tube* ($V2$). The carrier tube is biased class B. Loading and drive for the tube are set to provide maximum linear voltage swing at carrier level. The peak tube is biased class C; at carrier level it just begins to conduct plate current. Each tube delivers an output power equal to twice the carrier power when working into a load impedance of $R \Omega$. At carrier, the reflected load impedance at the plate of $V1$ is equal to $2R \Omega$. This is the correct value of load impedance for carrier-level output at full plate voltage swing. (The impedance-inverting property of the 90° network is similar to a $1/4$ -wave transmission line.)

A 90° phase lead network is included in the grid circuit of $V1$ to compensate for the 90° phase lag produced by the impedance-inverting network in the plate circuit. For negative modulation swings, the carrier tube performs as a linear amplifier with a load impedance of $2R \Omega$; the peak tube contributes no power to the output.

On positive modulation swings, the peak tube conducts and contributes power to the $R/2$ load resistance. This is equivalent to connecting a negative resistance in shunt with the $R/2$ load resistance, so that the value seen at the load end of the 90° network increases. This increase is reflected through the network as a decrease in load resistance at the plate of $V1$, causing an increase in output current from the tube, and an increase in output power. The drive levels on the tubes are adjusted so that each contributes the same power (equal to twice the carrier power) at a positive modulation peak. For this condition, a load of $R \Omega$ is presented to each tube.

The key aspect of the Doherty circuit is the change in load impedance on tube $V1$ with modulation. This property enables the device to deliver increased output power at a constant plate voltage swing. The result is high efficiency and good linearity. Overall efficiency of the Doherty amplifier ranges from 60% to 65%.

A practical application of the Doherty circuit is shown in Figure 5.13b. The shunt reactances of the phase-shift networks are achieved by detuning the related tuned circuits. The tuned circuits in the grid are tuned above the operating frequency; those in the plate are tuned below the operating frequency.

The Doherty amplifier has two important advantages in high-power applications. First, the peak anode voltage at either tube is (approximately) only one-half that required for an equivalent carrier power high-level PWM or class B anode-modulated transmitter, thus significantly increasing reliability and usable tube life. Second, no large modulation transformer or special filtering components are used in the final amplifier stages that might cause transient overshoot distortion.

Disadvantages of the Doherty amplifier include nonlinear distortion performance and increased complexity of tuning. The major sources of nonlinear distortion include the nonlinearity of the carrier tube at or near the 100% negative modulation crest, and the nonlinearity of the peak tube at or near carrier level when it is just beginning to conduct. Both sources of distortion may be effectively reduced by using moderate amounts of overall envelope feedback.

5.4.3 Screen-Modulated Doherty-Type Amplifier

A variation on the basic Doherty scheme can be found in the screen-modulated power amplifier. The design is unique in that the system does not have to function as a linear amplifier. Instead, modulation is applied to the screen grids of both the carrier and peak tubes. The peak tube is modulated upward during the positive half of the modulation cycle, and the carrier tube is modulated downward during the negative half of the cycle. This technique results in performance improvements over the classic Doherty design in that RF excitation voltages and modulating voltages are isolated from each other, thereby eliminating a troublesome source of tuning vs. modulation interaction.

The basic screen-modulation system is shown in Figure 5.14. The peak and carrier tubes are biased and driven in quadrature as class C amplifiers from the continuous wave RF drive source. At carrier level, the screen voltage of the carrier tube is adjusted so that the carrier device is near anode saturation and delivering approximately 96% of the carrier power. The screen voltage of the peak tube is adjusted so that it is just into conduction, and is supplying the remaining 4% (approximately) of carrier power. The combined anode efficiency at carrier level is greater than 77%.

Modulation of the carrier occurs when the screen voltage of the peak tube begins to rise during the positive modulation half-cycle, thus causing the peak tube to supply more RF current to the output load. This increase of current into the output network causes the resistance seen by the interanode network to increase and, because of the impedance-inverting characteristics of the 90° interanode network, causes a proportional decrease in the load impedance presented to the carrier tube.

The carrier-tube resonant anode voltage drop is fully saturated over the entire positive modulation half-cycle, and therefore is effectively a constant-voltage source. The power output of the carrier tube thus increases during the positive modulation half-cycle, caused by the modulated decreasing impedance at its anode, until both peak and carrier tubes deliver twice carrier power at the 100% positive modulation crest. During the negative

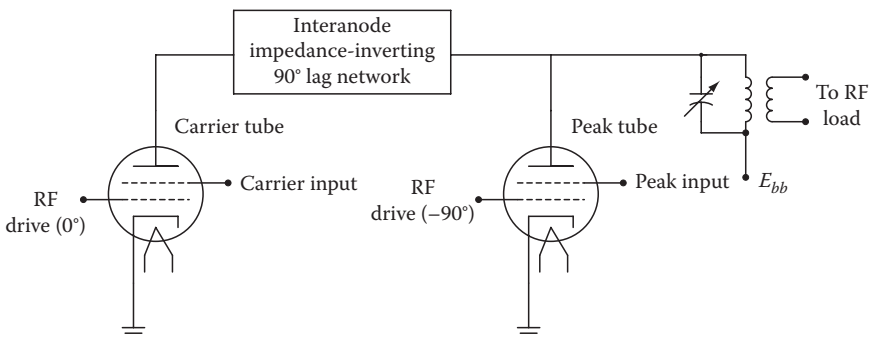


FIGURE 5.14
Screen-modulated Doherty-type circuit.

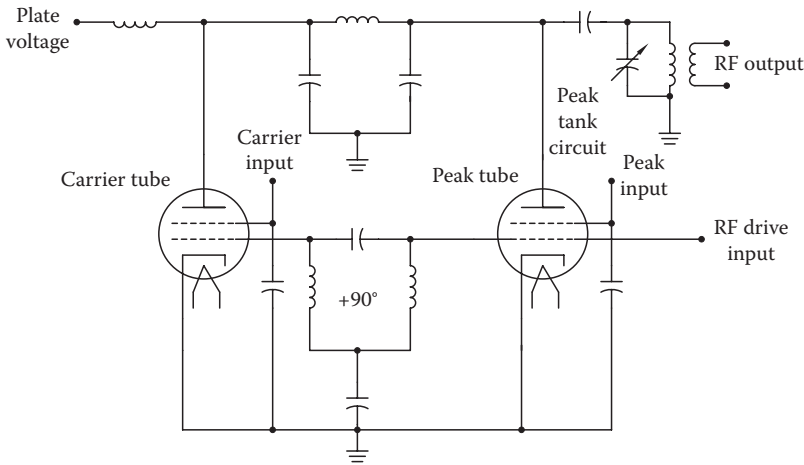


FIGURE 5.15

Practical implementation of a screen-modulated Doherty-type 1MW power amplifier. (Courtesy of Continental Electronics, Dallas, TX.)

modulation half-cycle, the peak tube is held out of conduction, while the carrier-tube output voltage decreases linearly to zero output at the 100% modulation trough.

A screen-modulated amplifier designed for MF operation at 1 MW is shown in Figure 5.15. The carrier tube delivers all of the 1 MW carrier power to the load through the 90° impedance-inverting network. When modulation is applied to the screen of the carrier tube (during the negative half-cycle), the carrier tube output follows the modulation linearly. Because the carrier tube is driven to full swing, its voltage excursion does not increase during the positive half-cycle. Positive modulation is needed on the screen only to maintain the full plate swing as the impedance load changes on the carrier tube.

The peak tube normally is operated with a negative screen voltage that maintains plate current at near cutoff. During the positive peak of modulation, the screen of the peak tube is modulated upward. As the peak tube is modulated positive, it delivers power into the output circuit. As the delivered power increases, the load seen by the interplate network changes as described for the general case.

The advantages of screen modulation are the same as those identified for the Doherty linear amplifier, plus higher efficiency at all depths of modulation. Furthermore, the screen modulation system is less sensitive to misadjustment of RF amplifier tuning.

5.4.4 Terman–Woodyard Modulated Amplifier

The *Terman–Woodyard* configuration uses the basic scheme of Doherty for achieving high efficiency (the impedance-inverting property of a 90°

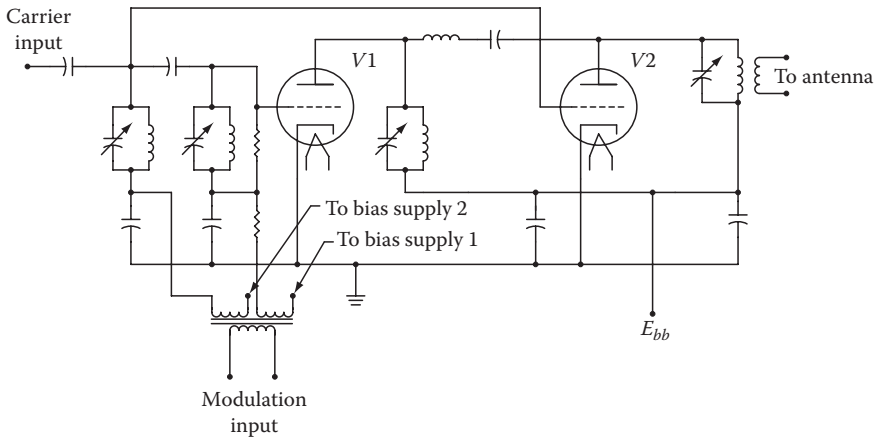


FIGURE 5.16
Terman–Woodyard high-efficiency modulated amplifier.

phase-shift network). However, the Terman–Woodyard design also employs grid modulation of both the carrier tube and the peak tube, allowing both tubes to be operated class C. The result is an increase in efficiency over the Doherty configuration.

The basic circuit is shown in Figure 5.16. With no modulation, $V1$ operates as a class C amplifier, supplying the carrier power, and $V2$ is biased so that it is just beginning to conduct. The efficiency at carrier is, therefore, essentially that of a class C amplifier.

During positive modulation swings, $V2$ conducts. At 100% modulation, both tubes supply equal amounts of power to the load, as in the Doherty design. During negative modulation swings, $V2$ is cut off, and $V1$ performs as a standard grid-modulated amplifier.

The Terman–Woodyard system rates efficiency as a function of modulation percentage for sinusoidal waveforms. Typical efficiency at carrier is 80%, falling to a minimum of 68% at 50% modulation and 73% at 100% modulation.

5.4.5 Dome Modulated Amplifier

The *Dome* high-efficiency amplifier employs three power output tubes driven by different input signals. Modulation is achieved by load-line modification during positive modulation excursions, and by linear amplification during negative modulation swings. Load-line modification is achieved by absorption of a portion of the generated RF power. However, most of the absorbed power is returned to the plate power supply, rather than being dissipated as heat. High efficiency is the result.

A basic Dome modulated amplifier is shown in Figure 5.17. Tube $V1$ is used in a plate-modulated configuration, supplying power to the grid of $V2$ (the power amplifier tube). The load impedance in the plate circuit of $V2$ is

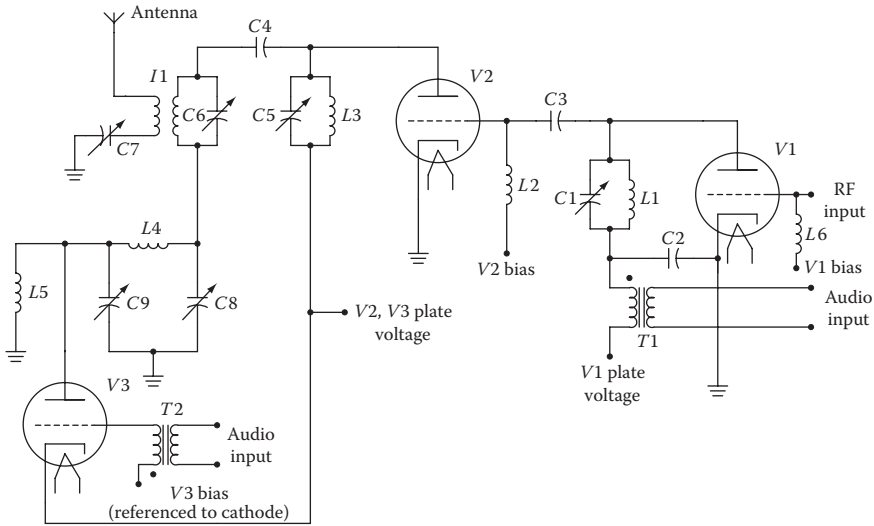


FIGURE 5.17
Dome high-efficiency modulated amplifier.

equivalent to the load impedance reflected into the primary of transformer $T1$ (from the antenna) in series with the impedance appearing across the $C8$ terminals of the 90° phase-shift network consisting of $C8$, $C9$, and $L4$. The impedance appearing across the $C8$ terminals of the 90° network is inversely related to the impedance across the other terminals of the network (in other words, the effective ac impedance of $V3$). Thus, with $V3$ cut off, a short circuit is reflected across the $C8$ terminals of the network. Tube $V3$ performs as a modulated rectifier with intelligence supplied to its grid. This tube is referred to as the *modifier*.

With no modulation (the carrier condition), tube $V3$ is biased to cutoff, and the drive to $V2$ is adjusted until power output is equal to four times the carrier power (corresponding to a positive modulation peak). Note that all the plate signal voltage of $V2$ appears across the primary of transformer $T1$ for this condition. The bias on $V3$ is then reduced, lowering the ac impedance of the tube and, therefore, reflecting an increasing impedance at the $C8$ terminals of the 90° network. The bias on $V3$ is reduced until the operating carrier power is delivered to the antenna.

An amount of power equal to the carrier is thus rectified by $V3$ and returned to the plate power supply, except for that portion dissipated on the $V3$ plate. The drive level to $V2$ is finally adjusted so that the tube is just out of saturation.

For positive modulation excursions, $V3$ grid voltage is driven negative, reaching cutoff for a positive modulation peak. For negative modulation swings, $V2$ acts as a linear amplifier. $V3$ does not conduct during the negative modulation swing because the peak RF voltage on the plate is less than the dc supply voltage on the cathode.

Because of the serial loss incurred in the power amplifier tube and in the modifier tube for energy returned to the power supply, the Dome circuit is not as efficient as the Doherty configuration. Typical efficiency at carrier ranges from 55% to 60%.

5.5 Television Power Amplifier Systems

Although VSBAM is used for a number of varied applications, TV transmission is the most common. An analog TV transmitter is divided into two basic subsystems:

- The *visual* section, which accepts the video input, amplitude-modulates an RF carrier, and amplifies the signal to feed the antenna system.
- The *aural* section, which accepts the audio input, frequency-modulates a separate RF carrier, and amplifies the signal to feed the antenna system.

The visual and aural signals usually are combined to feed a single radiating antenna. Different transmitter manufacturers have different philosophies with regard to the design and construction of a transmitter. Some generalizations are possible, however, with respect to basic system configurations. Transmitters can be divided into categories based on the following criteria:

- Output power
- Final-stage design
- Modulation system

5.5.1 System Considerations

The amount of output power required of the transmitter will have a fundamental effect on system design. Power levels dictate the following parameters:

- Whether the unit will be of solid-state or vacuum tube design
- Whether air, water, or vapor cooling must be used
- The type of power supply required
- The sophistication of the high-voltage control and supervisory circuitry
- Whether *common amplification* of the visual and aural signals (rather than separate visual and aural amplifiers) is practical

As solid-state technology advances, the power levels possible in a reasonable transmitter design steadily increase, making solid-state systems more attractive options.

In the realm of UHF transmitters, the klystron (and its related devices) reigns supreme. Klystrons use an *electron-bunching* technique to generate high power—55kW from a single tube is not uncommon—at ultrahigh frequencies. They are currently the first choice for high-power, high-frequency service. Klystrons, however, are not particularly efficient. A stock klystron with no special circuitry might be only 40% efficient. Various schemes have been devised to improve klystron efficiency, the best known of which is *beam pulsing*. Two types of pulsing are in common use:

- *Mod-anode pulsing*, a technique designed to reduce power consumption of the device in analog television service during the color burst and video portion of the signal (and thereby improve overall system efficiency)
- *Annular control electrode (ACE) pulsing*, which accomplishes basically the same goal by incorporating the pulsing signal into a low-voltage stage of the transmitter, rather than a high-voltage stage (as with mod-anode pulsing)

Several additional technologies offer ways to improve UHF transmitter efficiency, including

- The *inductive output tube (IOT)*, also known as the *Klystrode*. This device essentially combines the cathode/grid structure of the tetrode with the drift tube/collector structure of the klystron.
- The *multistage depressed collector (MSDC) klystron*, a device that achieves greater efficiency through a redesign of the collector assembly. A multistage collector is used to recover energy from the electron stream inside the klystron and return it to the beam power supply.

Improved tetrode devices, featuring higher operating power at UHF and greater efficiency, have also been developed.

A number of approaches may be taken to AM of the visual carrier, notably low-level intermediate-frequency (IF) modulation. This approach allows superior distortion correction, more accurate vestigial sideband shaping, and significant economic advantages to the transmitter manufacturer.

An analog TV transmitter can be divided into four major subsystems:

- The exciter
- Intermediate power amplifier (IPA)
- Power amplifier
- High-voltage power supply

Figure 5.18 shows the audio, video, and RF paths for a typical design.

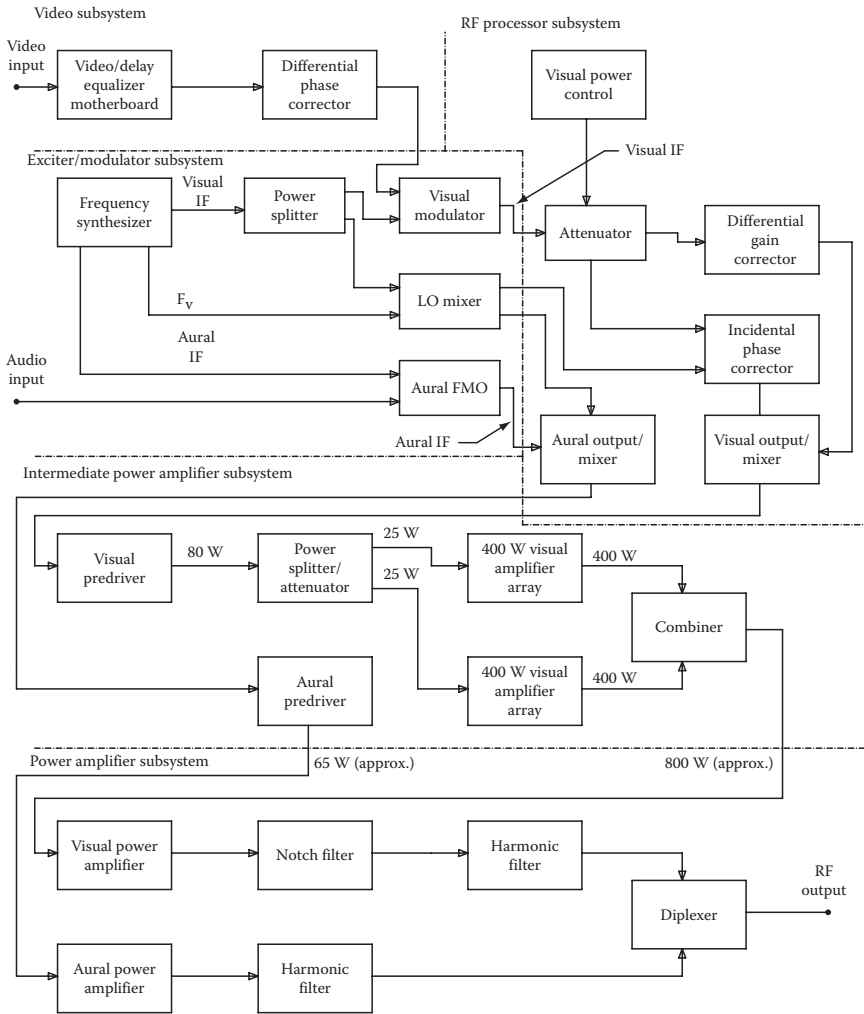


FIGURE 5.18 Basic block diagram of a TV transmitter. The three major subassemblies are the exciter, IPA, and PA. The power supply provides operating voltages to all sections and high voltage to the PA stage.

5.5.2 Power Amplifier

The power amplifier raises the output energy of the transmitter to the required RF operating level. Tetrodes in TV service are usually operated in the class B mode to obtain reasonable efficiency while maintaining a linear transfer characteristic. Class B amplifiers, when operated in tuned circuits, provide linear performance because of the flywheel effect of the resonance circuit. This allows a single tube to be used instead of two in push-pull fashion. The bias point of the linear amplifier must be chosen so that the

transfer characteristic at low modulation levels matches that at higher modulation levels. Even so, some nonlinearity is generated in the final stage, requiring differential gain correction. The plate (anode) circuit of a tetrode PA usually is built around a coaxial resonant cavity, providing a stable and reliable tank.

UHF transmitters using a klystron in the final output stage must operate class A, the most linear but also most inefficient operating mode for a vacuum tube. The basic efficiency of a nonpulsed klystron is approximately 40%. Pulsing, which provides full available beam current only when it is needed (during peak of sync), can improve device efficiency by as much as 25%, depending on the type of pulsing used.

Two basic types of klystrons are presently in service:

- Integral-cavity klystron
- External-cavity klystron

The basic theory of operation is identical for each tube, but the mechanical approach is radically different. In the integral-cavity klystron, the cavities are built into the klystron to form a single unit. In the external-cavity klystron, the cavities are outside the vacuum envelope and bolted around the tube when the klystron is installed in the transmitter.

A number of factors come into play in a discussion of the relative merits of integral-cavity vs. external-cavity designs. Primary considerations include operating efficiency, purchase price, and life expectancy.

The PA stage includes a number of sensors that provide input to supervisory and control circuits. Because of the power levels present in the PA stage, sophisticated fault-detection circuits are required to prevent damage to components in the event of a problem inside or outside the transmitter. An RF sample, obtained from a directional coupler installed at the output of the transmitter, is used to provide automatic power-level control.

The transmitter system discussed in this section assumes separate visual and aural PA stages. This configuration is normally used for high-power analog transmitters. A combined mode also may be used, however, in which the aural and visual signals are added prior to the PA. This approach offers a simplified system, but at the cost of additional *precorrection* of the input video signal.

PA stages often are configured so that the circuitry of the visual and aural amplifiers is identical, providing backup protection in the event of a visual PA failure. The aural PA can then be reconfigured to amplify both the aural and the visual signals, at reduced power.

The aural output stage of a TV transmitter is similar in basic design to an FM broadcast transmitter. Tetrode output devices generally operate class C, providing good efficiency. Klystron-based aural PAs are used in UHF transmitters.

5.5.2.1 Application Example

A 60 kW analog television transmitter is shown in block diagram form in Figure 5.19. A single high-power klystron is used in the visual amplifier, and another is used in the aural amplifier. The tubes are driven from solid-state IPA modules. The transmitter utilizes ACE-type beam control, requiring additional predistortion to compensate for the nonlinearities of the final visual stage. Predistortion is achieved by correction circuitry at an intermediate frequency in the modulator. Both klystrons are driven from the output of a circulator, which ensures a minimum of driver-to-load mismatch problems.

A block diagram of the beam modulator circuit is shown in Figure 5.20. The system receives input signals from the modulator, which synchronizes ACE pulses to the visual tube with the video information. The pulse waveform is developed through a pulse amplifier, rather than a switch. This permits more accurate adjustments of operating conditions of the visual amplifier.

Although the current demand from the beam modulator is low, the bias is near cathode potential, which is at a high voltage relative to ground. The modulator, therefore, must be insulated from the chassis. This is accomplished with optical transmitters and receivers connected via fiber optic cables. The fiber optic lines carry supervisory, gain control, and modulating signals.

The four-cavity external klystrons will tune to any channel in the UHF-TV band. An adjustable *beam perveance* feature enables the effective electrical length of the device to be varied by altering the beam voltage as a function of operating frequency. Electromagnetic focusing is used on both tubes. The cavities, body, and gun areas of the klystrons are air cooled. The collectors are vapor-phase cooled using an external heat exchanger system.

The outputs of the visual and aural klystrons are passed through harmonic filters to an RF combiner before being applied to the antenna system.

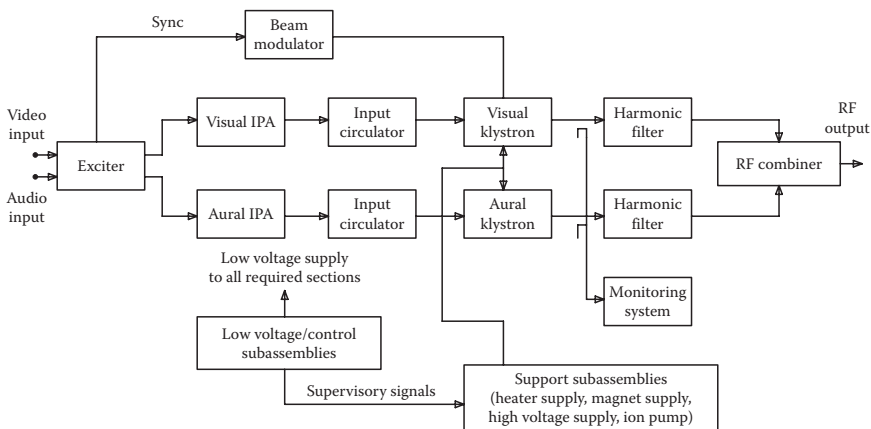


FIGURE 5.19
Schematic diagram of a 60 kW klystron-based analog TV transmitter.

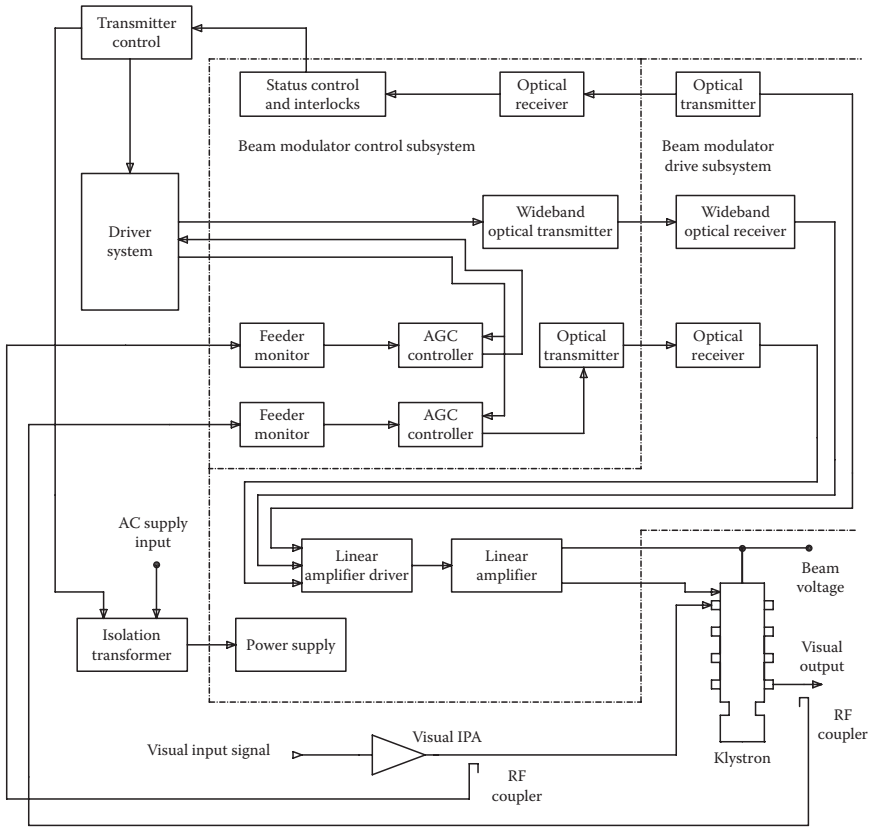


FIGURE 5.20

Schematic diagram of the modulator section of a 60kW analog TV transmitter.

5.6 FM Power Amplifier Systems

The amplitude of a frequency modulation (FM) signal remains constant with modulation so that efficient class B and C amplifiers can be used.* Two basic circuit types have evolved:

- Amplifiers based on a tetrode or pentode in a grid-driven configuration
- Amplifiers based on a high- μ triode or tetrode device in a cathode-driven configuration (grounded grid)

* This section was contributed by Geoffrey N. Mendenhall, P.E., Cincinnati, OH, and Warren B. Bruene, P.E., Dallas, TX.

High-power FM amplifiers operating in the VHF range typically employ resonant cavity systems.

5.6.1 Cathode-Driven Triode Amplifier

The characteristics of high- μ triodes are well adapted to FM power amplifier service. The grounded-grid circuit is simple, and no screen or grid bias power supplies are required. Figure 5.21 shows the basic circuit configuration. In this case, the grid is connected directly to chassis ground. The difference between dc cathode current and dc plate current is dc grid current. The output tank is a shorted coaxial cavity that is capacitively loaded by the tube output and stray circuit capacitance. A small capacitor is used for trimming the tuning, and another small variable capacitor is used for adjusting the loading. A pi network is used to match the 50 Ω input to the tube cathode input impedance.

The triode usually is operated in the class B mode to achieve maximum power gain, which is on the order of 20 (13 dB). The device can be driven into

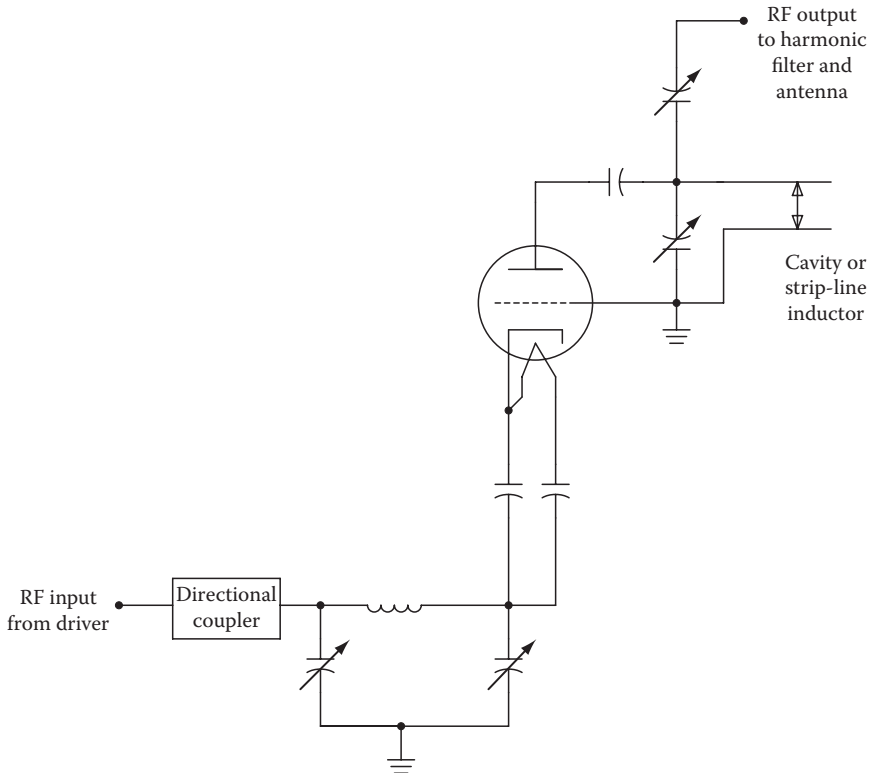


FIGURE 5.21

Cathode-driven triode power amplifier.

class C operation by providing negative grid bias or positive cathode bias. This increases the plate efficiency, but also requires increased drive power.

Most of the drive into a grounded-grid amplifier is fed through the tube and appears at the output. This increases the apparent efficiency so that the efficiency factor given by the transmitter manufacturer may be higher than the actual plate efficiency of the tube. The true plate efficiency is determined from the following:

$$Eff_p = \frac{P_o}{(I_p \times E_p) + (P_{in} \times \alpha)} \quad (5.2)$$

where

Eff_p is the actual plate efficiency

P_o is the measured output power

I_p is the dc plate input current to final stage

E_p is the dc plate input voltage to final stage

P_{in} is the RF input drive power

α is the fraction of grid-to-plate transfer (0.9 typical)

Because most of the drive power is fed through the tube, any changes in loading of the output circuit also will affect the input impedance and the driver stage.

Because there is RF drive voltage on the cathode (filament), some means of decoupling must be used to block the signal from the filament transformer. One method employs high-current RF chokes; another feeds the filament power through the input tank circuit inductor.

5.6.2 Grounded-Grid vs. Grid-Driven Tetrode

Tetrodes also may be operated in the grounded-grid configuration by placing both the control grid and the screen grid at RF ground. Higher efficiency and gain can be achieved by placing negative bias on the control grid while placing a positive voltage on the screen grid of a cathode-driven tetrode.

The input capacitance of a tetrode in a grounded-grid configuration is much less than in a grid-driven configuration, and the input impedance is lower, providing wider bandwidth. The approximate input capacitances of some common tube types are listed in Table 5.1.

The typical drive power requirements of a high-gain tetrode, as a function of plate voltage, for a 5kW power amplifier operating at 100MHz are as follows:

- Grounded-grid configuration: 340W input at 4.5kV plate voltage, 280W at 5.2kV
- Grid-driven configuration: 190W at 4.5kV plate voltage, 140W at 5.2kV

TABLE 5.1

Approximate Input Capacitance of Common Power Grid Tubes Used in Grounded-Grid and Grounded-Cathode Service

Tube Type	Grounded Grid	Grounded Cathode
4CX3,000A	67	140
4CX3,500A	59	111
4CX5,000A	53	115
4CX15,000A	67	161
4CX20,000A/8990	83	190

There are several trade-offs in performance between the grounded-grid and the grid-driven modes of a tetrode PA with respect to gain, efficiency, amplitude bandwidth, phase bandwidth, and incidental AM under equivalent operating conditions:

- When the PA is driven into saturation, its bandwidth is limited by the output cavity bandwidth in the grounded-grid amplifier. The PA bandwidth in the grid-driven amplifier is limited by the input circuit Q .
- Output bandwidth under saturation can be improved in either configuration by reducing the plate voltage. This involves a trade-off in efficiency with a smaller voltage swing. The bandwidth improvement can be obtained with a loss of PA gain and efficiency.
- A grounded-grid saturated PA improves bandwidth over a grid-driven saturated PA at the expense of amplifier gain. Best performance for FM operation is obtained when the amplifier is driven into saturation where little change in output power occurs with increasing drive power. Maximum efficiency also occurs at this point.
- The phase linearity in the 0.5 dB bandwidth of the amplifier is better in a grid-driven configuration. The class C grounded-grid PA exhibits a more nonlinear phase slope within the passband, yet has a wider amplitude bandwidth. This phenomenon is the result of interaction of the input and output circuits because they are effectively connected in series in the grounded-grid configuration. The neutralized grid-driven PA provides more isolation between these networks, so they behave more like independent filters.

5.6.3 Grid-Driven Tetrode/Pentode Amplifiers

RF generators using tetrode amplifiers throughout usually require fewer stages than those using triodes for the same power output. Because tetrodes have higher power gain, they are driven into class C operation for high plate efficiency. Countering these advantages is the requirement for neutralization, along with screen and bias power supplies. Figure 5.22 shows a

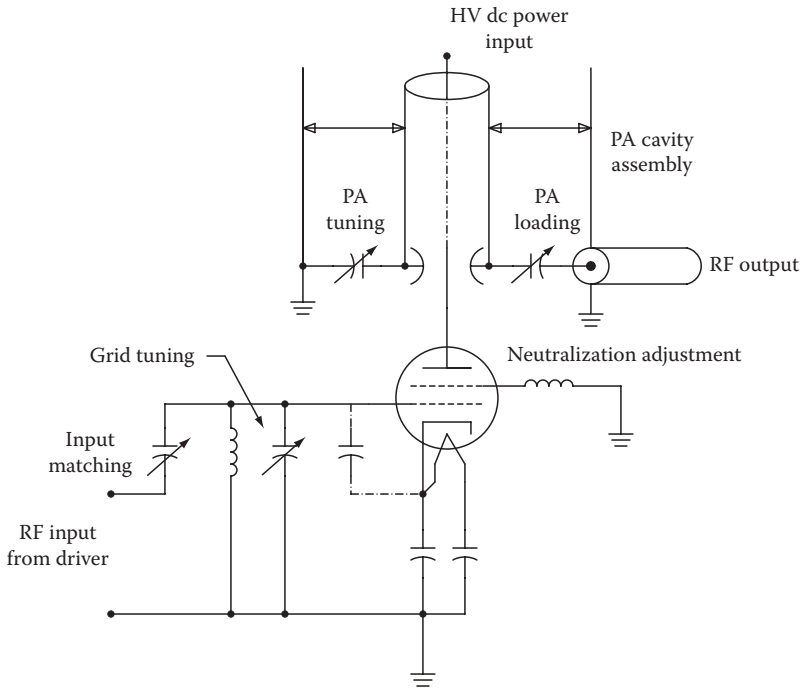


FIGURE 5.22
Grid-driven grounded-screen tetrode power amplifier.

schematic of a grid-driven tetrode amplifier. In this example, the screen is operated at dc ground potential, and the cathode (filament) is operated below ground by the amount of screen voltage required (grounded-screen operation). The advantage of this configuration is that stability problems resulting from undesired resonances in the screen bypass capacitors are eliminated. With this arrangement, however, the screen supply must be capable of handling the combined plate and screen currents.

With directly heated tubes, it is necessary to use filament bypass capacitors. During grounded-screen operation, these bypass capacitors must have a higher breakdown voltage rating (relative to grounded-grid operation) because they will have the dc screen voltage across them. The filament transformer must have additional insulation to withstand the dc screen voltage. The screen power supply provides a negative voltage in series with the cathode-to-ground, and must have the additional capacity to handle the sum of the plate and screen currents. A coaxial cavity is used in the output circuit so that the circulating current is spread over large surfaces, keeping the losses low. This cavity is a shorted $1/4$ -wavelength transmission line section that resonates the tube output capacitance. The length is preset to the desired carrier frequency, and a small-value variable capacitor is used to trim the system to resonance.

Pentode amplifiers have even higher gain than their tetrode counterparts. The circuit configuration and bias supply requirements for the pentode are similar to the tetrode's because the third (suppressor) grid is tied directly to ground. The additional isolating effect of the suppressor grid eliminates the need for neutralization in the pentode amplifier.

5.6.4 Impedance Matching into the Grid

The grid circuit usually is loaded (swamped) with added resistance. The purpose of this resistance is to broaden the bandwidth of the circuit by lowering the circuit Q and to provide a more constant load to the driver. The resistance also makes neutralization less critical so that the amplifier is less likely to become unstable with varying output circuit loading.

Cathode and filament lead inductance from inside the tube, through the socket and filament bypass capacitors to ground, can increase grid driver requirements. This effect is the result of RF current flowing from grid to filament through the tube capacitance and then through the filament lead inductance to ground. An RF voltage is developed on the filament that, in effect, causes the tube to be partly cathode driven. This undesirable extra drive power requirement can be minimized by series resonating the cathode return path with the filament bypass capacitors or by minimizing the cathode-to-ground inductance by using a tube socket incorporating thin-film dielectric *sandwich* capacitors for coupling and bypassing.

High-power grid-driven class C amplifiers require a swing of several hundred RF volts on the grid. To develop this high-voltage swing, the input impedance of the grid must be increased by the grid input matching circuit. Because the capacitance between the grid and the other tube elements may be 100 pF or more, the capacitive reactance at 100 MHz, for example, will be very low unless the input capacitance is parallel resonated with an inductor. Bandwidth can be maximized by minimizing any additional circuit capacitance and using a portion of the tube input capacitance as part of the impedance-transformation network. Figure 5.23 illustrates two popular methods of resonating and matching into the grid of a high-power tube. Both schemes can be analyzed by recognizing that the desired impedance transformation is produced by an equivalent L network.

In Figure 5.23a, a variable inductor (L_{in}) is used to raise the input reactance of the tube by bringing the tube input capacitance (C_{in}) almost to parallel resonance. Parallel resonance is not reached because a small amount of parallel capacitance (C_p) is required by the equivalent L network to transform the high impedance (Z_{in}) of the tube down to a lower value through the series matching inductor (L_s). This configuration has the advantage of providing a low-pass filter by using part of the tube's input capacitance to form C_p .

Figure 5.23b uses variable inductor L_{in} to take the input capacitance (C_{in}) past parallel resonance so that the tube's input impedance becomes slightly

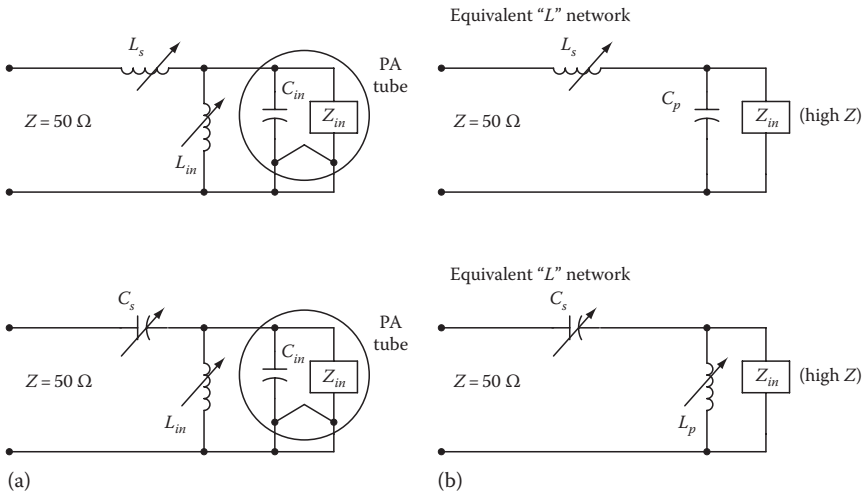


FIGURE 5.23

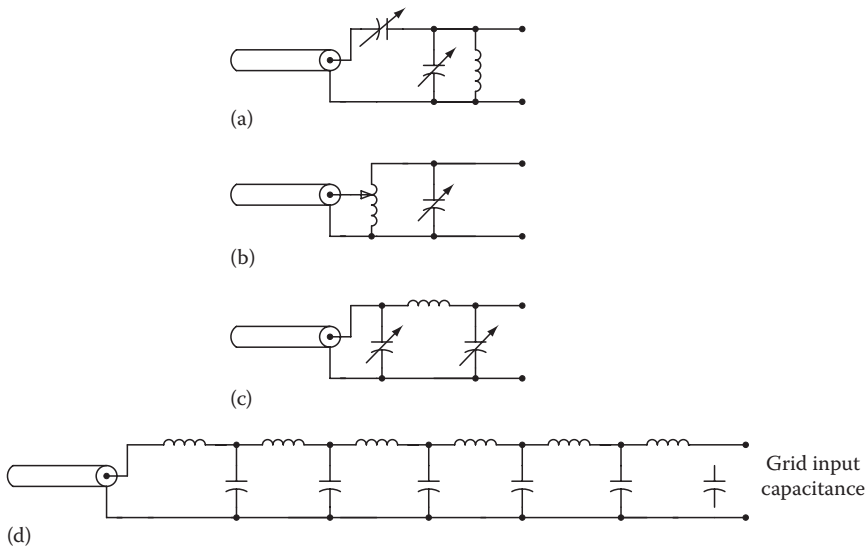
Input matching circuits: (a) inductive input matching and (b) capacitive input matching.

inductive. The variable series matching capacitor (C_s) forms the rest of the equivalent L network. This configuration constitutes a high-pass filter.

5.6.4.1 Interstage Coupling

The IPA output circuit and the final amplifier input circuit often are coupled by a coaxial transmission line. Impedance matching usually is accomplished at either end by one of the configurations shown in Figure 5.24. The circuits shown, except for Figure 5.24d, require some interactive adjustment of the tuning and loading elements to provide a satisfactory impedance match for each operating frequency and RF drive level. The circuit in Figure 5.24d utilizes multiple LC sections, with each section providing a small step in the total impedance transformation. This technique provides a broadband impedance match without adjustment, thereby improving transmitter stability, ease of operation, and maintainability. A single grid resonating control is sufficient to tune and match the 50Ω driver impedance to the high input impedance of the grid over a relatively wide band of frequencies and RF power levels.

The transmission line matching problem is eliminated in some transmitter designs by integrating an IPA stage, using one or more tubes, into the grid circuit of the final amplifier by having the plate of the IPA and the grid of the final tube share a common tuned circuit. This technique has the advantage of simplicity by not transforming the impedance down to 50Ω and then back up to the grid impedance level.

**FIGURE 5.24**

Interstage RF coupling circuits: (a) variable capacitance matching, (b) variable tank matching, (c) variable pi matching, and (d) fixed broadband impedance-transformation network.

Solid-state RF power devices possess a low load impedance at the device output terminal, so an impedance transformation that goes through the 50Ω intermediate impedance of a coaxial cable is necessary to couple these devices into the relatively high impedance of the final amplifier grid.

5.6.5 Neutralization

Most cathode-driven amplifiers using grounded-grid triodes do not require neutralization. It is necessary that the grid-to-ground inductance, both internal and external to the tube, be kept low to maintain this advantage. Omission of neutralization will allow a small amount of interaction between the output circuit and the input circuit through the plate-to-filament capacitance. This effect is not particularly noticeable because of the large coupling between the input and output circuits through the electron beam of the tube. Cathode-driven tetrodes have higher gain than triodes and, therefore, often require some form of neutralization.

Grid-driven high-gain tetrodes need accurate neutralization for best stability and performance. Self-neutralization can be accomplished simply by placing a small amount of inductance between the tube screen grid and ground. This inductance is usually in the form of several short adjustable-length straps. The RF current flowing from plate to screen in the tube also flows through this screen lead inductance. This develops a small RF

voltage on the screen, of the opposite phase, which cancels the voltage fed back through the plate-to-grid capacitance. This method of lowering the self-neutralizing frequency of the tube works only if the self-neutralizing frequency of the tube/socket combination is above the desired operating frequency before the inductance is added. Special attention must be given to minimizing the inductances in the tube socket by integrating distributed bypass capacitors into the socket and cavity deck assembly. Pentodes normally do not require neutralization because the suppressor grid effectively isolates the plate from the grid.

5.7 Special-Application Amplifiers

The operating environment of tubes used in commercial, industrial, and military applications usually is characterized by widely varying supply voltages, heavy and variable vibration, and significant changes in loading. Scientific applications are no less demanding. Research projects span a wide range of powers and frequencies. Typical uses for power vacuum tubes include the following:

- Super proton synchrotron delivering 2.4 MW at 200 MHz
- Fusion reactor using 33 MW of high-frequency heating in the 25–60 MHz band
- Plasma research using 12 MW in the 60–120 MHz range
- Fusion research using a 500 kW ion cyclotron resonance heating generator operating at 30–80 MHz

It can be seen from these examples that efficiency and the resulting heat dissipation are major considerations in the design of an amplifier or oscillator for scientific uses.

5.7.1 Distributed Amplification

Specialized research and industrial applications often require an amplifier covering several octaves of bandwidth. At microwave frequencies, a traveling wave tube (TWT) may be used. At MF and below, multiple tetrodes usually are combined to achieve the required bandwidth. A *distributed amplifier* consists of multiple tetrodes and lumped-constant transmission lines. The lines are terminated by load resistances with magnitudes equal to their characteristic impedances. A cyclical wave of current is present on the output transmission line circuit, with each tube contributing its part in proper phase. A typical configuration for a distributed amplifier includes 8–16 tubes. The efficiency of such a design is low.

5.7.2 Radar

The exciter stage of a radar set comprises oscillators, frequency multipliers, and mixers. The signals produced depend on whether the transmitter output device is operated as a power amplifier or an oscillator.

Transmitters using power oscillators such as *magnetrons* determine the RF by tuning of the device itself. In a conventional (*noncoherent*) radar system, the only frequency required is the *local oscillator* (LO). The LO differs from the magnetron frequency by an intermediate frequency, and this difference usually is maintained with an *automatic frequency control* (AFC) loop. Figure 5.25 shows a simple magnetron-based radar system with two methods of tuning:

- Slaving the magnetron to follow the *stable local oscillator* (STALO)
- Slaving the STALO to follow the magnetron

If the radar must use *coherent detection* (such as in Doppler applications), a second oscillator, called a *coherent oscillator* (COHO), is required. The COHO operates at the intermediate frequency and provides a reference output for signal processing circuits.

The synchronizer circuits in the exciter supply timing pulses to various radar subsystems. In a simple marine radar, this may consist of a single multivibrator that triggers the transmitter. In a larger system, 20–30 timing pulses may be required. These may turn the beam current on and off in various transmitter stages, start and stop RF pulse time attenuators, start display sweeps, and perform numerous other functions. Newer radar systems generate the required timing signals digitally. A digital synchronizer is illustrated in Figure 5.26.

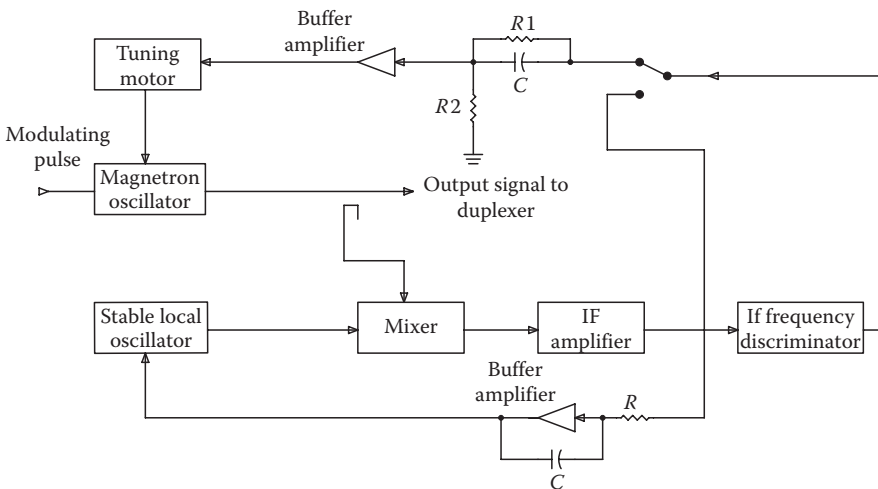


FIGURE 5.25

Two common approaches to automatic frequency control of a magnetron oscillator.

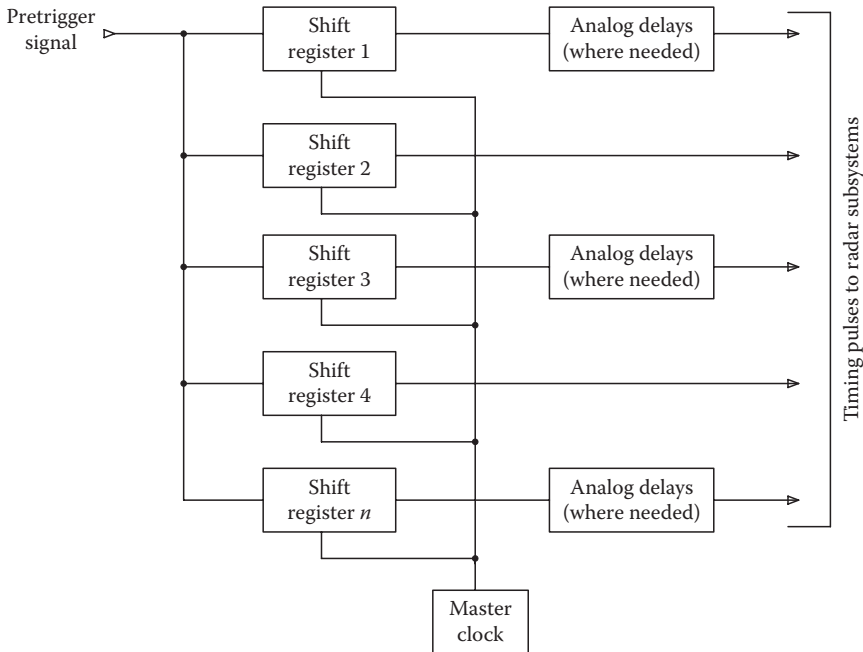


FIGURE 5.26
Digital synchronizer system for radar applications.

5.7.2.1 Modulator

A radar system RF amplifier usually centers on one of two microwave devices: a *crossed-field tube* or *linear-beam tube*. Both are capable of high peak output power at microwave frequencies. To obtain high efficiency from a pulsed radar transmitter, it is necessary to cut off the current in the output tube between pulses. The modulator performs this function. Some RF tubes include control electrodes or grids to achieve the same result. Three common types of modulators are used in radar equipment:

- *Line-type modulator* (Figure 5.27). This common radar modulator is used most often to pulse a magnetron. Between pulses, a charge is stored in a *pulse-forming network* (PFN). A trigger signal fires a *thyatron* tube, short-circuiting the input to the PFN, which causes a voltage pulse to appear at the primary of transformer *T1*. The PFN components are chosen to produce a rectangular pulse at the magnetron cathode, with the proper voltage and current to excite the magnetron to oscillation. An advantage of this design is its simplicity. A drawback is the inability to electronically change the width of the transmitted pulse.

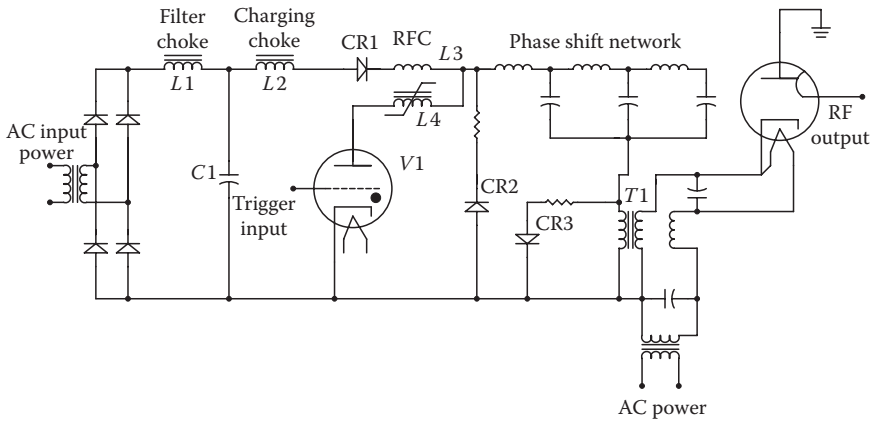


FIGURE 5.27
Line-type modulator for radar.

- *Active-switch modulator* (Figure 5.28). This system permits pulse width variation, within the limitations of the energy stored in the high-voltage power supply. A switch tube controls the generation of RF by completing the circuit path from the output tube to the power supply, or by causing stored energy to be dumped to the output device. The figure shows the basic design of an active-switch modulator and three variations on the scheme. The circuits differ in the method of coupling power supply energy to the output tube (capacitor-coupled, transformer-coupled, or a combination of the two methods).
- *Magnetic modulator* (Figure 5.29). This design is the simplest of the three modulators discussed. No thyatron or switching device is used. Operation of the modulator is based on the saturation characteristics of inductors $L1$, $L2$, and $L3$. A long-duration low-amplitude pulse is applied to $L1$, which charges $C1$. As $C1$ approaches its fully charged state, $L2$ saturates and the energy in $C1$ is transferred in a resonant fashion to $C2$. This process continues to the next stage ($L3$ and $C3$). The transfer time is set by selection of the components to be about one-tenth that of the previous stage. At the end of the chain, a short-duration high-amplitude pulse is generated, exciting the RF output tube.

Because the applications for radar vary widely, so do antenna designs. Sizes range from less than 1 ft to hundreds of feet in diameter. An antenna intended for radar applications must direct radiated power from the transmitter to the azimuth and elevation coordinates of the target. It also must serve as a receive antenna for the echo.

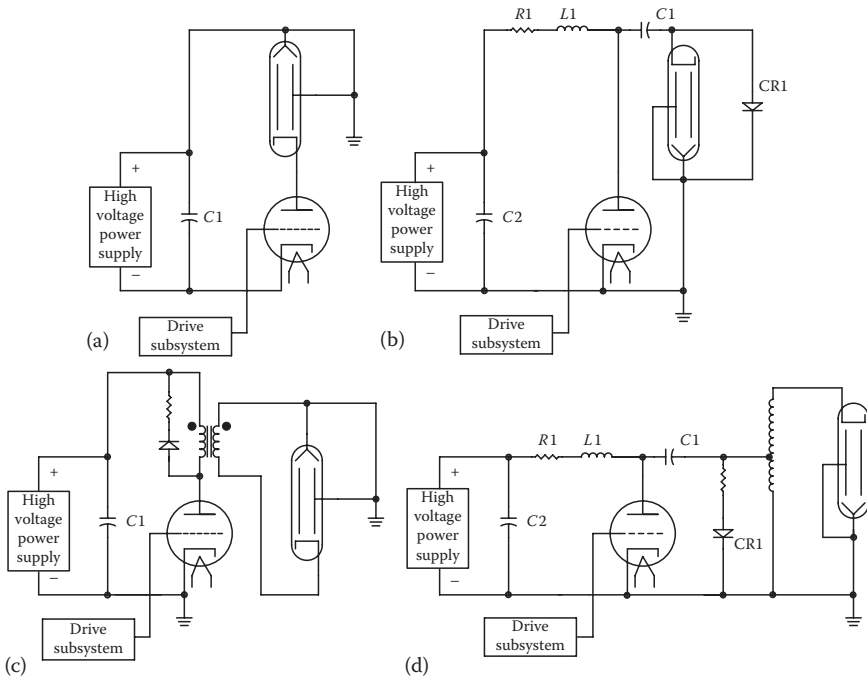


FIGURE 5.28 Active-switch modulator circuits: (a) direct-coupled system, (b) capacitor-coupled, (c) transformer-coupled, and (d) capacitor- and transformer-coupled.

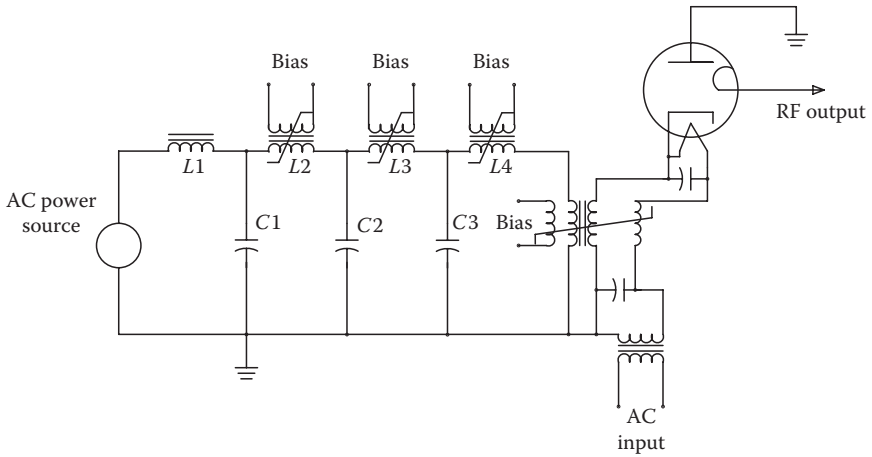


FIGURE 5.29 Magnetic modulator circuit.

References

1. Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Eimac, San Carlos, CA, 1984.
2. Woodard, G. W., AM transmitters, in *NAB Engineering Handbook*, 8th edn., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, p. 356, 1992.

Bibliography

- Chireix, H., High power outphasing modulation, *Proceedings of the IRE*, 23(11), 1370, 1935.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, 1992.
- Doherty, W. H., A new high efficiency power amplifier for modulated waves, *Proceedings of the IRE*, 24(9), 1163, 1936.
- Fink, D. and D. Christiansen, *Electronic Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.
- Gray, T. S., *Applied Electronics*, Massachusetts Institute of Technology, Cambridge, MA, 1954.
- Heising, R. A., Modulation in radio telephony, *Proceedings of the IRE*, 9(3), 305, 1921.
- Heising, R. A., Transmission system, U.S. Patent no. 1,655,543, January 1928.
- High Power Transmitting Tubes for Broadcasting and Research*, Philips Technical Publication, Eindhoven, the Netherlands, 1988.
- Honey, J. F., Performance of AM and SSB communications, *Tele-Tech.*, September 1953.
- Jordan, E. C., *Reference Data for Engineers: Radio, Electronics, Computers, and Communications*, 7th edn., Howard W. Sams, Indianapolis, IN, 1985.
- Martin, T. L., Jr., *Electronic Circuits*, Prentice Hall, Englewood Cliffs, NJ, 1955.
- Mina, A. and F. Parry, Broadcasting with megawatts of power: The modern era of efficient powerful transmitters in the Middle East, *IEEE Transactions on Broadcasting*, 35(2), 121–130, 1989.
- Pappenfus, E. W., W. B. Bruene, and E. O. Schoenike, *Single Sideband Principles and Circuits*, McGraw-Hill, New York, 1964.
- Ridgwell, J. F., A new range of beam-modulated high-power UHF television transmitters, *Communications & Broadcasting* (28), 27–35, 1988.
- Skolnik, M. I. (ed.), *Radar Handbook*, McGraw-Hill, New York, 1980.
- Terman, F. E. and J. R. Woodyard, A high efficiency grid-modulated amplifier, *Proceedings of the IRE*, 26(8), 929, 1938.
- Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.
- Weldon, J. O., Amplifiers, U. S. Patent no. 2,836,665, May 1958.
- Whitaker, J. C., *AC Power Systems*, 2nd edn., CRC Press, Boca Raton, FL, 1998.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1991.
- Woodard, G. W., AM transmitters, in *NAB Engineering Handbook*, 9th edn., J. C. Whitaker (ed.), National Association of Broadcasters, Washington, DC, pp. 353–381, 1998.

6

Microwave Power Tubes

6.1 Introduction

Microwave power tubes span a wide range of applications, operating at frequencies from 300 MHz to 300 GHz with output powers from a few hundred watts to more than 10 MW. Applications range from the familiar to the exotic. The following devices are included under the general description of microwave power tubes:

- Klystron, including the *reflex* and *multicavity* klystron
- *Multistage depressed collector* (MSDC) klystron
- *Inductive output tube* (IOT)
- *Traveling wave tube* (TWT)
- *Crossed-field tube*
- *Coaxial magnetron*
- *Gyrotron*
- *Planar triode*
- High-frequency tetrode
- *Diacrode*

This wide variety of microwave devices has been developed to meet a broad range of applications. Some common uses include the following:

- UHF-TV transmission
- Shipboard and ground-based radar
- Weapons guidance systems
- *Electronic countermeasure* (ECM) systems
- Satellite communications
- Tropospheric scatter communications
- Fusion research

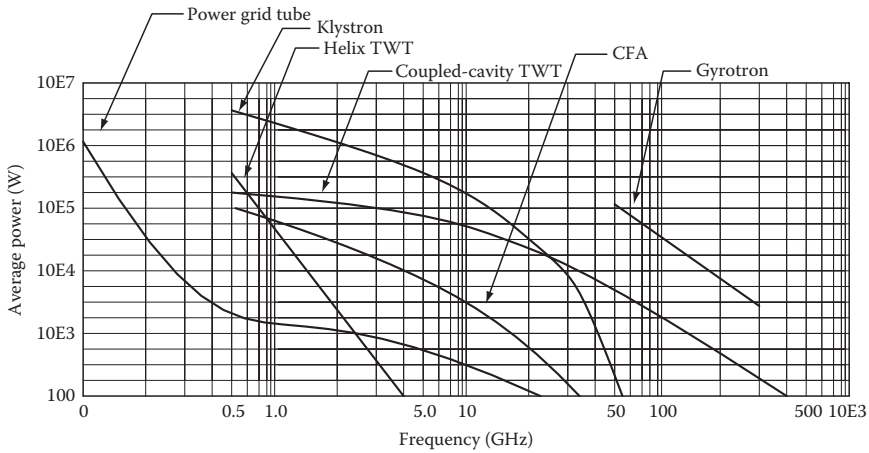


FIGURE 6.1

Microwave power tube type as a function of frequency and output power.

As new applications are identified, improved devices are designed to meet the needs. Microwave power tube manufacturers continue to push the limits of frequency, operating power, and efficiency. Microwave technology, therefore, is an evolving science. Figure 6.1 charts device type as a function of operating frequency and power output.

Two principal classes of microwave vacuum devices are in common use today:

- *Linear-beam* tubes
- *Crossed-field* tubes

Each class serves a specific range of applications. In addition to these primary classes, some power grid tubes also are used at microwave frequencies.

6.1.1 Linear-Beam Tubes

In a linear-beam tube, as the name implies, the electron beam and the circuit elements with which it interacts are arranged linearly. The major classifications of linear-beam tubes are shown in Figure 6.2. In such a device, a voltage applied to an anode accelerates electrons drawn from a cathode, creating a beam of kinetic energy. Power supply potential energy is converted to kinetic energy in the electron beam as it travels toward the microwave circuit. A portion of this kinetic energy is transferred to microwave energy as radio frequency (RF) waves slow down the electrons. The remaining beam energy is either dissipated as heat or returned to the power supply at the collector. Because electrons will repel one another, there usually is an applied magnetic focusing field to maintain the beam

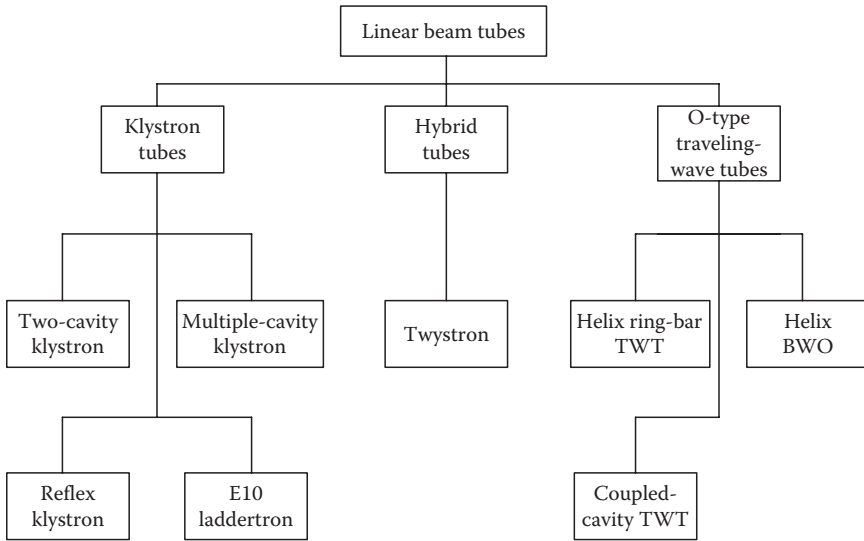


FIGURE 6.2
Types of linear-beam microwave tubes.

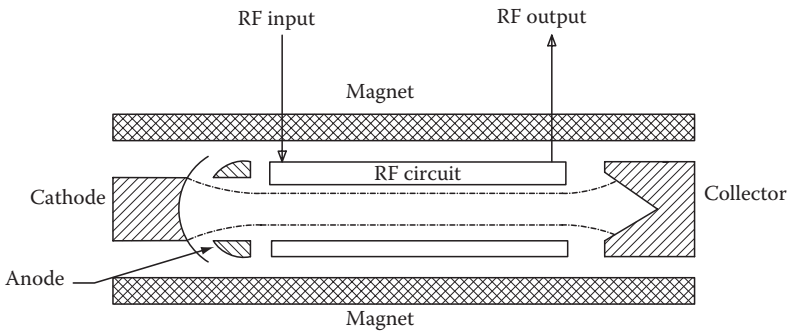


FIGURE 6.3
Schematic diagram of a linear-beam tube.

during the interaction process. The magnetic field is supplied either by a solenoid or permanent magnets. Figure 6.3 shows a simplified schematic of a linear-beam tube.

6.1.2 Crossed-Field Tubes

The magnetron is the pioneering device of the family of crossed-field tubes. The family tree of this class of devices is shown in Figure 6.4. Although the physical appearance differs from that of linear-beam tubes, which are usually circular in format, the major difference is in the interaction physics that requires a magnetic field at right angles to the applied electric field. Whereas the

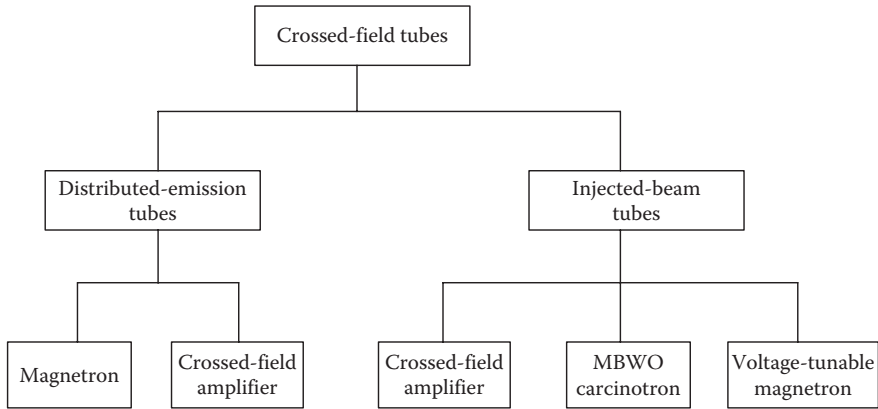


FIGURE 6.4
Types of crossed-field microwave tubes.

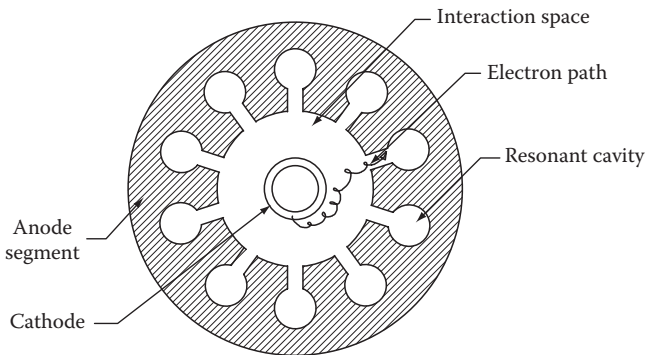


FIGURE 6.5
Magnetron electron path looking down into the cavity with the magnetic field applied.

linear-beam tube sometimes requires a magnetic field to maintain the beam, the crossed-field tube always requires a magnetic focusing field.

Figure 6.5 shows a cross section of the magnetron, including the magnetic field applied perpendicular to the cathode–anode plane. The device is basically a diode with the anode composed of a plurality of resonant cavities. The interaction between the electrons emitted from the cathode and the crossed electric and magnetic fields produces a series of space-charge spokes that travel around the anode–cathode space in a manner that transfers energy to the RF signal supported by the multicavity circuit. The mechanism is highly efficient.

6.1.2.1 Crossed-Field Amplifier

Figure 6.6 shows the family tree of the crossed-field amplifier (CFA). The configuration of a typical distributed emission amplifier is similar to that of the

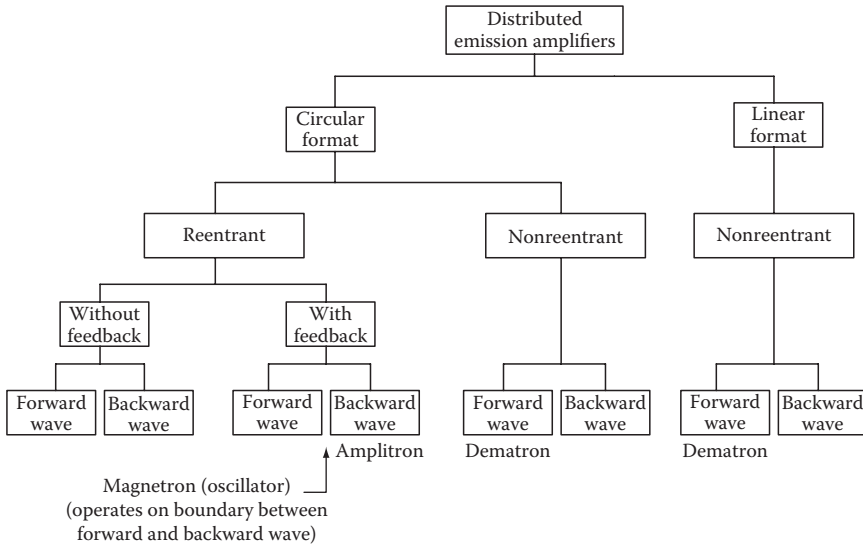


FIGURE 6.6
Family tree of the distributed emission CFA.

magnetron except that the device has an input for the introduction of RF energy into the circuit. Current is obtained primarily by secondary emission from the negative electrode that serves as a cathode throughout all or most of the interaction space. The earliest versions of this tube type were called *amplitrons*.

The CFA has been deployed in radar systems operating from ultrahigh frequency (UHF) to the Ku-band, and at power levels of several megawatts. In general, bandwidth ranges from a few percent to as much as 25% of the center frequency.

6.2 Grid Vacuum Tubes

The physical construction of a vacuum tube causes the output power and available gain to decrease with increasing frequency. The principal limitations faced by grid-based devices include the following:

- **Physical size.** Ideally, the RF voltages between electrodes should be uniform, but this condition cannot be realized unless the major electrode dimensions are significantly less than 1/4 wavelength at the operating frequency. This restriction presents no problems at very high frequency (VHF), but as the operating frequency increases into the microwave range, severe restrictions are placed on the physical size of individual tube elements.

- Electron transit time. Interelectrode spacing, principally between the grid and the cathode, must be scaled inversely with frequency to avoid problems associated with electron transit time. Possible adverse conditions include (1) excessive loading of the drive source, (2) reduction in power gain, (3) back heating of the cathode as a result of electron bombardment, and (4) reduced conversion efficiency.
- Voltage standoff. High-power tubes operate at high voltages. This presents significant problems for microwave vacuum tubes. For example, at 1 GHz the grid-cathode spacing must not exceed a few mils. This places restrictions on the operating voltages that may be applied to the individual elements.
- Circulating currents. Substantial RF currents may develop as a result of the inherent interelectrode capacitances and stray inductances/capacitances of the device. Significant heating of the grid, connecting leads, and vacuum seals may result.
- Heat dissipation. Because the elements of a microwave grid tube must be kept small, power dissipation is limited.

Still, a number of grid-based vacuum tubes find applications at high frequencies. For example, planar triodes are available that operate at several gigahertz, with output powers of 1–2 kW in pulsed service. Efficiency (again for pulsed applications) ranges from 30% to 60%, depending on the frequency.

6.2.1 Planar Triode

A cross-sectional diagram of a planar triode is shown in Figure 6.7. The envelope is made of ceramic, with metal members penetrating the ceramic to provide for connection points. The metal members are shaped either as disks or as disks with cylindrical projections.

The cathode is typically oxide coated and indirectly heated. The key design objective for a cathode is high emission density and long tube life. Low-temperature emitters are preferred because high cathode temperatures typically result in more evaporation and shorter life.

The grid of the planar triode is perhaps the greatest design challenge for tube manufacturers. Close spacing of small-sized elements is needed, at tight tolerances. Good thermal stability also is required, because the grid is subjected to heating from currents in the element itself, plus heating from the cathode and bombardment of electrons from the cathode.

The anode, usually made of copper, conducts the heat of electron bombardment to an external heat sink. Most planar triodes are air cooled.

Planar triodes designed for operation at 1 GHz and above are used in a variety of circuits. The grounded-grid configuration is most common.

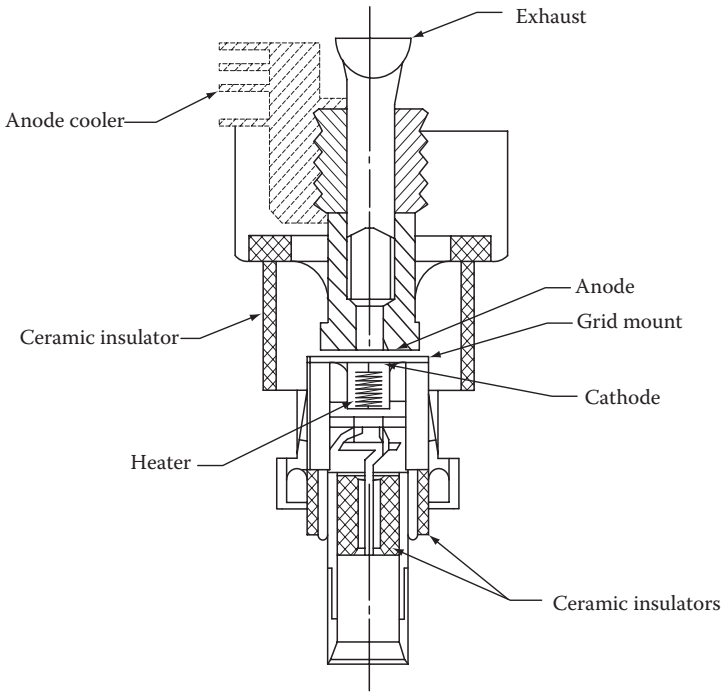


FIGURE 6.7
Cross section of a 7289 planar triode.

The plate resonant circuit is cavity based, using waveguide, coaxial line, or stripline. Electrically, the operation of the planar triode is much more complicated at microwave frequencies than at low frequencies. Figure 6.8a compares the elements at work for a grounded-grid amplifier operating at low frequencies and Figure 6.8b compares the situation at microwave frequencies. The equivalent circuit is made more complex by

- Stray inductance and capacitance of the tube elements
- Effects of the tube contact rings and socket elements
- Distributed reactance of cavity resonators and the device itself
- Electron transit-time effects, which result in resistive loading and phase shifts

Reasonable gains of 5–10 dB may be achieved with a planar triode. Increased gain is available by cascading stages. Interstage coupling may consist of waveguide or coaxial-line elements. Tuning is accomplished by varying the cavity inductance or capacitance. Additional bandwidth is possible by stagger tuning of cascaded stages.

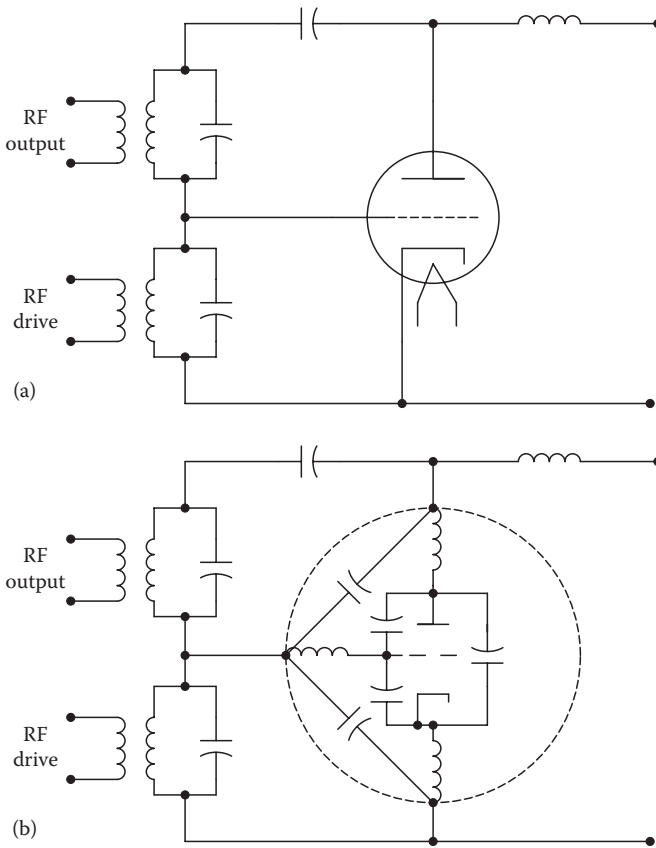


FIGURE 6.8 Grounded-grid equivalent circuits: (a) low-frequency operation and (b) microwave-frequency operation. The cathode-heating and grid-bias circuits are not shown.

6.2.2 High-Power UHF Tetrode

New advancements in vacuum tube technology have permitted the construction of high-power UHF transmitters based on tetrodes. Such devices are attractive because they inherently operate in a relatively efficient class AB mode. UHF tetrodes operating at high power levels provide essentially the same specifications, gain, and efficiency as tubes operating at lower powers. The anode power supply is much lower in voltage than the collector potential of a klystron- or IOT-based system (8 kV is common). Also, the tetrode does not require a focusing magnet system.

Efficient removal of heat is the key to making a tetrode practical at high power levels. Such devices typically use water or vapor-phase cooling. Air cooling at such levels is impractical because of the fin size that would be required.

Also, the blower for the tube would have to be quite large, reducing the overall transmitter ac-to-RF efficiency.

Another drawback inherent in tetrode operation is that the output circuit of the device appears electrically in series with the input circuit and the load [1]. The parasitic reactance of the tube elements, therefore, is a part of the input and output tuned circuits. It follows, then, that any change in the operating parameters of the tube as it ages can affect tuning. More importantly, the series nature of the tetrode places stringent limitations on internal element spacings and the physical size of those elements in order to minimize the electron transit time through the tube vacuum space. It is also fair to point out, however, that the tetrode's input-to-output circuit characteristic has at least one advantage: power delivered to the input passes through the tube and contributes to the total power output of the transmitter. Because tetrodes typically exhibit low gain compared to a klystron-based device, significant power can be required at the input circuit. The pass-through effect, therefore, contributes to the overall operating efficiency of the transmitter.

The expected lifetime of a tetrode in UHF service is usually shorter than a klystron of the same power level. Typical lifetimes of 8,000–15,000 h have been reported. Intensive work, however, has led to products that offer higher output powers and extended operating lifetime, while retaining the benefits inherent in tetrode devices.

6.2.3 Diacrode

The *Diacrode* (Thomson) is an adaptation of the high-power UHF tetrode. The operating principle of the Diacrode is basically the same as that of the tetrode. The anode current is modulated by an RF drive voltage applied between the cathode and the power grid. The main difference is in the position of the active zones of the tube in the resonant coaxial circuits, resulting in improved reactive current distribution in the electrodes of the device.

Figure 6.9 compares the conventional tetrode with the Diacrode. The Diacrode includes an electrical extension of the output circuit structure to an external cavity [2]. The small dc-blocked cavity rests on top of the tube, as illustrated in Figure 6.10.

The cavity is a quarter-wave transmission line, as measured from the top of the cavity to the vertical center of the tube. The cavity is short-circuited at the top, reflecting an open circuit (current minimum) at the vertical center of the tube and a current maximum at the base of the tube, like the conventional tetrode, and a second current maximum above the tube at the cavity short-circuit.

With two current maximums, the RF power capability of the Diacrode is double that of the equivalent tetrode, while the element voltages remain the same. All other properties and aspects of the Diacrode are basically

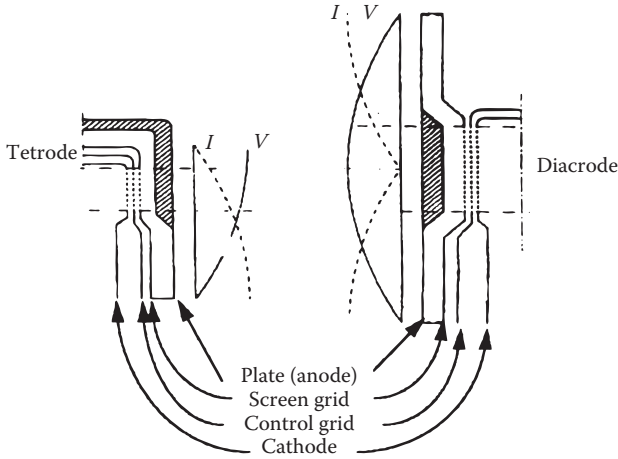


FIGURE 6.9

Cutaway view of the tetrode (*left*) and the Diacrode (*right*). Note that the RF current peaks above and below the Diacrode center while on the tetrode there is only one peak at the bottom. (After Hulick, T.P., 60kW Diacrode UHF TV transmitter design, performance and field report, in *Proceedings of the 1996 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, p. 442, 1996.)

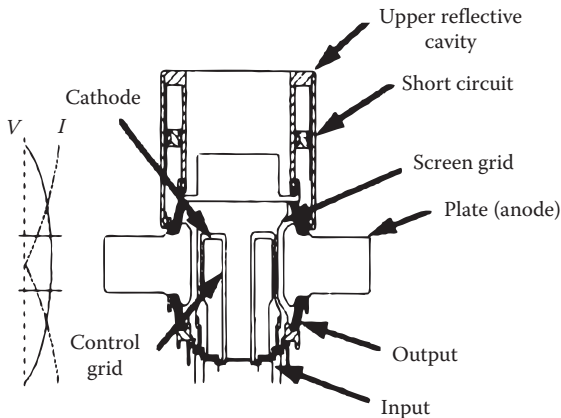


FIGURE 6.10

Elements of the Diacrode, including the upper cavity. Double current, and, consequently, double power, is achieved with the device because of the current peaks at the top and bottom of the tube, as shown. (After Hulick, T.P., 60kW Diacrode UHF TV transmitter design, performance and field report, in *Proceedings of the 1996 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, p. 442, 1996.)

identical to the TH563 high-power UHF tetrode (Thomson), upon which the Diacrode is patterned.

Some of the benefits of such a device, in addition to the robust power output available, are the low high-voltage requirements (low relative to a klystron/IOT-based system that is), small size, and simple replacement procedures.

6.3 Klystron

The klystron is a *linear-beam* device that overcomes the transit-time limitations of a grid-controlled tube by accelerating an electron stream to a high velocity before it is modulated. Modulation is accomplished by varying the velocity of the beam, which causes the drifting of electrons into *bunches* to produce RF *space current*. One or more cavities reinforce this action at the operating frequency. The output cavity acts as a transformer to couple the high-impedance beam to a low-impedance transmission line. The frequency response of a klystron is limited by the impedance-bandwidth product of the cavities, but may be extended through stagger tuning or the use of multiple-resonance filter-type cavities.

The klystron is one of the primary means of generating high power at UHF and above. Output powers for multicavity devices range from a few thousand watts to 10 MW or more. The klystron provides high gain and requires little external support circuitry. Mechanically, the klystron is relatively simple. It offers long life and requires minimal routine maintenance.

6.3.1 Reflex Klystron

The reflex klystron uses a single-cavity resonator to modulate the RF beam and extract energy from it. The construction of a reflex klystron is shown in Figure 6.11. In its basic form, the tube consists of the following elements:

- A cathode
- Focusing electrode at cathode potential
- Coaxial line or reentrant-type cavity resonator, which also serves as an anode
- *Repeller* or *reflector* electrode, which is operated at a moderately negative potential with respect to the cathode

The cathode is so shaped that, in relation to the focusing electrode and anode, an electron beam is formed that passes through a gap in the resonator, as shown in the figure, and travels toward the repeller. Because the repeller has a negative potential with respect to the cathode, it turns the electrons back toward the anode, where they pass through the anode gap a second time. By varying the applied voltage on the reflector electrode, phasing of the beam can be varied to produce the desired oscillating mode and to control the frequency of oscillation.

The variation of position with time for electrons in the anode-repeller space is illustrated in Figure 6.12. Path *a* corresponds to an electron that emerges from the anode with a velocity corresponding to the anode voltage. This electron follows a parabolic path, as shown, determined by the electric field in the anode-repeller space.

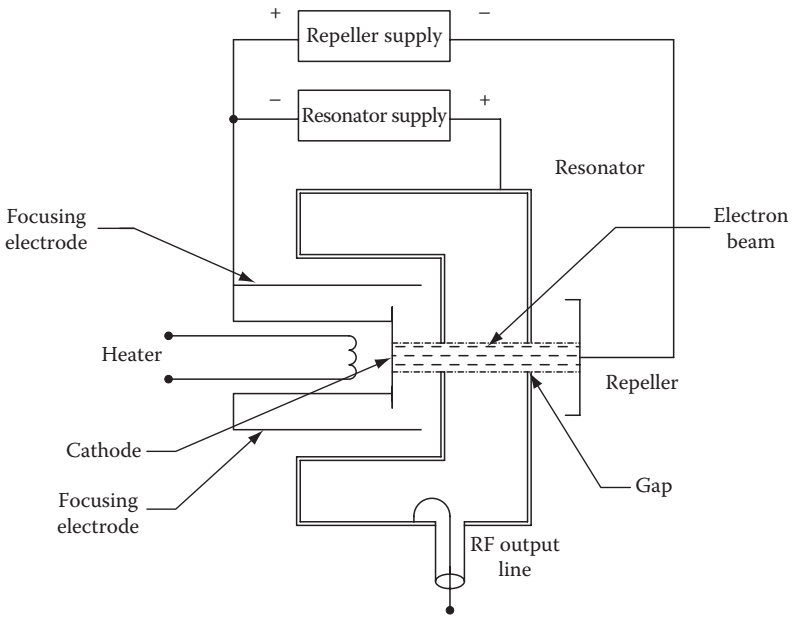


FIGURE 6.11
Schematic representation of a reflex klystron.

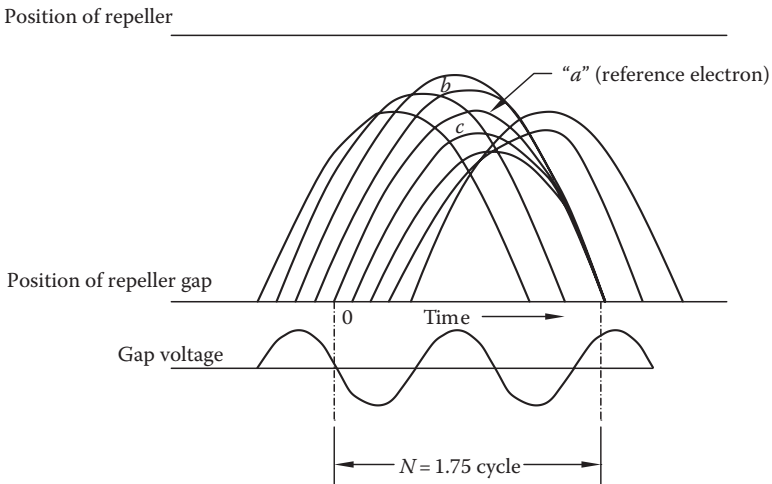


FIGURE 6.12
Position-time curves of electrons in the anode-repeller space, showing the tendency of the electrons to bunch around the electron passing through the anode at the time when the alternating gap voltage is zero and becoming negative.

Operation of the reflex klystron can best be understood by examining the movement of electrons inside the device. Assume that oscillations exist in the resonator so that an alternating voltage develops across the gap. Assume further that the electron corresponding to path *a* passed through the gap at the instant that this alternating voltage across the gap was zero and becoming negative. An electron passing through the gap just before electron *a* will encounter an accelerating voltage across the gap and, therefore, will emerge from the anode with greater velocity than the first or *reference* electron. This second electron, accordingly, penetrates farther toward the repeller against the retarding field and, as a result, takes longer to return to the anode. Consequently, this electron follows path *b*, as shown in Figure 6.12, and tends to arrive at the anode on its return path at the same time as the reference electron because its earlier start is more or less compensated for by increased transit time. In a similar manner, an electron passing through the anode gap slightly later than the reference electron will encounter a negative or retarding field across the gap, and so will emerge from the anode with less velocity than the electron that follows path *a*. This third electron will then follow trajectory *c* and return to the anode more quickly than electron *a*. Electron *c*, therefore, tends to return to the anode at about the same time as electron *a*, because the later start of electron *c* is compensated for by the reduced transit time.

This variation with time of the velocity of electrons emerging from the anode is termed *velocity modulation*. The effect of this phenomenon can be seen in Figure 6.12 to cause a *bunching* of electrons about the electron that passed through the gap when the resonator voltage was zero and became negative. This bunching causes the electrons that are returned toward the anode by the repeller to pass through the anode gap in bursts or pulses, one each cycle. When these pulses pass through the gap at such a time that the electrons in the pulse are slowed as a result of the alternating voltage existing across the gap at the instant of their return passage, energy will be delivered to the oscillations in the resonator, thereby assisting in maintaining the oscillations. This condition corresponds to a transit time *N* from the resonator toward the repeller and back to the resonator of approximately

$$N = n + \frac{3}{4} \quad (6.1)$$

where *n* is an integer (including zero).

The transit time in the anode-repeller space in any particular case depends upon the following:

- The anode voltage
- Repeller voltage
- Geometry of the anode-repeller space

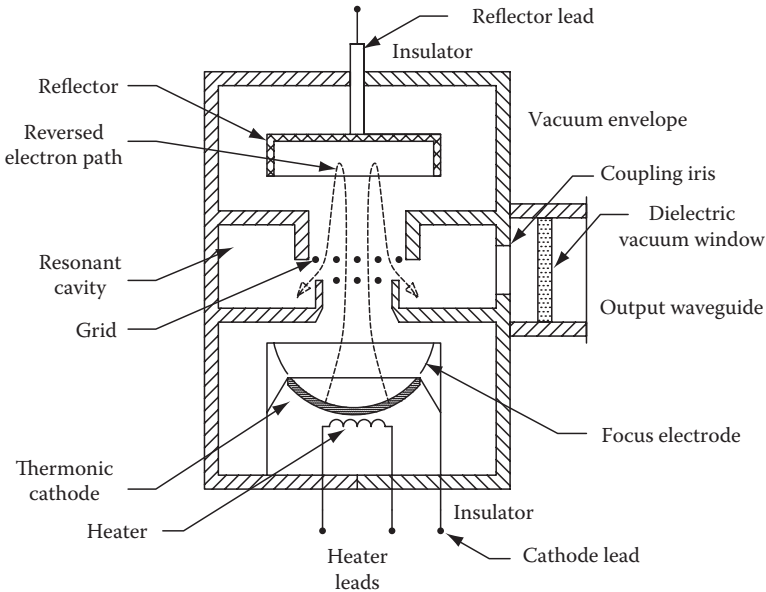


FIGURE 6.13
Schematic cross section of a reflex klystron oscillator.

The extent of the bunching action that takes place when the transit time of the reference electron has the correct value for sustaining oscillations is determined by the following:

- The amplitude of the alternating voltage across the resonator gap in relation to the anode and repeller voltage
- The geometry of the repeller space

The reflex klystron typically includes a grid to concentrate the electric field so that it may efficiently couple to the electron beam. Such a device is illustrated in Figure 6.13.

The reflex klystron may be used as a local oscillator, low-power FM transmitter, or test signal source. Reflex tubes are used primarily from 4 to 40 GHz. Power outputs of 1 W or less are common.

The reflex tube is the only klystron in which *beam feedback* is used to produce output energy. In klystrons with more than one cavity, the electron beam passes through each cavity in succession.

6.3.2 Two-Cavity Klystron

The two-cavity klystron operates on the same bunching principle as the reflex klystron, but incorporates two cavities connected by a *drift tube*. Figure 6.14 shows a cross section of a classic device. The heater/cathode element

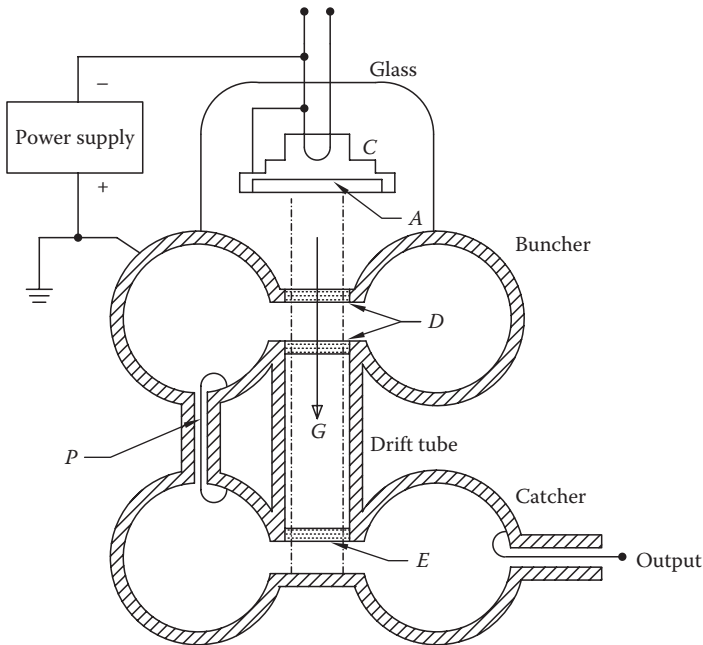


FIGURE 6.14
Cross section of a classic two-cavity klystron oscillator.

(shown as *A* in the figure) produces an electron beam in conjunction with a focusing electrode. The route taken by the electrons is as follows:

- The beam passes through grid elements *D* in the side of a reentrant cavity resonator (the buncher).
- The beam then passes through the drift tube, which is at the same electrical potential as the buncher.
- Finally, the beam enters a second resonator termed the *collector*, which is provided with a grid *E*.

The cathode and its associated focusing electrode are maintained at a high negative potential with respect to the remaining part of the structure, all of which is at the same dc potential. The entire arrangement illustrated in the figure is enclosed in a vacuum.

The operational principles of the two-cavity klystron are similar in nature to those of the reflex klystron. Assume, first, that oscillations exist in the buncher so that an alternating voltage is present across the gap *D*. When this voltage is zero but just becoming positive, an electron passing through the buncher travels through the grids *D*, down the drift tube, and into the collector resonator with unchanged velocity. However, an electron that passes through the buncher slightly later receives acceleration while passing

through, because of the positive alternating field that it encounters between grids D, and enters the drift tube with increased velocity. This later electron, therefore, tends to overtake the earlier electron. Similarly, an electron that arrives at the buncher slightly earlier than the first (reference electron) encounters a field between grids D that opposes its motion. Hence, this early electron enters the drift tube with reduced velocity and tends to drop back and be overtaken by the reference electron.

As a result of these actions, the electrons bunch together as they travel down the drift tube. This effect is more pronounced at certain distances from the buncher. If the collector is located at a distance where the bunching is pronounced, the electrons enter the element in pulses, one pulse per cycle.

With proper adjustment, the amount of power required to produce the bunching effect is relatively small compared with the amount of energy delivered by the electron beam to the collector. As a result, the klystron operates as an amplifying device.

6.3.2.1 Two-Cavity Klystron Oscillator

The two-cavity klystron oscillator is designed for applications requiring moderate power (up to 100 W), stable frequency output, and low sideband noise. The device has a coupling iris on the wall between the two cavities. The tube can be frequency-modulated by varying the cathode voltage about the center of the oscillating mode. Although it is more efficient and powerful than the reflex klystron, the two-cavity klystron requires more modulator power. Applications for the two-cavity klystron include Doppler radar systems.

6.3.2.2 Two-Cavity Klystron Amplifier

Similar in design to the two-cavity oscillator, the two-cavity klystron amplifier provides limited power output (10 W or less) and moderate gain (about 10 dB). A driving signal is coupled into the input cavity, which produces velocity modulation of the beam. After the drift space, the density-modulated beam induces current in the output resonator. Electrostatic focusing of the beam is common.

The two-cavity klystron finds only limited applications because of its restrictions on output power and gain.

6.3.3 Multicavity Klystron

The multicavity klystron is an important device for amplifying signals to high power levels at microwave frequencies [3]. Each cavity tuned to the operating frequency adds about 20 dB gain to the 10 dB gain offered by the basic two-cavity klystron amplifier. Overall gains of 60 dB are practical. Cavities may be tuned to either side of resonance to broaden the operating bandwidth of the device. Klystrons with up to eight cavities have been produced. Operating power for continuous wave (CW) klystrons ranges up to 1 MW per device, and as much as 50 MW per device for pulsed applications.

The primary physical advantage of the klystron over a grid-based power tube is that the cathode-to-collector structure is virtually independent of transit-time effects. Therefore, the cathode can be made large and the electron beam density kept low.

The operating frequency of a klystron may be fixed—determined by the mechanical characteristics of the tube and its cavities—or tunable. Cavities are tuned mechanically using one of several methods, depending on the operating power and frequency. Tuning is accomplished by changing the physical dimensions of the cavities using one or more of the following techniques:

- *Cavity wall deformation*, in which one wall of the cavity consists of a thin diaphragm that is moved in and out by a tuning mechanism. About 3% frequency shift may be accomplished using this method, which varies the inductance of the cavity.
- *Movable cavity wall*, in which one wall of the cavity is moved in or out by a tuning mechanism. About 10% frequency shift is possible with this approach, which varies the inductance of the cavity.
- *Paddle element*, in which an element inside the cavity moves perpendicularly to the beam and adds capacitance across the interaction gap. A tuning range of about 25% is provided by this approach.
- *Combined inductive-capacitive tuning*, which uses a combination of the previous methods. Tuning variations of 35% are possible.

Each of these tuning methods may be used whether the cavity is inside or outside the vacuum envelope of the tube. Generally speaking, however, tubes that use external cavities provide more adjustment range, usually on the order of 35%. Bandwidth may be increased by stagger tuning of the cavities, at the expense of gain.

High conversion efficiency requires the formation of electron bunches, which occupy a small region in velocity space, and the formation of *inter-bunch regions* with low electron density. The latter is particularly important because these electrons are phased to be accelerated into the collector at the expense of the RF field. The energy loss as a result of an electron accelerated into the collector may exceed the energy delivered to the field by an equal but properly phased electron. Therein lies a key in improving the efficiency of the klystron: recover a portion of this wasted energy.

Klystrons are cooled by air or liquid for powers up to 5kW. Tubes operating in excess of 5kW are usually water or vapor cooled.

6.3.3.1 Operating Principles

A high-velocity electron beam emitted from the cathode passes through the anode and into the RF interaction region, as shown in Figure 6.15 [3]. An external magnetic field is employed to prevent the beam from spreading as

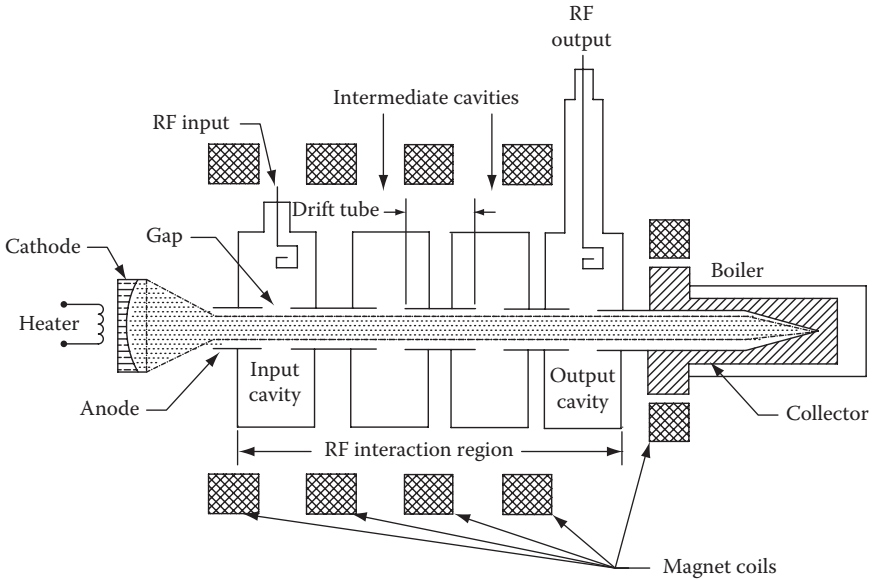


FIGURE 6.15
Principal elements of a klystron.

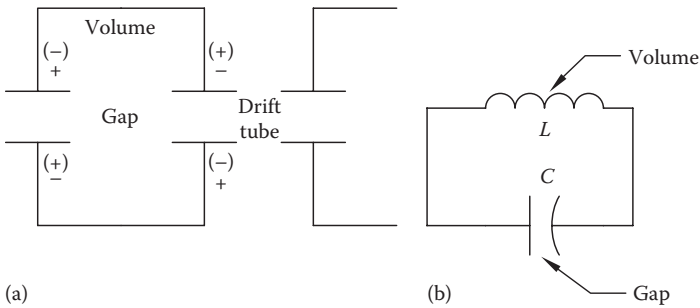


FIGURE 6.16
Klystron cavity: (a) physical arrangement and (b) equivalent circuit.

it passes through the tube. At the other end of the device, the electron beam impinges on the collector electrode, which dissipates the beam energy and returns the electron current to the beam power supply.

The RF interaction region, where amplification occurs, contains resonant cavities and field-free drift spaces. The first resonant cavity encountered by an electron in the beam (the input cavity) is excited by the microwave signal to be amplified, and an alternating voltage of the signal frequency is developed across the gap. This action can be best explained by drawing an analogy between a resonant cavity and a conventional parallel resonant LC circuit (see Figure 6.16). The cavity gap corresponds to the capacitor, and the volume of the cavity to the inductor. If the cavity is of the correct physical

dimensions, it will resonate at the desired microwave frequency. At resonance, opposite sides of the gap become alternately positive and negative at a frequency equal to the microwave input signal frequency.

An electron passing through the gap when the voltage across the gap is zero continues with unchanged velocity along the drift tube toward the next cavity gap; this electron is the *reference electron*. An electron passing through the same gap slightly later is accelerated by the positive field at the gap. This electron speeds up and tends to overtake the reference electron ahead of it in the drift tube. However, an electron that passes through the gap slightly ahead of the reference electron encounters a negative field and is slowed down. This electron tends to fall back toward the following reference electron. As a result of passing the alternating field of the input-cavity gap, the electrons gradually bunch together as they travel down the drift tube, as illustrated in Figure 6.17. Because electrons approach the input-cavity gap with equal velocities and emerge with different velocities, which are a function of the microwave signal, the electron beam is said to be velocity-modulated. As the electrons travel down the drift tube, bunching develops, and the density of electrons passing a given point varies cyclically with time. This bunching is identical in nature to the action in a two-cavity klystron, discussed previously.

The modulation component of the beam current induces current in each of the following cavities. Because each cavity is tuned near resonance, the resulting increase in field at each gap produces successively better-defined electron bunches and, consequently, amplification of the input signal. The RF energy produced in this interaction with the beam is extracted from the

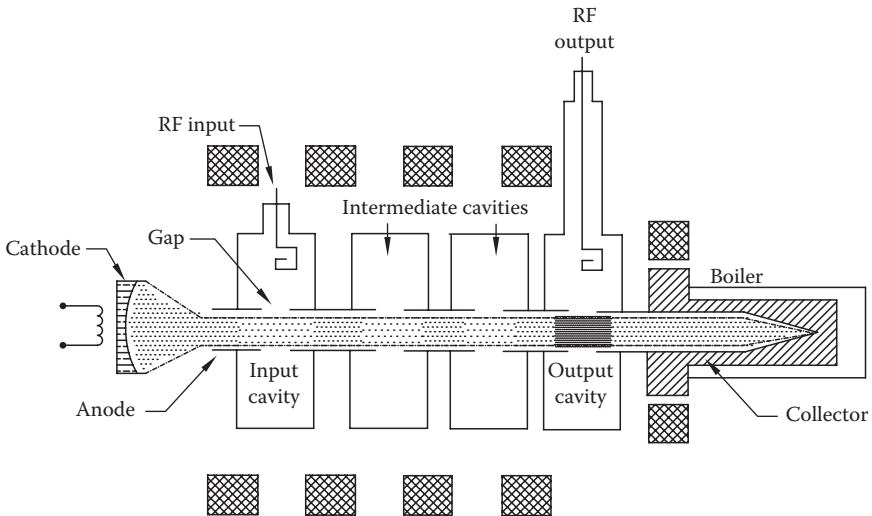


FIGURE 6.17
Bunching effect of a multicavity klystron.

beam and fed into a coaxial or waveguide transmission line by means of a coupling loop in the output cavity. The dc beam input power not converted to RF energy is dissipated in the collector.

6.3.3.2 Emission

The electron gun section of a klystron, shown in Figure 6.18, consists of the following elements [3]:

- A heater
- An emitter
- A beam-forming focusing electrode
- A modulating anode

When the emitter temperature is raised to the proper value by the heater, electrons are released from the emitter surface. The electrons are accelerated toward the modulating anode, which is at a positive potential with respect to the emitter. As the electrons travel between the emitter and the modulating anode, they are formed into a beam by the lens action of the focusing electrode and modulating anode. Figure 6.19 shows how this lens is formed.

All cathodes have a specific optimum range of operating temperature. The temperature of the cathode must be sufficiently high to prevent variations in heater power from affecting the electron emission current (beam current) in

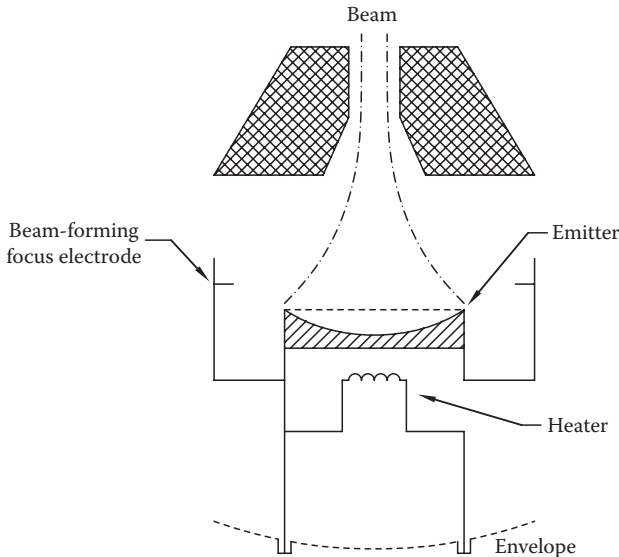


FIGURE 6.18

Diode section of a klystron electron gun.

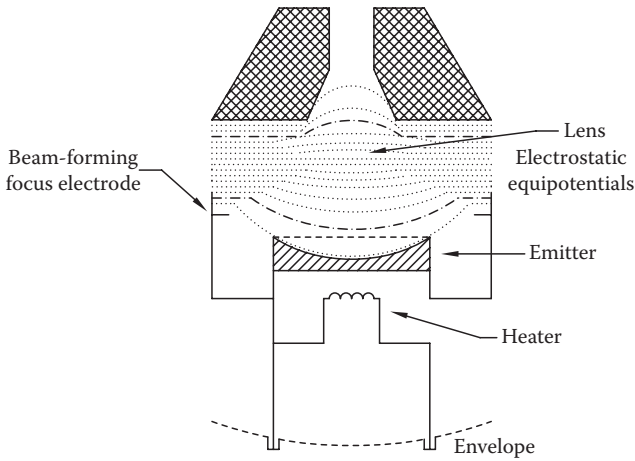


FIGURE 6.19
Beam forming in the diode section of a klystron electron gun.

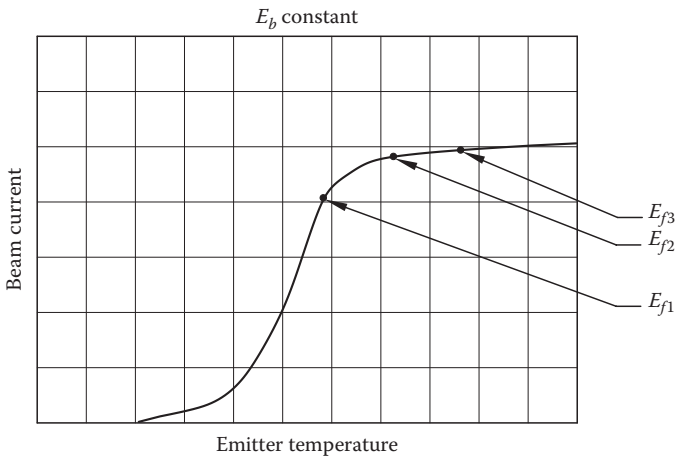


FIGURE 6.20
Klystron beam-current variations as a function of emitter temperature.

the klystron. However, the temperature of the emitting surface must not be higher than necessary because excessive temperature can shorten emission life. Figure 6.20 shows beam current (emission current) as a function of the emitter temperature, which varies directly with heater power. When the heater voltage (E_{f1}) is too low, the emitter will not be hot enough to produce the desired beam current. In addition, even small variations in heater voltage will change the beam current significantly. With the proper heater voltage (E_{f2}), constant beam current will be maintained even with minor variations in heater voltage. The same is true for a higher heater voltage value (E_{f3}), but in this case the emitter temperature is greater than that needed for the desired beam current.

Reduced tube life will result. The correct value of heater voltage and/or heater current is included in the data shipped with each klystron.

6.3.3.3 Modulating Anode

Because the modulating anode is electrically isolated from both the cathode and klystron body (the RF structure, between pole pieces) the voltage applied to it provides a convenient means for controlling beam current independently of the beam voltage applied between the cathode and body, as shown in Figure 6.21. When the cathode is operated in the space-charge-limited region, E_{f2} and E_{f3} of Figure 6.20, the emission current will be a specific function of the applied voltage:

$$I_b = kE^{3/2} \quad (6.2)$$

where

I_b is the beam current in A
 E is the beam potential in V

The constant k is a function of the geometry of the cathode–anode structure, and is termed *perveance*. Because the modulating anode is physically positioned between the RF structure (body) and the cathode, even if the full beam voltage is maintained between cathode and body, the actual beam current into the tube may be reduced at will by biasing the modulating anode to any voltage between cathode and body. Figure 6.22 shows the relationship between beam current and voltage described in the previous equation.

6.3.3.4 Magnetic Field

Beam focusing is used in multielement klystrons to keep the electron beam uniformly small. Focusing may be accomplished by one or more electrostatic *lenses* or external magnetic fields placed parallel to the beam.

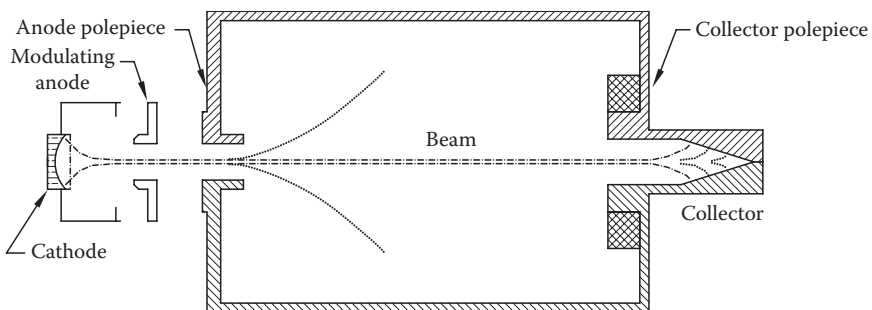


FIGURE 6.21
 Modulating-anode electrode in a multicavity klystron.

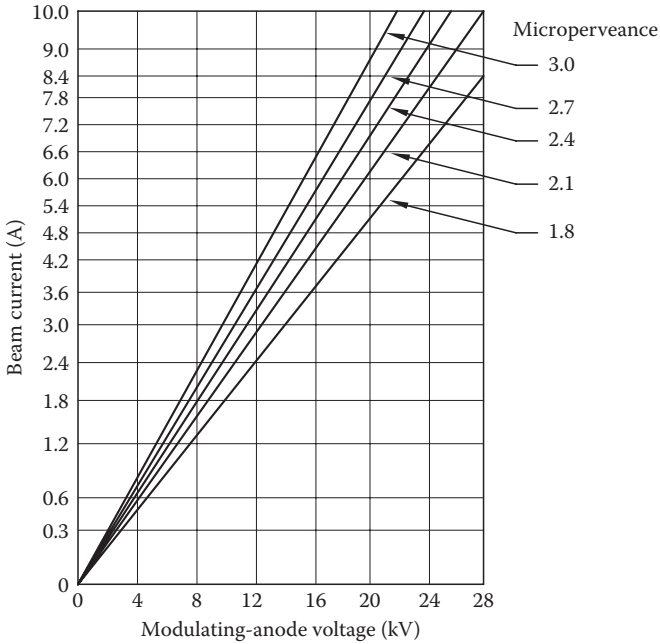


FIGURE 6.22 Beam-current variation as a function of modulating-anode voltage.

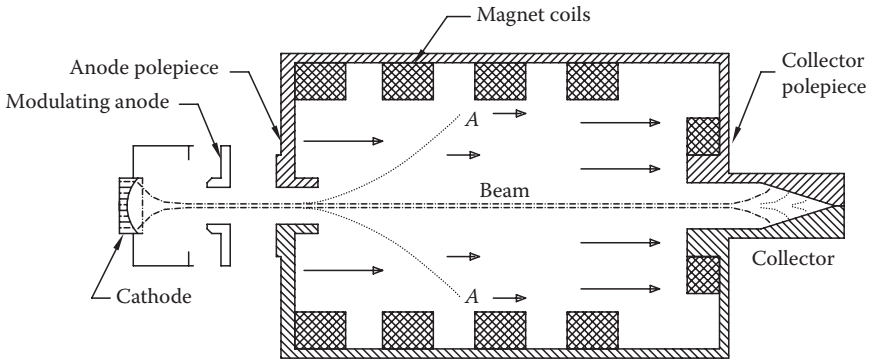


FIGURE 6.23 Effect of a magnetic field on the electron beam of a klystron.

Magnetic focusing is the most common method. Permanent magnets are used at operating powers of about 5 kW or less. Electromagnets are used at higher powers.

Electromagnetic coils typically are placed around the klystron to develop the magnetic field along the axis of the RF circuit. This field controls the size of the electron beam and keeps it aligned with the drift tubes. Figure 6.23 illustrates the beam-forming portion at the cathode end of the klystron and

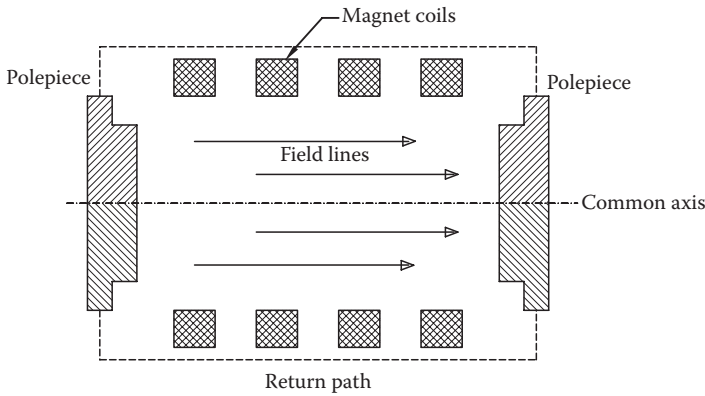


FIGURE 6.24
Field pattern of a klystron electromagnet.

RF section, where the magnetic field is developed between two cylindrical disks (pole pieces). The electron beam in this illustration is shown traveling two paths. One path shows the beam spreading out to points *A*; the other path shows the beam confined by the magnetic field to a constant size throughout the distance between pole pieces. The beam spreads toward points *A* when the magnetic field is inadequate.

Figure 6.24 shows the magnetic field pattern of a typical solenoid used for klystrons. When direct current passes through the magnetic coils, a magnetic field is generated along the axis of the tube. The strength of this field can be controlled by changing the current flow through the magnetic coils. The shape of the field is determined by pole-piece geometry and winding distribution inside the solenoid. Figure 6.25 illustrates the field pattern and the shape of the beam for a properly adjusted field.

Figure 6.26 demonstrates how the beam of a klystron is distorted when an external magnetic material is placed near the RF circuit of the tube.

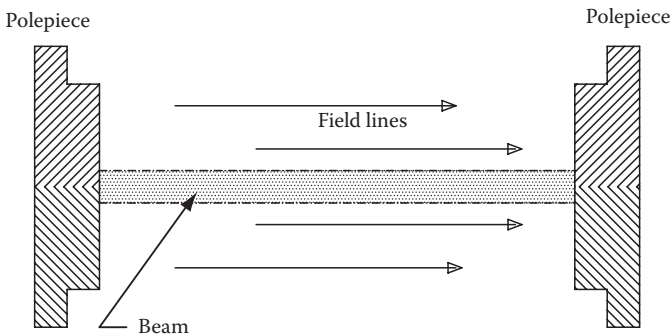


FIGURE 6.25
Field pattern and beam shape in a properly adjusted magnetic field.

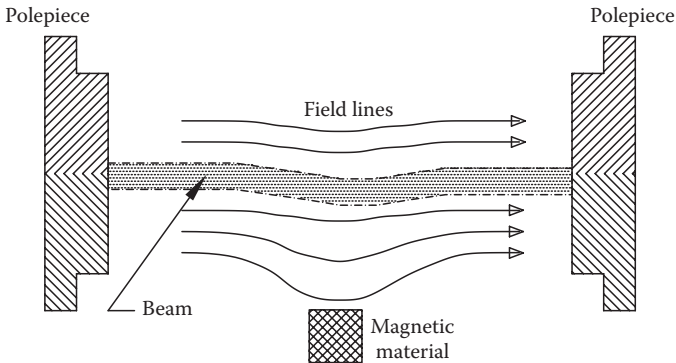


FIGURE 6.26
Distortion of the field pattern and beam shape due to magnetic material in the magnetic field.

Electrons in the beam will follow the bent magnetic field lines and may strike the walls of the drift tube. Klystron damage can result. Magnetic materials such as screwdrivers, wrenches, bolts, and nuts must not be left near the magnetic circuit or near the cathode or collector. Magnetic tools must not be used to tune a klystron.

6.3.3.5 RF Structure

The RF structure of a klystron amplifier consists of tunable resonant circuits (cavities) positioned along the axis of the electron beam. The electron beam traveling through the cavities provides the necessary coupling between each of the RF circuits.

The cavities of a klystron are high-frequency parallel resonant circuits constructed so that they provide an RF voltage across the capacitive component (gap), which interacts with the dc beam. Figure 6.27a illustrates the polarity near the drift-tube tips within a cavity excited by an alternating voltage of microwave signal frequency. Figure 6.27b is the equivalent circuit of a simple cavity. To achieve circuit resonance, the inductive and capacitive reactances of each of the components must be equal. The reactance of each of the components shown in the figure can be measured separately. However, the reactances of the components within a klystron cavity are difficult to determine, because they cannot be measured individually. Therefore, the regions of voltage maxima or minima are used to define each component of a klystron cavity in the following way:

- The capacitance of a cavity is developed across the gap at the drift tubes where the voltage is at maximum.
- The inductance of a cavity is located in the outer volume of the cavity where the voltage is at minimum.

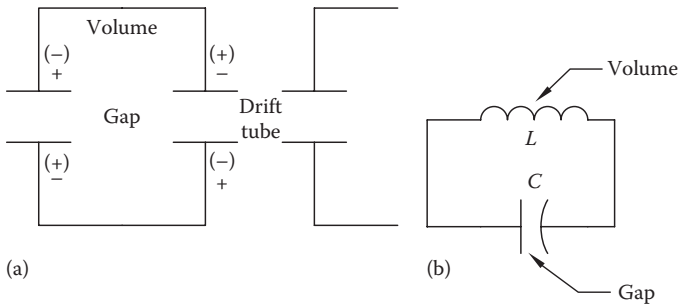


FIGURE 6.27
Klystron cavity: (a) physical element and (b) equivalent electric circuit.

By defining each component of a cavity in these terms, it is easy to visualize changes in the volume as changes in inductance, and changes affecting the gap as changes in capacitance.

The resonant frequency of each cavity can be adjusted to the operating frequency of the RF generator. This can be accomplished in one of two ways:

- Change the inductance by changing the volume of the cavity.
- Change the capacitance of the drift-tube gaps.

Figure 6.28 illustrates how the capacitance of the cavity gap can be modified by attaching a post to a thin-wall diaphragm with a paddle close to the drift-tube gap. In Figure 6.28a, the mechanical configuration of a cavity with this type of tuning is illustrated; Figure 6.28b shows the equivalent circuit capacitance formed between the paddle and the drift tubes at the gap. Moving the paddle away from the drift tubes decreases the gap capacitance and increases the resonant frequency of the cavity.

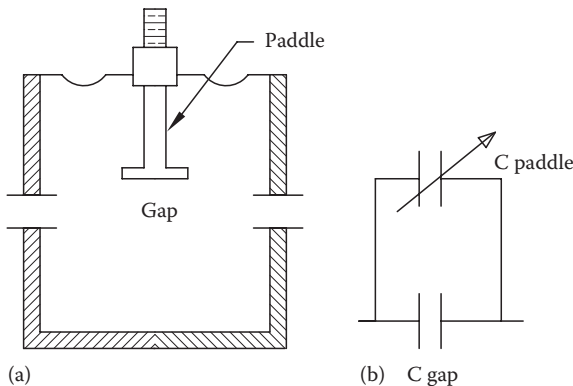


FIGURE 6.28
Capacitance-tuned klystron cavity: (a) physical element and (b) equivalent circuit.

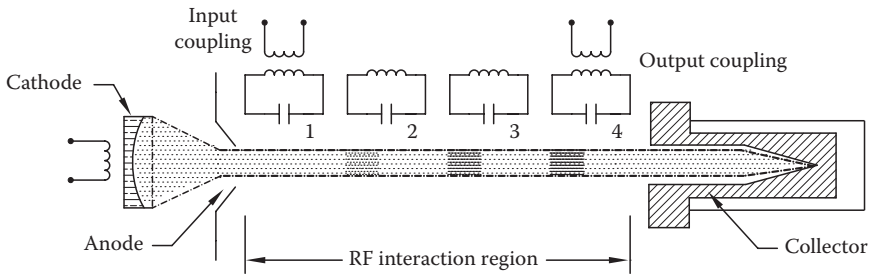


FIGURE 6.29
Schematic equivalent circuit of a four-cavity klystron.

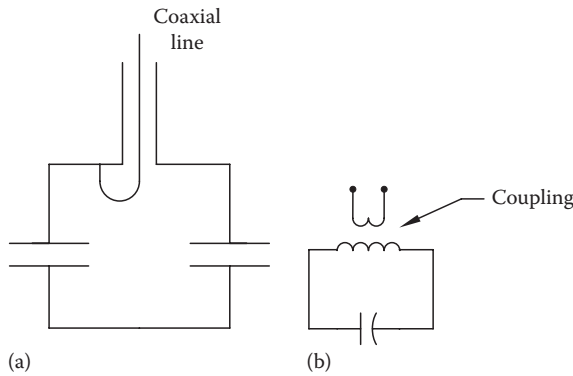


FIGURE 6.30
Klystron loop coupling: (a) mechanical arrangement and (b) equivalent electric circuit.

Figure 6.29 shows a schematic diagram of the equivalent circuits of a four-cavity klystron. Circuit 1 is the input, and circuit 4 is the output. Figure 6.30a illustrates magnetic-loop coupling, where the RF energy is fed through a coaxial line with its center conductor inserted into the klystron cavity. The end of the center conductor is formed into a loop. This forms a simple one-turn transformer that couples RF energy into or out of the cavity through a coaxial transmission line. Figure 6.30b shows the equivalent circuit. The transformer formed by the loop and cavity is an impedance-matching device between the transmission line and the cavity.

Klystron cavities may be externally loaded to improve their instantaneous bandwidth characteristics. These loads lower the Q of the cavities slightly and thereby increase the bandwidth of the klystron.

6.3.3.6 Phased Electron Operation

By properly phasing the second-harmonic fields of a klystron, a favorable electron density distribution pattern can be established at the output gap. The result is the generation of additional RF energy.

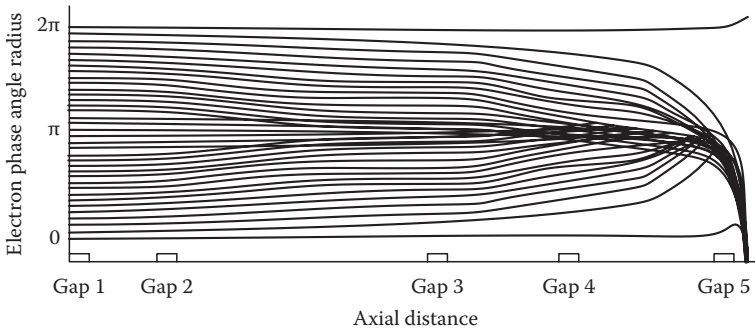


FIGURE 6.31

Plot of the relative phase of the reference electrons as a function of axial distance in a high-efficiency klystron.

A phase-space diagram for a high-power klystron is shown in Figure 6.31. The curves represent a plot of the relative phase of the reference electrons as a function of axial distance along the tube. Electrons having negative slope have been decelerated. Electrons having positive slope have been accelerated with respect to a nonaccelerated electron parallel to the axis. The diagram shows how the electrons are nicely grouped at the output cavity gap while the interbunch regions are relatively free of electrons.

This interaction can be viewed another way, as shown in Figure 6.32, which plots the normalized RF beam currents as a function of distance along the tube. The curves show that the fundamental component of the plasma wave has a negative slope at the third gap. This normally would be a poor condition, but because of the drift of the interbunch electrons, the fundamental current peaks at nearly 1.8 times the dc beam current. The theoretical limit for

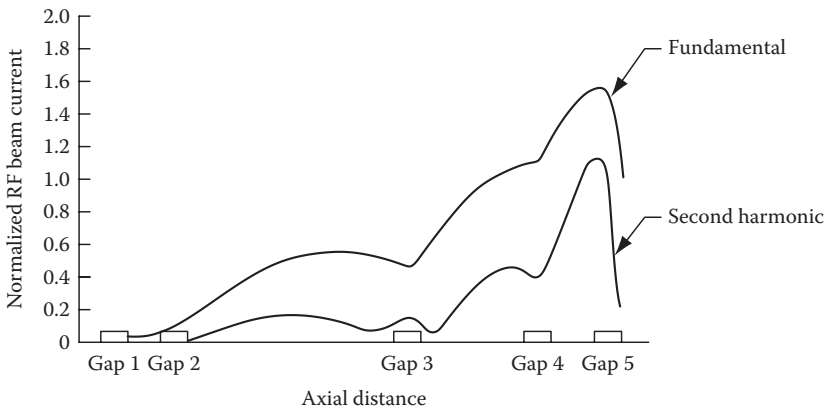


FIGURE 6.32

Plot of the normalized RF beam currents as a function of distance along the length of a high-efficiency klystron.

perfect bunching in a delta function is 2. The second harmonic of the plasma wave also peaks at the output gap, which adds to the conversion efficiency.

6.3.3.7 Types of Devices

Klystrons can be classified according to the following basic parameters:

- Power operating level. Klystrons are available ranging from a few hundred watts to more than 10MW.
- Operating frequency. Klystrons typically are used over the frequency range of 300MHz–40GHz.
- Number of cavities. The number of resonant cavities may range from one to five or more. Furthermore, the cavities may be *integral* or *external* to the vacuum envelope of the device.

The klystron is a true linear amplifier from zero signal level up to 2–3 dB below saturated output. Figure 6.33 shows a typical transfer characteristic for a klystron. RF modulation is applied to the input drive signal. Amplitude modulation typically is limited to the linear portion of the gain transfer characteristic (class A operation). The result is low efficiency, because the beam power is always on. For applications requiring frequency modulation, the drive power is set for saturated output.

Pulse modulation of the klystron may be obtained by applying a negative rectangular voltage, instead of a dc voltage, to the cathode. The RF drive, set to a saturation value, usually is pulsed on for a slightly shorter time than the beam pulse.

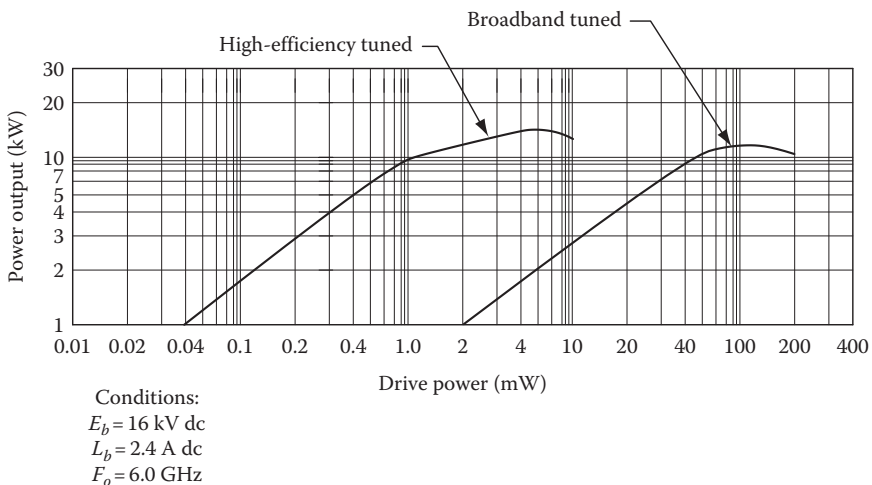


FIGURE 6.33 Typical gain, output power, and drive requirements for a klystron.

TABLE 6.1

Typical Operating Parameters for an Integral-Cavity Klystron for UHF-TV Service

Parameter	Typical Value
Operating frequency (visual)	519 MHz
Output power, peak-of-sync	64 kW
Drive power, peak-of-sync	15 W
Gain, peak-of-sync	36 dB
Efficiency, saturated	55%
Bandwidth, -1 dB	6 MHz
Beam voltage	24.5 kV
Beam current	4.8 A
Body current	10 mA dc
Modulating-anode voltage	17 kV dc
Modulating-anode current	0.5 mA dc
Focusing current	30 A dc
Load VSWR	>1.1:1
Collector temperature	130°C

Source: Courtesy of Varian, Palo Alto, CA.

Because of the high power levels typically used at UHF frequencies, device efficiency is a critical parameter. Klystrons usually are rated in terms of *saturated efficiency*, determined by dividing the saturated RF output power by the dc input power.

Table 6.1 lists typical operating parameters for an integral-cavity klystron.

6.3.4 Beam Pulsing

Beam pulsing is a common method of improving the efficiency of a broadband linear klystron amplifier. Depending on the transmitted waveform, efficiency may be boosted by 25% or more. This technique typically was used in analog UHF-TV transmitters to reduce visual klystron beam dissipation during video portions of the transmission. *Sync pulsing*, as the technique is commonly known, is accomplished by changing the operating point of the tube during the synchronizing interval, when peak power is required, and returning it to a linear transfer characteristic during the video portion of the transmission.

This control is accomplished through the application of a voltage to an electrode placed near the cathode of the klystron. Biasing toward cathode potential increases the beam current, and biasing toward ground (collector potential) decreases beam current.

In the analog composite TV waveform, video information occupies 75% of the amplitude, and sync occupies the remaining 25%. The *tip of sync* represents the peak power of the transmitted waveform. Black (the *blanking level*) represents 56% of the peak power. If the blanking level could be made

to represent 100% modulation and the sync pulsed in, as in a radar system, efficiency would be increased significantly. Unfortunately, the color-burst signal extends 50% into the sync region, and any attempt to completely pulse the sync component would distort the color-burst reference waveform. Sync pulsing is, therefore, limited to 12.5% above black to protect color burst. Two common implementations of beam pulsing have been used:

- Modulating-anode (mod-anode) pulsing
- Annular beam control electrode (ACE) pulsing

The mod-anode system utilizes an additional electrode after the cathode to control beam power. The ACE-type tube operates on a similar principle, but the annular ring is placed close to the cathode so that the ring encloses the electron beam. Because of the physical design, pulsing with the ACE-type gun is accomplished at a much lower voltage than with a mod-anode device. The ACE element, in effect, grid-modulates the beam.

In theory, the amount of beam-current reduction achievable and the resulting efficiency improvement are independent of whether mod-anode or ACE-type pulsing are used. In practice, however, differences are noted. With existing mod-anode pulsers, an efficiency improvement of about 19% in beam current over nonpulsed operation may be achieved. A beam reduction of 30%–35% may be achieved through use of an ACE-equipped tube. The effect of ACE voltage on beam current is shown in Figure 6.34 for an external-cavity klystron.

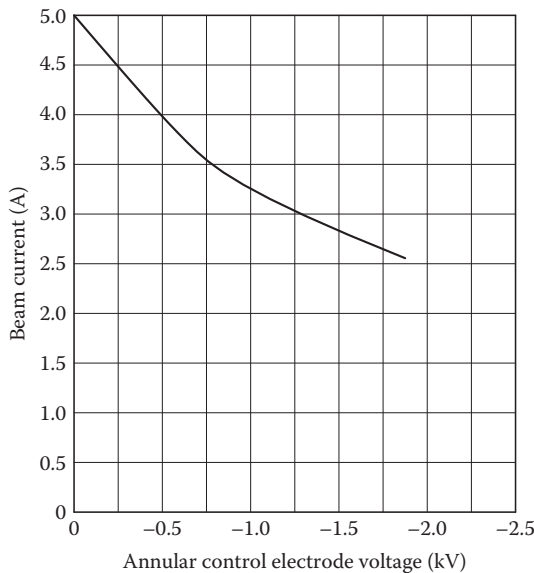


FIGURE 6.34 Klystron beam current as a function of annular control electrode voltage.

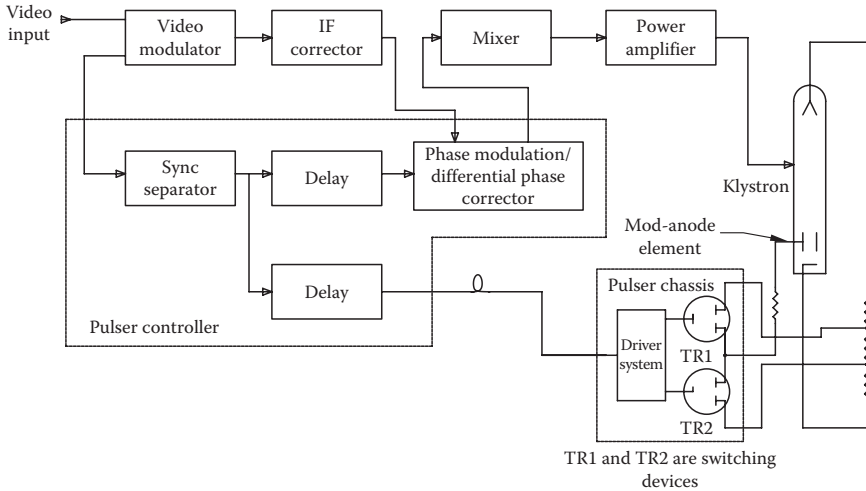


FIGURE 6.35
Block diagram of a switching mod-anode pulser.

The ACE electrode, positioned in the gun assembly, is driven by a negative-directed narrowband video signal of a few hundred volts peak. The annular ring varies the beam density through a *pinching action* that effectively reduces the cross-sectional area of the stream of electrons emitted by the cathode. The klystron thus operates in a quasi-class-AB condition rather than the normal class A (for linear analog TV service).

A peak-of-sync FOM for an integral-cavity UHF-TV klystron without ACE or equivalent control is 0.67–0.68. Through the use of an ACE-type tube, the FOM may be increased to 0.80. Similar improvements in efficiency can be realized for external-cavity klystrons. A typical switching-type mod-anode pulser is shown in Figure 6.35.

Pulsing is not without its drawbacks, however. The greater the pulsing, the greater the precorrection required from the modulator. Precorrection is needed to compensate for nonlinearities of the klystron transfer characteristic during the video modulation period. Level-dependent RF phase precorrection also may be required. Switching between different klystron characteristics produces phase modulation of the visual carrier. If not corrected, this phase modulation may result in intercarrier “buzz” in the received audio. These and other considerations limit the degree of pulsing that may be achieved on a reliable basis.

6.3.5 Integral vs. External Cavity

In an integral-cavity klystron, the resonant cavities are located within the vacuum envelope of the tube. In an external-cavity klystron, the cavities are located outside the envelope in a mechanical assembly that wraps around the drift tube.

Fundamentally, klystron theory applies equally to integral- and external-cavity tubes. In both cases, a velocity-modulated electron beam interacts with multiple resonant cavities to provide an amplified output signal. The resonant cavity *interaction gap* and drift length requirements for optimum performance—including conversion efficiency—are independent of whether the tuning mechanism is inside or outside the vacuum envelope. High-efficiency integral- and external-cavity klystrons have been designed to provide maximum conversion performance consistent with signal bandwidth requirements. The saturated conversion efficiency is essentially identical for integral- or external-cavity klystrons with equal numbers of resonant cavities.

6.3.5.1 Number of Cavities

The relative benefits of integral-cavity vs. external-cavity klystrons can be debated at length. Discussion typically centers on operating efficiency, life expectancy, and replacement device cost. In a comparison of the efficiency of integral-cavity vs. external-cavity klystrons for UHF-TV applications, however, the question really boils down to the number of cavities (four or five) used in the device. Four-cavity external tubes are standard. It is usually not practical to produce high-power five-cavity external tubes for mechanical and electrical (voltage standoff) reasons.

Integral devices permit the addition of a fifth cavity to the design for two primary reasons. First, the device itself is mechanically more robust. The addition of a fifth cavity to an external klystron increases the length and weight of the device. Because of the size of the drift tube in an external device designed for UHF-TV frequencies, the additional weight is difficult to support. Second, because the cavities are enclosed in a vacuum in the integral design, voltage standoff problems are greatly reduced.

The tuning mechanism of an integral-cavity klystron is enclosed in a rigid steel shell. In the external-cavity unit, the tuning mechanism is in air. Within the region of the tuning mechanism, RF fields can reach high levels, especially at the high end of the UHF-TV band. The highest energy field occurs in the area of the last cavity. Because of the high dielectric properties of a vacuum, dielectric breakdown is less of a problem in the integral design, which places the tuning mechanism within a vacuum envelope. This situation is of particular concern at the high end of the UHF band, where the spacing of tuning elements is closer.

The physics of the integral- and external-cavity klystrons are essentially the same. In a comparison of integral and external units of like design—that is, with the same number of cavities—performance should be identical. The two units follow the same laws of physics and use basically the same components up to the beam stick. It is when manufacturers take advantage of the relative merits of each design that differences in performance are realized.

6.3.5.2 Efficiency

For the sake of comparison, the data presented in this section will assume that

- The power level is 60 kW
- All integral-cavity devices utilize five cavities
- All external-cavity devices utilize four cavities

The five-cavity integral *S-tuned* klystron is inherently at least 20% more efficient than a four-cavity tube. (*S-tuning* refers to the method of stagger tuning the cavities.) The five-cavity klystron generally is specified by the manufacturer for a minimum efficiency of 52% (saturated efficiency). Typical efficiency is 55%. Four-cavity devices are characteristically specified at 42% minimum and 45% typical.

This efficiency advantage is possible because the fifth cavity of the integral design permits tuning patterns that allow maximum transfer of RF energy while maintaining adequate bandpass response. Tighter bunching of electrons in the beam stick, a function of the number of cavities, also contributes to the higher efficiency operation. In actuality, the fifth cavity allows design engineers to trade gain for efficiency. Still, the five-cavity tube has significantly more gain than the four-cavity device. The five-cavity unit, therefore, requires less drive, which simplifies the driving circuit. A five-cavity klystron requires approximately 25 W of drive power, while a four-cavity tube needs as much as 90 W for the same power output.

Under pulsed operation, approximately the same reduction in beam current is realized with both integral- and external-cavity klystrons. There is, fundamentally, no reason that one type of klystron should pulse differently than the other.

6.3.5.3 Performance Trade-Offs

It is the designer's choice whether to build a transmitter with the klystron cavities located inside or outside the vacuum envelope. There are benefits and drawbacks to each approach.

When the cavity resonators are a part of the tube, the device becomes more complicated and more expensive. However, the power generating system is all together in one package, which simplifies installation significantly.

When the resonator is separate from the tube, as in an external-cavity device, it can be made with more *compliance* (greater room for adjustment). Consequently, a single device may be used over a wider range of operating frequencies. In terms of UHF-TV, a single external-cavity device may be tuned for operation over the entire UHF-TV band. This feature is not possible if the resonant cavities are built into the device. To cover the entire UHF-TV band, three integral-cavity tubes are required. The operational divisions are as follows:

- Channels 14–29 (470–566 MHz)
- Channels 30–51 (566–698 MHz)
- Channels 52–69 (698–746 MHz)

This practical limitation to integral-cavity klystron construction may be a drawback for some facilities. For example, it is not uncommon for group operations to share one or more spare klystrons. If the facilities have operating frequencies outside the limits of a single integral device, it may be necessary to purchase more than one spare. Also, when the cavities are external, the resonators are in air and can be accessed to permit fine adjustments of the tuning stages for peak efficiency.

The advantages of tube changing are significant with an integral device. Typical tube change time for an integral klystron is 1 h, as opposed to 4–6 h for an external device. The level of experience of the technician is also more critical when an external-type device is being changed. Tuning procedures must be carefully followed by maintenance personnel to avoid premature device failure.

6.3.6 MSDC Klystron

Developmental work on the MSDC klystron began in the mid-1980s. The project* produced a working tube capable of efficiency in UHF service that had been impossible with previous klystron-based technology.

The MSDC device may be used in a number of varied applications. NASA originally became involved in the project as a way to improve the efficiency of satellite transmitters. With limited power available onboard a space vehicle, efficient operation is critically important. Such transmitters traditionally operate in a linear, inefficient mode. UHF-TV broadcasters were interested in the MSDC because it promised to reduce the huge operating costs associated with high-power operation.

Experimentation with depressed collector klystrons dates back to at least the early 1960s. Early products offered a moderate improvement in efficiency, but at the price of greater mechanical and electrical complexity. The MSDC design, although mechanically complex, offers a significant gain in efficiency.

6.3.6.1 Theory of Operation

MSDC tubes have been built around both integral-cavity and external-cavity klystrons. The devices are essentially identical to a standard klystron, except for the collector assembly. Mathematical models provided researchers with detailed information on the interactions of electrons in the collector region. Computer modeling also provided the basis for optimization of a

* A joint effort of the National Aeronautics and Space Administration (NASA), several UHF-TV transmitter manufacturers, Varian Associates, and other concerns.

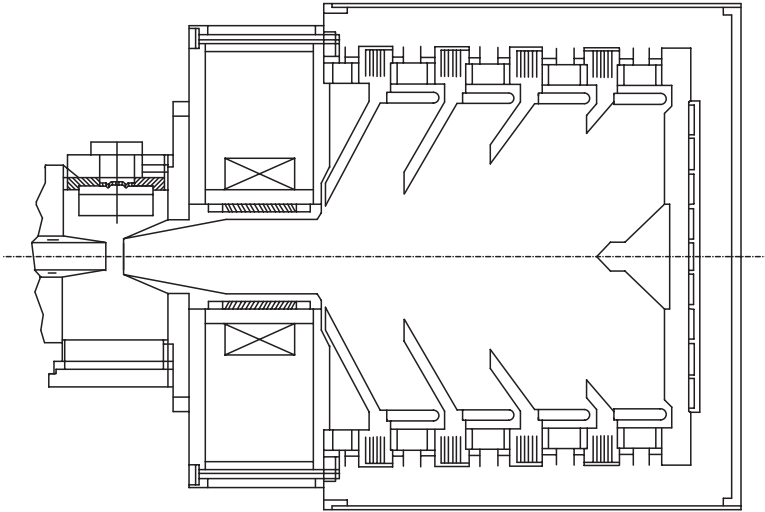


FIGURE 6.36

Mechanical design of the MSDC assembly. Note the “V” shape of the four-element system.

beam-reconditioning scheme incorporated into the device. Beam reconditioning is achieved by including a *transition region* between the RF interaction circuit and the collector under the influence of a magnetic field.

From the electrical standpoint, the more stages of a MSDC klystron, the better. The trade-off, predictably, is increased complexity and, therefore, increased cost for the product. There is also a point of diminishing returns that is reached as additional stages are added to the depressed collector system. A four-stage device was chosen for TV service because of these factors. As more stages are added (beyond four), the resulting improvement in efficiency is proportionally smaller.

Figure 6.36 shows the mechanical configuration of the four-stage MSDC klystron. Note the “V” shape that was found, through computer modeling, to provide the best *capture* performance, minimizing electron feedback. A partially assembled collector assembly is shown in Figure 6.37.

Because the MSDC device is identical to a conventional klystron except for the collector, efficiency improvement techniques used for klystrons can be incorporated into the MSDC. ACE-type pulsing commonly is used to improve efficiency in various applications. Figure 6.38 illustrates the effects of ACE voltage on beam current.

6.3.6.2 Electron Trajectories

The dispersion of electrons in the multistage collector is the key element in recovering power from the beam and returning it to the power supply. This is the mechanism that permits greater operating efficiency from the MSDC device.

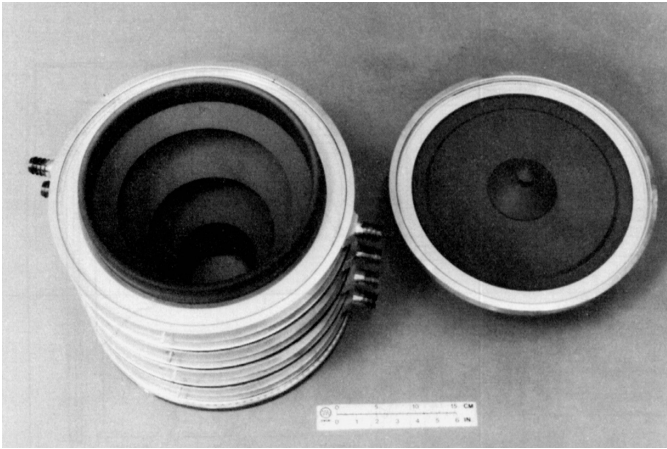


FIGURE 6.37
Partially assembled MSDC collector. (Courtesy of Varian, Palo alto, CA.)

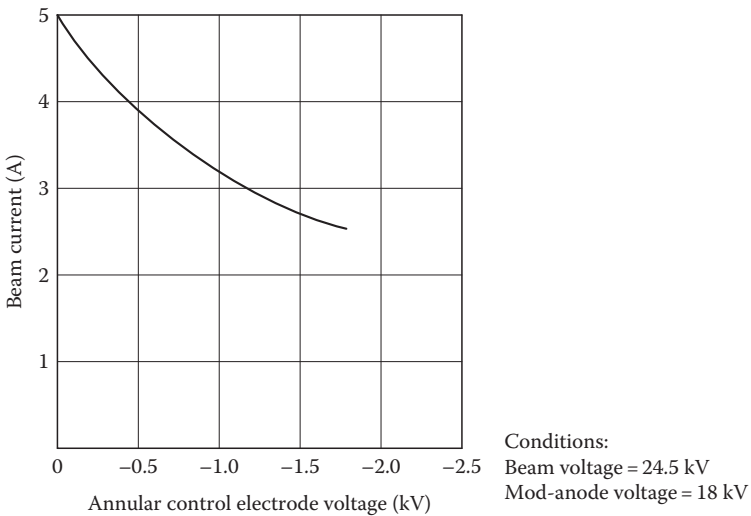
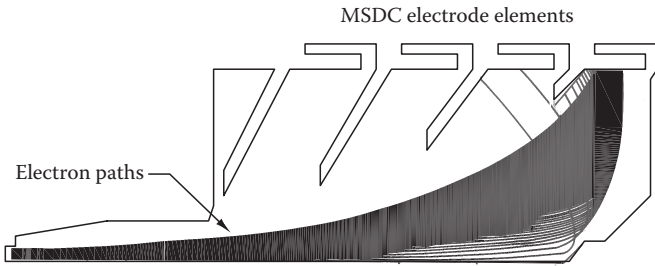


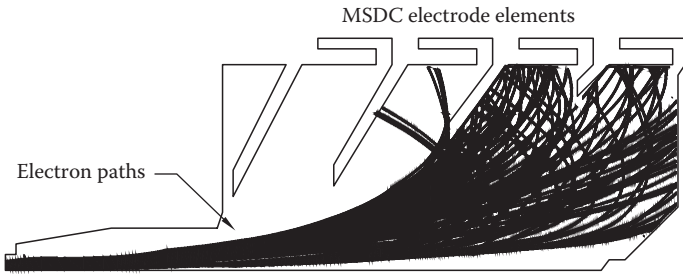
FIGURE 6.38
Beam current as a function of annular control electrode voltage for an MSDC klystron.

Figure 6.39 illustrates the dispersion of electrons in the collector region during a carrier-only operating mode. Note that there is little dispersion of electrons between stages of the MSDC. Most are attracted to electrode 4, the element at the lowest potential (6.125 kV), referenced to the cathode.

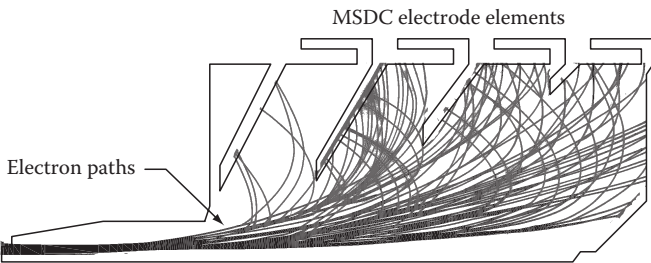
Figure 6.40 shows collector electron trajectories at 25% saturation. The electrons exhibit predictable dispersion characteristics during the application of modulation, which varies the velocity of the electrons.

**FIGURE 6.39**

Collector electron trajectories for the carrier-only condition. Note that nearly all electrons travel to the last electrode (4), producing electrode current I_4 .

**FIGURE 6.40**

Collector trajectories at 25% saturation. With the application of modulation, the electrons begin to sort themselves out.

**FIGURE 6.41**

Collector electron trajectories at 50% saturation, approximately the blanking level. The last three electrodes (2, 3, and 4) share electrons in a predictable manner, producing currents I_2 , I_3 , and I_4 .

Figure 6.41 shows electron trajectories at 50% saturation. Note the increased number of electrons attracted to electrodes 2 and 3, the higher-potential electrodes (referenced to the cathode).

Figure 6.42 illustrates electron dispersion at 90% saturation. As the modulation level is increased, more electrons are attracted to the higher-voltage electrodes. The dramatic increase in electron capture by electrode 1, the highest-potential element of the device (at a voltage of 24.5 kV, referenced to the cathode), can be observed.

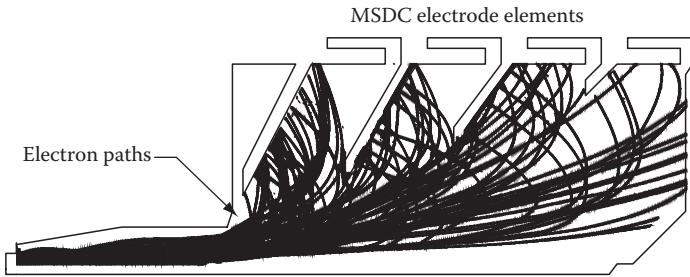


FIGURE 6.42

Collector trajectories at 90% saturation, the sync level. Note the significant increase in the number of electrons attracted to the first electrode, producing I_1 .

The electrons, thus, sort themselves out in a predictable manner. Notice the arc that is present on many electron traces. The electrons penetrate the electrostatic field of the collector and then are pulled back to their respective potentials.

A savings in power is realized because the electrostatic forces set up in the MSDC device slow down the electrons before they contact the copper collector electrode. The heat that would be produced in the collector is, instead, returned to the power supply in the form of electric energy. In theory, peak efficiency would occur if the electrons were slowed down to zero velocity. In practice, however, that is not possible.

Figure 6.43 shows the distribution of collector current as a function of drive power. With no RF drive, essentially all current goes to electrode 4, but as drive is increased, I_4 drops rapidly as collector current is distributed among the other elements. Note that the current to electrode 5 (cathode potential) peaks at about 10% of beam current. This suggests that the *secondary yield* of the collector surfaces is within acceptable limits.

Inserted between the klystron and the collector assembly is a refocusing electromagnet that controls the electron beam as it enters the collector region.

6.3.6.3 Mechanical Construction

The completed MSDC assembly is shown in Figure 6.44 with the collector shield partially removed to allow visibility of the collector elements. The collector of the four-stage MSDC design actually is composed of five elements mounted between ceramic rings for electrical insulation. The fifth electrode is at cathode potential.

Cooling for the MSDC is, not surprisingly, more complicated than for a conventional klystron. The trade-off, however, is that there is less heat to remove because of the higher efficiency of the device. Water cooling is provided on each electrode of the MSDC tube.

Figure 6.45 illustrates the overall mechanical design of an external-cavity version of the MSDC tube, including placement of the device in its cavity bay.

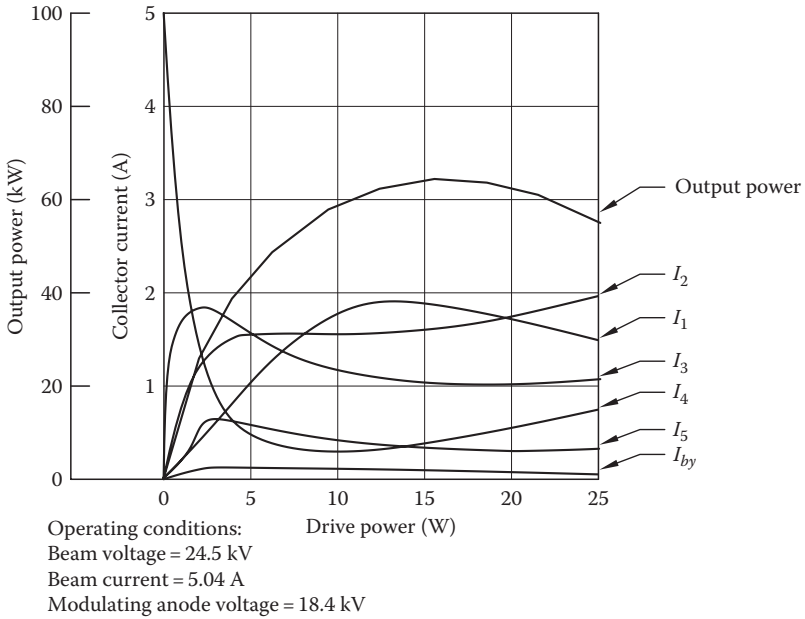


FIGURE 6.43 Distribution of electrode current as a function of drive power. Note the significant drop in I_4 as drive power is increased. I_5 is the electrode at cathode potential.

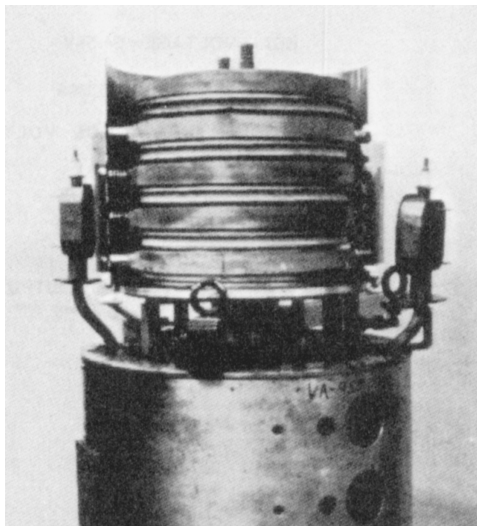


FIGURE 6.44 MSDC collector assembly with the protective shield partially removed. (Courtesy of Varian, Palo Alto, CA.)

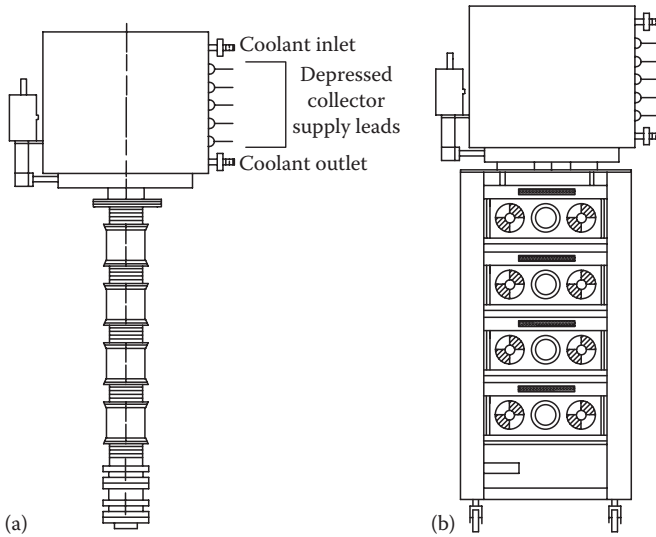


FIGURE 6.45 Mechanical construction of an external-cavity MSDC tube: (a) device and (b) device in cavity assembly.

6.3.6.4 MSDC Power Supply

Design criteria for the collector power supply system provide a mixed bag of requirements. The critical parameter is the degree of regulation between the cathode and anode. The relative differences between the elements of the collector are not, in most applications, significant. Consequently, the bulk of the power supplied to the tube does not need to be well regulated. This is in contrast with conventional klystron operation, in which the entire beam power supply must be regulated. This factor effectively decreases the amount of power that must be regulated to 1%–2% of the dc input, offsetting to some extent the additional cost involved in constructing multiple supplies to facilitate the four-stage MSDC design. Two approaches can be taken to collector power supply design:

- Parallel arrangement, shown in Figure 6.46
- Series arrangement, shown in Figure 6.47

Note that, in both cases, the collector electrodes are stepped at a 6.125 kV potential difference for each element.

6.3.6.5 Device Performance

The efficiency improvement of the MSDC klystron over the conventional klystron is impressive—in the neighborhood of 2×. The bandpass performance of the device is another critical parameter. Figure 6.48 charts power

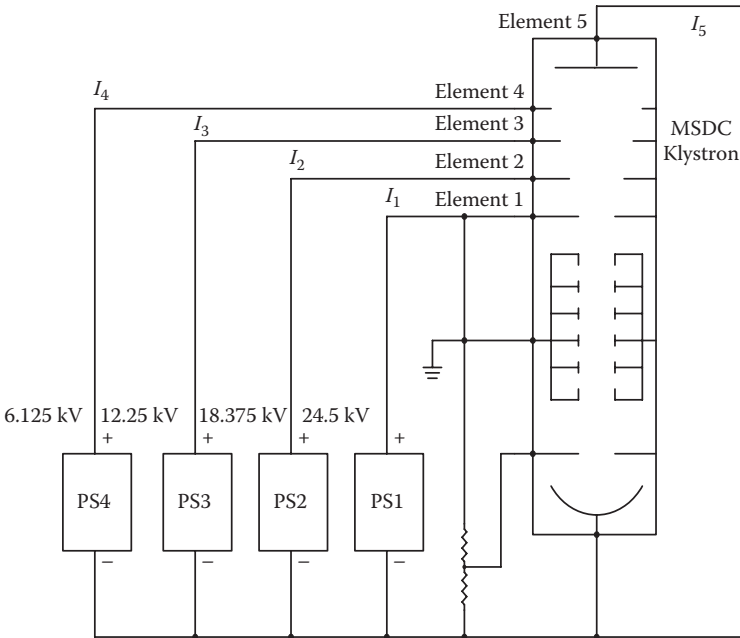


FIGURE 6.46 Parallel configuration of the MSDC power supply. Each supply section has an output voltage that is an integral multiple of 6.125 kV.

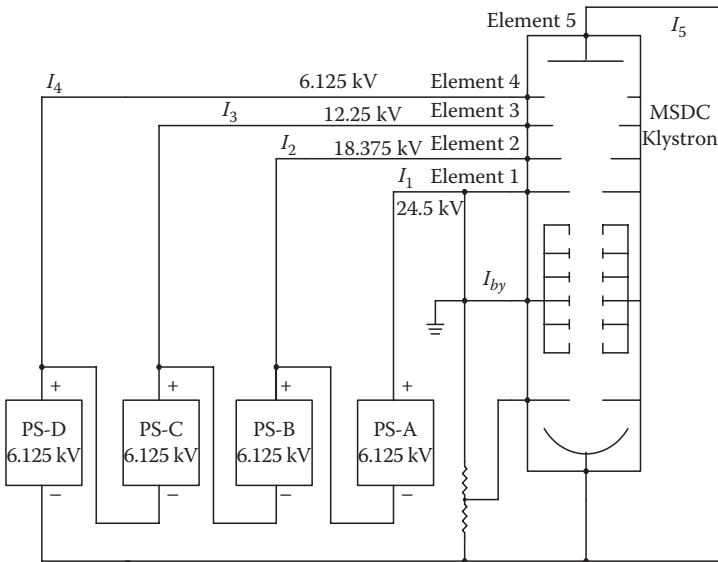


FIGURE 6.47 Series configuration of the MSDC power supply, which uses four identical 6.125 kV power supplies connected in series to achieve the needed voltages.

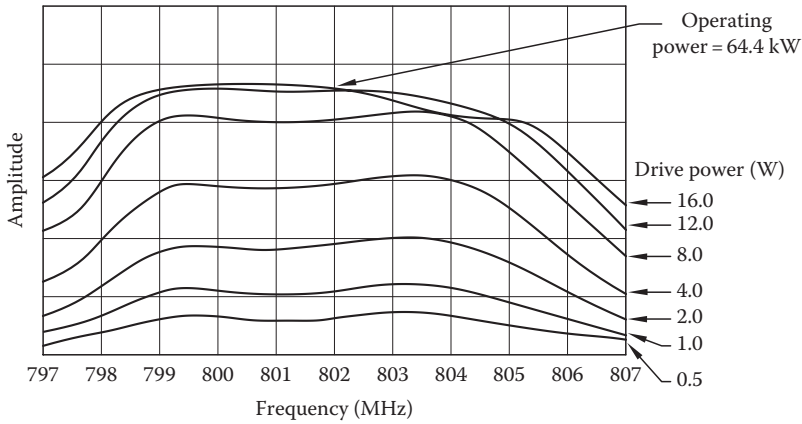


FIGURE 6.48 Device bandwidth as a function of frequency and drive power. Beam voltage is 24.5 kV, and beam current is 5.04 A for an output power of 64 kW. These traces represent the full-power test of the MSDC device.

output as a function of frequency and RF drive at full power. (Drive power is charted from 0.5 to 16 W.) Figure 6.49 charts power output as a function of frequency and RF drive with the tube in a beam-pulsing mode. (Drive power is charted from 0.5 to 32 W.) Note that the traces provide good linearity over a 6 MHz bandwidth. Gain, as a function of frequency and power, is essentially constant and undisturbed.

Table 6.2 lists typical operating parameters for an external-cavity MSDC device.

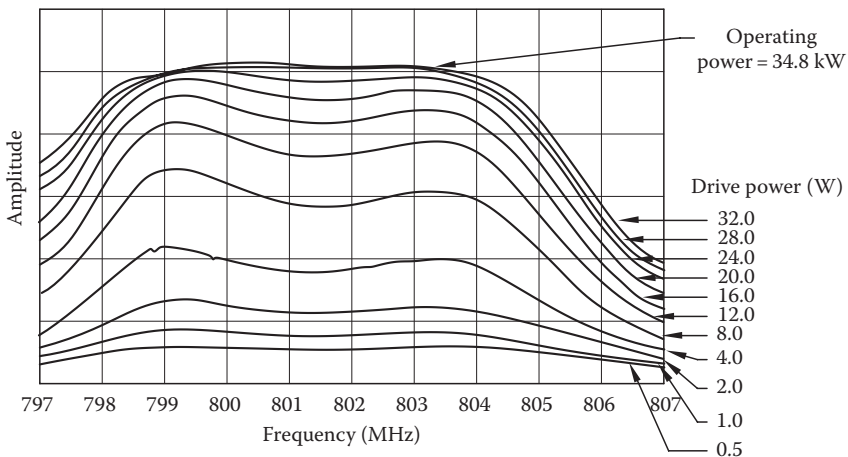


FIGURE 6.49 Device bandwidth in the beam-pulsing mode as a function of frequency and drive power. Beam voltage is 24.5 kV, and beam current is 3.56 A with an output power of 34.8 kW.

TABLE 6.2

Typical Operating Parameters for an External-Cavity
MSDC Klystron for UHF-TV Service

Parameter	Typical Value
Operating frequency (visual)	519 MHz
Output power, peak-of-sync	64 kW
Drive power, peak-of-sync	20 W
Gain, peak-of-sync	35 dB
Figure of merit	1.3 ^a
Bandwidth, -1 dB	6 MHz
Beam voltage	24.5 kV
Beam current	5.3 A
Body current	50 mA dc
Modulating-anode voltage	19.5 kV dc
Modulating-anode current	0.5 mA dc
Focusing current	11 A dc
Ion pump voltage	3.2 kV dc
Load VSWR	>1.1:1
Refocusing coil current	7 A dc
Refocusing coil voltage	10 V dc

Source: Courtesy of Varian, Palo Alto, CA.

^a *Figure of merit* is equal to the quotient of peak-of-sync output power and the average dc beam input power.

6.3.6.6 Applying the MSDC Klystron

Figure 6.50 shows a 60 kW analog TV transmitter incorporating two external-cavity MSDC klystrons, one for the visual and another for the aural. Design of the transmitter is basically identical to a non-MSDC system, with the exception of the power supply and cooling system. The efficiency available from the MSDC makes further device improvements subject to the law of diminishing returns, as mentioned previously. Support equipment begins to consume an increasingly large share of the power budget as the output device efficiency is improved. Tuning of the MSDC klystron is the same as for a conventional klystron, and the same magnetic circuit is typically used.

The power supply arrangement for the example MSDC system is shown in Figure 6.51. A series beam supply was chosen for technical and economic reasons. Although the current to each collector varies widely with instantaneous output level, the total current stays within narrow limits. Using a single transformer for all supplies, therefore, minimizes the size of the iron core required. A 12-pulse rectifier bank provides low ripple and reduces the need for additional filtering. The size and complexity of the rectifier stack are increased little beyond a normal beam supply because the total potential of the four supplies is similar to that of a normal klystron transmitter (24.5–27.5 kV).

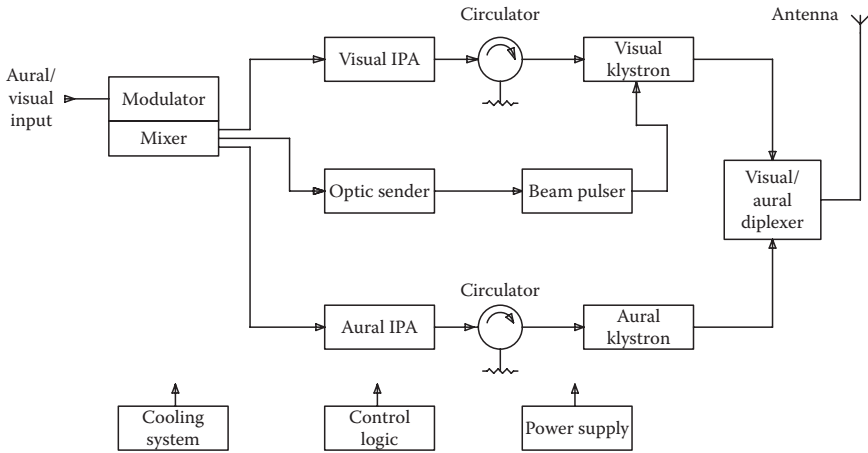


FIGURE 6.50 Simplified block diagram of a 60kW MSDC transmitter. The aural klystron may utilize a conventional or MSDC tube at the discretion of the user.

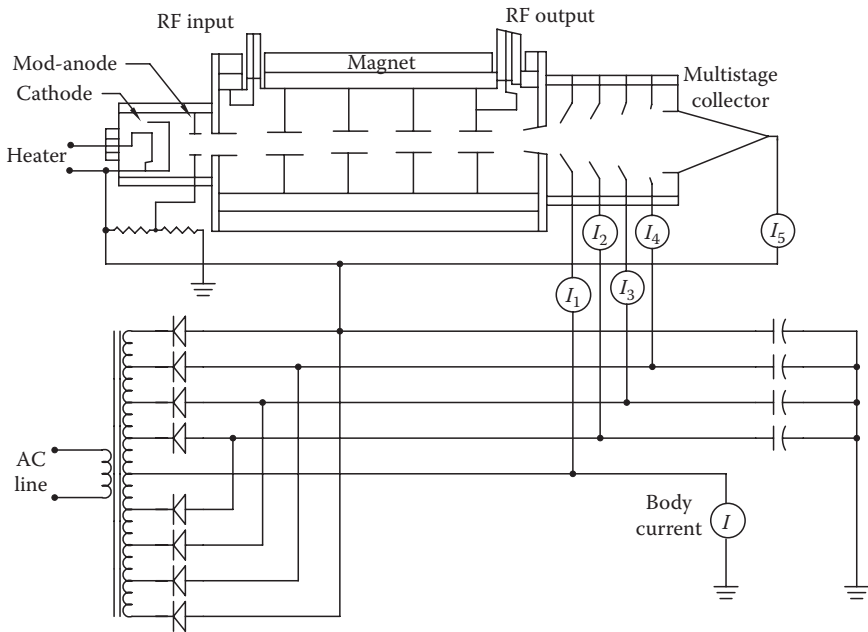


FIGURE 6.51 Simplified schematic diagram of the power supply for an MSDC klystron transmitter. A single high-voltage transformer with multiple taps is used to provide the needed collector voltage potentials.

The collector stages of the MSDC device are water cooled by a single water path that loops through each electrode element. Because high voltage is present on the individual elements, purity of the water is critical to proper operation. A two-stage system is used with a water-to-water plate heat exchanger separating the primary and secondary systems.

6.4 Klystrode/Inductive Output Tube

The Klystrode tube is the result of a development program started in 1980* with UHF-TV in mind as a primary application. The basic concept of the Klystrode dates back to the late 1930s,[†] but it was not until the early 1980s that serious engineering effort was put into the tube to make it a viable product for high-power UHF service. The fundamental advantage of the Klystrode, also known as the IOT, is its ability to operate class B. The result is higher efficiency when compared with a conventional klystron.

6.4.1 Theory of Operation

As its name implies, the Klystrode tube is a hybrid between a klystron and a tetrode. The high reliability and power-handling capability of the klystron is due, in part, to the fact that electron beam dissipation takes place in the collector electrode, quite separate from the RF circuitry. The electron dissipation in a tetrode is at the anode and the screen grid, both of which are an inherent part of the RF circuit and, therefore, must be physically small at UHF frequencies. The tetrode, however, has the advantage that modulation is produced directly at the cathode by a grid so that a long drift space is not required to produce density modulation. The Klystrode/IOT has a similar advantage over the klystron—high efficiency in a small package [4]. The Klystrode tube is a registered trademark of Varian. The IOT is the generic description for the device, and it will be used in this discussion.

The IOT is shown schematically in Figure 6.52. The electron beam is formed at the cathode, density-modulated with the input RF signals by a grid, then accelerated through the anode aperture. In its bunched form, the beam drifts through a field-free region and then interacts with the RF field in the output cavity. Power is extracted from the beam in the same way as in a klystron. The input circuit resembles that of a typical UHF power grid tube. The output circuit and collector resemble those in a klystron.

* Varian/Eimac, San Carlos, CA.

[†] Andrew Haeff (1938).

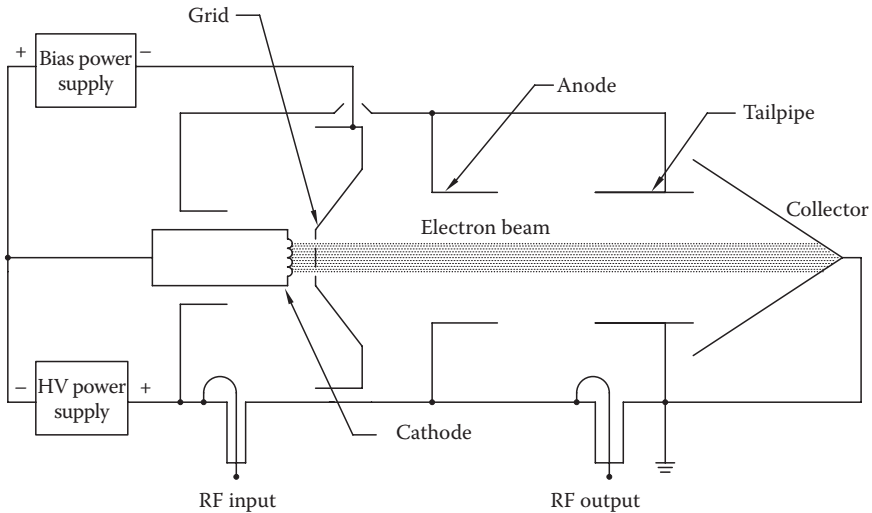


FIGURE 6.52
Simplified schematic diagram of the Klystrode tube.

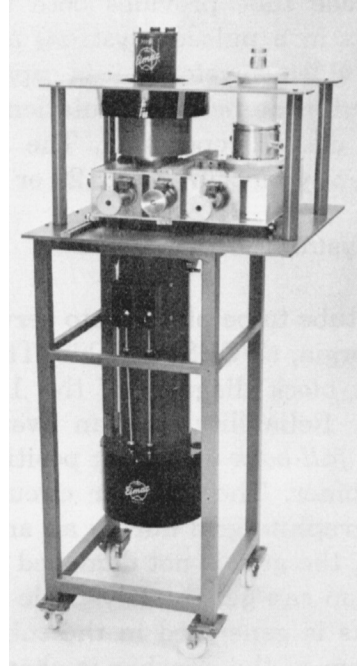
A production version of a 60kW device is shown in Figure 6.53. Double-tuned cavities are used to obtain the required operating bandwidth. The load is coupled at the second cavity, as shown in Figure 6.54. This arrangement has proved to be an attractive way to couple power out of the device because no coupling loop or probe is required in the primary cavity, which can be a problem at the high end of the UHF band.

Because the IOT provides both beam power variation during sync pulses (as in a pulsed klystron) and variation of beam power over the active modulating waveform, it is capable of high efficiency. The device provides full-time beam modulation as a result of its inherent structure and class B operation.

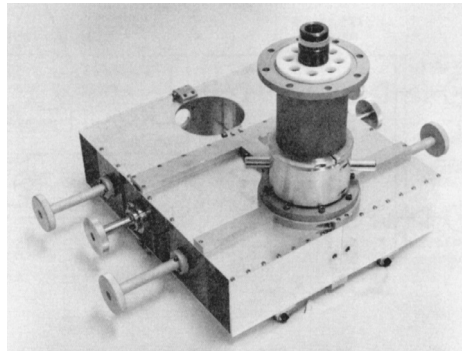
6.4.2 Electron Gun

Many design elements from the klystron have been borrowed for use in the IOT [4]. A barium aluminate cathode is used for the gun, taken from a conventional klystron together with its heater structure. The methods used to support the cathode, heater, and necessary heat shields in the klystron gun have been retained for use in the IOT. Figure 6.55 shows the construction of the IOT gun and output cavity.

In the IOT, the RF input voltage is applied between the cathode and a grid that allows extra electrons to be drawn from the cathode into a low-quietest-current electron beam according to the instantaneous RF voltage appearing between the grid and cathode. The resulting density-modulated beam is then passed into the klystron-like RF output interaction region of the tube.

**FIGURE 6.53**

A 60 kW Klystrode tube mounted in its support stand with the output cavity attached. (Courtesy of Varian, Palo Alto, CA.)

**FIGURE 6.54**

Close-up view of the double-tuned output cavity of a 60 kW Klystrode tube. (Courtesy of Varian, Palo Alto, CA.)

The grid is clamped in place in front of the cathode, supported on a metal cylinder, and isolated from the cathode by a ceramic insulator. A second ceramic insulator supports the complete gridded electron gun at the correct distance from the grounded anode. This ceramic insulator completes the vacuum envelope and holds off the full beam voltage of approximately 30 kV. The IOT gun operates at cathode-to-anode voltages approximately 50% higher than those of conventional klystron guns. Stray capacitance in the grid support structure has been minimized to reduce losses.

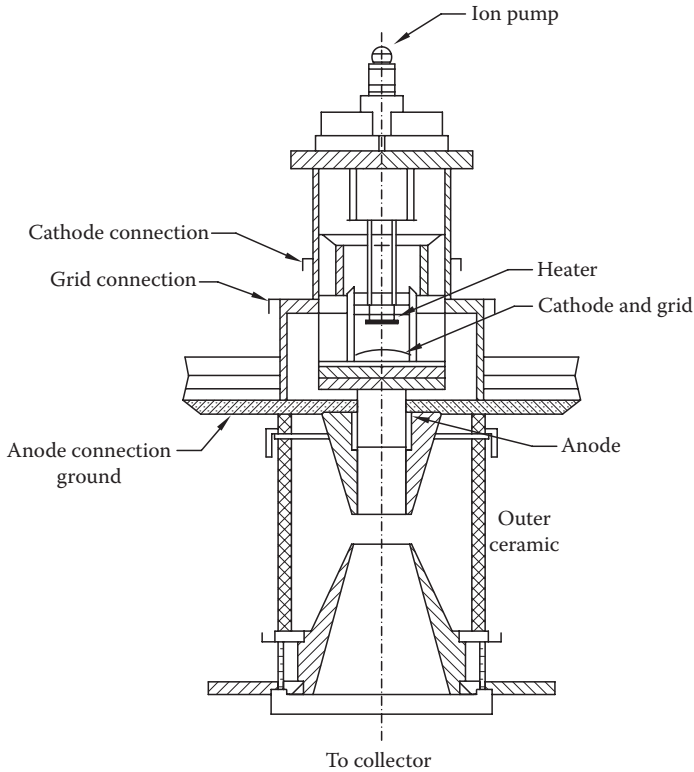


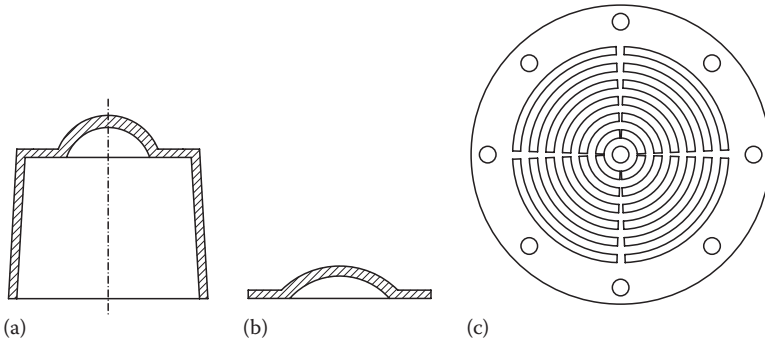
FIGURE 6.55

Overall structure of the IOT gun and output cavity. (After Clayworth, G.T. et al., *Some exciting adventures in the IOT business*, in *Proceedings of the 1992 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 200–208, 1992.)

The grid-to-cathode space of the electron gun forms the end of a long and complex RF transmission line from the RF input connector of the input cavity system. This design has a significant effect upon the final frequency range of the IOT input cavity. Because the grid-to-cathode distance is crucial to the physics of the device, the grid requires a rigid support structure to maintain proper spacing at high operating temperatures.

6.4.3 Grid Structure

A pyrolytic graphite grid is used in the IOT because of its strength and its ability to maintain desired specifications at elevated temperatures [4]. Pyrolytic graphite has a unique advantage over other common grid materials in that its strength increases as the temperature climbs to 2500°C and above, whereas the strength of pure metals universally decreases

**FIGURE 6.56**

Pyrolytic graphite grid: (a) shell, (b) grid blank, and (c) grid cut.

as the temperature increases. This allows the designer to produce a thin grid, with fine grid wires that may be accurately positioned and will retain their position and shape when raised to operating temperatures of approximately 1000°C.

During production of the grid, a hydrocarbon, typically methane, is fed into a low-pressure chamber containing a graphite rod of the correct form to produce the required graphite shell. A shell typically consists of a graphite cylinder with a closed, shaped end. The cold gas is passed into a hot zone in the reactor, which is heated by an RF eddy current system from outside the vessel. An ordered carbon structure, as opposed to an amorphous structure, is required. To provide the needed processing, the temperature of the reactor is set to approximately 2000°C at a pressure of 10 Torr. The graphite shells produced in this way have a layered structure with anisotropic properties and are physically durable.

The shells are then machined to the desired shape and size, and holes are cut into the resulting grid back using a laser under computer control. The process of producing a pyrolytic graphite grid is illustrated in Figure 6.56.

6.4.4 Input Cavity

Although similar, the Klystrode and the IOT are not identical. There are both subtle and significant differences in each design. The greatest departure can be seen in the configuration of the input cavity [4]. The design of the IOT is shown in Figure 6.57. A cylindrical resonant cavity containing an annular sliding tuning door is used. The cavity is folded at the IOT electron gun end to make contact with the tube via an RF choke structure. Because both the cathode and grid operate at beam potential (30 kV), it is necessary to maintain the body of the cavity and its tuning mechanism at ground potential by using RF chokes to prevent leakage of RF energy, while holding off the full beam voltage.

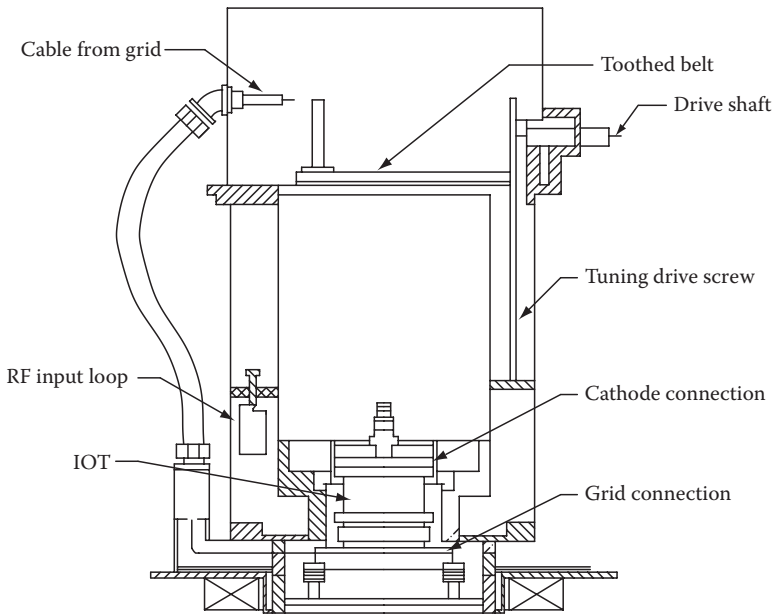


FIGURE 6.57

IOT input cavity. (After Clayworth, G.T. et al., Some exciting adventures in the IOT business, in *Proceedings of the 1992 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 200–208, 1992.)

At a given point on the grid connection within the RF choke, an insulated high-voltage cable is attached and exits so that the -30 kV dc can be supplied to the grid. This cable is screened and fitted with RF chokes to prevent UHF energy from passing into the beam power supply. RF input energy is fed via a coaxial cable to the input cavity, which is excited by means of a loop antenna carried by the annular tuning door.

Cooling air for the IOT electron gun and grid connection is fed through an insulated air pipe down the inside of the input cavity.

The tuning door of the cavity is driven by 3 two-stage tuning screws coupled to the outside via a rubber-toothed belt and bevel-gear drive. A mechanical turns counter is provided to relate cavity door position with operating frequency.

6.4.5 Output Cavity

The primary output cavity is clamped around the output ceramic of the IOT, as on a conventional klystron [4]. The output cavity, shown in Figure 6.58, contains an RF coupling loop that may be rotated about a horizontal axis to adjust the degree of coupling through a short transmission line section into a secondary cavity via a doorknob-type antenna. The secondary cavity contains a dome structure adjusted in size so that the cavity can be made to

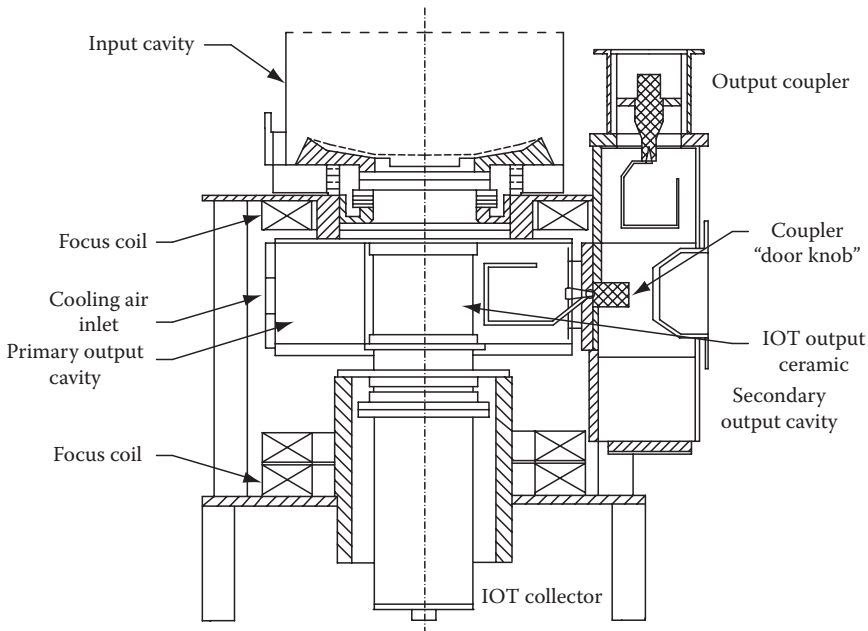


FIGURE 6.58

IOT output cavity. (After Clayworth, G.T. et al., *Some exciting adventures in the IOT business*, in *Proceedings of the 1992 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 200–208, 1992.)

cover the required frequency band. An output coupler, of standard klystron design incorporating a loop antenna, connects the secondary cavity of the RF output feeder system via a standard output interface. This design provides for instantaneous bandwidths of 8 MHz or greater.

Because RF voltages on the order of the IOT beam voltage can be expected in the output system, RF arc detectors are fitted into both cavities.

Cooling of the output cavities is accomplished by means of filtered forced air. Air enters the primary cavity, passes over the output ceramic and coupling loop, then exits via holes in the coupling hub assembly into the secondary cavity. After cooling the secondary cavity, some of the air exits via the contact fingers on the tuning doors, while a small portion traverses the output loop and coupler, exiting at a stub-pipe fitting for that purpose.

Typical operating parameters for the IOT are shown in Table 6.3.

6.4.6 Application Considerations

An active crowbar circuit is included in most applications of the Klystron/IOT to protect the pyrolytic graphite grid in the event of an arc condition inside the tube [4]. The crowbar provides the added benefit of reducing the amount of gas generated in the tube during an arc. A block diagram of a

TABLE 6.3
 Typical Operating Parameters of the IOT
 in Analog TV Service

Parameter	Typical Value
Beam voltage	28 kV
Mean beam current	1.15 A
Peak beam current	2.50 A
Grid bias	-72 V
Grid current	1 mA
Body current	6 mA
Input power	399 W
Output power	40.3 kW

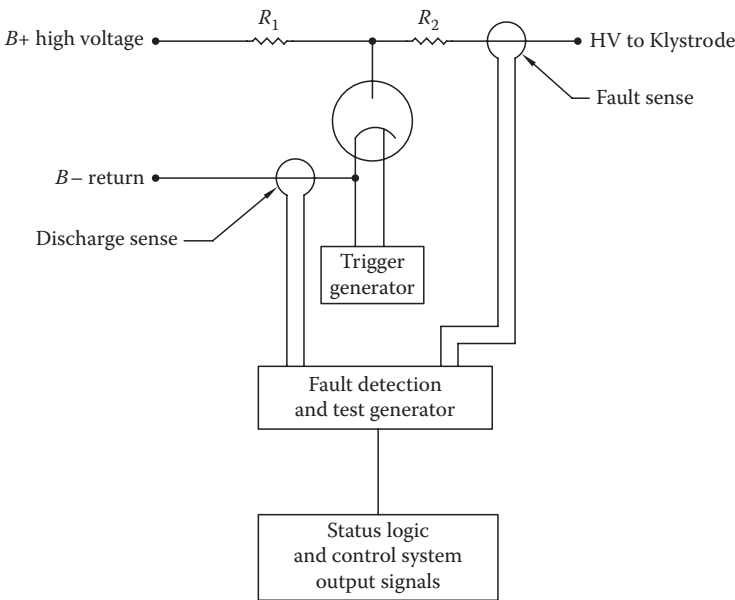


FIGURE 6.59
 Simplified block diagram of the crowbar circuit designed into the Klystron SK-series (Comark) transmitters. The circuit is intended to protect the tube from potentially damaging fault currents.

typical crowbar is shown in Figure 6.59. The response time of the crowbar is typically less than 10µs. The peak current permissible through the discharge tube may be 3000 A or greater.

Because of the IOT's class B operation, the response of the power supply to a varying load is an important design parameter. If a standard klystron beam supply were used in an IOT-equipped RF generator, performance under varying modulation levels could be unacceptable. The IOT power supply must be designed for tight transient regulation.

The IOT does not have the same hard saturation characteristic as a klystron. The IOT's high-power transfer curve flattens more slowly than a klystron's and continues to increase with increasing drive power, requiring a different approach to linearity correction.

6.4.7 Continuing Research Efforts

Several variations on the basic Klystrode/IOT theme have been developed, including air-cooled devices operating at 30 kW and above. Air cooling is practical at these power levels because of the improved efficiency that class B operation provides.

Research also is being conducted to extend the operating power of the IOT to 500 kW or more. Designed for scientific research applications, such devices offer numerous benefits over conventional klystron technology. The IOT is much smaller than a klystron of similar power, and it requires less support circuitry. Because of the improved efficiency, power supply requirements are reduced, and device cooling is simplified.

6.5 Constant Efficiency Amplifier

Because of the similarity between the spent electron beam in an IOT and that of a klystron or TWT, it is possible to consider the use of a MSDC on an IOT to improve the operating efficiency [5]. This had been considered by Priest and Shrader [6] and by Gilmore [7], but the idea was rejected because of the complexity of the MSDC assembly and because the IOT already exhibited fairly high efficiency. Subsequent development by Symons et al. [5,8] led to an IOT modified by the addition of a MSDC that has the interesting property of providing linear amplification with (approximately) constant efficiency.

6.5.1 Theory of Operation

Figure 6.60 shows a schematic representation of the *constant efficiency amplifier (CEA)* [5]. The cathode, control grid, anode and output gap, and external circuitry are essentially identical with those of the IOT amplifier. Drive power introduced into the input cavity produces an electric field between the control grid and cathode, which draws current from the cathode during positive half-cycles of the input RF signal. For operation as a linear amplifier, the peak value of the current—or more accurately, the fundamental component of the current—is made (as nearly as possible) proportional to the square root of the drive power, so that the product of this current and the voltage it induces in the output cavity will be proportional to the drive power.

Following the output cavity is a MSDC in which several typical electron trajectories are shown. These are identified by the letters *a–e*. The collector

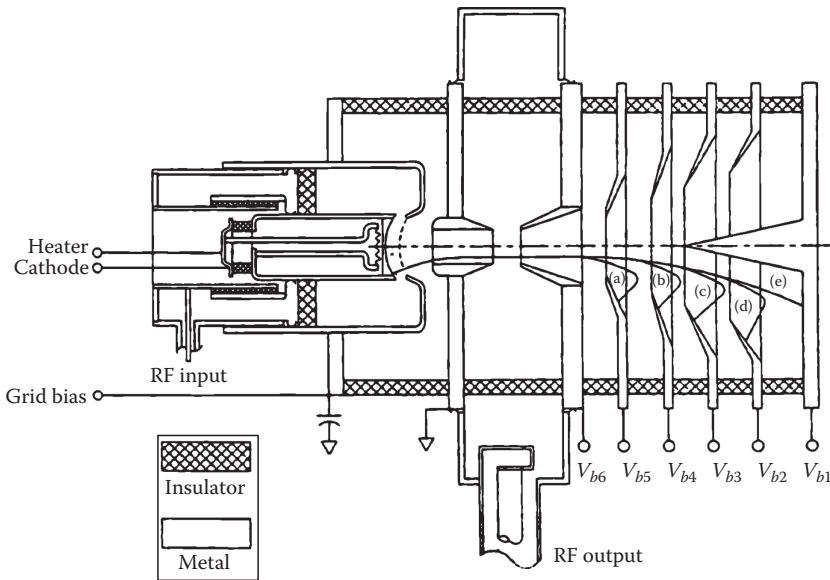


FIGURE 6.60 Schematic overview of the MSDC IOT or *constant efficiency amplifier*. (After Symons, R.S., The constant efficiency amplifier, in *Proceedings of the NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 523–530, 1997)

electrodes are connected to progressively lower potentials between the anode potential and the cathode potential so that more energetic electrons penetrate more deeply into the collector structure and are gathered on electrodes of progressively lower potentials.

In considering the difference between an MSDC IOT and an MSDC klystron, it is important to recognize that in a class B device, no current flows during the portion of the RF cycle when the grid voltage is below cutoff and the output gap fields are accelerating. As a result, it is not necessary to have any collector electrode at a potential equal to or below cathode potential. At low output powers, when the RF output gap voltage is just equal to the difference in potential between the lowest-potential collector electrode and the cathode, all the current will flow to that electrode. Full class B efficiency is thus achieved under these conditions.

As the RF output gap voltage increases with increased drive power, some electrons will have lost enough energy to the gap fields so they cannot reach the lowest potential collector, and so current to the next-to-the-lowest potential electrode will start increasing. The efficiency will drop slightly and then start increasing again until all the current is just barely collected by the two lowest-potential collectors, and so forth.

Maximum output power is reached when the current delivered to the output gap is sufficient to build up an electric field or voltage that will just stop a few electrons. At this output power, the current is divided between all of

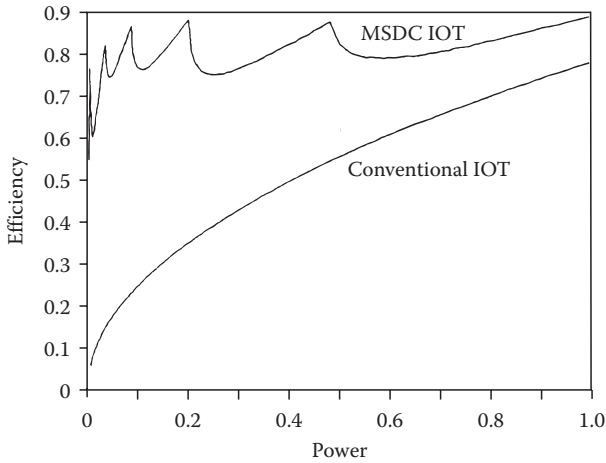


FIGURE 6.61

Calculated efficiency of an MSDC IOT compared to an unmodified IOT. (After Symons, R.S., The constant efficiency amplifier, in *Proceedings of the NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 523–530, 1997)

the collector electrodes and the efficiency will be somewhat higher than the efficiency of a single collector, class B amplifier [5].

The challenge of developing a MSDC for an IOT is not quite the same as that of developing a collector for a conventional klystron [8]. It is different because the dc component of beam current rises and falls in proportion to the square root of the output power of the tube. The dc beam current is not constant as it is in a klystron (or a TWT for that matter). As a result, the energy spread is low because the output cavity RF voltage is low at the same time that the RF and dc beam currents are low. Thus, there will be small space-charge forces, and the beam will not spread as much as it travels deep into the collector toward electrodes having the lowest potential. For this reason, the collector is likely to be rather long and thin when compared to the MSDC for a conventional klystron, as described previously.

Figure 6.61 charts the calculated efficiency for an IOT with six depressed collector stages, set at 0.1, 0.2, 0.3, 0.45, 0.7, and 1.0 times the beam voltage (compared to that of a conventional IOT).

6.6 Traveling Wave Tube

The TWT is a linear-beam device finding extensive applications in communications and research. Power levels range from a few watts to 10 MW. Gain ranges from 40 to 70 dB for small drive signals. The TWT consists of four basic elements:

- *Electron gun.* The gun forms a high-current-density beam of electrons that interact with a wave traveling along the RF circuit to increase the amplitude of the RF signal. In a typical application, electrons are emitted from a cathode and converged to the proper beam size by focusing electrodes.
- *RF interaction circuit.* The RF wave is increased in amplitude as a result of interaction with the electron beam from the gun. The fundamental principle on which the TWT operates is that an electron beam, moving at approximately the same velocity as an RF wave traveling along a circuit, gives up energy to the RF wave.
- *Magnetic electron beam focusing system.* The beam size is maintained at the proper dimensions through the interaction structure by the focusing system. This may be accomplished by using either a permanent magnet or an electromagnetic focusing element.
- *Collector.* The electron beam is received at the collector after it has passed through the interaction structure. The remaining beam energy is dissipated in the collector.

Figure 6.62 shows the basic elements of a TWT.

The primary differences between types of TWT devices involve the RF interaction structure employed. In Figure 6.62, the interaction structure is a helix. A variety of other structures may be employed, depending on the operating power and frequency. Three common approaches are used to provide the needed magnetic beam focusing. Illustrated in Figure 6.63, they are as follows:

- Electromagnetic focusing is used primarily on high-power tubes, where tight beam focusing is required.
- Permanent-magnet focusing is used where the interaction structure is short.

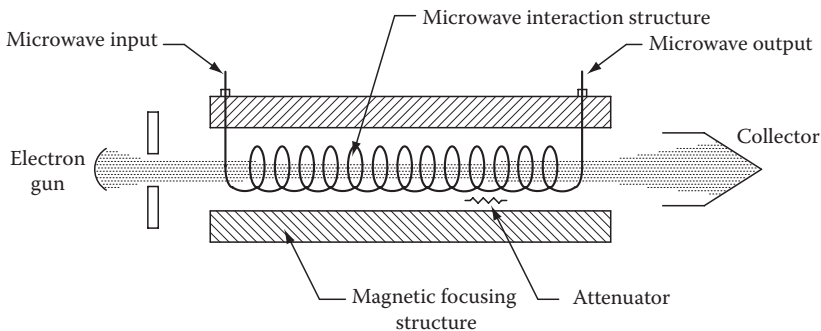


FIGURE 6.62

Basic elements of a TWT.

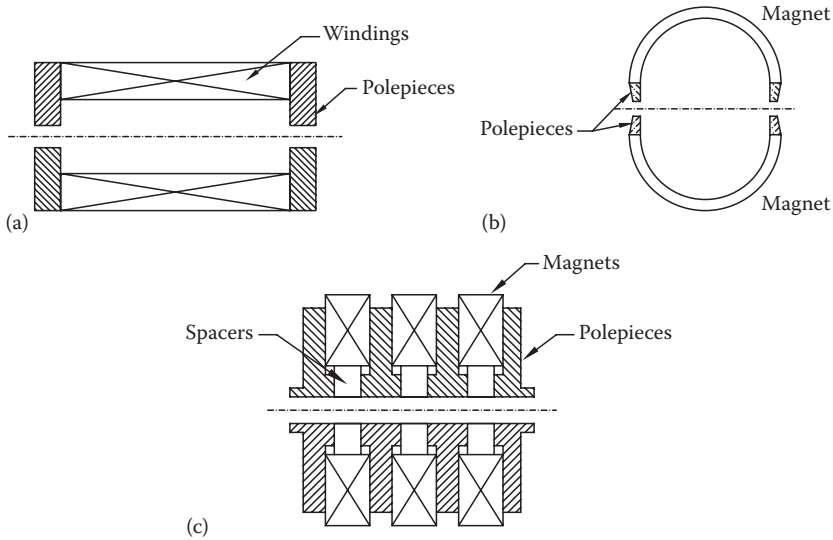


FIGURE 6.63

Magnetic focusing for a TWT: (a) solenoid type, (b) permanent-magnet type, and (c) periodic permanent-magnet structure.

- Periodic permanent-magnet focusing is used on most helix TWT and coupled-cavity tubes. The magnets are arranged with alternate polarity in successive cells along the interaction region.

6.6.1 Theory of Operation

The interaction structure acts to slow the RF signal so that it travels at approximately the same speed as the electron beam. Electrons enter the structure during both positive and negative portions of the RF cycle. Electrons entering during the positive portion are accelerated; those entering during the negative portion are decelerated. The result is the creation of *electron bunches* that produce an alternating current superimposed on the dc beam current. This alternating current induces the growth of an RF *circuit wave* that encourages even tighter electron bunching.

One or more *severs* are included to absorb reflected power that travels in a backward direction on the interaction circuit. This reflected power is the result of a mismatch between the output port and the load. Without the sever, regenerative oscillations could occur.

At a given frequency, a particular level of drive power will result in maximum bunching and power output. This operating point is referred to as *saturation*.

6.6.1.1 Interaction Circuit

The key to TWT operation lies in the interaction element. Because RF waves travel at the speed of light, a method must be provided to slow down the

forward progress of the wave to approximately the same velocity as the electron beam from the cathode. The beam speed of a TWT is typically 10%–50% the speed of light, corresponding to cathode voltages of 4–120 kV. Two mechanical structures are commonly used to slow the RF wave:

- *Helix circuit.* The helix is used where bandwidths of an octave or more are required. Over this range the velocity of the signal carried by the helix is basically constant with frequency. Typical operating frequencies range from 500 MHz to 40 GHz. Operating power, however, is limited to a few hundred watts. TWTs intended for higher-frequency operation may use a variation of the helix, shown in Figure 6.64. The *ring-loop* and *ring-bar* designs permit peak powers of hundreds of kilowatts. The average power, however, is about the same as that of a conventional helix because the structure used to support the interaction circuit is the same.
- *Coupled-cavity circuit.* The coupled-cavity interaction structure permits operation at high peak and average power levels, and moderate bandwidth (10% being typical). TWTs using coupled-cavity structures are available at operating frequencies from 2 to 100 GHz. The basic design of a coupled-cavity interaction circuit is shown in Figure 6.65. Resonant cavities, coupled through slots cut in the cavity end walls, resemble a folded waveguide. Two basic schemes are used: the *cloverleaf* and the *single-slot space harmonic* circuit.

The cloverleaf, also known as the *forward fundamental* circuit, is used primarily on high-power tubes. The cloverleaf provides operation at up to 3 MW peak power and 5 kW average at S-band frequencies. The single-slot space harmonic interaction circuit is more common than the cloverleaf. The mechanical design is simple, as shown in the figure. The single-slot space

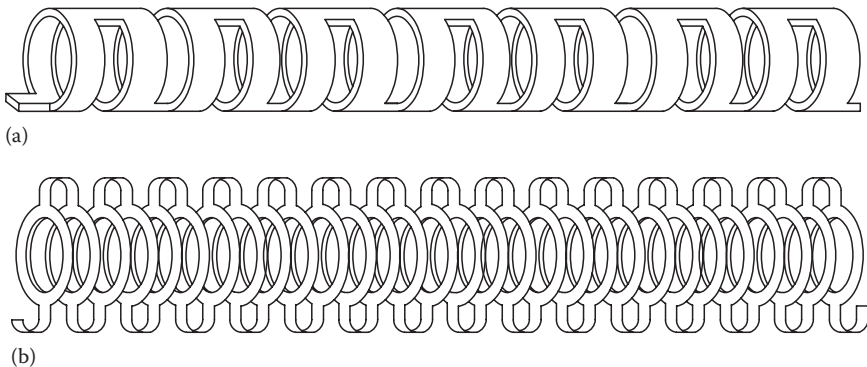
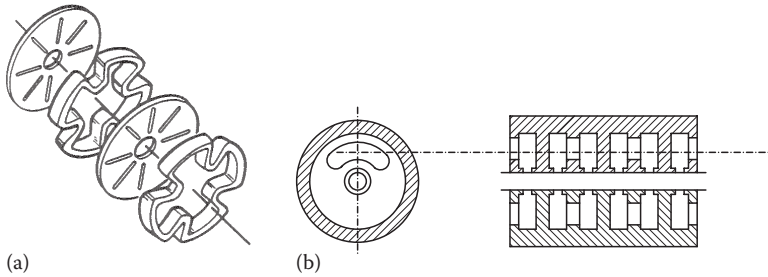


FIGURE 6.64

Helix structures for a TWT: (a) ring-loop circuit and (b) ring-bar circuit.

**FIGURE 6.65**

Coupled-cavity interaction structures: (a) forward fundamental circuit or “cloverleaf” and (b) single-slot space harmonic circuit.

harmonic structure typically provides peak power of up to approximately 50 kW and average power of 5 kW at X-band frequencies.

6.6.1.2 Pulse Modulation

The electron beam from the gun may be pulse-modulated using one of four methods:

- *Cathode pulsing.* The cathode is pulsed in a negative direction with respect to the grounded anode. This approach requires the full beam voltage and current to be switched.
- *Anode pulsing.* This approach is similar to cathode pulsing, except that the full beam voltage is switched between cathode potential and ground. The current switched, however, is only that value intercepted on the anode. Typically, the intercepted current is a few percent of the full beam potential.
- *Focus electrode pulsing.* If the focus electrode, which normally operates at or near cathode potential, is biased negatively with respect to the cathode, the beam will be turned off. The voltage swing required is typically one-third of the full cathode voltage. This approach is attractive because the focus electrode draws essentially no current, making implementation of a switching modulator relatively easy.
- *Grid pulsing.* The addition of a grid to the cathode region permits control of beam intensity. The voltage swing required for the grid, placed directly in front of the cathode, is typically 5% of the full beam potential.

6.6.1.3 Electron Gun

The electron gun of a TWT is a device that supplies the electron beam to the tube [9]. A schematic diagram of a generic electron gun is given in Figure 6.66. The device consists of a hot cathode heated by an electric heater, a negatively

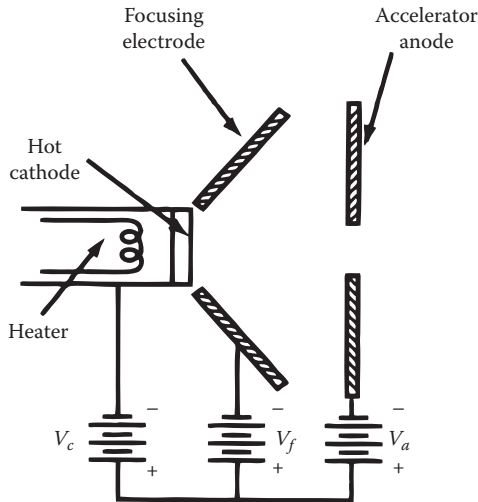


FIGURE 6.66

Generic TWT electron gun structure. (After Ishii, T.K., *Traveling wave tubes*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 428–443, 1996.)

biased focusing electrode or *focuser*, and a positively biased *accelerating anode*. The cross-sectional view given in the figure can be a two-dimensional or three-dimensional coaxial structure [10,11].

An axially symmetrical solid cylindrical electron beam is produced by the gun structure shown in Figure 6.66 if the structure is axially cylindrically symmetrical. If the middle of the hot cathode is made nonemitting and only the edge of the cathode is emitting, the cathode becomes an *annular cathode* [9]. The annular cathode produces a hollow beam. The annular electron beam can be used to reduce beam current for a given microwave output power.

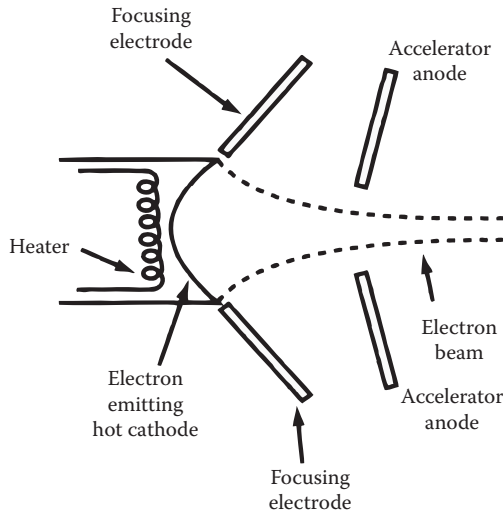
If the gun structure shown in Figure 6.66 is two dimensional, then a ribbon-shaped electron beam is produced. A ribbon-shaped beam is used for a TWT of a two-dimensional structure.

If the angle of the focusing electrode against the axis of the electron beam is 67.5° and the anode is also tilted forward to produce a *rectilinear flow* (electron flow parallel to the z axis in Figure 6.66), then such an electron gun is termed the *Pierce gun*.

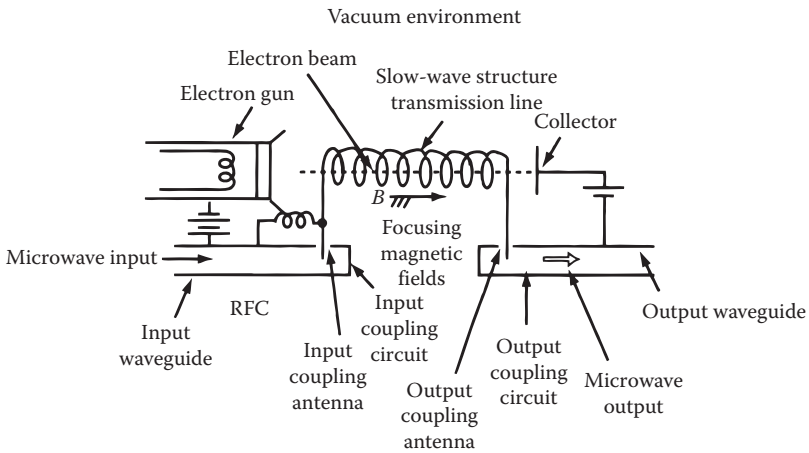
In practice, the hot cathode surface is curved as shown in Figure 6.67 to increase the electron emitting surface and to obtain a high-density electron beam.

6.6.1.4 Beam Focusing

Electrons in an electron beam mutually repel each other by the electron's own Coulomb force because of their negative charge [9]. In addition, the electron beam usually exists in proximity to the positively biased slow-wave structure, as shown in Figure 6.68. Therefore, the electron beam tends to diverge.

**FIGURE 6.67**

Cross-sectional view of a TWT electron gun with a curved hot cathode. (After Ishii, T.K., *Traveling wave tubes*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 428–443, 1996.)

**FIGURE 6.68**

Generic configuration of a TWT. (After Ishii, T.K., *Traveling wave tubes*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 428–443, 1996.)

The process of confining the electron beam within the desired trajectory against the mutual repulsion and diverging force from the slow-wave structure is termed *electron beam focusing*.

The electron beam in a TWT is usually focused by a dc magnetic flux applied parallel to the direction of the electron beam, which is coaxial to the slow-wave transmission line. Variations on this basic technique include

- *Brillouin flow*, where the output of the electron gun is not exposed to the focusing magnetic flux [12]
- *Immersed flow*, where the electron gun itself is exposed to and unshielded from the focusing para-axial longitudinal magnetic flux [12]
- *Generic flow*, where the electron gun is not shielded from the focusing magnetic flux, and the focusing flux is not para-axial (in other words, neither Brillouin nor immersed flow)

6.6.1.5 Collector Assembly

Various configurations are used for the collector assembly of a TWT. Figure 6.69 shows a selection of the more common, including [9]

- Plate collector
- Cone collector
- Curved cone
- Cylinder collector
- Depressed potential cylinder
- Two-stage collector
- Three-stage collector

Cooling options include conduction, air, and water.

Cooling of a low-power TWT is accomplished by clamping the tube to a metal baseplate, mounted in turn on an air- or liquid-cooled heat sink. Coupled-cavity tubes below 1 kW average power are convection cooled by circulating air over the entire length of the device. Higher-power coupled-cavity tubes are cooled by circulating liquid over the tube body and collector.

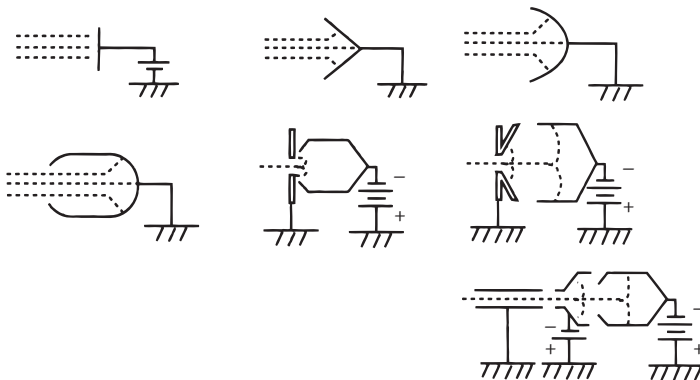


FIGURE 6.69 Cross-sectional view of various collector configurations for a TWT. (From Ishii, T.K., *Traveling wave tubes*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 428–443, 1996. Used with permission.)

6.6.2 Operating Efficiency

Efficiency is not one of the TWT's strong points. Early TWTs offered only about 10% dc-to-RF efficiency. Wide bandwidth and power output are where the TWT shines. TWT efficiency may be increased in two basic ways: (1) collector depression for a single-stage collector or (2) use of a multistage collector.

Collector depression refers to the practice of operating the collector at a voltage lower than the full beam voltage. This introduces a potential difference between the interaction structure and the collector, through which electrons pass. The amount by which a single-stage collector can be depressed is limited by the remaining energy of the slowest electrons. In other words, the potential reduction can be no greater than the amount of energy of the slowest electrons, or they will turn around and reenter the interaction structure, causing oscillations.

By introducing multiple depressed collector stages, still greater efficiency can be realized. This method provides for the collection of the slowest electrons at one collector potential, while allowing those with more energy to be collected on other stages that are depressed still further. This approach, similar to the MSDC design discussed previously, is illustrated in Figure 6.70.

6.6.3 Operational Considerations

Although TWTs offer numerous benefits to the end user, they are not without their potential drawbacks.

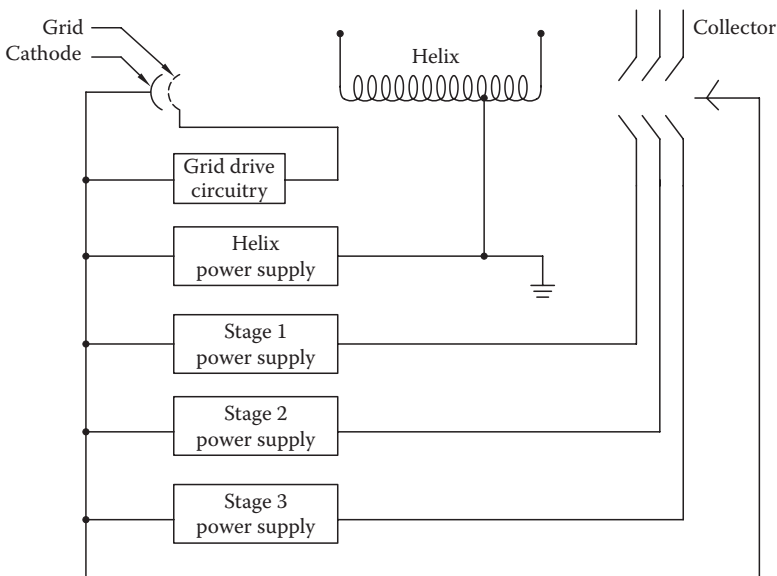


FIGURE 6.70
Power supply configuration for a MSDC TWT.

6.6.3.1 Intermodulation Distortion

TWTs are susceptible to intermodulation (IM) distortion when multiple carriers are introduced, as in double illuminated or multiple SCPC (*single channel per carrier*) satellite transponders. These IM products may be found at frequencies that are displaced from the fundamental carriers by the difference in frequency between them. When multiple carriers are present, the potential for IM exists. This potential is reduced by operating the TWT below saturation. Power must be reduced (backed off) in proportion to the number of carriers and their relative power.

6.6.3.2 Second-Harmonic Content

Because of the wide bandwidth and nonlinear operating characteristics under saturation conditions, a TWT may generate significant second-harmonic energy. It is not uncommon to measure second-harmonic energy at the output of a TWT that is down only 10 dB from the operating carrier.

Reduction of harmonic content usually involves injecting a coherent harmonic signal with controlled phase and amplitude along with the fundamental carrier so that they interact, minimizing harmonic energy at the output of the device.

6.6.3.3 AM/PM Conversion

AM/PM conversion is the change in phase angle between the input and output signals as the input varies in amplitude. The root cause of this distortion in a TWT centers on the reduction of electron beam velocity as the input signal level increases. This causes a greater energy exchange between the electron beam and the RF wave. At a level 20 dB below the input power required for saturation, AM/PM conversion is negligible. At higher levels, AM/PM distortion may increase.

6.6.3.4 Phase Variation

When the velocity of the electron beam in the TWT is changed, the phase of the output signal also will vary. The primary causes of beam velocity variations include changes in one or more of the following:

- Cathode temperature
- Grid voltage
- Anode voltage
- Cathode voltage

The TWT power supply must be well regulated (to less than 0.2%) to prevent beam velocity changes that may result in output signal phase variations.

6.7 Crossed-Field Tubes

A crossed-field microwave tube is a device that converts dc into microwave energy using an electronic energy-conversion process. These devices differ from *beam tubes* in that they are *potential-energy converters*, rather than kinetic-energy converters. The term *crossed field* is derived from the orthogonal characteristics of the dc electric field supplied by the power source and the magnetic field required for beam focusing in the interaction region. This magnetic field typically is supplied by a permanent-magnet structure. Such devices also are referred to as *M-tubes*.

Practical devices based on the crossed-field principles fall into two broad categories:

- *Injected-beam crossed-field tubes.* The electron stream is produced by an electron gun located external to the interaction region, similar to a TWT. The concept is illustrated in Figure 6.71.
- *Emitting-sole tubes.* The electron current for interaction is produced directly within the interaction region by secondary electron emissions, which result when some electrons are driven to the negative electrode and allowed to strike it. The negative electrode is formed using a material capable of producing significant numbers of secondary-emission electrons. The concept is illustrated in Figure 6.72.

6.7.1 Magnetron

The magnetron encompasses a class of devices finding a wide variety of applications. Pulsed magnetrons have been developed that cover frequency

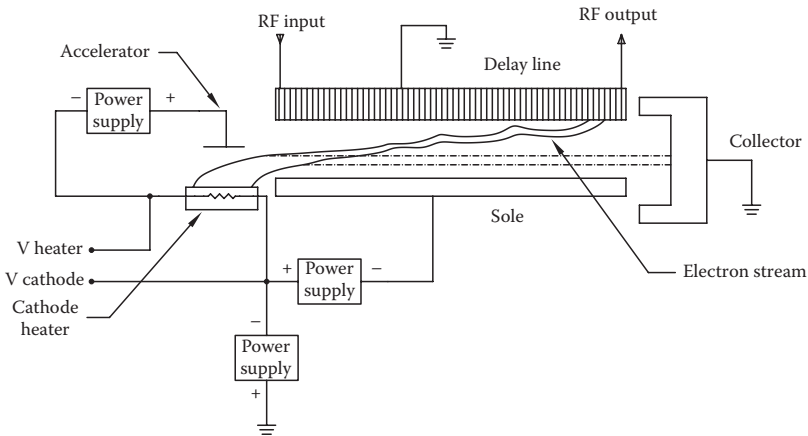


FIGURE 6.71

Linear injected-beam microwave tube.

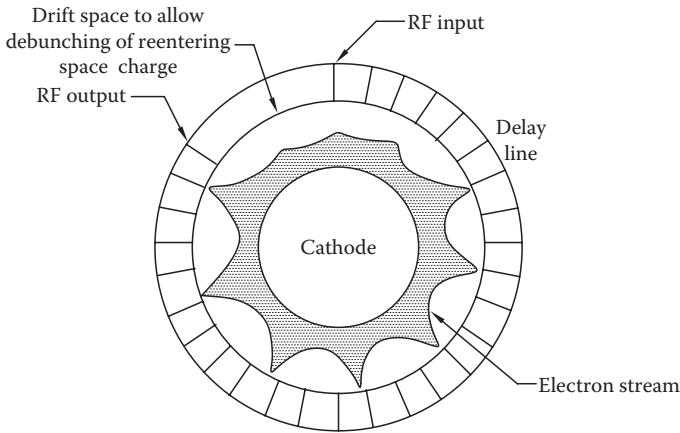


FIGURE 6.72
Reentrant emitting-sole CFA tube.

ranges from the low UHF band to 100 GHz. Peak power from a few kilowatts to several megawatts has been obtained. Typical overall efficiencies of 30%–40% may be realized, depending on the power level and operating frequency. CW magnetrons also have been developed, with power levels of a few hundred watts in a tunable tube, and up to 25 kW or more in a fixed-frequency device. Efficiencies range from 30% to as much as 70%.

The magnetron operates electrically as a simple diode. Pulsed modulation is obtained by applying a negative rectangular voltage waveform to the cathode with the anode at ground potential. Operating voltages are less critical than for beam tubes; line-type modulators often are used to supply pulsed electric power. The physical structure of a conventional magnetron is shown in Figure 6.73.

High-power pulsed magnetrons are used in radar systems. Low-power pulsed devices find applications as beacons. Tunable CW magnetrons are used in ECM applications. Fixed-frequency devices are used as microwave heating sources.

Tuning of conventional magnetrons is accomplished by moving capacitive tuners or by inserting symmetrical arrays of plungers into the inductive portions of the device. Tuner motion is produced by a mechanical connection through flexible bellows in the vacuum wall. Tuning ranges of 10%–12% of bandwidth are possible for pulsed tubes, and as much as 20% for CW tubes.

6.7.1.1 Operating Principles

Most magnetrons are built around a cavity structure of the type shown in Figure 6.74. The device consists of a cylindrical cathode and anode, with cavities in the anode that open into the cathode–anode space—the *interaction*

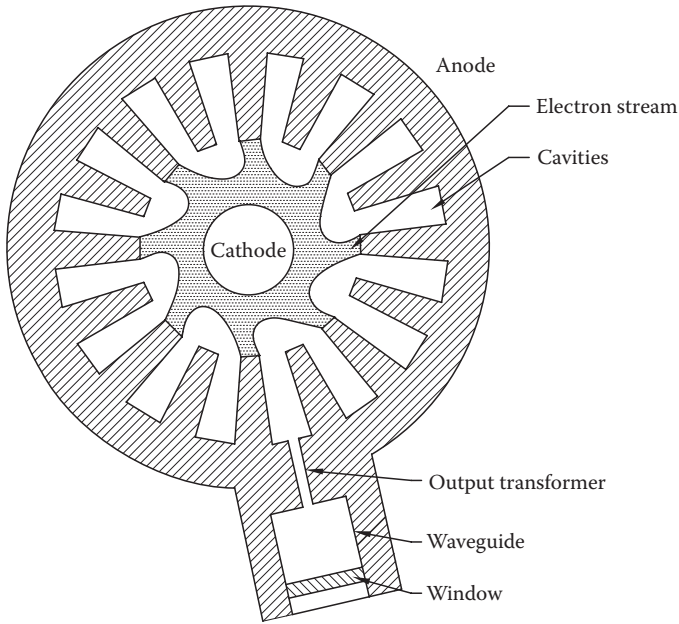


FIGURE 6.73
Conventional magnetron structure.

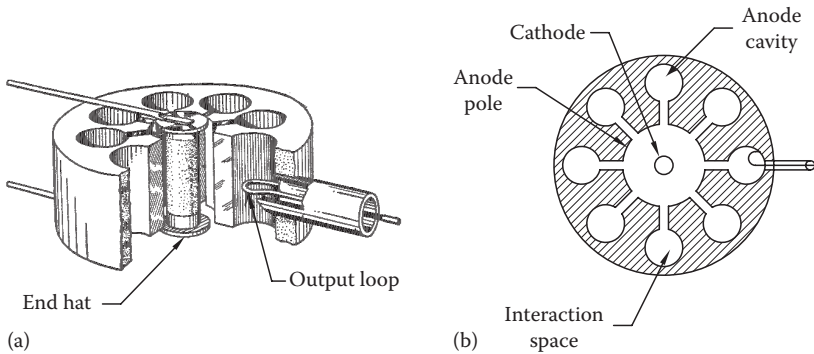
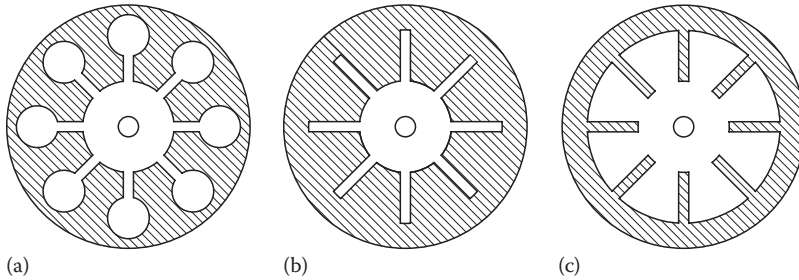


FIGURE 6.74
Cavity magnetron oscillator: (a) cutaway view and (b) cross-sectional view perpendicular to the axis of the cathode.

space—as shown. Power can be coupled out of the cavities by means of a loop or a tapered waveguide.

Cavities, together with the spaces at the ends of the anode block, form the resonant system that determines the frequency of the generated oscillations. The actual shape of the cavity is not particularly important, and various types are used, as illustrated in Figure 6.75. The oscillations associated with the cavities are of such a nature that alternating magnetic flux lines pass through

**FIGURE 6.75**

Cavity magnetron oscillator anode: (a) hole-and-slot type, (b) slot type, and (c) vane type.

the cavities parallel to the cathode axis, while the alternating electric fields are confined largely to the region where the cavities open into the interaction space. The most important factors determining the resonant frequency of the system are the dimensions and shape of the cavities in a plane perpendicular to the axis of the cathode. Frequency also is affected by other factors such as the end space and the axial length of the anode block, but to a lesser degree.

The magnetron requires an external magnetic field with flux lines parallel to the axis of the cathode. This field usually is provided by a permanent magnet or electromagnet.

The cathode is commonly constructed as a cylindrical disk.

6.7.1.2 Coaxial Magnetron

The frequency stability of a conventional magnetron is affected by variations in the load impedance and by cathode-current fluctuations. Depending on the extent of these two influences, the magnetron occasionally may fail to produce a pulse. The coaxial magnetron minimizes these effects by using the anode geometry shown in Figure 6.76. Alternate cavities are slotted to provide coupling to a surrounding coaxial cavity.

The oscillating frequency is controlled by the combined vane system and the resonant cavity. Tuning may be accomplished through the addition of a movable end plate in the cavity, as shown in Figure 6.77.

6.7.1.3 Frequency-Agile Magnetron

Tubes developed for specialized radar and ECM applications permit rapid tuning of the magnetron. A conventional device may be tuned using one of the following methods:

- A rapidly rotating capacitive element. Tubes of this type are referred to as *spin-tuned magnetrons*.
- A hydraulic-driven tuning mechanism. Tubes of this type are referred to as *mechanically tuned magnetrons*.

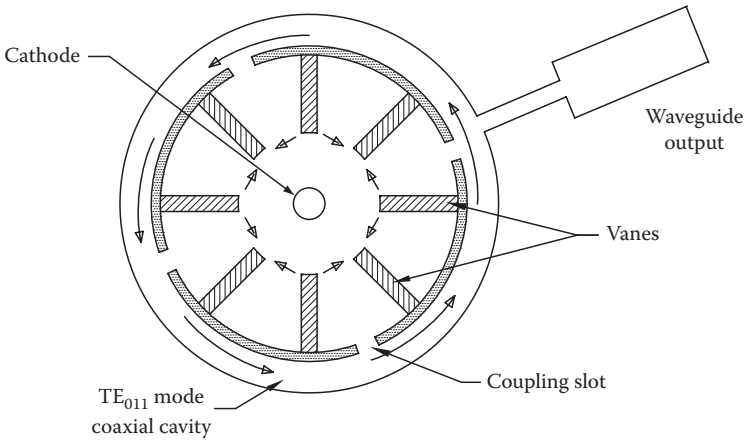


FIGURE 6.76
Structure of a coaxial magnetron.

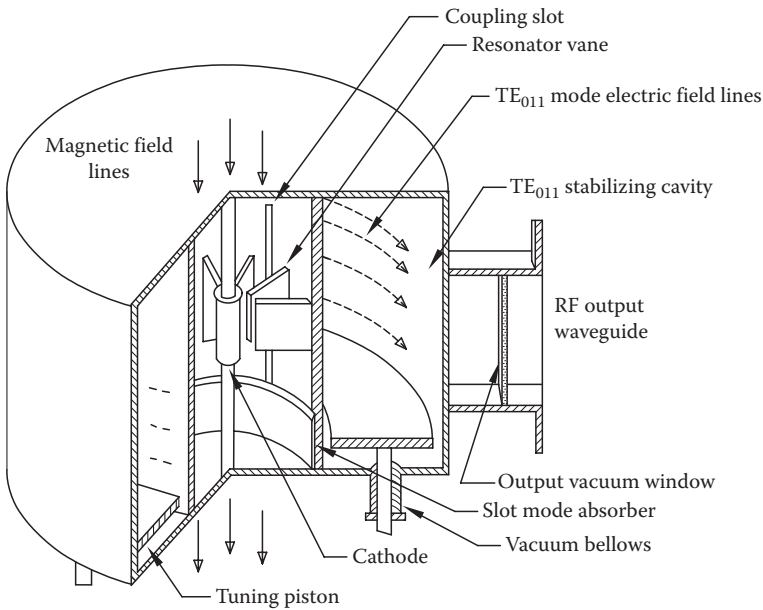


FIGURE 6.77
Structure of a tunable coaxial magnetron.

Electronic tuning of magnetrons is also possible, with tuning rates as high as several megahertz per microsecond.

6.7.1.4 Linear Magnetron

Although the most common types of magnetrons are radial in nature, the *linear magnetron* and *inverted magnetron* may also be used, depending upon

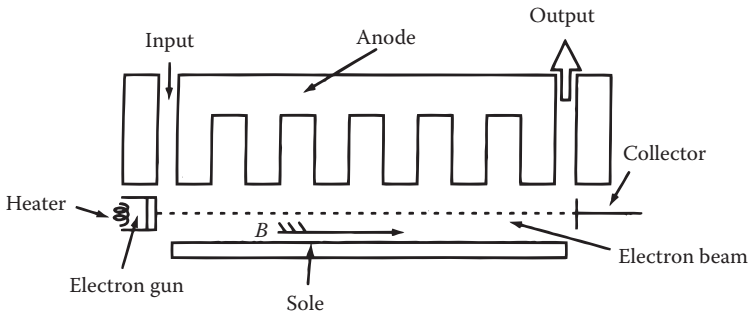


FIGURE 6.78

Cross-sectional view of a linear magnetron. (After Ishiik, T.K., *Other microwave vacuum devices*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

the application [13]. A cross-sectional view of a linear magnetron is given in Figure 6.78. Shown in the figure is the *O-type* linear magnetron, in which the electron beam emitted from the electron gun is focused by a longitudinally applied dc magnetic flux density (B), as in the case of the TWT.

As shown in the figure, a number of slots are included in the basic structure. These slots are cut $\frac{1}{4}$ -wavelength deep, functioning as quarter-wave cavity resonators. This structure forms a series of microwave cavity resonators coupling to an electron beam, in a similar manner to the multicavity klystron. The velocity-modulated electrons are bunched, and the tightly bunched electrons produce amplified microwave energy at the output cavity, which is coupled to an external circuit. The linear magnetron typically offers high gain, but narrow frequency bandwidth.

6.7.2 Backward Wave Oscillator

In a TWT, if the microwave signal to be amplified is propagating in the slow-wave structure backwardly to the direction of the electron beam, the device is termed a *backward wave oscillator* (BWO) [13]. Microwaves traveling in a backward direction carry positive feedback energy toward the electron gun and yield stronger velocity modulation and bunching. Thus, the system is inherently an oscillator rather than a stable amplifier. The input is typically terminated by an impedance-matched reflectionless termination device. The oscillation frequency is a function of the speed of the electrons and the time constant of the feedback mechanism. The speed of electron motion is controlled by the anode voltage.

An *M-type* radial BWO is shown in Figure 6.79. The direction of electron pole motion and the direction of microwave propagation along the annular reentrant-type slow-wave structure are opposite each other. It should be noted that the depths of the slits cut in the inner surface of the anode is very shallow—much less than $\frac{1}{4}$ -wavelength deep. In other words, the slits are

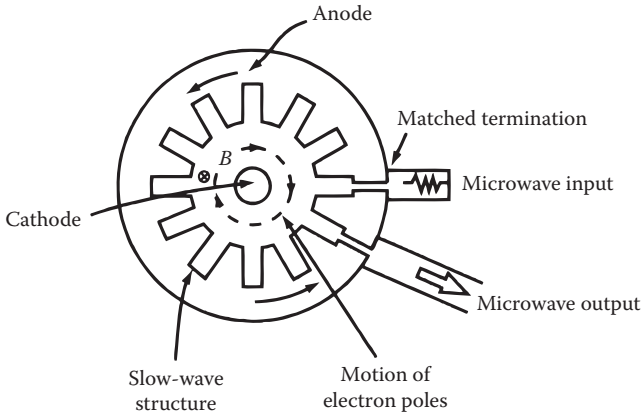


FIGURE 6.79

Functional schematic of the M-type radial BWO. (After Ishiik, T.K., *Other microwave vacuum devices*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

not in resonance; they are not cavity resonators, as is the case of a magnetron. Rather, the slits are nonresonating, as in the case of a TWT. In the M-type radial BWO, the electron beam is focused by a magnetic flux density applied perpendicular to the beam, as seen from Figure 6.79.

An M-type radial BWO is sometimes termed the *Carcinotron*, a trade name. A key feature of the Carcinotron is its wide voltage tunability over a broad frequency range.

6.7.3 Strap-Fed Devices

A radial magnetron can be configured so that every other pole of the anode resonators are conductively tied for microwave potential equalization, as shown in Figure 6.80a [13]. These conducting tie rings are termed *straps*; the technique of using strap rings is termed *strapping*. Strapping ensures good synchronization of microwaves in the magnetron resonators with the rotation of electron poles.

The technique of strapping is extended and modified for an M-type radial BWO, as shown in Figure 6.80b. Strapping rings tie every other pole of the radial slow-wave structure, as in the case of a strapped radial magnetron, but the strapping rings are no longer reentrant. Microwave energy to be amplified is fed to the strap at one end, and the amplified output is extracted from the other end. This type of electron tube is termed a *strap-fed device*.

If an M-type radial BWO is strapped, usually it does not start oscillation by itself. But, if microwave energy is fed through the strap from the outside using an external microwave power source to the microwave input, then the oscillation starts—and even if the exciter source is turned off, the oscillation continues. This type of M-type radial BWO is termed the *platinotron* [14].

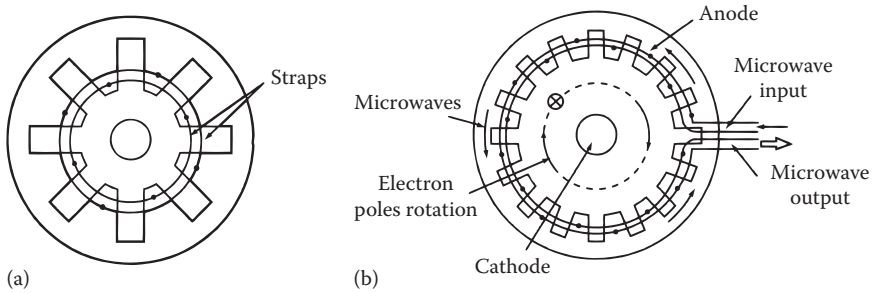


FIGURE 6.80

Strap-fed devices: (a) strapped radial magnetron and (b) nonreentrant strapping of a BWO. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

In a platinotron, if the output of the tube is fed back to the input through a high- Q cavity resonator, it becomes a self-starting oscillator. The oscillation frequency is stabilized by the high- Q cavity resonator. This type of high-power frequency-stabilized strapped radial BWO is termed the *stabilotron* [14]. The operating powers are at kilowatt and megawatt levels.

Performance of the platinotron depends on, among other things, the design of the slow-wave structure. For example, the interdigital slow-wave structure as shown in Figure 6.80 has a limited power-handling capability and frequency bandwidth. Design of a slow-wave structure with greater power-handling capacity and stability, with broader frequency bandwidth, is possible. For example, instead of an anode with an interdigital slow-wave structure, the anode can be made of an annular open conducting duct, loaded with a number of pairs of conducting posts across the open duct. Strapping is done at every other tip of the pairs of conducting posts. This type of strapping loads the slow-wave structure, stabilizing it and preventing oscillation. The structure of the anode with an annular duct and pairs of posts increases the power-handling capability. This type of loaded radial BWO is termed the *amplitron* [10,14]. The amplitron is capable of amplifying high-power microwave signals with pulses and CW. It is used for long-range pulsed radar transmitter amplifiers and industrial microwave heating generators. The operating power levels range from kilowatt to megawatt levels.

6.7.4 Gyrotron

The *gyrotron* is a *cyclotron resonance maser*.* The device includes a cathode, collector, and circular waveguide of gradually varying diameter. Electrons are emitted at the cathode with small variations in speed. The electrons

* Maser is an acronym for *microwave amplification by simulated emission of radiation*. Maser is a general class of microwave amplifiers based on molecular interaction with electromagnetic radiation.

then are accelerated by an electric field and guided by a static magnetic field through the device. The nonuniform induction field causes the rotational speed of the electrons to increase. The linear velocity of the electrons, as a result, decreases. The interaction of the microwave field within the waveguide and the rotating (helical) electrons causes bunching similar to the bunching within a klystron. A decompression zone at the end of the device permits decompression and collection of the electrons.

The power available from a gyrotron is 100 times greater than that possible from a classic microwave tube at the same frequency. Power output of 100 kW is possible at 100 GHz, with 30% efficiency. At 300 GHz, up to 1.5 kW may be realized, but with only 6% efficiency.

6.7.4.1 Theory of Operation

The trajectory of an electron in an electron beam focused by a longitudinally applied magnetic field is a helix [13]. If the electron velocity, electron injection angle, and applied longitudinal magnetic flux density are varied, then an electron beam of helical form with different size and pitch will be formed. A coil-shaped electron beam will be produced by adjusting the acceleration voltage, applied magnetic flux density, and the electron injection angle to the focusing magnetic field. The coil of the electron beam can be a simple single coil, or—depending on the adjustment of the aforementioned three parameters—it can be an electron beam of a double coil, or a large coil made of thin small coils. In the case of the double-coil trajectory, the large coil-shaped trajectory is termed the *major orbit* and smaller coil trajectory is termed the *minor orbit*.

If a single coil-shaped electron beam is launched in a waveguide, as shown in Figure 6.81, then microwaves in the waveguide will interact with the helical beam. This type of vacuum tube is termed the *helical beam tube* [14]. In this class of device, a single-coil helical beam is launched into a TE_{10} mode rectangular waveguide. Inside the waveguide, microwaves travel from right to left

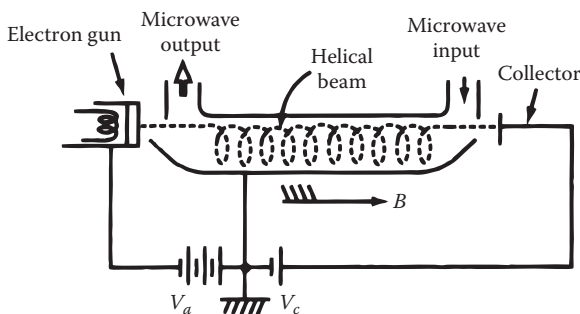


FIGURE 6.81

Basic structure of a helical beam tube. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

and the helical beam travels in an opposite direction. Therefore, the microwave-electron beam interaction is of the *backward wave* type. If the microwave frequency, the focusing magnetic flux density B , and the acceleration voltage V_a are properly adjusted, this device will function as a *backward wave amplifier*. Electrons in the helical beam interact with the transverse microwave electric fields and are velocity-modulated at the left-hand side of the waveguide as the beam enters into the waveguide. The velocity-modulated electrons in the helical beam are bunched as they travel toward the right. If the alternating microwave electric field synchronizes its period and phase with the helical motion of bunched electrons so that the electrons always receive retardation from microwave transverse electric fields, then the electrons lose their kinetic energy and the microwave signal gains in electric field energy according to the principle of kinetic energy conservation. Thus, the amplified microwave power emerges at the waveguide output at the left (because the microwaves travel backward).

In Figure 6.81, if the microwave input port and the output port are interchanged with each other, then the system becomes a *forward wave amplifier*. Such a forward wave amplifier is termed a *peniotron* [14].

If the electron gun is modified to incorporate a side emitting cathode and the waveguide is changed to TE_{11} mode oversized circular waveguide, as shown in Figure 6.82a, the gyrotron is formed. In this device, both ends of the waveguide are open and there are sufficient reflections in the waveguide for positive feedback. The gyrotron is, thus, a forward wave oscillator.

A *double-coil helical beam gyrotron* is shown in Figure 6.82b. The device is formed by readjusting the anode voltage and the focusing flux density so that the electron beam is made into a double helical coil (as shown in the figure), and operating the oversized circular waveguide in the TE_{01} mode. In the TE_{01} mode, the microwave transverse electric fields exist as concentric circles. Therefore, the tangential electric fields interact with electrons in the small coil trajectory. The alternating tangential microwave electric fields are made to synchronize with the tangential motion of electrons in the minor coil-shaped trajectory. Thus, electron velocity modulation takes place near the cathode and bunching takes place in the middle of the tube. Microwave kinetic energy transfer takes place as the beam approaches the right. The focusing magnetic flux density B is applied only in the interaction region. Therefore, if the electron beam comes out of the interaction region, it is defocused and collected by the anode waveguide (as depicted in Figure 6.82b). If the circular waveguide is operated in an oversized TE_{11} mode, with the double-coil helical beam, then the device is referred to as a *tornadotron* [14]. Microwave-electron interaction occurs between the parallel component of tangential motion of the small helical trajectory and the TE_{11} mode microwave electric field. If the phase of microwave electric field decelerates bunched electrons, then the lost kinetic energy of the bunched electrons is transferred to the microwave signal and oscillation begins.

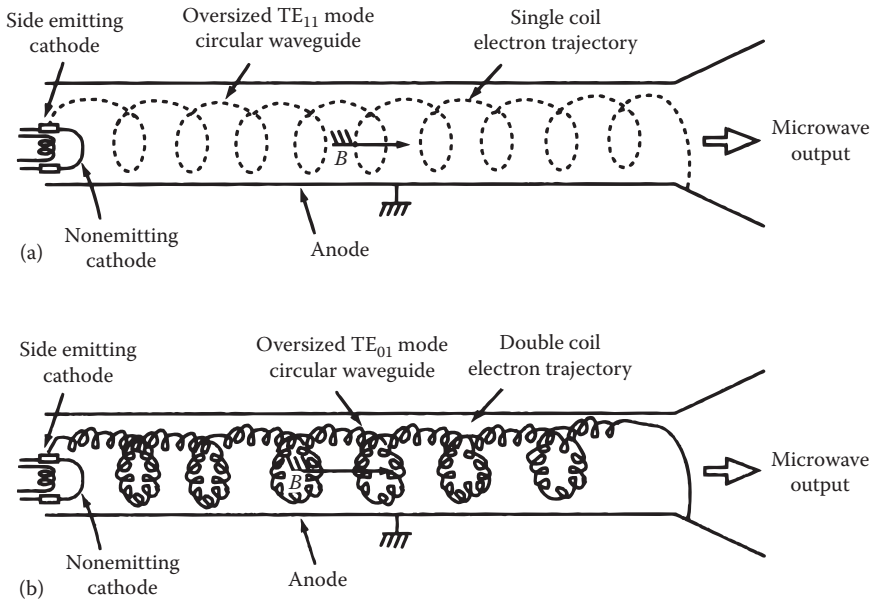


FIGURE 6.82

Functional schematic diagram of the gyrotron: (a) single coiled helical beam gyrotron and (b) electron trajectory of double-coil helical beam gyrotron. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

6.7.4.2 Gyrotron Design Variations

The gyrotron exists in a number of design variations, each optimized for a particular feature or application [13,15,16].

When the gyrotron circular waveguide is split as shown in Figure 6.83, the tube is termed the *gyroklystron amplifier* [13]. Both waveguides resonate to the input frequency, and there are strong standing waves in both waveguide resonators. The input microwave signal to be amplified is fed through a side opening to the input waveguide resonator. This is the *buncher resonator*, which functions in a manner similar to the klystron. The buncher resonator imparts velocity modulation to gyrating electrons in the double helical coil-shaped electron beam. There is a drift space between the buncher resonator and the catcher resonator at the output. While drifting electrons bunch and bunched electrons enter into the output waveguide catcher resonator, electron speed is adjusted in such a manner that electrons are decelerated by the resonating microwave electric field. This lost kinetic energy in bunched electrons is transformed into microwave energy and microwaves in the catcher resonator are, thus, amplified. The amplified power appears at the output of the tube.

If the gyrotron waveguide is an unsplit one-piece waveguide that is impedance-matched and not resonating, as shown in Figure 6.84, the tube is termed the *gyrotron traveling wave tube amplifier* [13]. In this tube, the input

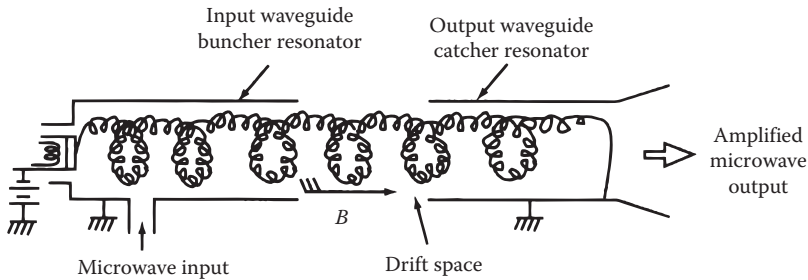


FIGURE 6.83 Basic structure of the gyrokystron amplifier. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

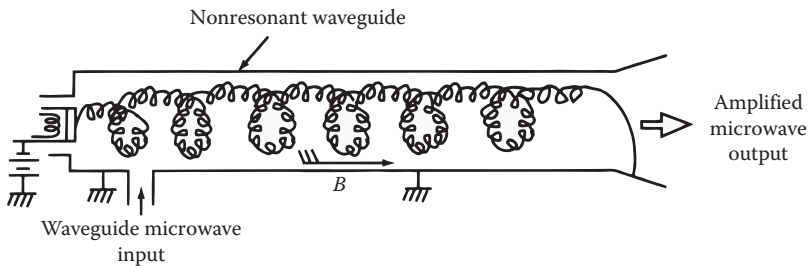


FIGURE 6.84 Basic structure of the gyrokystron TWT amplifier. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

microwaves are fed through an opening in the waveguide near the electron gun. Microwaves in the waveguide are amplified gradually as they travel toward the output port by interacting with the double-coiled helical electron beam, which is velocity-modulated and bunched. There are no significant standing waves in the waveguide. Microwaves grow gradually in the waveguide as they travel toward the output port as a result of interaction with electrons.

If the electron gun of the gyrotron is moved to the side of the waveguide and microwave power is extracted from the waveguide opening in proximity to the electron gun, as shown in Figure 6.85, then the device is termed a *gyrotron backward oscillator* [13]. The principle involved is similar to the BWO, and the process of velocity modulation, drifting, bunching, and catching is similar to that of the klystron. Microwave energy induced in the waveguide travels in both directions, but the circuit is adjusted to emphasize the waves traveling in a backward direction. The backward waves become the output of the tube and, at the same time, carry the positive feedback energy to the electrons just emitted and to be velocity-modulated. The system, thus, goes into oscillation.

If the gyrotron waveguide is split into two again, but this time the input side waveguide is short and the output side waveguide is long, as shown in

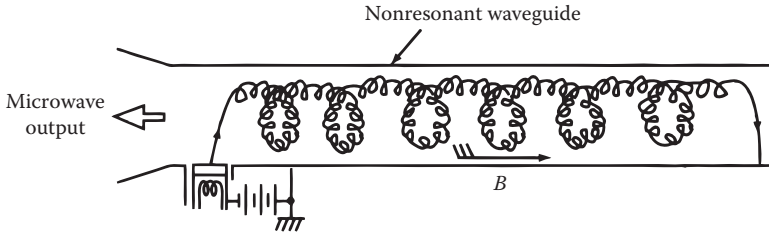


FIGURE 6.85

Basic structure of the gyrotron backward oscillator. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

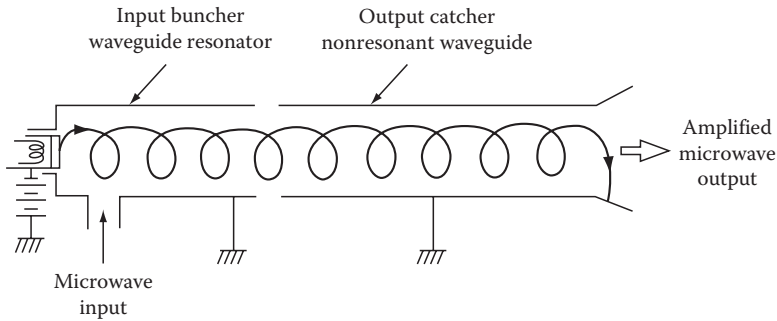


FIGURE 6.86

Basic structure of the gyrotwystron amplifier. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.)

Figure 6.86, then the tube is termed a *gyrotwystron amplifier* [13]. This device is a combination of the gyroklystron and gyrotron TWT amplifier, thus the name gyrotwystron amplifier. The input side waveguide resonator is the same as the input resonator waveguide of a gyroklystron. There are strong standing waves in the input bunched-waveguide resonator. There is no drift space between the two waveguides. The output side waveguide is a long impedance-matched waveguide and there is no microwave standing wave in the waveguide (a traveling-wave waveguide). As microwaves travel in the waveguide, they interact with bunched electrons and the microwaves grow as they move toward the output port.

6.8 Other Microwave Devices

There are a number of variations on the basic microwave devices outlined previously in this chapter. Some find widespread usage, others have little commercial interest. Among the important specialty device classifications are *quasiquantum devices* and modified klystrons.

6.8.1 Quasiquantum Devices

In a gyrotron, if electrons are accelerated by extremely high voltage, they become relativistic [13]. That is, the mass of an electron becomes a function of the velocity as

$$m = \frac{m_0}{\sqrt{1 - \left(\frac{c}{v}\right)^2}} \tag{6.3}$$

where

- m is the relativistic mass of an electron
- m_0 is the static mass of an electron
- c is the speed of light in vacuum
- v is the speed of the electron in question

The energy states of electrons are specified by a set of quantum numbers, and the transition of energy states occurs only between those numbers so described. A set of quantum numbers specifies the high-energy orbit and low-energy orbit of an electron.

One device based on this principle is the *free electron laser* or the *ubitron* [13]. A schematic diagram of the device is given in Figure 6.87. A high-speed relativistic electron beam is emitted from an electron gun focused by a longitudinally applied dc magnetic flux density B and is periodically deflected by magnetic means. The repetitious deflections create among relativistic electrons high-energy states and low-energy states. If the high-energy electrons in the deflection waveguides are stimulated by the resonance frequency of the waveguide resonator, then the *downward transition* (the transition of electrons

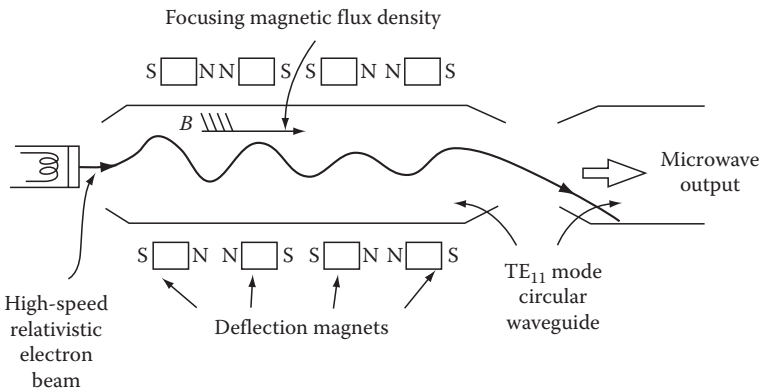


FIGURE 6.87 Basic structure of the ubitron. (From Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996. Used with permission.)

in a high-energy state to a low-energy state) occurs. Microwave emissions result because of the energy transitions at the stimulation frequency.

The ubitron operates at millimeter or submillimeter wave frequencies, and typically at high power levels.

6.8.2 Variations on the Klystron

Because the multiple cavities of a conventional klystron are difficult to fabricate at millimeter wavelengths, variations on the basis design have been produced [13]. Instead of using multiple cavity resonators, the cavities can instead be combined into a single cavity. Such a design can be accomplished through the use of ladder-shaped grids, as shown in Figure 6.88a by the dashed lines with the single cavity. This type of klystron is referred to as the *laddertron* [14]. Because of the ladder structure, a standing wave will exist on the ladder line. Velocity modulation results near the electron gun and bunching begins near the middle of the electron stream. Catching is done in the ladder structure near the collector. Because this device is a one-cavity resonator, there is built-in positive feedback. Oscillation, therefore, begins in a manner similar to a reflex klystron, also a single cavity tube.

At millimeter wavelengths, a dominant mode cavity resonator becomes very small and loses its power-handling capability. In a laddertron, however, the cavity resonator is a multiple higher-mode resonator. Therefore, the Q of the cavity is high and the size of the cavity resonator is reasonably large. This being the case, the laddertron can handle tens of watts of power at frequency as high as 50GHz.

The catcher of a two-cavity klystron amplifier can be replaced by the slow-wave structure of a TWT, as shown in Figure 6.88b. This type of tube is termed the *twystron* [14]. The slow-wave structure provides a broader frequency bandwidth than a regular two-cavity klystron. A microwave input fed into the buncher cavity results in velocity modulation of the electron beam. Bunching occurs while the beam drifts, and the bunched electrons induce a microwave voltage in the slow-wave structure. The electron beam is focused by the use of longitudinally applied magnetic flux density B .

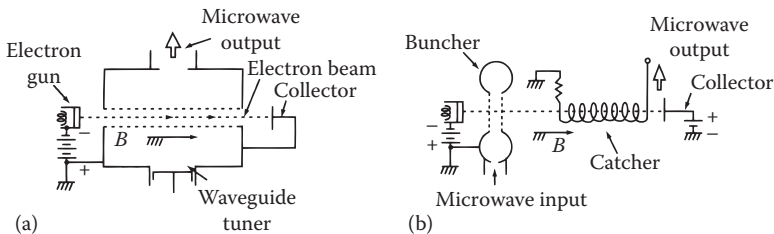


FIGURE 6.88

Functional schematic diagrams of modified klystrons: (a) laddertron and (b) twystron. (After Ishiik, T.K., Other microwave vacuum devices, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996. Used with permission.)

Commercial twystrons typically operate at 2–6 GHz and are used in pulsed radar transmitter power amplifiers. Pulsed peak output powers range from 1 to 7 MW (pulse width of 10–50 μ s) [10].

6.9 Microwave Tube Life

Any analysis of microwave tube life must first identify the parameters that define *life*. The primary *wear-out* mechanism in a microwave power tube is the electron gun at the cathode. In principle, the cathode eventually will evaporate the activating material and cease to produce the required output power. Tubes, however, rarely fail because of low emission, but for a variety of other reasons that are usually external to the device.

Power tubes designed for microwave applications provide long life when operated within their designed parameters. The point at which the device fails to produce the required output power can be predicted with some accuracy, based on design data and in-service experience. Most power tubes, however, fail because of mechanisms other than predictable chemical reactions inside the device itself. External forces, such as transient overvoltages caused by lightning, cooling system faults, and improper tuning, more often than not lead to the failure of a microwave tube.

6.9.1 Life-Support System

Transmitter control logic usually is configured for two states of operation:

- An *operational level*, which requires all of the “life-support” systems to be present before the high-voltage (HV) command is enabled
- An *overload level*, which removes HV when one or more fault conditions occur

The cooling system is the primary life-support element in most RF generators. The cooling system should be fully operational before the application of voltages to the tube. Likewise, a cool-down period usually is recommended between the removal of beam and filament voltages and shutdown of the cooling system.

Most microwave power tubes require a high-voltage removal time of less than 100 ms after the occurrence of an overload. If the trip time is longer, damage to the device may result. *Arc detectors* often are installed in the cavities of high-power tubes to sense fault conditions and shut down the high-voltage power supply before damage can be done to the tube. Depending on the circuit parameters, arcs can be sustaining, requiring

removal of high voltage to squelch the arc. A number of factors can cause RF arcing, including

- Overdrive condition
- Mistuning of one or more cavities
- Poor cavity fit (applies to external types only)
- Undercoupling of the output to the load
- Lightning strike at the antenna
- High VSWR

Regardless of the cause, arcing can destroy internal elements or the vacuum seal if drive and/or high voltage are not removed quickly. A lamp usually is included with each arc detector photocell for test purposes.

6.9.2 Protection Measures

A microwave power tube must be protected by control devices in the amplifier system. Such devices offer either visual indications or aural alarm warnings, or they actuate interlocks within the system. Figure 6.89 shows a klystron amplifier and the basic components associated with its

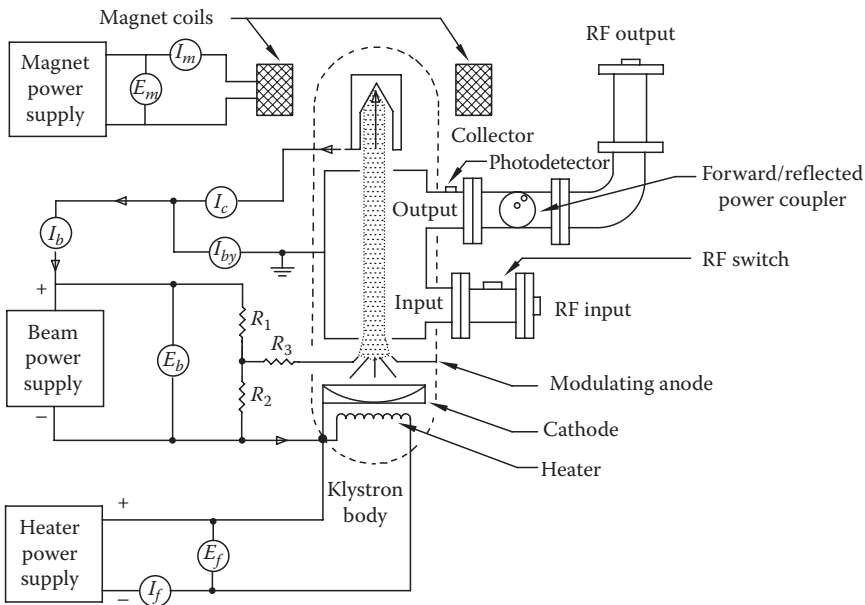


FIGURE 6.89

Protection and metering system for a klystron amplifier. (From Clayworth, G.T. et al., *Some exciting adventures in the IOT business*, in *NAB 1992 Broadcast Engineering Conference Proceedings*, National Association of Broadcasters, Washington, DC, pp. 200–208, 1992.)

operation, including metering for each of the power supplies. Other types of microwave power devices use similar protection schemes. Sections of coaxial transmission line, representing essential components, are shown in the figure attached to the RF input and RF output ports of the tube. A single magnetic coil is shown to represent any coil configuration that may exist; its position in the drawing is for convenience only and does not represent its true position in the system.

6.9.2.1 Heater Supply

The heater power supply can be either ac or dc. If it is dc, the positive terminal must be connected to the common heater-cathode terminal and the negative terminal to the second heater terminal. The amount of power supplied to the heater is important because it establishes the cathode operating temperature. The temperature must be high enough to provide ample electron emission, but not so high that emission life is jeopardized.

Because the cathode and heater are connected to the negative side of the beam supply, they must be insulated to withstand the full beam potential. This dictates the use of a filament transformer designed specifically for the specific application.

6.9.2.2 Beam Supply

The high-voltage beam supply furnishes the dc input power to the microwave device. In a klystron, the positive side of the beam supply is connected to the body and collector. The negative terminal is connected to the common heater-cathode terminal. The voltmeter, E_b in Figure 6.89, measures the beam voltage applied between the cathode and the body of the klystron.

Current meter I_c measures collector current, typically 95% or more of the total device current. Current meter I_{by} measures body current. An interlock should interrupt the beam supply if the body current exceeds a specified maximum value.

The sum of the body current (I_{by}) and collector current (I_c) is equal to the beam current (I_b), which should stay constant as long as the beam voltage and modulating-anode voltage are held constant.

6.9.2.3 Magnet Supply

Electrical connections to the dc magnet supply typically include two meters, one for measuring current through the circuit (I_m in Figure 6.89), and one for measuring voltage (E_m). When a microwave device is installed in its magnet assembly, both parameters should be measured and recorded for future reference. If excessive body current or other unusual symptoms occur, these data will be valuable for system analysis.

Undercurrent protection should be provided to remove beam voltage if the magnetic circuit current falls below a preset value. The interlock also should prevent the beam voltage from being applied if the magnetic circuit is not energized. This scheme, however, will not provide protection if windings of the magnet are short-circuited. Short-circuited conditions can be determined by measuring the normal values of voltage and current and recording them for future reference.

The body-current overload protection should actuate if the magnetic field is reduced for any reason.

6.9.2.4 RF Circuits

In Figure 6.89, monitoring devices are shown on the RF input and output of the klystron. These monitors protect the device in case of a failure in the RF output circuit. Two directional couplers and a photodetector are attached to the output of the device. These components and an RF switching device on the input form a protective network against output transmission line mismatch. The RF switch is activated by the photodetector or the reflected power monitor and must be capable of removing RF drive power from the tube in less than 10 ms (typically).

In the RF output circuit, the forward power coupler is used to measure the relative power output of the microwave device. The reflected power coupler measures the RF energy reflected by the output circuit components or antenna. Damaged components or foreign material in the RF line will increase the RF reflected power. The amount of reflected power should be no more than 5% of the actual forward RF output power of the tube in most applications. An interlock monitors the reflected power and removes RF drive to the device if the reflected energy reaches an unsafe level. To protect against arcs occurring between the monitor and the output window, a photodetector is placed between the monitor and the window. Light from an arc will trigger the photodetector, which actuates the interlock system to remove RF drive before the window is damaged.

6.9.3 Filament Voltage Control

Extending the life of a microwave tube begins with accurate adjustment of filament voltage, as noted in the previous section. The filament should not be operated at a reduced voltage in an effort to extend tube life, as in the case of the thoriated tungsten grid tube. In a microwave tube, reduced filament voltage may cause uneven emission from the surface of the cathode with little or no improvement in cathode life.

Voltage should be applied to the filament for a specified warm-up period before the application of beam current to minimize thermal stress on the cathode/gun structure. However, voltage should not be applied to the filaments for extended periods (typically 2h or more) if no beam voltage

is present. The net rate of evaporation of emissive material from the cathode surface is greater without beam voltage. Subsequent condensation of material on gun components may lead to voltage standoff problems.

6.9.4 Cooling System

The cooling system is vital to any RF generator. In a high-power microwave transmitter, the cooling system may need to dissipate as much as 70% of the input ac power in the form of waste heat. For vapor-phase-cooled devices, pure (distilled or demineralized) water must be used. Because the collector is usually only several volts above ground potential, it is generally not necessary to use deionized water.

6.9.5 Reliability Statistics

Determination of the *mean time between failures (MTBF)* of a microwave tube provides valuable information on the operation of a given device in a given installation. The following formulas can be used to predict tube life in operating years and to estimate the number of replacement tubes that will be needed during the life of an RF generator:

$$Y = \frac{MTBF}{H} \tag{6.4}$$

$$N = \frac{L \times S}{Y} \tag{6.5}$$

$$R = N - S \tag{6.6}$$

where

Y is the tube life in operating years

MTBF is the tube mean time between failures (gathered from manufacturer literature or on-site experience)

H are the hours of operation per year (= 365 × hours per operating day)

N are the number of tubes needed over the life of the transmitter

L is the anticipated life of the transmitter in years

S are the number of tubes per transmitter

R are the number of replacement tubes needed over the life of the transmitter

6.9.5.1 MTTR

The *mean time to repair (MTTR)* of a given piece of RF equipment is an important consideration for any type of facility. MTTR defines the *maintainability* of a system. In the case of a microwave generator, the time required to change a tube is an important factor, especially if standby equipment

is not available. The example time-change estimates given previously in this chapter (1 h for an integral-cavity klystron and 4–6 h for a 60 kW UHF external-cavity tube) assume that no preparation work has been performed on the spare device. Pretuning a tube is one way to shorten the MTTR for an external-cavity device. The spare tube is installed during a maintenance period, and the system is tuned for proper operation. After the positions of all tuning controls are documented, the tube is removed and returned to its storage container, along with the list of tuning control settings. In this way, the external-cavity tube can be placed into service much faster during an emergency. This procedure probably will not result in a tube tuned for optimum performance, but it may provide a level of performance that is acceptable on a temporary basis.

Consideration of MTTR is important for a facility because most microwave tube failure is caused by mechanisms other than reduced cathode emission, which is a *soft failure* that can be anticipated with some degree of accuracy. Catastrophic failures, on the other hand, offer little, if any, warning.

References

1. Ostroff, N. S., A unique solution to the design of an ATV transmitter, in *Proceedings of the 1996 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, p. 144, 1996.
2. Hulick, T. P., 60 kW Diacrode UHF TV transmitter design, performance and field report, in *Proceedings of the 1996 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, p. 442, 1996.
3. Varian Associates, *Integral Cavity Klystrons for UHF-TV Transmitters*, Palo Alto, CA, 1971.
4. Clayworth, G. T., H. P. Bohlen, and R. Heppinstall, Some exciting adventures in the IOT business, in *Proceedings of the 1992 NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 200–208, 1992.
5. Symons, R. S., The constant efficiency amplifier, in *Proceedings of the NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 523–530, 1997.
6. Priest, D. H. and M. B. Shrader, The klystrode—An unusual transmitting tube with potential for UHF-TV, *Proceedings of the IEEE*, 70(11), 1318–1325, 1982.
7. Gilmore, A. S., *Microwave Tubes*, Artech House, Dedham, MA, pp. 196–200, 1986.
8. Symons, R., M. Boyle, J. Cipolla, H. Schult, and R. True, The constant efficiency amplifier—A progress report, in *Proceedings of the NAB Broadcast Engineering Conference*, National Association of Broadcasters, Washington, DC, pp. 77–84, 1998.
9. Ishii, T. K., Traveling wave tubes, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 428–443, 1996.

10. Liao, S. Y., *Microwave Electron Tube Devices*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
11. Sims, G. D. and I. M. Stephenson, *Microwave Tubes and Semiconductor Devices*, Interscience, London, U.K., 1963.
12. Hutter, R. G. E., *Beam and Wave Electronics in Microwave Tubes*, Interscience, London, U.K., 1960.
13. Ishiik, T. K., Other microwave vacuum devices, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 444–457, 1996.
14. Ishii, T. K., *Microwave Engineering*, Harcourt-Brace-Jovanovich, San Diego, CA, 1989.
15. Coleman, J. T., *Microwave Devices*, Reston Publishing, Reston, VA, 1982.
16. McCune, E. W., Fission plasma heating with high-power microwave and millimeter wave tubes, *Journal of Microwave Power*, 20, 131–136, 1985.

Bibliography

- Badger, G., The klystron: A new high-efficiency UHF-TV power amplifier, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1986.
- Collins, G. B., *Radar System Engineering*, McGraw-Hill, New York, 1947.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, 1992.
- Dick, B., New developments in RF technology, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1986.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.
- Fisk, J. B., H. D. Hagstrum, and P. L. Hartman, The magnetron as a generator of centimeter waves, *Bell System Technical Journal*, 25, 167, 1946.
- Ginzton, E. L. and A. E. Harrison, Reflex klystron oscillators, *Proceedings of the IRE*, 34, 97, 1946.
- Institute of Electrical and Electronics Engineers, Inc., *IEEE Standard Dictionary of Electrical and Electronics Terms*, New York, 1984.
- McCune, E., Final report: The multi-stage depressed collector project, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1988.
- Ostrov, N., A. Kiesel, A. Whiteside, and A. See, Klystron-equipped UHF-TV transmitters: Report on the initial full service station installations, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1989.
- Ostrov, N., A. Whiteside, and L. Howard, An integrated exciter/pulsar system for ultra high-efficiency klystron operation, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1985.
- Ostrov, N., A. Whiteside, A. See, and A. Kiesel, A 120 kW klystron transmitter for full broadcast service, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1988.

- Pierce, J. R., Reflex oscillators, *Proceedings of the IRE*, 33, 112, 1945.
- Pierce, J. R., Theory of the beam-type traveling wave tube, *Proceedings of the IRE*, 35, 111, 1947.
- Pierce, J. R. and L. M. Field, Traveling-wave tubes, *Proceedings of the IRE*, 35, 108, 1947.
- Pond, N. H. and C. G. Lob, Fifty years ago today or on choosing a microwave tube, *Microwave Journal*, 226–238, September 1988.
- Shrader, M. B., Klystron technology update, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1988.
- Spangenberg, K., *Vacuum Tubes*, McGraw-Hill, New York, 1947.
- Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.
- Varian, R. and S. Varian, A high-frequency oscillator and amplifier, *Journal of Applied Physics*, 10, 321, 1939.
- Webster, D. L., Cathode bunching, *Journal of Applied Physics*, 10, 501, 1939.
- Whitaker, J. C. and T. Blankenship, Comparing integral and external cavity klystrons, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, November 1988.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1991.
- Whitaker, J. C. (ed.), *NAB Engineering Handbook*, 9th edn., National Association of Broadcasters, Washington, DC, 1998.

7

RF Interconnection and Switching

7.1 Introduction

The components that connect, interface, transfer, and filter radio frequency (RF) energy within a given system—or between systems—are critical elements in the operation of vacuum tube devices. Such hardware, usually passive, determines to a large extent the overall performance of the RF generator. To optimize the performance of power vacuum devices, first it is necessary to understand and optimize the components upon which the tube depends.

The mechanical and electrical characteristics of the transmission line, waveguide, and associated hardware that carry power from a power source (usually a transmitter) to the load (usually an antenna) are critical to proper operation of any RF system. Mechanical considerations determine the ability of the components to withstand temperature extremes, lightning, rain, and wind. In other words, they determine the overall reliability of the system.

At high frequencies where the wavelength becomes small with respect to feature size, it is necessary to consider an electronic signal as an electromagnetic wave and the structure where this signal exists as a waveguide.

7.1.1 Skin Effect

The effective resistance offered by a given conductor to RFs is considerably higher than the ohmic resistance measured with direct current. This is because of an action known as the *skin effect*, which causes the currents to be concentrated in certain parts of the conductor and leaves the remainder of the cross section to contribute little or nothing toward carrying the applied current.

When a conductor carries an alternating current, a magnetic field is produced that surrounds the wire. This field continually expands and contracts as the ac wave increases from zero to its maximum positive value and back to zero, then through its negative half-cycle. The changing magnetic lines of force cutting the conductor induce a voltage in the conductor in a direction

that tends to retard the normal flow of current in the wire. This effect is more pronounced at the center of the conductor. Thus, current within the conductor tends to flow more easily toward the surface of the wire; the higher the frequency, the greater the tendency for current to flow at the surface. The depth of current flow is a function of frequency, and it is determined from the equation

$$d = \frac{2.6}{\sqrt{\mu f}} \quad (7.1)$$

where

d is the depth of current in mil

μ is the permeability (copper = 1, steel = 300)

f is the frequency of signal in MHz

It can be calculated that at a frequency of 100 kHz, current flow penetrates a conductor by 8 mils. At 1 MHz, the skin effect causes current to travel in only the top 2.6 mils in copper and even less in almost all other conductors. Therefore, the series impedance of conductors at high frequencies is significantly higher than at low frequencies. Figure 7.1 shows the distribution of current in a radial conductor.

When a circuit is operating at high frequencies, the skin effect causes the current to be redistributed over the conductor cross section in such a way as to make most of the current flow where it is encircled by the smallest number of flux lines. This general principle controls the distribution of current regardless of the shape of the conductor involved. With a flat-strip conductor, the current flows primarily along the edges, where it is surrounded by the smallest amount of flux.

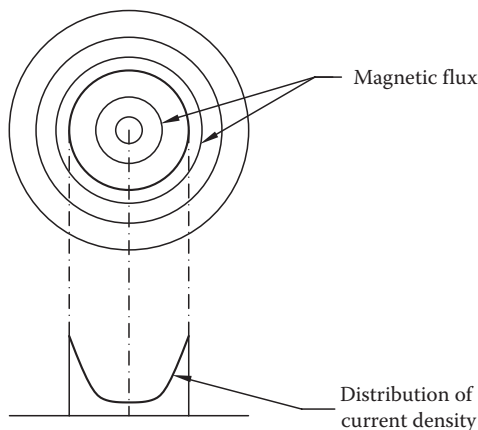


FIGURE 7.1

Skin effect on an isolated round conductor carrying a moderately high frequency signal.

7.2 Coaxial Transmission Line

Two types of coaxial transmission line are in common use today in moderate-to-high power systems: *rigid line* and corrugated (*semiflexible*) line. Rigid coaxial cable is constructed of heavy-wall copper tubes with Teflon or ceramic spacers. (Teflon is a registered trademark of DuPont.) Rigid line provides electrical performance approaching an ideal transmission line, including

- High power-handling capability
- Low loss
- Low VSWR (*voltage standing wave ratio*)

Rigid transmission line is, however, expensive to purchase and install.

The primary alternative to rigid coax is semiflexible transmission line made of corrugated outer and inner conductor tubes with a spiral polyethylene (or Teflon) insulator. The internal construction of a semiflexible line is shown in Figure 7.2. Semiflexible line has four primary benefits:

- It is manufactured in a continuous length, rather than the 20 ft sections typically used for rigid line.
- Because of the corrugated construction, the line may be shaped as required for routing from the transmitter to the antenna.

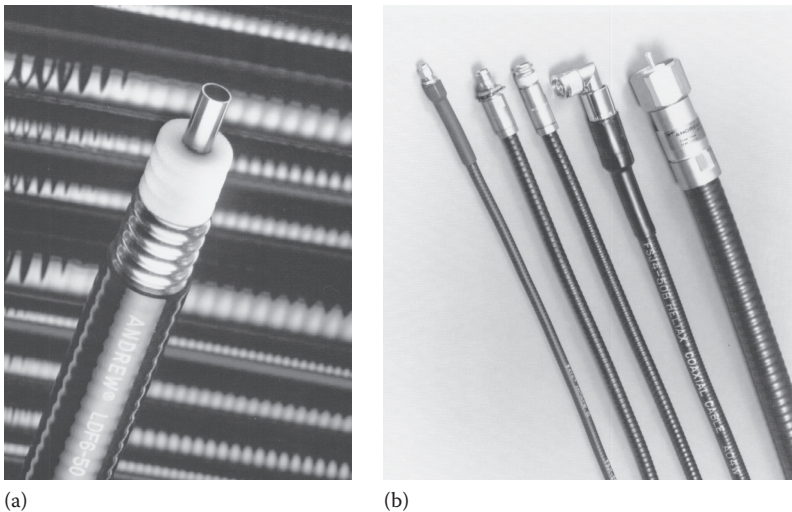


FIGURE 7.2

Semiflexible coaxial cable: (a) a section of cable showing the basic construction and (b) cable with various terminations. (Courtesy of Andrew, Orland Park, IL.)

- The corrugated construction permits differential expansion of the outer and inner conductors.
- Each size of line has a minimum bending radius. For most installations, the flexible nature of corrugated line permits the use of a single piece of cable from the transmitter to the antenna, with no elbows or other transition elements. This speeds installation and provides for a more reliable system, all other things being equal.

7.2.1 Electrical Parameters

A signal traveling in free space is unimpeded; it has a free-space velocity equal to the speed of light. In a transmission line, capacitance and inductance slow the signal as it propagates along the line. The degree to which the signal is slowed is represented as a percentage of the free-space velocity. This quantity is called the *relative velocity of propagation* and is described by the equation

$$V_p = \frac{1}{\sqrt{L \times C}} \quad (7.2)$$

where

L is the inductance in H/ft

C is the capacitance in F/ft

Furthermore,

$$V_r = \frac{V_p}{c} \times 100\% \quad (7.3)$$

where

V_p is the velocity of propagation

c is 9.842×10^8 ft/s (free-space velocity)

V_r is the velocity of propagation as a percentage of free-space velocity

Losses in a coaxial cable arise from two primary sources:

- The resistance of the conductors and the currents flowing in the center and outer conductors
- The conduction current in the dielectric

The conductor losses are ohmic and, due to the skin effect, increase with frequency. The dielectric losses result from the resistance of the dielectric material to the conduction currents. A third type of loss in a coaxial cable is radiation. This effect is usually minimal.

7.2.1.1 Transverse Electromagnetic Mode

The principal mode of propagation in a coaxial line is the *transverse electromagnetic mode* (TEM). This mode will not propagate in a waveguide, and

that is why coaxial lines can propagate a broadband of frequencies efficiently. The cutoff frequency for a coaxial transmission line is determined by the line dimensions. Above cutoff, modes other than TEM can exist and the transmission properties are no longer defined. The cutoff frequency is equivalent to

$$F_c = \frac{7.50 \times V_r}{D_i + D_o} \quad (7.4)$$

where

F_c is the cutoff frequency in GHz

V_r is the velocity (%)

D_i is the inner diameter of outer conductor in in.

D_o is the outer diameter of inner conductor in in.

At dc, current in a conductor flows with uniform density over the cross section of the conductor. At high frequencies, the current is displaced to the conductor surface. The effective cross section of the conductor decreases, and the conductor resistance increases because of the skin effect.

Center conductors are made from copper-clad aluminum or high-purity copper and can be solid, hollow tubular, or corrugated tubular. Solid center conductors are found on semiflexible cable with 1/2 in. or smaller diameter. Tubular conductors are found in 7/8 in. or larger-diameter cables. Although the tubular center conductor is used primarily to maintain flexibility, it also can be used to pressurize an antenna through the feeder.

7.2.1.2 Dielectric

Coaxial lines use two types of dielectric construction to isolate the inner conductor from the outer conductor. The first is an air dielectric, with the inner conductor supported by a dielectric spacer and the remaining volume filled with air or nitrogen gas. The spacer, which may be constructed of spiral or discrete rings, typically is made of Teflon or polyethylene. Air-dielectric cable offers lower attenuation and higher average power ratings than foam-filled cable but requires pressurization to prevent moisture entry.

Foam-dielectric cables are ideal for use as feeders with antennas that do not require pressurization. The center conductor is surrounded completely by foam-dielectric material, resulting in a high dielectric breakdown level. The dielectric materials are polyethylene-based formulations, which contain antioxidants to reduce dielectric deterioration at high temperatures.

7.2.1.3 Impedance

The expression *transmission line impedance* applied to a point on a transmission line signifies the vector ratio of line voltage to line current at that

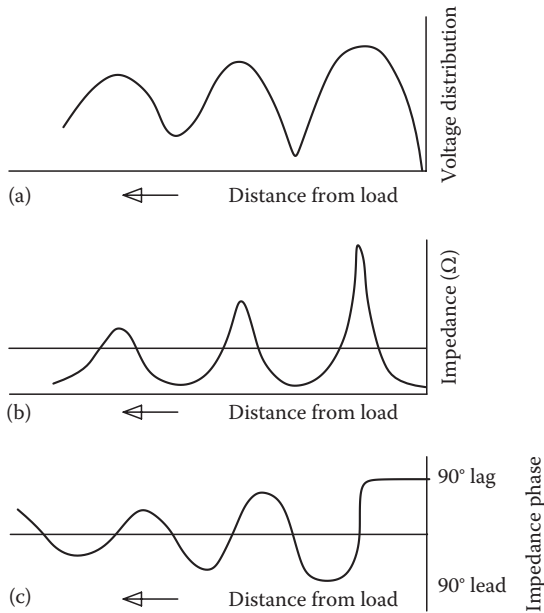


FIGURE 7.3

Magnitude and power factor of line impedance with increasing distance from the load for the case of a short-circuited receiver and a line with moderate attenuation: (a) voltage distribution, (b) impedance magnitude, and (c) impedance phase.

particular point. This is the impedance that would be obtained if the transmission line was cut at the point in question, and the impedance looking toward the receiver was measured.

Because the voltage and current distribution on a line are such that the current tends to be small when the voltage is large (and vice versa), as shown in Figure 7.3, the impedance will, in general, be oscillatory in the same manner as the voltage (large when the voltage is high and small when the voltage is low). Thus, in the case of a short-circuited receiver, the impedance will be high at distances from the receiving end that are odd multiples of 1/4 wavelength, and it will be low at distances that are even multiples of 1/4 wavelength.

The extent to which the impedance fluctuates with distance depends on the standing wave ratio (ratio of reflected to incident waves), being less as the reflected wave is proportionally smaller than the incident wave. In the particular case where the load impedance equals the characteristic impedance, the impedance of the transmission line is equal to the characteristic impedance at all points along the line.

The power factor of the impedance of a transmission line varies according to the standing waves present. When the load impedance equals the characteristic impedance, there is no reflected wave, and the power factor of the

impedance is equal to the power factor of the characteristic impedance. At RFs, the power factor under these conditions is accordingly resistive. However, when a reflected wave is present, the power factor is unity (resistive) only at the points on the line where the voltage passes through a maximum or a minimum. At other points the power factor will be reactive, alternating from leading to lagging at intervals of $1/4$ wavelength. When the line is short-circuited at the receiver, or when it has a resistive load less than the characteristic impedance so that the voltage distribution is of the short-circuit type, the power factor is inductive for lengths corresponding to less than the distance to the first voltage maximum. Thereafter, it alternates between capacitive and inductive at intervals of $1/4$ wavelength. Similarly, with an open-circuited receiver or with a resistive load greater than the characteristic impedance so that the voltage distribution is of the open-circuit type, the power factor is capacitive for lengths corresponding to less than the distance to the first voltage minimum. Thereafter, the power factor alternates between capacitive and inductive at intervals of $1/4$ wavelength, as in the short-circuited case.

7.2.1.4 Resonant Characteristics

A transmission line can be used to perform the functions of a resonant circuit. If the line, for example, is short-circuited at the receiver, at frequencies in the vicinity of a frequency at which the line is an odd number of $1/4$ wavelengths long, the impedance will be high and will vary with frequency in the vicinity of resonance. This characteristic is similar in nature to a conventional parallel resonant circuit. The difference is that with the transmission line, there are a number of resonant frequencies, one for each of the infinite number of frequencies that make the line an odd number of $1/4$ wavelengths long. At very high frequency (VHF), the parallel impedance at resonance and the circuit Q obtainable are far higher than can be realized with lumped circuits. Behavior corresponding to that of a series resonant circuit can be obtained from a transmission line that is an odd number of $1/4$ wavelengths long and open-circuited at the receiver.

Transmission lines also can be used to provide low-loss inductances or capacitances if the proper combination of length, frequency, and termination is employed. Thus, a line short-circuited at the receiver will offer an inductive reactance when less than $1/4$ wavelength and a capacitive reactance when between $1/4$ and $1/2$ wavelengths. With an open-circuited receiver, the conditions for inductive and capacitive reactances are reversed.

7.2.2 Electrical Considerations

VSWR, attenuation, and power-handling capability are key electrical factors in the application of coaxial cable. High VSWR can cause power loss,

voltage breakdown, and thermal degradation of the line. High attenuation means less power delivered to the antenna, higher power consumption at the transmitter, and increased heating of the transmission line itself.

VSWR is a common measure of the quality of a coaxial cable. High VSWR indicates nonuniformities in the cable that can be caused by one or more of the following conditions:

- Variations in the dielectric core diameter
- Variations in the outer conductor
- Poor concentricity of the inner conductor
- Nonhomogeneous or periodic dielectric core

Although each of these conditions may contribute only a small reflection, they can add up to a measurable VSWR at a particular frequency.

Rigid transmission line typically is available in a standard length of 20 ft, and in alternative lengths of 19.5 and 19.75 ft. The shorter lines are used to avoid VSWR buildup caused by discontinuities resulting from the physical spacing between line section joints. If the section length selected and the operating frequency have a 1/2-wave correlation, the connector junction discontinuities will add. This effect is known as *flange buildup*. The result can be excessive VSWR. The critical frequency at which a 1/2-wave relationship exists is given by

$$F_{cr} = \frac{490.4 \times n}{L} \quad (7.5)$$

where

F_{cr} is the critical frequency

n is any integer

L is the transmission line length in ft

For most applications, the critical frequency for a chosen line length should not fall closer than ± 2 MHz of the passband at the operating frequency.

Attenuation is related to the construction of the cable itself and varies with frequency, product dimensions, and dielectric constant. Larger-diameter cable exhibits lower attenuation than smaller-diameter cable of similar construction when operated at the same frequency. It follows, therefore, that larger-diameter cables should be used for long runs.

Air-dielectric coax exhibits less attenuation than comparable-size foam-dielectric cable. The attenuation characteristic of a given cable also is affected by standing waves present on the line resulting from an impedance mismatch. Table 7.1 shows a representative sampling of semiflexible coaxial cable specifications for a variety of line sizes.

TABLE 7.1

Representative Specifications for Various Types of Flexible Air-Dielectric Coaxial Cable

Cable Size (in.)	Maximum Frequency (MHz)	Velocity (%)	Peak Power 1 MHz (kW)	Average Power		Attenuation ^a	
				100 MHz (kW)	1 MHz (kW)	100 MHz (dB)	1 MHz (dB)
1-5/8	2.7	92.1	145	145	14.4	0.020	0.207
3	1.64	93.3	320	320	37	0.013	0.14
4	1.22	92	490	490	56	0.010	0.113
5	0.96	93.1	765	765	73	0.007	0.079

^a Attenuation specified in dB/100 ft.

7.2.3 Coaxial Cable Ratings

Selection of a type and size of transmission line is determined by a number of parameters, including power-handling capability, attenuation, and phase stability.

7.2.3.1 Power Rating

Both *peak* and *average* power ratings are required to fully describe the capabilities of a transmission line. In most applications, the peak power rating limits the low frequency or pulse energy, and the average power rating limits high-frequency applications, as shown in Figure 7.4. Peak power ratings usually are stated for the following conditions:

- VSWR = 1.0
- Zero modulation
- One atmosphere of absolute dry air pressure at sea level

The peak power rating of a selected cable must be greater than the following expression, in addition to satisfying the average-power-handling criteria:

$$E_{pk} > P_t \times (1 + M)^2 \times \text{VSWR} \quad (7.6)$$

where

E_{pk} is the cable peak power rating in kW

P_t is the transmitter power in kW

M is the amplitude modulation percentage expressed decimally (100% = 1.0)

VSWR is the voltage standing wave ratio

From this equation, it can be seen that 100% amplitude modulation will increase the peak power in the transmission line by a factor of 4. Furthermore, the peak power in the transmission line increases directly with VSWR.

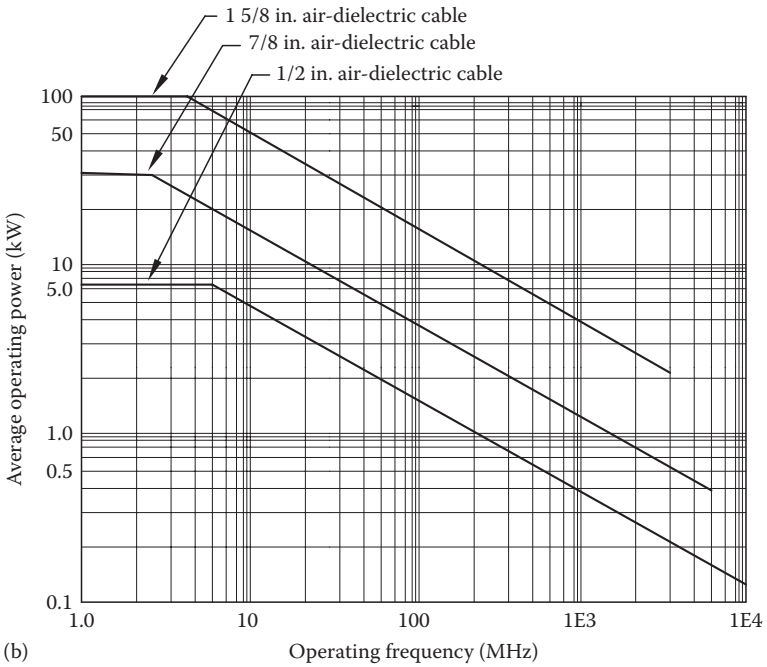
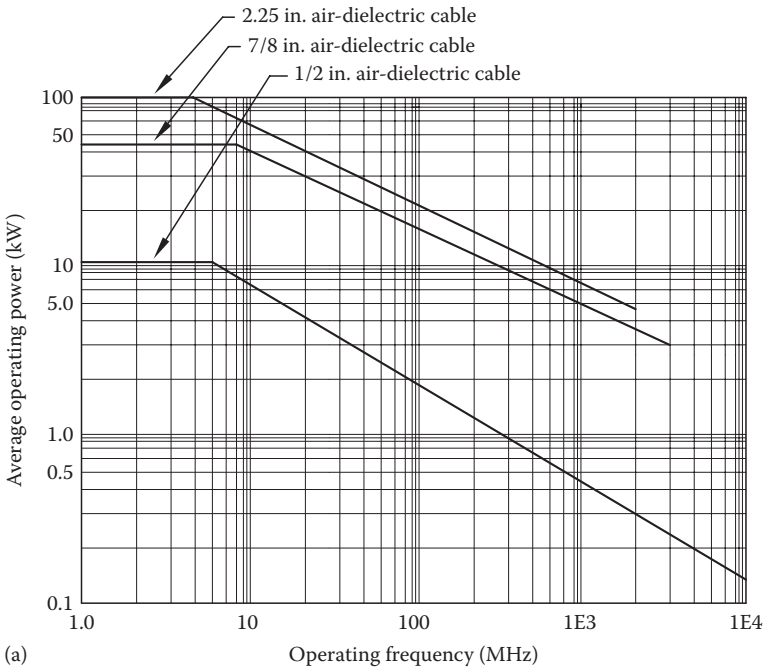


FIGURE 7.4 Power rating data for a variety of coaxial transmission lines: (a) 50 Ω line and (b) 75 Ω line.

The peak power rating is limited by the voltage breakdown potential between the inner and outer conductors of the line. The breakdown point is independent of frequency. It varies, however, with the line pressure (for an air-dielectric cable) and the type of pressurizing gas.

The average power rating of a transmission line is limited by the safe long-term operating temperature of the inner conductor and the dielectric. Excessive temperatures on the inner conductor will cause the dielectric material to soften, leading to mechanical instability inside the line.

The primary purpose of pressurization of an air-dielectric cable is to prevent the ingress of moisture. If allowed to accumulate in the line, moisture can increase attenuation and reduce the breakdown voltage between the inner and outer conductors. Pressurization with high-density gases can be used to increase both the average power and the peak power ratings of a transmission line. For a given line pressure, the increased power rating is more significant for peak power than for average power. High-density gases used for such applications include Freon 116 and sulfur hexafluoride. Figure 7.5 illustrates the effects of pressurization on cable power rating.

An adequate safety factor is necessary for peak and average power ratings. Most transmission lines are tested at two or more times their rated peak power before shipment to the customer. This safety factor is intended as a provision for transmitter transients, lightning-induced effects, and high-voltage excursions resulting from unforeseen operating conditions.

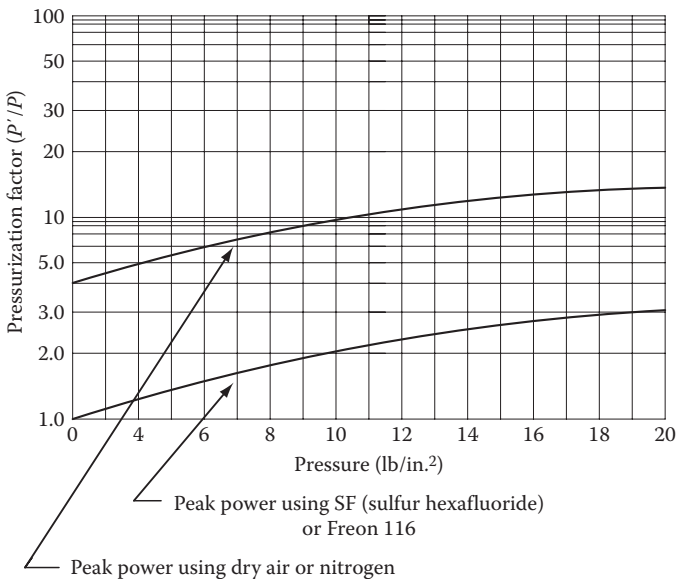


FIGURE 7.5

Effects of transmission line pressurization on peak power rating. Note that P' is the rating of the line at the increased pressure, and P is the rating of the line at atmospheric pressure.

7.2.3.2 Connector Effects

Foam-dielectric cables typically have a greater dielectric strength than air-dielectric cables of similar size. For this reason, foam cables would be expected to exhibit higher peak power ratings than air lines. Higher values, however, usually cannot be realized in practice because the connectors commonly used for foam cables have air spaces at the cable/connector interface that limit the allowable RF voltage to "air cable" values.

The peak-power-handling capability of a transmission line is the smaller of the values for the cable and the connectors attached to it. Table 7.2 lists the peak power ratings of several common RF connectors at standard conditions (defined in the previous section).

7.2.3.3 Attenuation

The attenuation characteristics of a transmission line vary as a function of the operating frequency and the size of the line itself. The relationships are shown in Figure 7.6.

The *efficiency* of a transmission line dictates how much power output by the transmitter actually reaches the antenna. Efficiency is determined by the length of the line and the attenuation per unit length.

The attenuation of a coaxial transmission line is defined by the equation

$$\alpha = 10 \times \log \left\{ \frac{P_1}{P_2} \right\} \quad (7.7)$$

where

α is attenuation in dB/100 m

P_1 is the input power into a 100 m line terminated with the nominal value of its characteristic impedance

P_2 is the power measured at the end of the line

TABLE 7.2
Electrical Characteristics of Common RF Connectors

Connector Type	DC Test Voltage (kW)	Peak Power (kW)
SMA	1.0	1.2
BNC, TNC	1.5	2.8
N, UHF	2.0	4.9
GR	3.0	11
HN, $\frac{7}{16}$	4.0	20
LC	5.0	31
7/8 EIA, F Flange	6.0	44
1-5/8 EIA	11.0	150
3-1/8 EIA	19.0	4

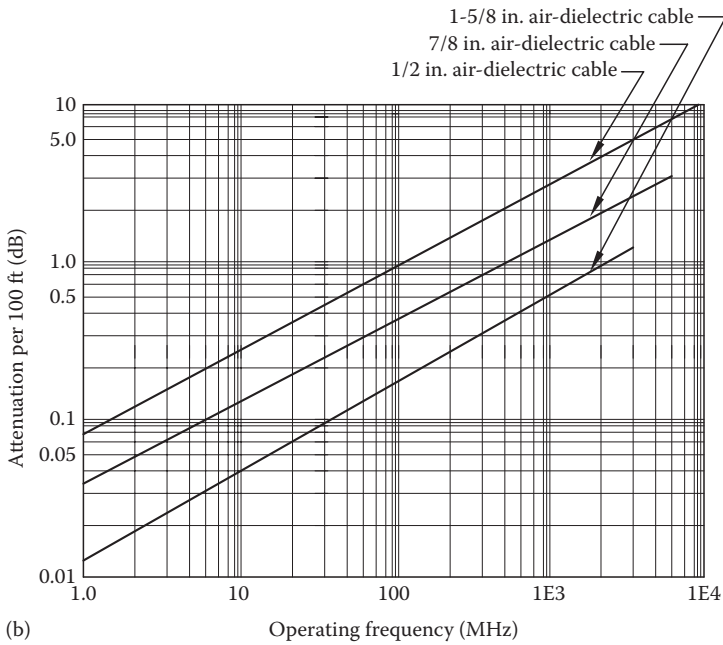
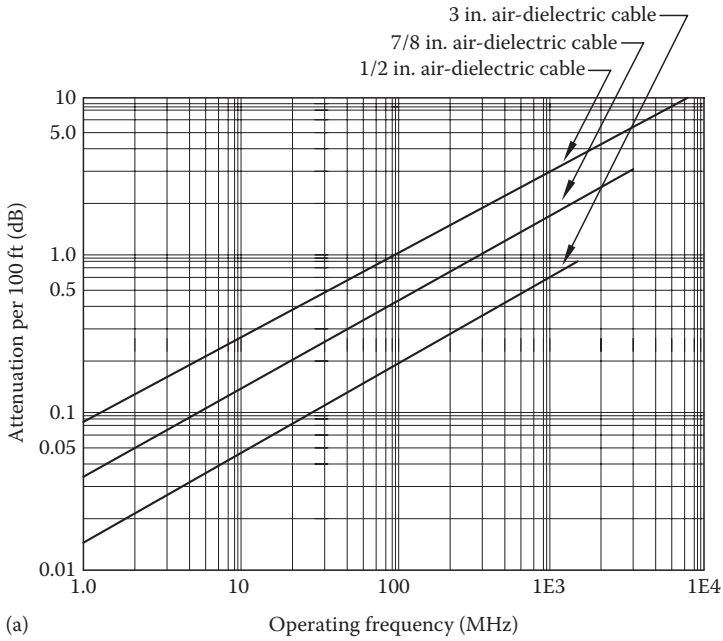


FIGURE 7.6 Attenuation characteristics for a selection of coaxial cables: (a) 50 Ω line and (b) 75 Ω line.

Stated in terms of efficiency (E , %),

$$E = 100 \times \left\{ \frac{P_o}{P_i} \right\} \tag{7.8}$$

where

P_i is the power delivered to the input of the transmission line

P_o is the power delivered to the antenna

The relationship between efficiency and loss in decibels (*insertion loss*) is illustrated in Figure 7.7.

Manufacturer-supplied attenuation curves typically are guaranteed to within approximately $\pm 5\%$. The values given usually are rated at 24°C (75°F)

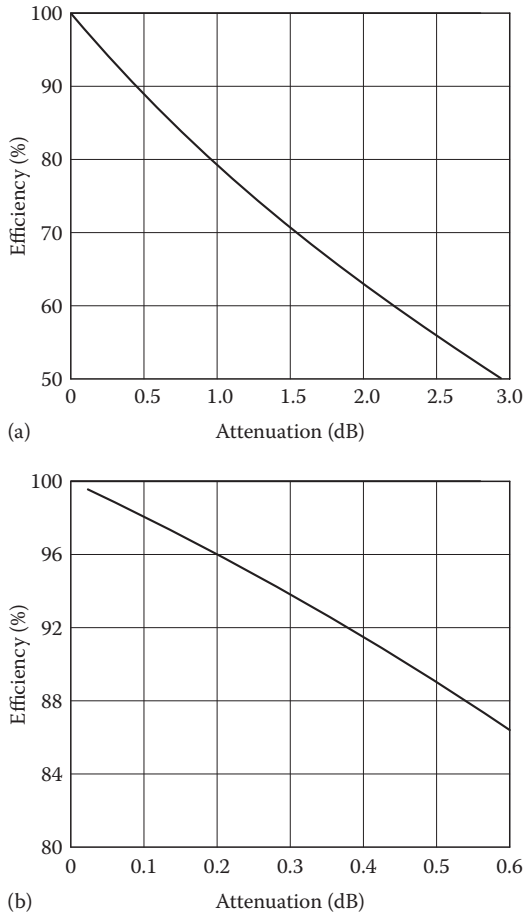


FIGURE 7.7

Conversion chart showing the relationship between decibel loss and efficiency of a transmission line: (a) high-loss line and (b) low-loss line.

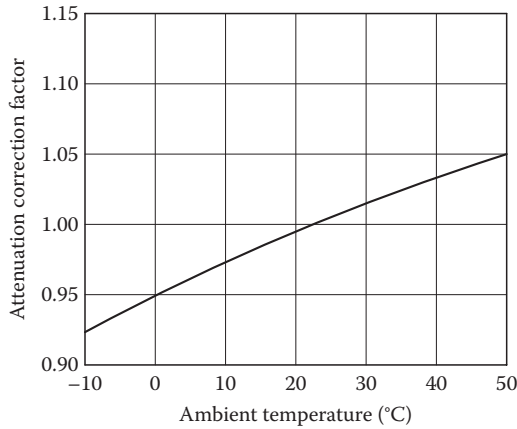


FIGURE 7.8
Variation of coaxial cable attenuation as a function of ambient temperature.

ambient temperature. Attenuation increases slightly with higher temperature or applied power. The effects of ambient temperature on attenuation are illustrated in Figure 7.8.

Loss in connectors is negligible, except for small (SMA and TNC) connectors at frequencies of several gigahertz and higher. Small connectors used at high frequencies typically add 0.1 dB of loss per connector.

When a transmission line is attached to a load, such as an antenna, the VSWR of the load increases the total transmission loss of the system. This effect is small under conditions of low VSWR. Figure 7.9 illustrates the interdependence of these two elements.

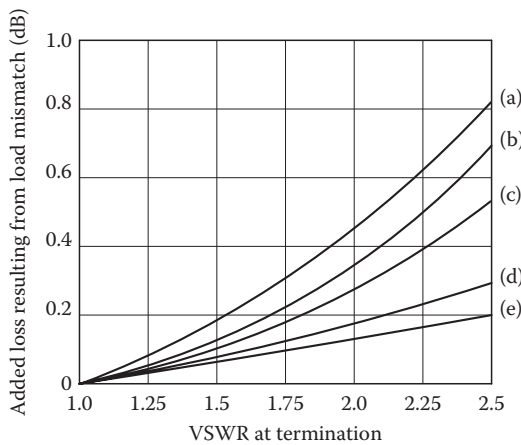


FIGURE 7.9
Effect of load VSWR on transmission line loss (a) 6 dB normal line attenuation, (b) 3 dB normal line attenuation, (c) 2 dB normal line attenuation, (d) 1 dB normal line attenuation, and (e) 0.5 dB normal line attenuation.

7.2.3.4 Phase Stability

A coaxial cable expands as the temperature of the cable increases, causing the electrical length of the line to increase as well. This factor results in phase changes that are a function of operating temperature. The phase relationship can be described by the equation

$$\theta = 3.66 \times 10^{-7} \times P \times L \times T \times F \quad (7.9)$$

where

θ is the phase change in deg.

P is the phase-temperature coefficient of the cable

L is the length of coax in ft

T is the temperature range (minimum-to-maximum operating temperature)

F is the frequency in MHz

Phase changes that are a function of temperature are important in systems utilizing multiple transmission lines, such as a directional array fed from a single phasing source. To maintain proper operating parameters, the phase changes of the cables must be minimized. Specially designed coaxial cables that offer low-phase-temperature characteristics are available. Two types of coax are commonly used for this purpose:

- *Phase-stabilized* cables, which have undergone extensive temperature cycling until such time as they exhibit their minimum phase-temperature coefficient
- *Phase-compensated* cables, in which changes in the electrical length have been minimized through adjustment of the mechanical properties of the dielectric and inner/outer conductors

7.2.4 Mechanical Parameters

Corrugated copper cables are designed to withstand bending with no change in properties. Low-density foam- and air-dielectric cables generally have a minimum bending radius of 10 times the cable diameter. Super flexible versions provide a smaller allowable bending radius.

Rigid transmission line will not tolerate bending. Instead, transition elements (elbows) of various sizes are used. Individual sections of rigid line are secured by multiple bolts around the circumference of a coupling flange.

When a large cable must be used to meet attenuation requirements, short lengths of a smaller cable (jumpers or *pigtails*) may be used on either end for ease of installation in low-power systems. The trade-off is slightly higher attenuation and some additional cost.

The *tensile strength* of a cable is defined as the axial load that may be applied to the line with no more than 0.2% permanent deformation after the load is released. When divided by the weight per foot of cable, this gives

an indication of the maximum length of cable that is self-supporting and therefore can be installed readily on a tower with a single hoisting grip. This consideration usually applies only to long runs of corrugated line; rigid line is installed one section at a time.

The *crush strength* of a cable is defined as the maximum force per linear inch that may be applied by a flat plate without causing more than a 5% deformation of the cable diameter. Crush strength is a good indicator of the ruggedness of a cable and its ability to withstand rough handling during installation.

Cable jacketing affords mechanical protection during installation and service. Semiflexible cables typically are supplied with a jacket consisting of low-density polyethylene blended with 3% carbon black for protection from the sun's ultraviolet rays, which can degrade plastics over time. This approach has proved to be effective, yielding a life expectancy of more than 20 years. Rigid transmission line has no covering over the outer conductor.

For indoor applications, where fire-retardant properties are required, cables can be supplied with a fire-retardant jacket, usually listed by Underwriters Laboratories. Note that under the provisions of the National Electrical Code, outside plant cables such as standard black polyethylene-jacketed coaxial line may be run a certain distance inside a building with no additional protection. The line also may be placed in conduit for longer runs.

Low-density foam cables are designed to prevent water from traveling along their length, should it enter through damage to the connector or the cable sheath. This is accomplished by mechanically locking the outer conductor to the foam dielectric by annular corrugations. Annular or ring corrugations, unlike helical or screw-thread-type corrugations, provide a water block at each corrugation. Closed-cell polyethylene dielectric foam is bonded to the inner conductor, completing the moisture seal.

A coaxial cable line is only as good as the connectors used to tie it together. The connector interface must provide a weatherproof bond with the cable to prevent water from penetrating the connection. This is ensured by using O-ring seals. The cable-connector interface also must provide a good electrical bond that does not introduce a mismatch and increase VSWR. Good electrical contact between the connector and the cable ensures that proper RF shielding is maintained.

7.3 Waveguide

As the operating frequency of a system reaches into the ultrahigh frequency (UHF) band, waveguide-based transmission line systems become practical, and depending on the frequency, necessary for high-power operation. A waveguide is any structure that guides an electromagnetic wave. The term

commonly refers to a closed metallic structure with a fixed cross section within which a guided wave propagates. From the mechanical standpoint, waveguide is simplicity itself. There is no inner conductor; RF energy is *launched* into the structure and propagates to the load. Several types of waveguide are available, including rectangular, square, circular, and elliptical. Waveguide offers several advantages over coax. First, unlike coax, waveguide can carry more power as the operating frequency increases. Second, efficiency is significantly better with waveguide at higher frequencies.

Rectangular waveguide commonly is used in high-power transmission systems. Circular waveguide also may be used, especially for applications requiring a cylindrical member, such as a rotating joint for an antenna feed. The physical dimensions of the guide are selected to provide for propagation in the *dominant* (lowest-order) mode.

Waveguide is not without its drawbacks, however. Rectangular or square guide constitutes a large windload surface, which places significant structural demands on a tower. Because of the physical configuration of rectangular and square guide, pressurization is limited, depending on the type of waveguide used (0.5 psi is typical). Excessive pressure can deform the guide shape and result in increased VSWR. Wind also may cause deformation and ensuing VSWR problems. These considerations have led to the development of circular and elliptical waveguide.

7.3.1 Propagation Modes

In an enclosed waveguide, an infinite number of distinct solutions exist, each one of these solutions being referred to as a *waveguide mode*. At any given operating frequency, the cross section of the waveguide and the type of material in the waveguide determine the characteristics of these modes. Propagation modes for waveguide most commonly fall into two broad categories:

- *Transverse-electric* (TE) waves
- *Transverse-magnetic* (TM) waves

With TE waves, the electric vector (*E vector*) is perpendicular to the direction of propagation. With TM waves, the magnetic vector (*H vector*) is perpendicular to the direction of propagation. These propagation modes take on integers (from 0 or 1 to infinity) that define field configurations. Only a limited number of these modes can be propagated, depending on the dimensions of the guide and the operating frequency.

Energy cannot propagate in waveguide unless the operating frequency is above the *cutoff frequency*. The cutoff frequency for rectangular guide is

$$F_c = \frac{c}{2 \times a} \quad (7.10)$$

where

- F_c is the waveguide cutoff frequency
- c is 1.179×10^{10} in./s (the velocity of light)
- a is the wide dimension of the guide

The cutoff frequency for circular waveguide is defined by the equation

$$F_c = \frac{c}{3.14 \times a'} \quad (7.11)$$

where a' is the radius of the guide.

There are four common propagation modes in waveguide:

- $TE_{0,1}$, the principal mode in rectangular waveguide.
- $TE_{1,0}$, also used in rectangular waveguide.
- $TE_{1,1}$, the principal mode in circular waveguide. $TE_{1,1}$ develops a complex propagation pattern with electric vectors curving inside the guide. This mode exhibits the lowest cutoff frequency of all modes, which allows a smaller guide diameter for a specified operating frequency.
- $TM_{0,1}$, which has a slightly higher cutoff frequency than $TE_{1,1}$ for the same size guide. Developed as a result of discontinuities in the waveguide, such as flanges and transitions, $TM_{0,1}$ energy is not coupled out by either dominant or cross-polar transitions. The parasitic energy must be filtered out, or the waveguide diameter chosen carefully to reduce the unwanted mode.

The field configuration for the dominant mode in rectangular waveguide is illustrated in Figure 7.10. Note that the electric field is vertical, with intensity maximum at the center of the guide and dropping off sinusoidally to zero intensity at the edges. The magnetic field is in the form of loops that lie in planes that are at right angles to the electric field (parallel to the top and bottom of the guide). The magnetic field distribution is the same for all planes perpendicular to the Y -axis. In the X direction, the intensity of the component of magnetic field that is transverse to the axis of the waveguide (the component in the direction of X) is at any point in the waveguide directly proportional to the intensity of the electric field at that point. This entire configuration of fields travels in the direction of the waveguide axis (the Z direction in Figure 7.10).

The field configuration for the $TE_{1,1}$ mode in circular waveguide is illustrated in Figure 7.11. The $TE_{1,1}$ mode has the longest cutoff wavelength and is, accordingly, the dominant mode. The next higher mode is the $TM_{0,1}$, followed by $TE_{2,1}$.

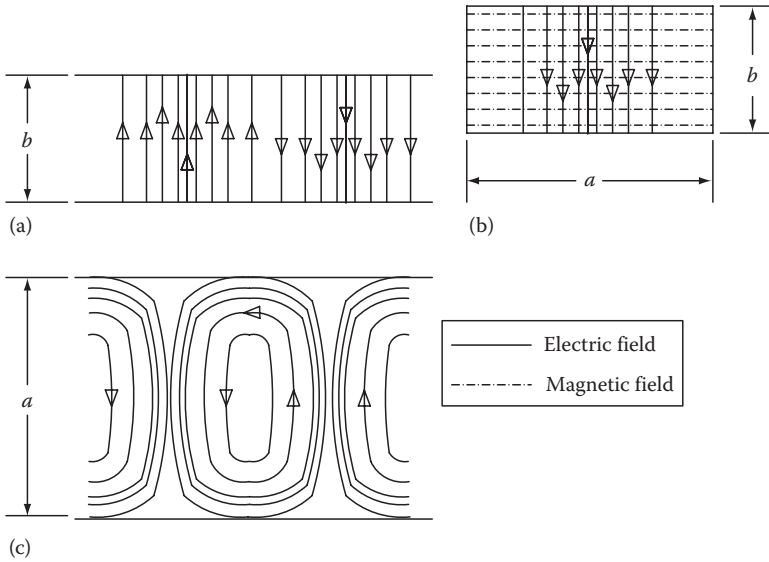


FIGURE 7.10 Field configuration of the dominant or $TE_{1,0}$ mode in a rectangular waveguide: (a) side view, (b) end view, and (c) top view.

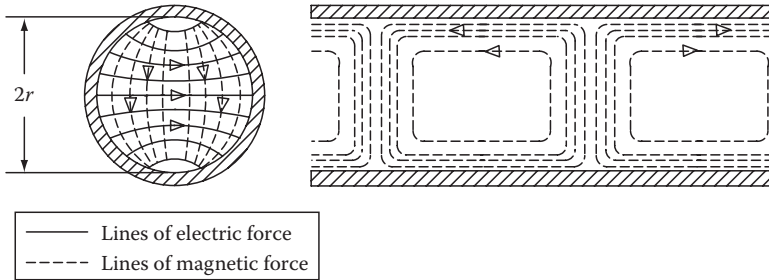


FIGURE 7.11 Field configuration of the dominant mode in circular waveguide.

7.3.1.1 Dual-Polarity Waveguide

Waveguide will support dual-polarity transmission within a single run of line. A combining element (*dual-polarized transition*) is used at the beginning of the run, and a splitter (*polarized transition*) is used at the end of the line. Square waveguide has found numerous applications in such systems. Theoretically, the $TE_{1,0}$ and $TE_{0,1}$ modes are capable of propagation without cross-coupling, at the same frequency, in lossless waveguide of square cross section. In practice, surface irregularities, manufacturing tolerances, and wall losses give rise to $TE_{1,0}$ - and $TE_{0,1}$ -mode cross-conversion. Because this conversion occurs continuously along the waveguide, long guide runs usually are avoided in dual-polarity systems.

7.3.1.2 Efficiency

Waveguide losses result from the following:

- Power dissipation in the waveguide walls and the dielectric material filling the enclosed space
- Leakage through the walls and transition connections of the guide
- Localized power absorption and heating at the connection points

The operating power of waveguide may be increased through pressurization. Sulfur hexafluoride commonly is used as the pressurizing gas.

7.3.2 Ridged Waveguide

Rectangular waveguide may be ridged to provide a lower cutoff frequency, thereby permitting use over a wider frequency band. As illustrated in Figure 7.12, one- and two-ridged guides are used. Increased bandwidth comes at the expense of increased attenuation, relative to an equivalent section of rectangular guide.

7.3.3 Circular Waveguide

Circular waveguide offers several mechanical benefits over rectangular or square guide. The windload of circular guide is two-thirds that of an equivalent run of rectangular waveguide. It also presents lower and more uniform windloading than rectangular waveguide, reducing tower structural requirements.

The same physical properties of circular waveguide that give it good power handling and low attenuation also result in electrical complexities. Circular waveguide has two potentially unwanted modes of propagation: the cross-polarized $TE_{1,1}$ and $TM_{0,1}$ modes.

Circular waveguide, by definition, has no short or long dimension and, consequently, no method to prevent the development of cross-polar or *orthogonal* energy. Cross-polar energy is formed by small ellipticities in the

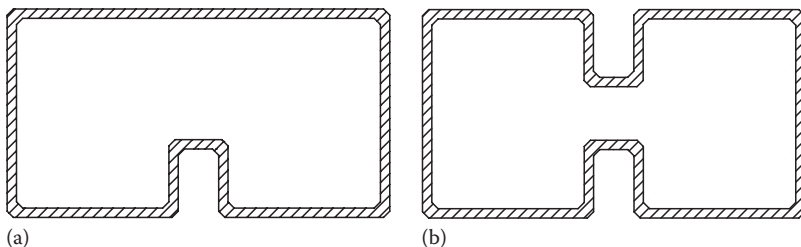


FIGURE 7.12

Ridged waveguide: (a) single ridged and (b) double ridged.

waveguide. If the cross-polar energy is not trapped out, the parasitic energy can recombine with the dominant-mode energy.

7.3.3.1 Parasitic Energy

Hollow circular waveguide works as a high-Q resonant cavity for some energy and as a transmission medium for the rest. The parasitic energy present in the cavity formed by the guide will appear as increased VSWR if not disposed of. The polarization in the guide meanders and rotates as it propagates from the source to the load. The end pieces of the guide, typically circular-to-rectangular transitions, are polarization sensitive. See Figure 7.13a. If the polarization of the incidental energy is not matched to the transition, energy will be reflected.

Several factors can cause this undesirable polarization. One cause is out-of-round guides that result from nonstandard manufacturing tolerances. In Figure 7.13, the solid lines depict the situation at launching: perfectly circular guide with perpendicular polarization. The dashed lines show how

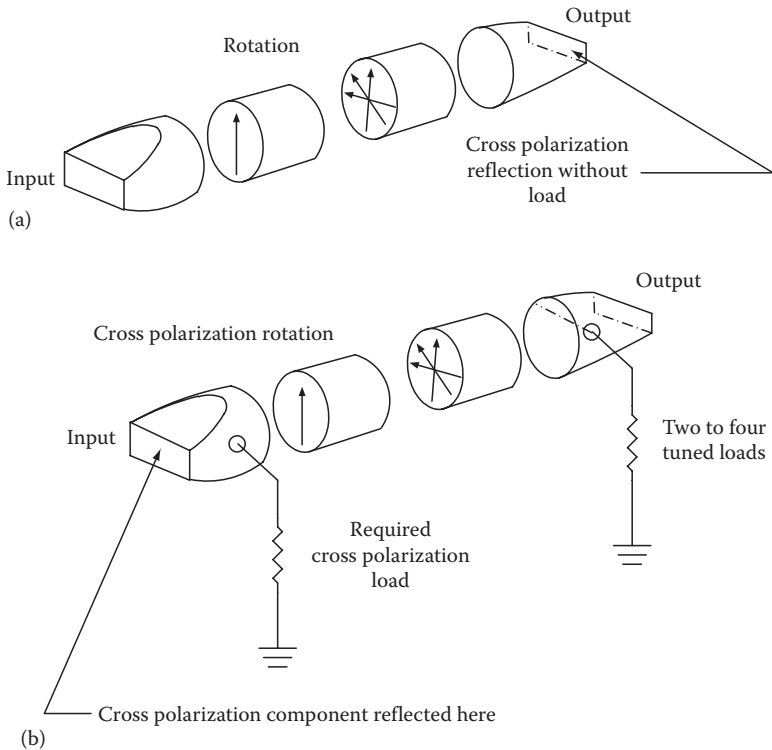


FIGURE 7.13 Effects of parasitic energy in circular waveguide: (a) trapped cross-polarization energy and (b) delayed transmission of the trapped energy.

certain ellipticities cause polarization rotation into unwanted states, while others have no effect. A 0.2% change in diameter can produce a -40 dB cross-polarization component per wavelength. This is roughly 0.03 in for 18 in. of guide length.

Other sources of cross-polarization include twisted and bent guides, out-of-roundness, offset flanges, and transitions. Various methods are used to dispose of this energy trapped in the cavity, including absorbing loads placed at the ground and/or antenna level.

7.3.4 Doubly Truncated Waveguide

The design of *doubly truncated waveguide* (DTW) is intended to overcome the problems that may result from parasitic energy in a circular waveguide. As shown in Figure 7.14, DTW consists of an almost elliptical guide inside a circular shell. This guide does not support cross-polarization; tuners and absorbing loads are not required. The low windload of hollow circular guide is maintained, except for the flange area.

Each length of waveguide is actually two separate pieces: a doubly truncated center section and a circular outer skin, joined at the flanges on each end. A large hole in the broadwall serves to pressurize the circular outer skin. Equal pressure inside the DTW and inside the circular skin ensures that the guide will not “breathe” or buckle as a result of rapid temperature changes.

DTW exhibits about 3% higher windloading than an equivalent run of circular waveguide (because of the transition section at the flange joints), and 32% lower loading than comparable rectangular waveguide.

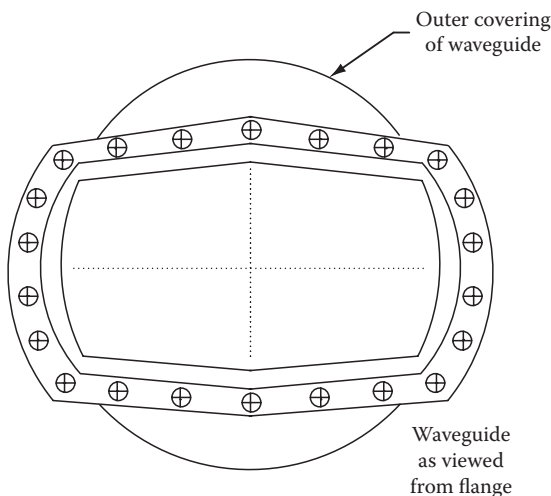


FIGURE 7.14

Physical construction of DTW.

7.3.5 Impedance Matching

The efficient flow of power from one type of transmission medium to another requires matching of the field patterns across the boundary to launch the wave into the second medium with a minimum of reflections. Coaxial line typically is matched into rectangular waveguide by extending the center conductor of the coax through the broadwall of the guide, parallel to the electric field lines across the guide. Alternatively, the center conductor can be formed into a loop and oriented to couple the magnetic field to the guide mode.

Standing waves are generally to be avoided in waveguide for the same reasons that they are to be avoided in transmission lines. Accordingly, it is usually necessary to provide impedance-matching systems in waveguides to eliminate standing waves. One approach involves the introduction of a compensating reflection in the vicinity of a load that neutralizes the standing waves that would exist in the system because of an imperfect match. A probe or tuning screw commonly is used to accomplish this, as illustrated in Figure 7.15. The tuning screw projects into the waveguide in a direction parallel to the electric field. This is equivalent to shunting a capacitive load across the guide. The susceptance of the load increases with extension into the guide up to $1/4$ wavelength. When the probe is exactly $1/4$ wavelength long, it becomes resonant and causes the guide to behave as though there were an open circuit at the point of the resonant probe. Probes longer than $1/4$ wavelength but shorter than $3/4$ wavelength introduce inductive loading. The extent to which such a probe projects into the waveguide determines the compensating reflection, and the position of the probe with respect to the standing wave pattern to be eliminated determines the phasing of the reflected wave.

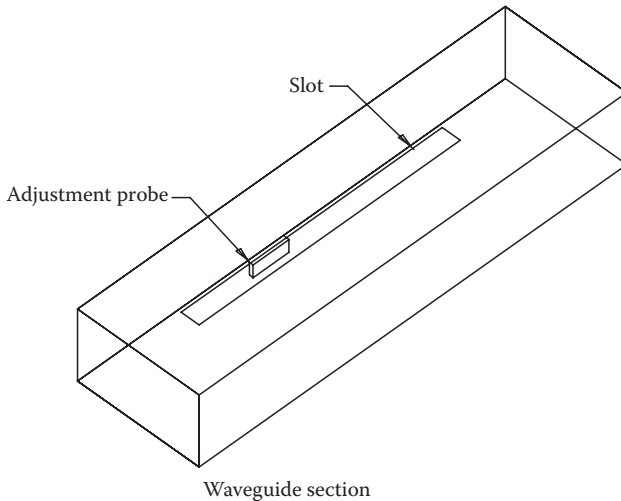
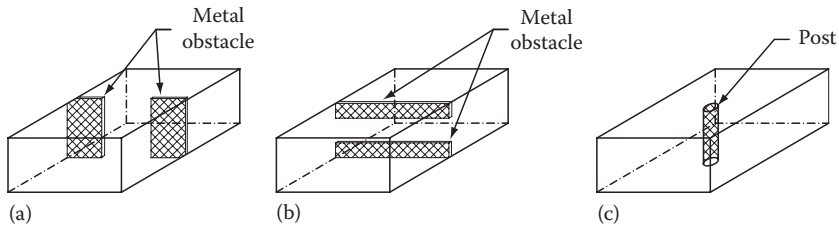


FIGURE 7.15

Probe configured to introduce a reflection in a waveguide that is adjustable in magnitude and phase.

**FIGURE 7.16**

Waveguide obstructions used to introduce compensating reflections: (a) inductive window, (b) capacitive window, and (c) post (inductive) element.

Dielectric slugs produce an effect similar to that of a probe. The magnitude of the effect depends upon the following considerations:

- The dielectric constant of the slug
- Thickness of the slug in an axial direction
- Whether the slug extends entirely across the waveguide

The phase of the reflected wave is controlled by varying the axial position of the slug.

There are several alternatives to the probe and slug for introducing controllable irregularity for impedance matching, including a metallic barrier or *window* placed at right angles to the axis of the guide, as illustrated in Figure 7.16. Three configurations are shown:

- The arrangement illustrated in Figure 7.16a produces an effect equivalent to shunting the waveguide with an inductive reactance.
- The arrangement shown in Figure 7.16b produces the effect of a shunt capacitive susceptance.
- The arrangement shown in Figure 7.16c produces an inductive shunt susceptance.

The waveguide equivalent of the coaxial cable tuning stub is a *tee* section, illustrated in Figure 7.17. The magnitude of the compensating effect is controlled by the position of the short-circuiting plug in the branch. The phase of the compensating reflected wave produced by the branch is determined by the position of the branch in the guide.

7.3.5.1 Waveguide Filters

A section of waveguide beyond cutoff constitutes a simple high-pass reflective filter. Loading elements in the form of posts or stubs may be employed to supply the reactances required for conventional lumped-constant filter designs.

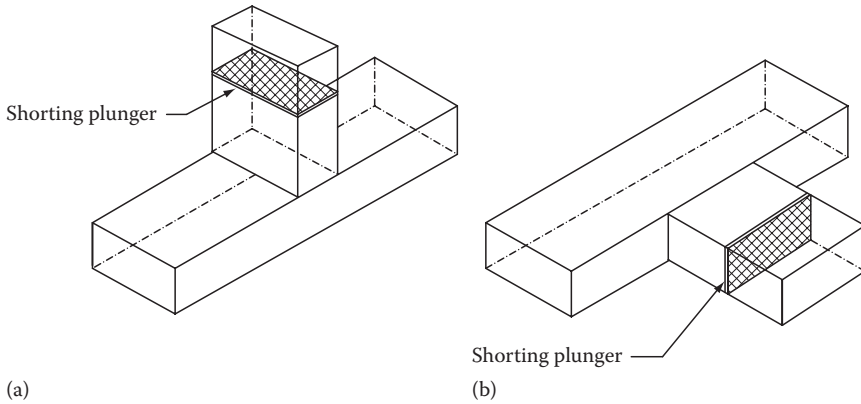


FIGURE 7.17

Waveguide stub elements used to introduce compensating reflections: (a) series *tee* element and (b) shunt *tee* element.

Absorption filters prevent the reflection of unwanted energy by incorporating lossy material in secondary guides that are coupled through *leaky walls* (small sections of guide beyond cutoff in the passband). Such filters typically are used to suppress harmonic energy.

7.3.6 Installation Considerations

Waveguide system installation is both easier and more difficult than traditional transmission line installation. There is no inner conductor to align, but alignment pins must be set and more bolts are required per flange. Transition hardware to accommodate loads and coax-to-waveguide interfacing also is required.

Flange reflections can add up in phase at certain frequencies, resulting in high VSWR. The length of the guide must be chosen so that flange reflection buildup does not occur within the operating bandwidth.

Flexible sections of waveguide are used to join rigid sections or components that cannot be aligned otherwise. Flexible sections also permit controlled physical movement resulting from thermal expansion of the line. Such hardware is available in a variety of forms. Corrugated guide commonly is produced by shaping thin-wall seamless rectangular tubing. Flexible waveguide can accommodate only a limited amount of mechanical movement. Depending on the type of link, the manufacturer may specify a maximum number of bends.

7.3.6.1 Tuning

Circular waveguide must be tuned. This requires a two-step procedure. First, the cross-polar $TE_{1,1}$ component is reduced, primarily through *axial ratio compensators* or *mode optimizers*. These devices counteract the net system ellipticity and indirectly minimize cross-polar energy. The cross-polar filters also may be rotated to achieve maximum isolation between the dominant and cross-polar modes.

Cross-polar energy manifests itself as a net signal rotation at the end of the waveguide run. A perfect system would have a net rotation of zero.

In the second step, tuning slugs at both the top and bottom of the waveguide run are adjusted to reduce the overall system VSWR. Tuning waveguide can be a complicated and time-consuming procedure. Once set, however, tuning normally does not drift and must be repeated only if major component changes are made.

7.3.6.2 Waveguide Hardware

Increased use of waveguide has led to the development of waveguide-based hardware for all elements from the output of the RF generator to the load. Waveguide-based filters, elbows, directional couplers, switches, combiners, and diplexers are readily available. The RF performance of a waveguide component is usually better than the same item in coax. This is especially true in the case of diplexers and filters. Waveguide-based hardware provides lower attenuation and greater power-handling capability for a given physical size.

7.3.7 Cavity Resonators

Any space completely enclosed by conducting walls can contain oscillating electromagnetic fields. Such a cavity possesses certain frequencies at which it will resonate when excited by electrical oscillations. These *cavity resonators* find extensive use as resonant circuits at VHF and above. The advantages of cavity resonators over conventional *LC* circuits include the following:

- Simplicity in design.
- Relatively large physical size compared with alternative methods of obtaining resonance. This attribute is important in high-power, high-frequency applications.
- High *Q*.
- The capability to configure the cavity to develop extremely high shunt impedance.

Cavity resonators commonly are used at wavelengths on the order of 10cm or less.

The simplest cavity resonator is a section of waveguide shorted at each end with a length *l* equal to

$$l = \frac{\lambda_g}{2} \quad (7.12)$$

where λ_g is the guide wavelength.

This configuration results in a resonance similar to that of a 1/2-wavelength transmission line short-circuited at the receiving end.

A sphere or any other enclosed surface (irrespective of how irregular the outline) also may be used to form a cavity resonator.

Any given cavity is resonant at a number of frequencies, corresponding to the different possible field conditions that can exist within the space. The resonance having the longest wavelength (lowest frequency) is termed the *dominant* or *fundamental resonance*. The resonant wavelength is proportional to the size of the resonator. If all dimensions are doubled, the wavelength corresponding to resonance will likewise be doubled. The resonant frequency of a cavity can be changed by incorporating one or more of the following mechanisms:

- Altering the mechanical dimensions of the cavity. Small changes can be achieved by flexing walls, but large changes require some form of sliding member.
- Coupling reactance into the resonator through a coupling loop.
- Introducing a movable copper paddle into the cavity. A paddle placed inside the resonator will affect the normal distribution of flux and tend to raise the resonant frequency by an amount determined by the orientation of the paddle.

The Q of a cavity resonator has the same significance as for a conventional resonant circuit. Q can be defined for a cavity by the relationship

$$Q = 2\pi \frac{E_s}{E_l} \quad (7.13)$$

where

E_s is the energy stored

E_l is the energy lost per cycle

The energy stored is proportional to the square of the magnetic flux density integrated throughout the volume of the resonator. The energy lost in the walls is proportional to the square of the magnetic flux density integrated over the surface of the cavity. To obtain high Q , the resonator should have a large ratio of volume to surface area, because it is the volume that stores energy and the surface area that dissipates energy.

Coupling can be obtained from a resonator by means of a coupling loop or coupling electrode. Magnetic coupling is accomplished through the use of a loop oriented so as to enclose magnetic flux lines existing in the desired mode of operation. This technique is illustrated in Figure 7.18. A current passed through the loop will excite oscillations of this mode. Conversely, oscillations existing in the resonator will induce a voltage in the coupling loop. The magnitude of the coupling can be controlled by rotating the loop; the coupling reduces to zero when the plane of the loop is parallel to the magnetic flux.

Coupling of a resonator also may be accomplished through the use of a probe or opening in one wall of the cavity.

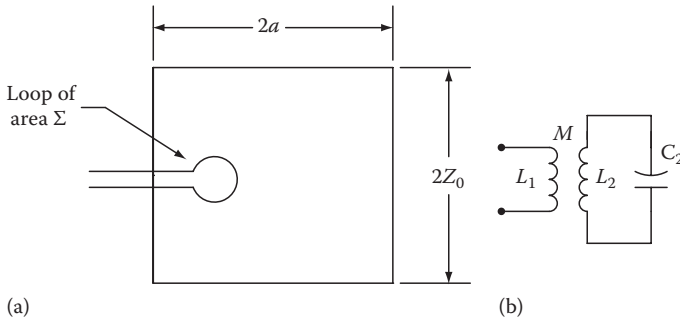


FIGURE 7.18
Cavity resonator coupling: (a) coupling loop and (b) equivalent circuit.

7.4 RF Combiner and Diplexer Systems

The basic purpose of an RF combiner is to add two or more signals to produce an output signal that is a composite of the inputs. The combiner performs this signal addition while providing isolation between inputs. Combiners perform other functions as well, and can be found in a wide variety of RF equipment utilizing power vacuum tubes. Combiners are valuable devices because they permit multiple amplifiers to drive a single load. The isolation provided by the combiner permits tuning adjustments to be made on one amplifier—including turning it on or off—without significantly affecting the operation of the other amplifier. In a typical application, two amplifiers drive the hybrid and provide two output signals:

- A combined output representing the sum of the two input signals, typically directed toward the antenna.
- A difference output representing the difference in amplitude and phase between the two input signals. The difference output typically is directed toward a dummy (reject) load.

For systems in which more than two amplifiers must be combined, two or more combiners are cascaded.

Diplexers are similar in nature to combiners but permit the summing of output signals from two or more amplifiers operating at different frequencies. This allows, for example, the outputs of several transmitters operating on different frequencies to utilize a single broadband antenna.

7.4.1 Passive Filters

A *filter* is a multiport network designed specifically to respond differently to signals of different frequency [1]. This definition excludes *networks*, which

incidentally behave as filters, sometimes to the detriment of their main purpose. Passive filters are constructed exclusively with passive elements (i.e., resistors, inductors, and capacitors). Filters are generally categorized by the following general parameters:

- Type
- Alignment (or class)
- Order

7.4.1.1 Filter Type

Filters are categorized by type, according to the magnitude of the frequency response, as one of the following [1]:

- *Low-pass* (LP)
- *High-pass* (HP)
- *Bandpass* (BP)
- *Band-stop* (BS)

The terms *band-reject* or *notch* are also used as descriptive of the BS filter. The term *all-pass* is sometimes applied to a filter whose purpose is to alter the phase angle without affecting the magnitude of the frequency response. Ideal and practical interpretations of the types of filters and the associated terminology are illustrated in Figure 7.19.

In general, the voltage gain of a filter in the *stop band* (or *attenuation band*) is less than $\sqrt{2}/2$ (≈ 0.707) times the maximum voltage gain in the passband. In logarithmic terms, the gain in the stop band is at least 3.01 dB less than the maximum gain in the passband. The cutoff (*break* or *corner*) frequency separates the passband from the stop band. In BP and BS filters, there are two cutoff frequencies, sometimes referred to as the *lower* and *upper* cutoff frequencies. Another expression for the cutoff frequency is *half-power frequency*, because the power delivered to a resistive load at cutoff frequency is one-half the maximum power delivered to the same load in the passband. For BP and BS filters, the center frequency is the frequency of maximum or minimum response magnitude, respectively, and bandwidth is the difference between the upper and lower cutoff frequencies. *Rolloff* is the transition from passband to stop band and is specified in gain unit per frequency unit (e.g., gain unit/Hz, dB/decade, dB/octave, etc.)

7.4.1.2 Filter Alignment

The *alignment* (or class) of a filter refers to the shape of the frequency response [1]. Fundamentally, filter alignment is determined by the coefficients

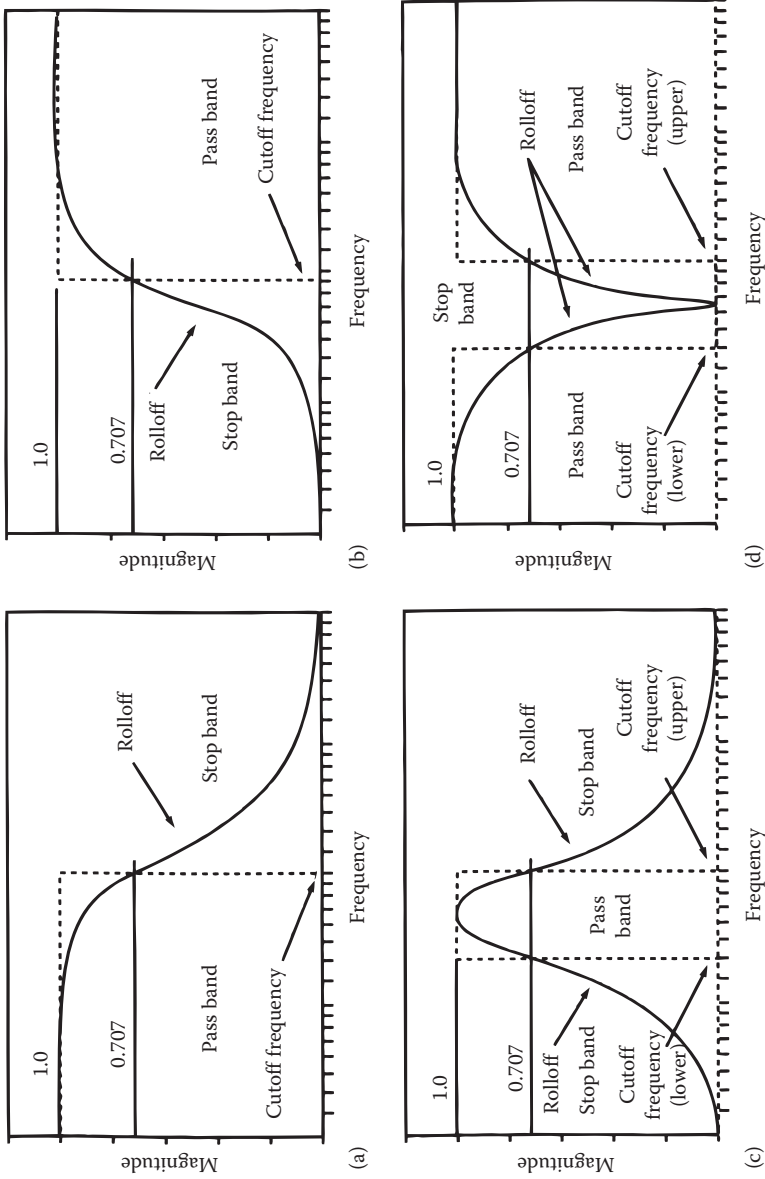


FIGURE 7.19 Filter characteristics by type: (a) low-pass, (b) high-pass, (c) bandpass, and (d) band-stop. (After Harrison, C., *Passive filters*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 279–290, 1996.)

of the filter network transfer function, so there is an indefinite number of filter alignments, some of which may not be realizable. The more common alignments are

- Butterworth
- Chebyshev
- Bessel
- Inverse Chebyshev
- Elliptic (or Cauer)

Each filter alignment has a frequency response with a characteristic shape, which provides some particular advantage (see Figure 7.20). Filters with Butterworth, Chebyshev, or Bessel alignment are called *all-pole filters* because their low-pass transfer functions have no zeros. Table 7.3 summarizes the characteristics of the standard filter alignments.

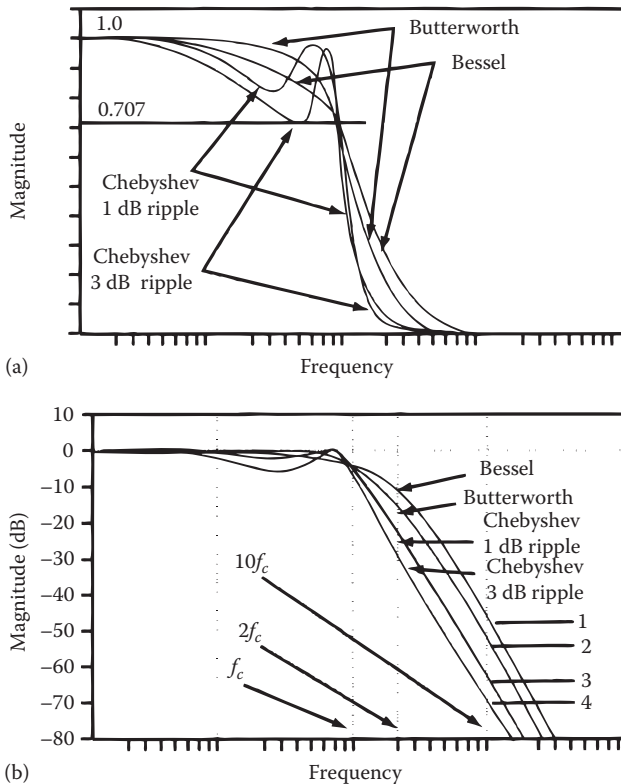


FIGURE 7.20

Filter characteristics by alignment, third-order, all-pole filters: (a) magnitude and (b) magnitude in decibels. (After Harrison, C., *Passive filters*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 279–290, 1996.)

TABLE 7.3
Summary of Standard Filter Alignments

Alignment	Passband Description	Stop Band Description	Comments
Butterworth	Monotonic	Monotonic	All-pole; maximally flat
Chebyshev	Rippled	Monotonic	All-pole
Bessel	Monotonic	Monotonic	All-pole; constant phase shift
Inverse Chebyshev	Monotonic	Rippled	
Elliptic (or Cauer)	Rippled	Rippled	

Source: Harrison, C., Passive filters, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 279–290, 1996.

7.4.1.3 Filter Order

The *order* of a filter is equal to the number of poles in the filter network transfer function [1]. For a lossless LC filter with resistive (nonreactive) termination, the number of reactive elements (inductors or capacitors) required to realize an LP or HP filter is equal to the order of the filter. Twice the number of reactive elements are required to realize a BP or a BS filter of the same order. In general, the order of a filter determines the slope of the rolloff—the higher the order, the steeper the rolloff. At frequencies greater than approximately one octave above cutoff (i.e., $f > 2f_c$), the rolloff for all-pole filters is $20 n$ dB/decade (or approximately $6 n$ dB/octave), where n is the order of the filter (Figure 7.21). In the vicinity of f_c , both filter alignment and filter order determine rolloff.

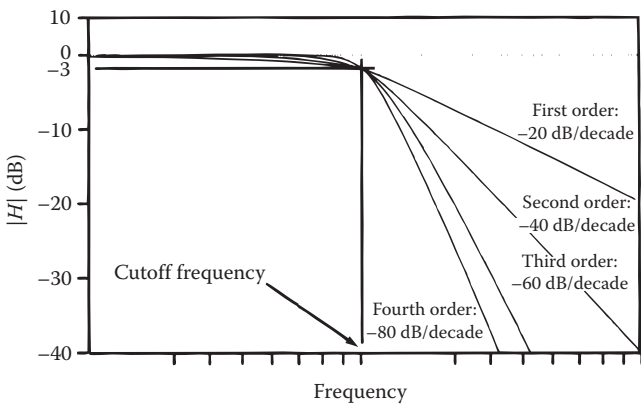


FIGURE 7.21
Effects of filter order on rolloff (Butterworth alignment). (After Harrison, C., Passive filters, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 279–290, 1996.)

7.4.2 Four-Port Hybrid Combiner

A hybrid combiner (coupler) is a reciprocal four-port device that can be used for either splitting or combining RF energy over a wide range of frequencies. An exploded view of a typical 3 dB 90° hybrid is illustrated in Figure 7.22. The device consists of two identical parallel transmission lines coupled over a distance of approximately one-quarter wavelength and enclosed within a single outer conductor. Ports at the same end of the coupler are in phase, and ports at the opposite end of the coupler are in quadrature (90° phase shift) with respect to each other.

The phase shift between the two inputs or outputs is always 90° and is virtually independent of frequency. If the coupler is being used to combine two signals into one output, these two signals must be fed to the hybrid in phase quadrature. When the coupler is used as a power splitter, the division is equal (half-power between the two ports). The hybrid presents a constant impedance to match each source.

Operation of the combiner can best be understood through observation of the device in a practical application. Figure 7.23 shows a four-port hybrid combiner used to add the outputs of two transmitters to feed a single load. The combiner accepts one RF source and splits it equally into two parts. One part arrives at output port C with 0° phase (no phase delay; it is the *reference phase*). The other part is delayed by 90° at port D. A second RF source connected to input port B, but with a phase delay of 90° , also will split in two, but the signal arriving at port C now will be in phase with source 1, and the signal arriving at port D will cancel, as shown in the figure.

Output port C, the summing point of the hybrid, is connected to the load. Output port D is connected to a resistive load to absorb any residual power resulting from slight differences in amplitude and/or phase between the two

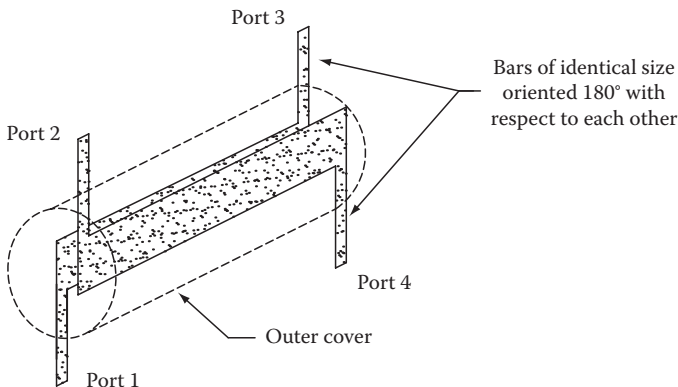


FIGURE 7.22

Physical model of a 90° hybrid combiner.

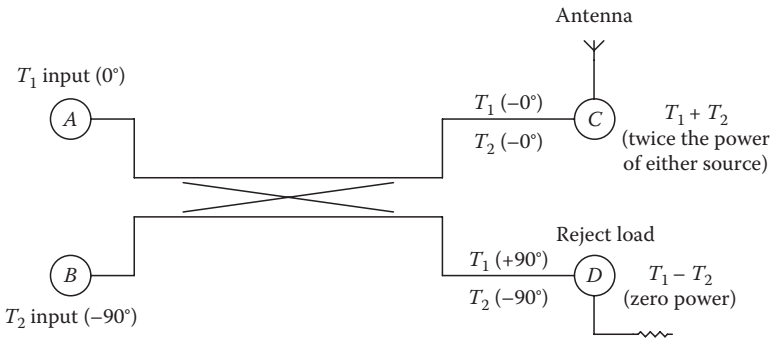


FIGURE 7.23

Operating principles of a hybrid combiner. This circuit is used to add two identical signals at inputs *A* and *B*.

input sources. If one of the RF inputs fails, half of the remaining transmitter output will be absorbed by the resistive load at port *D*.

The four-port hybrid works only when the two signals being mixed are identical in frequency and amplitude, and when their relative phase is 90° .

Operation of the hybrid can best be described by a *scattering matrix* in which vectors are used to show how the device operates. Such a matrix is shown in Table 7.4. In a 3 dB hybrid, two signals are fed to the inputs. An input signal at port 1 with 0° phase will arrive in phase at port 3, and at port 4 with a 90° lag (-90°) referenced to port 1. If the signal at port 2 already contains a 90° lag (-90° referenced to port 1), both input signals will combine in phase at port 4. The signal from port 2 also experiences another 90° change in the hybrid as it reaches port 3. Therefore, the signals from ports 1 and 2 cancel each other at port 3.

If the signal arriving at port 2 leads by 90° (mode 1 in the table), the combined power from ports 1 and 2 appears at port 4. If the two input signals are matched in phase (mode 4), the output ports (3 and 4) contain one-half of the power from each of the inputs.

If one of the inputs is removed, which would occur in a transmitter failure, only one hybrid input receives power (mode 5). Each output port then would receive one-half the input power of the remaining transmitter, as shown.

The input ports present a predictable load to each amplifier with a VSWR that is lower than the VSWR at the output port of the combiner. This characteristic results from the action of the difference port, typically connected to a dummy load. Reflected power coming into the output port will be directed to the reject load, and only a portion will be fed back to the amplifiers. Figure 7.24 illustrates the effect of output port VSWR on input port VSWR and on the isolation between ports.

As noted previously, if the two inputs from the separate amplifiers are not equal in amplitude and not exactly in phase quadrature, some power will be dissipated in the difference port reject load. Figure 7.25 plots the effect of

TABLE 7.4
Single 90° Hybrid System Operating Modes

Mode	Input		Schematic	Output	
	1	2		3	4
1	$P_1/0^\circ$	$P_2/-90^\circ$		0	$P_1 + P_2$
2	$P_1/0^\circ$	$P_2/90^\circ$		$P_1 + P_2$	0
3	$P_1/0^\circ$	$P_2/0^\circ$		$P_1/2 + P_2/2$	$P_1/2 + P_2/2$
4	$P_1/0^\circ$	$P_2 = 0$		$P_1/2$	$P_1/2$
5	$P_1 = 0$	$P_2/0^\circ$		$P_2/2$	$P_2/2$

Legend: \uparrow = Unit vector port 1
 \uparrow = Unit vector port 2
 $\uparrow \leftarrow$ = Vector cancellation
 $\uparrow \leftarrow$ = Vector addition
 $\nearrow \nwarrow$ = Half power from each vector

power imbalance, and Figure 7.26 plots the effects of phase imbalance. The power lost in the reject load can be reduced to a negligible value by trimming the amplitude and/or phase of one (or both) amplifiers.

7.4.3 Nonconstant-Impedance Diplexer

Diplexers are used to combine amplifiers operating on different frequencies (and at different power levels) into a single output. Such systems typically are utilized to sum different transmitter outputs to feed a single broadband antenna.

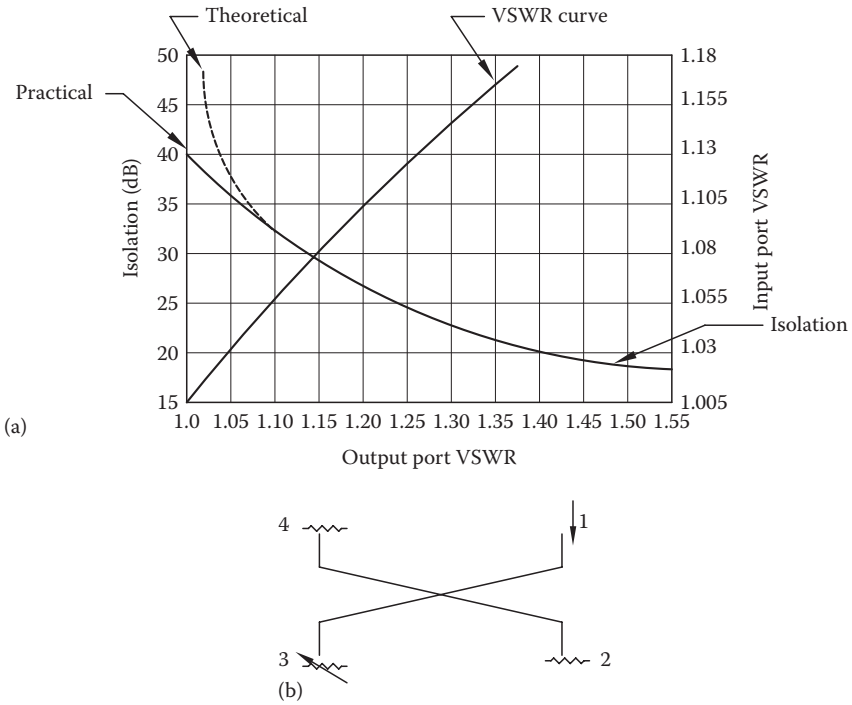
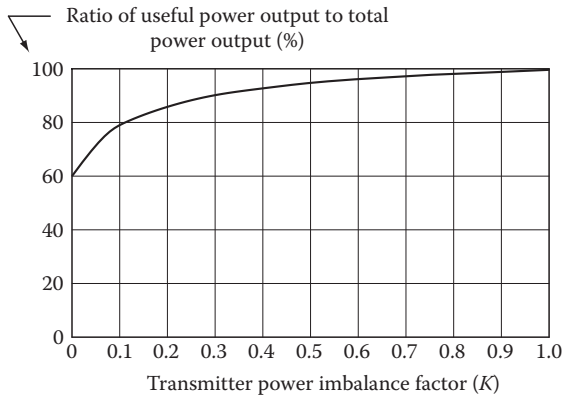


FIGURE 7.24 Effects of load VSWR on input VSWR and isolation: (a) respective curves and (b) coupler schematic.



$$K = \sqrt{\frac{P_a}{P_b}}$$

P_a = Output power of transmitter 1 (lower power system)

P_b = Output power of transmitter 2

FIGURE 7.25 Effects of power imbalance at the inputs of a hybrid coupler.

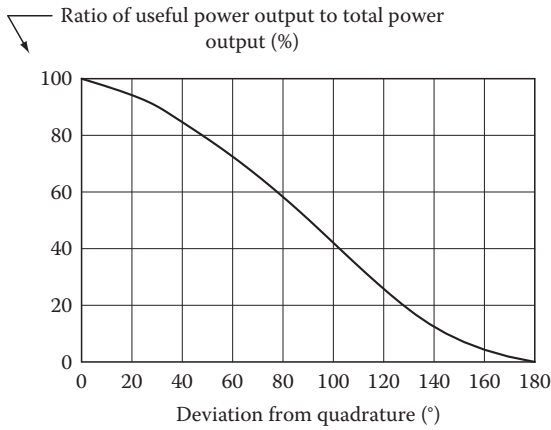


FIGURE 7.26
Phase sensitivity of a hybrid coupler.

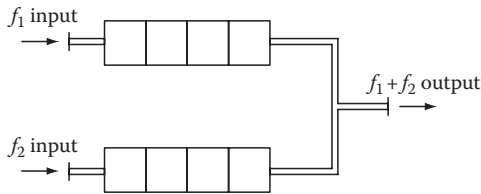


FIGURE 7.27
Nonconstant-impedance branch diplexer. In this configuration, two banks of filters feed into a coaxial tee.

The *branch diplexer* is the typical configuration for a diplexer that does not exhibit constant-impedance inputs. As shown in Figure 7.27, the branch diplexer consists of two banks of filters, each feeding into a coaxial tee. The electrical length between each filter output and the centerline of the tee is frequency sensitive, but this fact is more of a tuning nuisance than a genuine user concern.

For this type of diplexer, all of the electrical parameters are a function of the filter characteristics. The VSWR, insertion loss, group delay, and rejection/isolation will be the same for the overall system as they are for the individual banks of cavities. The major limitation of this type of combiner is the degree of isolation that can be obtained for closely spaced channels.

7.4.4 Constant-Impedance Diplexer

The constant-impedance diplexer employs 3 dB hybrids and filters with a terminating load on the isolated port. The filters in this type of combiner can be either notch type or bandpass type. The performance characteristics are noticeably different for each design.

7.4.4.1 Band-Stop Diplexer

The band-stop (notch) constant-impedance diplexer is configured as shown in Figure 7.28. For this design, the notch filters must have a high Q response to keep insertion loss low in the passband skirts. The high Q characteristic results in a sharp notch. Depending on the bandwidth required of the diplexer, two or more cavities may be located in each leg of the diplexer. They are typically stagger tuned, one high and one low for the two-cavity case. With this dual-cavity reject response in each leg of the band-stop diplexer system, the following analysis explains the key performance specifications.

If frequency f_1 is fed into the top left port of Figure 7.28, it will be split equally into the upper and lower legs of the diplexer. Both of these signals will reach the filters in their respective legs and be rejected/reflected back toward the input hybrid, recombine, and emerge through the lower left port, also known as the *wideband output*.

The VSWR looking into the f_1 input is near 1:1 at all frequencies in the band. Within the bandwidth of the reject skirts, the observed VSWR is equal to the termination of the wideband output. Outside of the passband, the signals will pass by the cavities, enter the rightmost hybrid, recombine, and emerge into the dummy load. Consequently, the out-of-band VSWR is, in fact, the VSWR of the load.

The insertion loss from the f_1 input to the wideband output is low, typically on the order of 0.1 dB at carrier. This insertion loss depends on perfect reflection from the cavities. As the rejection diminishes on the skirts of the filters, the insertion loss from f_1 to the wideband output increases.

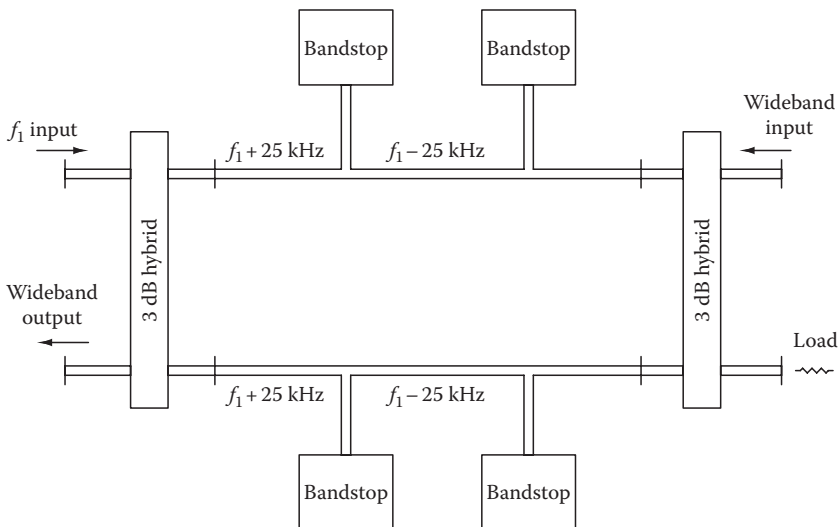


FIGURE 7.28

Band-stop (notch) constant-impedance diplexer module. This design incorporates two 3 dB hybrids and filters, with a terminating load on the isolated port.

The limitation in reject bandwidth of the cavities causes the insertion loss to rise at the edges of the passband. The isolation from f_1 to the wideband input consists of a combination of the reject value of the cavities plus the isolation of the rightmost hybrid.

A signal entering at f_1 splits and proceeds in equal halves rightward through both the upper and lower legs of the diplexer. It is rejected by the filters at carrier and, to a lesser extent, on both sides of the carrier. Any residual signal that gets by the cavities reaches the rightmost hybrid. There it recombines and emerges from the *load port*. The hybrid provides a specified isolation from the load port to the wideband input port. This hybrid isolation must be added to the filter rejection to obtain the total isolation from the f_1 input to the wideband input.

If a signal is fed into the wideband input of the combiner shown in Figure 7.28, the energy will split equally and proceed leftward along the upper and lower legs of the diplexer. Normally, f_1 is not fed into the wideband input. If it were, f_1 would be rejected by the notch filters and recombine into the load. All other signals sufficiently removed from f_1 will pass by the cavities with minimal insertion loss and recombine into the wideband output.

The VSWR looking into the wideband input is equal to the VSWR at the output for frequencies other than f_1 . If f_1 were fed into the wideband input, the VSWR would be equal to the VSWR of the load.

Isolation from the wideband input to the f_1 input is simply the isolation available in the leftmost 3 dB hybrid. This isolation is usually inadequate for high-power applications. To increase the isolation from the wideband input to the f_1 input, it is necessary to use additional cavities between the f_1 transmitter and the f_1 input that will reject the frequencies fed into the wideband input. Unfortunately, adding these cavities to the input line also cancels the constant-impedance input. Although the notch diplexer as shown in the figure is truly a constant-impedance type of diplexer, the constant impedance is presented to the two inputs by virtue of using the hybrids at the respective inputs.

The hybrids essentially cause the diplexer to act as an absorptive type of filter to all out-of-band signals. The out-of-band signals generated by the transmitter are absorbed by the load rather than reflected to the transmitter.

If a filter is added to the input to supplement isolation, the filter will reflect some out-of-band signals back at the transmitter. The transmitter then will be seeing the passband impedance of this supplemental filter rather than the constant impedance of the notch diplexer. This is a serious deficiency for applications that require a true constant-impedance input.

7.4.4.2 Bandpass Constant-Impedance Diplexer

The bandpass constant-impedance diplexer is shown in Figure 7.29. This system takes all of the best features of diplexers and combines them into one unit. It also provides a constant-impedance input that need not be supplemented with input cavities that rob the diplexer of its constant-impedance input.

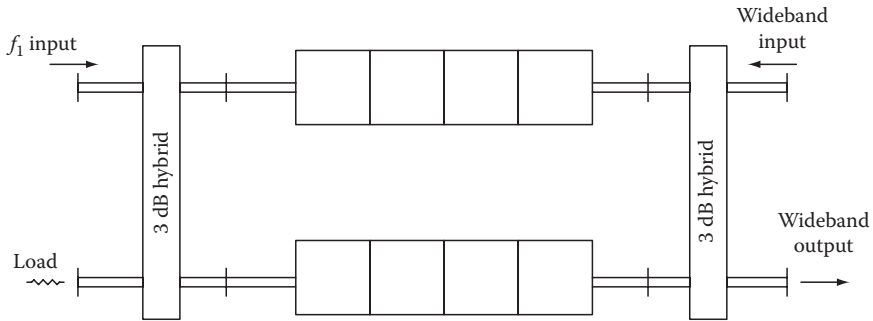


FIGURE 7.29
Bandpass constant-impedance diplexer.

The bandpass filters exhibit good bandwidth, providing near 1:1 VSWR across the operating bandpass. Insertion is low at carrier (0.28 dB is typical), rising slightly at bandpass extremes. Diplexer rejection, when supplemented by isolation of the hybrids, provides ample transmitter-to-transmitter isolation. Group delay is typically exceptional, providing performance specifications similar to those of a branch-style bandpass system. This configuration has the additional capability of providing high port-to-port isolation between closely spaced operating channels, as well as a true constant-impedance input.

The hybrids shown in Figure 7.29 work in a manner identical to those described for a band-stop diplexer. However, the bandpass filters cause the system to exhibit performance specifications that exceed the band-stop system in every way. Consider a signal entering at the f_1 input.

Within the passbands of the filters, which are tuned to f_1 , the VSWR will be near 1:1 at carrier, rising slightly at the bandpass extremes. Because of the characteristic of the leftmost hybrid, the VSWR is, in fact, a measure of the similarity of response of the top and bottom bands of filters. The insertion loss looking from the f_1 input to the wideband output will be similar to the insertion loss of the top and bottom filters individually. A value of approximately 0.28 dB at carrier is typical.

Both the insertion loss and group delay can be determined by the design bandwidth of the filters. Increasing bandwidth causes the insertion loss and group delay deviation to decrease. Unfortunately, as the bandwidth increases with a given number of cavities, the isolation suffers for closely spaced channels because the reject skirt of the filter decreases with increasing bandwidth.

Isolation of f_1 to the wideband input is determined as follows: A signal enters at the f_1 input, splits equally into the upper and lower banks of filters, passes with minimal loss through the filters, and recombines into the wideband output of the rightmost hybrid. Both the load and the wideband input ports are isolated by their respective hybrids to some given value below the

f_1 input level. Isolation of the f_1 input to the wideband input is supplemented by the reject skirt of the next module.

A signal fed into the wideband input could be any frequency removed from f_1 by some minimal amount. As the wideband signal enters, it will be split into equal halves by the hybrid, then proceed to the left until the two components reach the reject skirts of the filters. The filters will shunt all frequencies removed from f_1 by some minimal amount. If the shunt energy is in phase for the given frequency when the signal is reflected back to the right hybrid, it will recombine into the wideband output. The VSWR under these conditions will be equal to the termination at the wideband output. If the reject skirts of the filter are sufficient, the insertion loss from wideband input to wideband output will be minimal (0.03 dB is typical).

The isolation from the wideband input to f_1 can be determined as follows: A signal enters at the wideband input, splits equally into upper and lower filters, and is rejected by the filters. Any residual signal that passes through the filters in spite of the rejection still will be in the proper phase to recombine into the load, producing additional isolation to the f_1 input port. Thus, the isolation of the wideband input to the f_1 input is the sum of the rejection from the filters and from the left hybrid.

7.4.4.3 Isolation of f_1 to f_2

Extending the use of the diplexer module into a multiplexer application supplements the deficient isolation described previously (narrowband input performance), while maintaining the constant-impedance input. In a multiplexer system, the wideband input of one module is connected to the wideband output of the next module, as illustrated in Figure 7.30.

It has already been stated that the isolation from the f_1 input to the wideband input is deficient, but additional isolation is provided by the isolation of the wideband output to the f_2 input of the next module. Consider that f_1 has already experienced 30 dB isolation to the wideband input of the same module. When this signal continues to the next module through the wideband output of module 2, it will be split into equal halves and proceed to the left of module 2 until it reaches the reject skirts of the filters in module 2. Assume that these filters are tuned to f_2 and reject f_1 by at least 25 dB. The combined total isolation of f_1 to f_2 is the sum of the 30 dB of the right hybrid in module 1, plus the 25 dB of the reject skirts of module 2, for a total of 55 dB.

7.4.4.4 Intermodulation Products

The isolation just described is equal in magnitude to that for a band-stop module, but it provides further protection against the generation of intermodulation products. The most troublesome intermodulation (*intermod*) products usually occur when an incoming (secondary) signal mixes with the second harmonic of a primary transmitter.

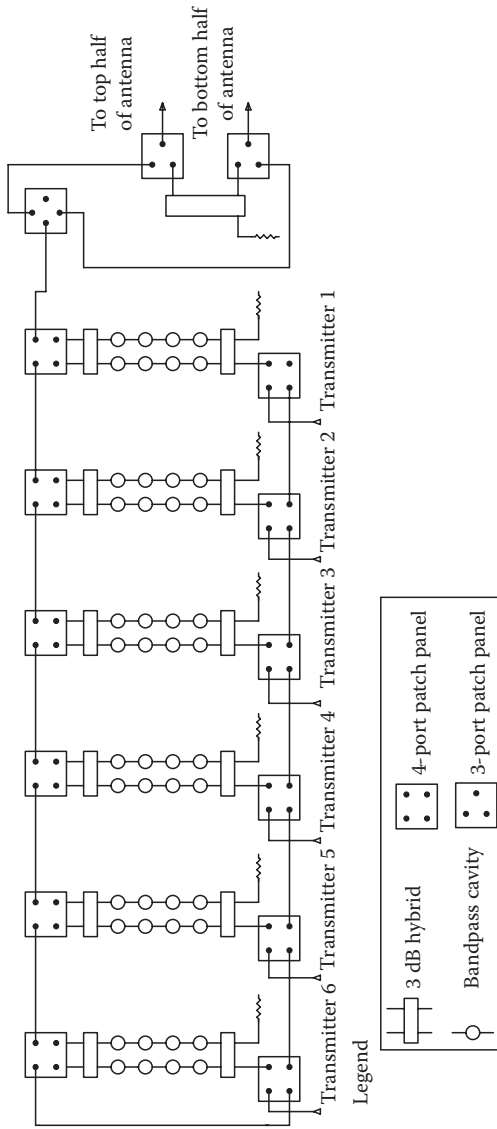


FIGURE 7.30 Schematic diagram of a six-module bandpass multiplexer. This configuration accommodates a split antenna design and incorporates patch panels for bypass purposes.

When the primary transmitter is operating on frequency A , the intermod will occur at the frequency $(2 \times A) - B$. This formula invariably places the intermod from the primary transmitter symmetrically about the operating frequency. By an interesting coincidence, the bandpass filters in the bandpass module also provide symmetrical reject response on both sides of the primary operating frequency.

Assume again that the incoming signal is attenuated by 30dB in the respective hybrid and by 25dB in the filter, for a total of 55 dB. If an intermod product still is generated in spite of this isolation, it will emerge on the other skirt of the filter attenuated by 25dB. In the bandpass system, the incoming signal is attenuated by 55 dB and the resulting outgoing spur by 25 dB, for a total of 80 dB suppression.

Interestingly, the entire 80dB of attenuation is supplied by the diplexer regardless of the *turnaround loss* of the transmitter. The tendency toward wideband final stage amplifiers in transmitters requires constant-impedance inputs. The transmitters also require increased isolation because they usually offer limited turnaround loss.

7.4.4.5 Group Delay

Group delay in the bandpass multiplexer module is equal to the sum of the narrowband input group delay and the wideband input group delay of all modules between the input and the load (antenna). The narrowband input group delay is a U-shaped response, with minimum at center and rising to a maximum on both sides at the frequency where the reject rises to 3 dB. Group delay then decreases rapidly at first, then more slowly.

If the bandwidth of the passband is made such that the group delay is ± 25 ns over ± 150 kHz (in an example system operating near 100 MHz), the 3 dB points will be at ± 400 kHz and the out-of-band group delay will fall rapidly at ± 800 kHz and possibly ± 1.0 MHz. If there are no frequencies 800 kHz or 1.0 MHz removed upstream in other modules, this poses no problem.

If modules upstream are tuned to 800 kHz or 1.0 MHz on either side, then the group delay (when viewed at the upstream module) will consist of its own narrowband input group delay plus the rapidly falling group delay of the wideband input of the closely spaced downstream module. Under these circumstances, if good group delay is desired, it is possible to utilize a *group delay compensation* module.

A group delay compensation module consists of a hybrid and two cavities used as notch cavities. It typically provides a group delay response that is inverted, compared with a narrowband input group delay. Because group delay is additive, the inverted response subtracts from the standard response, effectively reducing the group delay deviation.

It should be noted that the improvement in group delay is obtained at a cost of insertion loss. In large systems (8–10 modules), the insertion loss can be high because of the cumulative total of all wideband losses. Under these

conditions, it may be more prudent to accept higher group delay and retain minimal insertion loss.

7.4.5 Microwave Combiners

Hybrid combiners typically are used in microwave amplifiers to add the output energy of individual power modules to provide the necessary output from an RF generator. Quadrature hybrids effect a VSWR-canceling phenomenon that results in well-matched power amplifier inputs and outputs that can be broadbanded with proper selection of hybrid tees. Several hybrid configurations are possible, including the following:

- *Split-tee*
- *Branch-line*
- *Magic-tee*
- *Backward-wave*

Key design parameters include coupling bandwidth, isolation, and ease of fabrication. The equal-amplitude quadrature-phase reverse-coupled TEM $1/4$ -wave hybrid is particularly attractive because of its bandwidth and amenability to various physical implementations. Such a device is illustrated in Figure 7.31.

7.4.6 Hot-Switching Combiners

Switching RF is nothing new. Typically, the process involves coaxial switches, coupled with the necessary logic to ensure that the “switch” takes place with no RF energy on the contacts. This process usually takes the system off-line for a few seconds while the switch is completed. Through the use of hybrid combiners, however, it is possible to redirect RF signals without turning the carrier off. This process is referred to as *hot switching*. Figure 7.32 illustrates two of the most common switching functions (SPST and DPDT) available from hot switchers.

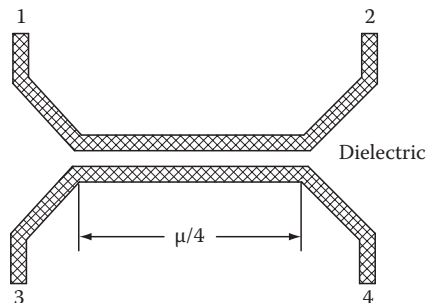


FIGURE 7.31
Reverse-coupled $1/4$ -wave hybrid combiner.

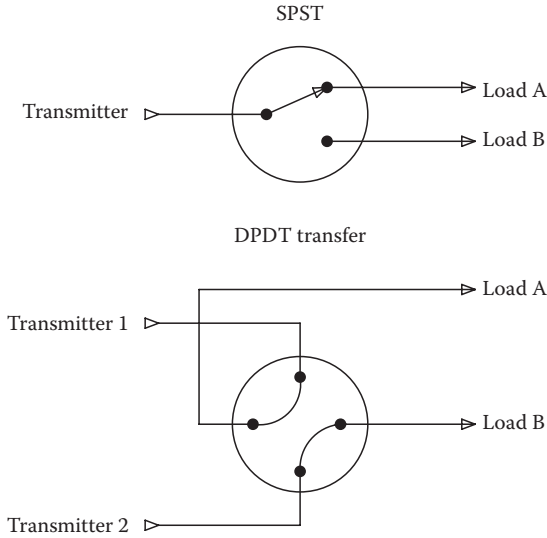


FIGURE 7.32
Common RF switching configurations.

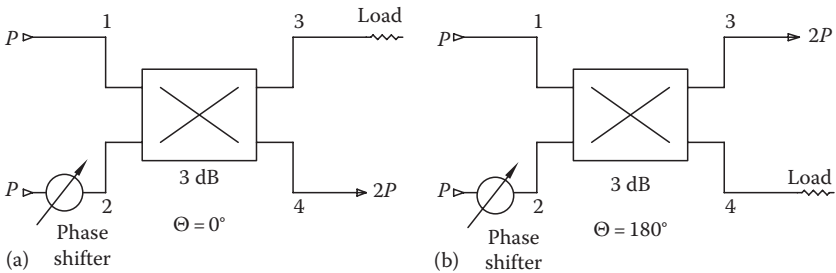
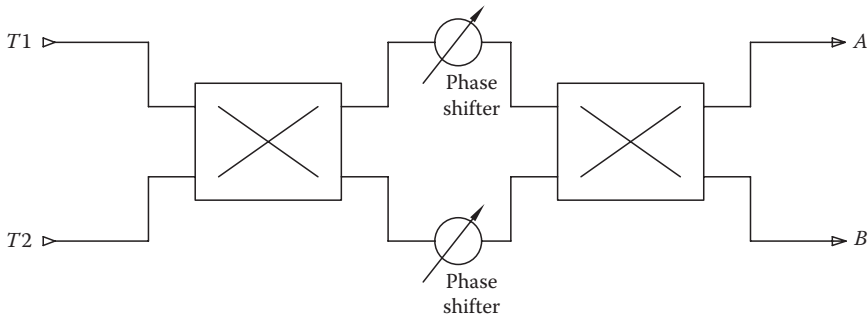


FIGURE 7.33
Hybrid switching configurations: (a) phase set so that the combined energy is delivered to port 4 and (b) phase set so that the combined energy is delivered to port 3.

The unique phase-related properties of an RF hybrid make it possible to use the device as a switch. The input signals to the hybrid in Figure 7.33a are equally powered but differ in phase by 90° . This phase difference results in the combined signals being routed to the output terminal at port 4. If the relative phase between the two input signals is changed by 180° , the summed output then appears on port 3, as shown in Figure 7.33b. The 3 dB hybrid combiner, thus, functions as a switch.

This configuration permits the switching of two RF generators to either of two loads. Remember, however, that the switch takes place when the phase difference between the two inputs is 90° . To perform the switch in a useful way requires adding a high-power phase shifter to one input leg of

**FIGURE 7.34**

Additional switching and combining functions enabled by adding a second hybrid and another phase shifter to a hot-switching combiner.

the hybrid. The addition of the phase shifter permits the full power to be combined and switched to either output. This configuration of hybrid and phase shifter, however, will not permit switching a main or standby generator to a main or auxiliary load (DPDT function). To accomplish this additional switch, a second hybrid and phase shifter must be added, as shown in Figure 7.34. This configuration then can perform the following switching functions:

- RF source 1 routed to output *B*
- RF source 2 routed to output *A*
- RF source 1 routed to output *A*
- RF source 2 routed to output *B*

The key element in developing such a switch is a high-power phase shifter that does not exhibit reflection characteristics. In this application, the phase shifter allows the line between the hybrids to be electrically lengthened or shortened. The ability to adjust the relative phase between the two input signals to the second hybrid provides the needed control to switch the input signal between the two output ports.

If a continuous analog phase shifter is used, the transfer switch shown in Figure 7.34 also can act as a hot switchless combiner where RF generators 1 and 2 can be combined and fed to either output *A* or *B*. The switching or combining functions are accomplished by changing the physical position of the phase shifter.

Note that it does not matter whether the phase shifter is in one or both legs of the system. It is the phase difference ($\theta_1 - \theta_2$) between the two input legs of the second hybrid that is important. With two-phase shifters, dual drives are required. However, the phase shifter needs only two positions. In a one-phase shifter design, only a single drive is required, but the phase shifter must have four fixed operating positions.

7.4.6.1 Phase Relationships

To better understand the dual-hybrid switching and combining process, it is necessary to examine the primary switching combinations. Table 7.5 lists the various combinations of inputs, relative phase, and output configurations that are possible with the single-phase shifter design.

Using vector analysis, note that when two input signals arrive in phase (mode 1) at ports 1 and 2 with the phase shifter set to 0° , the circuit acts like a crossover network with the power from input port 1 routed to output port 4. Power from input port 2 is routed to output port 3. If the phase shifter is set to 180° , the routing changes, with port 1 being routed to port 3 and port 2 being routed to port 4.

Mode 2 represents the case where one of the dual-input RF generators has failed. The output signal from the first hybrid arrives at the input to the second hybrid with a 90° phase difference. Because the second hybrid introduces a 90° phase shift, the vectors add at port 4 and cancel at port 3. This effectively switches the working transmitter connected to port 1 to output port 4, the load.

By introducing a 180° phase shift between the hybrids, as shown in modes 4 and 5, it is possible to reverse the circuit. This allows the outputs to be on the same side of the circuit as the inputs. This configuration might be useful if generator 1 fails, and all power from generator 2 is directed to a diplexer connected to output 4.

Normal operating configurations are shown in modes 6 and 7. When both generators are running, it is possible to have the combined power routed to either output port. The switching is accomplished by introducing a $\pm 90^\circ$ phase shift between the hybrids.

As shown in the table, it is possible to operate in all the listed modes through the use of a single-phase shifter. The phase shifter must provide four different phase positions. A similar analysis would show that a two-phase shifter design, with two positions for each shifter, is capable of providing the same operational modes.

The key to making hybrid switches work in the real world lies in the phase shifter. The dual 90° hybrid combiner just discussed requires a phase shifter capable of introducing a fixed phase offset of -90° , 0° , $+90^\circ$, and $+180^\circ$. This can be accomplished easily at low-power levels through the use of a sliding short-circuit (trombone-type) line stretcher. However, when high-frequency and high-power signals are being used, the sliding short circuit is not an appropriate design choice. Among the practical methods available are

- Variable-dielectric vane
- Dielectric post
- Variable-phase hybrid

TABLE 7.5

Operating Modes of the Dual 90° Hybrid/Single-Phase Shifter Combiner System

Mode	Input		ϕ	Input Vector	Schematic	Output Vector	Output	
	1	2					3	4
1	$P_1/0^\circ$ $P_1/0^\circ$	$P_2/0^\circ$ $P_2/0^\circ$	0° 180°				P_2 P_1	P_1 P_2
2	$P_1/0^\circ$	$P_2 = 0$	0°				0	P_1
3	$P_1 = 0$	$P_2/0^\circ$	0°				P_2	0
4	$P_1/0^\circ$	$P_2 = 0$	180°				P_1	0
5	$P_1 = 0$	$P_2/0^\circ$	180°				0	P_2
6	$P_1/0^\circ$	$P_2/0^\circ$	90°				0	$P_1 + P_2$
	$P_1/0^\circ$	$P_2/0^\circ$	-90°				$P_1 + P_2$	0

Legend:

- \uparrow = Unit vector port 1
- \uparrow = Unit vector port 2
- $\uparrow\uparrow$ = Paths 1-3, 2-4 $\uparrow\uparrow$ = Paths 1-4, 2-3
- \uparrow or \uparrow 0° phase \leftarrow or \leftarrow -90° phase
- \dots or \downarrow -180° phase
- $\uparrow\downarrow$ $\uparrow\downarrow$ $\leftarrow\rightarrow$ $\rightarrow\leftarrow$ $\downarrow\uparrow$ $\downarrow\uparrow$ = Vector cancellation
- $\leftarrow\rightarrow$ $\uparrow\downarrow$ $\rightarrow\leftarrow$ $\downarrow\uparrow$ $\downarrow\uparrow$ $\downarrow\uparrow$ = Vector addition

7.4.6.2 Variable-Dielectric Vane

The variable-dielectric vane consists of a long dielectric sheet mounted in a section of rectangular waveguide, as illustrated in Figure 7.35. The dielectric sheet is long enough to introduce a 270° phase shift when located in the center of the waveguide. As the dielectric sheet is moved toward the wall, into the lower field, the phase shift decreases. A single-sided phase shifter easily can provide the needed four positions. A two-stage 1/4-wave transformer is used on each end of the sheet to maintain a proper match for any position over the desired operating band. The performance of a typical switchless combiner, using the dielectric vane, is given in Table 7.6.

7.4.6.3 Dielectric Posts

Dielectric posts, shown in Figure 7.36, operate on the same principle as the dielectric vane. The dielectric posts are positioned 1/4-wavelength apart from each other to cancel any mismatch, and to maintain minimal VSWR.

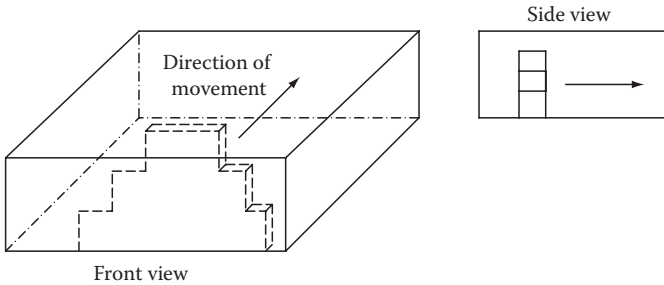


FIGURE 7.35 Dielectric vane switcher, which consists of a long dielectric sheet mounted within a section of rectangular waveguide.

TABLE 7.6

Typical Performance of a Dielectric Vane Phase Shifter

Type	Input	Phase Change (°)	VSWR	Input Attenuation		Output Attenuation	
				1 (dB)	2 (dB)	3 (dB)	4 (dB)
Single input	T1	180	1.06	—	39	0.1	39
	T1	0	1.05	—	39	39	0.1
	T2	180	1.05	39	—	39	0.1
	T2	0	1.06	39	—	0.1	39
Dual input	T1 + T2	270	1.06	—	—	0.1	36
	T1 + T2	90	1.06	—	—	36	0.1

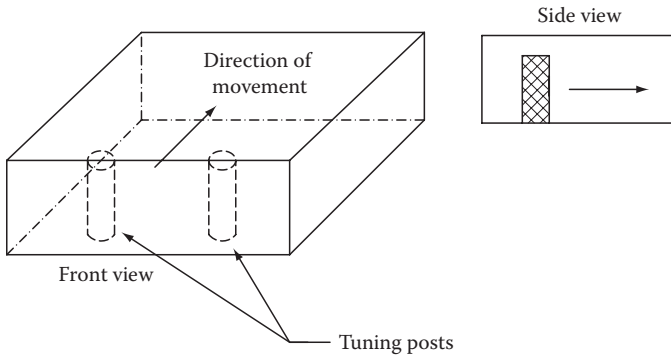


FIGURE 7.36
Dielectric post waveguide phase shifter.

7.4.6.4 Variable-Phase Hybrid

The variable-phase hybrid, shown in Figure 7.37, relies on a 90° hybrid, similar to those used in a combiner. With a unit vector incident on port 1, the power is split by the 90° hybrid. The signal at ports 3 and 4 is reflected by the short circuit. These reflected signals are out of phase at port 1 and in phase at port 2. The relative phase of the hybrid can be changed by moving the short circuit.

The variable-phase hybrid is linear with respect to position. Noncontacting choke-type short circuits, with high front-to-back ratios, are typically used in the device. The performance available from a typical high-power variable-phase switchless combiner is given in Table 7.7.

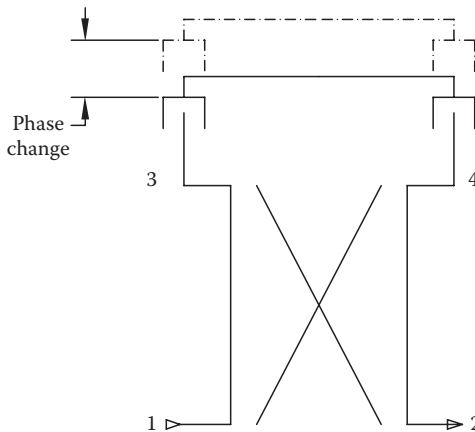


FIGURE 7.37
Variable-phase hybrid phase shifter.

TABLE 7.7

Typical Performance of a Variable-Phase Hybrid Phase Shifter

Type	Input	Phase Change (deg)	VSWR	Input Attenuation		Output Attenuation	
				1 (dB)	2 (dB)	3 (dB)	4 (dB)
Single input	T1	180	1.06	—	36	0.1	52
	T1	0	1.04	—	36	50	0.1
	T2	180	1.06	36	—	52	0.1
	T2	0	1.07	36	—	0.1	50
	T1 + T2	90	1.06	—	—	36	0.1

7.4.7 Phased-Array Antenna Systems

The phased-array antenna, commonly used in radar applications, is a variation on the hot-switching combiner systems discussed in the previous section. The hot switching combiner is used to control the flow of RF energy from one or more generators to two or more loads, while the phased-array antenna is used to modify—in a controlled fashion—the radiating characteristics of an antenna.

Phased-array antennas are *steered* by tilting the phase front independently in two orthogonal directions called the *array coordinates*. Scanning in either array coordinate causes the beam to move along a cone whose center is at the center of the array. As the beam is steered away from the array normal, the projected aperture in the beam's direction varies, causing the beamwidth to vary proportionately.

Arrays can be classified as either active or passive. Active arrays contain duplexers and amplifiers behind every element or group of elements of the array. Passive arrays are driven from a single feed point. Active arrays are capable of high-power operation. Both passive and active arrays must divide the signal from a single transmission line among all the elements of the system. This can be accomplished through one of the following methods:

- *Optical feed*: a single feed, usually a monopulse horn, is used to illuminate the array with a spherical phase front, illustrated in Figure 7.38. Power collected by the rear elements of the array is transmitted through the phase shifters to produce a planar front and steer the array. The energy then may be radiated from the other side of the array or reflected and reradiated through the collecting elements. In the latter case, the array acts as a *steerable reflector*.
- *Corporate feed*: a system utilizing a series-feed network (Figure 7.39) or parallel-feed network (Figure 7.40). Both designs use transmission-line

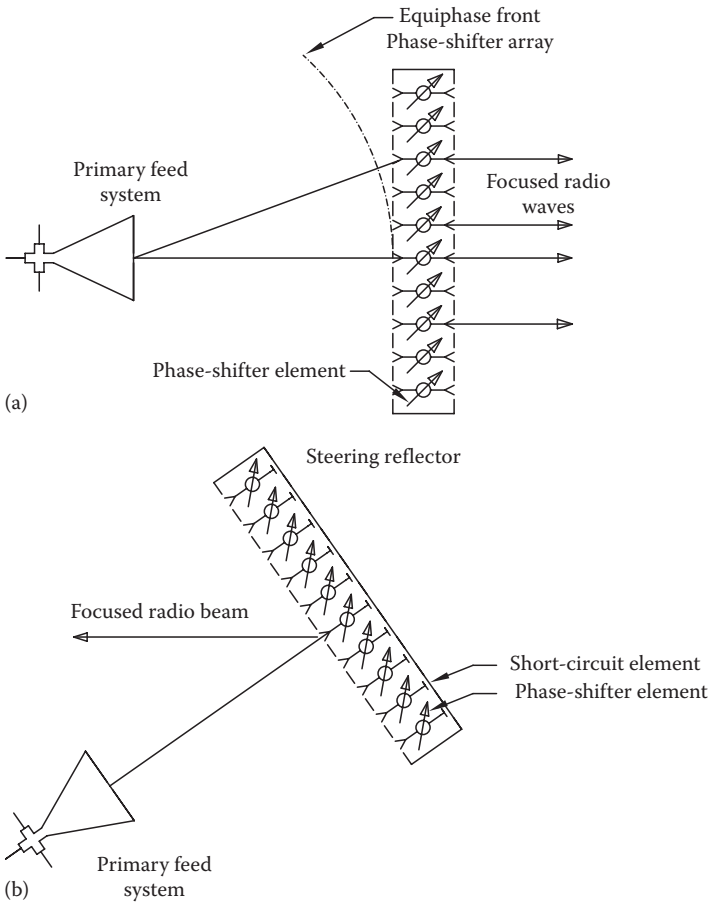


FIGURE 7.38 Optical antenna feed systems: (a) lens and (b) reflector.

components to divide the signal among the elements. Phase shifters can be located at the elements or within the dividing network. Both the series- and parallel-feed systems have several variations, as shown in the figures.

- *Multiple-beam network*: a system capable of forming simultaneous beams with a given array. The *Butler matrix*, shown in Figure 7.41, is one such technique. It connects the N elements of a linear array to N feed points corresponding to N beam outputs. The phase shifter is one of the most critical components of the system. It produces controllable phase shift over the operating band of the array. Digital and analog phase shifters have been developed using both ferrites and *pin* diodes.

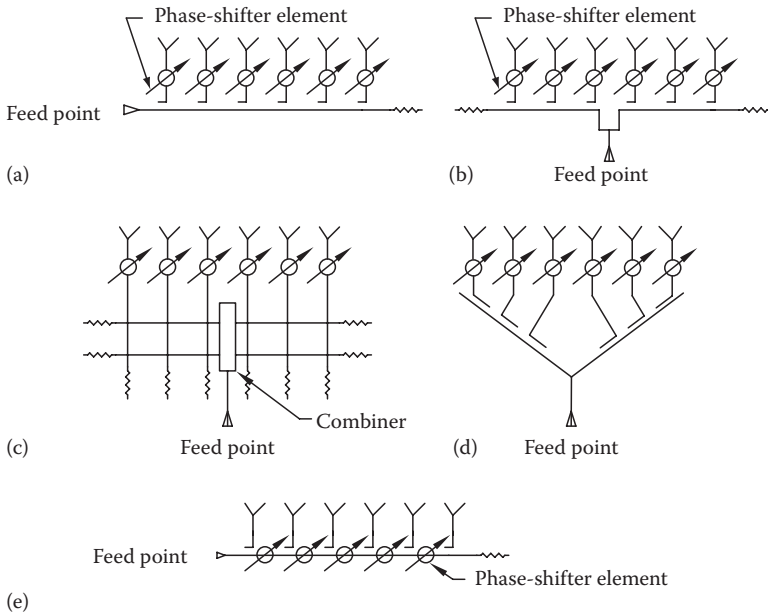


FIGURE 7.39 Series-feed networks: (a) end feed, (b) center feed, (c) separate optimization, (d) equal path length, and (e) series phase shifters.

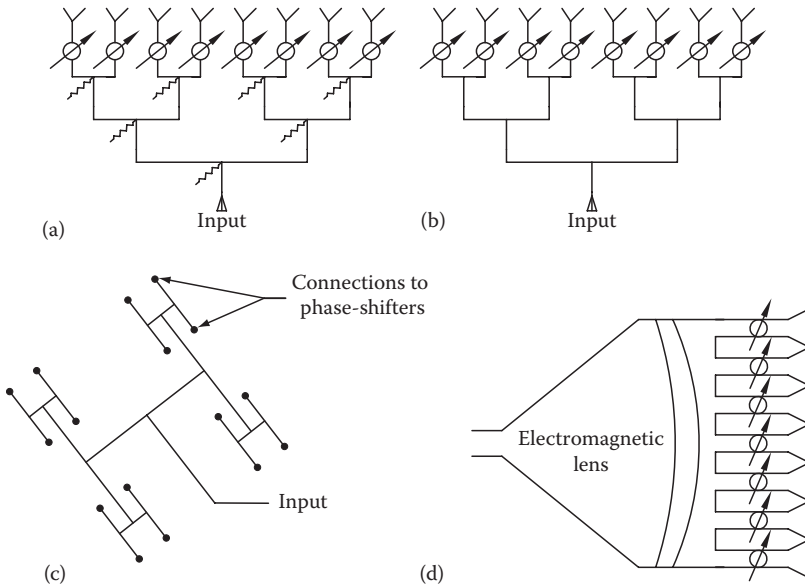


FIGURE 7.40 Types of parallel-feed networks: (a) matched corporate feed, (b) reactive corporate feed, (c) reactive stripline, and (d) multiple reactive divider.

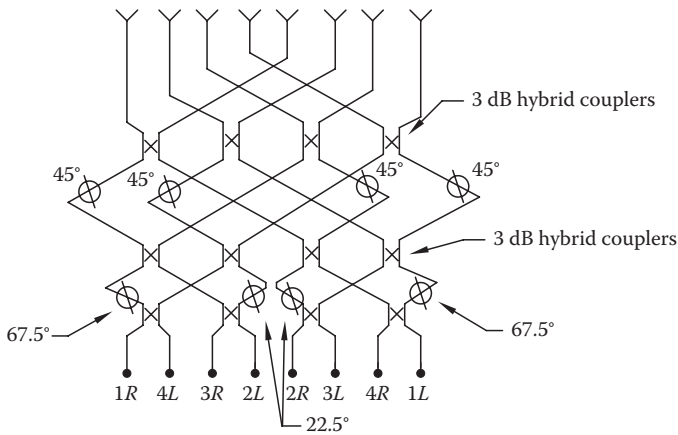


FIGURE 7.41
Butler beam-forming network.

Frequency scan is another type of multiple-beam network, but one that does not require phase shifters, dividers, or beam-steering computers. Element signals are coupled from points along a transmission line. The electrical path length between elements is longer than the physical separation, so a small frequency change will cause a phase change between elements that is large enough to steer the beam. This technique can be applied only to one array coordinate. If a two-dimensional array is required, phase shifters normally are used to scan the other coordinate.

7.4.7.1 Phase-Shift Devices

The design of a phase shifter must meet two primary criteria:

- Low transmission loss
- High power-handling capability

The *Reggia–Spencer* phase shifter is a representative device. As illustrated in Figure 7.42, the device consists of a ferrite rod mounted inside a waveguide that delays the RF signal passing through the waveguide, permitting the array to be steered. The amount of phase shift can be controlled by the current in the solenoid, because of the effect a magnetic field has on the permeability of the ferrite. This design is a *reciprocal phase shifter*, meaning that the device exhibits the same phase shift for signals passing in either direction (forward or reverse). Nonreciprocal phase shifters also are available, where phase-shift polarity reverses with the direction of propagation.

Phase shifters also may be implemented using pin diodes in transmission line networks. One configuration, shown in Figure 7.43, uses diodes as switches

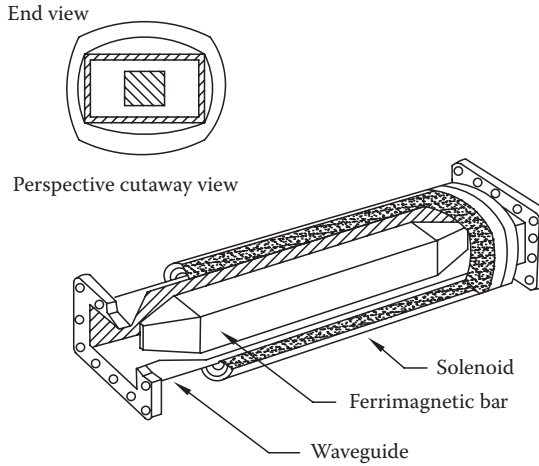


FIGURE 7.42
Basic concept of a Reggia–Spencer phase shifter.

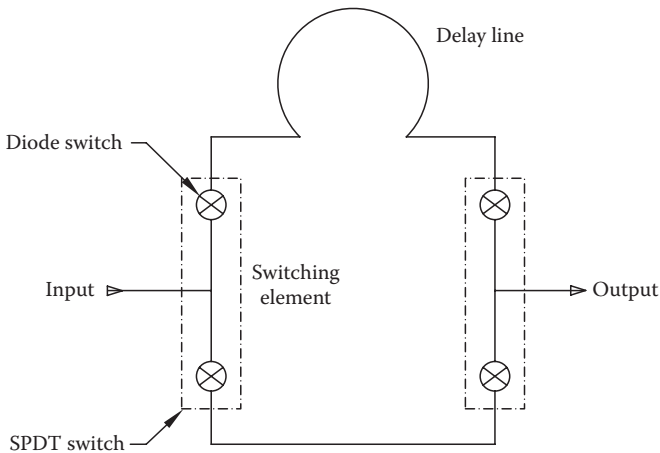


FIGURE 7.43
Switched-line phase shifter using *pin* diodes.

to change the signal path length of the network. A second type uses pin diodes as switches to connect reactive loads across a transmission line. When equal loads are connected with $1/4$ -wave separation, a pure phase shift results.

7.4.7.2 Radar System Duplexer

The duplexer is an essential component of any radar system. The switching elements used in a duplexer include gas tubes, ferrite circulators, and pin diodes. Gas tubes are the simplest. A typical gas-filled *TR tube* is shown in Figure 7.44. Low-power RF signals pass through the tube with little

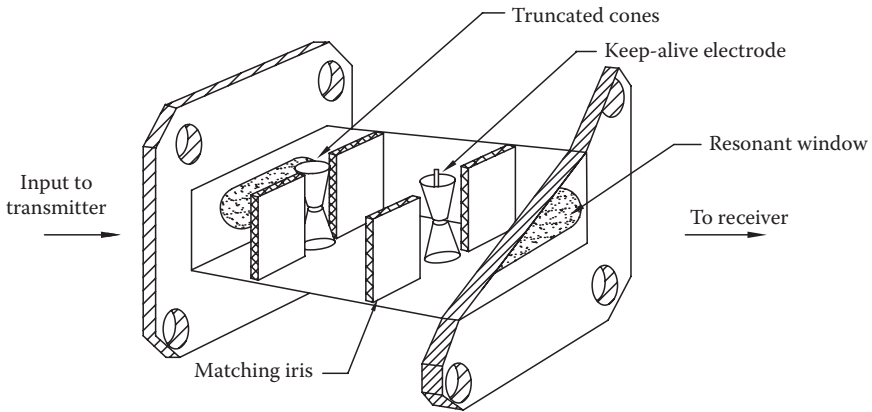


FIGURE 7.44
Typical construction of a TR tube.

attenuation. Higher-power signals, however, cause the gas to ionize and present a short circuit to the RF energy.

Figure 7.45 illustrates a *balanced duplexer* using hybrid junctions and TR tubes. When the transmitter is on, the TR tubes fire and reflect the RF power to the antenna port of the input hybrid. During the receive portion of the radar function, signals picked up by the antenna are passed through the TR tubes and on to the receiver port of the output hybrid.

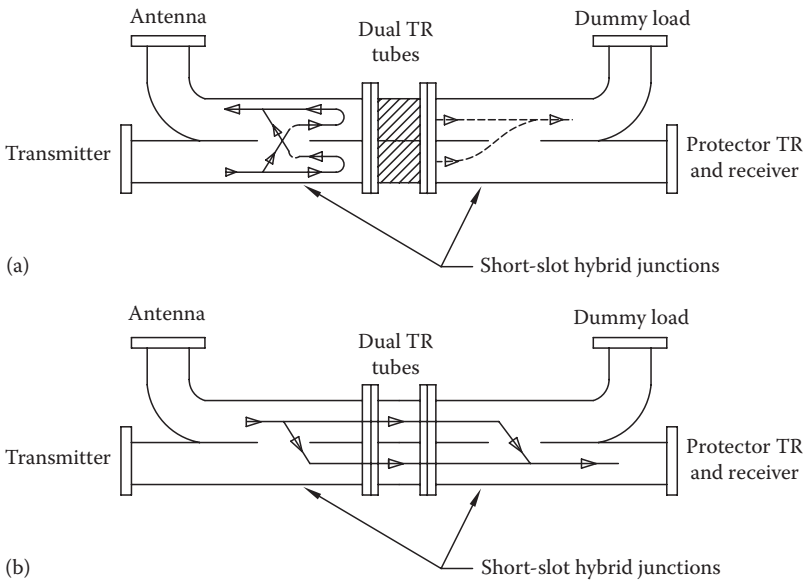


FIGURE 7.45
Balanced duplexer circuit using dual TR tubes and two short-slot hybrid junctions: (a) transmit mode and (b) receive mode.

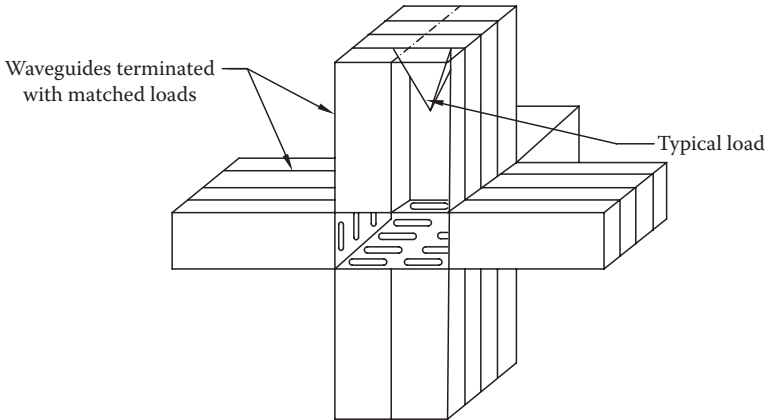


FIGURE 7.46
Construction of a dissipative waveguide filter.

Newer radar systems often use a ferrite circulator as the duplexer. A TR tube is required in the receiver line to protect input circuits from transmitter power reflected by the antenna because of an imperfect match. A four-port circulator generally is used with a load between the transmitter and receiver ports so that power reflected by the TR tube is properly terminated.

Pin diode switches also have been used in duplexers to perform the protective switching function of TR tubes. Pin diodes are more easily applied in coaxial circuitry, and at lower microwave frequencies. Multiple diodes are used when a single diode cannot withstand the expected power.

Microwave filters sometimes are used in the transmit path of a radar system to suppress spurious radiation, or in the receive signal path to suppress spurious interference.

Harmonic filters commonly are used in the transmission chain to absorb harmonic energy output by the system, preventing it from being radiated or reflected back from the antenna. Figure 7.46 shows a filter in which harmonic energy is coupled out through holes in the walls of the waveguide to matched loads.

Narrowband filters in the receive path, often called *preselectors*, are built using mechanically tuned cavity resonators or electrically tuned *TIG resonators*. Preselectors can provide up to 80 dB suppression of signals from other radar transmitters in the same RF band, but at a different operating frequency.

7.5 High-Power Isolators

The high-power ferrite isolator offers the ability to stabilize impedance, isolate the RF generator from load discontinuities, eliminate reflections from the load, and absorb harmonic and intermodulation products. The isolator

also can be used to switch between an antenna or load under full power, or to combine two or more generators into a common load.

Isolators commonly are used in microwave transmitters at low power to protect the output stage from reflections. Ferrite isolators are also available that can handle 500kW or more of forward power with less than 0.1 dB of forward power loss.

7.5.1 Theory of Operation

High-power isolators are three-port versions of a family of devices known as *circulators*. The circulator derives its name from the fact that a signal applied to one of the input ports can travel in only one direction, as shown in Figure 7.47. The input port is isolated from the output port. A signal entering port 1 appears only at port 2; it does not appear at port 3 unless reflected from port 2. An important benefit of this one-way power transfer is that the input VSWR at port 1 is dependent only on the VSWR of the load placed at port 3. In most applications, this load is a resistive (dummy) load that presents a “perfect” load to the RF generator.

The unidirectional property of the isolator results from magnetization of a ferrite alloy inside the device. Through correct polarization of the magnetic field of the ferrite, RF energy will travel through the element in only one direction (port 1 to 2, port 2 to 3, and port 3 to 1). Reversing the polarity of the magnetic field makes it possible for RF flow in the opposite direction.

In the basic design, the ferrite is placed in the center of a Y-junction of three transmission lines, either waveguide or coax. Sections of the material are bonded together to form a thin cylinder perpendicular to the electric field. Even though the insertion loss is low, the resulting power dissipated in the cylinder can be as high as 2% of the forward power. Special provisions must be made for heat removal. It is efficient heat-removal capability that makes high-power operation possible.

The insertion loss of the ferrite must be kept low so that minimal heat is dissipated. Values of ferrite loss on the order of 0.05 dB have been produced. This equates to an efficiency of 98.9%. Additional losses from the transmission line and matching structure contribute slightly to loss. The overall loss is typically less than 0.1 dB, or 98% efficiency. The ferrite element in a high-power system is usually water cooled in a closed-loop path that uses an external radiator.

The two basic circulator implementations are shown in Figure 7.47a and b. These designs consist of Y-shaped conductors sandwiched between magnetized ferrite discs [2]. The final shape, dimensions, and type of material vary according to frequency of operation, power-handling requirements, and the method of coupling. The *distributed constant circulator* is a broadband device, not quite as efficient in terms of insertion loss and leg-to-leg isolation, and considerably more expensive to produce. It is useful,

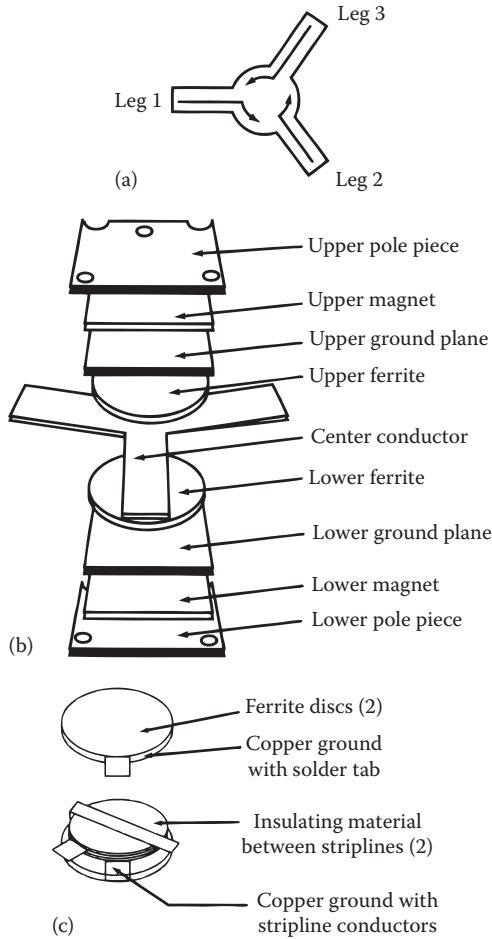


FIGURE 7.47

Basic characteristics of a circulator: (a) operational schematic, (b) distributed constant circulator, and (c) lump constant circulator. (After Surette, R.A., *Combiners and combining networks*, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1368–1381, 1996.)

however, in applications where broadband isolation is required. More common is the *lump constant circulator*, a less expensive and more efficient, but narrowband, design.

At least one filter is always installed directly after an isolator, because the ferrite material of the isolator generates harmonic signals. If an ordinary band-pass or band-reject filter is not to be used, a harmonic filter will be needed.

7.5.2 Applications

The high-power isolator permits an RF generator to operate with high performance and reliability despite a load that is less than optimum. The problems

presented by ice formations on a transmitting antenna provide a convenient example. Ice buildup will detune an antenna, resulting in reflections back to the transmitter and high VSWR. If the VSWR is severe enough, transmitter power will have to be reduced to keep the system on the air. An isolator, however, permits continued operation with no degradation in signal quality. Power output is affected only to the extent of the reflected energy, which is dissipated in the resistive load.

A high-power isolator also can be used to provide a stable impedance for devices that are sensitive to load variations, such as klystrons. This allows the device to be tuned for optimum performance regardless of the stability of the RF components located after the isolator. Figure 7.48 shows the output of a wideband (6 MHz) klystron operating into a resistive load, and into an antenna system. The power loss is the result of an impedance difference. The periodicity of the ripple shown in the trace is a function of the distance of the reflections from the source.

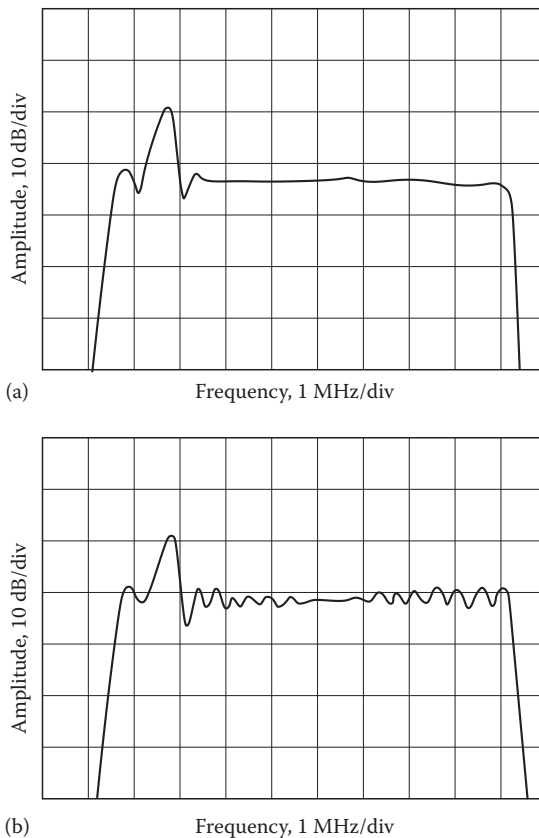


FIGURE 7.48

Output of a klystron operating into different loads through a high-power isolator: (a) resistive load and (b) an antenna system.

7.5.2.1 Hot Switch

The circulator can be made to perform a switching function if a short circuit is placed at the output port. Under this condition, all input power will be reflected back into the third port. The use of a high-power stub on port 2, therefore, permits redirecting the output of an RF generator to port 3.

At odd $1/4$ -wave positions, the stub appears as a high impedance and has no effect on the output port. At even $1/4$ -wave positions, the stub appears as a short circuit. Switching between the antenna and a test load, for example, can be accomplished by moving the shorting element $1/4$ wavelength.

7.5.2.2 Diplexer

An isolator may be configured to combine the output signals of two transmitters into a single output for the antenna. The approach is shown in Figure 7.49. A single notch cavity is added (as shown). The two signals will thus be routed to the antenna.

7.5.2.3 Multiplexer

A multiplexer can be formed by cascading multiple circulators, as illustrated in Figure 7.50. Filters must be added, as shown. The primary drawback of this approach is the increased power dissipation that occurs in circulators nearest the antenna (or load).

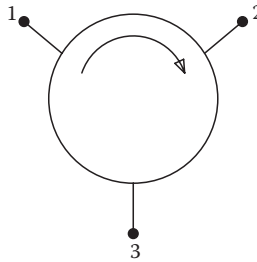


FIGURE 7.49
Use of a circulator in a diplexer application.

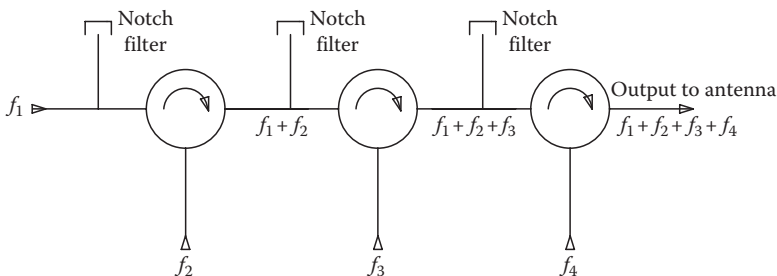


FIGURE 7.50
Using multiple circulators to form a multiplexer.

References

1. Harrison, C., Passive filters, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 279–290, 1996.
2. Surette, R. A., Combiners and combining networks, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 1368–1381, 1996.

Bibliography

- Andrew Corporation, Circular waveguide: System planning, installation and tuning, Technical Bulletin 1061 H, Orland Park, IL, 1980.
- Andrew Corporation, Broadcast transmission line systems, Technical Bulletin 1063 H, Orland Park, IL, 1982.
- Ben-Dov, O. and C. Plummer, Doubly truncated waveguide, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, January 1989.
- Benson, K. B. and J. C. Whitaker, *Television and Audio Handbook for Technicians and Engineers*, McGraw-Hill, New York, 1989.
- Cablewave Systems, The broadcaster's guide to transmission line systems, Technical Bulletin 21A, North Haven, CT, 1976.
- Cablewave Systems, Rigid coaxial transmission lines, cablewave systems catalog 700, North Haven, CT, 1989.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, 1992.
- DeComier, B., Inside FM multiplexer systems, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1988.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 2nd edn., McGraw-Hill, New York, 1982.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.
- Golio, M., *RF and Microwave Applications and Systems*, CRC Press, Boca Raton, FL, 2008.
- Golio, M., *The RF and Microwave Handbook*, 2nd edn., CRC Press, Boca Raton, FL, 2008.
- Heymans, D., Hot switches and combiners, *Broadcast Engineering*, Overland Park, KS, December 1987.
- Jordan, E. C., *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 7th edn., Howard W. Sams, Indianapolis, IN, 1985.
- Krohe, G. L., Using circular waveguide, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1986.
- Perelman, R. and T. Sullivan, Selecting flexible coaxial cable, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1988.
- Stenberg, J. T., Using super power isolators in the broadcast plant, in *Proceedings of the Broadcast Engineering Conference*, Society of Broadcast Engineers, Indianapolis, IN, 1988.
- Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.

- Vaughan, T. and E. Pivit, High power isolator for UHF television, in *Proceedings of the NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1989.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1990.
- Whitaker, J. C., G. DeSantis, and C. Paulson, *Interconnecting Electronic Systems*, CRC Press, Boca Raton, FL, 1993.

8

Properties of Materials

8.1 Metals*

Metals serve many different functions in the realization of radio frequency (RF) and microwave devices, each of which results in different requirements that must be met. Thus, the optimum metal for each application will vary widely. Considerations include a wide range of electrical, mechanical, chemical, and thermal properties.

A first-order consideration in the choice of metals for many electrical applications is the electrical resistance of the metal conductor. DC resistance of a metal rod is given by

$$R = \frac{\rho L}{A} \quad (8.1)$$

where

R is the resistance of the rod

ρ is the resistivity of the metal

L is the length of the rod

A is the cross-sectional area of the rod

The dc electrical properties of metals are also sometimes discussed in terms of conductivity. Conductivity is the inverse of resistivity given by

$$\sigma = \frac{1}{\rho} \quad (8.2)$$

where σ is the conductivity of the material. For most applications, high conductivity, or conversely low resistivity, is desirable. The resistivity of a number of metals is listed in Table 8.1. The tabulated data represent an average resistivity value from a number of sources. Values for resistivity of a metal can vary significantly because of small impurity differences between samples. Pure aluminum may exhibit a 5% increase in resistivity with the introduction of only 0.01% impurities, for example. Metals such as brass and

* This section is adapted from Golio, M., Metals, in *The RF and Microwave Handbook*, 2nd edn., M. Golio (ed.), CRC Press, Boca Raton, FL, 2008.

TABLE 8.1

Electrical Resistivity of Several Metals at Room Temperature
in $10^{-8} \Omega\text{m}$

Metal	Electrical Resistivity
Aluminum	2.7
Antimony	38.6
Beryllium	3.6
Brass	5.6
Chromium	12.7
Copper	1.7
Gold	2.2
Kovar	49.0
Lead	21.1
Magnesium	4.4
Manganese	147.2
Molybdenum	5.3
Nickel	7.0
Palladium	10.4
Platinum	10.4
Silver	1.6
Tantalum	13.0
Titanium	44.9
Tungsten	5.4
Vanadium	22.9
Zinc	6.0

Sources: Smith, A., *Radio Frequency Principles and Applications*, IEEE Press, New York, 1998; Cutnell, J. and Johnson, K.: *Physics*. 1995. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission; James, A., and Lord, M., *Macmillan's Chemical and Physical Data*, Macmillan, London, U.K., 1992; Kaye, G., and Laby, T., *Tables of Physical and Chemical Constants*, 15th edn., Longman, London, U.K., 1993; Samsonov, G. (ed.), *Handbook of the Physicochemical Properties of the Elements*, IFI-Plenum, New York, 1968; Lide, D. (ed.), *Chemical Rubber Company Handbook of Chemistry and Physics*, 79th edn., CRC Press, Boca Raton, FL, 1998; Ellis, H. (ed.), *Nuffield Advanced Science Book of Data*, Longman, London, U.K., 1972; Oberg, E. et al., *Machinery Handbook*, 24th edn., Industrial Press, New York, 2000; Mitchell, M., *J. Appl. Phys.*, 46, 4742, 1975; Dean, J. (ed.), *Lange's Handbook of Chemistry*, 14th edn., McGraw-Hill, New York, 1992; Seitz, F., *The Modern Theory of Solids*, Dover Publications, New York, 1987; Jung, W. et al., *Phys. Rev. B (Solid State)*, 15(2), 659, 1977.

Kovar are alloys and may exhibit even wider variations in properties based on the percentage of elemental metals in the composition. Brass is an alloy of copper and zinc, while Kovar is a nickel–cobalt ferrous alloy.

The resistivity of metal is also a function of temperature. For small perturbations in temperature, this temperature dependence may be characterized by the equation

$$\alpha = \frac{1}{\rho} \frac{\partial \rho}{\partial T} \quad (8.3)$$

where α is the temperature coefficient of resistivity. The measured temperature dependence of resistivity changes as a function of temperature and can vary significantly at temperatures near 0 K or well above room temperature. For most applications, a low value for the temperature coefficient of resistivity is desirable. Table 8.2 presents the temperature coefficient of resistivity at room temperature for several metals that might be chosen for their low resistivity values [1,2].

TABLE 8.2

Temperature Coefficient of Metal Resistivity at Room Temperature in (1/K)

Metal	Temperature Coefficient of Resistivity
Aluminum	4.38×10^{-3}
Copper	3.92×10^{-3}
Gold	3.61×10^{-3}
Molybdenum	4.60×10^{-3}
Silver	3.71×10^{-3}

Sources: Smith, A., *Radio Frequency Principles and Applications*, IEEE Press, New York, 1998; Cutnell, J. and Johnson, K.: *Physics*. 1995. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission; James, A., and Lord, M. *Macmillan's Chemical and Physical Data*, Macmillan, London, U.K., 1992; Kaye, G., and Laby, T., *Tables of Physical and Chemical Constants*, 15th edn., Longman, London, U.K., 1993; Samsonov, G. (ed.), *Handbook of the Physicochemical Properties of the Elements*, IFI-Plenum, New York, 1968; Lide, D. (ed.), *Chemical Rubber Company Handbook of Chemistry and Physics*, 79th edn., CRC Press, Boca Raton, FL, 1998; Ellis, H. (ed.), *Nuffield Advanced Science Book of Data*, Longman, London, U.K., 1972; Oberg, E. et al., *Machinery Handbook*, 24th edn., Industrial Press, New York, 2000; Mitchell, M., *J. Appl. Phys.*, 46, 4742, 1975; Dean, J. (ed.), *Lange's Handbook of Chemistry*, 14th edn., McGraw-Hill, New York, 1992; Seitz, F., *The Modern Theory of Solids*, Dover Publications, New York, 1987; Jung, W. et al., *Phys. Rev. B (Solid State)*, 15(2), 659, 1977.

8.1.1 Skin Depth

As discussed in Chapter 7, an electromagnetic field can penetrate into a conductor only a minute distance at microwave frequencies. The field amplitude decays exponentially from its surface value according to

$$A = e^{-x/\delta_s} \quad (8.4)$$

where

x is the normal distance into the conductor measured from the surface

δ_s is the skin depth

The skin depth or depth of penetration into a metal is defined as the distance the wave must travel in order to decay by an amount equal to $e^{-1} = 0.368$ or 8.686 dB. The skin depth δ_s is given by

$$\delta_s = \frac{1}{\sqrt{\pi f \mu \sigma}} \quad (8.5)$$

where

f is the frequency

σ is the metal conductivity

μ is the permeability of the metal given as

$$\mu = \mu_0 \mu_r \quad (8.6)$$

where

μ_0 equal to the permeability of free space

μ_r the relative permeability of the metal

For most metals used as conductors for microwave and RF applications, the relative permeability, $\mu_r = 1$. The relative permeability of ferroelectric materials such as iron and steel are typically on the order of several hundred.

Skin depth is closely related to the shielding effectiveness of a metal since the attenuation of electric field strength into a metal can be expressed as Equation 8.4. For static or low-frequency fields, the only method of shielding a space is by surrounding it with a high-permeability material. For RF and microwave frequencies, however, a thin sheet or screen of metal serves as an effective shield from electric fields.

Skin depth can be an important consideration in the development of guided wave and reflecting structures for high-frequency work. For the best conductors, skin depth is on the order of microns for 1 GHz fields. Since electric fields cannot penetrate very deeply into a conductor, all current is concentrated near the surface. As conductivity or frequency approach infinity, skin depth approaches zero and the current is contained in a narrower and narrower region. For this reason, only the properties of the surface metal affect RF or microwave resistance. A poor conductor with

a thin layer of high-conductivity metal will exhibit the same RF conduction properties as a solid high-conductivity structure [1–3].

8.1.2 Heat Conduction

One of the uses of metal in the development of RF and microwave parts and modules is as a heat spreader. For many applications that involve high-power-density electronic components, the efficient removal of heat is of great importance in order to preserve component reliability.

The one-dimensional heat flow equation that applies to metals (as well as other media) is given as

$$q = kA \frac{\partial T}{\partial x} \quad (8.7)$$

where

q is the heat flow

k is the thermal conductivity of the metal

A is the cross-sectional area for heat flow

$\partial T/\partial x$ is the temperature gradient across the metal

For applications where good heat-sinking characteristics are desired, high thermal conductivity, k , is desirable.

8.1.3 Thermal Expansion

Because RF and microwave components are often exposed to significant temperature excursions, consideration of the coefficient of linear expansion of the metals must be included in making a metal selection for many applications. When temperature is increased, the average distance between atoms is also increased. This leads to an expansion of the whole solid body with increasing temperature. The changes to the linear dimension of the metal can be characterized by

$$\beta = \frac{1}{l} \frac{\Delta l}{\Delta T} \quad (8.8)$$

where

β is the coefficient of linear expansion

l is the linear dimension of the material

ΔT is the change in temperature

Δl is the change in linear dimension arising from the change in temperature

Linear expansion properties of metals are important whenever metallic structures are bonded to other materials in an electronic assembly. When two materials with dissimilar thermal expansion characteristics are bonded

together, significant stress is experienced during temperature excursions. This stress can result in one or both of the materials breaking or cracking and this can result in degraded electrical performance. The best choice of metals to match thermal linear expansion properties is, therefore, determined by the thermal coefficient of linear expansion of the material that is used with the metal. Table 8.3 presents the coefficients of linear expansion for several metals as well as other materials that are often used for RF and microwave circuits.

TABLE 8.3

Thermal Coefficient of Linear Expansion of Some of the Materials Used in Microwave and RF Packaging Applications (at Room Temperature, in $10^{-6}/K$)

Material	Thermal Coefficient of Expansion
<i>Dielectrics</i>	
Aluminum nitride	4
Alumina 96%	6
Beryllia	6.5
Diamond	1
Glass-ceramic	4-8
Quartz (fused)	0.54
<i>Metals</i>	
Aluminum	23.1
Antimony	10.9
Beryllium	11.6
Brass	18.9
Chromium	4.9
Copper	16.5
Gold	14.1
Kovar	5.5
Lead	29.1
Magnesium	25.4
Manganese	22.5
Molybdenum	5.0
Nickel	13.0
Palladium	11.5
Platinum	8.9
Silver	18.9
Tantalum	6.4
Titanium	8.7
Tungsten	4.4
Vanadium	8.1
Zinc	30.1

TABLE 8.3 (continued)

Thermal Coefficient of Linear Expansion of Some of the Materials Used in Microwave and RF Packaging Applications (at Room Temperature, in $10^{-6}/\text{K}$)

Material	Thermal Coefficient of Expansion
<i>Semiconductors</i>	
GaAs	5.9
Silicon	2.6
Silicon carbide	2.2

Sources: Smith, A., *Radio Frequency Principles and Applications*, IEEE Press, New York, 1998; Cutnell, J. and Johnson, K.: *Physics*. 1995. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission; James, A., and Lord, M. *Macmillan's Chemical and Physical Data*, Macmillan, London, U.K., 1992; Kaye, G., and Laby, T. *Tables of Physical and Chemical Constants*, 15th edn., Longman, London, U.K., 1993; Samsonov, G. (ed.), *Handbook of the Physicochemical Properties of the Elements*, IFI-Plenum, New York, 1968; Lide, D. (ed.), *Chemical Rubber Company Handbook of Chemistry and Physics*, 79th edn., CRC Press, Boca Raton, FL, 1998; Ellis, H. (ed.), *Nuffield Advanced Science Book of Data*, Longman, London, U.K., 1972; Oberg, E. et al., *Machinery Handbook*, 24th edn., Industrial Press, New York, 2000; Mitchell, M., *J. Appl. Phys.*, 46, 4742, 1975; Dean, J. (ed.), *Lange's Handbook of Chemistry*, 14th edn., McGraw-Hill, New York, 1992; Seitz, F., *The Modern Theory of Solids*, Dover Publications, New York, 1987; Jung, W. et al., *Phys. Rev. B (Solid State)*, 15(2), 659, 1977.

8.2 Dielectrics*

When electrical circuits are built, the physical environment consists of conductors separated from each other by insulating materials. Depending on the application, the properties of the insulating materials are chosen to satisfy requirements of low conductivity, a high or low dielectric permittivity, and desirable loss properties. In most cases the high field properties of the materials do not play a role. This is also true with respect to electrical breakdown. However, there are exceptions. For instance, in transmitters the fields can be very large and the choice of insulators has to reflect those conditions.

In most applications it is desirable that the material properties be independent of the strength of the applied field. This is not always true and there are now materials becoming available that can be used as transmission line phase shifters for steerable antennas, taking advantage of the field-dependent dielectric permittivity.

* This section is adapted from Etzold, K. F., Dielectrics, in *The RF and Microwave Handbook*, 2nd edn., M. Golio (ed.), CRC Press, Boca Raton, FL, 2008.

Probably the most important property of dielectrics in circuit applications is the dielectric permittivity, often colloquially called the dielectric constant. This quantity relates the electric field in the material to the free charge on the surfaces. Superficially, the connection between this property and the application of dielectrics in a circuit seems rather tenuous. Consider now, however, the capacity C of a device

$$C = \frac{q}{V} \quad (8.9)$$

where

q is the charge on the plates

V is the voltage across the device

Usually in a discrete capacitor (understood here as a circuit element) we would like to store as much charge as possible in a given volume. Therefore, if we can somehow double the stored charge, the capacity C will double. What allows us to do this is the choice of the material that fills the device. Thus, the material in this example has twice the charge storage capability or dielectric permittivity. This quantity is typically labeled by the Greek letter epsilon (ϵ).

Consider, as an example, a parallel plate capacitor. This simple geometry is easily analyzed, and in general we can calculate the capacity of a device exactly if it is possible to calculate the electric field. For the parallel plate device the capacity is

$$C = \epsilon \frac{A}{d} \quad (8.10)$$

where

A is the area

d is the separation between the plates

If the material between the plates is vacuum, $\epsilon = 8.85 \times 10^{-12}$ F/m or in units that are a little easier to remember, 8.85 pF/m. The unit of capacity is the F. A capacitor with square 1 m \times 1 m plates and a separation of 1 m has a capacity of 8.85 pF, using Equation 8.10. This assumes parallel field lines between the plates that clearly will not be the case here unless a special, guarded geometry is chosen (to guard a circuit, conductors at the same potential are placed nearby. There are no field lines between equipotentials, and thus there are no contributions to the capacity). We can consider Equation 8.10 for the capacity as a definition of dielectric permittivity. In general, the capacity is proportional to ϵ and the capacity of a device is given by

$$C = \epsilon \times \text{Geometry factor} \quad (8.11)$$

Also for practical applications, the dielectric permittivity is usually specified relative to that of vacuum:

$$\epsilon = \epsilon_0 \times k \quad (8.12)$$

Thus, ϵ_0 is defined as the dielectric permittivity of vacuum and k is the relative dielectric permittivity, most often referred to as simply, but somewhat inaccurately, the dielectric permittivity or also dielectric constant; k of course is dimensionless. As an example, the dielectric permittivity of typical glass-epoxy circuit board is 4.5, where this is understood to be the relative dielectric permittivity. On physical grounds all materials have a dielectric permittivity greater than 1. Also, for the same reason the dielectric permittivity of vacuum has no frequency dependence. In Table 8.4, the important properties of selected insulating materials used in engineering application are given.

So far, we have treated dielectrics as essentially perfect and lossless insulators that allow energy storage greater than that of vacuum. However, as can be seen from the fact that the table has a loss factor column, frequency-dependent energy losses will occur in all materials (except again in vacuum) [16]. As could be anticipated, there are a number of loss mechanisms. Perhaps the easiest loss mechanism to deal with conceptually is the low-frequency conductivity of the material. If there is a current flow as a result of an applied field, energy will be dissipated. The dc conductivity can be due to the basic properties of the materials such as the presence of impurities and their associated energy levels. It can also be due to crystalline defects, and conduction can take place along the surfaces of crystallites that constitute many materials (e.g., ceramics). Most often suitable materials can be chosen selected for the desired properties. In Table 8.4, the conductivity for selected materials is presented. It is seen that some natural and man-made materials have an

TABLE 8.4

Properties of Some Typical Engineering Insulating Materials

Material	k	Loss	Frequency	Conductivity
Vacuum	1.00	0	All	Zero
Air	1.0006	0		
Glass Vycor 7910	3.8	9.1×10^{-4}		
Glass Corning 0080	6.75	5.8×10^{-2}		
Al_2O_3	8.5	10^{-3}	1 MHz	
Teflon (PTFE)	2.0	2×10^{-4}	1 MHz	10^{-17}
Arlon 25N circuit board	3.28	2.5×10^{-3}	1 MHz	
Epoxy-glass circuit board	4.5			
Beryllium oxide	7.35			
Diamond	5.58			10^{-16}
PZT (lead zirconium oxide)	~ 1000			
Undoped silicon	11.8			
TaO_5	28			
Quartz (SiO_2)	3.75–4.1	2×10^{-4}		
Mica (Ruby)	6.5–8.7	3.5×10^{-4}		
Water	78.2	0.04	1 MHz	

extremely low conductivity. Indeed, the time constant for loss of charge for a free-floating (disconnected) device can be weeks. The time constant τ (in s) of an RC network is defined as

$$\tau = RC \quad (8.13)$$

where R is the resistance in parallel with the capacitor C (R is in Ω and C in F). Consider Teflon (PTFE). The dc conductivity at room temperature (25°C) is 10^{-17} .

Therefore, for this material the time constant for the charge to decay to $1/e$ of its original value is 23 days using the dielectric permittivity of 2.0 from Table 8.4. This demonstrates that there exist insulating materials for the cabling and supporting structures that are suitable for applications where only the smallest leakage can be tolerated.

8.2.1 Frequency Dependence of the Properties

We need now to consider the behavior at all user frequencies. These extend from dc into the microwave and optical range. In this range a number of different loss mechanisms have to be considered. But, before this can be done, the definition of the dielectric permittivity needs to be extended, so that ac effects can be properly considered. The losses are accounted for if an imaginary component is added to the definition of the dielectric permittivity. Thus,

$$\epsilon = \epsilon' - j\epsilon'' \quad (8.14)$$

where

ϵ' is the real part of the dielectric permittivity

j is the square root of -1

The existence of losses is now explicitly given by the imaginary quantity ϵ'' . It is also obvious that the term "permittivity" as a property of dielectrics is somewhat of a misnomer and is partly historical. In fact both the real and imaginary parts are functions of frequency. As is often the case in describing physical parameters that are complex (in the mathematical sense), the two quantities in Equation 8.14 are actually related. This is a consequence of the fact that the electromagnetic behavior is governed by Maxwell's equations, the absence of discontinuities and also that causality holds, a statement that relates to the time evolution of electric fields. Thus, the so-called Kramers-Kronig equations relate the two quantities [17]

$$\begin{aligned} \epsilon'(\omega) &= \frac{1}{\pi} \int \frac{\epsilon''}{x - \omega} dx \\ \epsilon''(\omega) &= \frac{1}{\pi} \int \frac{\epsilon'}{x - \omega} dx \end{aligned} \quad (8.15)$$

$$-\infty < \text{limits} < +\infty$$

$$\omega = 2\pi f$$

where f is the frequency in Hz and where the dependence of the real and imaginary parts of ϵ on frequency is now explicitly stated and where the integrals are the Cauchy principal values. Thus, we only need to know the real or imaginary parts and the corresponding complimentary part can be calculated. For instance, if we know the real part of the dielectric response over the full frequency range, we can calculate the losses by integration. Even if our knowledge of the real part (capacity) is only over a limited frequency range, the complementary part (loss) can often still be estimated, albeit with reduced accuracy.

8.3 Tabular Data*

8.3.1 Electrical Resistivity of Pure Metals

Table 8.5 gives the electrical resistivity, in units of $10^{-8} \Omega \text{ m}$, for 28 common metallic elements as a function of temperature. The data refer to polycrystalline samples. The number of significant figures indicates the accuracy of the values. However, at low temperatures (especially below 50 K) the electrical resistivity is extremely sensitive to sample purity. Thus, the low-temperature values refer to samples of specified purity and treatment.

Table 8.6 gives resistivity values in the neighborhood of room temperature for other metallic elements that have not been studied over an extended temperature range.

8.3.2 Electrical Resistivity of Selected Alloys

Values of the resistivity contained in Table 8.7 are given in units of $10^{-8} \Omega \text{ m}$. General comments in the preceding table for pure metals also apply here.

8.3.3 Resistance of Wires

Table 8.8 gives the approximate resistance of various metallic conductors. The values have been computed from the resistivities at 20°C , except as otherwise stated, and for the dimensions of wire indicated. Owing to differences in purity in the case of elements and of composition in alloys, the values can be considered only as approximations.

8.3.4 Dielectric Constants of Common Materials

The data in Tables 8.9–8.11 refer to temperatures in the range 17°C – 22°C .

* This section is adapted from Golio, M. (ed.), Appendix A—Mathematics, symbols, and physical constants, in *The RF and Microwave Handbook*, 2nd edn., CRC Press, Boca Raton, FL, 2008.

TABLE 8.5
Electrical Resistivity in $10^{-8} \Omega \text{ m}$

T (K)	Aluminum	Barium	Beryllium	Calcium	Cesium	Chromium	Copper
1	0.000100	0.081	0.0332	0.045	0.0026		0.00200
10	0.000193	0.189	0.0332	0.047	0.243		0.00202
20	0.000755	0.94	0.0336	0.060	0.86	0.00280	
40	0.0181	2.91	0.0367	0.175	1.99		0.0239
60	0.0959	4.86	0.067	0.40	3.07		0.0971
80	0.245	6.83	0.075	0.65	4.16		0.215
100	0.442	8.85	0.133	0.91	5.28	1.6	0.348
150	1.006	14.3	0.510	1.56	8.43	4.5	0.699
200	1.587	20.2	1.29	2.19	12.2	7.7	1.046
273	2.417	30.2	3.02	3.11	18.7	11.8	1.543
293	2.650	33.2	3.56	3.36	20.5	12.5	1.678
298	2.709	34.0	3.70	3.42	20.8	12.6	1.712
300	2.733	34.3	3.76	3.45	21.0	12.7	1.725
400	3.87	51.4	6.76	4.7		15.8	2.402
500	4.99	72.4	9.9	6.0		20.1	3.090
600	6.13	98.2	13.2	7.3		24.7	3.792
700	7.35	130	16.5	8.7		29.5	4.514
800	8.70	168	20.0	10.0	34.6	5.262	
900	10.18	216	23.7	11.4		39.9	6.041

T (K)	Gold	Hafnium	Iron	Lead	Lithium	Magnesium	Manganese
1	0.0220	1.00	0.0225		0.007	0.0062	7.02
10	0.0226	1.00	0.0238		0.008	0.0069	18.9
20	0.035	1.11	0.0287		0.012	0.0123	54
40	0.141	2.52	0.0758		0.074	0.074	116
60	0.308	4.53	0.271		0.345	0.261	131
80	0.481	6.75	0.693	4.9	1.00	0.557	132
100	0.650	9.12	1.28	6.4	1.73	0.91	132
150	1.061	15.0	3.15	9.9	3.72	1.84	136
200	1.462	21.0	5.20	13.6	5.71	2.75	139
273	2.051	30.4	8.57	19.2	8.53	4.05	143
293	2.214	33.1	9.61	20.8	9.28	4.39	144
298	2.255	33.7	9.87	21.1	9.47	4.48	144
300	2.271	34.0	9.98	21.3	9.55	4.51	144
400	3.107	48.1	16.1	29.6	13.4	6.19	147
500	3.97	63.1	23.7	38.3		7.86	149
600	4.87	78.5	32.9			9.52	151
700	5.82		44.0			11.2	152
800	6.81		57.1			12.8	
900	7.86					14.4	

(continued)

TABLE 8.5 (continued)

Electrical Resistivity in $10^{-8} \Omega \cdot \text{m}$

T (K)	Molybdenum	Nickel	Palladium	Platinum	Potassium	Rubidium	Silver
1	0.00070	0.0032	0.0200	0.002	0.0008	0.0131	0.00100
10	0.00089	0.0057	0.0242	0.0154	0.0160	0.109	0.00115
20	0.00261	0.0140	0.0563	0.0484	0.117	0.444	0.0042
40	0.0457	0.068	0.334	0.409	0.480	1.21	0.0539
60	0.206	0.242	0.938	1.107	0.90	1.94	0.162
80	0.482	0.545	1.75	1.922	1.34	2.65	0.289
100	0.858	0.96	2.62	2.755	1.79	3.36	0.418
150	1.99	2.21	4.80	4.76	2.99	5.27	0.726
200	3.13	3.67	6.88	6.77	4.26	7.49	1.029
273	4.85	6.16	9.78	9.6	6.49	11.5	1.467
293	5.34	6.93	10.54	10.5	7.20	12.8	1.587
298	5.47	7.12	10.73	10.7	7.39	13.1	1.617
300	5.52	7.20	10.80	10.8	7.47	13.3	1.629
400	8.02	11.8	14.48	14.6			2.241
500	10.6	17.7	17.94	18.3			2.87
600	13.1	25.5	21.2	21.9			3.53
700	15.8	32.1	24.2	25.4			4.21
800	18.4	35.5	27.1	28.7			4.91
900	21.2	38.6	29.4	32.0			5.64

T (K)	Sodium	Strontium	Tantalum	Tungsten	Vanadium	Zinc	Zirconium
1	0.0009	0.80	0.10	0.00016		0.0100	0.250
10	0.0015	0.80	0.102	0.000137	0.0145	0.0112	0.253
20	0.016	0.92	0.146	0.00196	0.039	0.0387	0.357
40	0.172	1.70	0.751	0.0544	0.304	0.306	1.44
60	0.447	2.68	1.65	0.266	1.11	0.715	3.75
80	0.80	3.64	2.62	0.606	2.41	1.15	6.64
100	1.16	4.58	3.64	1.02	4.01	1.60	9.79
150	2.03	6.84	6.19	2.09	8.2	2.71	17.8
200	2.89	9.04	8.66	3.18	12.4	3.83	26.3
273	4.33	12.3	12.2	4.82	18.1	5.46	38.8
293	4.77	13.2	13.1	5.28	19.7	5.90	42.1
298	4.88	13.4	13.4	5.39	20.1	6.01	42.9
300	4.93	13.5	13.5	5.44	20.2	6.06	43.3
400		17.8	18.2	7.83	28.0	8.37	60.3
500		22.2	22.9	10.3	34.8	10.82	76.5
600		26.7	27.4	13.0	41.1	13.49	91.5
700		31.2	31.8	15.7	47.2		104.2
800		35.6	35.9	18.6	53.1		114.9
900			40.1	21.5	58.7		123.1

TABLE 8.6

Electrical Resistivity of Pure Metals

Element	T (K)	Electrical Resistivity
		$10^{-8} \Omega \text{ m}$
Antimony	273	39
Bismuth	273	107
Cadmium	273	6.8
Cerium	290–300	82.8
Cobalt	273	5.6
Dysprosium	290–300	92.6
Erbium	290–300	86.0
Europium	290–300	90.0
Gadolinium	290–300	131
Gallium	273	13.6
Holmium	290–300	81.4
Indium	273	8.0
Iridium	273	4.7
Lanthanum	290–300	61.5
Lutetium	290–300	58.2
Mercury	273	94.1
Neodymium	290–300	64.3
Niobium	273	15.2
Osmium	273	8.1
Polonium	273	40
Praseodymium	290–300	70.0
Promethium	290–300	75
Protactinium	273	17.7
Rhenium	273	17.2
Rhodium	273	4.3
Ruthenium	273	7.1
Samrium	290–300	94.0
Scandium	290–300	56.2
Terbium	290–300	115
Thallium	273	15
Thorium	273	14.7
Thulium	290–300	67.6
Tin	273	11.5
Titanium	273	39
Uranium	273	28
Ytterbium	290–300	25.0
Yttrium	290–300	59.6

TABLE 8.7
Electrical Resistivity of Selected Alloys

	273 K	293 K	300 K	350 K	400 K	273 K	293 K	300 K	350 K	400 K	
<i>Alloy—Aluminum—Copper</i>											
Wt % Al						<i>Alloy—Copper—Nickel</i>					
						<i>Wt % Cu</i>					
99 ^a	2.51	2.74	2.82	3.38	3.95	99 ^c	2.71	2.85	2.91	3.27	3.62
95 ^a	2.88	3.10	3.18	3.75	4.33	95 ^c	7.60	7.71	7.82	8.22	8.62
90 ^b	3.36	3.59	3.67	4.25	4.86	90 ^c	13.69	13.89	13.96	14.40	14.81
85 ^b	3.87	4.10	4.19	4.79	5.42	85 ^c	19.63	19.83	19.90	20.32	20.70
80 ^b	4.33	4.58	4.67	5.31	5.99	80 ^c	25.46	25.66	25.72	26.12 ^m	26.44 ^m
70 ^b	5.03	5.31	5.41	6.16	6.94	70 ⁱ	36.67	36.72	36.76	36.85	36.89
60 ^b	5.56	5.88	5.99	6.77	7.63	60 ⁱ	45.43	45.38	45.35	45.20	45.01
50 ^b	6.22	6.55	6.67	7.55	8.52	50 ⁱ	50.19	50.05	50.01	49.73	49.50
40 ^c	7.57	7.96	8.10	9.12	10.2	40 ^c	47.42	47.73	47.82	48.28	48.49
30 ^c	11.2	11.8	12.0	13.5	15.2	30 ⁱ	40.19	41.79	42.34	44.51	45.40
25 ^f	16.3 ^m	17.2	17.6	19.8	22.2	25 ^c	33.46	35.11	35.69	39.67 ^m	42.81 ^m
15 ^h	—	12.3	—	—	—	15 ^c	22.00	23.35	23.85	27.60	31.38
19 ^g	10.8 ^m	11.0	11.1	11.7	12.3	10 ^c	16.65	17.82	18.26	21.51	25.19
5 ^e	9.43	9.61	9.68	10.2	10.7	5 ^c	11.49	12.50	12.90	15.69	18.78
1 ^b	4.46	4.60	4.65	5.00	5.37	1 ^c	7.23	8.08	8.37	10.63 ^m	13.18 ^m
<i>Alloy—Aluminum—Magnesium</i>											
Wt % Al						<i>Alloy—Copper—Palladium</i>					
						<i>Wt % Cu</i>					
99 ^c	2.96	3.18	3.26	3.82	4.39	99 ^c	2.10	2.23	2.27	2.59	2.92
95 ^c	5.05	5.28	5.36	5.93	6.51	95 ^c	4.21	4.35	4.40	4.74	5.08
90 ^c	7.52	7.76	7.85	8.43	9.02	90 ^c	6.89	7.03	7.08	7.41	7.74
85	—	—	—	—	—	85 ^c	9.48	9.61	9.66	10.01	10.36

(continued)

Table 8.7 (continued)

Electrical Resistivity of Selected Alloys

	273 K	293 K	300 K	350 K	400 K	273 K	293 K	300 K	350 K	400 K
70	—	—	—	—	—	16.87	17.01	17.06	17.41	17.78
60	—	—	—	—	—	21.73	21.87	21.92	22.30	22.69
50	—	—	—	—	—	27.62	27.79	27.86	28.25	28.64
40	—	—	—	—	—	35.31	35.51	35.57	36.03	36.47
30	—	—	—	—	—	46.50	46.66	46.71	47.11	47.47
25	—	—	—	—	—	46.25	46.45	46.52	46.99 ^m	47.43 ^m
15	—	—	—	—	—	36.52	36.99	37.16	38.28	39.35
10 ^b	17.1	17.4	17.6	18.4	19.2	28.90	29.51	29.73	31.19 ^m	32.56 ^m
5 ^b	13.1	13.4	13.5	14.3	15.2	20.00	20.75	21.02	22.84 ^m	24.54 ^m
1 ^a	5.92	6.25	6.37	7.20	8.03	11.90	12.67	12.93 ^m	14.82 ^m	16.68 ^m

Alloy—Copper—Gold

Wt % Cu

99 ^c	1.73	1.86 ^m	1.91 ^m	2.24 ^m	2.58 ^m	1.84	1.97	2.02	2.36	2.71
95 ^c	2.41	2.54 ^m	2.59 ^m	2.92 ^m	3.26 ^m	2.78	2.92	2.97	3.33	3.69
90 ^c	3.29	4.42 ^m	3.46 ^m	3.79 ^m	4.12 ^m	3.66	3.81	3.86	4.25	4.63
85 ^c	4.20	4.33	4.38 ^m	4.71 ^m	5.05 ^m	4.37	4.54	4.60	5.02	5.44
80 ^c	5.15	5.28	5.32	5.65	5.99	5.01	5.19	5.26	5.71	6.17
70 ^c	7.12	7.25	7.30	7.64	7.99	5.87	6.08	6.15	6.67	7.19
60 ^c	9.18	9.13	9.36	9.70	10.05	—	—	—	—	—
50 ^c	11.07	11.20	11.25	11.60	11.94	—	—	—	—	—
40 ^c	12.70	12.85	12.90 ^m	13.27 ^m	13.65 ^m	—	—	—	—	—
30 ^c	13.77	13.93	13.99 ^m	14.38 ^m	14.78 ^m	—	—	—	—	—
25 ^c	13.93	14.09	14.14	14.54	14.94	—	—	—	—	—
15 ^c	12.75	12.91	12.96 ^m	13.36 ^m	13.77	—	—	—	—	—

Alloy—Copper—Zinc

Wt % Cu

99 ^b	1.84	1.97	2.02	2.36	2.71
95 ^b	2.78	2.92	2.97	3.33	3.69
90 ^b	3.66	3.81	3.86	4.25	4.63
85 ^b	4.37	4.54	4.60	5.02	5.44
80 ^b	5.01	5.19	5.26	5.71	6.17
70 ^b	5.87	6.08	6.15	6.67	7.19
60	—	—	—	—	—
50	—	—	—	—	—
40	—	—	—	—	—
30	—	—	—	—	—
25	—	—	—	—	—
15	—	—	—	—	—

10 ^c	10.70	10.86	10.91	11.31	11.72	10	—	—	—	—
5 ^c	7.25	7.41 ^m	7.46	7.87	8.28	5	—	—	—	—
1 ^c	3.40	3.57	3.62	4.03	4.45	1	—	—	—	—
<i>Alloy—Gold—Palladium</i>										
<i>Wt % Au</i>										
99 ^c	2.69	2.86	2.91	3.32	3.73	99 ^a	10.9	12.0	12.4	18.7
95 ^c	5.21	5.35	5.41	5.79	6.17	95 ^c	18.7	19.9	20.2	26.8
90 ⁱ	8.01	8.17	8.22	8.56	8.93	90 ^c	24.2	25.5	25.9	33.2
85 ^b	10.50 ^m	10.66	10.72 ^m	11.100 ^m	11.48 ^m	85 ^c	27.8	29.2	29.7	37.3
80 ^b	12.75	12.93	12.99	13.45	13.93	80 ^c	30.1	31.6	32.2	40.0
70 ^c	18.23	18.46	18.54	19.10	19.67	70 ^b	32.3	33.9	34.4	42.4
60 ^b	26.70	26.94	27.01	27.63 ^m	28.23 ^m	60 ^c	53.8	57.1	58.2	73.9
50 ^a	27.23	27.63	27.76	28.64 ^m	29.42 ^m	50 ^d	28.4	30.6	31.4	43.7
40 ^a	24.65	25.23	25.42	26.74	27.95	40 ^d	19.6	21.6	22.5	34.0
30 ^b	20.82	21.49	21.72	23.35	24.92	30 ^c	15.3	17.1	17.7	27.4
25 ^b	18.86	19.53	19.77	21.51	23.19	25 ^b	14.3	15.9	16.4	25.1
15 ^a	15.08	15.77	16.01	17.80	19.61	15 ^c	12.6	13.8	14.2	21.1
10 ^a	13.25	13.95	14.20 ^m	16.00 ^m	17.81 ^m	10 ^c	11.4	12.5	12.9	18.9
5 ^a	11.49 ^m	12.21	12.46 ^m	14.26 ^m	16.07 ^m	5 ^c	9.66	10.6	10.9	16.1 ^m
1 ^a	10.07	10.85 ^m	11.12 ^m	12.99 ^m	14.80 ^m	1 ^b	7.17	7.94	8.12	12.8
<i>Alloy—Gold—Silver</i>										
<i>Wt % Au</i>										
99 ^b	2.58	2.75	2.80 ^m	3.22 ^m	3.63 ^m	99 ^b	1.891	2.007	2.049	2.35
95 ^a	4.58	4.74	4.79	5.19	5.59	95 ^b	3.58	3.70	3.74	4.04
90 ⁱ	6.57	6.73	6.78	7.19	7.58	90 ^b	5.82	5.94	5.98	6.28
85 ⁱ	8.14	8.30	8.36 ^m	8.75	9.15	85 ^k	7.92 ^m	8.04 ^m	8.08	8.38 ^m
80 ⁱ	9.34	9.50	9.55	9.94	10.33	80 ^k	10.01	10.13	10.17	10.47
<i>Alloy—Silver—Palladium</i>										
<i>Wt % Ag</i>										
99 ^b	1.891	2.007	2.049	2.35	2.66	99 ^b	1.891	2.007	2.049	2.35
95 ^b	3.58	3.70	3.74	4.04	4.34	95 ^b	3.58	3.70	3.74	4.04
90 ^b	5.82	5.94	5.98	6.28	6.59	90 ^b	5.82	5.94	5.98	6.28
85 ^k	7.92 ^m	8.04 ^m	8.08	8.38 ^m	8.68 ^m	85 ^k	7.92 ^m	8.04 ^m	8.08	8.38 ^m
80 ^k	10.01	10.13	10.17	10.47	10.78	80 ^k	10.01	10.13	10.17	10.47

(continued)

Table 8.7 (continued)

Electrical Resistivity of Selected Alloys

	273 K	293 K	300 K	350 K	400 K		273 K	293 K	300 K	350 K	400 K
70 ^j	10.70	10.86	10.91	11.29	11.68 ^m	70 ^k	14.53	14.65	14.69	14.99	15.30
60 ^j	10.92	11.07	11.12	11.50	11.87	60 ^k	20.9	21.1	21.2	21.6	22.0
50 ^j	10.23	10.37	10.42	10.78	11.14	50 ^k	31.2	31.4	31.5	32.0	32.4
40 ^j	8.92	9.06	9.11	9.46 ^m	9.81	40 ^k	42.2	42.2	42.2	42.3	42.3
30 ^a	7.34	7.47	7.52	7.85	8.19	30 ^b	40.4	40.6	40.7	41.3	41.7
25 ^a	6.46	6.59	6.63	6.96	7.30 ^m	25 ^k	36.67 ^m	37.06	37.19	38.1 ^m	38.8 ^m
15 ^a	4.55	4.67	4.72	5.03	5.34	15 ⁱ	27.08 ^m	26.68 ^m	27.89 ^m	29.3 ^m	30.6 ^m
10 ^a	3.54	3.66	3.71	4.00	4.31	10 ⁱ	21.69	22.39	22.63	24.3	25.9
5 ⁱ	2.52	2.64 ^m	2.68 ^m	2.96 ^m	3.25 ^m	5 ^b	15.98	16.72	16.98	18.8 ^m	20.5 ^m
1 ^b	1.69	1.80	1.84 ^m	2.12 ^m	2.42 ^m	1 ^a	11.06	11.82	12.08 ^m	13.92 ^m	15.70 ^m

^a Uncertainty in resistivity is $\pm 2\%$.^b Uncertainty in resistivity is $\pm 3\%$.^c Uncertainty in resistivity is $\pm 5\%$.^d Uncertainty in resistivity is $\pm 7\%$ below 300 K and $\pm 5\%$ at 300 and 400 K.^e Uncertainty in resistivity is $\pm 7\%$.^f Uncertainty in resistivity is $\pm 8\%$.^g Uncertainty in resistivity is $\pm 10\%$.^h Uncertainty in resistivity is $\pm 12\%$.ⁱ Uncertainty in resistivity is $\pm 4\%$.^j Uncertainty in resistivity is $\pm 1\%$.^k Uncertainty in resistivity is $\pm 3\%$ up to 300 K and $\pm 4\%$ above 300 K.^l Uncertainty in resistivity is $\pm 2\%$ up to 300 K and $\pm 4\%$ above 300 K.^m In temperature range where no experimental data are available.

TABLE 8.8

Resistance of Selected Wires

B. and S. No.	Ω/cm	Ω/ft	B. and S. No.	Ω/cm	Ω/ft
<i>Advance (0°C) Q = 48.0 × 10⁻⁶ Ω cm</i>			<i>Brass Q = 7.00 × 10⁻⁶ Ω cm</i>		
10	0.000912	0.0278	10	0.000133	0.00406
12	0.00145	0.0442	12	0.000212	0.00645
14	0.00231	0.0703	14	0.000336	0.0103
16	0.00367	0.112	16	0.000535	0.0163
18	0.00583	0.178	18	0.000850	0.0259
20	0.00927	0.283	20	0.00135	0.0412
22	0.0147	0.449	22	0.00215	0.0655
24	0.0234	0.715	24	0.00342	0.104
26	0.0373	1.14	26	0.00543	0.166
27	0.0470	1.43	27	0.00686	0.209
28	0.0593	1.81	28	0.00864	0.263
30	0.0942	2.87	30	0.0137	0.419
32	0.150	4.57	32	0.0219	0.666
34	0.238	7.26	34	0.0348	1.06
36	0.379	11.5	36	0.0552	1.68
40	0.958	29.2	40	0.140	4.26
<i>Aluminum Q = 2.828 × 10⁻⁶ Ω cm</i>			<i>Climax Q = 87.0 × 10⁻⁶ Ω cm</i>		
10	0.0000538	0.00164	10	0.00165	0.0504
12	0.0000855	0.00260	12	0.00263	0.0801
14	0.000136	0.00414	14	0.00418	0.127
16	0.000216	0.00658	16	0.00665	0.203
18	0.000344	0.0105	18	0.0106	0.322
20	0.000546	0.0167	20	0.0168	0.512
22	0.000869	0.0265	22	0.0267	0.815
24	0.00138	0.0421	24	0.0425	1.30
26	0.00220	0.0669	26	0.0675	2.06
27	0.00277	0.0844	27	0.0852	2.60
28	0.00349	0.106	28	0.107	3.27
30	0.00555	0.169	30	0.171	5.21
32	0.00883	0.269	32	0.272	8.28
34	0.0140	0.428	34	0.432	13.2
36	0.0223	0.680	36	0.687	20.9
40	0.0564	1.72	40	1.74	52.9
<i>Constantan (0°C) Q = 44.1 × 10⁻⁶ Ω cm</i>			<i>Excello Q = 92.0 × 10⁻⁶ Ω cm</i>		
10	0.000838	0.0255	10	0.00175	0.0533
12	0.00133	0.0406	12	0.00278	0.0847
14	0.00212	0.0646	14	0.00442	0.135

(continued)

TABLE 8.8 (continued)

Resistance of Selected Wires

B. and S. No.	Ω/cm	Ω/ft	B. and S. No.	Ω/cm	Ω/ft
16	0.00337	0.103	16	0.00703	0.214
18	0.00536	0.163	18	0.0112	0.341
20	0.00852	0.260	20	0.0178	0.542
22	0.0135	0.413	22	0.0283	0.861
24	0.0215	0.657	24	0.0449	1.37
26	0.0342	1.04	26	0.0714	2.18
27	0.0432	1.32	27	0.0901	2.75
28	0.0545	1.66	28	0.114	3.46
30	0.0866	2.64	30	0.181	5.51
32	0.138	4.20	32	0.287	8.75
34	0.219	6.67	34	0.457	13.9
36	0.348	10.6	36	0.726	22.1
40	0.880	26.8	40	1.84	56.0
<i>Copper, annealed $Q = 1.724 \times 10^{-6} \Omega cm$</i>			<i>German silver $Q = 33.0 \times 10^{-6} \Omega cm$</i>		
10	0.0000328	0.000999	10	0.000627	0.0191
12	0.0000521	0.00159	12	0.000997	0.0304
14	0.0000828	0.00253	14	0.00159	0.0483
16	0.000132	0.00401	16	0.00252	0.0768
18	0.000209	0.00638	18	0.00401	0.122
20	0.000333	0.0102	20	0.00638	0.194
22	0.000530	0.0161	22	0.0101	0.309
24	0.000842	0.0257	24	0.0161	0.491
26	0.00134	0.0408	26	0.0256	0.781
27	0.00169	0.0515	27	0.0323	0.985
28	0.00213	0.0649	28	0.0408	1.24
30	0.00339	0.103	30	0.0648	1.97
32	0.00538	0.164	32	0.103	3.14
34	0.00856	0.261	34	0.164	4.99
36	0.0136	0.415	36	0.260	0.794
40	0.0344	1.05	40	0.659	20.1
<i>Eureka (0°C) $Q = 47.0 \times 10^{-6} \Omega cm$</i>			<i>Gold $Q = 2.44 \times 10^{-6} \Omega cm$</i>		
10	0.000893	0.0272	10	0.0000464	0.00141
12	0.00142	0.0433	12	0.0000737	0.00225
14	0.00226	0.688	14	0.000117	0.00357
16	0.00359	0.109	16	0.000186	0.00568
18	0.00571	0.174	18	0.000296	0.00904
20	0.00908	0.277	20	0.000471	0.0144
22	0.0144	0.440	22	0.000750	0.0228
24	0.0230	0.700	24	0.00119	0.0363

TABLE 8.8 (continued)

Resistance of Selected Wires

B. and S. No.	Ω/cm	Ω/ft	B. and S. No.	Ω/cm	Ω/ft
26	0.0365	1.11	26	0.00189	0.0577
27	0.0460	1.40	27	0.00239	0.0728
28	0.0580	1.77	28	0.00301	0.0918
30	0.0923	2.81	30	0.00479	0.146
32	0.147	4.47	32	0.00762	0.232
34	0.233	7.11	34	0.0121	0.369
36	0.371	11.3	36	0.0193	0.587
40	0.938	28.6	40	0.0487	1.48
<i>Iron Q = 10.0 × 10⁻⁶ Ω cm</i>			<i>Manganin Q = 44.0 × 10⁻⁶ Ω cm</i>		
10	0.000190	0.00579	10	0.000836	0.0255
12	0.000302	0.00921	12	0.00133	0.0405
14	0.000481	0.0146	14	0.00211	0.0644
16	0.000764	0.0233	16	0.00336	0.102
18	0.00121	0.0370	18	0.00535	0.163
20	0.00193	0.0589	20	0.00850	0.259
22	0.00307	0.0936	22	0.0135	0.412
24	0.00489	0.149	24	0.0215	0.655
26	0.00776	0.237	26	0.0342	1.04
27	0.00979	0.299	27	0.0431	1.31
28	0.0123	0.376	28	0.0543	1.66
30	0.0196	0.598	30	0.0864	2.63
32	0.0312	0.952	32	0.137	4.19
34	0.0497	1.51	34	0.218	6.66
36	0.789	2.41	36	0.347	10.6
40	0.200	6.08	40	0.878	26.8
<i>Lead Q = 22.0 × 10⁻⁶ Ω cm</i>			<i>Molybdenum Q = 5.7 × 10⁻⁶ Ω cm</i>		
10	0.000418	0.0127	10	0.000108	0.00330
12	0.000665	0.0203	12	0.000172	0.00525
14	0.00106	0.0322	14	0.000274	0.00835
16	0.00168	0.0512	16	0.000435	0.0133
18	0.00267	0.0815	18	0.000693	0.0211
20	0.00425	0.130	20	0.00110	0.0336
22	0.00676	0.206	22	0.00175	0.0534
24	0.0107	0.328	24	0.00278	0.0849
26	0.0171	0.521	26	0.00443	0.135
27	0.0215	0.657	27	0.00558	0.170
28	0.0272	0.828	28	0.00704	0.215
30	0.0432	1.32	30	0.0112	0.341

(continued)

TABLE 8.8 (continued)

Resistance of Selected Wires

B. and S. No.	Ω/cm	Ω/ft	B. and S. No.	Ω/cm	Ω/ft
32	0.0687	2.09	32	0.0178	0.542
34	0.109	3.33	34	0.0283	0.863
36	0.174	5.29	36	0.0450	1.37
40	0.439	13.4	40	0.114	3.47
<i>Magnesium</i> $Q = 4.6 \times 10^{-6} \Omega cm$			<i>Monel metal</i> $Q = 42.0 \times 10^{-6} \Omega cm$		
10	0.0000874	0.00267	10	0.000798	0.0243
12	0.000139	0.00424	12	0.00127	0.0387
14	0.000221	0.00674	14	0.00202	0.0615
16	0.000351	0.0107	16	0.00321	0.0978
18	0.000559	0.0170	18	0.00510	0.156
20	0.000889	0.0271	20	0.00811	0.247
22	0.00141	0.0431	22	0.0129	0.393
24	0.00225	0.0685	24	0.0205	0.625
26	0.00357	0.109	26	0.0326	0.994
27	0.00451	0.137	27	0.0411	1.25
28	0.00568	0.173	28	0.0519	1.58
30	0.00903	0.275	30	0.0825	2.51
32	0.0144	0.438	32	0.131	4.00
34	0.0228	0.696	34	0.209	6.36
36	0.0363	1.11	36	0.331	10.1
40	0.0918	2.80	40	0.838	25.6
<i>Nichrome</i> $Q = 150.0 \times 10^{-6} \Omega cm$			<i>Silver (18°C)</i> $Q = 1.629 \times 10^{-6} \Omega cm$		
10	0.0021281	0.06488	10	0.0000310	0.000944
12	0.0033751	0.1029	12	0.0000492	0.00150
14	0.0054054	0.1648	14	0.0000783	0.00239
16	0.0085116	0.2595	16	0.000124	0.00379
18	0.0138383	0.4219	18	0.000198	0.00603
20	0.0216218	0.6592	20	0.000315	0.00959
22	0.0346040	1.055	22	0.000500	0.0153
24	0.0548088	1.671	24	0.000796	0.0243
26	0.0875760	2.670	26	0.00126	0.0386
28	0.1394328	4.251	27	0.00160	0.0486
30	0.2214000	6.750	28	0.00201	0.0613
32	0.346040	10.55	30	0.00320	0.0975
34	0.557600	17.00	32	0.00509	0.155
36	0.885600	27.00	34	0.00809	0.247
38	1.383832	42.19	36	0.0129	0.392
40	2.303872	70.24	40	0.0325	0.991

TABLE 8.8 (continued)

Resistance of Selected Wires

B. and S. No.	Ω/cm	Ω/ft	B. and S. No.	Ω/cm	Ω/ft
<i>Nickel</i> $Q = 7.8 \times 10^{-6} \Omega cm$			<i>Steel, piano wire</i> ($0^\circ C$) $Q = 11.8 \times 10^{-6} \Omega cm$		
10	0.000148	0.00452	10	0.000224	0.00684
12	0.000236	0.00718	12	0.000357	0.0109
14	0.000375	0.0114	14	0.000567	0.0173
16	0.000596	0.0182	16	0.000901	0.0275
18	0.000948	0.0289	18	0.00143	0.0437
20	0.00151	0.0459	20	0.00228	0.0695
22	0.00240	0.0730	22	0.00363	0.110
24	0.00381	0.116	24	0.00576	0.176
26	0.00606	0.185	26	0.00916	0.279
27	0.00764	0.233	27	0.0116	0.352
28	0.00963	0.294	28	0.0146	0.444
30	0.0153	0.467	30	0.0232	0.706
32	0.0244	0.742	32	0.0368	1.12
34	0.0387	1.18	34	0.0586	1.79
36	0.0616	1.88	36	0.0931	2.84
40	0.156	4.75	40	0.236	7.18
<i>Platinum</i> $Q = 10.0 \times 10^{-6} \Omega cm$			<i>Steel, invar</i> (35% Ni) $Q = 81.0 \times 10^{-6} \Omega cm$		
10	0.000190	0.00579	10	0.00154	0.0469
12	0.000302	0.00921	12	0.00245	0.0746
14	0.000481	0.0146	14	0.00389	0.119
16	0.000764	0.0233	16	0.00619	0.189
18	0.00121	0.0370	18	0.00984	0.300
20	0.00193	0.0589	20	0.0156	0.477
22	0.00307	0.0936	22	0.0249	0.758
24	0.00489	0.149	24	0.0396	1.21
26	0.00776	0.237	26	0.0629	1.92
27	0.00979	0.299	27	0.0793	2.42
28	0.0123	0.376	28	0.100	3.05
30	0.0196	0.598	30	0.159	4.85
32	0.0312	0.952	32	0.253	7.71
34	0.0497	1.51	34	0.402	12.3
36	0.0789	2.41	36	0.639	19.5
40	0.200	6.08	40	1.62	49.3
<i>Tantalum</i> $Q = 15.5 \times 10^{-6} \Omega cm$			<i>Tungsten</i> $Q = 5.51 \times 10^{-6} \Omega cm$		
10	0.000295	0.00898	10	0.000105	0.00319
12	0.000468	0.0143	12	0.000167	0.00508

(continued)

TABLE 8.8 (continued)

Resistance of Selected Wires

B. and S. No.	Ω/cm	Ω/ft	B. and S. No.	Ω/cm	Ω/ft
14	0.000745	0.0227	14	0.000265	0.00807
16	0.00118	0.0361	16	0.000421	0.0128
18	0.00188	0.0574	18	0.000669	0.0204
20	0.00299	0.0913	20	0.00106	0.0324
22	0.00476	0.145	22	0.00169	0.0516
24	0.00757	0.231	24	0.00269	0.0820
26	0.0120	0.367	26	0.00428	0.130
27	0.0152	0.463	27	0.00540	0.164
28	0.0191	0.583	28	0.00680	0.207
30	0.0304	0.928	30	0.0108	0.330
32	0.0484	1.47	32	0.0172	0.524
34	0.0770	2.35	34	0.0274	0.834
36	0.122	3.73	36	0.0435	1.33
40	0.309	9.43	40	0.110	3.35
<i>Tin</i> $Q = 11.5 \times 10^{-6} \Omega cm$			<i>Zinc</i> ($0^\circ C$) $Q = 5.75 \times 10^{-6} \Omega cm$		
10	0.000219	0.00666	10	0.000109	0.00333
12	0.000348	0.0106	12	0.000174	0.00530
14	0.000553	0.0168	14	0.000276	0.00842
16	0.000879	0.0268	16	0.000439	0.0134
18	0.00140	0.0426	18	0.000699	0.0213
20	0.00222	0.0677	20	0.00111	0.0339
22	0.00353	0.108	22	0.00177	0.0538
24	0.00562	0.171	24	0.00281	0.0856
26	0.00893	0.272	26	0.00446	0.136
27	0.0113	0.343	27	0.00563	0.172
28	0.0142	0.433	28	0.00710	0.216
30	0.0226	0.688	30	0.0113	0.344
32	0.0359	1.09	32	0.0180	0.547
34	0.0571	1.74	34	0.0286	0.870
36	0.0908	2.77	36	0.0454	1.38
40	0.230	7.00	40	0.115	3.50

TABLE 8.9
Dielectric Constants of Solids

Material	Freq. (Hz)	Dielectric Constant	Material	Freq. (Hz)	Dielectric Constant
Acetamide	4×10^8	4.0	Diphenylmethane	4×10^8	2.7
Acetanilide	—	2.9	Dolomite \perp optic axis	10^8	8.0
Acetic acid (2°C)	4×10^8	4.1	Dolomite \parallel	10^8	6.8
Aluminum oleate	4×10^8	2.40	Ferrous oxide (15°C)	10^8	14.2
Ammonium bromide	10^8	7.1	Iodine	10^8	4
Ammonium chloride	10^8	7.0	Lead acetate	10^8	2.6
Antimony trichloride	10^8	5.34	Lead carbonate (15°C)	10^8	18.6
Apatite \perp optic axis	3×10^8	9.50	Lead chloride	10^8	4.2
Apatite \parallel optic axis	3×10^8	7.41	Lead monoxide (15°C)	10^8	25.9
Asphalt	$<3 \times 10^4$	2.68	Lead nitrate	6×10^7	37.7
Barium chloride (anhyd.)	6×10^7	11.4	Lead oleate	4×10^8	3.27
Barium chloride (2H ₂ O)	6×10^7	9.4	Lead sulfate	10^4	14.3
Barium nitrate	6×10^7	5.9	Lead sulfide (15°C)	16^8	17.9
Barium sulfate (15°C)	10^8	11.40	Malachite (mean)	10^{12}	7.2
Beryl \perp optic axis	10^4	7.02	Mercuric chloride	10^8	3.2
Beryl \parallel optic axis	10^4	6.08	Mercurous chloride	10^8	9.4
Calcite \perp optic axis	10^4	8.5	Naphthalene	4×10^8	2.52
Calcite \parallel optic axis	10^4	8.0	Phenanthrene	4×10^8	2.80
Calcium carbonate	10^4	6.14	Phenol (10°C)	4×10^8	4.3
Calcium fluoride	10^4	7.36	Phosphorus, red	10^8	4.1
Calcium sulfate (2H ₂ O)	10^4	5.66	Phosphorus, yellow	10^8	3.6
Cassiterite \perp optic axis	10^{12}	23.4	Potassium aluminum sulfate	10^8	3.8
Cassiterite \parallel optic axis	10^{12}	24	Potassium carbonate (15°C)	10^8	5.6

(continued)

TABLE 8.9 (continued)
Dielectric Constants of Solids

Material	Freq. (Hz)	Dielectric Constant	Material	Freq. (Hz)	Dielectric Constant
<i>d</i> -Cocaine	5×10^8	3.10	Potassium chlorate	6×10^7	5.1
Cupric oleate	4×10^8	2.80	Potassium chloride	10^4	5.03
Cupric oxide (15°C)	10^8	18.1	Potassium chromate	6×10^7	7.3
Cupric sulfate (anhyd.)	6×10^7	10.3	Potassium iodide	6×10^7	5.6
Cupric sulfate (5H ₂ O)	6×10^7	7.8	Potassium nitrate	6×10^7	5.0
Diamond	10^8	5.5	Potassium sulfate	6×10^7	5.9
Quartz ⊥ optic axis	3×10^7	4.34	Sodium carbonate (10H ₂ O)	6×10^7	5.3
Quartz ∥ optic axis	3×10^7	4.27	Sodium chloride	10^4	6.12
Resorcinol	4×10^8	3.2	Sodium nitrate	—	5.2
Ruby ⊥ optic axis	10^4	13.27	Sodium oleate	4×10^8	2.75
Ruby ∥ optic axis	10^4	11.28	Sodium perchlorate	6×10^7	5.4
Rutile ⊥ optic axis	10^8	86	Sucrose (mean)	3×10^8	3.32
Rutile ∥ optic axis	10^8	170	Sulfur (mean)	—	4.0
Selenium	10^8	6.6	Thallium chloride	10^4	46.9
Silver bromide	10^4	12.2	<i>p</i> -Toluidine	4×10^8	3.0
Silver chloride	10^4	11.2	Tourmaline ⊥ optic axis	10^4	7.10
Silver cyanide	10^4	5.6	Tourmaline ∥ optic axis	10^4	6.3
Smithsonite ⊥ optic axis	10^{12}	9.3	Urea	4×10^8	3.5
Smithsonite ∥ optic axis	10^{10}	9.4	Zircon ⊥ ∥	10^8	12
Sodium carbonate (anhyd.)	6×10^7	8.4			

TABLE 8.10

Dielectric Constants of Ceramics

Material	Dielectric Constant 10 ⁴ Hz	Dielectric Strength V/mil	Volume Resistivity Ω cm (23°C)	Loss Factor
Alumina	4.5–8.4	40–160	10 ¹¹ –10 ¹⁴	0.0002–0.01
Cordierite	4.5–5.4	40–250	10 ¹² –10 ¹⁴	0.004–0.012
Forsterite	6.2	240	10 ¹⁴	0.0004
Porcelain (dry process)	6.0–8.0	40–240	10 ¹² –10 ¹⁴	0.0003–0.02
Porcelain (wet process)	6.0–7.0	90–400	10 ¹² –10 ¹⁴	0.006–0.01
Porcelain, zircon	7.1–10.5	250–400	10 ¹³ –10 ¹⁵	0.0002–0.008
Steatite	5.5–7.5	200–400	10 ¹³ –10 ¹⁵	0.0002–0.004
Titanates (Ba, Sr, Ca, Mg, and Pb)	15–12.000	50–300	10 ⁸ –10 ¹³	0.0001–0.02
Titanium dioxide	14–110	100–210	10 ¹³ –10 ¹⁸	0.0002–0.005

TABLE 8.11

Dielectric Constants of Glasses

Type	Dielectric Constant at 100 MHz (20°C)	Volume (350°C) MΩ cm	Loss Factor ^a
Corning 0010	6.32	10	0.015
Corning 0080	6.75	0.13	0.058
Corning 0120	6.65	100	0.012
Pyrex 1710	6.00	2,500	0.025
Pyrex 3320	4.71	—	0.019
Pyrex 7040	4.65	80	0.013
Pyrex 7050	4.77	16	0.017
Pyrex 7052	5.07	25	0.019
Pyrex 7060	4.70	13	0.018
Pyrex 7070	4.00	1,300	0.0048
Vycor 7230	3.83	—	0.0061
Pyrex 7720	4.50	16	0.014
Pyrex 7740	5.00	4	0.040
Pyrex 7750	4.28	50	0.011
Pyrex 7760	4.50	50	0.0081
Vycor 7900	3.9	130	0.0023
Vycor 7910	3.8	1,600	0.00091
Vycor 7911	3.8	4,000	0.00072
Corning 8870	9.5	5,000	0.0085
G.E. Clear (silica glass)	3.81	4,000–30,000	0.00038
Quartz (fused)	3.75 4.1 (1 MHz)	—	0.0002 (1 MHz)

^a Power factor × dielectric constant equals loss factor.

References

1. Halliday, D. and R. Resnick, *Fundamentals of Physics*, John Wiley & Sons, New York, 1970.
2. Plonus, M., *Applied Electromagnetics*, McGraw-Hill, New York, 1978.
3. Collin, R., *Foundations for Microwave Engineering*, McGraw-Hill, New York, 1966.
4. Smith, A., *Radio Frequency Principles and Applications*, IEEE Press, New York, 1998.
5. Cutnell, J. and K. Johnson, *Physics*, 3rd edn., John Wiley & Sons, New York, 1995.
6. James, A. and M. Lord, *Macmillan's Chemical and Physical Data*, Macmillan, London, U.K., 1992.
7. Kaye, G. and T. Laby, *Tables of Physical and Chemical Constants*, 15th edn., Longman, London, U.K., 1993.
8. Samsonov, G. (ed.), *Handbook of the Physicochemical Properties of the Elements*, IFI-Plenum, New York, 1968.
9. Lide, D. (ed.), *Chemical Rubber Company Handbook of Chemistry and Physics*, 79th edn., CRC Press, Boca Raton, FL, 1998.
10. Ellis, H. (ed.), *Nuffield Advanced Science Book of Data*, Longman, London, U.K., 1972.
11. Oberg, E., F. Jones, H. Horton, and H. Ryffell, *Machinery Handbook*, 24th edn., Industrial Press, New York, 2000.
12. Mitchell, M., Electrical resistivity of beryllium, *Journal of Applied Physics*, 46, 4742–4746, 1975.
13. Dean, J. (ed.), *Lange's Handbook of Chemistry*, 14th edn., McGraw-Hill, New York, 1992.
14. Seitz, F., *The Modern Theory of Solids*, Dover Publications, New York, 1987.
15. Jung, W., F. Schmidt, and G. Danielson, Thermal conductivity of high-purity vanadium, *Physical Review B (Solid State)*, 15(2), 659–665, 1977.
16. Kittel, D., *Introduction to Solid State Physics*, 3rd edn., John Wiley & Sons, New York, 1967.
17. Jonscher, A. K., *Dielectric Relaxation in Solids*, Chelsea Dielectric Press, London, U.K., 1983.

Bibliography

- Anikin, M. N., A. S. Zubrilov, A. A. Lebedev, A. M. Sterlchuck, and A. E. Cherenkov, *Fiuzika i Tekhnika Polurpovodnikov*, 24, 467, 1992.
- Barker, A. S. and M. Ilegems, Infrared lattice vibrations and free-electron dispersion in GaN, *Physical Review B*, 7, 743, 1973.
- Benjamin, M. C., C. Wang, R. F. Davis, and R. J. Nemanich, *Applied Physics Letters*, 64, 3288, 1994.
- Beyer, W. H. (ed.), *CRC Standard Mathematical Tables and Formulae*, 29th edn., CRC Press, Boca Raton, FL, 1991.

- Choyke, W. J., *The Physics and Chemistry of Carbides, Nitrides and Borides*, Kluwer, Dordrecht, the Netherlands, NATO ASI Series E, Applied Sciences, vol. 185, and references therein, 1990.
- Shackelford, J. F., *CRC Materials Science and Engineering Handbook*, 3rd edn., CRC Press, Boca Raton, FL, p. 304, 2000.
- Elliott, D., *Integrated Circuit Fabrication Technology*, McGraw-Hill, New York, 1982.
- Garland, C. W. and K. C. Parks, Low-temperature elastic constants of gallium arsenide, *Journal of Applied Physics*, 33, 759, 1962.
- Lide, D. R. (ed.), *CRC Handbook of Chemistry and Physics*, 76th edn., CRC Press, Boca Raton, FL, 1992.
- Lide, D. R., Properties of semiconductors, *CRC Handbook of Chemistry and Physics*, 1999–2000 edn., CRC Press, Boca Raton, FL, pp. 12–98, 2000.
- Maruska, H. P. and J. J. Tietjen, The preparation and properties of vapor-deposited single-crystalline GaN, *Applied Physics Letters*, 15, 327, 1969.
- Mohammad, S. N. and H. Morkoc, Progress and prospects of group III–V nitride semiconductors, *Progress in Quantum Electronics*, 20, 361, 1996.
- Pankow, J. F., *Aquatic Chemistry Concepts*, Lewis Publishers, Chelsea, MI, 1991.
- Schaffer, P. T. B. and R. G. Naum, Refractive index and dispersion of beta silicon carbide, *Journal of Optical Society of America*, 59, 1498, 1969.
- Schroder, D., *Semiconductor Material and Device Characterization*, John Wiley & Sons, New York, 1990.
- Shackelford, J. and W. Alexander (eds.), *CRC Materials Science and Engineering Handbook*, CRC Press, Boca Raton, FL, 1992.
- Suzuki, A., A. Ogura, K. Furukawa, Y. Fuji, M. Shigeta, and S. Nakajima, *Journal of Applied Physics*, 64, 2818, 1988.
- Tallarida, R. J., *Pocket Book of Integrals and Mathematical Formulas*, 2nd edn., CRC Press, Boca Raton, FL, 1991.
- Wessels, B. W. and J. C. Gatos, *Journal of Physics Chemical Solids*, 15, 83, 1977.
- Yoder, M. N., *Wide Bandgap Semiconductor Materials and Devices*, 43(10), 1634, 1966.

9

Cooling Considerations

9.1 Introduction

Adequate cooling of the tube envelope and seals is one of the principal factors affecting vacuum tube life [1]. Deteriorating effects increase directly with the temperature of the tube envelope and ceramic/metal seals. Inadequate cooling is almost certain to invite premature failure of the device.

Tubes operated at VHF and above are inherently subjected to greater heating action than tubes operated at lower frequencies. This results directly from the following:

- The flow of larger RF charging currents into the tube capacitances at higher frequencies
- Increased dielectric losses at higher frequencies
- The tendency of electrons to bombard parts of the tube structure other than the normal elements as the operating frequency is increased

Greater cooling, therefore, is required at higher frequencies. The technical data sheet for a given tube type specifies the maximum allowable operating temperature. For forced-air- and water-cooled tubes, the recommended amount of air or water is also specified in the data sheet. Both the temperature and quantity of coolant should be measured to be certain that cooling is adequate.

9.1.1 Thermal Properties

In the commonly used model for materials, heat is a form of energy associated with the position and motion of the molecules, atoms, and ions of the material [2]. The position is analogous with the state of the material and is *potential energy*, while the motion of the molecules, atoms, and ions is *kinetic energy*. Heat added to a material makes it hotter, and heat withdrawn from a material makes it cooler. Heat energy is measured in *calories* (cal), *British Thermal Units* (Btu), or *joules*. A calorie is the amount of energy required to raise the temperature of one gram of water by 1°C (14.5°C–15.5°C). A Btu is the unit of energy necessary to raise the temperature of one pound of water

by 1°F. A joule is an equivalent amount of energy equal to the work done when a force of one newton acts through a distance of one meter.

Temperature is a measure of the average kinetic energy of a substance. It can also be considered a relative measure of the difference of the heat content between bodies.

Heat capacity is defined as the amount of heat energy required to raise the temperature of one mole or atom of a material by 1°C without changing the state of the material. Thus, it is the ratio of the change in heat energy of a unit mass of a substance to its change in temperature. The heat capacity, often referred to as *thermal capacity*, is a characteristic of a material and is measured in cal/g/°C or Btu/lb/°F.

Specific heat is the ratio of the heat capacity of a material to the heat capacity of a reference material, usually water. Because the heat capacity of water is 1 Btu/lb and 1 cal/g, the specific heat is numerically equal to the heat capacity.

Heat transfers through a material by conduction resulting when the energy of atomic and molecular vibrations is passed to atoms and molecules with lower energy. As heat is added to a substance, the kinetic energy of the lattice atoms and molecules increases. This, in turn, causes an expansion of the material that is proportional to the temperature change, over normal temperature ranges. If a material is restrained from expanding or contracting during heating and cooling, internal stress is established in the material.

9.1.2 Heat Transfer Mechanisms

The process of heat transfer from one point or medium to another is a result of temperature differences between the two. Thermal energy may be transferred by any of three basic modes:

- Conduction
- Convection
- Radiation

A related mode is the convection process associated with the change of phase of a fluid, such as condensation or boiling. A vacuum tube amplifier or oscillator may use several of these mechanisms to maintain the operating temperature of the device within specified limits.

9.1.2.1 Conduction

Heat transfer by conduction in solid materials occurs whenever a hotter region with more rapidly vibrating molecules transfers a portion of its energy to a cooler region with less rapidly vibrating molecules. Conductive heat transfer is the most common form of thermal exchange in electronic equipment. Thermal conductivity for solid materials used in electronic equipment spans a wide range of values, from excellent (high conductivity) to poor

TABLE 9.1
Thermal Conductivity of Common Materials

Material	Btu/(h ft °F)	W/(m °C)
Silver	242	419
Copper	228	395
Gold	172	298
Beryllia	140	242
Phosphor bronze	30	52
Glass (borosilicate)	0.67	1.67
Mylar	0.11	0.19
Air	0.015	0.026

(low conductivity). Generally speaking, metals are the best conductors of heat, whereas insulators are the poorest. Table 9.1 lists the thermal conductivity of materials commonly used in the construction (and environment) of power vacuum tubes. Table 9.2 compares the thermal conductivity of various substances as a percentage of the thermal conductivity of copper.

With regard to vacuum tube devices, conduction is just one of the cooling mechanisms involved in maintaining element temperatures within specified limits. At high power levels, there is always a secondary mechanism that works to cool the device. For example, heat generated within the tube elements is carried to the cooling surfaces of the device by conduction. Convection or radiation then is used to remove heat from the tube itself.

Thermal conduction has been used successfully to cool vacuum tubes through direct connection of the anode to an external heat sink, but only at low power levels [1].

TABLE 9.2
Relative Thermal Conductivity of Various Materials as a Percentage of the Thermal Conductivity of Copper

Material	Relative Conductivity
Silver	105
Copper	100
Berlox high-purity BeO	62
Aluminum	55
Beryllium	39
Molybdenum	39
Steel	9.1
High-purity alumina	7.7
Steatite	0.9
Mica	0.18
Phenolics, epoxies	0.13
Fluorocarbons	0.05

9.1.2.2 Convection

Heat transfer by natural convection occurs as a result of a change in the density of a fluid (including air), which causes fluid motion. Convective heat transfer between a heated surface and the surrounding fluid is always accompanied by a mixing of fluid adjacent to the surface. Power vacuum tube devices relying on convective cooling invariably utilize forced air or water passing through a heat-transfer element, typically the anode in a grid tube or the collector in a microwave tube [1]. This *forced convection* provides for a convenient and relatively simple cooling system. In such an arrangement, the temperature gradient is confined to a thin layer of fluid adjacent to the surface so that the heat flows through a relatively thin *boundary layer*. In the main stream outside this layer, isothermal conditions exist.

9.1.2.3 Radiation

Cooling by radiation is a function of the transfer of energy by electromagnetic wave propagation. The wavelengths between 0.1 and 100m are referred to as *thermal radiation wavelengths*. The ability of a body to radiate thermal energy at any particular wavelength is a function of the body temperature and the characteristics of the surface of the radiating material. Figure 9.1 charts the

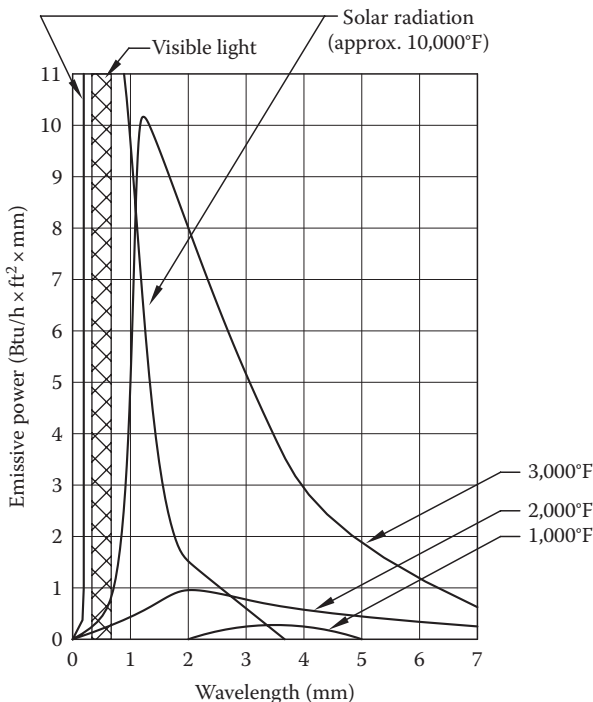


FIGURE 9.1

Blackbody energy distribution characteristics.

TABLE 9.3

Emissivity Characteristics of Common Materials at 80°F

Surface Type	Finish	Emissivity
Metal	Copper (polished)	0.018
Metal	Nickel	0.21
Metal	Silver	0.10
Metal	Gold	0.04–0.23
Glass	Smooth	0.9–0.95
Ceramic	Cermet ^a	0.58

^a Ceramic containing sintered metal.

ability to radiate energy for an ideal radiator, a *blackbody*, which, by definition, radiates the maximum amount of energy at any wavelength. Materials that act as perfect radiators are rare. Most materials radiate energy at a fraction of the maximum possible value. The ratio of the energy radiated by a given material to that emitted by a blackbody at the same temperature is termed *emissivity*. Table 9.3 lists the emissivity of various common materials.

9.1.3 Physics of Boiling Water

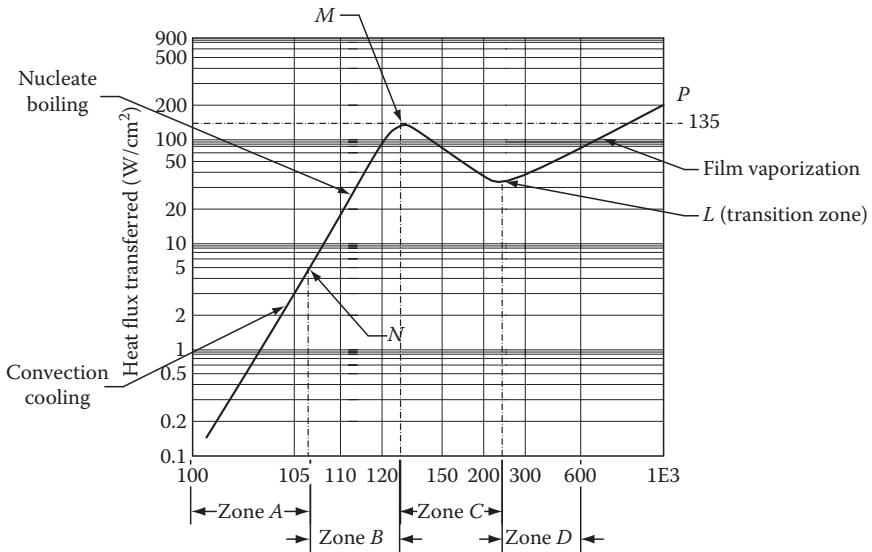
The *Nukiyama curve* shown in Figure 9.2 charts the heat-transfer capability (measured in W/cm^2) of a heated surface, submerged in water at various temperatures [1]. The first portion of the curve—*zone A*—indicates that from 100°C to about 108°C, heat transfer is a linear function of the temperature differential between the hot surface and the water, reaching a maximum of about $5 W/cm^2$ at 108°C. This linear area is known as the *convection cooling zone*. Boiling takes place in the heated water at some point away from the surface.

From 108°C to 125°C—*zone B*—heat transfer increases as the fourth power of ΔT until, at 125°C, it reaches $135 W/cm^2$. This zone is characterized by *nucleate boiling*; that is, individual bubbles of vapor are formed at the hot surface, break away, and travel upward through the water to the atmosphere.

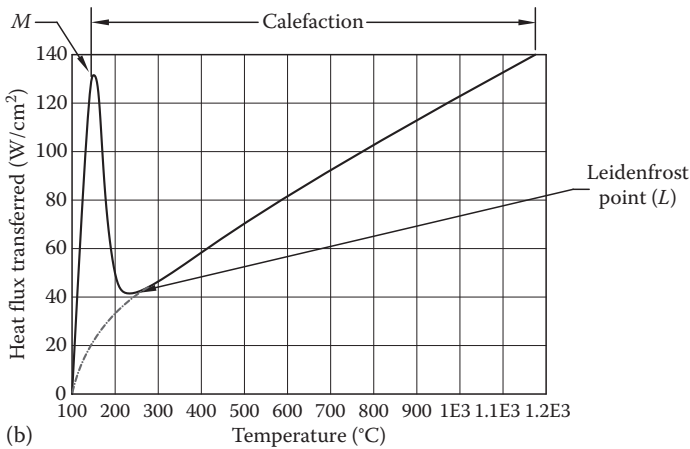
Above 125°C, an unstable portion of the Nukiyama curve is observed, where increasing the temperature of the heated surface actually reduces the unit thermal conductivity. At this area—*zone C*—the vapor partially insulates the heated surface from the water until a temperature of approximately 225°C is reached on the hot surface. At this point—called the *Leidenfrost point*—the surface becomes completely covered with a sheath of vapor, and all heat transfer is accomplished through this vapor cover. Thermal conductivity of only $30 W/cm^2$ is realized at this region.

From the Leidenfrost point on through *zone D*—the *film vaporization zone*—heat transfer increases with temperature until at about 1000°C the value of $135 W/cm^2$ again is reached.

The linear plot of the Nukiyama curve indicates that zones A and B are relatively narrow areas and that a heated surface with unlimited heat capacity



(a)



(b)

FIGURE 9.2

Nukiyama heat-transfer curves: (a) logarithmic and (b) linear.

will tend to pass from zone A to zone D in a short time. This irreversible superheating is known as *calefaction*. For a cylindrical vacuum tube anode, the passing into total calefaction would not be tolerated, because any unit heat-transfer density above 135 W/cm² would result in temperatures above 1000°C, well above the safe limits of the tube.

These properties dictate many of the physical parameters of water- and vapor-cooled power vacuum tubes. For a given dissipation, the anode structure that comes in contact with the coolant must be sufficient to keep the tube operating within the desired cooling zone of the Nukiyama curve and, in any event, at less than 135 W/cm² dissipation on the anode structure.

9.2 Application of Cooling Principles

Excessive dissipation is perhaps the single greatest cause of catastrophic failure in a power tube [1]. Power amplifier (PA) tubes used in communications, industrial, and research applications can be cooled using one of three methods: forced-air, liquid, and vapor-phase cooling. Forced-air cooling is the simplest and most common method used at low to moderate power levels.

The critical points of almost every PA tube type are the metal-to-ceramic junctions or seals. At temperatures below 250°C these seals remain secure, but above that temperature, the bonding in the seal may begin to disintegrate. Warping of grid structures also may occur at temperatures above the maximum operating level of the device. The result of prolonged overheating is shortened tube life or catastrophic failure. Several precautions usually are taken to prevent damage to tube seals under normal operating conditions. Air directors or sections of tubing may be used to provide spot cooling at critical surface areas of the device. Airflow sensors typically prevent operation of the RF system in the event of a cooling system failure.

Temperature control is important for vacuum tube operation because the properties of many of the materials used to build a tube change with increasing temperature. In some applications, these changes are insignificant. In others, however, such changes can result in detrimental effects, leading to—in the worst case—catastrophic failure. Table 9.4 details the variation of electrical and thermal properties with temperature for various substances.

TABLE 9.4

Variation of Electrical and Thermal Properties of Common Insulators as a Function of Temperature

Parameters		20°C	120°C	260°C	400°C	538°C
Thermal conductivity ^a	99.5% BeO	140	120	65	50	40
	99.5% Al ₂ O ₃	20	17	12	7.5	6
	95.0% Al ₂ O ₃	13.5				
	Glass	0.3				
Power dissipation ^b	BeO	2.4	2.1	1.1	0.9	0.7
Electrical resistivity ^c	BeO	10 ¹⁶	10 ¹⁴	5 × 10 ¹²	10 ¹²	10 ¹¹
	Al ₂ O ₃	10 ¹⁴	10 ¹⁴	10 ¹²	10 ¹²	10 ¹¹
	Glass	10 ¹²	10 ¹⁰	10 ⁸	10 ⁶	
Dielectric constant ^d	BeO	6.57	6.64	6.75	6.90	7.05
	Al ₂ O ₃	9.4	9.5	9.6	9.7	9.8
Loss tangent ^d	BeO	0.00044	0.00040	0.00040	0.00049	0.00080

^a Heat transfer in Btu/ft²/h/°F.

^b Dissipation in W/cm/°C.

^c Resistivity in Ω cm.

^d At 8.5GHz.

9.2.1 Forced-Air Cooling Systems

Air cooling is the simplest and most common method of removing waste heat from a vacuum tube device [1]. The basic configuration is illustrated in Figure 9.3. The normal flow of cooling air is upward, making it consistent with the normal flow of convection currents. In all cases, the socket is an open structure or has adequate vent holes to allow cooling of the base end of the tube. Cooling air enters at the grid circuit compartment below the socket through a screened opening, passes through the socket to cool the base end of the tube, sweeps upward to cool the envelope, and enters the output circuit compartment.

The output compartment is provided with a mesh-covered opening that permits the air to vent out readily. High-power tubes typically include a chimney to direct airflow, as shown in Figure 9.4. For a forced-air design, a suitable fan or blower is used to pressurize the compartment below the tube. No holes should be provided for the passage of air from the lower to the upper compartment other than the passages through the socket and tube base. A certain amount of pressure must be built up to force the proper amount of air through the socket to cool the anode.

Attention must be given to airflow efficiency and turbulence in the design of a cooling system. Consider the case shown in Figure 9.5. Improper layout has resulted in inefficient movement of air because of circulating thermal currents. Anode cooling will be insufficient in this case.

The cooling arrangements illustrated in Figures 9.3 and 9.4 provide for the uniform passage of cooling air over the tube base and anode. An arrangement that forces cooling air transversely across the tube base

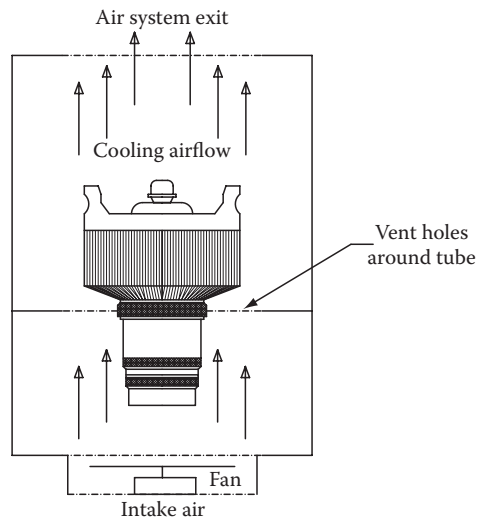


FIGURE 9.3
Cooling system design for a power grid tube.

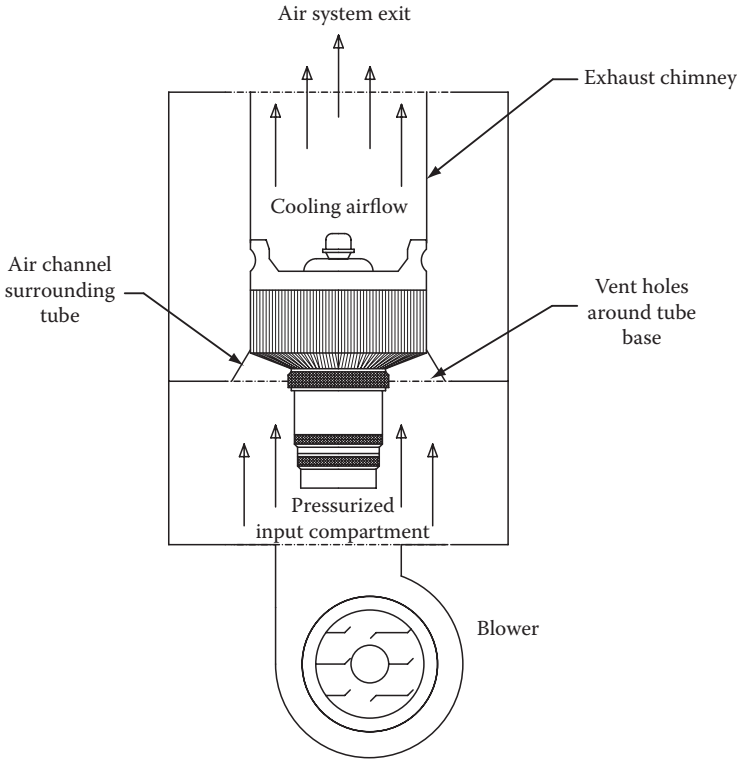


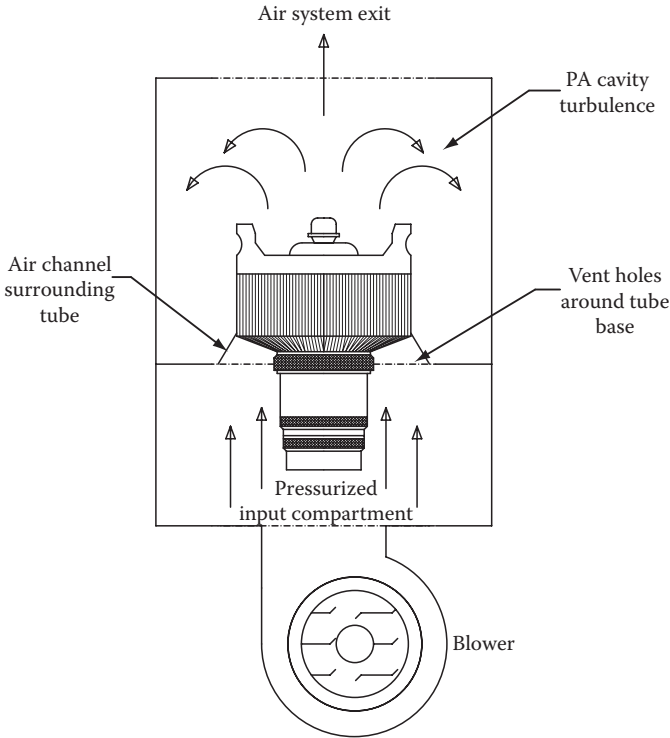
FIGURE 9.4
Use of a chimney to improve cooling of a power grid tube.

and/or anode will not provide the same cooling effectiveness and may result in hot spots on the device.

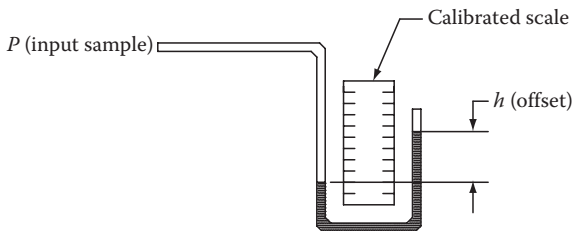
In many cases, packaged air system sockets and chimneys are designed specifically for a tube or family of tube types. The technical data sheet specifies the recommended socketing for adequate cooling. The tube data sheet also will specify the back pressure, in inches of water, and the cubic feet per minute (cfm) required for adequate cooling. In an actual application, the back pressure may be measured by means of a simple manometer. This device consists of a U-shaped glass tube partially filled with water, as shown in Figure 9.6.

9.2.1.1 Cooling Airflow Data

Power tube manufacturers typically outline minimum cooling airflow requirements for large external-anode tubes in the form of one or more charts [1]. Figure 9.7 plots $P_t/\Delta T$ as a function of A_m , where P_t is the total

**FIGURE 9.5**

Inefficient cooling system in which circulating air in the output compartment reduces the effectiveness of the heat-removal system.

**FIGURE 9.6**

Manometer used to measure air pressure. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

power dissipated in W , ΔT is the tube temperature rise in $^{\circ}\text{C}$, and A_m is the mass airflow rate in lb/min . This type of graph is used in calculating the cooling requirements of a vacuum tube device. The graph applies to a specified tube and socket–chimney combination; furthermore, the direction of airflow is specified. When reverse airflow (from anode to base) is used, the cooling requirements are sharply increased because the air applied to the

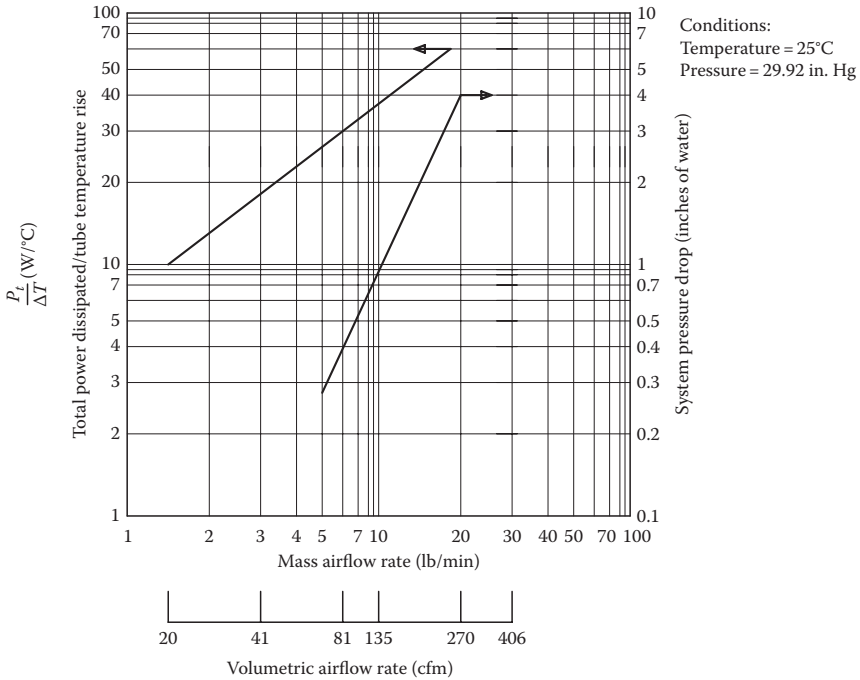


FIGURE 9.7 Cooling airflow requirements for a power grid tube. (Data courtesy of Varian, Palo Alto, CA.)

base seals already will have been heated by its passage through the anode cooler, losing much of its cooling effectiveness.

To use the type of graph shown in Figure 9.7, first determine the minimum cooling requirements using the following steps:

1. Calculate the total power dissipated (P_t) by adding all of the power dissipated by the tube during operation in a given installation. This value includes plate and filament dissipations, plus maximum anticipated grid and screen dissipations (as applicable).
2. Calculate the maximum tube temperature rise (ΔT) by taking the difference between the maximum rated tube temperature specified in the appropriate data sheet and the maximum expected air-inlet temperature.
3. To convert the mass airflow rate M (lb/min) to volumetric airflow rate Q (cfm) at 25°C and at sea level, divide the mass airflow rate by the density of air at 25°C and 29.9 in mercury.
4. The curve on the right side of the graph is the pressure drop (ΔP) in inches of water across the tube and its specified socket–chimney combination. This value is valid at 25°C at sea level only.

To adjust the 25°C sea-level laboratory test conditions to any other atmospheric (socket-inlet) condition, multiply both the Q and ΔP values by the ratio of the laboratory standard density (0.074 lb/ft³; 25°C at sea level) to the density at the new socket-inlet condition.

A shorter method may be used to approximately correct the 25°C sea-level requirements to both a different temperature and/or barometric socket-inlet condition. These corrections are made by multiplying the Q and ΔP values by the appropriate correction factors listed in Table 9.5.

Figure 9.8 is a graph of the combined correction factors that can be applied to the 25°C sea-level information for land-based installations located at elevations up to 10,000 ft, and for socket-inlet air temperatures between 10°C and 50°C. Figure 9.9 provides a method to convert the mass airflow rate in pounds per minute into volumetric airflow rate (cfm) at 25°C and sea level.

TABLE 9.5Correction Factors for Q and ΔP

Socket-Inlet Air Temperature (°C)	Q and ΔP Correction Factor	
0	0.917	
5	0.933	
10	0.950	
15	0.967	
20	0.983	
25	1.000	
30	1.017	
35	1.034	
40	1.051	
45	1.067	
50	1.084	
Socket-inlet Air Pressure ^a	Altitude (ft)	Q and ΔP Correction Factor
29.92	0	1.00
24.90	5,000	1.20
20.58	10,000	1.46
16.89	15,000	1.77
13.75	20,000	2.17
11.10	25,000	2.69
8.89	30,000	3.37
7.04	35,000	4.25

Source: After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.

^a Pressure in inches of mercury.

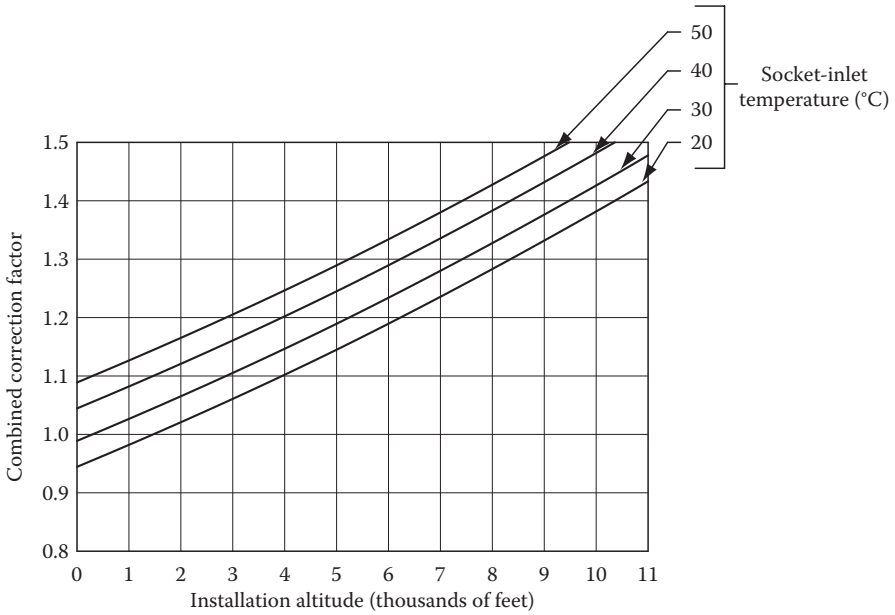


FIGURE 9.8 Combined correction factors for land-based tube applications. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

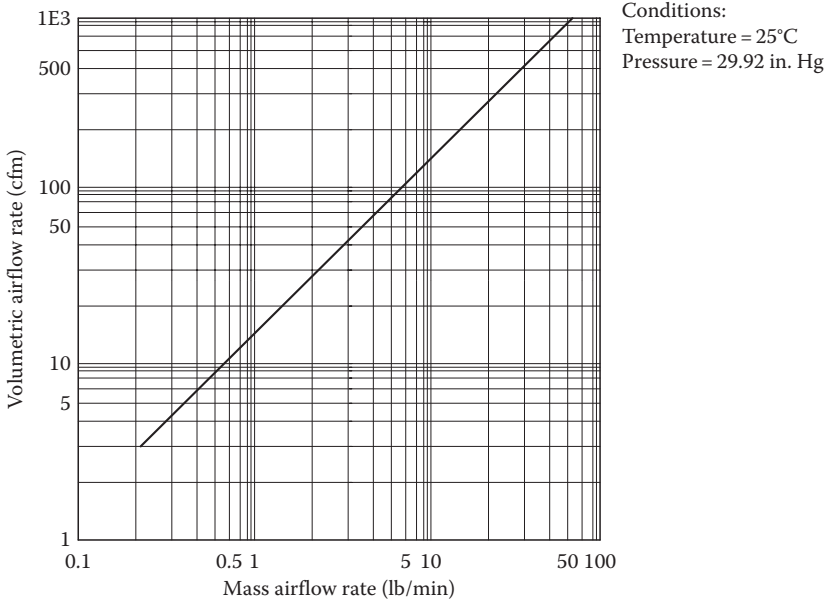


FIGURE 9.9 Conversion of mass airflow rate to volumetric airflow rate. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

Good engineering practice must be used when applying altitude and temperature corrections to the 25°C sea-level cooling requirement for airborne applications. Although the air outside the aircraft may be very cold at high altitudes, the air actually entering the tube socket may be many degrees warmer. This inlet temperature (and pressure) is affected by the installation design (compressed, ram, static, or recirculating air in a pressurized heat exchanger).

9.2.1.2 Blower Selection

In the previous section, a method of determining minimum air cooling requirements for external-anode tubes was described, pertaining to any altitude and air temperature [1]. Because most blower manufacturers furnish catalog data on their products in the form of volumetric airflow (Q , cfm) versus operating back pressure (ΔP , inches of water) for sea-level conditions only, the information gained from the foregoing procedure cannot be compared directly with data furnished by manufacturers for the purpose of selecting the proper device. The following method is recommended for use in selecting a blower for applications above sea level from existing blower catalog data:

1. Determine the Q and ΔP requirements for the tube socket–chimney combination for an ambient air temperature of 25°C at sea level. Include the estimated ΔP of ducting and filters in the system.
2. Determine the corrected Q and ΔP system requirements for the actual inlet temperature and altitude conditions by multiplying by the correction factor shown in Figure 9.7.
3. Multiply the ΔP —but not the Q —requirement by the correction factor cited in step 2.
4. Use the corrected Q factor and doubly corrected ΔP value to select a blower from the manufacturer's published sea-level curves. Although this blower will overcool the tube at sea level when operated in an ambient temperature of 25°C, it will provide adequate cooling at the actual inlet temperature and at high-altitude conditions.

9.2.2 Water Cooling

A water-cooled tube depends upon an adequate flow of fluid to remove heat from the device and transport it to an external heat sink [1]. The recommended flow as specified in the technical data sheet should be maintained at all times when the tube is in operation. Inadequate water flow at high temperature may cause the formation of steam bubbles at the anode surface where the water is in direct contact with the anode. This condition can contribute to premature tube failure.

Circulating water can remove about 1.0 kW/cm² of effective internal-anode area. In practice, the temperature of water leaving the tube is limited to 70°C

to preclude the possibility of spot boiling. The water then is passed through a heat exchanger where it is cooled to 30°C–40°C before being pumped over the tube anode again.

9.2.2.1 Cooling System Design

A liquid cooling system consists of the following principal components:

- A source of coolant
- Circulation pump
- Heat exchanger
- Coolant purification loop
- Various connection pipes, valves, and gauges
- Flow interlocking devices (required to ensure coolant flow anytime the equipment is energized)

Such a system is shown schematically in Figure 9.10. In most cases the liquid coolant will be water, but if there is a danger of freezing, it will be necessary to use an antifreeze solution such as ethylene glycol. In these cases, coolant flow must be increased or plate dissipation reduced to compensate for the poorer heat capacity of the ethylene glycol solution. A mixture of 60%

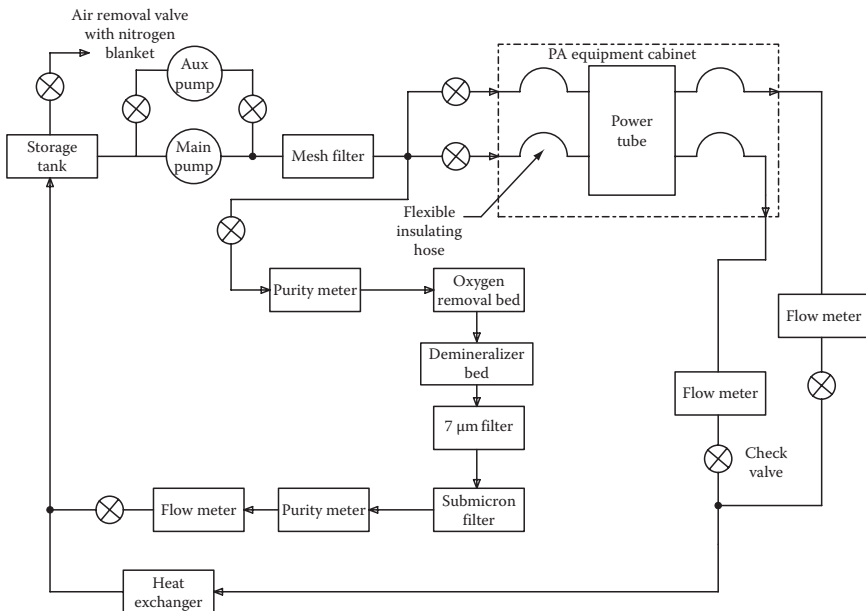


FIGURE 9.10

Functional schematic of a water cooling system for a high-power amplifier or oscillator.

ethylene glycol to 40% water by weight will be about 75% as efficient as pure water at 25°C. Regardless of the choice of liquid, the system volume must be maintained above the minimum required to ensure proper cooling of the vacuum tube(s).

The main circulation pump must be of sufficient size to ensure necessary flow and pressure as specified on the tube data sheet. Care must be taken when connecting the coolant lines to the tube to be certain that flow is in the direction specified. Improper direction of flow may result in inadequate cooling as well as excessive pressure and possible mechanical deformation of the anode. Precautions should be taken during system design to protect against water *hammering* of the anode. A filter screen of at least 60 mesh usually is installed in the pump outlet line to trap any circulating debris that might clog coolant passages within the tube.

The heat exchanger system is sized to maintain the outlet temperature such that the outlet water from the tube at full plate dissipation does not exceed 70°C. Filament or grid coolant courses may be connected in parallel or series with the main supply as long as the maximum outlet temperature is not exceeded.

Valves and pressure meters are installed on the inlet lines to the tube to permit adjustment of flow and measurement of pressure drop, respectively. A pressure meter and check valve are employed in the outlet line. In addition, a flow meter—sized in accordance with the tube data sheet—and a thermometer are included in the outlet line of each coolant course. These flow meters are equipped with automatic interlock switches, wired into the system electrical controls, so that the tube will be completely deenergized in the event of a loss of coolant flow in any one of the coolant passages. In some tubes, filament power alone is sufficient to damage the tube structure in the absence of proper water flow.

The lines connecting the plumbing system to the inlet and outlet ports of the tube are made of a flexible insulating material configured so as to avoid excessive strain on the tube flanges. Polypropylene tubing is the best choice for this service, but chlorinated polyvinyl chloride (CPVC) pipe, which is stronger, also is acceptable. Reinforced polypropylene, such as *Nylobraid*, is expensive, but excellent in this application. The coolant lines must be of sufficient length to keep electrical leakage below about 4 mA at full operating power.

The hoses are coiled or otherwise supported so that they do not contact each other or any conducting surface between the high-voltage end and ground. Conducting hose *barbs*, connected to ground, are provided at the low-potential end so that the insulating column of water is broken and grounded at the point where it exits the equipment cabinet.

All metallic components within the water system, including the pump, must be of copper, stainless steel, or unleaded brass or bronze; copper or stainless steel is preferred. If brass or bronze is used, some zinc may be leached out of the metal into the system. Any other material, such as iron or cold rolled steel, will grossly contaminate the water.

Even if the cooling system is constructed using the recommended materials, and is filled with distilled or deionized water, the solubility of the metals, carbon dioxide, and free oxygen in the water make the use of a coolant purification (*regeneration*) loop essential. The integration of the purification loop into the overall cooling system is shown in Figure 9.10. The regeneration loop typically taps 5%–10% of the total cooling system capacity, circulating it through oxygen scavenging and deionization beds, and submicron filters before returning it into the main system. Theoretically, such a purification loop can process water to 18 MΩ cm resistivity at 25°C. In practice, resistivity will be somewhat below this value.

Figure 9.11 shows a typical purification loop configuration. Packaged systems such as the one illustrated are available from a number of manufacturers. In general, such systems consist of replaceable cartridges that perform the filtering, ion exchange, and organic-solid-removal functions. The system usually will include flow and pressure gauges and valves, and conductivity cells for continuous evaluation of the condition of the water and filters.

9.2.2.2 Water Purity and Resistivity

The purity of the cooling water is an important operating parameter for any water-cooled amplifier or oscillator. The specific resistivity must be maintained at 1 MΩ cm minimum at 25°C. Distilled or deionized water should be used and the purity and flow protection periodically checked to ensure

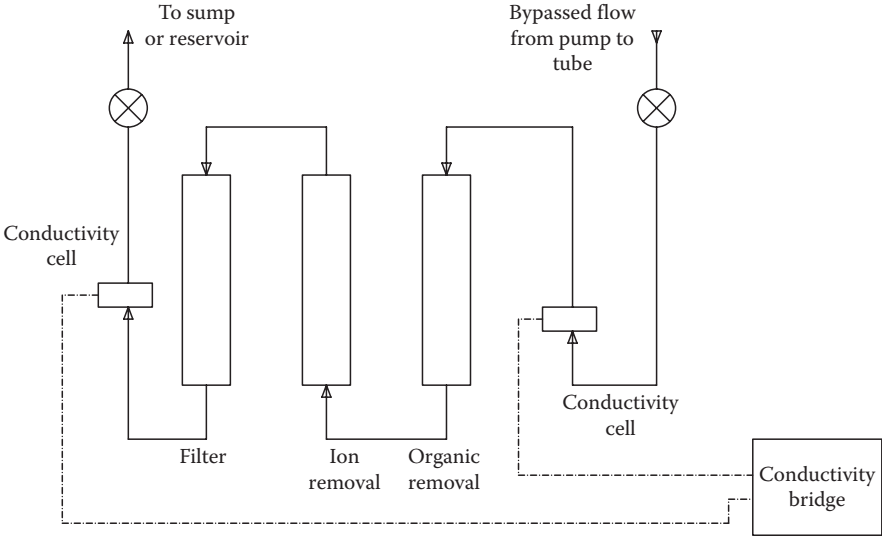


FIGURE 9.11
Typical configuration of a water purification loop.

against degradation. Oxygen and carbon dioxide in the coolant will form copper oxide, reducing cooling efficiency, and electrolysis may destroy the coolant passages. In addition, a filter screen should be installed in the tube inlet line to trap any circulating debris that might clog coolant passages within the tube.

After normal operation for an extended period, the cooling system should be capable of holding 3–4 M Ω cm until the filter beds become contaminated and must be replaced. Although the resistivity measurement is not a test for free oxygen in the coolant, the oxygen filter bed always should be replaced when the deionizing bed is replaced.

The resistivity of the coolant also is affected by the temperature of the water. The temperature dependence of the resistivity of pure water is charted in Figure 9.12.

It is recommended that the coolant water be circulated at all times. This procedure provides the following benefits:

- Maintains high resistivity
- Reduces bacteria growth
- Minimizes oxidation resulting from coolant stagnation

If it is undesirable to circulate the coolant at the regular rate when the tube is deenergized, a secondary circulating pump can be used to move the coolant at a lower rate to purge any air that might enter the system and to

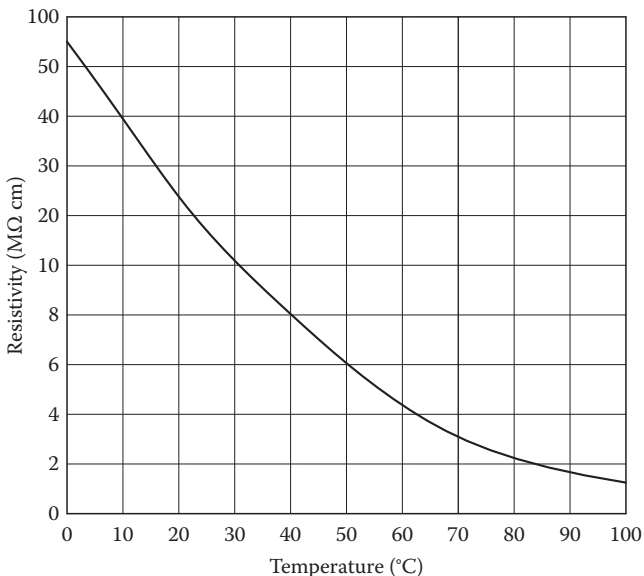


FIGURE 9.12

Effect of temperature on the resistivity of ultrapure water.

prevent stagnation. Approximate recommended minimum circulation rates within the coolant lines are as follows:

- 2 m/s during normal operation
- 30 cm/s during standby mode

It is recommended that a circulation rate of 10 m/s be established within the tube itself. When it is necessary to turn the coolant flow completely off, the cooling system should be restarted, and allowed to return to a minimum of 1 M Ω cm resistivity before energizing the tube.

The regeneration loop is typically capable of maintaining the cooling system such that the following maximum contaminant levels are not exceeded:

- Copper: 0.05 ppm by weight
- Oxygen: 0.5 ppm by weight
- CO₂: 0.5 ppm by weight
- Total solids: 3 ppm by weight

These parameters represent maximum levels; if the precautions outlined in this section are taken, actual levels will be considerably lower.

If the cooling system water temperature is allowed to reach 50°C, it will be necessary to use cartridges in the coolant regeneration loop that are designed to operate at elevated temperatures. Most ordinary cartridges will decompose in high-temperature service.

9.2.2.3 Condensation

The temperature of the input cooling water is an important consideration for proper operation of the cooling system [1]. If the air is humid and the cooling water is cold, condensation will accumulate on the surfaces of pipes, tube jackets, and other parts carrying water. This condensation may decrease the surface leakage resistance, or drops of water may fall on electric components, causing erratic operation or system failure. Some means is necessary, therefore, to control the temperature of the incoming water to keep it above the dew point. Such control is rather easy in a closed cooling system, but in a system that employs tap water and drains the exhaust water into a sewer, control is difficult.

9.2.2.4 Preventive Maintenance

Through electrolysis and scale formation, hard water may cause a gradual constriction of some parts of the water system [1]. Therefore, water flow and plumbing fittings must be inspected regularly. The fittings on the positive-potential end of an insulating section of hose or ceramic water coil or column are particularly subject to corrosion or electrolysis unless they have protective *targets*. The target elements should be checked periodically and replaced when they have disintegrated.

To extend the life of the resin beds in the purification loop, all the coolant lines should be flushed with a nonsudsing detergent and a citric acid solution,* then rinsed repeatedly with tap water before initially connecting the resin beds and filling with distilled or filtered deionized water. It is also good practice to sterilize the coolant lines with a chlorine solution† before filling to prevent algae and/or bacteria growth.

9.2.3 Vapor-Phase Cooling

Vapor-phase cooling offers several advantages over conventional water cooling systems by exploiting the latent heat of the evaporation of water [1]. Raising the temperature of 1 g of water from 40°C to 70°C (as in a water system) requires 30 cal of energy. Transforming 1 g of water at 100°C to steam vapor requires 540 cal. In a vapor cooling system, then, a given quantity of water will remove nearly 20 times as much energy as in a water cooling system. Power densities as high as 135 W/cm² of effective internal-anode surface can be attained through vapor cooling.

A typical vapor-phase installation consists of a tube with a specially designed anode immersed in a *boiler* filled with distilled water. When power is applied to the tube, anode dissipation heats the water to 100°C; further applied energy causes the water to boil and be converted into steam vapor. The vapor is passed through a condenser where it gives up its energy and is converted back into the liquid state. This condensate is then returned to the boiler, completing the cycle.

To achieve the most efficient heat transfer, the anode must be structured to provide for optimum contact with the water in the boiler. Figure 9.13 shows several examples.

Because the boiler usually is at a high potential relative to ground, it must be insulated from the rest of the system. The boiler typically is mounted

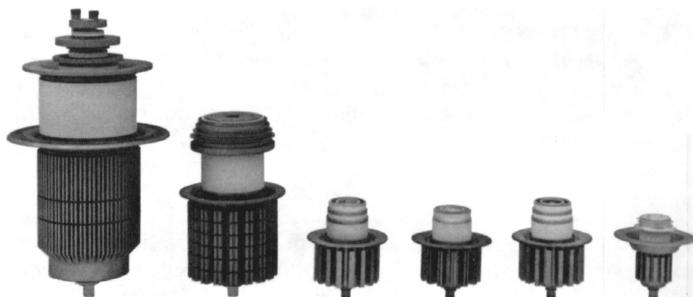


FIGURE 9.13

Vapor-phase-cooled tubes removed from their companion boilers. (Courtesy of Varian/Eimac, San Carlos, CA.)

* Citric acid solution mixtures vary; consult tube manufacturer.

† Chlorine solution: sodium hypochlorite bleach added in an amount sufficient to give the odor of chlorine in the circulating water.

on insulators, and the steam and water connections are made through insulated tubing. High-voltage standoff is more easily accomplished in a vapor-phase system than in a water-cooled system because of the following factors:

- There is a minimum of contamination because the water is constantly being redistilled.
- There is inherently higher resistance in the system because of the lower water-flow rate. The water inlet line is of a smaller diameter, resulting in greater resistance.

In a practical system, a 2 ft section of insulating tubing on the inlet and outlet ports is capable of 20 kV standoff.

Because of the effects of *thermosiphoning*, natural circulation of the water eliminates the need for a pump in most systems.

The dramatic increase in heat absorption that results from converting hot water to steam is repeated in reverse in the condenser. As a result, a condenser of much smaller thermal capacity is required for a vapor-phase system as opposed to a water-cooled system. For example, in a practical water-cooled system, water enters the heat exchanger at 70°C and exits at 40°C; the mean temperature is 55°C. For an ambient external temperature of 25°C, the mean differential between the water and the heat sink (air) is 30°C. The greater the differential, the more heat transferred. In a practical vapor-phase-cooled system, water enters the heat exchanger as steam at 100°C and exits as water at 100°C; the mean temperature is 100°C. For an ambient external temperature of 25°C, the mean differential between the vapor/water and the heat sink is 75°C. In this example, the vapor-phase condenser is nearly three times more efficient than its water-cooled counterpart.

Note that the condenser in a vapor-phase system may use either air or water as a heat sink. The water-cooled condenser provides for isolation of the PA tube cooling system from the outside world. This approach may be necessary because of high voltage or safety reasons. In such a system, the water is considered a *secondary coolant*.

Where air-cooled condensers are preferred, the higher thermal gradient can be exploited in reducing the size of condenser equipment and in lowering the blower horsepower requirement. In some instances where sufficient area is available, natural convection alone is used to cool the steam condensers, resulting in complete elimination of the condenser blower.

A typical vapor-phase cooling system is shown in Figure 9.14. It consists of the following elements:

- Power tube
- Boiler
- Condenser

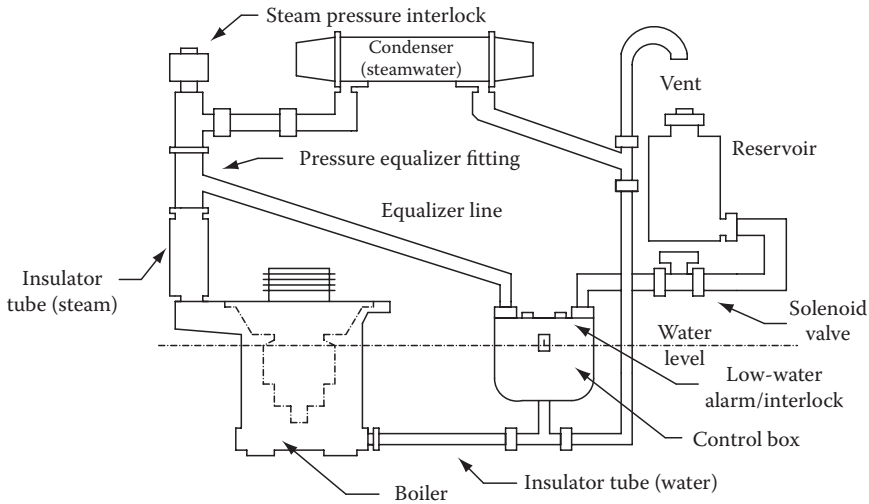


FIGURE 9.14

Vapor-phase cooling system for a power tube. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

- Insulated tubing
- Control box
- Reservoir
- Associated plumbing

9.2.3.1 Anode Design

The success of vapor-phase cooling is dependent on anode and boiler designs that allow a tube to operate at a temperature that results in maximum heat dissipation [1]. The most common approach has been to incorporate thick vertical fins on the exterior of the anode to achieve a radial temperature gradient on the surfaces submerged in water. In this way, a hot spot does not cause instantaneous runaway of the tube; only the average fin temperature increases, and this merely shifts the temperature range to a somewhat higher level. The temperatures of the fins typically vary from approximately 110°C at the tip to about 180°C at the root. With an average overall fin temperature of approximately 115°C, for example, the average transferred heat flux is on the order of 60 W/cm².

When operating at low dissipation levels, boiling takes place at the root of the fin. Increasing power density causes this boiling area to move out toward the end of the fin until, at rated dissipation levels, boiling takes place on the outer half of the fins. Good design dictates that the anode fin outer edge always remain at less than 125°C.

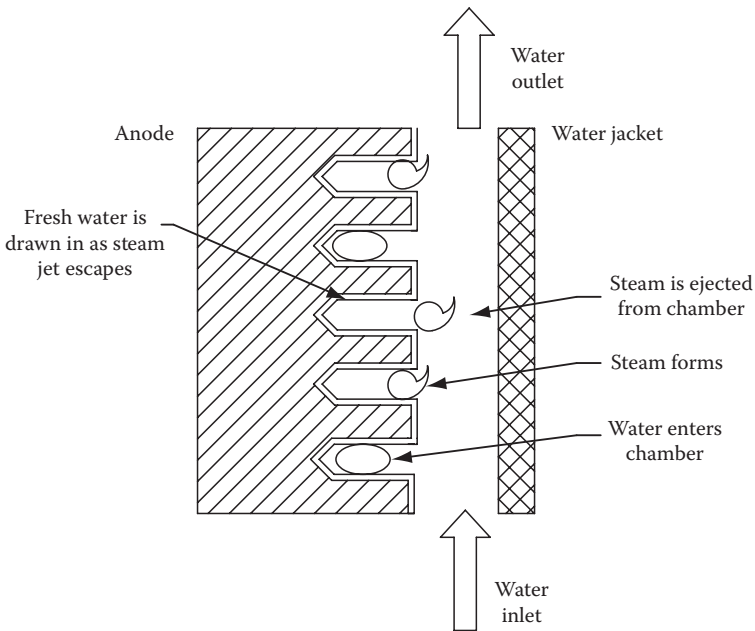


FIGURE 9.15

Cross section of a *holed anode* design for vapor-phase cooling. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

Anode shape is also important in assuring good cooling by breaking up the sheath of vapor that might tend to form at the surfaces. To a point, the more complicated the shape, the more efficient the anode cooler. Horizontal slots often are milled into the vertical fins to provide more area and to break up bubbles. The more vigorous the boiling action, the lower the possibility that an insulating vapor sheath will form. One design—known as the “pineapple”—carries this idea to the extreme, and incorporates square “knobs” all around the anode to provide dozens of heat-radiating surfaces. The more conventional finned anode, however, is generally adequate and less costly to fabricate.

The *holed anode* is another popular design for lower power densities (up to 100 W/cm^2). In this design, a heavy cylindrical anode is made with vertical holes in the outside wall. Vapor, formed by the boiling of water within the holes, is siphoned upward through the holes to the top of the boiler. This type of cooler always operates below the calefaction point and is used in smaller tube types. The concept is illustrated in Figure 9.15.

9.2.3.2 Boiler

The boiler supports the power tube and contains the water used for cooling [1]. In addition, it acts as the high-voltage anode connector. The boiler should

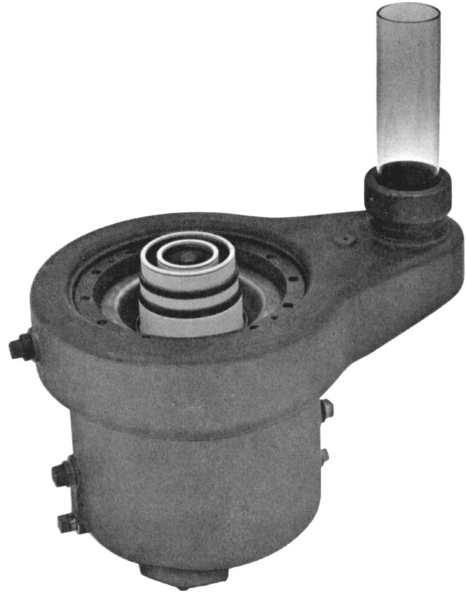


FIGURE 9.16

4CV35,000A tetrode mounted in its boiler assembly. (Courtesy of Varian, Palo Alto, CA.)

be mounted so that the axis of the tube is vertical. For effective cooling, tilt should be limited to less than 2° to ensure that the anode is covered with water and that the steam outlet is clear. Figure 9.16 shows a 4CV35,000A tetrode mounted in a boiler.

The anode flange of the tube must seal securely against the O-ring provided on the boiler. A locking flange presses the anode flange against the O-ring for a vapor-tight seal. The steam outlet at the top of the steam separation chamber on the boiler and the water inlet at the bottom of the boiler are equipped with fittings for attaching the Pyrex-insulated tubing. A *target* to inhibit electrolytic action is provided in the inlet water fitting.

In most cases the boiler is at a high potential relative to ground. Therefore, it must be electrically insulated from the rest of the system. The boiler is mounted on insulators, and the steam and water connections are made through insulating tubing. Boilers can be constructed with provisions for mounting multiple tubes in parallel. In such an arrangement, a single water inlet and steam outlet fitting typically would be used.

Common anode/boiler hardware is well suited to RF amplifiers operating up into the HF band, which typically use conventional tuned circuits. Operation in systems at higher frequencies, however, becomes more complicated than water cooling, and certainly more complicated than air cooling. The cooling system in an air-cooled RF generator does not interfere with the proper operation of the cavity; it is invisible to the cavity. Water cooling introduces some compromises and/or adjustments in the cavity, which can be readily overcome. Vapor-cooled tubes, however, require special design attention for successful use.

9.2.3.3 Insulating Tubing

The length of the steam and water insulating lines varies with individual installation requirements, but will always be shorter than would be needed in a circulating water system [1]. The length of the insulating tubing is determined by the following considerations:

- The voltage to be applied to the tube anode
- Purity of the water
- Volume of returned cooling water

9.2.3.4 Control Box

The function of the control box is to monitor and adjust the water level in the boiler [1]. A typical box is shown in Figure 9.17. The control box, an airtight vessel containing an overflow siphon and two float switches, also serves as a partial reservoir. When the water level drops approximately 1/4 in. below the recommended level, the first switch is closed. This signal may be used, at the same time, to activate a solenoid-controlled water valve to admit more distilled water from an external reservoir; it also may actuate a warning alarm.

The second float switch is operated if the water level drops approximately 1/2 in. below the optimum level. This would be tantamount to a water system failure; the switch would be used to open the control circuit interlocks and remove power from the tube.

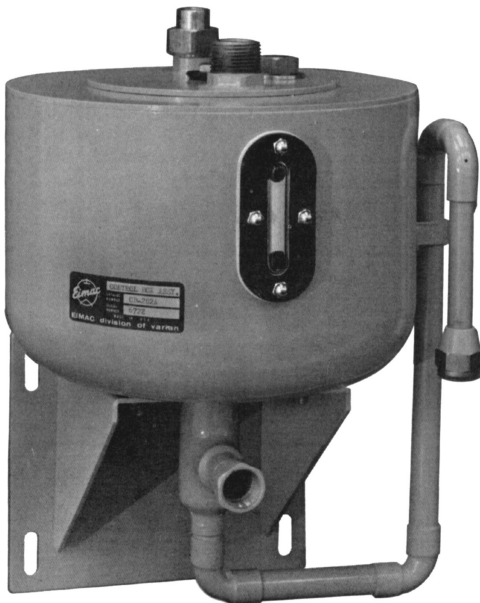


FIGURE 9.17
Control box for a vapor-phase cooling system. (Courtesy of Varian, Palo Alto, CA.)

For the control box to perform its protective function properly, the water-level mark must be precisely in line with the water-level mark on the boiler. For electrical reasons, the control box generally is mounted some distance from the boiler, and therefore leveling of the two components must be carefully checked during installation. Figure 9.18 shows a cutaway drawing of a classic boiler and tube combination. Figure 9.19 is a

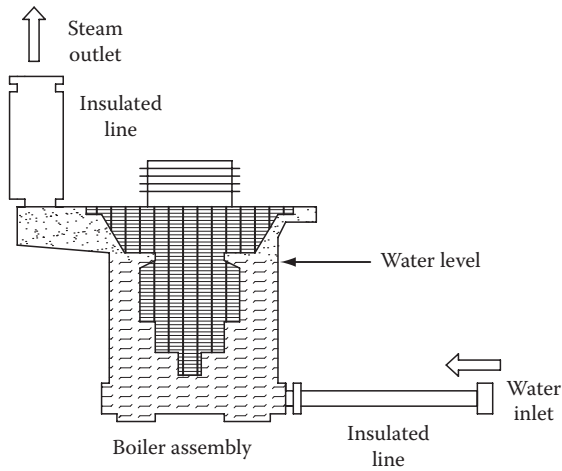


FIGURE 9.18

Cutaway view of a boiler and tube combination. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

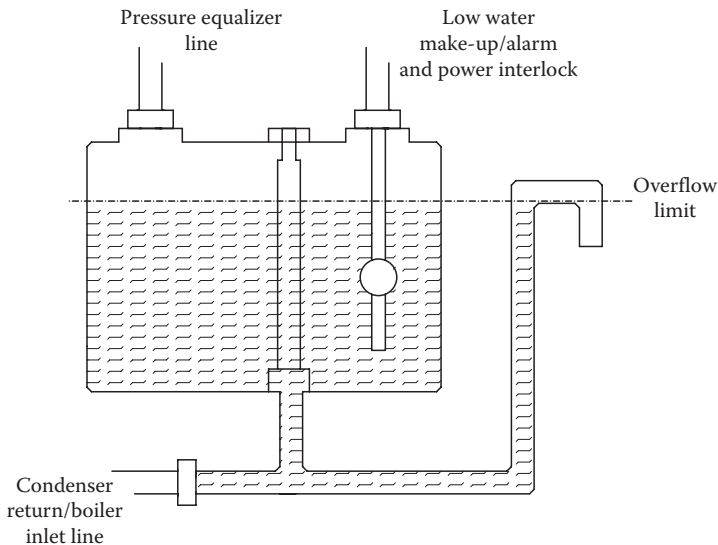


FIGURE 9.19

Cutaway view of a control box. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

cutaway drawing of a control box, showing the position of the float switches and the overflow pipe.

During extended operation, some quantity of water and steam being circulated through the condenser will be lost. The amount is dependent on the size of the system. The water level in the boiler, therefore, will gradually drop. The use of the control box as a reservoir minimizes this effect. In large or multiple-tube installations, an auxiliary reservoir is connected to the control box to increase the ratio of stored water to circulating water and steam. Where it may be necessary to operate multiple tubes at different physical elevations, individual control boxes are required. A multiple-tube system is shown in Figure 9.20.

9.2.3.5 Equalizer Line

For the control box to duplicate the same pressure conditions that exist in the boiler, the vapor-phase system must be fitted with an equalizer line [1]. This length of tubing connects the steam side of the system with the top of the control box. As partial steam pressure begins to build up in the boiler, the equalizer line allows this same pressure to appear in the control box. Although the steam pressure is low, less than 0.5 psi above atmosphere, an error in the operation of the control box would be introduced unless the vessel is equalized.

The fitting used to connect the equalizer line to the steam outlet tube must be constructed to prevent a venturi effect from developing because of the velocity of the vapor. This is best accomplished by directing an elbow within the adapter fitting toward the boiler, as shown in Figure 9.21.

9.2.3.6 Condenser

Both air-cooled and water-cooled condensers are available for vapor cooling systems [1]. Choose condensers with good reserve capabilities and low pressure drop. Air- and water-cooled condensers may be mounted in any position, provided they allow the condensed water to flow freely by gravity to the boiler return line. Water must not be allowed to stand in the condenser where it might cause back pressure to the entering steam.

The condenser should be mounted above the level of the boiler(s) so that water will drain from it to the boiler return line. Where it is necessary to mount the condenser at a lower physical level than the system water level, an auxiliary pump must be used to return water to the boiler. This arrangement is recommended for the "steam-out-the-bottom" boiler system, as illustrated in Figure 9.22.

9.2.3.7 Pressure Interlock

Most tube manufacturers recommend the use of a steam pressure interlock switch on the steam or inlet side of the condenser [1]. This switch, set to about

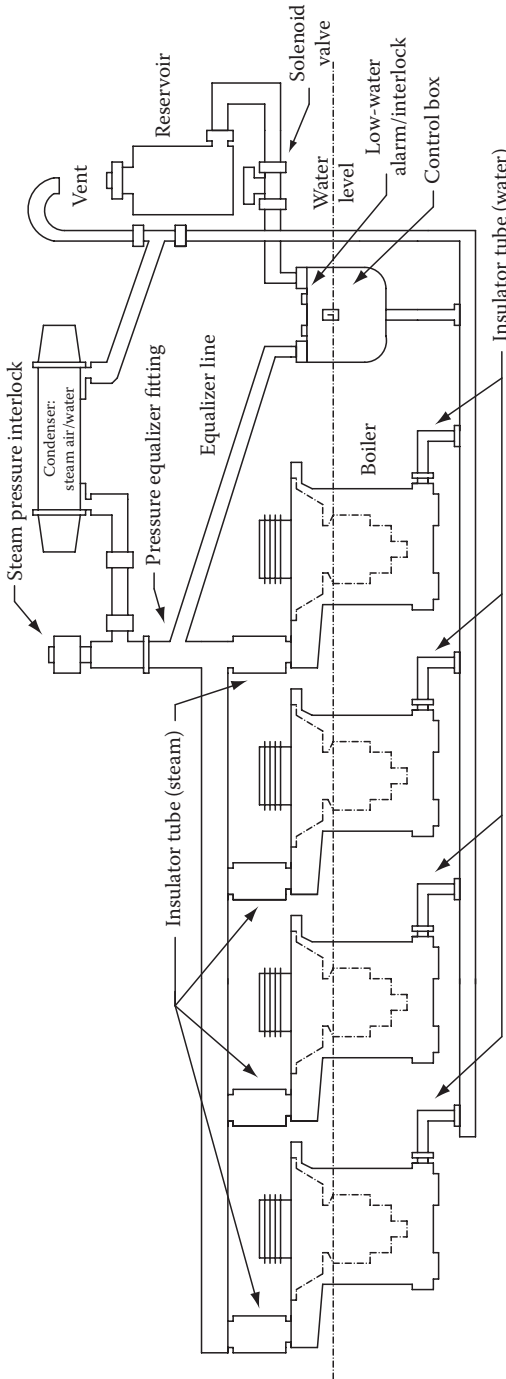
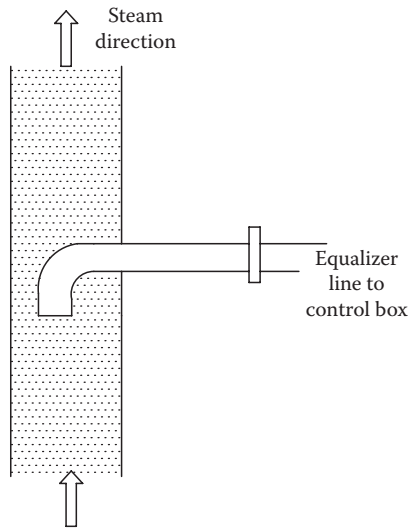
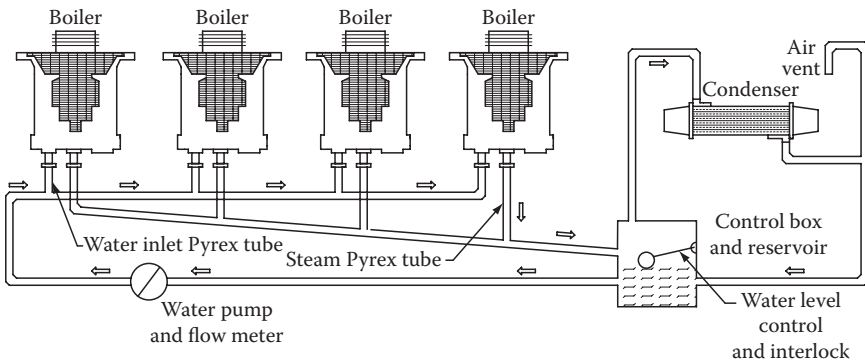


FIGURE 9.20 Typical four-tube vapor cooling system with a common water supply. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

**FIGURE 9.21**

Cutaway view of a pressure equalizer fitting. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

**FIGURE 9.22**

Typical four-tube vapor cooling system using "steam-out-the-bottom" boilers. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

0.5lb/in.², is used as a power interlock that senses abnormal steam pressure resulting from constrictions in the condenser or piping.

9.2.3.8 Piping

Piping should be of copper or glass throughout the system [1]. The steam piping should be the same diameter as the Pyrex tube from the boiler. The size is dependent on the power level and the volume of generated steam, and it will range from 1 to 1/4 in. at the 8kW level to 6 in. for the 250kW

level of dissipation. The steam path should be as direct as practical and must be sloped to prevent condensate from collecting at a low point where it might cause back pressure. All low spots should be drained back to the inlet water line.

The diameter of water return piping from the condenser to the control box will vary from 3/4 to 1-3/4 in., depending again on the power level. This tubing should be the same diameter as the boiler inlet water fitting. It also should be sloped so that water or vapor pockets are not allowed to form, and it must allow the condensate to return by gravity to the control box and the boiler. A vent to air on the outlet side of the condenser should be incorporated to maintain the water side of the system at atmospheric pressure. Provisions for draining the distilled water should be made at the lowest physical level of the system.

The equalizer line also should be sloped from the adapter fitting on the steam line to the top of the control box. This will allow the condensate to return to the control box.

9.2.3.9 Automatic Refilling System

Figure 9.23 shows a typical vapor cooling system with provisions to allow additional water into the control box [1]. An auxiliary reservoir

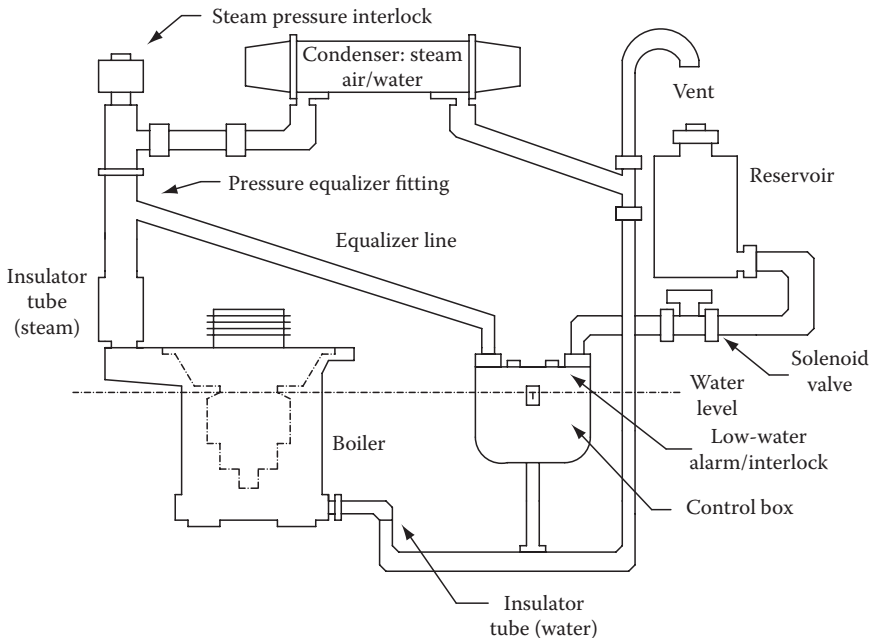


FIGURE 9.23

Vapor cooling system incorporating a reservoir of distilled water. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

is connected through a solenoid-operated water valve to the controller. When water loss resulting from evaporation causes the water level in the boiler and the control box to drop about $\frac{1}{4}$ in. below normal, the first float switch in the control box closes and actuates the solenoid-controlled valve to permit makeup water to enter the system. When the proper level is restored, the switch opens, the valve closes, and the flow of makeup water is stopped.

9.2.3.10 Alternative Vapor Cooling Systems

The systems described thus far are “classic” designs in which a separate tube, boiler, condenser, and level control box are used [1]. Variations on this basic scheme are numerous. One such alternative system, offered for use with large power tubes, utilizes a “steam-out-the-bottom” boiler. This configuration makes it possible to keep the steam and water systems, plus the plumbing, below the tubes. This configuration often simplifies the electrical design of the amplifier or oscillator. Figure 9.24 shows a typical boiler used in this cooling technique.

A small water pump circulates a continuous flow of water over a weir—in or baffle—in the boiler, maintaining a constant water level. Generated steam is forced under slight pressure out the bottom of the boiler, through an insulator tube in the condenser. Water from the condenser flows into the control box before being pumped back into the boiler. Protective devices include a water-flow interlock and water-level switch in the control box to ensure an adequate supply of coolant.

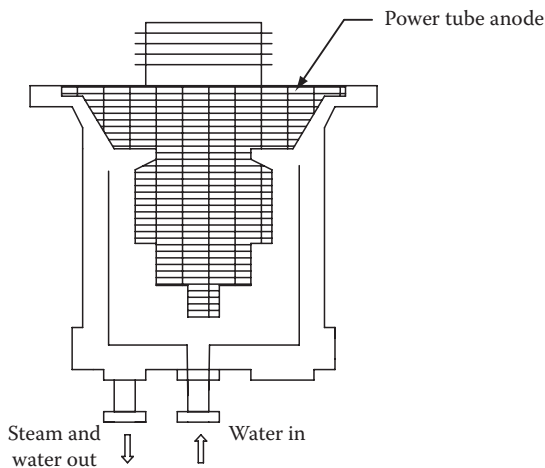


FIGURE 9.24

Cutaway view of a “steam-out-the-bottom” boiler. (After Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.)

9.2.3.11 Maintenance

Maintenance problems associated with circulating water systems are practically eliminated in a vapor cooling system [1]. As mentioned previously, systems can be designed to eliminate all rotating machinery or moving parts. Good engineering practice does, however, dictate periodic maintenance. The glass insulator tubes should be inspected occasionally to be sure they contain no deposits that might cause high-voltage flashover. Water conductivity can be checked by measuring the dc resistance, as in a typical circulating water system. The water should be replaced if the dc resistance drops below $1.0\text{M}\Omega\text{ cm}$.

In practice, a vapor-cooled system will remain cleaner longer than a water-cooled system. In the vapor-cooled boiler, the water is continuously being redistilled, and only pure water is introduced at the bottom of the boiler. Any contaminants will tend to remain in the boiler itself, where they can be removed easily. The periods between equipment shutdown for draining and cleaning will be at least twice as long for a vapor cooling system because of this inherent self-cleaning action.

Each time a tube is removed or replaced, the rubber O-ring between the boiler and the tube should be inspected and, if necessary, replaced. At the same time, the inside of the boiler and the control box can be inspected and cleaned as needed. The electrolytic *target* should be replaced whenever its metallic end is no longer visible in the inlet water line.

9.2.4 Temperature Measurements

Before completing any design employing power tubes, check temperatures in critical areas, such as metal/ceramic seals and (except for water- and vapor-cooled types) the anode just above and below the cooling fins [1]. Another critical area on many medium and large tubes of coaxial design is the central part of the base, which is often recessed and may need special cooling provisions. Electrical contact to the tube terminals is usually by copper-beryllium collets, which should not exceed a temperature of 150°C for any extended period of time. A number of manufacturers make temperature-sensitive paints that are useful for such measurements.* “Crayon-type” indicators are also available, in a wide range of temperatures. Conventional temperature measuring-techniques may be affected by the radio frequency energy concentrated at the tube and nearby components. Furthermore, because all power tubes operate at high voltage, there is danger of electric shock with directly connected devices.

When a temperature-sensitive paint or crayon is used for measurement, it must be remembered that many of the areas of concern are vacuum seals

* Tempilaq is one such product (Tempil Division, Big Three Industrial Gas and Equipment Co., South Plainfield, NJ).

involving relatively thin metal. It is important, therefore, that the indicator material be noncorrosive.

Considering the importance of tube element temperatures, every design must be evaluated carefully. All power tubes carry an absolute maximum temperature rating for the seals and envelope, and in the case of an external-anode forced-air-cooled tube, for the anode core itself. Operation above the rated maximum can cause early tube failure. Where long life and consistent performance are important factors, it is normally desirable to maintain tube temperatures comfortably below the rated maximum.

The equipment operator has the ultimate responsibility of making sure that the power tubes used in the system are operating within specified ratings. Most current tube designs use a ceramic envelope and an externally mounted anode or collector. The only visible sign of excessive temperature may be the cosmetic plating beginning to darken or turn black (oxidize). By then, significant damage may have occurred.

9.2.4.1 “Crayon” Temperature Measurement

When a crystalline solid is heated, a temperature will be reached at which the solid changes sharply to a liquid [1]. This melting point has a definite, reproducible value that is virtually unaffected by ambient conditions that may cause errors in other temperature-sensing methods. For example, fusible temperature indicators are practical in induction heating and in the presence of static electricity or ionized air around electric equipment, where electronic means of measuring temperatures often function erratically.

The most popular type of fusible indicator is a temperature-sensitive stick closely resembling a crayon, with a calibrated melting point. These crayon-type indicators are made in 100 or so different specified temperature ratings in the range of 100°F–2500°F. Typically, each crayon has a temperature-indicating accuracy within 1% of its rating. The workpiece to be tested is marked with the crayon. When the workpiece attains the predetermined melting point of the crayon, the mark changes from a solid to a liquid phase (liquefies), indicating that the workpiece has reached that temperature.

Under certain circumstances, premarking with crayon is not practical. This is the case if one or more of the following is true:

- A prolonged heating period is experienced.
- The surface is highly polished and does not readily accept a crayon mark.
- The material being marked is one that gradually absorbs the liquid phase of the crayon.

Note that a melted mark, upon cooling, will not solidify at the exact temperature at which it melted. Solidification of a melted crayon mark, therefore, cannot be relied on for exact temperature indication.

9.2.4.2 Phase-Change Fluid

A phase-changing fluid, or *fusible temperature-indicating lacquer*, offers the greatest flexibility in temperature measurement of a power tube [1]. The lacquer-type fluid contains a solid material of calibrated melting point suspended in an inert, volatile, nonflammable vehicle. As with crayon-type indicators, there are approximately 100 different temperature ratings, covering a range from 100°F to 2500°F. Accuracy is typically within 1% of the specified value.

The lacquer is supplied in the proper consistency for brushing. If spraying or dipping is desirable as the mode of application, a thinner is available to alter the viscosity without impairing the temperature-indicating accuracy.

Phase-changing lacquers often are used when a smooth or soft surface must be tested, or in situations where the surface is not readily accessible for application of a crayon mark during the heating process. Within seconds after application, the lacquer dries to a dull matte finish, and it responds rapidly when the temperature to be indicated is reached. The response delay of a lacquer mark is only a fraction of a second. This time can be reduced to milliseconds if a mark of minimal thickness is applied.

Upon reaching its rated temperature, the lacquer mark will liquefy. On subsequent cooling, however, the fluid will not revert to its original unmelted appearance but, rather, to a glossy or crystalline coating, which is evidence of its having reached the required temperature. Temperature-indicating lacquers, upon cooling, will not resolidify at the same temperature at which they melted.

If spraying of the phase-change lacquer is to be the sole means of application, the operator may find it more convenient to use an aerosol-packaged form of this material. An aerosol phase-change temperature indicator is identical to the brush-on material in performance and interpretation.

The first commercial form of the fusible indicator was the *pellet*, which continues to be useful in certain applications. Pellets are most frequently employed when extended heating periods are involved, or where a greater bulk of indicator material is necessary. They are also useful when observations must be made from a distance and when air-space temperatures are to be monitored.

Phase-change temperature-indicator pellets are available in flat, 7/16 in. diameter × 1/8 in. thick tablets. For special applications, miniature pellets, 1/8 in. × 1/8 in. are also available. The range of 100°F–2500°F covered by the crayon and lacquer type is, likewise, covered by the pellets. Pellets with coverage extending to 3200°F also may be obtained.

Another variation of the phase-change indicator is the temperature-sensitive label.* These self-adhesive-backed monitors consist of one or more heat-sensitive indicators sealed under transparent heat-resistant windows. The centers of the indicator circles turn from white to black at the temperature ratings shown on the label. The change to black is irreversible, representing an absorption of the temperature-sensitive substance into its backing material. After registering the temperature history of the workpiece, the exposed monitor label can be removed and affixed to a service report to remain part of a permanent record.

9.2.4.3 Selection Process

Fusible temperature indicators and lacquers have at least three major advantages over other methods of determining surface temperature [1]:

- The temperature indications obtained are unquestionably those of the surface being tested. It is not necessary to equilibrate a relatively massive probe with the surface (a probe may conduct heat away from the region being tested), which requires the use of a correction factor to obtain the actual surface temperature.
- There is no delay in obtaining an indication. Because a mark left by a crayon or a lacquer is of extremely small mass, it attains rapid equilibrium with the surface. There is no conduction of heat away from the surface, which would prolong response time and result in erroneously low temperature readings.
- The technique is simple and economical. Determination of surface temperature by most other means requires technical competence and skill and, in some cases, sophisticated instrumentation. Accurate surface temperature readings can be obtained with fusible indicators and lacquers with little effort, training, and expense.

There are numerous instances, particularly in determining heat distribution in complex systems, in which the relative simplicity of fusible indicators and lacquers makes surface temperature investigations feasible.

9.2.5 Air-Handling System

All modern air-cooled PA tubes use an air-system socket and matching chimney for cooling [1]. The chimney is designed to be an integral part of the tube cooling system. Operation without the chimney may significantly reduce airflow through the tube and result in excessive dissipation of the device. It also is possible that operation without the proper chimney could damage other components in the circuit because of excessive radiated heat. Normally, the tube socket is mounted in a pressurized compartment so that cooling air

* Temp-Plate is one such product (Williams-Wahl Corp., Santa Monica, CA).

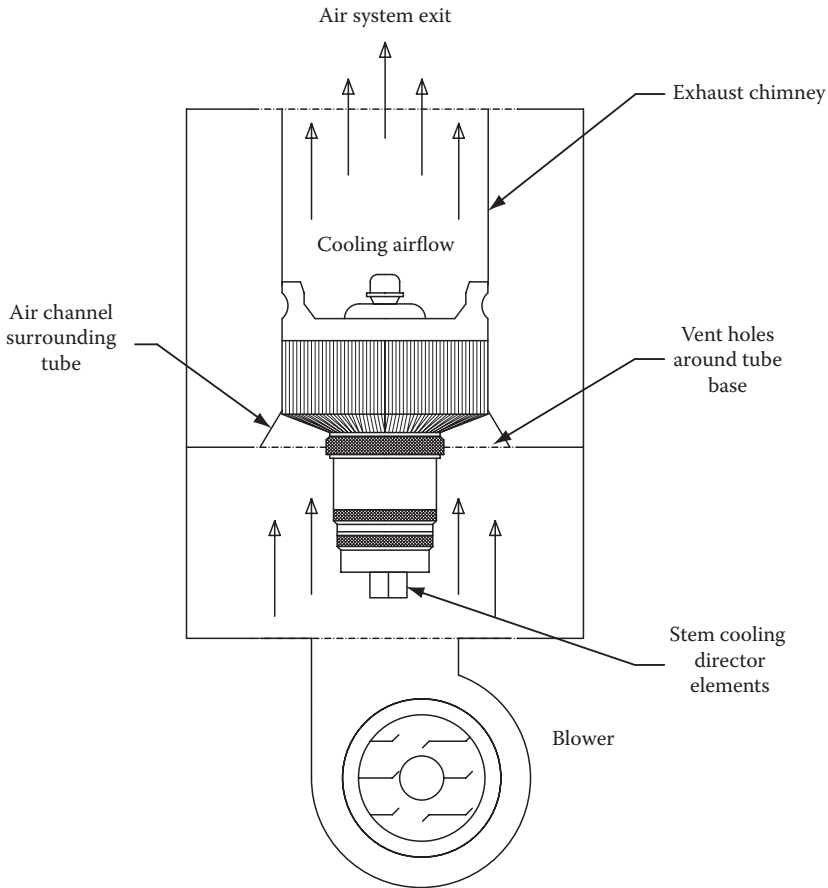


FIGURE 9.25
Airflow system for an air-cooled power tube.

passes through the socket and is guided to the anode cooling fins, as illustrated in Figure 9.25.

Cooling of the socket assembly is important for proper cooling of the tube base and for cooling the contact rings of the tube itself. The contact fingers used in the collet assembly of a socket typically are made of beryllium copper. If subjected to temperatures above 150°C for an extended period of time, the beryllium copper will lose its temper (springy characteristic) and will no longer make good contact with the base rings of the device. In extreme cases, this type of socket problem may lead to arcing, which can burn through the metal portion of the tube base ring. Such an occurrence ultimately may lead to catastrophic failure of the device because of a loss of the vacuum envelope. Other failure modes for a tube socket include arcing between the collet and tube ring, which can weld a part of the socket and tube together. The result is failure of both the tube and the socket.

9.3 Operating Environment

Long-term reliability of a power vacuum tube requires regular attention to the operating environment. Periodic tests and preventive maintenance are important components of this effort. Optimum performance of the cooling system can be achieved only when all elements of the system are functioning properly.

9.3.1 Air-Handling System

The temperature of the intake air supply is a parameter that is usually under the control of the end user. The preferred cooling air temperature is typically no higher than 75°F, and no lower than the room dew point. The air temperature should not vary because of an oversized air-conditioning system or because of the operation of other pieces of equipment at the facility. Monitoring the PA tube exhaust stack temperature is an effective method of evaluating overall RF system performance. This can be easily accomplished. It also provides valuable data on the cooling system and final stage tuning.

Another convenient method for checking the efficiency of the transmitter cooling system over a period of time involves documenting the back pressure that exists within the PA cavity. This measurement is made with a *manometer*, a simple device that is available from most heating, ventilation, and air-conditioning (HVAC) suppliers. The connection of a simplified manometer to a transmitter PA output compartment is illustrated in Figure 9.26.

By charting the manometer readings, it is possible to accurately measure the performance of the transmitter cooling system over time. Changes resulting from the buildup of small dust particles (*microdust*) may be too gradual to be detected except through back-pressure charting. Deviations from the typical back-pressure value, either higher or lower, could signal a problem with the air-handling system. Decreased PA input or output compartment back pressure could indicate a problem with the blower motor or an accumulation of dust and dirt on the blades of the blower assembly. Increased back pressure, on the other hand, could indicate dirty PA tube anode cooling fins (for the input compartment case) or a buildup of dirt on the PA exhaust ducting (for the output compartment case). Either condition is cause for concern. A system suffering from reduced air pressure into the PA compartment must be serviced as soon as possible. Failure to restore the cooling system to proper operation may lead to premature failure of the PA tube or other components in the input or output compartments. Cooling problems do not improve with time; they invariably get worse.

Failure of the PA compartment air-interlock switch to close reliably may be an early indication of impending trouble in the cooling system. This situation could be caused by normal mechanical wear or vibration of the switch assembly, or it may signal that the PA compartment air pressure has dropped.

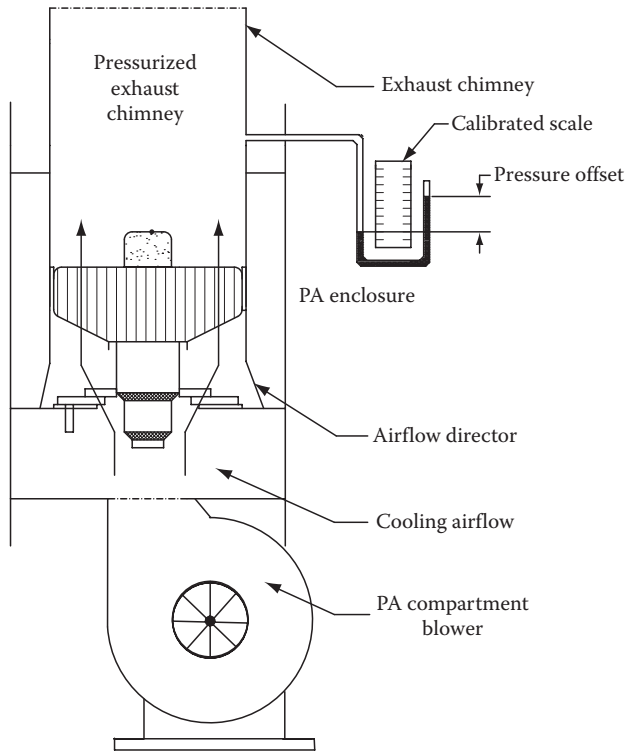


FIGURE 9.26

Manometer device used for measuring back pressure in the PA compartment of an RF generator.

In such a case, documentation of manometer readings will show whether the trouble is caused by a failure of the air pressure switch or a decrease in the output of the air-handling system.

9.3.2 Air Cooling System Design

Cooling system performance in an RF generator is not necessarily related to airflow volume. The cooling capability of air is a function of its mass, not its volume. The designer must determine an appropriate airflow rate within the equipment and establish the resulting resistance to air movement. A specified static pressure that should be present within the ducting of the transmitter can be a measure of airflow. For any given combination of ducting, filters, heat sinks, radio frequency interference (RFI) honeycomb shielding, tubes, tube sockets, and other elements in the transmitter, a specified system resistance to airflow can be determined. It is important to realize that any changes in the position or number of restricting elements within the system will change the system resistance and, therefore, the effectiveness of the cooling. The altitude of operation is also a consideration in cooling

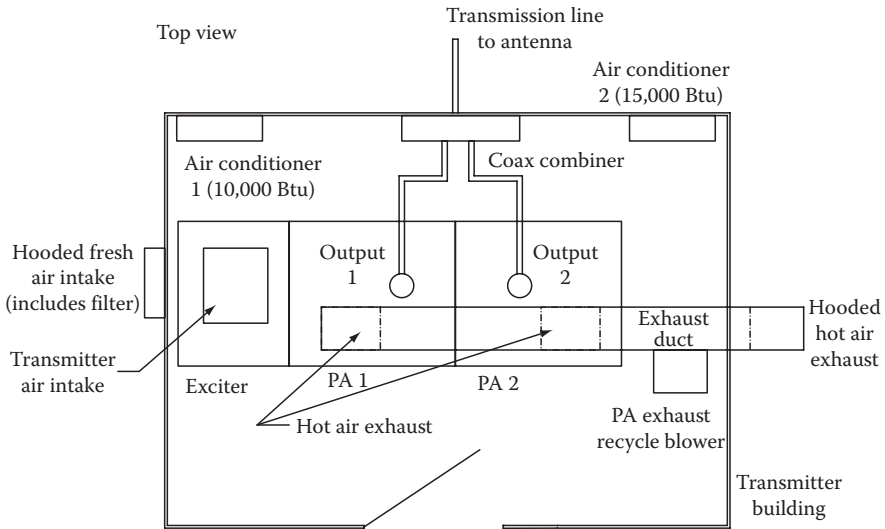


FIGURE 9.27

Typical heating and cooling arrangement for a high-power transmitter installation. Ducting of PA exhaust air should be arranged so that it offers minimum resistance to airflow.

system design. As altitude increases, the density (and cooling capability) of air decreases. A calculated increase in airflow is required to maintain the cooling effectiveness that the system was designed to achieve.

Figure 9.27 shows a typical high-power transmitter plant. The building is oriented so that the cooling activity of the blowers is aided by normal wind currents during the summer months. Air brought in from the outside for cooling is filtered in a hooded air-intake assembly. The building includes a heater and air conditioner.

The layout of a transmitter room HVAC system can have a significant impact on the life of the PA tube(s) and the ultimate reliability of the RF generator. Air intake and output ports must be designed with care to avoid airflow restrictions and back-pressure problems. This process, however, is not as easy as it may seem. The science of airflow is complex and generally requires the advice of a qualified HVAC consultant.

To help illustrate the importance of proper cooling system design and the real-world problems that some facilities have experienced, consider the following examples taken from actual case histories.

9.3.2.1 Case 1

A fully automatic building ventilation system (Figure 9.28) was installed to maintain room temperature at 20°C during the fall, winter, and spring. During the summer, however, ambient room temperature would increase to

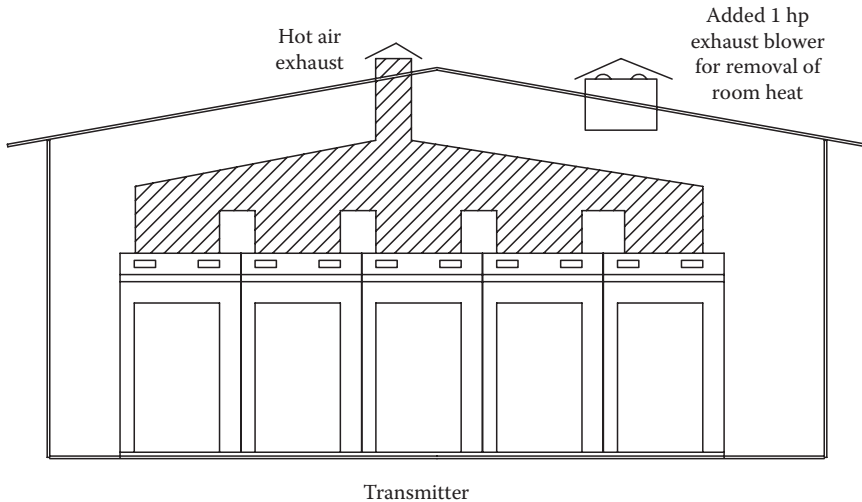


FIGURE 9.28

Case study in which excessive summertime heating was eliminated through the addition of a 1 hp exhaust blower to the building.

as much as 60°C. A field survey showed that the only building exhaust route was through the transmitter. Therefore, air entering the room was heated by test equipment, people, solar radiation on the building, and radiation from the transmitter itself before entering the transmitter. The problem was solved through the addition of an exhaust fan (3000 cfm). The 1 hp fan lowered room temperature by 20°C.

9.3.2.2 Case 2

A simple remote installation was constructed with a heat-recirculating feature for the winter (Figure 9.29). Outside supply air was drawn by the transmitter cooling system blowers through a bank of air filters, and hot air was exhausted through the roof. A small blower and damper were installed near the roof exit point. The damper allowed hot exhaust air to blow back into the room through a tee duct during the winter months. For summer operation, the roof damper was switched open and the room damper closed. For winter operation, the arrangement was reversed. The facility, however, experienced short tube life during winter operation, even though the ambient room temperature during winter was not excessive.

The solution involved moving the roof damper 12 ft down to just above the roof tee. This eliminated the stagnant “air cushion” above the bottom heating duct damper and significantly improved airflow in the region. Cavity back pressure was, therefore, reduced. With this relatively simple modification, the problem of short tube life disappeared.

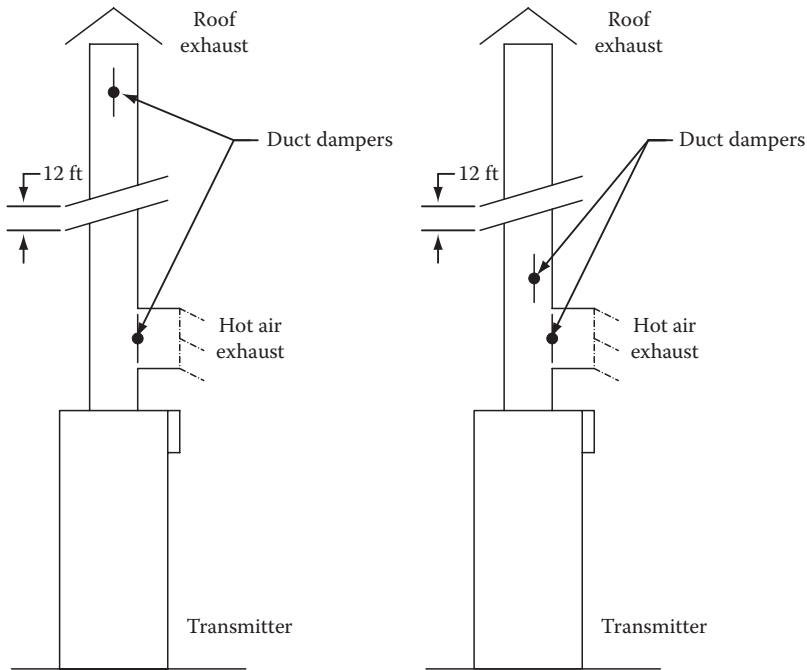


FIGURE 9.29

Case study in which excessive back pressure to the PA cavity was experienced during winter periods, when the rooftop damper was closed. The problem was eliminated by repositioning the damper as shown.

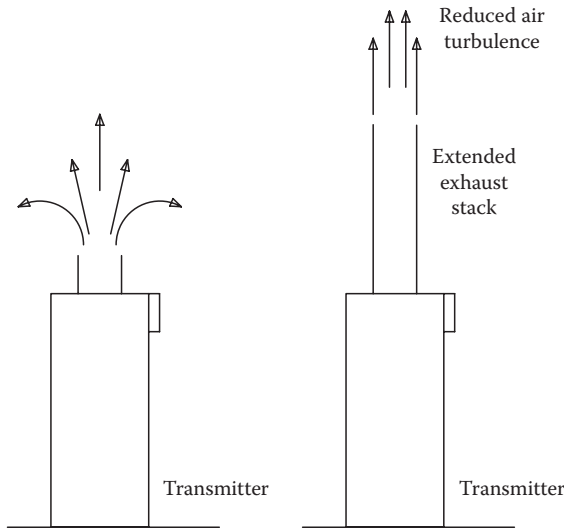
9.3.2.3 Case 3

An inconsistency regarding test data was discovered within a transmitter manufacturer's plant. Units tested in the engineering lab typically ran cooler than those at the manufacturing test facility. Figure 9.30 shows the test station difference, a 4 ft exhaust stack that was used in the engineering lab. The addition of the stack increased airflow by up to 20% because of reduced air turbulence at the output port, resulting in a 20°C decrease in tube temperature.

These examples point out how easily a cooling problem can be caused during HVAC system design. All power delivered to the transmitter either is converted to RF energy and sent to the antenna or becomes heated air (or water). Proper design of a cooling system, therefore, is a part of transmitter installation that should not be taken lightly.

9.3.3 Site Design Guidelines

There are any number of physical plant designs that will provide for reliable operation of high-power RF systems [3]. One constant, however, is the

**FIGURE 9.30**

Case study in which air turbulence at the exhaust duct resulted in reduced airflow through the PA compartment. The problem was eliminated by adding a 4 ft extension to the output duct.

requirement for tight temperature control. Cooling designs can be divided into three broad classifications:

- Closed site design
- Open site design
- Hybrid design

If the equipment user is to provide adequately for hot air exhaust and fresh air intake, the maximum and minimum environmental conditions in which the equipment will operate must be known. In addition, the minimum cooling requirements of the equipment must be provided by the manufacturer. The following parameters should be considered:

- Site altitude
- Maximum expected outside air temperature
- Minimum expected outside air temperature
- Total airflow through the transmitter
- Air temperature rise through the transmitter
- Air exhaust area

Keep in mind that the recommended cooling capacity for a given RF system applies only to cooling the transmitter; any additional cooling load in the building must be considered separately when selecting the air system components.

The transmitter exhaust should not be the only exhaust port in the room because heat from the peripheral equipment would then be forced to go out through the transmitter.

The *sensible-heat load*, then, is the sum of all additional heat loads, including

- Solar radiation
- Heat gains from equipment and lights
- Heat gains from personnel in the area that is to be cooled

9.3.3.1 Closed Site Design

Figure 9.31 illustrates a site layout that works well in most climates, as long as the transmitter is small and the building is sealed and well insulated [3]. In fact, this closed configuration will ensure the longest possible transmitter life and lowest maintenance cost. No outside air laden with moisture and contaminants circulates through the transmitter.

At sites using this arrangement, it has been observed that periodic transmitter cleaning is seldom required. In a typical closed system, the air conditioner is set to cool when the room temperature reaches 75°F–80°F. The closed system also uses a louvered emergency intake blower, which is set by its own thermostat to pull in outside air if the room temperature reaches excessive levels (above 90°F). This blower is required in a closed configuration to prevent the possibility of thermal runaway if the air conditioner fails. Without such an emergency ventilation system, the transmitter would recirculate its own heated air, further heating the room. System failure probably would result.

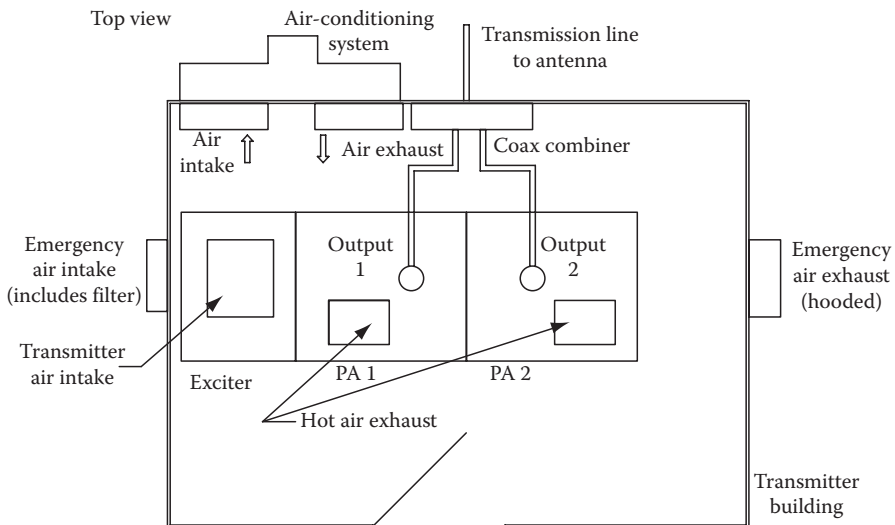


FIGURE 9.31

Closed site ventilation design, with a backup inlet/outlet system.

During winter months, the closed system is self-heating (unless the climate is harsh, or the transmitter power output is low), because the transmitter exhaust is not ducted outside but simply empties into the room. Also during these months, the emergency intake blower can be used to draw cold outside air into the room instead of using the air conditioner, although this negates some of the cleanliness advantages inherent to the closed system.

An exhaust blower should not be substituted for an intake blower, because positive room pressure is desired for venting the room. This ensures that all air in the building has passed through the intake air filter. Furthermore, the louvered emergency exhaust vent(s) should be mounted high in the room, so that hot air is pushed out of the building first.

The closed system usually makes economic sense only if the transmitter exhaust heat load is relatively small.

Periodic maintenance of a closed system involves the following activities:

- Checking and changing the air conditioner filter periodically
- Cleaning the transmitter air filter as needed
- Checking that the emergency vent system works properly
- Keeping the building sealed from insects and rodents

9.3.3.2 Open Site Design

Figure 9.32 depicts a site layout that is the most economical to construct and operate [3]. The main attribute of this approach is that the transmitter air

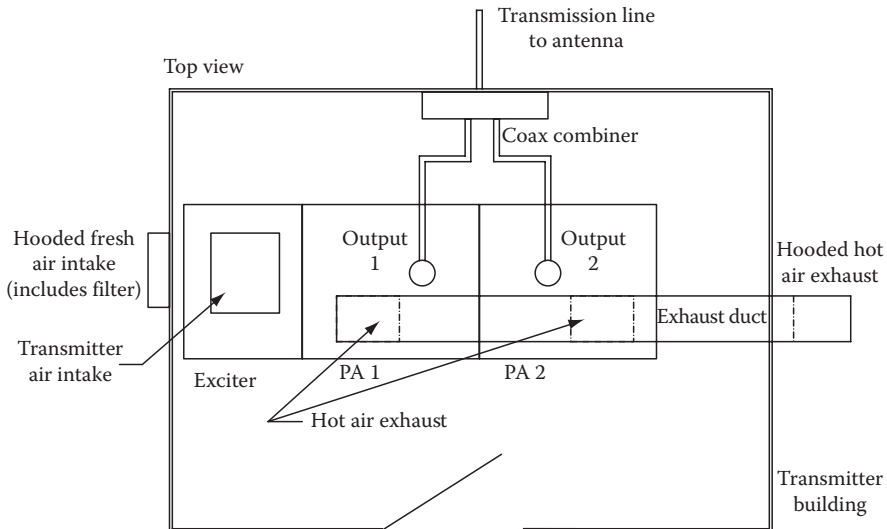


FIGURE 9.32
Open site ventilation design, using no air conditioner.

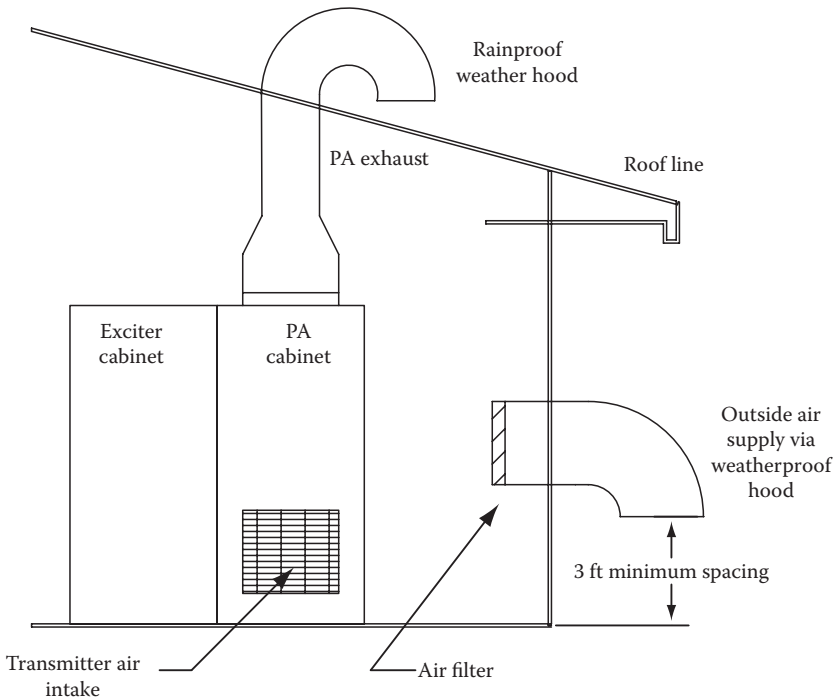


FIGURE 9.33
Transmitter exhaust-collection hood.

supply is not heated or cooled, resulting in cost savings. This is not a closed system; outdoor air is pumped into the transmitter room through air filters. The transmitter then exhausts the hot air. If the duct work is kept simple and the transmitter has a dedicated exhaust port, such a direct exhaust system works adequately. Many transmitters do not lend themselves to a direct exhaust connection, however, and a hood mounted over the transmitter may be required to collect the hot air, as illustrated in Figure 9.33. With a hooded arrangement, it may be necessary to install a booster fan in the system, typically at the wall or roof exit, to avoid excessive back pressure.

When building an open air-circulation system, there are some important additional considerations. The room must be positively pressurized so that only filtered air, which has come through the intake blower and filter, is available to the transmitter. With a negatively pressurized room, air will enter through every hole and crack in the building, and will not necessarily be filtered. Under negative pressure, the transmitter blower also will have to work harder to exhaust air.

The intake blower should be a “squirrel-cage” type rather than a fan type. A fan is meant to move air within a pressurized environment; it cannot compress the air. A squirrel-cage blower will not only move air, but also will pressurize the area into which the air is directed. Furthermore, the intake air blower

must be rated for more cfm airflow than the transmitter will exhaust outside. A typical 20kW FM broadcast transmitter, for example, will exhaust 500–1000 cfm. A blower of 1200 cfm, therefore, would be an appropriate size to replenish the transmitter exhaust and positively pressurize the room.

In moderate and warm climates, the intake blower should preferably be located on a north-facing outside wall. If the air intake is on the roof, it should be elevated so that it does not pick up air heated by the roof surface.

High-quality pleated air filters are recommended. Home-style fiberglass filters are not sufficient. Local conditions may warrant using a double filtration system, with coarse and fine filters in series.

Secure the advice of a knowledgeable HVAC shop when designing filter boxes. For a given cfm requirement, the larger the filtration area, the lower the required air velocity through the filters. This lower velocity results in better filtration than forcing more air through a small filter. In addition, the filters will last longer. Good commercial filtration blowers designed for commercial installation are available from industrial supply houses and are readily adapted to RF facility use.

The transmitter exhaust may be ducted through a nearby wall or through the roof. Avoid ducting straight up, however. Many facilities have suffered water damage to transmitters in such cases because of the inevitable deterioration of roofing materials. Normally, ducting the exhaust through an outside wall is acceptable. The duct work typically is bent downward a foot or two outside the building to keep direct wind from creating back pressure in the exhaust duct. Minimize all bends in the duct work. If a 90° bend must be made, it should be a large-radius bend with curved *helper* vanes inside the duct to minimize turbulence and back pressure. A 90° L- or T-bend is not recommended, unless oversized and equipped with internal vanes to assist the turning airflow.

For moderate and cool climates, an automatic damper can be employed in the exhaust duct to direct a certain amount of hot exhaust air back into the transmitter building as needed for heating. This will reduce outside air requirements, providing clean, dry, heated air to the transmitter during cold weather.

9.3.3.3 Hybrid Design

Figure 9.34 depicts a hybrid site layout that is often used for high-power transmitters or sites supporting multiple transmitters [3]. As with the layout in Figure 9.32, the room is positively pressurized with clean, filtered air from the outside. A portion of this outside air is then drawn through the air conditioner for cooling before delivery to the transmitter area. Although not all of the air in the room goes through the air conditioner, enough does to make a difference in the room temperature. In humid areas, much of the moisture is removed by the air conditioner. The transmitter exhaust is directed outside.

Such a hybrid system is often the choice for larger transmitter sites where a closed system would prove too costly, but an unconditioned system would run excessively hot during the summer months.

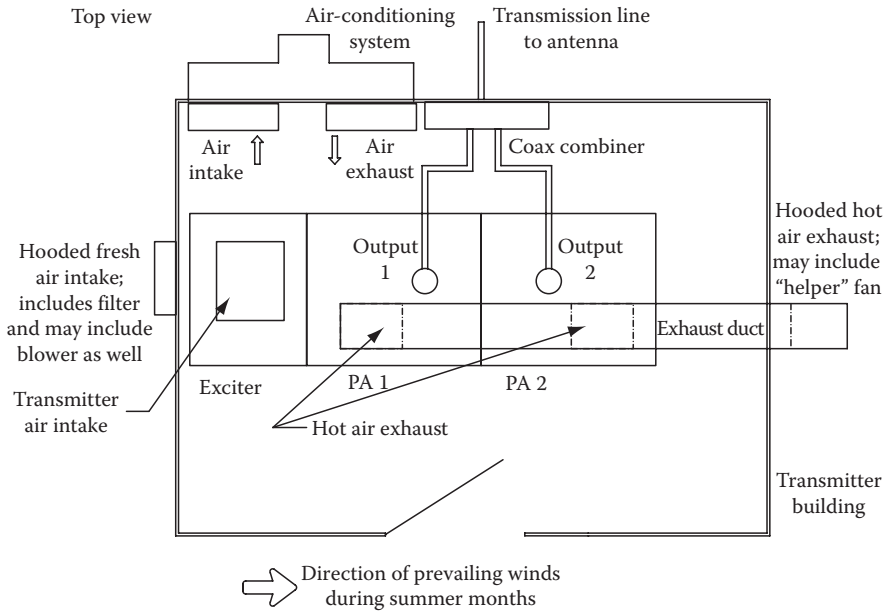


FIGURE 9.34

Hybrid site ventilation design, using an air conditioner in addition to filtered outside air.

Some closed or hybrid systems use two parallel air conditioners. Most of the time, only one runs. The thermostats of the units are staggered so that if one cannot keep the air below the ideal operating point, the other will turn on to assist.

9.3.4 Water/Vapor Cooling System Maintenance

The cooling system is vital to any RF generator. In a klystron-based unit, for example, the cooling system may dissipate a large percentage of the input ac power in the form of waste heat in the collector. Pure, distilled water should be used. Deionized water is required unless the isolation of high voltages is unnecessary, as in the case of the collector of a klystron (the collector typically operates at or near ground potential).

Any impurities in the cooling system water eventually will find their way into the water jacket and cause corrosion of the plate or collector. It is essential to use high-purity water with low conductivity, and to replace the water in the cooling system as needed. Efficient energy transfer from the heated surface into the water is necessary for long vacuum tube life. Oil, grease, soldering flux residue, and pipe sealant containing silicone compounds must be excluded from the cooling system. This applies to both vapor- and liquid-conduction cooling systems, although it is usually more critical in the vapor-phase type. The sight glass in a vapor-phase water jacket provides a convenient checkpoint for coolant condition.

In general, greater flows and greater pressures are inherent in liquid-cooled versus vapor-phase systems, and, when a leak occurs, large quantities of coolant can be lost before the problem is discovered. The condition of gaskets, seals, and fittings, therefore, should be checked regularly. Many liquid-cooled tubes use a distilled water and ethylene glycol mixture. (Do not exceed a 50:50 mix by volume.) The heat transfer of the mixture is lower than that of pure water, requiring the flow to be increased, typically by 25%. Greater coolant flow suggests closer inspection of the cooling system after adding the glycol.

The action of heat and air on ethylene glycol causes the formation of acidic products. The acidity of the coolant can be checked with litmus paper. Buffers can be added to the glycol mixture. Buffers are alkaline salts that neutralize acid forms and prevent corrosion. Because they are ionizable chemical salts, the buffers cause the conductivity of the coolant to increase. Consult the power tube manufacturer before adding buffers.

The following general guidelines are recommended for power tube cooling systems:

- The resistivity of the water must be maintained above the minimum level established by the power tube manufacturer. Although recommended values vary, generally speaking, resistivity should be maintained at or above $1\text{ M}\Omega\text{ cm}$ (at 30°C) for power tubes operating at a high voltage, referenced to ground. A resistivity value of $30\text{ k}\Omega\text{ cm}$ or greater (at 30°C) is usually sufficient for tubes whose collector operates at or near ground potential (such as a klystron).
- The pH factor must be within the range of 6.0–8.0.
- The particulate matter size must not be greater than $50\text{ }\mu\text{m}$ (325 mesh).
- The inlet water temperature must not exceed 70°C , and this temperature should be regulated to $\pm 5^\circ\text{C}$.

When the water in the cooling system fails to satisfy any one of these requirements, prompt action is necessary to correct the deficiency. If the water is contaminated, the system must be flushed and replaced with clean water.

9.3.4.1 Tests for Purity

The recommended method for measuring resistivity, pH, and particulate matter is by use of laboratory instruments. However, the following simple tests should provide sufficient information to determine the general condition of the coolant.

- The appearance of the water is a good general indicator of coolant condition. If the water looks turbid or tastes or smells brackish, it is good practice to change the water and flush the system.

- The pH factor is easily checked by using pH paper or litmus paper. Sold under a variety of trade names, pH paper is available at laboratory supply stores. Directions for use usually are printed on the package.
- Particulate matter and impurities can be checked by using the “foaming test” (described in the next section).
- Resistivity can be measured accurately with a resistance bridge. Generally speaking, if the water does not pass the foaming test, the resistivity will be low as well.

9.3.4.2 Foaming Test for Water Purity

One of the many factors affecting the life and operating efficiency of vapor-cooled tubes is the purity of the water in the cooling system. If impurities are present, foaming may occur that will inhibit heat transfer, thereby lowering the cooling efficiency of the system. The impurities that most frequently produce foaming are

- Cleaning-compound residue
- Detergents
- Joint-sealing compounds
- Oily rust preventives in pumps and other components
- Valve-stem packing
- Impurities in tap water

Tests should be conducted periodically to confirm the quality of the coolant in a vapor-phase system. Conduct such tests after each water change, system cleaning, or modification. The following items are required:

- A 1/2 in. × 4 in. glass test tube with rubber stopper
- A 1-pint glass or polypropylene bottle with cap

The procedure is as follows:

- Fill the cooling system with water and circulate it until thoroughly mixed (about 30 min).
- Drain a sample of water into the bottle, and cool to room temperature.
- If the water sample stands for more than 1 h, slowly invert the capped bottle about 10 times. Avoid shaking the bottle because this will create air bubbles in the water. (When the water is static, foaming impurities tend to collect at the surface. This step mixes the sample without generating foam.)
- Using the sample water, rinse the test tube and stopper three times.
- Fill the test tube to the halfway mark with the sample water.

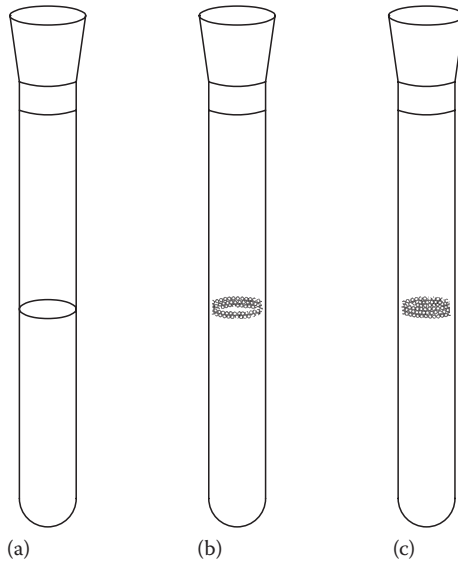


FIGURE 9.35

Results of the “foaming test” for water purity: (a) pure sample, (b) sample that is acceptable on a short-term basis, and (c) unacceptable sample.

- Shake the test tube vigorously for 15 s.
- Let the sample stand for 15 s.
- Observe the amount of foam remaining on top of the water and compare it with the drawings shown in Figure 9.35.

A completely foam-free water surface and test tube, as illustrated in Figure 9.35a, indicates no foam-producing impurities. If the water surface and test tube wall are partially covered with foam, but a circle of clear water appears in the center, the impurity level is temporarily acceptable (Figure 9.35b). A second test should be made in about 1 week. If the foam layer completely bridges the inside of the test tube (Figure 9.35c), the system should be flushed and cleaned.

References

1. Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian Associates, San Carlos, CA, 1984.
2. Besch, D. F., Thermal properties, in *The Electronics Handbook*, J. C. Whitaker (ed.), CRC Press, Boca Raton, FL, pp. 127–134, 1996.
3. Harnack, K., Airflow and cooling in RF facilities, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, pp. 33–38, November 1992.

Bibliography

- Cleaning and flushing klystron water and vapor cooling systems, Application Engineering Bulletin no. AEB-32, Varian Associates, Palo Alto, CA, February 1982.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, pp. 547–550, 1992.
- Foaming test for water purity, Application Engineering Bulletin no. AEB-26, Varian Associates, Palo Alto, CA, September 1971.
- Harper, C. A., *Electronic Packaging and Interconnection Handbook*, McGraw-Hill, New York, 1991.
- High power transmitting tubes for broadcasting and research, Philips Technical Publication, Eindhoven, the Netherlands, 1988.
- Integral cavity klystrons for UHF-TV transmitters, Varian Associates, Palo Alto, CA.
- Kimmel, E., Temperature sensitive indicators: How they work, when to use them, *Heat Treating*, March 1983.
- Power grid tubes for radio broadcasting, Thomson-CSF Publication #DTE-115, Thomson-CSF, Dover, NJ, 1986.
- Protecting integral-cavity klystrons against water leakage, Technical Bulletin no. 3834, Varian Associates, Palo Alto, CA, July 1978.
- Temperature measurements with Eimac power tubes, Application Bulletin no. 20, Varian Associates, San Carlos, CA, January 1984.
- Water purity requirements in liquid cooling systems, Application Bulletin no. 16, Varian Associates, San Carlos, CA, July 1977.
- Water purity requirements in water and vapor cooling systems, Application Engineering Bulletin no. AEB-31, Varian Associates, Palo Alto, CA, September 1991.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1990.
- Whitaker, J. C., *Maintaining Electronic Systems*, CRC Press, Boca Raton, FL, 1992.
- Wozniak, J., Using tetrode power amplifiers in high power UHF-TV transmitters, in *NAB 1992 Broadcast Engineering Conference Proceedings*, National Association of Broadcasters, Washington, DC, pp. 209–211, 1992.

10

Reliability Considerations

10.1 Introduction

The ultimate goal of any design engineer or maintenance department is zero downtime. This is an elusive goal, but one that can be approximated by examining the vulnerable areas of plant operation and taking steps to prevent a sequence of events that could result in system failure. In cases where failure prevention is not practical, a reliability assessment should encompass the stocking of spare parts, circuit boards, or even entire systems. A large facility may be able to cost-justify the purchase of backup gear that can be used as spares for the entire complex. Backup hardware is expensive, but so is downtime. Because of the finite lifetime of power vacuum tubes, these considerations are of special significance.

Failures can, and do, occur in electronic systems. The goal of product quality assurance (QA) at every step in the manufacturing and operating chain is to ensure that failures do not produce a systematic or repeatable pattern. The ideal is to eliminate failures altogether. Short of that, the secondary goal is to end up with a random distribution of failure modes. This indicates that the design of the system is fundamentally optimized and that failures are caused by random events that cannot be predicted. In an imperfect world, this is often the best that end users can hope for. Reliability and maintainability must be built into products or systems at every step in the design, construction, and maintenance process. They cannot be treated as an afterthought.

10.1.1 Terminology

To understand the principles of reliability engineering, the following basic terms must be defined:

- *Availability.* The probability that a system subject to repair will operate satisfactorily on demand.
- *Average life.* The mean value for a normal distribution of product or component lives. This term is generally applied to mechanical failures resulting from “wear-out.”

- *Burn-in*. The initially high failure rate encountered when a component is placed on test. Burn-in failures usually are associated with manufacturing defects and the debugging phase of early service.
- *Defect*. Any deviation of a unit or product from specified requirements. A unit or product may contain more than one defect.
- *Degradation failure*. A failure that results from a gradual change, over time, in the performance characteristics of a system or part.
- *Downtime*. Time during which equipment is not capable of doing useful work because of malfunction. This does not include preventive maintenance time. Downtime is measured from the occurrence of a malfunction to its correction.
- *Failure*. A detected cessation of ability to perform a specified function or functions within previously established limits. A failure is beyond adjustment by the operator by means of controls normally accessible during routine operation of the system.
- *Failure mode and effects analysis (FMEA)*. An iterative documented process performed to identify basic faults at the component level and determine their effects at higher levels of assembly.
- *Failure rate*. The rate at which failure occurs during an interval of time as a function of the total interval length.
- *Fault tree analysis (FTA)*. An iterative documented process of a systematic nature performed to identify basic faults, determine their causes and effects, and establish their probabilities of occurrence.
- *Lot size*. A specific quantity of similar material or a collection of similar units from a common source; in inspection work, the quantity offered for inspection and acceptance at any one time. This may be a collection of raw material, parts, subassemblies inspected during production, or a consignment of finished products to be sent out for service.
- *Maintainability*. The probability that a failure will be repaired within a specified time after it occurs.
- *Mean time between failures (MTBF)*. The measured operating time of a single piece of equipment divided by the total number of failures during the measured period of time. This measurement normally is made during that period between early life and wear-out failures.
- *Mean time to repair (MTTR)*. The measured repair time divided by the total number of failures of the equipment.
- *Mode of failure*. The physical description of the manner in which a failure occurs and the operating condition of the equipment or part at the time of the failure.

- *Part failure rate.* The rate at which a part fails to perform its intended function.
- *Quality assurance.* All those activities, including surveillance, inspection, control, and documentation, aimed at ensuring that a product will meet its performance specifications.
- *Reliability.* The probability that an item will perform satisfactorily for a specified period of time under a stated set of use conditions.
- *Reliability growth.* Actions taken to move a hardware item toward its reliability potential, during development, subsequent manufacturing, and/or operation.
- *Reliability predictions.* Compiled failure rates for parts, components, subassemblies, assemblies, and systems. These generic failure rates are used as basic data to predict a value for reliability.
- *Sample.* One or more units selected at random from a quantity of product to represent that product for inspection purposes.
- *Sequential sampling.* Sampling inspection in which, after each unit is inspected, the decision is made to accept, reject, or inspect another unit. (Note: Sequential sampling as defined here is sometimes called *unit sequential sampling* or *multiple sampling*.)
- *System.* A combination of parts, assemblies, and sets joined together to perform a specific operational function or functions.
- *Test to failure.* Testing conducted on one or more items until a predetermined number of failures have been observed. Failures are induced by increasing electrical, mechanical, and/or environmental stress levels, usually in contrast to *life tests*, in which failures occur after extended exposure to predetermined stress levels. A life test can be considered a test to failure using age as the stress.

10.2 Quality Assurance

Electronic component and system manufacturers design and implement QA procedures for one fundamental reason: nothing is perfect. The goal of a QA program is to ensure, for both the manufacturer and the customer, that all but some small, mutually acceptable percentage of devices or systems shipped will be as close to perfection as economics and the state of the art allow. There are trade-offs in this process. It would be unrealistic, for example, to perform extensive testing to identify potential failures if the cost of that testing exceeded the cost savings that would be realized by not having to replace the devices later in the field.

The focal points of any QA effort are *quality* and *reliability*. These terms are not synonymous. They are related, but they do not provide the same measure of a product:

- Quality is the measure of a product's performance relative to some established criteria.
- Reliability is the measure of a product's life expectancy.

Stated from a different perspective, quality answers the question of whether the product meets applicable specifications *now*; reliability answers the question of *how long* the product will continue to meet its specifications.

10.2.1 Inspection Process

QA for components normally is performed through sampling rather than through 100% inspection. The primary means used by QA departments for controlling product quality at the various processing steps include the following:

- *Gates*. A mandatory sampling of every lot passing through a critical production stage. Material cannot move on to the next operation until QA has inspected and accepted the lot.
- *Monitor points*. A periodic sampling of some attribute of the component. QA personnel sample devices at a predetermined frequency to verify that machines and operators are producing material that meets preestablished criteria.
- *Quality audit*. An audit carried out by a separate group within the QA department. This group is charged with ensuring that all production steps throughout the manufacturer's facility are in accordance with current specifications.
- *Statistical quality control*. A technique, based on computer modeling, that incorporates data accumulated at each gate and monitor point to construct statistical profiles for each product, operation, and piece of equipment within the plant. Analysis of these data over time allows QA engineers to assess trends in product performance and failure rates.

QA for a finished subassembly or system may range from a simple go/no-go test to a thorough operational checkout that may take days to complete.

10.2.2 Reliability Evaluation

Reliability prediction is the process of quantitatively assessing the reliability of a component or system during development, before large-scale fabrication and field operation. During product development, predictions serve as a guide by which design alternatives can be judged for reliability. To be effective, the prediction technique must relate engineering variables to reliability variables.

A prediction of reliability is obtained by determining the reliability of each critical item at the lowest system level and proceeding through intermediate levels until an estimate of overall reliability is obtained. This prediction method depends on the availability of accurate evaluation models that reflect the reliability of lower-level components. Various formal prediction procedures are used, based on theoretical and statistical concepts.

10.2.2.1 Parts-Count Method

Although the parts-count method does not relate directly to vacuum tube design, it is worthwhile to discuss the technique briefly in the context of overall system reliability. It is fair to point out that this technique has considerable validity when comparisons of vacuum tube vs. solid-state amplifier designs are being considered. To achieve high power levels, far more individual components are necessary for a solid-state design than for a vacuum tube-based system.

The parts-count approach to reliability prediction provides an estimate of reliability based on a count by part type (ICs, transistors, vacuum tube devices, resistors, capacitors, and other components). This method is useful during the early design stage of a product, when the amount of available detail is limited. The technique involves counting the number of components of each type, multiplying that number by a generic failure rate for each part type, and summing the products to obtain the failure rate of each functional circuit, subassembly, assembly, and/or block depicted in the system block diagram. The parts-count method is useful in the design phase because it provides rapid estimates of reliability, permitting assessment of the feasibility of a given concept.

10.2.2.2 Stress-Analysis Method

The stress-analysis technique is similar to the parts-count method, but utilizes a detailed parts model plus calculation of circuit stress values for each part before determining the failure rate. Each part is evaluated in its electric circuit and mechanical assembly application based on an electrical and thermal stress analysis. After part failure rates have been established, a combined failure rate for each functional block is determined.

10.2.3 Failure Analysis

FMEA can be performed with data taken from actual failure modes observed in the field or from hypothetical failure modes derived from one of the following:

- Design analysis
- Reliability prediction activities
- Experience with how specific parts fail

In the most complete form of FMEA, failure modes are identified at the component level. Failures are induced analytically into each component, and failure effects are evaluated and noted, including the severity and frequency (or probability) of occurrence. Using this approach, the probability of various system failures can be calculated, based on the probability of lower-level failure modes.

FTA is a tool commonly used to analyze failure modes found during design, factory test, or field operations. The approach involves several steps, including the development of a detailed logic diagram that depicts basic faults and events that can lead to system failure and/or safety hazards. These data are used to formulate corrective suggestions that, when implemented, will eliminate or minimize faults considered critical. An example FTA chart is shown in Figure 10.1.

10.2.4 Standardization

Standardization and reliability go hand in hand. Standardization of electronic components began with military applications in mind; the first

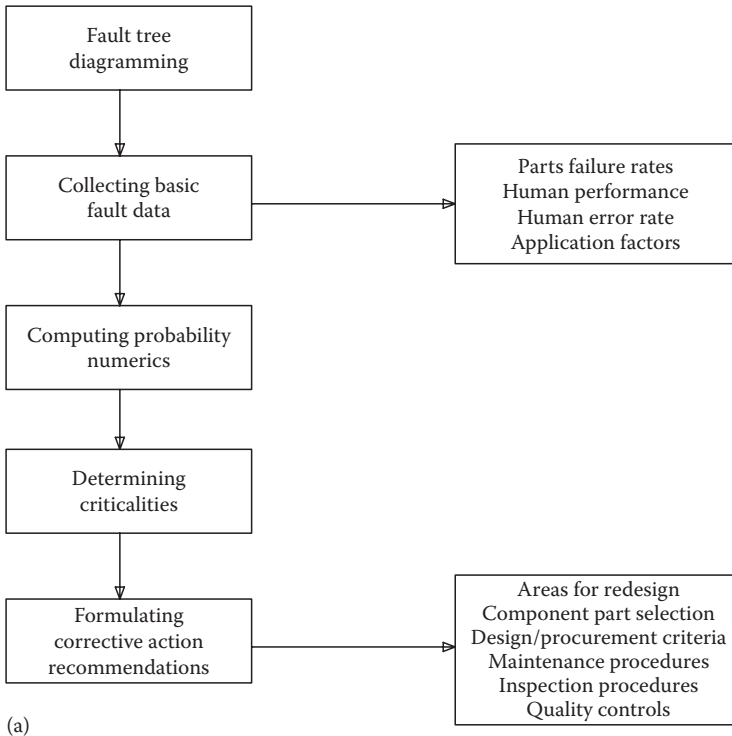







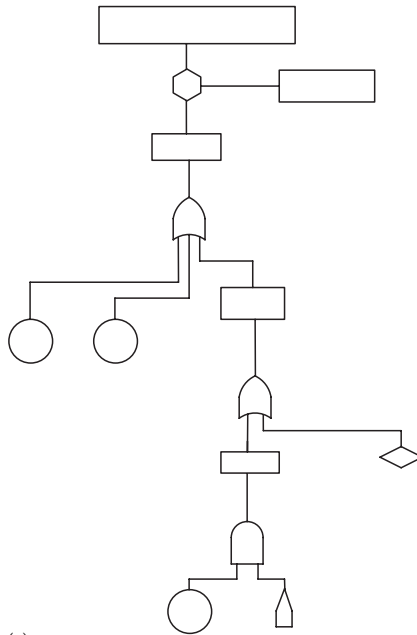


FIGURE 10.1

Example FTA diagram: (a) process steps.

-  An event or fault resulting from the combination of more basic faults, which can be further developed.
-  A basic fault (usually a specific circuit, part, or human error) that can be assigned a probability of occurrence.
-  A fault not developed further because of a lack of information, time, or value in doing so.
-  AND gate: The output event occurs only when all of the input events are present.
-  OR gate: The output occurs when one or more of the input events are present.
-  Inhibit gate: Similar to an AND gate, but used to include applications of a conditional event.
-  An event expected to occur in normal operation.

(b)



(c)

FIGURE 10.1 (continued)
 (b) fault tree symbols, and (c) example diagram.

recorded work was performed by the U.S. Navy with vacuum tubes. The navy recognized that some control at the component level was essential to the successful incorporation of electronics into naval systems.

Standardization and reliability are closely related, although there are many aspects of standardization whose reliability implications are subtle. The primary advantages of standardization include

- *Total product interchangeability.* Standardization ensures that products of the same part number provide the same physical and electrical characteristics. There have been innumerable instances of a replacement device bearing the same part number as a failed device, but not functioning identically to it. In many cases, the differences in performance were so great that the system would not function at all with the new device.
- *Consistency and configuration control.* Component manufacturers constantly redefine their products to improve yields and performance. Consistency and configuration control assure the user that product changes will not affect the interchangeability of the part.
- *Efficiencies of volume production.* Standardization programs usually result in production efficiencies that reduce the costs of parts, relative to components with the same level of reliability screening and control.
- *Effective spares management.* The use of standardized components makes the stocking of spare parts a much easier task. This aspect of standardization is not a minor consideration.
- *Multiple product sources.* Standardization encourages second-sourcing. Multiple sources help hold down product costs and encourage manufacturers to strive for better product performance.

10.3 Reliability Analysis

The science of reliability and maintainability matured during the 1960s with the development of sophisticated computer systems and complex military and spacecraft electronics. Components and systems never fail without a reason. That reason may be difficult to find, but determination of failure modes and weak areas in system design or installation is fundamental to increasing the reliability of any component or system, whether it is a power vacuum tube, integrated circuit, aircraft autopilot, or broadcast transmitter.

All equipment failures are logical; some are predictable. A system failure usually is related to poor-quality components or to abuse of the system or

a part within, either because of underrating or environmental stress. Even the best-designed components can be badly manufactured. A process can go awry, or a step involving operator intervention may result in an occasional device that is substandard or likely to fail under normal stress. Hence, the process of screening and/or *burn-in* to weed out problem parts is a universally accepted quality control tool for achieving high reliability.

10.3.1 Statistical Reliability

Figure 10.2 illustrates what is commonly known as the *bathtub curve*. It divides the expected lifetime of a class of parts into three segments: *infant mortality*, *useful life*, and *wear-out*. A typical burn-in procedure consists of the following steps:

- The parts are electrically biased and loaded; that is, they are connected in a circuit representing a typical application.
- The parts are cycled on and off (power applied, then removed) for a predetermined number of times. The number of cycles may range from 10 to several thousand during the burn-in period, depending on the component under test.
- The components under load are exposed to high operating temperatures for a selected time (typically 72–168 h). This constitutes an accelerated life test for the part.

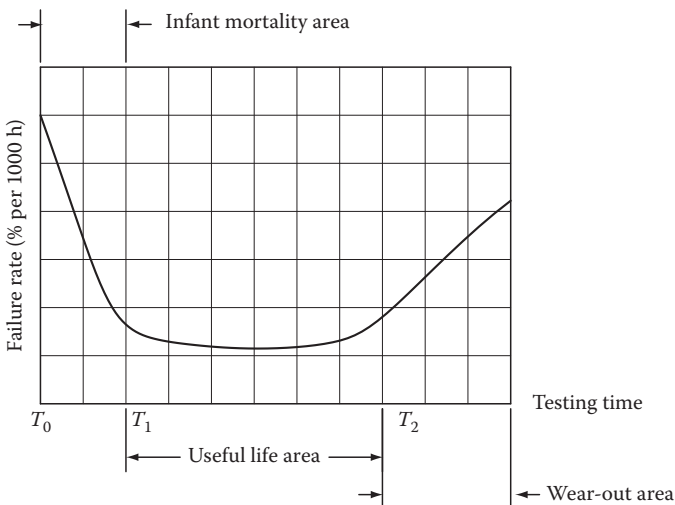


FIGURE 10.2

Statistical distribution of equipment or component failures vs. time for electronic systems and devices.

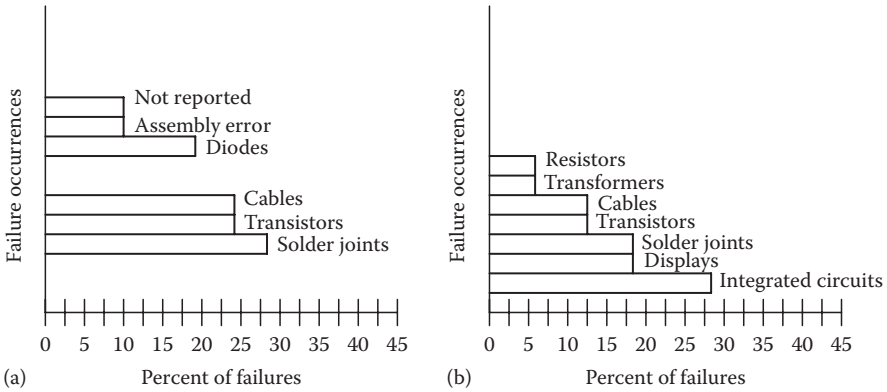


FIGURE 10.3

Distribution of component failures identified through burn-in testing: (a) steady-state high-temperature burn-in and (b) temperature cycling.

An alternative approach involves temperature shock testing, in which the component product is subjected to temperature extremes, with rapid changes between the *hot-soak* and *cold-soak* conditions. After the stress period, the components are tested for adherence to specifications. Parts meeting the established specifications are accepted for shipment to customers. Parts that fail to meet them are discarded.

Figure 10.3 illustrates the benefits of temperature cycling to product reliability. The charts compare the distribution of component failures identified through steady-state high-temperature burn-in vs. temperature cycling. Note that cycling screened out a significant number of failures. The distribution of failures under temperature cycling usually resembles the distribution of field failures. Temperature cycling simulates real-world conditions more closely than steady-state burn-in. The goal of burn-in testing is to ensure that the component lot is advanced beyond the infant mortality stage (T_1 on the bathtub curve). This process is used not only for individual components, but for entire systems as well.

Such a systems approach to reliability is effective, but not foolproof. The burn-in period is a function of statistical analysis; there are no absolute guarantees. The natural enemies of electronic parts are heat, vibration, and excessive voltage. Figure 10.4 documents failures vs. hours in the field for a piece of avionics equipment. The conclusion is made that a burn-in period of 200 h or more will eliminate 60% of the expected failures. However, the burn-in period for another system using different components may well be a different number of hours.

The goal of burn-in testing is to catch system problems and potential faults before the device or unit leaves the manufacturer. The longer the burn-in period, the greater the likelihood of catching additional failures. The problems with extended burn-in, however, are time and money. Longer burn-in translates to longer delivery delays and additional costs for the equipment

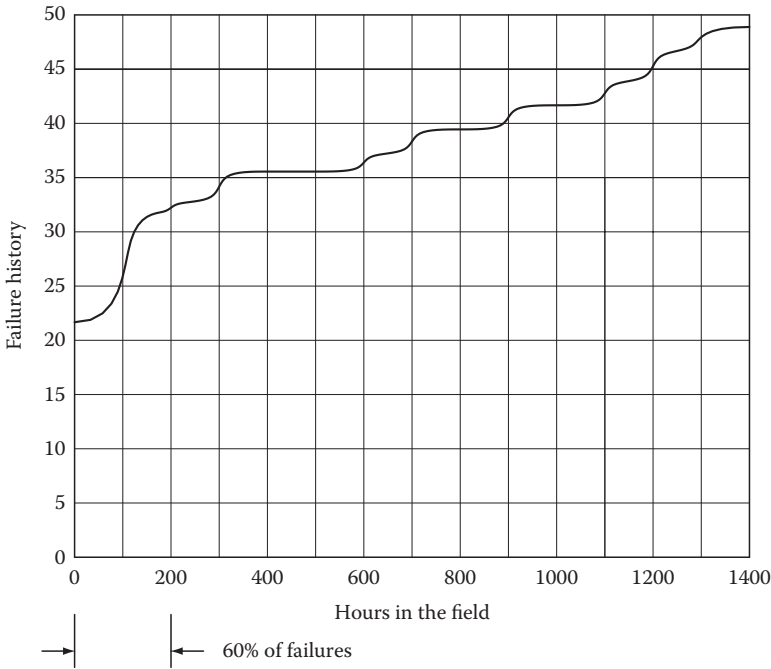


FIGURE 10.4 Failure history of a piece of avionics equipment vs. time. Note that 60% of the failures occurred within the first 200 h of service. (After Capitano, J. and J. Feinstein, Environmental stress screening demonstrates its value in the field, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.)

manufacturer, which are likely to be passed on to the end user. The point at which a product is shipped is based largely on experience with similar components or systems and the financial requirement to move products to customers.

10.3.1.1 Roller-Coaster Hazard Rate

The bathtub curve has been used for decades to represent the failure rate of an electronic system. A variation on this measure, known in the trade as the *hazard rate*, represents failure rate as a roller-coaster track, as illustrated in Figure 10.5. Hazard rate calculations require analysis of the number of failures of the system under test, as well as the number of survivors. Advocates of this approach point out that previous estimating processes and averaging tended to smooth the roller-coaster curve so that the humps were less pronounced, leading to an incorrect conclusion insofar as the hazard rate was concerned. The testing environment also has a significant effect on the shape of the hazard curve, as illustrated in Figure 10.6. Note that at the higher operating temperature (greater environmental stress), the roller-coaster hump has moved to an earlier age.

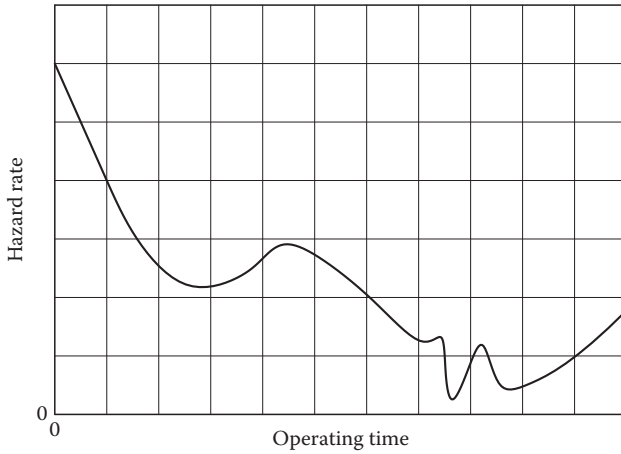


FIGURE 10.5 Roller-coaster hazard rate curve for electronic systems. (After Wong, K.L., Demonstrating reliability and reliability growth with environmental stress screening data, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.)

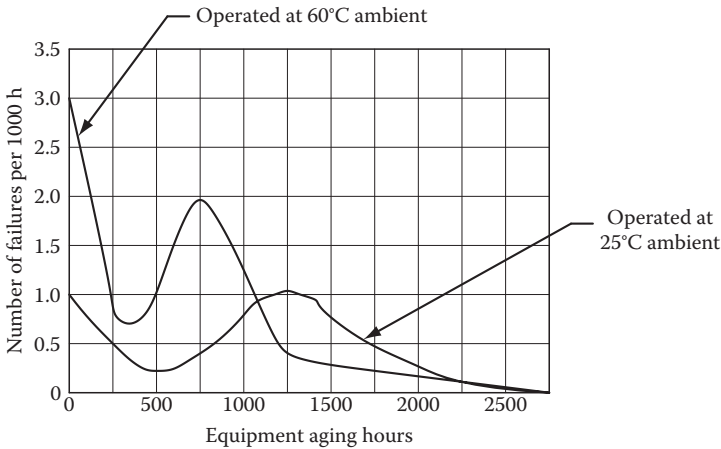


FIGURE 10.6 Effects of environmental conditions on the roller-coaster hazard rate curve. (After Wong, K.L., Demonstrating reliability and reliability growth with environmental stress screening data, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.)

10.3.2 Environmental Stress Screening

The science of reliability analysis is rooted in the understanding that there is no such thing as a random failure; every failure has a cause. For reasonably designed and constructed electronic equipment, failures not caused by outside forces result from built-in flaws or *latent defects*. Because different flaws are sensitive to different stresses, a variety of environmental forces must be

applied to a unit under test to identify any latent defects. This is the underlying concept behind *environmental stress screening* (ESS).

ESS, which has come into widespread use for aeronautics and military products, takes the burn-in process a step further by combining two of the major environmental factors that cause parts or units to fail: heat and vibration. Qualification testing for products at a factory practicing ESS involves a carefully planned series of checks for each unit off the assembly line. Units are subjected to random vibration and temperature cycling during production (for subassemblies and discrete components) and upon completion (for systems). The goal is to catch product defects at the earliest possible stage of production. ESS also can lead to product improvements in design and manufacture if feedback from the qualification stage to the design and manufacturing stages is implemented. Figure 10.7 illustrates the improvement in reliability that typically can be achieved through ESS over simple burn-in screening and through ESS with feedback to earlier production stages. Significant reductions in equipment failures in the field testing can be gained. Table 10.1 compares the merits of conventional reliability testing and ESS.

Designing an ESS procedure for a given product is no easy task. The environmental stresses imposed on the product must be great enough to cause fallout of marginal components during qualification testing. The stresses must not be so great, however, as to cause failures in good products. Any unit that is stressed beyond its design limits eventually will fail. The proper selection of stress parameters—generally, random vibration on a vibration generator and temperature cycling in an environmental chamber—can, in minutes,

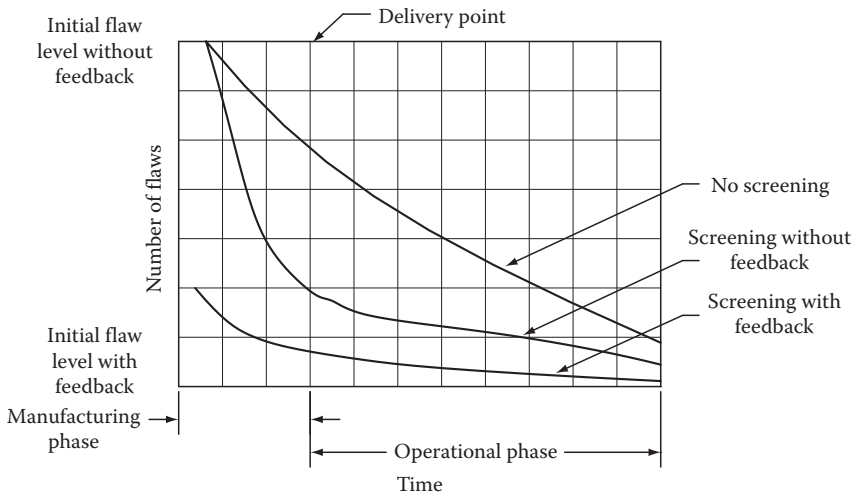


FIGURE 10.7 Effects of ESS on the reliability bathtub curve. (After Tustin, W., Recipe for reliability: Shake and bake, *IEEE Spectrum*, IEEE, New York, December 1986.)

TABLE 10.1

Comparison of Conventional Reliability Testing and Environmental Stress Screening

Parameter	Conventional Testing	Environmental Stress Screening
Test condition	Simulates operational equipment profile	Accelerated stress condition
Test sample size	Small	100% of production
Total test time	Limited	High
Number of failures	Small	Large
Reliability growth	Potential for gathering useful data small	Potential for gathering useful data good

Source: After Wong, K.L., Demonstrating reliability and reliability growth with environmental stress screening data, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.

uncover product flaws that might take weeks or months to manifest themselves in the field. The result is greater product reliability for the user.

The ESS concept requires that every product undergo qualification testing before integration into a larger system for shipment to an end user. The flaws uncovered by ESS vary from one unit to the next, but types of failures tend to respond to particular environmental stresses. Available data clearly demonstrate that the burn-in screens must match the flaws sought; otherwise, the flaws will probably not be found.

The concept of flaw–stimulus relationships can be presented in Venn diagram form. Figure 10.8 shows a Venn diagram for a hypothetical, but specific, product.

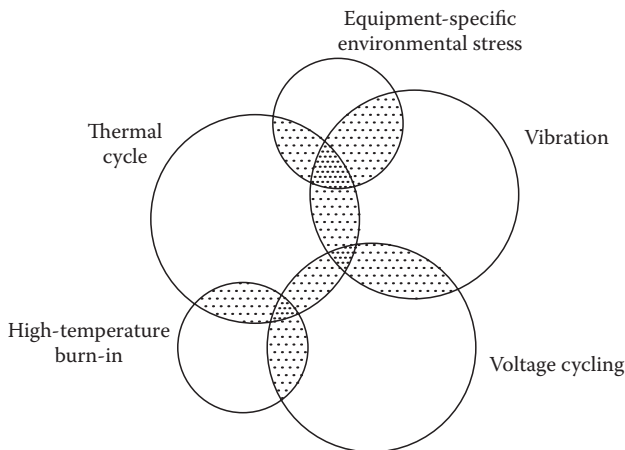


FIGURE 10.8

Venn diagram representation of the relationship between flaw precipitation and applied environmental stress. (After Hobbs, G.K., Development of stress screens, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.)

The required screen would be different for a different product. For clarity, not all stimuli are shown. Note that there are many latent defects that will not be uncovered by any one stimulus. For example, a solder splash that is just barely clinging to a circuit element probably would not be broken loose by high-temperature burn-in or voltage cycling, but vibration or thermal cycling probably would break the particle loose. Remember also that the defect may be observable only during stimulation and not during a static bench test.

The levels of stress imposed on a product during ESS should be greater than the stress to which the product will be subjected during its operational lifetime, but still be below the maximum design parameters. This rule of thumb is pushed to the limits under an *enhanced screening* process. Enhanced screening places the component or system at well above the expected field environmental levels. This process has been found to be useful and cost effective for many programs and products. Enhanced screening, however, requires the development of screens that are carefully identified during product design and development so that the product can survive the qualification tests. Enhanced screening techniques often are required for cost-effective products on a cradle-to-grave basis; that is, early design changes for screenability save tremendous costs over the lifetime of the product.

The types of products that can be checked economically through ESS break down into two categories: high-dollar items and mass-produced items. Units that are physically large in size, such as RF generators, are difficult to test in the finished state. Still, qualification tests using more primitive methods, such as cam-driven truck-bed shakers, are practical. Because most large systems generate a large amount of heat, subjecting the equipment to temperature extremes also may be accomplished. Sophisticated ESS for large systems, however, must rely on qualification testing at the subassembly stage.

The basic hardware complement for an ESS test station includes a thermal chamber shaker and controller/monitor. A typical test sequence includes 10 min of exposure to random vibration, followed by 10 cycles between temperature minimum and maximum. To save time, the two tests may be performed simultaneously.

10.3.3 Latent Defects

The cumulative failure rate observed during the early life of an electronic system is dominated by the latent defect content of the product, not its inherent failure rate. Product design is the major determinant of inherent failure rate. A product design will show a higher-than-expected inherent rate if the system contains components that are marginally overstressed, have inadequate functional margin, or contain a subpopulation of components that exhibit a wear-out life shorter than the useful life of the product. Industry has grown to expect the high instantaneous failure rate observed when a

new product is placed into service. The burn-in process, whether ESS or conventional, is aimed at shielding customers from the detrimental effects of infant mortality. The key to reducing early-product-life failures lies in reducing the number of latent defects.

A latent defect is some abnormal characteristic of the product or its parts that is likely to result in failure at some point, depending on

- The degree of abnormality
- The magnitude of applied stress
- The duration of applied stress

For example, consider a solder joint on the connecting pin of a vacuum tube device. One characteristic of the joint is the degree to which the pin hole is filled with solder, characterized as “percent fill.” All other characteristics being acceptable, a joint that is 100% filled offers the maximum mechanical strength, minimum resistance, and greatest current carrying capacity. Conversely, a joint that is 0% filled has no mechanical strength, and only if the lead is touching the barrel does it have any significant electrical properties. Between these two extreme cases are degrees of abnormality. For a fixed magnitude of applied stress

- A grossly abnormal solder joint probably will fail in a short time.
- A moderately abnormal solder joint probably will fail, but after a longer period of time than a grossly abnormal joint.
- A mildly abnormal solder joint probably will fail, but after a much longer period of time than in either of the preceding conditions.

Figure 10.9 illustrates this concept. A similar time–stress relationship holds for a fixed degree of abnormality and variable applied stress.

A latent defect eventually will advance to a *patent defect* when exposed to environmental, or other, stimuli. A patent defect is a flaw that has advanced to the point at which an abnormality actually exists. To carry on the solder example, a cold solder joint represents a flaw (latent defect). After vibration and/or thermal cycling, the joint (it is assumed) will crack. The joint will now have become a detectable (patent) defect. Some latent defects can be stimulated into patent defects by thermal cycling, some by vibration, and some by voltage cycling. Not all flaws respond to all stimuli.

There is strong correlation between the total number of physical and functional defects found per unit of product during the manufacturing process, and the average latent defect content of shipping product. Product- and process-design changes aimed at reducing latent defects not only improve the reliability of shipping product, but also result in substantial manufacturing cost savings.

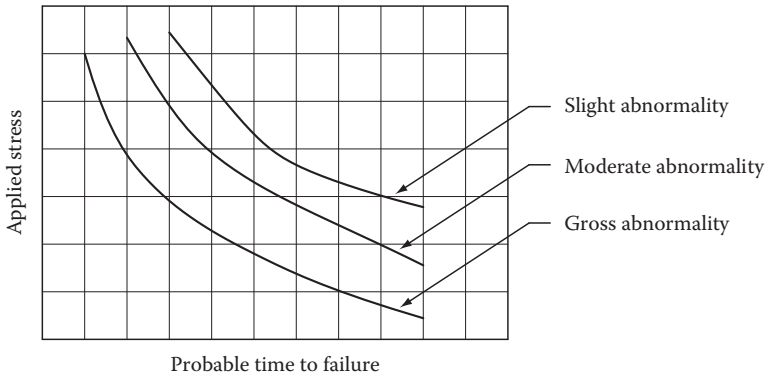


FIGURE 10.9

Estimation of the probable time to failure from an abnormal solder joint. (After Smith, W.B., *Integrated product and process design to achieve high reliability in both early and useful life of the product*, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.)

10.3.4 Operating Environment

The operating environment of an electronic system, either because of external environmental conditions or unintentional component underrating, may be significantly more stressful than the system manufacturer or the component supplier anticipated. Unintentional component underrating represents a design fault, but unexpected environmental conditions are possible for many applications, particularly in remote locations.

Conditions of extreme low or high temperatures, high humidity, and vibration during transportation may have a significant impact on long-term reliability of the system. For example, it is possible—and more likely, probable—that the vibration stress of the truck ride to a remote transmitting site will represent the worst-case vibration exposure of the transmitter and all components within it during the lifetime of the product.

Manufacturers report that most of the significant vibration and shock problems for land-based products arise from the shipping and handling environment. Shipping tends to be an order of magnitude more severe than the operating environment with respect to vibration and shock. Early testing for these problems involved simulation of actual shipping and handling events, such as end-drops, truck trips, side impacts, and rolls over curbs and cobblestones. Although unsophisticated by today's standards, these tests are capable of improving product resistance to shipping-induced damage.

10.3.5 Failure Modes

Latent failures aside, the circuit elements most vulnerable to failure in any piece of electronic hardware are those exposed to the outside world. In most

systems, the greatest threat typically involves one or more of the following components or subsystems:

- The ac-to-dc power supply
- Sensitive signal-input circuitry
- High-power output stages and devices
- Circuitry operating into an unpredictable load or into a load that may be exposed to lightning and other transient effects (such as an antenna)

Derating of individual components is a key factor in improving the overall reliability of a given system. The goal of derating is the reduction of electrical, mechanical, thermal, and other environmental stresses on a component to decrease the degradation rate and prolong expected life. Through derating, the margin of safety between the operating stress level and the permissible stress level for a given part is increased. This adjustment provides added protection from system overstress, unforeseen during design.

10.3.6 Maintenance Considerations

The reliability and operating costs over the lifetime of an RF system can be affected significantly by the effectiveness of the preventive maintenance program designed and implemented by the engineering staff. In the case of a *critical-system* unit that must be operational continuously or during certain periods, maintenance can have a major impact—either positive or negative—on downtime.

The reliability of any electronic system may be compromised by an *enabling event phenomenon*. This is an event that does not cause a failure by itself, but sets up (or enables) a second event that can lead to failure of the system. Such a phenomenon is insidious because the enabling event may not be self-revealing. Examples include the following:

- A warning system that has failed or has been disabled for maintenance
- One or more controls that are set incorrectly, providing false read-outs for operations personnel
- Redundant hardware that is out of service for maintenance
- Remote metering that is out of calibration

10.3.6.1 Common-Mode Failure

A *common-mode failure* is one that can lead to the failure of all paths in a redundant configuration. In the design of redundant systems, therefore, it is

important to identify and eliminate sources of common-mode failures, or to increase their reliability to at least an order of magnitude above the reliability of the redundant system. Common-mode failure points in a high-power RF system include the following:

- Switching circuits that activate standby or redundant hardware
- Sensors that detect a hardware failure
- Indicators that alert personnel to a hardware failure
- Software that is common to all paths in a redundant system

The concept of software reliability in control and monitoring has limited meaning in that a good program will always run, and copies of a good program will always run. On the other hand, a program with one or more errors will always fail, and so will the copies, given the same input data. The reliability of software, unlike hardware, cannot be improved through redundancy if the software in the parallel path is identical to that in the primary path.

10.3.6.2 Spare Parts

The spare parts inventory is a key aspect of any successful equipment maintenance program. Having adequate replacement components on hand is important not only to correct equipment failures, but to identify those failures as well. Many parts—particularly in the high-voltage power supply and RF chain—are difficult to test under static conditions. The only reliable way to test the component may be to substitute one of known quality. If the system returns to normal operation, then the original component is defective. Substitution is also a valuable tool in troubleshooting intermittent failures caused by component breakdown under peak power conditions.

10.4 Vacuum Tube Reliability

The end user should keep an accurate record of performance for each tube at the facility. Shorter-than-normal tube life could point to a problem in the RF amplifier or related hardware. The average life that may be expected from a power tube is a function of many operational parameters, including

- Filament voltage
- Ambient operating temperature
- RF power output
- Operating frequency
- Operating efficiency

The best estimate of life expectancy for a given system at a particular location comes from on-site experience. As a general rule of thumb, however, at least 4000h of service can be expected from most power grid tubes. Klystrons and other microwave devices typically provide in excess of 15,000h. Possible causes of short tube life include the following:

- Improper power stage tuning
- Inaccurate panel meters or external wattmeter, resulting in more demand from the tube than is actually required
- Poor filament voltage regulation
- Insufficient cooling system airflow (or water flow)
- Improper stage neutralization

An examination of the data sheet for a given vacuum tube will show that a number of operating conditions are possible, depending upon the class of service required by the application. As long as the maximum ratings of the device are not exceeded, a wide choice of operating parameters is possible. When studying the characteristic curves of each tube, remember that they represent the performance of a *typical* device. All electronic products have some tolerance among devices of a single type. Operation of a given device in a particular system may be different than that specified on the data sheet. This effect is more pronounced at very high frequency (VHF) and above.

10.4.1 Thermal Cycling

Most power tube manufacturers recommend a warm-up period between the application of *filament-on* and *plate-on* commands; about 5 min is typical. As a general rule, the minimum warm-up time is 2 min. Some RF generators include a time-delay relay to prevent the application of a plate-on command until a predetermined warm-up cycle is completed.

Most manufacturers also specify a recommended cool-down period between the application of *plate-off* and *filament-off* commands. This cool down, generally about 10 min, is designed to prevent excessive temperatures on the power amplifier (PA) tube surfaces when the cooling air is shut off. Large vacuum tubes contain a significant mass of metal, which stores heat effectively. Unless cooling air is maintained at the base of the tube and through the anode cooling fins, excessive temperature rise can occur. Again, the result can be shortened tube life, or even catastrophic failure because of seal cracks caused by thermal stress.

Most tube manufacturers suggest that cooling air continue to be directed toward the tube base and anode cooling fins after filament voltage has been removed to further cool the device. Unfortunately, however, not all control circuits are configured to permit this mode of operation.

10.4.2 Tube-Changing Procedure

Plug-in power tubes must be seated firmly in their sockets, and the connections to the anodes of the tubes must be tight. Once in place, the device typically requires no maintenance during its normal operating lifetime.

Insertion of a replacement tube must be performed properly. Gently rock and slightly rotate the tube as it is being inserted into the socket. This will help prevent bending and breaking the fingerstock. Be certain to apply sufficient force to seat the tube completely into the socket. Never use a lever or other tool on the tube to set it in place. Manual pressure should be adequate.

An intermediate point is reached when the grid contact fingerstock slides up the tube sides and encounters the contact area. It is important that the tube is fully inserted in the socket beyond this initial point of resistance.

Some sockets have adjustable stops that are set so that the tube grid contacts rest in the middle of the contact area when the tube is fully inserted. This positioning can be checked by inserting, then removing, a new tube. The scratch marks on the grid contacts will indicate the position of the tube relative to the socket elements. Figure 10.10 shows the minute scoring in the base contact rings of a power grid tube. A well-maintained socket will score the tube contacts when the device is inserted. If all fingers do not make contact, more current will flow through fewer contact fingers, causing overheating and burning, as shown in Figure 10.11.

Poor contact between a tube and its socket will always lead to problems. If the socket connectors fail to provide the proper electrical contact with the cylindrical tube elements, concentration of RF charging currents will result, leading to local overheating that may be destructive. After the connector loses its spring action, the heating is aggravated, and damage to the tube is likely. Some socket designs include a wire-wound spring encircling the outer

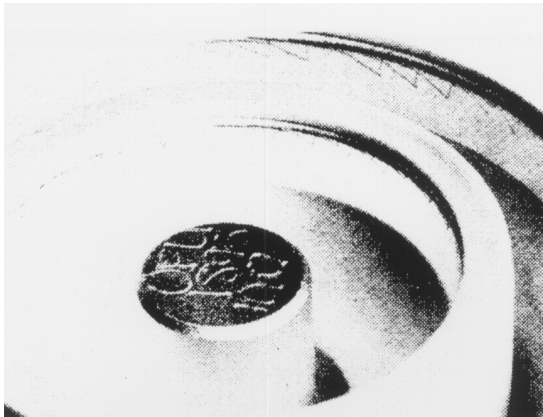


FIGURE 10.10

Proper fingerstock scoring of filament and control grid contacts on a power tube. (Courtesy of Varian, Palo Alto, CA.)

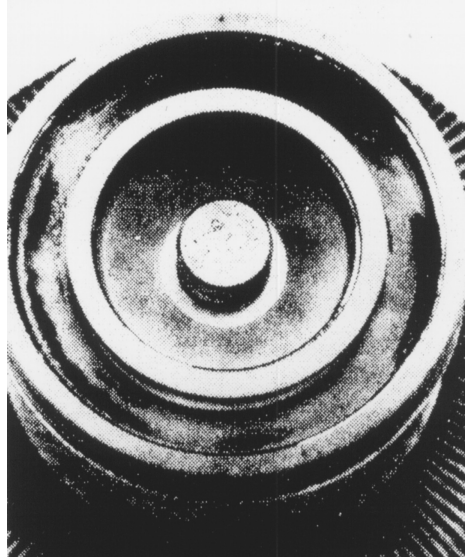


FIGURE 10.11

Damage to the grid contact of a power tube caused by insufficient fingerstock pressure. (Courtesy of Varian, Palo Alto, CA.)

circumference of the fingerstock to increase the contact pressure of the individual fingers. These springs should be replaced if they break or lose tension.

When a tube is removed from its socket, the socket should be blown or wiped clean and carefully inspected. Any discoloration caused by overheating of the socket fingerstock could signal impending tube/socket failure.

Many industrial and low-frequency tubes are not socketed, but installed by bolted or clamped connections. Stainless steel has a much lower coefficient of thermal expansion than copper. Therefore, clamped stainless steel anode connections should have some method of strain relief to prevent excess pressure from collapsing the tube anode as it heats.

All bolted or screwed connections should be tight. Verify that the clamps are snug and provide a good electrical contact around the entire circumference of the device. Because of the RF fields present, all clamps and bolts should be made from nonmagnetic materials. Copper, brass, and nonmagnetic series 300 stainless steel fasteners are preferred. Stainless steel is not a good conductor of electricity, so, although it may be used for clamping, it should not be part of the current path.

Before a new tube is installed in a system, it should be inspected for cracks or loose connections (in the case of tubes that do not socket-mount). Manufacturers typically recommend that after a new tube—or one that has been on the shelf for some time—is installed in the transmitter, it should be run with *filaments only* for at least 30 min, after which plate voltage may be applied. Next, the drive (modulation) is slowly brought up, in the case of an amplitude-modulated system. Residual gas inside the tube may cause an interelectrode arc (usually indicated as a plate overload) unless it is burned off in such a warm-up procedure.

10.4.3 Power Tube Conditioning

Large power tubes are subjected to rigorous processing during exhaust pumping at the time of manufacture [6]. Active elements are processed at temperatures several hundred degrees Celsius higher than those expected during normal operation. This procedure is designed to drive off surface and subsurface gas from the metals in the device to lower the possibility of these gases being released during the service life of the tube. Free gas molecules, however, always will be present to some degree in a fully processed tube. Gas, particularly oxygen-containing compounds, can combine chemically with cathode material to either permanently or temporarily destroy the electron emission capability of the cathode. Free gas molecules, when struck by electrons moving from the cathode to the anode, also may be ionized by having one or more electrons knocked from their systems. If enough of these ions, plus freed electrons, are present, a conduction path will be established that is not subject to control by the grid. This can result in runaway arcing that may involve all elements of the device. Current then will be limited primarily by the source voltage and impedance. (The space charge, to some degree, is neutralized by the presence of both electrons and positive ions.)

Electrons from sources other than the heated cathode provide low-current paths between elements when the voltage gradient is sufficiently high at the negative element for pure field emission. The voltage gradient at the negative element is governed by the following parameters:

- The applied voltage between elements
- The spacing between elements
- The surface contour of the negative element

A large voltage gradient can exist in the following areas:

- In front of a point on the negative element
- In front of a particle adhering to the negative element
- In front of a clump of gas molecules on the surface of the negative element

Field emission occurs readily from a cold surface if the operating conditions provide the necessary voltage gradient.

Ionization of free gas can result from bombardment by field-emitted electrons. Arcing is likely to occur as a result of field emission in operating equipment because the plate voltage is maximum during that part of the signal cycle when ordinary plate current from the cathode is shut off by the control grid. For this reason, an important part of tube processing is high-voltage conditioning to remove sharp points or small particles from tube elements. This part of tube processing may, and sometimes should,

be repeated in the field after shipment or storage if the tube is intended for use at a plate voltage in excess of 10 kV.

High-voltage conditioning (sometimes called *spot-knocking* or *debarnacling*) consists of applying successively higher voltages between tube elements, permitting the tube to spark internally at each voltage level until stable (no arcing). The voltage then is raised to the next higher potential until the tube is stable at a voltage approximately 15% higher than the peak signal voltage it will encounter in service.

The equipment used for tube conditioning is simple but specialized. It may provide dc or ac voltage, or both. The current required is small. The voltage should be continuously variable from practically zero to the highest value required for proper conditioning of the tube. The energy per spark is controlled by the internal resistance of the supply plus any external series resistor used. In dc conditioning, a milliammeter typically is included to measure the level of field emission current prior to sparking, or simply to determine whether the field emission current is within the specified range for a tube being tested. Also in dc conditioning, a capacitor may be placed across the tube under test to closely control the energy released for each spark. If the conditioning is to proceed effectively, the amount of energy delivered by the capacitor must be great enough to condition the surfaces, but not so large as to cause permanent damage.

10.4.3.1 Power Supply

The power supply is the heart of any high-voltage tube conditioning station [6]. Current-limiting resistance is desirable; the power supply may have an internal resistance on the order of 100 k Ω . In addition to the power supply, a current-limiting resistor of 100–300 k Ω should be provided in the circuit to the tube being conditioned. In dc processing, a storage capacitor of 1000–3000 pF can be placed directly across the device under test (DUT), with a high-voltage noninductive resistor of approximately 50 Ω connected between the tube and the capacitor. Naturally, this shunt capacitor must be capable of withstanding the full applied potential of the power supply.

Using both polarities on the tube elements is advantageous in high-voltage conditioning, suggesting that an ac supply might be the best choice for field tube conditioning. The advantage of the dc system, on the other hand, is that exact levels of field emission current and breakdown voltage may be determined for comparison with known “clean” tubes, and with the manufacturer’s specification. The polarity of the applied voltage may be reversed by mounting the DUT on an insulating frame so that the positive and negative high-voltage leads may be interchanged during processing. Also, the power supply rectifier and dc meters may be bypassed with suitable switching so that ac is obtained from the supply for initial tube cleanup. The storage capacitor across the tube should be disconnected if ac is used.

10.4.3.2 Conditioning Procedure

Manufacturers of large power tubes typically specify an upper limit for field emission between the filament, control grid, and screen grid connected together, and the anode [6]. Such devices should be allowed to spark until stable at the voltage where sparking first occurs. The voltage then is increased in steps determined by the voltage that initiates sparking again. This process is continued until the tube is stable at the maximum test voltage. It is advisable to reverse the polarity of the supply if dc testing is employed.

Usually, no field emission specifications exist for applied voltage between grids and between grid and filament, but these elements normally should hold off 10 kV internally. The need for conditioning between these elements is indicated if internal sparking occurs at a significantly lower voltage. Consult the tube manufacturer for specific recommendations on conditioning voltages.

During high-voltage processing, particularly between grids and between grid and filament, some of the redistributed gas molecules may be deposited on the cold filament, causing a temporary loss of emission. If this happens, operate the tube for approximately 1 h, with normal filament power only, to drive off the volatile material. Normal electron emission from the filament will be reestablished by this procedure. It is recommended that routine practice include operation of the filament for approximately 1 h as an immediate follow-up to high-voltage conditioning.

After conditioning, the tube should be stored in the same position (orientation) as that used for high-voltage conditioning and installation in the equipment. Movement of the tube after conditioning should be minimized to prevent any redistribution of particles within the device.

10.4.3.3 Considerations for Very Large Tubes

High-voltage ac conditioning is not recommended for very large tubes, such as the 8971/X2177, 8972/X2176, 8973/X2170, 8974/2159, 4CM400000A, and similar devices [6]. High-potential 50 or 60 Hz voltage can excite natural mechanical resonances in the large filament and grid structures of these tube types. Excitation of these resonances can break the filament and/or grids, thus destroying the device. If ac conditioning must be used, the following precautions should be taken:

- Connect all filament terminals together (external to the tube).
- Connect all terminals of the tube to a terminal of the high-voltage power supply. For example, if high voltage is to be applied between the screen and anode, the filament, grid, and screen should be connected together, then to one terminal of the supply. If high voltage is to be applied between the filament and grid, connect the grid, screen, and anode together and to one terminal of the supply.

No terminal of the tube should be allowed to float electrically during high-voltage conditioning.

- As the test voltage is increased, listen carefully for vibration inside the tube. If any is detected, immediately remove the ac high voltage.

These precautions will reduce the risk of damage from ac conditioning of very large tubes. However, because of the special procedures required and the risk of damage resulting from mechanical resonances, it is recommended that only dc conditioning be used for such devices.

10.4.3.4 Safety

Personnel protection dictates that the tube being processed, as well as the power supply, be enclosed in a metal cabinet with access doors interlocked so that no high voltage can be applied until the doors are closed [6]. The enclosure material, wall thickness, and dimensions must be chosen to limit x-ray radiation from the DUT to safe levels. A steel-wall thickness of 3/32 in. has been found adequate for dc testing up to approximately 70 kV. For conditioning up to 100 kV dc, it is necessary to add 0.05 in. of lead sheet as a lining to the steel enclosure.

The field of health standards regarding x-radiation dosage for a given period of time is complex and should be discussed with experts on the subject. Calibrated film badges or other reliable measuring devices normally are used to determine human exposure to x-rays over time.

10.4.4 Filament Voltage

A *true-reading* root-mean-square (rms) voltmeter is required for accurate measurement of filament voltage. The measurement is taken directly from the tube socket connections. A true-reading rms meter, instead of the more common *average responding* rms meter, is suggested because it can accurately measure a voltage despite an input waveform that is not a pure sine wave. Some filament voltage regulators use silicon controlled rectifiers (SCRs) to regulate the output voltage. Saturated core transformers, which often deliver nonsine waveforms, also are commonly used to regulate filament input power.

Long tube life requires filament voltage regulation. A tube whose filament voltage is allowed to vary along with the primary line voltage will not achieve the life expectancy possible with a tightly regulated supply. This problem is particularly acute at mountaintop installations, where utility regulation is generally poor.

To extend tube life, some end users leave the filaments on at all times, not shutting down at sign-off. If the sign-off period is 3h or less, this practice can be beneficial. Filament voltage regulation is a must in such situations

because the primary line voltages may vary substantially from the carrier-on to carrier-off value.

When a cold filament is turned on, damage can be caused by two effects [7]:

- Current inrush into the cold filament, which can be up to 10 times the normal operating current if the filament supply is of low impedance.
- Grain reorientation, which occurs at about 600°C–700°C, known as the *Miller–Larson effect*. This grain reorientation will result in a momentary plastic state that can cause the filament wire to grow in length and, therefore, become thinner.

The Miller–Larson effect is aggravated by variations in the cross-sectional area of the filament wires along their length. This will cause hot spots in the thinner sections because of the greater current densities there. The higher temperatures at the hot spots cause increased growth during warm-up through the 600°C temperature range, when they have a higher current density than the rest of the wire and a much higher power dissipation per unit length as a result of the higher resistance. Each time the filament is turned on, the wire becomes thinner until the hot spot temperature enters a runaway condition.

Filament voltage for a klystron should not be left on for a period of more than 2 h if no beam voltage is applied. The net rate of evaporation of emissive material from the cathode surface of a klystron is greater without beam voltage. Subsequent condensation of the material on gun components may lead to voltage hold-off problems and an increase in body current.

10.4.4.1 Black Heat

One practice used occasionally to extend the life of the thoriated-tungsten power tube filament is to operate the device at a low percentage of rated voltage during sign-off periods [7]. This technique, known as *black heat*, involves operating the filament at 25%–40% of the rated filament voltage. This range produces a temperature high enough to maintain the filament above the 600°C region where the Miller–Larson effect occurs, and sometimes low enough so that forced cooling is not required during standby.

Operating histories of several black heat installations indicate that the trauma of raising the filament voltage from 25% up to the operating voltage is still potentially damaging. For this reason, a higher standby voltage of about 80%–90% of the full rated value (*orange heat*) has been used with some success.

The voltage range of 40%–80% should be avoided because the filament can act as a getter within this range, and absorb gases that will raise the work function of the thoriated-tungsten filament. For orange heat operation, cooling system operation is required. Before attempting to implement a black

heat or orange heat filament management program, consult with the tube and transmitter manufacturers.

10.4.5 Filament Voltage Management

The thoriated-tungsten power grid tube lends itself to life extension through management of the applied filament voltage. Accurate adjustment of the filament voltage can extend the useful life of the device considerably, sometimes to twice the normal life expectancy. The following procedure typically is recommended:

- Install the tube, and tune for proper operation.
- Operate the filament at its full rated voltage for the first 200h following installation.
- Following the burn-in period, reduce the filament voltage by 0.1 V per step until system power output begins to fall (for frequency-modulated systems) or until modulating waveform distortion begins to increase (for amplitude-modulated systems).
- When the *emissions floor* has been reached, increase the filament voltage approximately 0.2 V peak-to-peak (P-P).

Long-term operation at this voltage can result in a substantial extension of the useful life of the tube, as illustrated in Figure 10.12.

The filament voltage should not, in any event, be operated at or below 90% of its rated value. At regular intervals, about every 3 months, the filament voltage is checked and increased if power output begins to fall or waveform distortion begins to rise. Filament voltage never should be increased to more than 105% of the rated voltage.

Note that some tube manufacturers place the minimum operating point at 94%. Others recommend that the tube be set for 100% filament voltage and left there. The choice is left to the user, after consultation with the manufacturer.

The filament current should be checked when the tube is installed, and at annual intervals thereafter, to ensure that the filament draws the expected current. Tubes can fail early in life because of an open filament bar—a problem that could be discovered during the warranty period if a current check is made upon installation.

For 1 week of each year of tube operation, the filament should be run at full rated voltage. This will operate the *getter* and clean the tube of gas.

10.4.5.1 Klystron Devices

Filament voltage is an equally important factor in achieving long life in a klystron. The voltage recommended by the manufacturer must be accurately set and checked on a regular basis. Attempts have been made to implement a filament

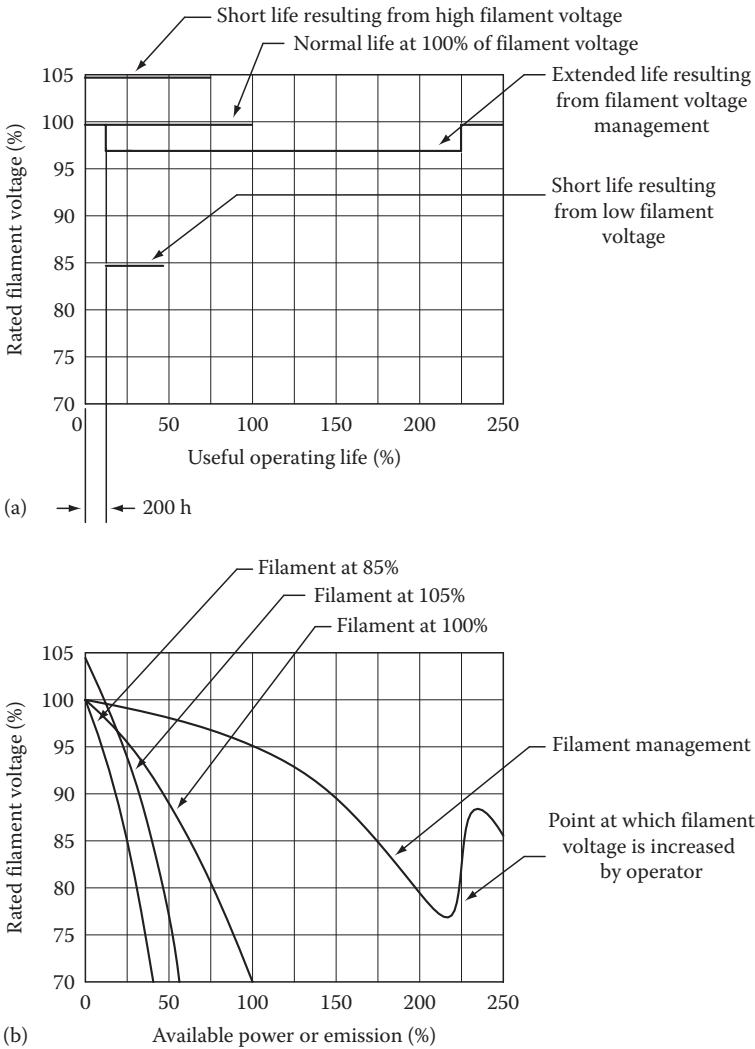


FIGURE 10.12

Effects of filament voltage management on a thoriated-tungsten filament power tube: (a) useful life as a function of filament voltage and (b) available power as a function of filament voltage. Note the dramatic increase in emission hours when filament voltage management is practiced.

voltage management program—as outlined previously for a thoriated-tungsten tube—on a klystron, with varying degrees of success. If filament management is attempted, the following guidelines should be observed:

- Install the tube, and tune for proper operation.
- Operate the filament at its full rated voltage for the first 200h following installation.

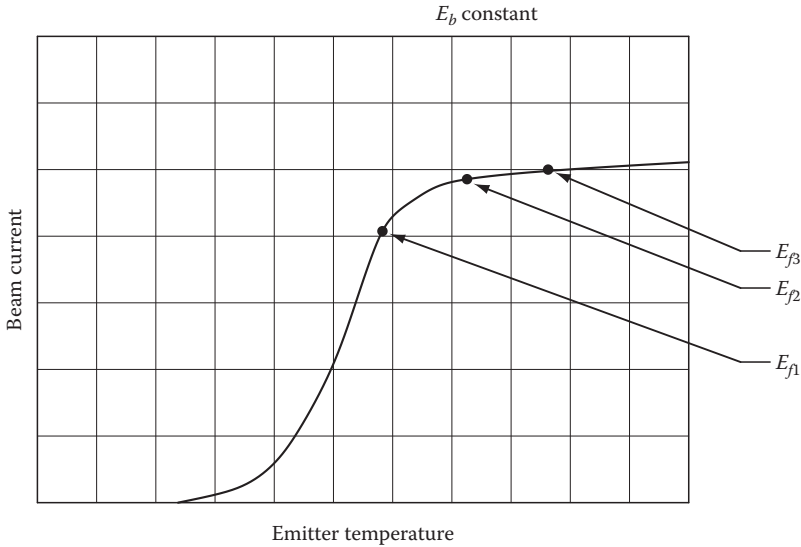


FIGURE 10.13
Effect of klystron emitter temperature on beam current.

- Following the burn-in period, reduce the filament voltage by 0.1 V per step until collector current begins to decrease. After each filament voltage change, wait at least 5 min to allow for cathode cooling.
- When the emissions floor has been reached, increase the filament voltage approximately 0.25 V P-P.

Figure 10.13 charts the relationship between emitter temperature and beam current. The filament should be adjusted to a point between E_{f1} and E_{f2} , which will vary with each klystron. Note that continued operation below the E_{f1} level, or at a point that causes emission instability, can result in beam defocusing, leading to increased modulating anode current and body current.

For operation at a reduced filament voltage, it may be necessary to allow a longer warm-up period when bringing the tube up from a cold start. If loss of emission is noted at turn-on, it will be necessary to either extend the warm-up time or raise the filament voltage to overcome the loss.

10.4.6 PA Stage Tuning

The PA stage of an RF generator may be tuned in several ways. Experience is the best teacher when it comes to adjusting for peak efficiency and performance. Compromises often must be made among various operating parameters.

Tuning can be affected by any number of changes in the PA stage. Replacing the final tube in a low- to medium-frequency RF generator usually does not

significantly alter stage tuning. It is advisable, however, to run through a touch-up tuning procedure just to be sure. Replacing a tube in a high-frequency RF generator, on the other hand, can significantly alter stage tuning. At high frequencies, normal tolerances and variations in tube construction result in changes in element capacitance and inductance. Likewise, replacing a component in the PA stage may cause tuning changes because of normal device tolerances.

Stability is one of the primary objectives of tuning. Adjust for broad peaks or dips, as required. Tune so that the system is stable from a cold start-up to normal operating temperature. Readings should not vary measurably after the first minute of operation.

Tuning is adjusted not only for peak efficiency, but also for peak performance. These two elements, unfortunately, do not always coincide. Trade-offs must sometimes be made to ensure proper operation of the system.

10.4.7 Fault Protection

An arc is a self-sustained discharge of electricity between electrodes with a voltage drop at the cathode on the order of the minimum ionizing potential of the environment [8]. The arc supports large currents by providing its own mechanism of electron emission from the negative terminal, plus space-charge neutralization.

All power vacuum tubes operate at voltages that can cause severe damage in the event of an internal arc unless the tube is properly protected. This damage can be relatively minor in tubes operating from low-energy power supplies, but it will become catastrophic in cases where a large amount of stored energy or follow-on current is involved. Because any high-voltage vacuum device may arc at one time or another, some means of protection is advised in all cases; it is mandatory in those instances where destructive quantities of energy are involved. In addition to protecting the tube in the event of a tube or circuit malfunction, such measures also protect the external circuitry.

Most power grid tubes employing thoriated-tungsten filaments are capable of withstanding relatively high energy arcs. However, while employing substantial materials, such devices usually incorporate fine wire grids that must be protected. In large tube applications, it is often found that the screen grid or bias supply, in addition to the anode supply, is capable of delivering destructive energy to the tube structure in the event of an internal arc. Consequently, protection must be provided in all cases where potentially destructive energy is involved. Large tubes usually are capable of withstanding up to 50 joules (J) total energy without permanent damage.

In the design of a PA, the stage should be tested to ensure that the 50J limit is not exceeded by any of the power supplies feeding the tube. Each supply, whether anode or grid, should be short circuit tested. A similar test should be made from plate-to-screen grid to ensure that the tube is protected in the

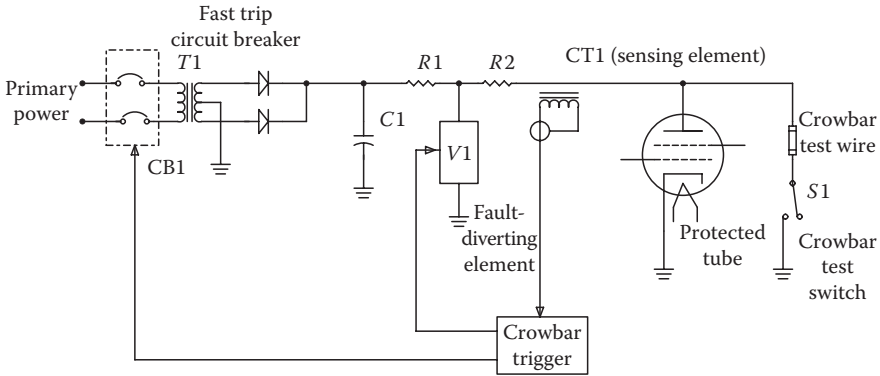


FIGURE 10.14
Arc protection circuit utilizing an energy-diverting element.

event of a plate-to-screen arc. A protective spark gap between the screen grid and cathode is recommended so that, if an arc occurs, the resulting spark from screen to cathode is external to the tube structure. The 50J test includes total energy delivered from the energy storage device and/or from the follow-on current from the power supply prior to the opening of the primary circuit.

Each power supply in the system must pass this test, or some form of energy-diverting or energy-limiting circuit must be employed. Figure 10.14 shows an energy-diverting circuit for a high-power tube. Energy diverters at high-power levels are complex. They typically employ large high-speed diverting elements such as thyratrons, ignitrons, triggered spark gaps, and, in some cases, a high-speed vacuum switch. These devices may be required to handle thousands of amperes of peak current. Resistor $R1$ in Figure 10.14 is used to limit the peak current into the energy-diverting switch $V1$, as well as to limit peak current demand on the transformer and rectifiers in the power supply. Resistor $R2$ must be large enough to keep the voltage across $V1$ from dropping so low under load-arc conditions as to inhibit the proper firing of the fault-diverting elements. $R1$ and $R2$ must also damp circuit ringing sufficiently to prevent current reversal in $V1$, which could cause $V1$ to open before the fault conditions are cleared. $R1$ and $R2$ are sized to minimize average power losses and to handle the required voltage and power transients while properly performing their functions. The sensing element $CT1$ detects any abnormal peak-current levels in the system. Signals from the current transformer are used to trigger $V1$ in the event of an overcurrent condition. A crowbar test switch, $S1$, can be used to perform testing, to ensure that the total energy dissipated in an arc is less than 50J. Typically, $S1$ is a remote-controlled high-voltage switch of either air or vacuum construction.

In a typical test, when $S1$ is closed, current transformer $CT1$ senses the fast rising current and fires the energy-diverting element $V1$; it simultaneously sends a fast trip signal to the main circuit breaker. Most energy

diverters fire within a few microseconds of the time the fault is detected; all voltages and currents are removed from the protected device in tens of microseconds.

By using this technique, it is possible to limit the energy delivered during an arc to much less than the recommended maximum. In a properly operating system, essentially no damage will occur within the tube even under severe arcing conditions.

10.4.8 Vacuum Tube Life

The vacuum tube suffers wear-out because of a predictable chemical reaction. Life expectancy is one of the most important factors to be considered in the use of vacuum tubes. In general, manufacturers specify maximum operating parameters for power grid tubes so that operation within the ratings will provide for a minimum useful life of 4000h. Considerable variation can be found with microwave power tubes, but 15,000h is probably a fair average useful life. It should be noted that some microwave devices, such as certain klystrons, have an expected useful life of 30,000h or more.

The cathode is the heart of a power tube. The device is said to *wear out* when filament emissions are inadequate for full power output or acceptable waveform distortion. In the case of the common thoriated-tungsten filament tube, three primary factors determine the number of hours a device will operate before reaching wear-out:

- The rate of evaporation of thorium from the cathode
- The quality of the tube vacuum
- The operating temperature of the filament

In the preparation of thoriated tungsten, 1%–2% of thorium oxide (thoria) is added to the tungsten powder before it is sintered and drawn into wire form. After being mounted in the tube, the filament usually is *carburized* by being heated to a temperature of about 2000K in a low-pressure atmosphere of hydrocarbon gas or vapor until its resistance increases by 10%–25%. This process allows the reduction of the thoria to metallic thorium. The life of the filament as an emitter is increased because the rate of evaporation of thorium from the carburized surface is several times smaller than from a surface of pure tungsten.

Despite the improved performance obtained by carburization of a thoriated-tungsten filament, the element is susceptible to deactivation by the action of positive ions. Although the deactivation process is negligible for anode voltages below a critical value, a trace of residual gas pressure too small to affect the emission from a pure tungsten filament can cause rapid deactivation of a thoriated-tungsten filament. This restriction places stringent requirements on vacuum processing of the tube.

These factors, taken together, determine the wear-out rate of the device. Catastrophic failures due to interelectrode short circuits or failure of the vacuum envelope are considered abnormal and are usually the result of some external influence. Catastrophic failures not the result of the operating environment usually are caused by a defect in the manufacturing process. Such failures generally occur early in the life of the component.

The design of the RF equipment can have a substantial impact on the life expectancy of a vacuum tube. Protection circuitry must remove applied voltages rapidly to prevent damage to the tube in the event of a failure external to the device. The filament turn-on circuit also can affect PA tube life expectancy. The surge current of the filament circuit must be maintained below a certain level to prevent thermal cycling problems. This consideration is particularly important in medium- and high-power PA tubes. When the heater voltage is applied to a cage-type cathode, the tungsten wires expand immediately because of their low thermal inertia. However, the cathode support, which is made of massive parts (relative to the tungsten wires), expands more slowly. The resulting differential expansion can cause permanent damage to the cathode wires. It also can cause a modification of the tube operating characteristics and, occasionally, arcs between the cathode and the control grid.

10.4.8.1 Catastrophic Failures

Catastrophic failures can be divided into two primary categories:

- The short-circuiting of broken or warped elements within the tube
- A loss of vacuum in the device

Air in the tube causes a loss of dielectric standoff between the internal elements. It is not possible to distinguish between an interelectrode short circuit and a loss of vacuum in the operating equipment.

Catastrophic failures that occur during initial installation are usually the result of broken elements. Failures that occur after an initial burn-in period are more likely to be caused by a loss of vacuum. In either case, continued efforts to operate the tube can result in considerable damage to the device and to other components in the system.

Note that a loss of vacuum can result from any of several anomalies. The two most common are

- A latent manufacturing defect in one of the ceramic-to-metal seals
- Circulating currents resulting from parasitic oscillations and/or improper tuning

Intermittent overloads are the most difficult to identify. They can be caused by circuit-operating conditions or internal tube failures. A broken or warped

filament element, for example, can move around inside the device and occasionally short circuit to the grid, causing a loss of grid bias and subsequent plate overload. Intermittent overloads also can be caused by a short circuit or high voltage standing wave ratio (VSWR) across the load.

10.4.9 Examining Tube Performance

Examination of a power tube after it has been removed from a transmitter or other type of RF generator can reveal a great deal about how well the equipment-tube combination is working. Contrast the appearance of a new power tube, shown in Figure 10.15, with a component at the end of its useful life. If a power tube fails prematurely, the device should be examined to determine whether an abnormal operating condition exists within the transmitter. Consider the following examples:

- Figure 10.16. Two 4CX15000A power tubes with differing anode heat-dissipation patterns. The tube shown in Figure 10.16a experienced excessive heating because of a lack of PA compartment cooling air or excessive dissipation because of poor tuning. The tube shown in Figure 10.16b shows a normal thermal pattern for a silver-plated 4CX15000A. Nickel-plated tubes do not show signs of heating because of the high heat resistance of nickel.
- Figure 10.17. Base-heating patterns on two 4CX15000A tubes. The tube shown in Figure 10.17a shows evidence of excessive heating because of high filament voltage or lack of cooling air directed

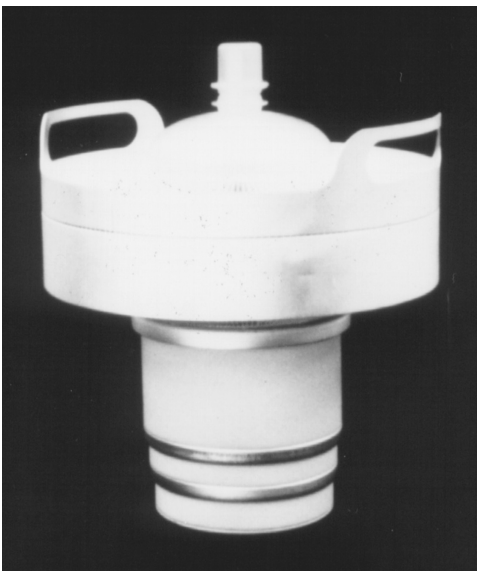
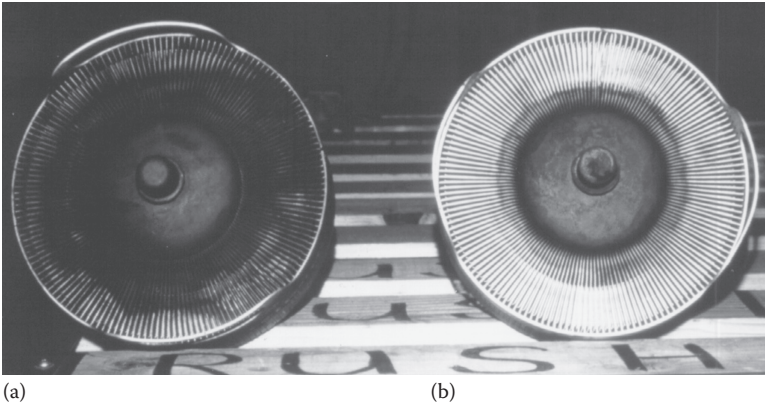
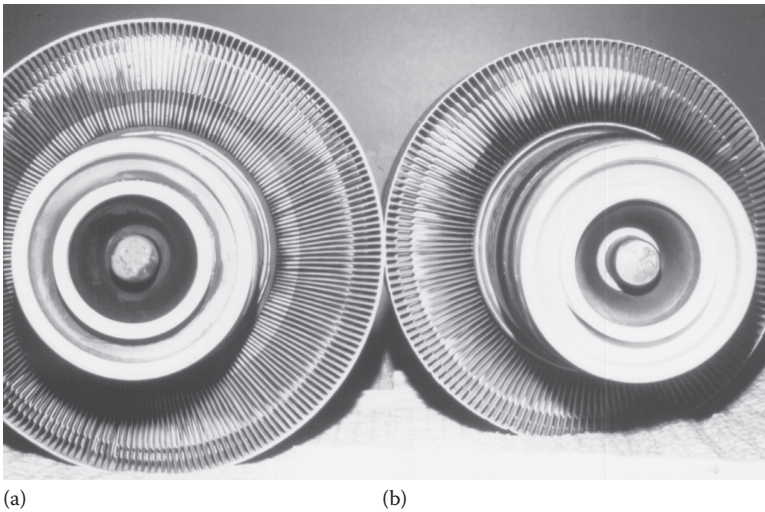


FIGURE 10.15

New, unused 4CX15000A tube. Contrast the appearance of this device with the tubes shown in the following figures. (Courtesy of Varian, Palo Alto, CA.)

**FIGURE 10.16**

Anode dissipation patterns on two 4CX15000A tubes: (a) excessive heating and (b) normal wear. (Courtesy of Econco Broadcast Service, Woodland, CA.)

**FIGURE 10.17**

Base heating patterns on two 4CX15000A tubes: (a) excessive heating and (b) normal wear. (Courtesy of Econco, Woodland, CA.)

toward the base of the device. The tube shown in Figure 10.17b shows a typical heating pattern with normal filament voltage.

- Figure 10.18. A 4CX5000A tube with burning on the screen-to-anode ceramic. Exterior arcing of this type generally indicates a socketing problem or another condition external to the tube.
- Figure 10.19. The stem portion of a 4CX15000A tube that had gone down to air while the filament was on. Note the deposits of tungsten oxide formed when the filament burned up. The grids are burned

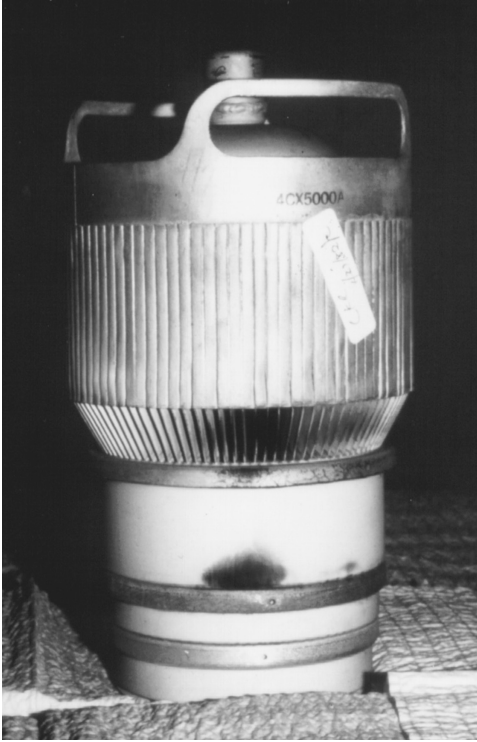


FIGURE 10.18
4CX5000A tube that appears to have suffered socketing problems. (Courtesy of Econco, Woodland, CA.)

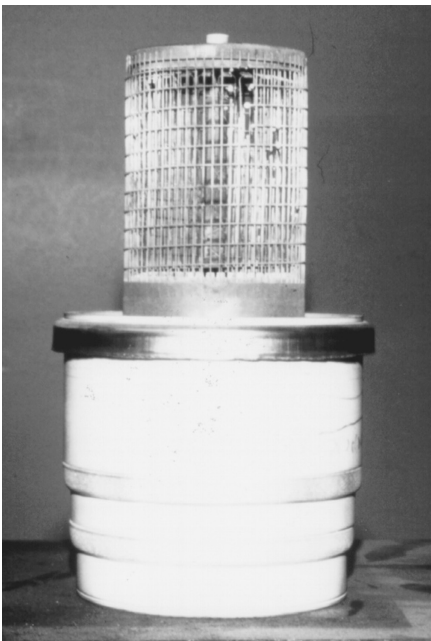


FIGURE 10.19
Interior elements of a 4CX15000A tube that had gone to air while the filament was lit. (Courtesy of Econco, Woodland, CA.)



FIGURE 10.20

4CX15000A tube showing signs of external arcing. (Courtesy of Econco, Woodland, CA.)

and melted because of the ionization arcs that subsequently occurred. A failure of this type will trip overload breakers in the RF generator. It is indistinguishable from a short-circuited tube in operation.

- Figure 10.20. A 4CX15000A tube that experienced arcing typical of a bent fingerstock, or exterior arcing caused by components other than the tube.

Note that the previous examples apply to silver-plated power grid tubes.

The external metal elements of a tube typically are plated with nickel or silver. Tubes that go into sockets are usually silver plated. The soft silver provides a better contact interface than the much harder nickel, and it deforms slightly under contact pressure, providing greater contact area. Silver plating has a dull, whitish cast, whereas nickel has a hard, metallic appearance.

Nickel is resistant to discoloration due to heat at normal tube operating temperatures, but silver will tarnish easily. As demonstrated in this section, the heat patterns on silver-plated tubes are helpful in problem analysis. If a nickel-plated tube shows signs of heat discoloration, a significant cooling or operating problem exists. Nickel will not discolor until it reaches a temperature much higher than a tube normally achieves under typical operating conditions.

10.4.10 Shipping and Handling Vacuum Tubes

Because of their fragile nature, vacuum tubes are packaged for shipment in foam-filled or spring-supported shipping containers. When it is necessary to transport a tube from one location to another, use the original packing material and box or crate. The tube should be removed whenever the RF equipment is relocated. Never leave a tube installed in its socket during an equipment move.

During installation or service, always keep the tube in its typical operating position (most commonly horizontal). Never allow a tube to roll along a surface. Damage to the filament or to other elements in the device may result.

It is occasionally necessary for an RF generator to be moved from one location to another within a plant. Equipment used in this manner must be equipped with air-filled casters; do not use solid casters.

When it is necessary or desirable to store a tube for an extended period of time, place it in a plastic bag to protect it from moisture, and return the device to its shipping box. Maintenance engineers are discouraged from placing markings on a power tube. Some engineers note service dates using ink or pencil on tubes. This practice is not recommended. Instead, attach a separate service record to the tube using a tie-wrap or similar removable device. Do not use self-adhesive labels for this purpose. When removed, adhesive labels usually leave some residue on the attachment surface.

Although it is not recommended that a power tube be stored indefinitely, most modern devices utilizing metal and ceramic envelopes can be stored for long periods of time without deterioration. It usually is not necessary to rotate a spare tube through an operating socket to degas the device.

Older vacuum tubes utilizing glass as the insulating medium tend to leak gas over time. The glass is not the problem, but, instead, the Kovar alloy used to seal the glass-to-metal parts of the tube. Kovar also is subject to rusting when moisture is present. Glass-insulated vacuum tubes should be sealed in a plastic bag for storage and rotated through the equipment at least once every 12 months. The larger the physical size of the tube, the greater the problem of gassing.

10.5 Klystron Reliability

Klystron amplifiers are well known as reliable, long-life microwave devices that provide the user thousands of hours of trouble-free service. Many of the guidelines discussed previously for power vacuum tubes in general also apply to klystrons and related microwave devices. However, several considerations are specific to the klystron, including

- Cooling system maintenance
- Ceramic element cleanings
- Gun element reconditioning
- Focusing electromagnet maintenance
- Protection during prime power interruption
- Shutoff precautions

10.5.1 Cleaning and Flushing the Cooling System

One of the major problems encountered in field operation of a klystron cooling system is the formation of corrosion products [9]. Corrosion frequently occurs in the cooling channels of the klystron body on the surface of the water/vapor-cooled collector, and in the cooling channels of the electromagnet. It is the result of a chemical reaction between free oxygen-laden coolant and the hot copper channel wall. Corrosion, as described here, is different from *scale formation*, which can be caused by the use of a coolant other than clean distilled water, or may be the by-product of electrolysis between dissimilar metals in the heat-exchange system. Scale formation can be controlled through the use of online purification loops.

New transmitter water lines frequently contain contaminants, including

- Solder
- Soldering or brazing flux
- Oils
- Metal chips or burrs
- Pieces of Teflon sealing tape

When the water lines are installed, these contaminants must be flushed and cleaned from the system before the klystron and magnet are connected.

The following guidelines are suggested for cleaning the cooling passages of water- and vapor-cooled klystrons and associated RF generation equipment. The procedures should be performed before a klystron is installed, then repeated on a periodic basis while the tube is in service.

10.5.1.1 Transmitter Flushing

Proper cleaning of the circulating water system in an RF generator is vital to long-term reliability. Although each design is unique, the following general procedures are appropriate for most installations [9]:

- Disconnect the tube and the magnet.
- Add jumper hoses between the input and output of the klystron and the electromagnet water lines.
- Disconnect or bypass the pump motor.
- Fill the system with hot tap water, if available.
- Open the drain in the transmitter cabinet and flush for 15 min or until it is clean.
- Flush the water lines between the tank and pump separately with hot tap water, if available, until they are clean.

- Connect all water lines, and fill the system with hot tap water, if available, and one cup of nonsudsing detergent. Trisodium phosphate is recommended.
- Operate the water system with hot tap water for 30 min. An immersion heater may be used to maintain hot water in the system.
- Drain the system, and flush with hot tap water for 30 min.
- Remove and clean the filter element.
- Refill the water system with tap water (at ambient temperature).
- Operate the water system. Maintain the water level while draining and flushing the system until no detergent, foam, or foreign objects or particles are visible in the drained or filtered element. To test for detergents in the water, drain a sample of water into a small glass test tube, and allow it to stand for 5 min. To generate foam, shake the test tube vigorously for 15 s and then allow it to stand 15 s. A completely foam-free surface indicates no foam-producing impurities.
- Repeat the two previous steps if detergent foam is still present.
- Drain the system and refill with distilled water when the tube and transmitter water lines are clean.
- Remove, clean, and replace the filter before using.

10.5.1.2 Flushing Klystron Water Lines

Although most klystrons are shipped from the manufacturer with water passages clean and dry, it is good engineering practice to flush all cooling passages before installing the tube. Occasionally, tubes that have been in service for some time develop scale on the collector and also must be flushed. Contaminated water further contributes to dirty water lines, requiring flushing to clear. The following general procedure is recommended for tubes with contaminated water lines, corrosion, scale, or blocked passages [9]:

- Remove the input water fitting (Hansen type), and add a straight pipe extension, approximately 1–2 in long, to the tube.
- Attach a hose to this fitting, using a hose clamp to secure it. This connection will serve as the drain line; it should be discharged into a convenient outlet.
- On some tube types, the normal body-cooling output line is fed to the base of the vapor-phase boiler. Remove the hose at the base of the boiler. Do not damage this fitting; it must be reused and must seal tightly.
- Attach a 2–3 in. long extension pipe that will fit the small hose at one end and a garden hose at the other end. Secure the extension with hose clamps.

- Connect the garden hose to a tap water faucet. Hot tap water is preferable, if available.
- Back-flush the klystron-body cooling passages for 10–15 min at full pressure until they are clean.
- Reconnect the input and output water lines to the klystron.
- If scale is present on the vapor-phase collector, use a solution of trisodium phosphate for the first cleaning. Perform this step after cleaning the transmitter water system (as described previously).
- Connect the klystron to the transmitter water lines, and fill the system with tap water (hot, if available) and one cup of nonsudsing detergent.
- Operate the water system for 15 min. Make certain that the water level completely covers the collector core.
- Drain the system, and flush for 30 min or until no detergent foam is present.
- Remove and clean the filter element.
- Fill the system with distilled water.
- Drain the system, and refill with distilled water.

10.5.1.3 Cleaning Klystron Water Lines

If heavy scaling has formed on the vapor-phase collector and/or blocked water passages, they may be cleaned using a stronger cleaning agent. A solution of *Ty-Sol** may be used to remove scale and corrosion. The following procedure is recommended [9]:

- Clean the transmitter water system (as described previously).
- Connect the tube to the transmitter water system.
- Fill the system with hot tap water (if available) and add 2 gal of cleaning solution for every 50 gal of water.
- Operate the water system for 15 min or until scale has been removed and the collector has a clean copper color.
- Make certain that the water level covers the top of the collector during cleaning. Evidence of cleaning action may be observed in the flow meters.
- Drain the system, and flush with tap water.
- Remove and clean the filter element, then replace.
- Refill the system with distilled water and flush for 30 min to 1 h.

* Chemical Research Association, Chicago, IL.

- Check for detergent foam at the end of the flushing period; also check the pH factor. Continue flushing with distilled water if foam is present or if the pH factor is not within the specified range.
- Refill the system with clean distilled water.

Chlorine present in common tap water is harmful to the water passages of the klystron. Thorough flushing with distilled water will remove all chlorine traces. Never use tap water for final refill or for makeup water.

10.5.1.4 Flushing and Cleaning Magnet Water Lines

The magnet water lines may be flushed and cleaned in the same manner as described for the klystron.

10.5.1.5 General Cleaning

Two remaining items must be clean to achieve efficient operation: the sight glass and float of the water-flow indicators. The water-flow indicators often become contaminated during use. This contamination collects on the sight glass and on the float, making flow-reading difficult. If too much contamination is present on the glass and float, sticking or erroneous readings may result. The detergent and cleaning solutions may not remove all of this contamination. If this is the case, remove and clean the flow meter using a brush to clean the glass surface.

10.5.2 Cleaning Ceramic Elements

Foreign deposits on the ceramic elements of a klystron should be removed periodically to prevent ceramic heating or arc paths that can lead to klystron failure. The recommended procedure is as follows:

- Remove the klystron from the RF generator, and mount it to a stand or other suitable fixture.
- Remove loose debris with a soft-fiber bristle brush.
- Clean the ceramic with a mild abrasive cleanser, preferably one that includes a bleaching agent. (Household cleansers such as Comet and Ajax are usually acceptable.) While cleaning, do not apply excessive force to the ceramic. Normal hand pressure is quite sufficient.
- Flush the ceramic with clean water to remove the cleanser.
- Flush a second time with clean isopropyl alcohol.
- Air dry the device. Be certain that all moisture has evaporated from the tube before returning it to service.

10.5.3 Reconditioning Klystron Gun Elements

During the operational life of a klystron, voltage standoff problems may be experienced when the operating mode of the device is changed from one collector voltage to a higher voltage or from continuous to pulsed service [10]. In both of these cases, the difference of potential between the internal elements is increased over values used in the previous mode of operation. In some instances, this elevated difference of potential can lead to internal arcing as a result of long-term buildup of electrical leakage between elements. This is a part of the normal aging process in klystrons. In many cases, however, this particular process can be reversed through high-voltage conditioning. Maintenance engineers have found it useful to clear any residual leakage between the various elements before pulsing an older klystron, or when increasing the operating potentials applied to an older klystron.

Klystron electron gun leakage can be checked, and, in many cases, the gun elements reconditioned, by applying a high dc potential between these elements. The process is termed *hi-potting*. It requires a current-limited (5 mA maximum), continuously variable dc supply with a range of 0–30 kV. The procedure is as follows:

- Turn off the heater, and allow the cathode sufficient time to cool completely (approximately 1–2 h).
- Remove all electrical connections from the klystron, except for the body and collector grounds.
- Connect the negative high-voltage output of the hi-potter to the heater-cathode terminal, and connect the positive ground lead to the mod-anode as shown in Figure 10.21. Increase the voltage output gradually, noting the leakage current.
- If high levels of current are observed, reduce the voltage and allow it to stand at the 1–2 mA level until the leakage current reduces (burns off).
- Continue to raise the hi-potter output until the voltage applied is equal to the normal operating value for the klystron being checked. The typical leakage level is 200–400 μA .
- Check the mod-anode-to-body ground leakage; reduce, if necessary, in the same fashion. For this purpose, the negative high-voltage output is connected to the mod-anode terminal and the positive ground lead to the klystron body.

The leakage level indicated as optimum for this procedure (200–400 μA) is a typical value. Tubes still should operate normally above this level, but they may be more susceptible to internal arcing.

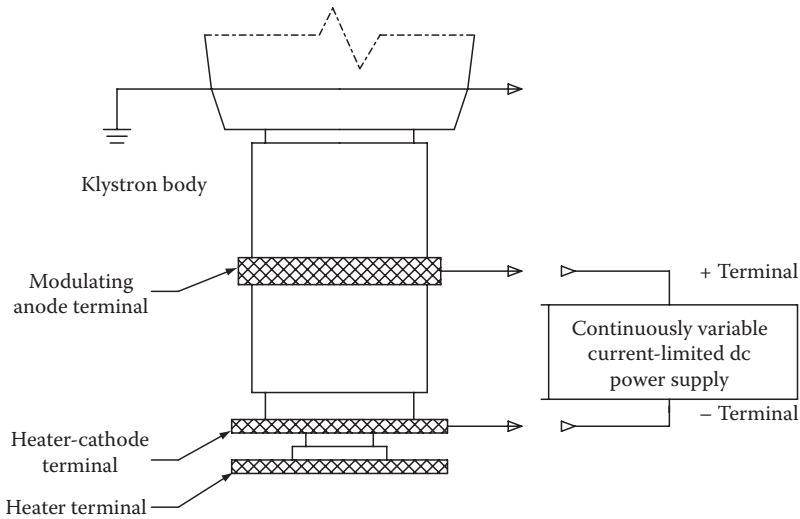


FIGURE 10.21

Electrical connections for reconditioning the gun elements of a klystron.

10.5.4 Focusing Electromagnet Maintenance

The focusing electromagnet of a klystron, given sufficient preventive maintenance, will provide many years of service [11]. Maintenance personnel can extend the lifetime of the device by reducing the operating temperature of the electromagnet windings and by minimizing thermal cycling. Field experience has shown that the major failure modes for a focusing electromagnet are short-circuited turns in the coil windings and/or short circuits from the windings to the coolant tubes.

Short-circuited turns result in lower magnetic flux density available to focus the klystron beam and are indicated by higher klystron body current. To counteract this effect, the transmitter operator often will increase the electromagnet current to a point above the rating necessary to obtain the magnetic field strength needed for rated klystron body current. This increases the heat dissipation in the device, which may lead to further problems. Such failures are the result of insulation breakdown between turns in the windings, and/or between the windings and the coolant tubes. Insulation breakdown usually is caused by excessive operating temperatures in the individual windings. Excessive temperature is the result of one or more of the following operational conditions:

- Excessive inlet coolant temperature
- Insufficient coolant flow
- Excessive ambient temperature
- Excessive operating current

Insulation breakdown also may result from excessive thermal cycling of the coil windings. Abrasion of internal insulation may occur during expansion and contraction of the winding assembly. Generally speaking, the life of a focusing electromagnet can be doubled for every 10°C drop in operating temperature (up to a point).

To minimize the effects of thermal cycling, the electromagnet temperature should be maintained at or below its operating value during standby periods by running at rated or reduced current, and at rated coolant flow. If neither running nor reduced current is available during standby conditions, heated coolant can be circulated through the electromagnet.

10.5.5 Power Control Considerations

Under certain conditions, high-voltage transients that exceed the specified maximum ratings of a klystron can occur in an RF generator. These transients can cause high-voltage arcing external to the tube across the klystron modulating anode and cathode seals. In some cases, such arcs can puncture sealing rings, letting the tube down to air. High-voltage transient arcing is not common to all installations, but when it does occur, it can result in a costly klystron failure. Overvoltage transients in power circuits can be traced to any of several causes, including lightning.

Collection of moisture on the tube is another potential cause of arc-induced failure. Integral-cavity klystrons, which are installed with the cathode down, may collect moisture in the high-voltage (cathode) area because of water leakage or condensation, leading to external arcing and loss of vacuum.

To provide a measure of protection against these types of failures, arc shields may be used on some tubes to protect the gun components that are the most likely to be damaged by arcs. The shield typically consists of an aluminum cover fitted on a portion of the tube. Although arc shields provide considerable protection, they cannot prevent all failures; the magnitude of a potential arc is an unknown quantity. It must be understood that an arc shield does not eliminate the cause of arcing, which should be addressed separately by the maintenance engineer. It is, instead, intended as a stop-gap measure to prevent destruction of the tube until the root cause of the arcing is eliminated.

Installation of a gas-filled spark-gap transient voltage suppressor also can serve to protect the klystron in the event of an arc or other transient discharge. The rating of the suppressor is determined by the type of tube being protected. The device is connected between ground and the cathode, installed as close as possible to the cathode. The suppressor will prevent the application of transient voltages that exceed the standoff rating of the klystron ceramic.

Power-line suppression components are useful for protecting transmission equipment from incoming line transients and other disturbances, but may not be effective in protecting against a system-induced arc within the transmitter or power supply that could damage or destroy the klystron. Therefore, protection of the klystron itself is recommended.

10.5.5.1 Primary Power Interruption

Unexpected interruptions of primary power, which occur seasonally during winter storms and summer lightning strikes, can damage or destroy a klystron unless adequate protection measures are taken [12]. There are three main factors involved in such failures:

- Mechanical and/or electrical malfunction of the main circuit breaker of the transmitter. This condition results when the breaker remains closed after the primary power is interrupted. With the loss of primary power, all excitation to the transmitter is removed. If the main circuit breaker remains closed when the power is restored, high voltage can be applied to the klystron before the magnet supply can provide adequate focusing. This allows full beam power dissipation in the body of the klystron, which can melt the drift tube.
- Loss of focusing during shutdown resulting from a power interruption at the transmitter in a system that does not employ magnet power supply filtering. When a loss of primary power occurs, the various power supplies start to discharge. The time constant of the associated filter circuit determines the rate of discharge. If there is no filter capacitance in the magnet supply, it will decay almost instantaneously while the beam filtering allows the high voltage to decay comparatively slowly. This permits defocusing of the electron beam and dissipation of the beam power in the drift tube region. Melting of the drift tube will result.
- Loss of one phase of the primary power supply. In some transmitters, the magnet supply and the protection circuits are fed from a common leg of the three-phase incoming power line. A loss of this phase will cause a loss of focusing and disable the protection circuits, including body-current protection. If the low-voltage-sensing relay in the main breaker is misadjusted or nonoperational, the high-voltage supply will remain on (although somewhat reduced) and, without focusing or protection, may melt the tube body.

To prevent these types of failures, the following steps are recommended:

- Conduct regular tests of the main circuit breaker in accordance with the manufacturer's instructions. Note that turning the breaker on and off using the transmitter control circuitry checks only half of the system. Tests also should be conducted by interrupting the primary feed to the breaker.
- Compute or measure the discharge time constant of the filtering circuits of both the beam and magnet supplies. Capacitance should be added to the magnet supply if the time constant is less than that of the beam supply. Be careful, however, not to increase the magnet

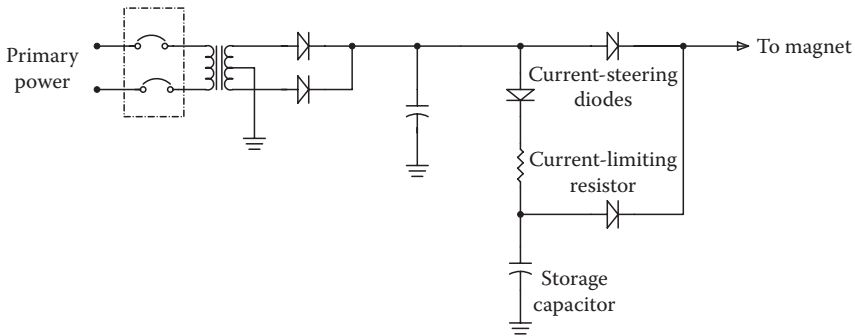


FIGURE 10.22

Diode-steering filtered power supply for a klystron focusing electromagnet.

supply filtering too much because it may also increase the ramp-up time of the magnet voltage. This could result in the beam supply being applied before the magnet supply is at its full potential. A simple diode-steering arrangement, such as the one shown in Figure 10.22, will prevent such problems in most cases.

- Consider installing a multipole vacuum relay at the high-voltage input power lines (between the main breaker and the high-power supply) to prevent instantaneous application of high voltage in the event of a malfunction of the main breaker. The primary or low-voltage circuit configuration of the vacuum relays should be such that it will prevent reenergizing of the relays before the magnetic field has built up when primary power is restored after a line loss.
- Install a solid-state phase-loss sensor if this protection is not already provided.
- Install an electromagnet undercurrent protection circuit that will remove the high voltage if the magnet current falls below a preset level.

An end user who is contemplating any of these changes should consult the original equipment manufacturer before proceeding.

References

1. Capitano, J. and J. Feinstein, Environmental stress screening demonstrates its value in the field, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.
2. Wong, K. L., Demonstrating reliability and reliability growth with environmental stress screening data, *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.

3. Tustin, W., Recipe for reliability: Shake and bake, *IEEE Spectrum*, IEEE, New York, December 1986.
4. Hobbs, G. K., Development of stress screens, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.
5. Smith, W. B., Integrated product and process design to achieve high reliability in both early and useful life of the product, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.
6. Conditioning of large radio-frequency power tubes, Application Bulletin no. 21, Varian/Eimac, San Carlos, CA, 1984.
7. Johnson, W., Techniques to extend the service life of high-power vacuum tubes, in *NAB Engineering Conference Proceedings*, National Association of Broadcasters, Washington, DC, pp. 285–293, 1994.
8. Fault protection, Application Bulletin no. 17, Varian/Eimac, San Carlos, CA, 1987.
9. Cleaning and flushing klystron water and vapor cooling systems, Application Engineering Bulletin no. AEB-32, Varian Associates, Palo Alto, CA, 1982.
10. Reconditioning of UHF-TV klystron electron gun elements, Technical Bulletin no. 4867, Varian Associates, Palo Alto, CA, July 1985.
11. UHF-TV klystron focusing electromagnets, Technical Bulletin no. 3122, Varian Associates, Palo Alto, CA, November 1973.
12. Recommendations for protection of 4KM series klystrons during prime power interruptions, Technical Bulletin no. 3542, Varian Associates, Palo Alto, CA.

Bibliography

- Anderson, L. R., *Electric Machines and Transformers*, Reston Publishing Company, Reston, VA.
- Brauer, D. C. and G. D. Brauer, Reliability-centered maintenance, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.
- Cameron, D. and R. Walker, Run-in strategy for electronic assemblies, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.
- Clark, R. J., Electronic packaging and interconnect technology working group report (IDA/OSD R&M study), IDA Record Document D-39, Institute of Defense Analysis, August 1983.
- Cleaning UHF-TV klystron ceramics, Technical Bulletin no. 4537, Varian Associates, Palo Alto, CA, May 1982.
- DesPlas, E., Reliability in the manufacturing cycle, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.
- Devaney, J., Piece parts ESS in lieu of destructive physical analysis, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.
- Dick, B., Caring for high-power tubes, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, pp. 48–58, November 1991.
- Doyle, E., Jr., How parts fail, *IEEE Spectrum*, IEEE, New York, October 1981.

- Extending transmitter tube life, Application Bulletin no. 18, Varian/Eimac, San Carlos, CA, 1990.
- 55kW integral cavity klystron arc shields, Technical Bulletin no. 4441, Varian Associates, Palo Alto, CA, April 1981.
- Fink, D. and D. Christiansen (eds.), *Electronics Engineers' Handbook*, 3rd edn., McGraw-Hill, New York, 1989.
- Fortna, H., R. Zavada, and T. Warren, An integrated analytic approach for reliability improvement, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.
- Frey, R., G. Ratchford, and B. Wendling, Vibration and shock testing for computers, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.
- Griffin, P., Analysis of the F/A-18 hornet flight control computer field mean time between failure, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, IEEE, New York, 1985.
- Hansen, M. D. and R. L. Watts, Software system safety and reliability, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1988.
- Hermason, S. E., Major, USAF. Letter dated December 2, 1988. From: Yates, W. and D. Shaller, Reliability engineering as applied to software, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Los Angeles, CA, IEEE, New York, 1990.
- High Power Transmitting Tubes for Broadcasting and Research*, Philips Technical Publication, Eindhoven, the Netherlands, 1988.
- Horn, R. and F. Hall, Maintenance centered reliability, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, IEEE, New York, 1983.
- Irland, E. A., Assuring quality and reliability of complex electronic systems: Hardware and software, *Proceedings of the IEEE*, 76(1), 5-18, 1988.
- Jordan, E. C. (ed.), *Reference Data for Engineers: Radio, Electronics, Computers, and Communications*, 7th edn., Howard W. Sams, Indianapolis, IN, 1985.
- Kenett, R. and M. Pollak, A semi-parametric approach to testing for reliability growth, with application to software systems, in *IEEE Transactions on Reliability*, IEEE, New York, August 1986.
- Laboratory Staff, *The Care and Feeding of Power Grid Tubes*, Varian EIMAC, San Carlos, CA, 1984.
- Maynard, E., OUSDRE working group, VHSIC technology working group report (IDA/OSD R&M study), Document D-42, Institute of Defense Analysis, November 1983.
- Meeldijk, V., Why do components fail? *Electronic Servicing and Technology*, Intertec Publishing, Overland Park, KS, November 1986.
- Nenoff, L., Effect of EMP hardening on system R&M parameters, *Proceedings of the 1986 Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.
- Neubauer, R. E. and W. C. Laird, Impact of new technology on repair, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.
- Pearman, R., *Power Electronics*, Reston Publishing Company, Reston, VA, 1980.
- Powell, R., Temperature cycling vs. steady-state burn-in, *Circuits Manufacturing*, Benwill Publishing, September 1976.

- Power grid tubes for radio broadcasting, Publication no. DTE-115, Thomson-CSF, Dover, NJ, 1986.
- Reduced heater voltage operation, Technical Bulletin no. 4863, Varian Associates, Palo Alto, CA, June 1985.
- Robinson, D. and S. Sauve, Analysis of failed parts on naval avionic systems, Report no. D180-22840-1, Boeing Company, October 1977.
- Spradlin, B. C., Reliability growth measurement applied to ESS, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Las Vegas, NE, IEEE, New York, 1986.
- Svet, F. A., Factors affecting on-air reliability of solid state transmitters, in *Proceedings of the SBE Broadcast Engineering Conference*, Society of Broadcast Engineers, Indianapolis, IN, October 1989.
- Technical Staff, Military/Aerospace Products Division, *The Reliability Handbook*, National Semiconductor, Santa Clara, CA, 1987.
- Tube Topics: A Guide for Vacuum Tube Users in a Transistorized World*, Econco Broadcast Service, Woodland, CA, 1990.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1990.
- Whitaker, J. C., *Maintaining Electronic Systems*, CRC Press, Boca Raton, FL, 1992.
- Wilson, S., What do you know about electronics, *Electronic Servicing and Technology*, Intertec Publishing, Overland Park, KS, September 1985.
- Wilson, M. F. and M. L. Woodruff, Economic benefits of parts quality knowledge, *Proceedings of the IEEE Reliability and Maintainability Symposium*, IEEE, New York, 1985.
- Wong, K. L., I. Quart, L. Kallis, and A. H. Burkhard, Culprits causing avionics equipment failures, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.
- Worm, C. M., The real world: A maintainer's view, in *Proceedings of the IEEE Reliability and Maintainability Symposium*, Philadelphia, PA, IEEE, New York, 1987.

11

Device Performance Criteria

11.1 Introduction

Many shortcomings in the received audio, video, and/or data stream of a transmission resulting from noise, reflections, and terrain limitations cannot be controlled by the equipment design engineer or the operating maintenance engineer. However, other problems can be prevented through proper application of vacuum tube devices and proper equipment adjustment. The critical performance parameters for a transmission system include

- Output power
- Transmission line/antenna voltage standing wave ratio (VSWR)
- Frequency stability/accuracy
- Occupied bandwidth
- Distortion mechanisms, including *total harmonic distortion* (THD), *intercarrier phase distortion*, and *quadrature distortion*
- Differential phase and gain
- Frequency response errors
- Envelope delay

The adjustment schedule required to maintain a satisfactory level of performance depends upon a number of factors, including

- Operating environment
- Equipment quality
- Age of the equipment
- Type of service being performed

11.2 Measurement Parameters

Most measurements in the radio frequency (RF) field involve characterizing fundamental parameters. These include signal level, phase, and frequency. Most other tests consist of measuring these fundamental parameters and displaying the results in combination by using some convenient format [1]. For example, signal-to-noise ratio (S/N) consists of a pair of level measurements made under different conditions and expressed as a logarithmic ratio. Modern test instruments perform the necessary manipulation of the measured parameters to yield a direct readout of the quantity of interest.

When characterizing a vacuum tube device or stage, it is common to view it as a box with input and output terminals. In normal use, a signal is applied to the input of the box, and the signal—modified in some way—appears at the output. Some measurements are *one-port* tests, such as impedance or noise level, and are not concerned with input/output signals, but with only one or the other.

Measurements are made on equipment to check performance under specified conditions and to assess suitability for use in a particular application. The measurements may be used to verify specified system performance or as a way of comparing several pieces of equipment for use in a system. Measurements also may be used to identify components in need of adjustment or repair.

Many parameters are important in vacuum tube devices and merit attention in the measurement process. Some common measurements include frequency response, gain or loss, harmonic distortion, intermodulation distortion (IMD), noise level, phase response, and transient response.

Measurement of level is fundamental to most RF specifications. Level can be measured either in absolute terms or in relative terms. Power output is an example of an absolute level measurement; it does not require any reference. S/N and gain or loss are examples of relative, or ratio, measurements; the result is expressed as a ratio of two measurements. Although it may not appear so at first, frequency response is also a relative measurement. It expresses the gain of the device under test as a function of frequency, with the midband gain (typically) as a reference.

Distortion measurements are a way of quantifying the unwanted components added to a signal by a piece of equipment. The most common technique is THD, but others are often used. Distortion measurements express the unwanted signal components relative to the desired signal, usually as a percentage or decibel value. This is another example of multiple level measurements that are combined to give a new measurement figure.

11.2.1 Power Measurements

The simplest definition of a level measurement is the alternating current amplitude at a particular place in the system under test [1]. However, in

contrast to direct current measurements, there are many ways of specifying ac voltage in a circuit. The most common methods include

- Average response
- Root mean square (rms)
- Peak

Strictly speaking, the term *level* refers to a logarithmic, or decibel, measurement. However, common parlance employs the term for an ac amplitude measurement, and that convention will be followed in this chapter.

11.2.1.1 Root Mean Square

The rms technique measures the effective power of the ac signal [1]. It specifies the value of the dc equivalent that would dissipate the same power if either were applied to a load resistor. This process is illustrated in Figure 11.1 for voltage measurements. The input signal is squared, and the average value is found. This is equivalent to finding the average power. The square root of

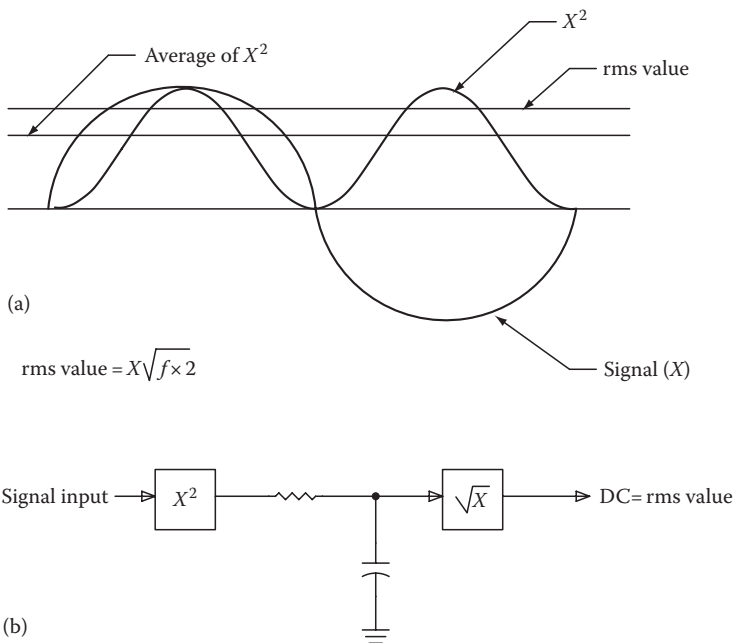


FIGURE 11.1 Root-mean-square (rms) voltage measurements: (a) the relationship of rms and average values and (b) the rms measurement circuit. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

this value is taken to transfer the signal from a power value back to a voltage. For the case of a sine wave, the rms value is 0.707 of its maximum value.

Assume that the signal is no longer a sine wave but rather a sine wave and several harmonics. If the rms amplitude of each harmonic is measured individually and added, the resulting value will be the same as an rms measurement of the signals together. Because rms voltages cannot be added directly, it is necessary to perform an rms addition. Each voltage is squared, and the squared values are added as follows:

$$V_{rms\ total} = \sqrt{V_{rms\ 1}^2 + V_{rms\ 2}^2 + \dots + V_{rms\ n}^2} \tag{11.1}$$

Note that the result is not dependent on the phase relationship of the signal and its harmonics. The rms value is determined completely by the amplitude of the components. This mathematical predictability is useful in practical applications of level measurement, enabling correlation of measurements made at different places in a system. It is also important in correlating measurements with theoretical calculations.

11.2.1.2 Average-Response Measurement

The average-responding voltmeter measures ac voltage by rectifying it and filtering the resulting waveform to its average value, as shown in Figure 11.2 [1].

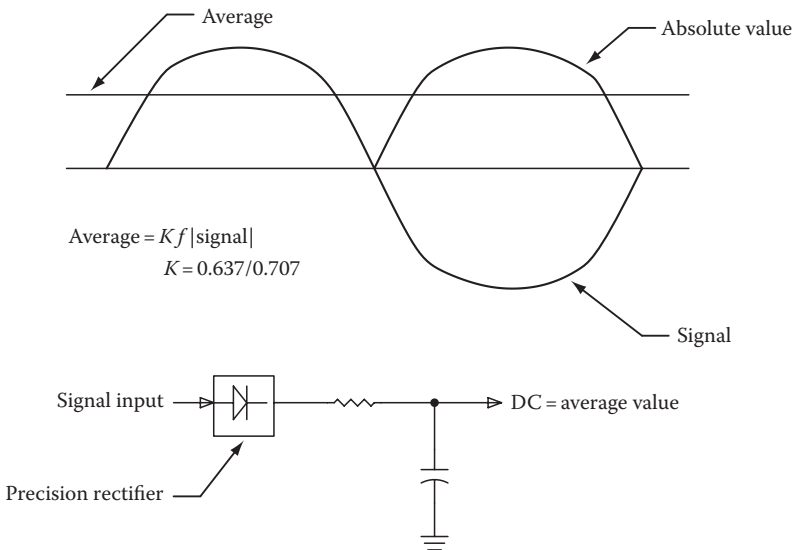


FIGURE 11.2 Average voltage measurements: (a) illustration of average detection and (b) average measurement circuit. (After Cabot, R.C., Audio test and measurements, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

This results in a dc voltage that can be read on a standard dc voltmeter. As shown in the figure, the average value of a sine wave is 0.637 of its maximum amplitude. Average-responding meters usually are calibrated to read the same as an rms meter for the case of a single sine wave signal. This results in the measurement being scaled by a constant K of 0.707/0.637, or 1.11. Meters of this type are called *average-responding, rms calibrated*. For signals other than sine waves, the response will be different and difficult to predict.

If multiple sine waves are applied, the reading will depend on the phase shift between the components and will no longer match the rms measurement. A comparison of rms and average-response measurements is made in Figure 11.3 for various waveforms. If the average readings are adjusted as described previously to make the average and rms values equal for a sine wave, all the numbers in the *average* column would be increased by 11.1%, and the *rms-average* numbers would be reduced by 11.1%.

11.2.1.3 Peak-Response Measurement

Peak-responding meters measure the maximum value that the ac signal reaches as a function of time [1]. This approach is illustrated in Figure 11.4. The signal is full-wave rectified to find its absolute value, then passed through a diode to a storage capacitor. When the absolute value of the voltage rises above the value stored on the capacitor, the diode will conduct and increase the stored voltage. When the voltage decreases, the capacitor will maintain the old value. Some method of discharging the capacitor is required so that a new peak value can be measured. In a true peak detector, this is accomplished by a solid-state switch. Practical peak detectors usually include a large-value resistor to discharge the capacitor gradually after the user has had a chance to read the meter.

The ratio of the true peak to the rms value is called the *crest factor*. For any signal but an ideal square wave, the crest factor will be greater than 1, as demonstrated in Figure 11.5. As the measured signal becomes more peaked, the crest factor increases.

By introducing a controlled charge and discharge time, a *quasi-peak* detector is achieved. The charge and discharge times may be selected, for example, to simulate the transmission pattern for a digital carrier. The gain of a quasi-peak detector is normally calibrated so that it reads the same as an rms detector for sine waves.

The *peak-equivalent sine* is another method of specifying signal amplitude. This value is the rms level of a sine wave having the same peak-to-peak amplitude as the signal under consideration. This is the peak value of the waveform scaled by the correction factor 1.414, corresponding to the peak-to-rms ratio of a sine wave. Peak-equivalent sine is useful when specifying test levels of waveforms in distortion measurements. If the distortion of a device is measured as a function of amplitude, a point will be reached at which the output level cannot increase any further. At this

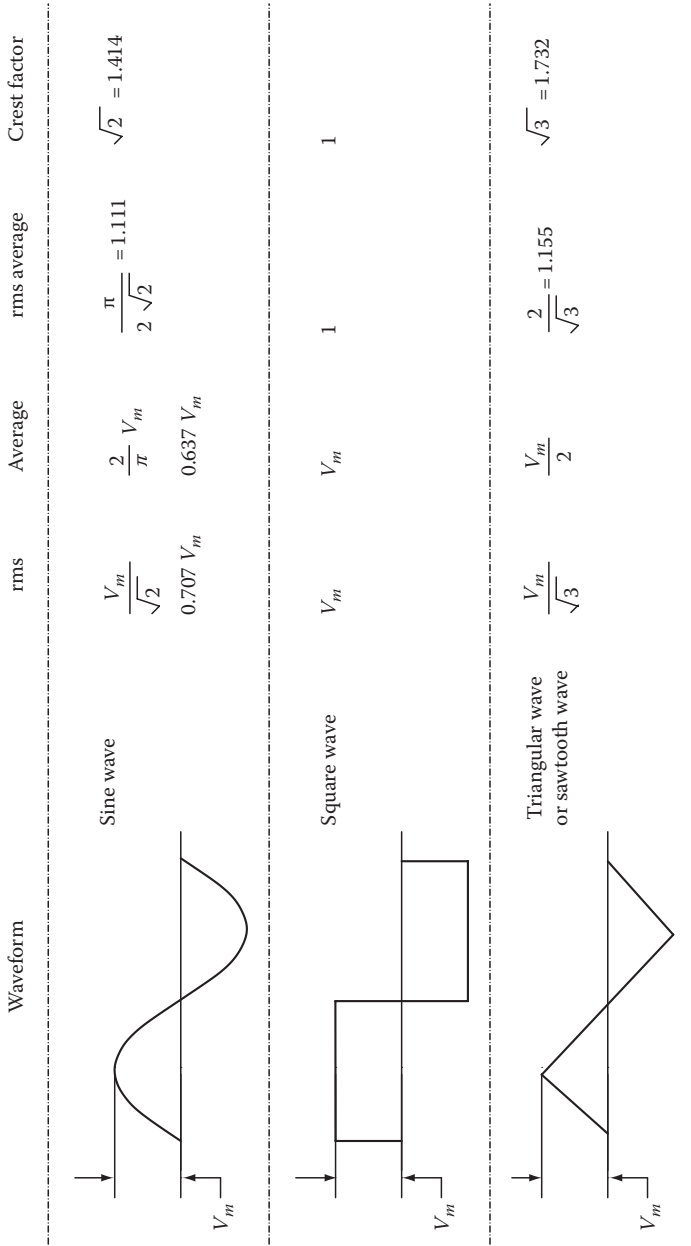


FIGURE 11.3 Comparison of rms and average voltage characteristics. (After Cabot, R.C., Audio test and measurements, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

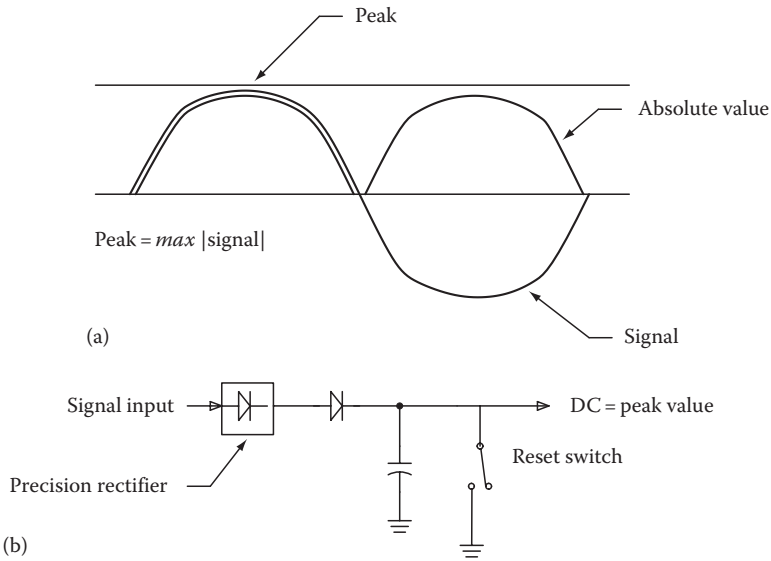


FIGURE 11.4 Peak voltage measurements: (a) illustration of peak detection and (b) peak measurement circuit. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

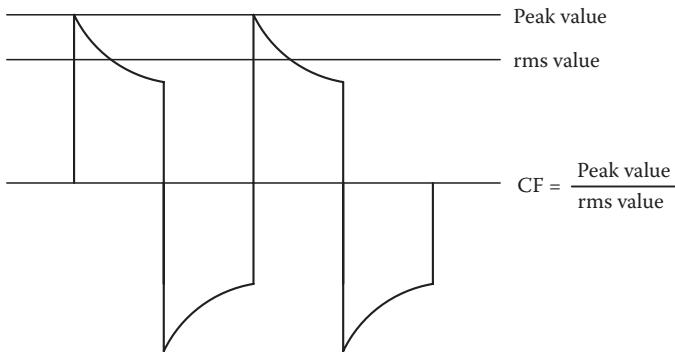


FIGURE 11.5 Illustration of the crest factor in voltage measurements. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

point the peaks of the waveform will be clipped, and the distortion will rise rapidly with further increases in level. If another signal is used for distortion testing on the same device, it is desirable that the levels at which clipping is reached correspond. Signal generators are normally calibrated in this way to allow changing between waveforms without clipping or readjusting levels.

11.2.1.4 Measurement Bandwidth

The bandwidth of the level-measuring instrument can have a significant effect on the accuracy of the reading [1]. For a meter with a single-pole roll-off (one bandwidth-limiting component in the signal path), significant measurement errors can occur. Such a meter with a specified bandwidth of 1 MHz, for example, will register a 10% error in the measurement of signals at 500 kHz. To obtain 1% accurate measurements (disregarding other error sources in the meter), the signal frequency must be less than 100 kHz.

Figure 11.6 illustrates another problem associated with limited-bandwidth measuring devices. In the figure, a distorted sine wave is measured by 2 m with different bandwidths. The meter with the narrower bandwidth does not respond to all the harmonics and gives a lower reading. The severity of this effect varies with the frequency being measured and the bandwidth of the meter; it can be especially severe in the measurement of wideband noise. Peak measurements are particularly sensitive to bandwidth effects. Systems with restricted low-frequency bandwidth will produce *tilt* in a square wave, and bumps in the high-frequency response will produce an *overshoot*. The effect of either will be an increase in the peak reading.

11.2.1.5 Meter Accuracy

Accuracy is a measure of how well an instrument quantifies a signal at a mid-band frequency [1]. This sets a basic limit on the performance of the meter in establishing the absolute amplitude of a signal. It is also important to look at the *flatness* specification to see how well this performance is maintained with changes in frequency. Flatness describes how well the measurements at any other frequency track those at the reference. If a meter has an accuracy of 2% at 1 MHz and a flatness of 1 dB (10%) from 20 kHz to 20 MHz, the inaccuracy can be as great as 12% at 20 MHz.

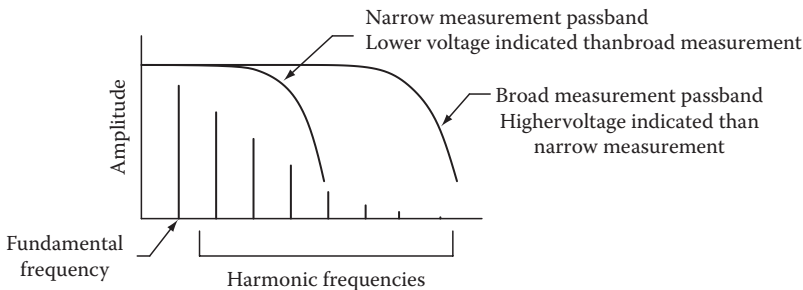


FIGURE 11.6
Effects of instrument bandwidth on voltage measurements.

Meters often have a specification of accuracy that changes with voltage range, being most accurate only in the range in which the instrument was calibrated. A meter with 1% accuracy on the 2V range and 1% accuracy per step would be 3% accurate on the 200V scale. Using the flatness specification given previously, the overall accuracy for a 100V, 20MHz sine wave is 14%. In many instruments, an additional accuracy derating is given for readings as a percentage of full scale, making readings at less than full scale less accurate.

However, the accuracy specification is not normally as important as the flatness. When performing frequency response or gain measurements, the results are relative and are not affected by the absolute voltage used. When measuring gain, however, the attenuator accuracy of the instrument is a direct error source. Similar comments apply to the accuracy and flatness specifications for signal generators. Most are specified in the same manner as voltmeters, with the inaccuracies adding in much the same manner.

11.2.1.6 Radio Frequency Power Measurement

Measurement of RF power output typically is performed with a directional wattmeter, calibrated for use over a specified range of frequencies and power levels. Various grades of accuracy are available, depending upon the requirements of the application. The most accurate measurements usually are made by determining the temperature rise of the coolant through a calibrated dummy load. The power absorbed by the coolant, usually water, can be calculated from the equation

$$P = K \times Q \times \Delta T \quad (11.2)$$

where

P is the power dissipated in kW

K is a constant, determined by the coolant (for pure water at 30°C, $K = 0.264$)

Q is the coolant flow in gal/min

ΔT is the difference between inlet and outlet water temperature in °C

This procedure often is used to verify the accuracy of in-line directional wattmeters, which are more convenient to use but typically offer less accuracy.

11.2.2 Decibel Measurement

Measurements in RF work often are expressed in decibels [1]. RF signals span a wide range of levels, too wide to be practically accommodated on a linear scale. The decibel is a logarithmic unit that compresses this wide range down to one that is easier to handle. Order-of-magnitude (factor of 10) changes result in equal increments on a decibel scale. A decibel may

be defined as the logarithmic ratio of two power measurements or as the logarithmic ratio of two voltages:

$$\text{dB} = 20 \log \left\{ \frac{E_1}{E_2} \right\} \quad (11.3)$$

$$\text{dB} = 10 \log \left\{ \frac{P_1}{P_2} \right\} \quad (11.4)$$

Decibel values from power measurements and decibel values from voltage measurements are equal if the impedances are equal. In both equations, the denominator variable is usually a stated reference, as illustrated by the example in Figure 11.7. Whether the decibel value is computed from the power-based equation or from the voltage-based equation, the same result is obtained.

A doubling of voltage will yield a value of 6.02 dB, and a doubling of power will yield 3.01 dB. This is true because doubling of voltage results in an increase in power by a factor of 4. Table 11.1 lists the decibel values for some common voltage and power ratios.

RF engineers often express the decibel value of a signal relative to some standard reference instead of another signal. The reference for decibel measurements may be predefined as a power level, as in dBk (decibels above 1 kW), or it may be a voltage reference. Often, it is desirable to specify levels in terms of a reference transmission level somewhere in the system under test. These measurements are designated dB_r, where the reference point or level must be separately conveyed.

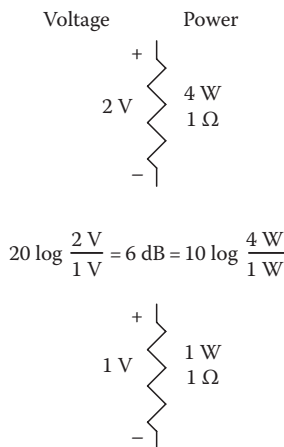


FIGURE 11.7

Example of the equivalence of voltage and power decibels. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

TABLE 11.1

Common Decibel Values and Conversion Ratios

dB Value	Voltage Ratio	Power Ratio
0	1	1
+1	1.122	1.259
+2	1.259	1.586
+3	1.412	1.995
+6	1.995	3.981
+10	3.162	10
+20	10	100
+40	100	10,000
-1	0.891	0.794
-2	0.794	0.631
-3	0.707	0.501
-6	0.501	0.251
-10	0.3163	0.1
-20	0.1	0.01
-40	0.01	0.0001

11.2.3 Noise Measurement

Noise measurements are specialized level measurements [1]. Noise may be expressed as an absolute level by simply measuring the voltage at the desired point in the system. This approach, however, is often not very meaningful. Specifying the noise performance as the S/N is usually a better approach. S/N is a decibel measurement of the noise level using the signal level measured at the same point as a reference. This makes measurements at different points in a system or in different systems directly comparable. A signal with a given S/N can be amplified with a perfect amplifier or attenuated with no change in the S/N. Any degradation in S/N at a later point in the system is the result of limitations of the equipment that follows.

Noise performance is an important parameter in the operation of any power vacuum tube amplifier or oscillator. All electric conductors contain free electrons that are in continuous random motion. It can be expected that, by pure chance, more electrons will be moving in one direction than in another at any instant. The result is that a voltage will be developed across the terminals of the conductor if it is an open circuit, or a current will be delivered to any connected circuit. Because this voltage (or current) varies in a random manner, it represents noise energy distributed throughout the frequency spectrum, from the lowest frequencies well into the microwave range. This effect is commonly referred to as *thermal-agitation noise*, because the motion of electrons results from thermal action. It is also

referred to as *resistance noise*. The magnitude of the noise depends upon the following:

- The resistance across which the noise is developed
- Absolute temperature of the resistance
- Bandwidth of the system involved

Random noise, similar in character to that produced in a resistance, is generated in vacuum tubes as a result of irregularities in electron flow. Tube noise can be divided into the following general classes:

- *Shot effect*, representing random variations in the rate of electron emission from the cathode
- *Partition noise*, arising from chance variations in the division of current between two or more positive electrodes
- *Induced grid noise*, produced as a result of variations in the electron stream passing adjacent to a grid
- *Gas noise*, generated by random variations in the rate of ion production by collision
- *Secondary emission noise*, arising from random variations in the rate of production of secondary electrons
- *Flicker effect*, a low-frequency variation in emission that occurs with oxide-coated cathodes

Shot effect in the presence of space charge, partition noise, and induced grid noise is the principal source of tube noise that must be considered in RF work.

11.2.4 Phase Measurement

When a signal is applied to the input of a device, the output will appear at some later point in time [1]. For sine wave excitation, this delay between input and output may be expressed as a proportion of the sine wave cycle, usually in degrees. One cycle is 360°, one half-cycle is 180°, and so on. This measurement is illustrated in Figure 11.8. The phasemeter input signal number 2 is delayed from, or is said to be *lagging*, input number 1 by 45°.

Most RF test equipment checks phase directly by measuring the proportion of one signal cycle between zero crossings of the signals. Phase typically is measured and recorded as a function of frequency over a specified range. For most vacuum tube devices, phase and amplitude responses are closely coupled. Any change in amplitude that varies with frequency will produce a corresponding phase shift.

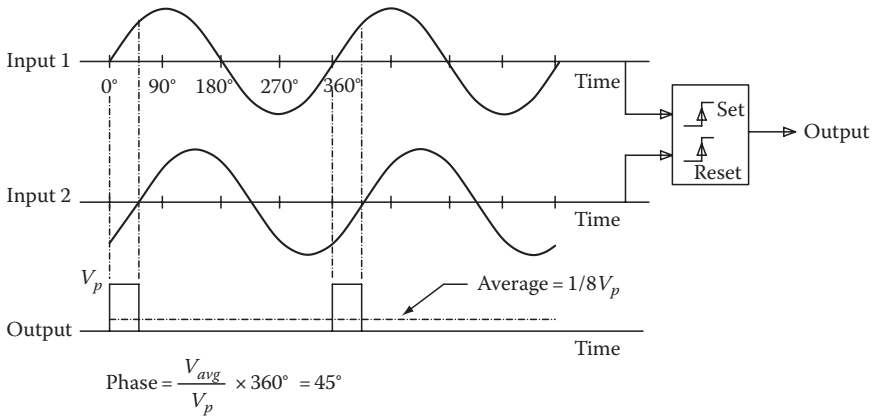


FIGURE 11.8 Illustration of the measurement of a phase shift between two signals. (After Cabot, R.C., Audio test and measurements, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

11.2.4.1 Relation to Frequency

When dealing with complex signals, the meaning of phase can become unclear [1]. Viewing the signal as the sum of its components according to Fourier theory, a different value of phase shift is found at each frequency. With a different phase value on each component, the one to be used as the reference is unclear. If the signal is periodic and the waveshape is unchanged passing through the device under test, a phase value still may be defined. This may be done by using the shift of the zero crossings as a fraction of the waveform period. Indeed, most commercial phase-measuring instruments will display this value. However, in the case of differential phase shift with frequency, the waveshape will be changed. It is then impossible to define any phase-shift value, and phase must be expressed as a function of frequency.

Group delay is another useful expression of the phase characteristics of an RF device. Group delay is the slope of the phase response. It expresses the relative delay of the spectral components of a complex waveform. If the group delay is flat, all components will arrive at a given point together. A peak or rise in the group delay indicates that those components will arrive later by the amount of the peak or rise. Group delay is computed by taking the derivative of the phase response vs. frequency. Mathematically,

$$\text{Group delay} = - \frac{(\text{phase at } f_2 - \text{phase at } f_1)}{f_2 - f_1} \tag{11.5}$$

This definition requires that phase be measured over a range of frequencies to give a curve that can be differentiated. It also requires that the phase measurements be performed at frequencies sufficiently close to provide a smooth and accurate derivative.

11.2.5 Nonlinear Distortion

Distortion is a measure of signal impurity, a deviation from ideal performance of a device, stage, or system [1]. Distortion usually is expressed as a percentage or decibel ratio of the undesired components to the desired components of a signal. Distortion of a vacuum tube device or stage is measured by inputting one or more sine waves of various amplitudes and frequencies. In simplistic terms, any frequencies at the output that were not present at the input are distortion. However, strictly speaking, components caused by power line interference or another spurious signal are not distortion but, rather, noise. Many methods of measuring distortion are in common use, including harmonic distortion and several types of IMD.

11.2.5.1 Harmonic Distortion

The transfer characteristic of a typical vacuum tube amplifier is shown in Figure 11.9 [1]. The transfer characteristic represents the output voltage at any point in the signal waveform for a given input voltage; ideally this is a straight line. The output waveform is the projection of the input sine wave on the device transfer characteristic. A change in the input produces

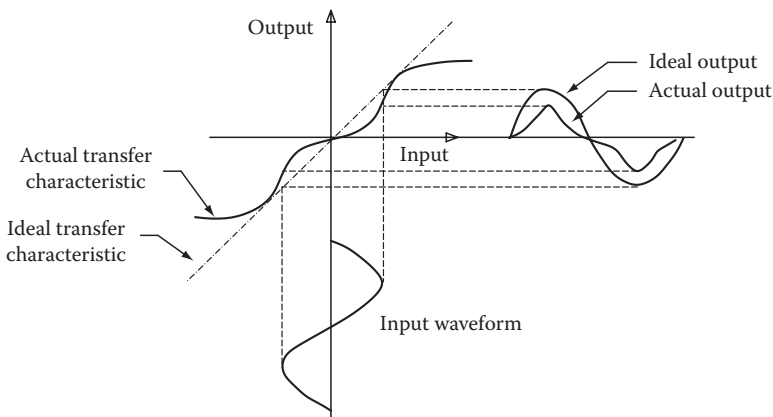


FIGURE 11.9

Illustration of THD measurement of an amplifier transfer characteristic. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

a proportional change in the output. Because the actual transfer characteristic is nonlinear, a distorted version of the input waveshape appears at the output.

Harmonic distortion measurements excite the device under test with a sine wave and measure the spectrum of the output. Because of the nonlinearity of the transfer characteristic, the output is not sinusoidal. By using Fourier series, it can be shown that the output waveform consists of the original input sine wave plus sine waves at integer multiples (harmonics) of the input frequency. The spectrum of a distorted signal is shown in Figure 11.10. The harmonic amplitudes are proportional to the amount of distortion in the device under test. The percentage of harmonic distortion is the rms sum of the harmonic amplitudes divided by the rms amplitude of the fundamental.

Harmonic distortion may be measured with a spectrum analyzer or a distortion test set. Figure 11.11 shows the setup for a spectrum analyzer.

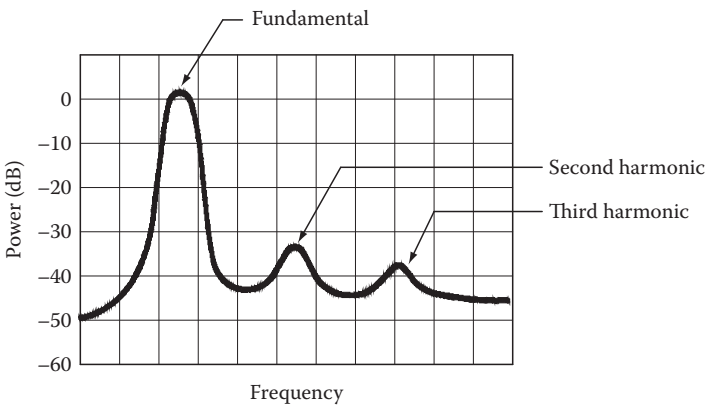


FIGURE 11.10
Example of reading THD from a spectrum analyzer.

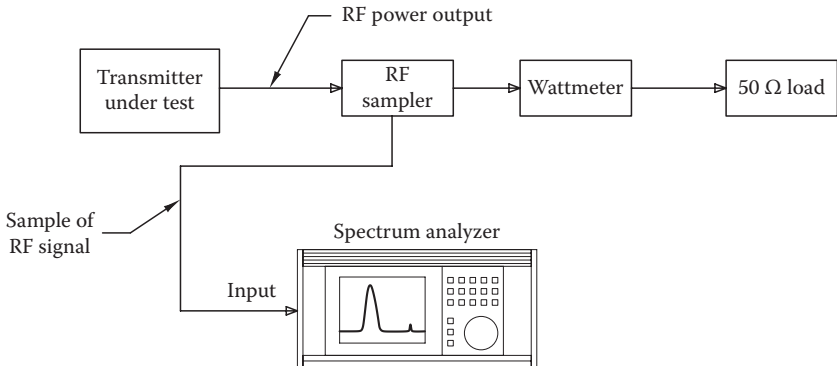


FIGURE 11.11
Common test setup to measure harmonic distortion with a spectrum analyzer.

As shown in Figure 11.10, the fundamental amplitude is adjusted to the 0 dB mark on the display. The amplitudes of the harmonics are then read and converted to linear scale. The rms sum of these values is taken, which represents the THD.

A simpler approach to the measurement of harmonic distortion can be found in the notch-filter distortion analyzer, illustrated in Figure 11.12. This device, commonly referred to as simply a distortion analyzer, removes the fundamental of the signal to be investigated and measures the remainder. Figure 11.13 shows the notch-filter approach applied to a spectrum analyzer for distortion measurements.

The correct method of representing percentage distortion is to express the level of the harmonics as a fraction of the fundamental level [2]. However, many commercial distortion analyzers use the total signal level as the reference voltage. For small amounts of distortion, these two quantities are essentially equivalent. At large values of distortion, however, the total signal level will be greater than the fundamental level. This makes distortion measurements on such units lower than the actual value. The relationship between

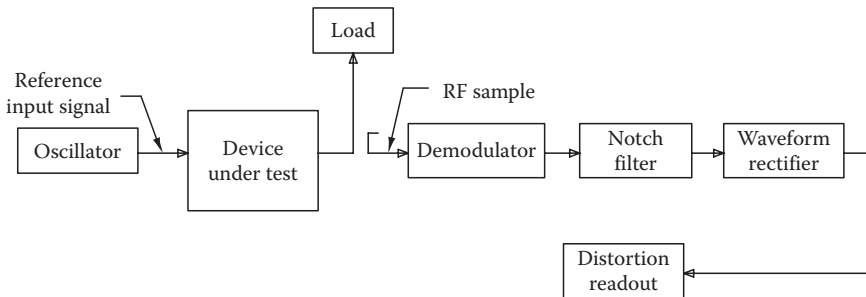


FIGURE 11.12
Simplified block diagram of a harmonic distortion analyzer.

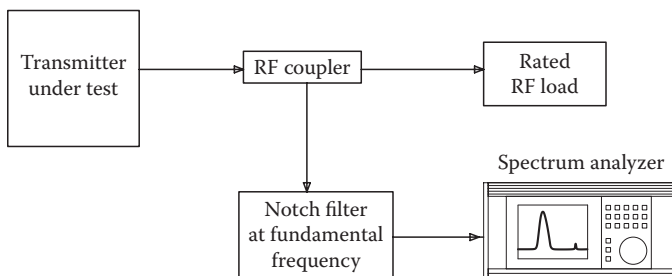


FIGURE 11.13
Typical test setup for measuring the harmonic and spurious output of a transmitter. The notch filter is used to remove the fundamental frequency to prevent overdriving the spectrum analyzer input and to aid in the evaluation of distortion components.

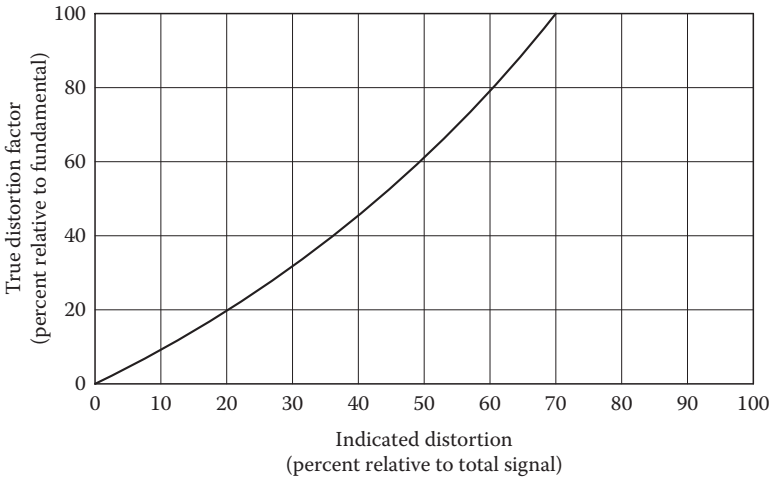


FIGURE 11.14 Conversion graph for indicated distortion and true distortion. (After Termaine, H.W., *Audio Cyclopedia*, Howard W. Sams, Indianapolis, IN, 1975.)

the measured distortion and true distortion is illustrated in Figure 11.14. The errors are not significant until about 20% measured distortion.

Because of the notch-filter response, any signal other than the fundamental will influence the results, not just harmonics. Some of these interfering signals are illustrated in Figure 11.15. Any practical signal contains some noise, and the distortion analyzer will include this noise in the reading. Because of these added components, the correct term for this measurement is *total harmonic distortion and noise* (THD + N). Additional filters are

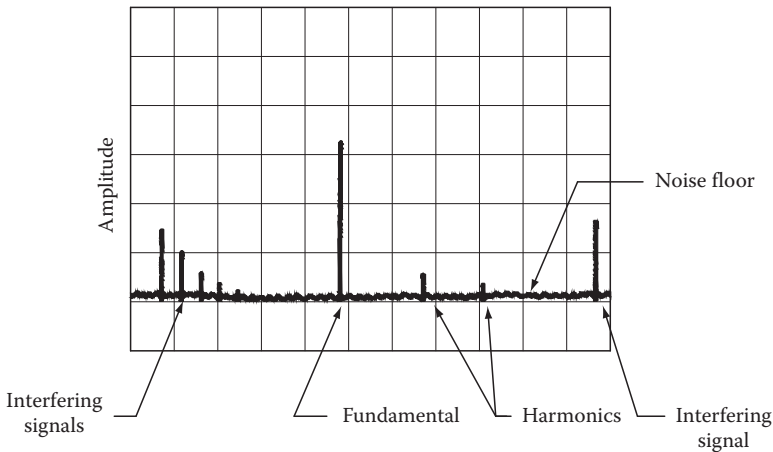
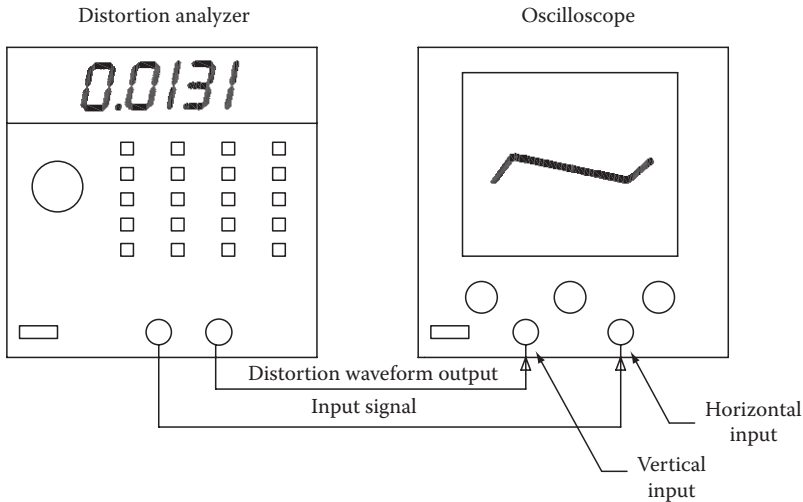


FIGURE 11.15 Example of interference sources in distortion and noise measurements.

**FIGURE 11.16**

Transfer-function monitoring configuration using an oscilloscope and distortion analyzer. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

included on most distortion analyzers to reduce unwanted noise and permit a more accurate reading.

The use of a sine wave test signal and a notch-type distortion analyzer provides the distinct advantage of simplicity in both design and use. This simplicity has an additional benefit in ease of interpretation. The shape of the output waveform from a notch-type analyzer indicates the slope of the nonlinearity. Displaying the residual components on the vertical axis of an oscilloscope and the input signal on the horizontal axis provides a plot of the deviation of the transfer characteristic from a best-fit straight line. This technique is diagrammed in Figure 11.16. The trace will be a horizontal line for a perfectly linear device. If the transfer characteristic curves upward on positive input voltages, the trace will bend upward at the right-hand side.

Examination of the distortion components in real time on an oscilloscope allows observation of oscillation on the peaks of the signal and clipping. This is a valuable tool in the design and development of RF circuits, and one that no other distortion measurement method can fully match. Viewing the residual components in the frequency domain by using a spectrum analyzer also reveals much information about the distortion mechanism inside the device or stage under test.

When measuring distortion at high frequencies, bandwidth limitations are an important consideration, as illustrated in Figure 11.17. Because the components being measured are harmonics of the input frequency, they may fall outside the passband of the device under test.

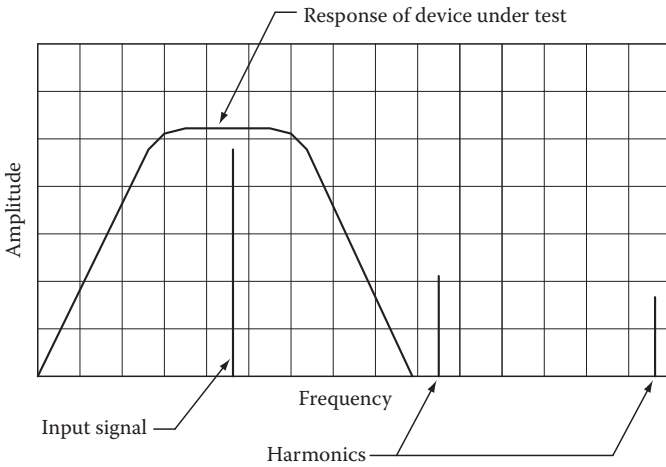


FIGURE 11.17 Illustration of problems that occur when measuring harmonic distortion in band-limited systems. (After Cabot, R.C., *Audio test and measurements*, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

11.2.5.2 Intermodulation Distortion

IMD describes the presence of unwanted signals that are caused by interactions between two or more desired signals [1]. IMD, as it applies to an RF amplifier, refers to the generation of spurious products in a nonlinear amplifying stage. IMD can be measured using a distortion monitor or a spectrum analyzer.

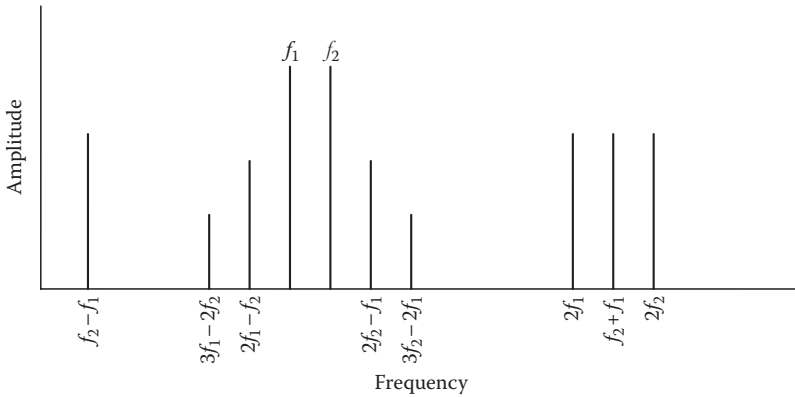
Nonlinearities in a circuit can cause it to act like a mixer, generating the sum and difference frequencies of two signals that are present. These same imperfections generate harmonics of the signals, which then can be mixed with other fundamental or harmonic frequencies. Figure 11.18 shows the relationship of these signals in the frequency domain. The *order* of a particular product of IMD is defined as the number of “steps” between it and a single fundamental signal. Harmonics add steps. For example, the second harmonic = $2f_1 = f_1 + f_1$, which is two steps, making it a *second-order* product. The product resulting from $2f_1 - f_2$ is, then, a *third-order* product. Some of the IMD products resulting from two fundamental frequencies include

$$\text{second-order} = 2f_1 \tag{11.6}$$

$$= 2f_2 \tag{11.7}$$

$$= f_1 + f_2 \tag{11.8}$$

$$= f_1 - f_2 \tag{11.9}$$

**FIGURE 11.18**

Frequency-domain relationships of the desired signals and low-order IMD products.

$$\text{third-order} = 2f_1 + f_2 \quad (11.10)$$

$$= 2f_1 - f_2 \quad (11.11)$$

$$= 2f_2 + f_1 \quad (11.12)$$

$$= 2f_2 - f_1 \quad (11.13)$$

and so on.

The order is important because, in general, the amplitudes of IMD products fall off as the order increases. Therefore, low-order IMD products have the greatest potential to cause problems if they fall within the band of interest, or on frequencies used by other services or equipment. Odd-order IMD products are particularly troublesome because they fall closest to the signals that cause them, usually within the operating frequency band.

11.2.5.3 Measurement Techniques

A number of techniques have been devised to measure the intermodulation (IM) of two or more signals passing through a device simultaneously. The most common of these involves the application of a test waveform consisting of a low-frequency signal and a high-frequency signal mixed in a specified amplitude ratio. The signal is applied to the device under test, and the output waveform is examined for modulation of the upper frequency by the low-frequency signal. The amount by which the low-frequency signal modulates the high-frequency signal indicates the degree of nonlinearity. As with harmonic distortion measurement, this test may be done with a spectrum analyzer or a dedicated distortion analyzer. The test setup for a spectrum analyzer is shown in Figure 11.19. Two independent signal sources are connected using a power combiner to drive the device under test. The sources

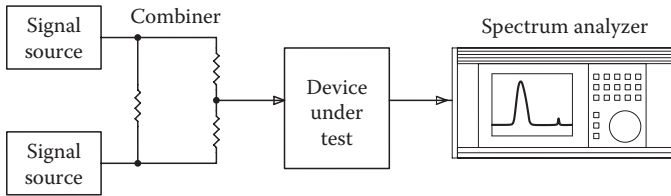


FIGURE 11.19 Test setup for measuring the IMD of an RF amplifier using a spectrum analyzer.

Range: -10 dBm Sweep time 1.23 s
 Res BW: 200 Hz
 Marker: 10,010,500 Hz -15.07 dBm

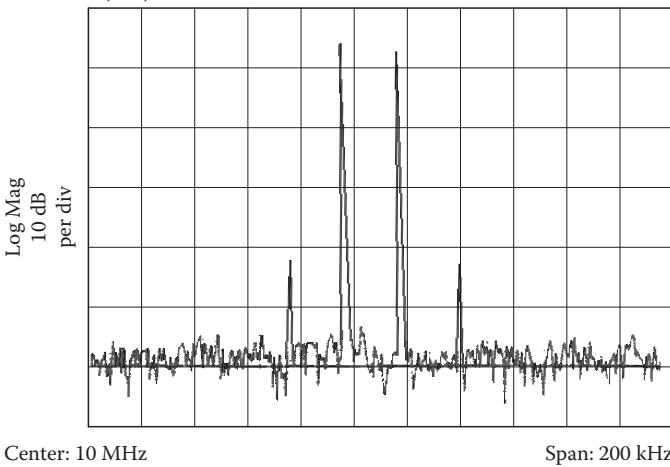


FIGURE 11.20 Typical two-tone IMD measurement, which evaluates the third-order products within the bandpass of the original tones.

are set at the same or different output level(s), but at different frequencies. A typical spectrum analyzer display of the two-tone distortion test is shown in Figure 11.20. As shown, the third-order products that fall close to the original two tones are measured. This measurement is a common, and important, one because such products are difficult to remove by filtering.

The mechanisms of IM as applied to interfering radiated signals are discussed later in this chapter.

11.2.5.4 Addition and Cancellation of Distortion Components

The addition and cancellation of distortion components in test equipment or the device under test is an often-overlooked problem in measurement [1]. Consider the examples in Figures 11.21 and 11.22. Assume that one device under test has a transfer characteristic similar to that diagrammed at the top left of Figure 11.21a, and another has the characteristic diagrammed at the center.

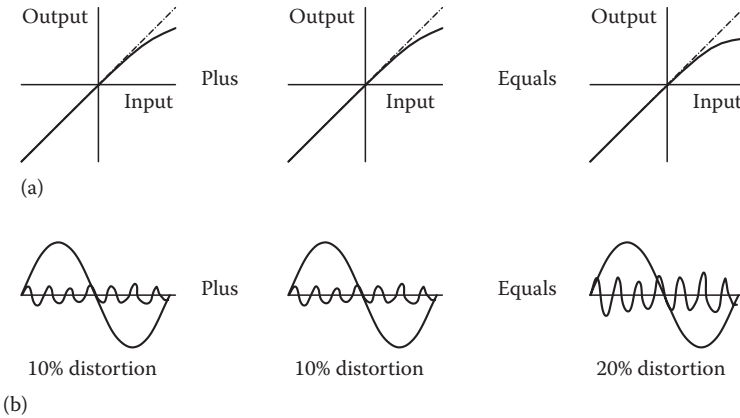


FIGURE 11.21 Illustration of the addition of distortion components: (a) addition of transfer-function nonlinearities and (b) addition of distortion components. (After Cabot, R.C., Audio test and measurements, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

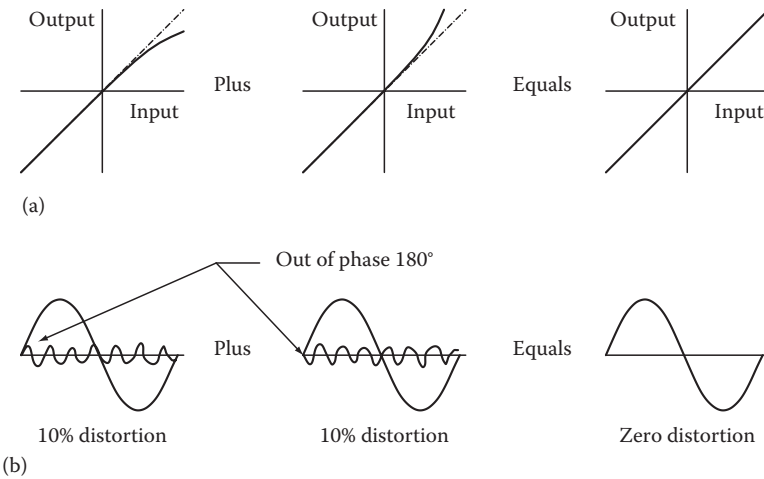


FIGURE 11.22 Illustration of the cancellation of distortion components: (a) cancellation of transfer-characteristic nonlinearities and (b) cancellation of the distortion waveform. (After Cabot, R.C., Audio test and measurements, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.)

If the devices are cascaded, the resulting transfer-characteristic nonlinearity will be magnified as shown. The effect on sine waves in the time domain is illustrated in Figure 11.21b. The distortion components generated by each nonlinearity are in phase and will sum to a component of twice the original magnitude. However, if the second device under test has a complementary transfer characteristic, as shown in Figure 11.22, quite a different result is obtained.

When the devices are cascaded, the effects of the two curves cancel, yielding a straight line for the transfer characteristic (Figure 11.22a). The corresponding distortion products are out of phase, resulting in no measured distortion components in the final output (Figure 11.22b).

This problem is common at low levels of distortion, especially between the test equipment and the device under test. For example, if the test equipment has a residual of 0.02% when connected to itself, and readings of 0.01% are obtained from the circuit under test, cancellations are occurring. It is also possible for cancellations to occur in the test equipment itself, with the combined analyzer and signal generator giving readings lower than the sum of their individual residuals. If the distortion is the result of even-order (asymmetrical) nonlinearity, reversing the phase of the signal between the offending devices will change a cancellation to an addition. If the distortion is the result of an odd-order (symmetrical) nonlinearity, phase inversions will not affect the cancellation.

11.2.5.5 Intermodulation Precorrection Techniques

In a practical system, amplifiers cannot be absolutely linear; all linear amplifiers exhibit IMD to some extent, relative to the amplitudes of the parent signals [3]. Adequate linearity is judged by the application, but in general, the in-band odd-order products should be as low as the state of the art permits because they are in-band and thus cannot be filtered out.

All IM products are the result of undesirable time domain multiplication of the matrix of input signals to a linear amplifier. All harmonic components are the result of the same process, only here, an individual signal is multiplied by itself. For second-order IM, the ideal linear amplifier is modified by modeling a single mixer feeding a time domain multiplied signal around the amplifier. For third-order IM, the ideal linear amplifier is modified by modeling two mixers feeding a time domain multiplied signal around the amplifier, as shown in Figure 11.23.

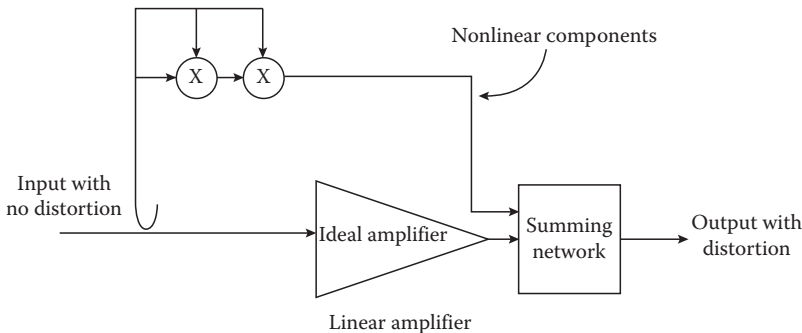


FIGURE 11.23

Basic model of an IM correction circuit. (After Hulick, T.P., Very simple out-of-band IMD correctors for adjacent channel NTSC/DTV transmitters, *Digital Television '98*, Intertec Publishing, Overland Park, KS, 1998.)

Because conventional amplitude modulation (AM) contains at least three parent signals (the carrier plus two sidebands), AM amplification has always been fraught with in-band IMD. A common technique to correct for these distortion products is to precorrect the power amplifier using the following process:

- Intentionally generate an anti-IM signal by mixing the three input signals together.
- Adjust the amplitude of the anti-IM product to match that of the real IM product.
- Invert the phase.
- Delay the correction signal or the parent signals in time so as to permit proper alignment.
- Add the correction signal into the main parent signal path to cause cancellation in the output circuit of the amplifier.

This is a well-known technique used in numerous applications.

Each IM product may have its own anti-IM product intentionally generated to cause targeted cancellation. Even if one or more parent signals contains modulation of some type—AM, frequency modulation (FM), or phase modulation (PM)—the legitimate modulation signal is often included in the intended mixing circuit so that even the sidebands produced around an IMD product are canceled.

The precorrection circuit typically consists of mixer devices with a local oscillator (LO) port, an RF port, and an IF port. The LO port is fed with the parent signal, such as the carrier, which contains no modulation. Its purpose is to hard-switch the mixer diodes, causing frequency translation of the signal fed to the RF port to the sum and difference frequencies of the LO and RF input signals exiting at the IF port. The IM product to be canceled is either a result of the sum or the difference signal so that the product of no concern is filtered out. The resulting intended mixer product is fed to the IF port of a second mixer where LO injection is from the third parent signal. The third-order intended mixing product is then available at the second mixer RF port and will be at the frequency of the offending IM and coherent with it.

Digital signals suitable for radio transmission actually remain analog, but are termed *digital* because of the nature of the information that they carry. A modern digital radio transmission using quadrature amplitude modulation (QAM) or n -level vestigial sideband modulation (e.g., 8-VSB) consists of a vector signal constantly moving in phase angle and magnitude relative to some reference. The vector signal is smooth and without discontinuity so that it is well contained within a finite bandwidth. Its digital nature stems from decoding the proper point or spatial area around a point (QAM) of a finite number of legitimate points, or the proper line (VSB) or spatial area around a line of a finite number of legitimate lines, that the vector is passing through.

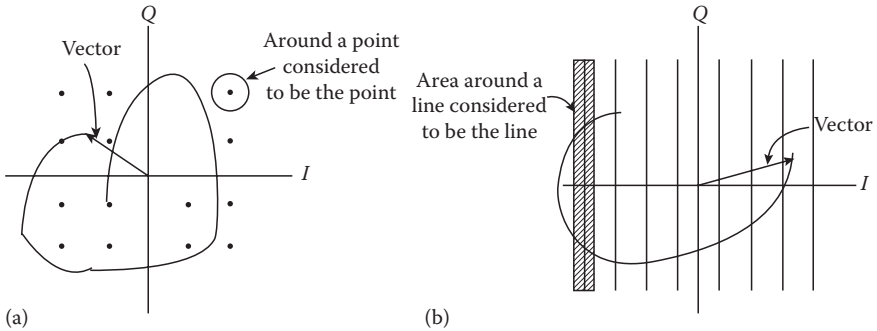


FIGURE 11.24 Digital decoding process: (a) 16-QAM and (b) 8-VSB. (After Hulick, T.P., *Very simple out-of-band IMD correctors for adjacent channel NTSC/DTV transmitters*, Digital Television '98, Intertec Publishing, Overland Park, KS, 1998.)

Figure 11.24 illustrates this process. In either case, the area is assigned a discrete digital value, which represents a portion of the information being transmitted. Spatial location is generally expressed in terms of rectangular coordinates with I (in phase) and Q (quadrature phase) components.

Because the vector appears to move at random from point to point, especially resulting from the use of an intentional randomizer circuit, the frequency distribution within the passband of the signal is constant, flat, and favors no particular frequency. Figure 11.25 illustrates the frequency distribution with the sidebands shown before and after amplification. As a result, the transmitted signal is said to be “noise-like.” As such, it takes on the nature of noise and need not be periodic.

For this case, there is no identifiable set of parent components upon which to obtain signals for the purpose of mixing to generate an anti-IM signal for cancellation purposes. There is no carrier signal to feed into the LO port of a mixer for hard switching to generate sum and difference frequencies. Because the signal is noise-like, there is no correlation of any

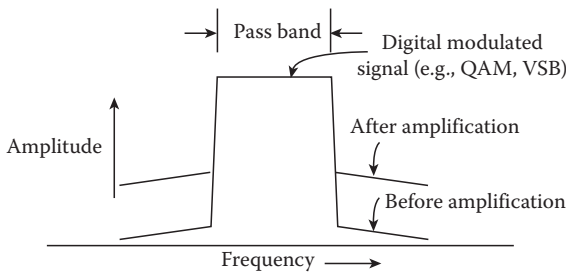


FIGURE 11.25 Digitally modulated RF signal spectrum before and after amplification. (After Hulick, T.P., *Very simple out-of-band IMD correctors for adjacent channel NTSC/DTV transmitters*, Digital Television '98, Intertec Publishing, Overland Park, KS, 1998.)

piece of passband spectrum with any other. Generating a precorrection signal, thus, is not possible using the process described for the discrete signal analog case.

11.3 Vacuum Tube Operating Parameters

Direct current input power to the final RF output stage is a function of the applied plate or collector voltage and the plate or collector current of the final tube(s) [4]. Test points or built-in meters usually are provided for checking the dc voltage and current levels of the final amplifier.

Transmitter deviation or modulation must be set properly. A deviation level set too high or too low adversely affects radio range. If the deviation is set too low, the S/N of the overall system is diminished. If the deviation is set too high, the transmitter signal may exceed the receiver's modulation acceptance bandwidth, causing severe distortion on modulation peaks and out-of-band spurious radiation.

The difference between the positive and negative excursions of the deviated waveform is defined as *modulation dissymmetry*. Theoretically, positive and negative excursions should be perfectly symmetrical. In practice, however, there is usually some dissymmetry. In a multichannel phase-modulated transmitter, modulation dissymmetry may be greater on some channels than others as a result of modulator tuning. If the modulator tuning is peaked on one channel, for example, response on another channel may be compromised.

11.3.1 Stage Tuning

A power output stage can be tuned in any number of ways [4]. Determining factors include the type of device, circuit design, and operational objectives. A PA can be tuned to optimize bandwidth, linearity, efficiency, and other parameters. It is often necessary to compromise one parameter to enhance another. This trade-off is dictated by the application.

11.3.1.1 Power Grid Tubes

When optimizing a tetrode or pentode RF amplifier for proper excitation and loading, the general tuning procedure varies depending upon whether the screen voltage is taken from a fixed supply or a dropping resistor supply with poor regulation [4]. If both the screen supply and grid bias are from fixed sources with good regulation, the plate current is almost entirely controlled by RF excitation. The loading is then varied until the maximum power output is obtained. Following these adjustments, the excitation is trimmed, along with the loading, until the desired control and screen grid currents are obtained.

In the case of an RF amplifier where both the screen and grid bias are taken from sources with poor regulation, the stage will tune in a manner similar to a triode RF power amplifier. The plate current will be controlled principally by varying the loading, and the excitation trimmed to yield the desired control grid current. In this case, the screen current will be almost entirely set by the choice of the dropping resistor. Excitation and loading will vary the screen voltage considerably, and these should be trimmed to yield a desired screen voltage.

Adjustment of the excitation and loading of a grounded-grid RF amplifier requires a slightly different procedure. The plate voltage (plate and screen voltage in the case of a tetrode or pentode) must be applied before the excitation. If this precaution is not followed, the control grid may sustain damage. The loading is increased as the excitation is increased. When the desired plate current is reached, note the power output. The loading can be reduced slightly and the excitation increased until the plate current is the same as before. If the power output is less than before, a check can be made with increased loading and less excitation. By proper trimming, the proper grid current, plate current, and optimum power output can be obtained. In a grounded-grid circuit, the cathode or input circuit is in series with the plate circuit. Because of this arrangement, any change made in the plate circuit will affect the input circuit. Therefore, the driver amplifier does not see its designed load until the driven stage is up to full plate current.

Typical operating values provide a starting point for tuning efforts. As long as the maximum limits are not exceeded, a wide range of tuning options is available.

11.3.1.2 Screen Voltage

Typical operating values given on the data sheet for tetrode and pentode devices include recommended screen voltages [4]. These values are not critical for most applications; they are chosen as a convenient value consistent with low driving power and reasonable screen dissipation. If lower values of screen voltage are used, more driving voltage will be required on the grid to obtain the same plate current. If higher values of screen voltage are used, less driving voltage will be required. Thus, high power gain can be achieved, provided the circuit exhibits adequate stability. Be careful not to exceed the screen dissipation limit. The value of screen voltage can be chosen to suit available power supplies or amplifier conditions.

The published characteristic curves of tetrodes and pentodes are shown for commonly used screen voltages. Occasionally, it is desirable to operate the tetrode or pentode at some screen voltage other than that shown on the characteristic curves. It is a relatively simple matter to convert the published curves to corresponding curves at a different screen voltage. This conversion is based on the fact that if all interelectrode voltages are either raised or lowered by the same relative amount, the shape of the voltage field pattern is not

TABLE 11.2

Three-Halves Current Values of Commonly Used Factors

Voltage Factor	Current Factor
0.25	0.125
0.50	0.35
0.75	0.65
1.00	1.00
1.25	1.40
1.50	1.84
1.75	2.30
2.00	2.80
2.25	3.40
2.50	4.00
2.75	4.60
3.00	5.20

Source: After Laboratory Staff, *Care and Feeding of Power Grid Tubes*, Varian/Eimac, San Carlos, CA, 1984.

altered, and neither is the current distribution. The current lines simply take on new proportionate values in accordance with the *three-halves power law*. This method fails only where insufficient cathode emission or high secondary emission affects the current values.

For example, if the characteristic curves are shown at a screen voltage of 250 V and it is desirable to determine conditions at 500 V on the screen, all voltage scales should be multiplied by the same factor that is applied to the screen voltage (2 in this case). The 1 kV plate voltage line thus becomes 2 kV, the 50 V grid voltage line becomes 100 V, and so on. All current lines then assume new values in accordance with the $3/2$ power law. Because the voltage was increased by a factor of 2, the current lines all will be increased in value by a factor of $2^{3/2}$ or 2.8. All current values are then multiplied by the factor of 2.8. The 100 mA line thus becomes a 280 mA line, and so on.

Likewise, if the screen voltage given on the characteristic curve is higher than the conditions desired, the voltage should be reduced by the same factor that is used to obtain the desired screen voltage. Correspondingly, the current values all will be reduced by an amount equal to the $3/2$ power of this factor. Commonly used factors are given in Table 11.2.

11.3.1.3 Back-Heating by Electrons

Back-heating of the cathode is an undesirable action involving the motion of electrons within the tube at VHF and UHF [4]. Because the time of flight of electrons from the cathode through the grid structure to the plate becomes

an appreciable part of the cycle at high frequencies, electrons can be stopped in flight and turned back by the rapidly changing grid voltage. Under these conditions, the electrons are deflected from their normal paths and given excess energy with which they bombard the cathode and other portions of the tube structure. This effect can be aggravated by the choice of operating conditions to the extent that destructive effects occur. The tube can be destroyed within a few minutes under severe conditions.

Fortunately, the conditions that tend to minimize this back bombardment by electrons also characterize minimum driving voltage. The tendency of electrons to turn back in flight is reduced by using the lowest possible RF grid voltage on the tube. This is achieved by using the lowest possible dc grid bias. In tetrodes, this effect is inherently lower because of the action of the screen in accelerating the electrons toward the anode. A tetrode also permits the use of smaller grid voltages. Consequently, under favorable conditions, the number of electrons turned back to heat the cathode and tube structure can be kept to a practical, low level. In addition to low dc grid bias, high screen voltage is desirable.

The plate circuit should always operate with heavy loading (low external plate impedance) so that the minimum instantaneous value of plate voltage will stay sufficiently positive to continue accelerating electrons to the anode. For this reason, the longest tube life is realized when the tetrode amplifier is heavily loaded, as indicated by having small values of dc screen and dc control grid current.

Never operate a tube with light plate loading. If the plate load is removed so that the minimum instantaneous plate voltage tends to fall to values approximating cathode potential (as it must do when the loading is removed completely and excitation is present), the number of electrons turned back can be destructive to the tube. It has been found that under conditions of no loading, the electron bombardment and increased electric field heating of the insulating portion of the tube is often sufficient to cause suck-in of the glass or cracking of the ceramic envelope. Automatic protection should be provided to remove all voltages from the tube if the plate circuit loading becomes too light for the amount of excitation applied.

Note that parasitic oscillations seldom are loaded heavily, as indicated by the high grid currents often observed during such self-oscillation. Thus, excessive RF plate voltages are developed that, at VHF, can be as damaging as unloaded operation on a VHF fundamental frequency. If unloaded VHF parasitic oscillations are present simultaneously with apparently satisfactory operation on the fundamental, unexplained reduction of tube life may result.

Occasionally, an output line circuit may resonate simultaneously to a harmonic frequency as well as to the fundamental frequency. The higher resonant modes of practical line circuits are not normally harmonically related. Sometimes, however, the tuning curve of a mode will cross the fundamental tuning curve and, at that point, the circuit will build up resonant voltages at both the harmonic and the fundamental frequencies. The harmonic

resonance usually is lightly loaded and the damaging action is similar to that of lightly loaded parasitic or fundamental. Again, while operation of the tube and circuit on the fundamental may appear normal, but with lower than expected efficiency, damage can occur to the tube.

In addition to operating the tube with minimum bias, high screen voltage, and heavy loading on the plate circuit, some degree of compensation for the remaining back-heating of the cathode may be required. This can be accomplished by lowering the filament voltage or heater voltage until the cathode operates at a normal temperature. In the case of tetrodes and pentodes, by taking precautions necessary to minimize back bombardment by electrons, the required compensation for back-heating of the cathode is not significant and may often be neglected.

11.3.2 Amplifier Balance

In a push-pull RF amplifier, lack of balance in the plate circuit or unequal plate dissipation usually is caused by a lack of symmetry in the RF circuit [4]. Normally, the tubes are similar enough that such imbalance is not associated with the tubes and their characteristics. This can readily be checked by interchanging the tubes in the sockets (provided both tubes have common dc voltages to plate, screen, and grid) and observing whether the unbalanced condition remains at the socket location or moves with the tube. If it remains at the socket location, the circuit requires adjustment. If appreciable imbalance is associated with the tube, it is likely that one tube is abnormal and should be replaced.

The basic indicators of balance are the plate current and plate dissipation of each tube. It is assumed that the circuit applies the same dc plate voltage, dc screen voltage (in the case of a tetrode or pentode), and dc grid bias to each tube from common supplies. Also, it is initially assumed that the plate circuit is mechanically and electrically symmetrical (or approximately so).

Imbalance in a push-pull RF amplifier is usually the result of unequal RF voltages applied to the grids of the tubes, or the RF plate circuit applying unequal RF voltages to the plates of the tubes. Correction of this situation involves first balancing grid excitation until equal dc plate currents flow in each tube. Next, the RF plate circuit is adjusted until equal plate dissipation appears on each tube, or equal RF plate voltage is observed. The balance of plate current is a more important criterion than equality of screen current (in a tetrode or pentode) or grid current. This results from the fact that tubes tend to be more uniform in plate current characteristics. However, the screen current also is sensitive to a lack of voltage balance in the RF plate circuit and may be used as an indicator. If the tubes differ somewhat in screen current characteristics, and the circuit has common dc supply voltages, final trimming of the plate circuit balance may be made by interchanging tubes and adjusting the circuit to give the same screen current for each tube, regardless of its location.

Note that the dc grid current has not been used as an indicator of balance of the RF power amplifier. It is probable that after following the foregoing procedure, the grid currents will be fairly well balanced, but this condition, in itself, is not a safe indicator of grid excitation balance.

11.3.3 Parallel Tube Amplifiers

The previous discussion has been oriented toward the RF push-pull amplifier. The same comments can be directed to parallel tube RF amplifiers [4]. Balance—the condition in which each tube carries its fair share of the load—still must be considered. In low-frequency amplifiers operating in class AB₁ or class AB₂, the idle dc plate current per tube should be balanced by separate bias adjustments for each tube. In many cases, some lack of balance of the plate currents will have negligible effect on the overall performance of the amplifier.

When tubes are operating in the idle position close to cutoff, an exact balance is difficult to achieve because plate current cannot be held to a close percentage tolerance. At this point, the action of the plate and screen voltages is in delicate balance with the opposing negative grid voltage. The state of balance is indicated by the plate current. Minor variations of individual grid wires or variations in the diameter of grid wires can upset the balance. It is practically impossible to control such minor variations in manufacture.

11.3.4 Harmonic Energy

A pulse of plate current delivered by the tube to the output circuit contains components of the fundamental and most harmonic frequencies [4]. To generate output power that is an harmonic of the exciting voltage applied to the control grid, it is merely necessary to resonate the plate circuit to the desired harmonic frequency. To optimize the performance of the amplifier, it is necessary to adjust the angle of plate current flow to maximize the desired harmonic. The shorter the length of the current pulse, in the case of a particular harmonic, the higher the plate efficiency. However, the bias, exciting voltage, and driving power must be increased. Also, if the pulse is too long or too short, the output power will decrease appreciably.

The plate circuit efficiency of tetrode and pentode harmonic amplifiers is quite high. In triode amplifiers, if feedback of the output harmonic occurs, the phase of the voltage feedback usually reduces the harmonic content of the plate pulse, and thereby lowers the plate circuit efficiency. Because tetrodes and pentodes have negligible feedback, the efficiency of a harmonic amplifier usually is comparable to that of other amplifier types. Also, the high amplification factor of a tetrode or pentode causes the plate voltage to have little effect on the flow of plate current. A well-designed tetrode or pentode also permits large RF voltages to be developed in the plate circuit while

still passing high peaks of plate current in the RF pulse. These two factors help to further increase plate efficiency.

11.3.4.1 Controlling Harmonics

The previous discussion of harmonics has been for the case where harmonic power in the load is the objective. The generation and radiation of harmonic energy must be kept to a minimum in a *fundamental frequency* RF amplifier [4]. The ability of the tetrode and pentode to isolate the output circuit from the input circuit over a wide range of frequencies is important in avoiding feedthrough of harmonic voltages from the grid. An important part of this shielding is the result of the basic physical design of the devices. The following steps permit reduction of unwanted harmonic energy in the output circuit:

- Minimize the circuit impedance between plate and cathode. This objective may be achieved by having some or all of the tuning capacitance of the resonant circuit close to the tube.
- Completely shield the input and output compartments of the stage.
- Use inductive output coupling from the resonant plate circuit, and possibly a capacitive or Faraday shield between the coupling coil and the tank coil. A high-frequency attenuating circuit, such as a *Pi* or *Pi-L* network, also may be used.
- Incorporate low-pass filters on all supply leads and wires leading to the output and input compartments.
- Use resonant traps for particular frequencies as required.
- Install a low-pass filter in series with the output transmission line.

11.3.5 Klystron Tuning Considerations

The output power stability of a klystron is sensitive to changes in the principal operating parameters, including

- Beam input power
- RF drive power
- RF drive frequency
- Device tuning
- Magnetic fields in the vicinity of the tube
- VSWR of the load

Varying any of these parameters will affect the RF output power response.

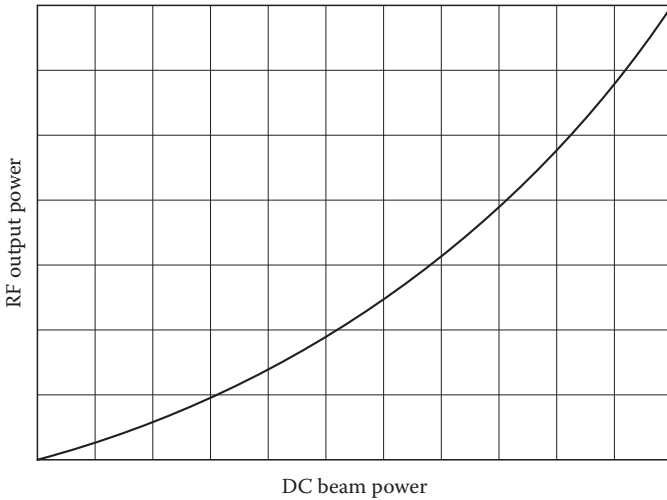


FIGURE 11.26
RF output power variation as a function of dc beam input power for a klystron.

Figure 11.26 illustrates how changes in the dc beam input power affect the RF output power of a klystron under constant RF drive conditions. Small changes in dc input power produce marked changes in RF output power. Figure 11.27 shows RF output power as a function of RF drive power applied to the tube. From this curve, it can be seen that when the RF drive power is low,

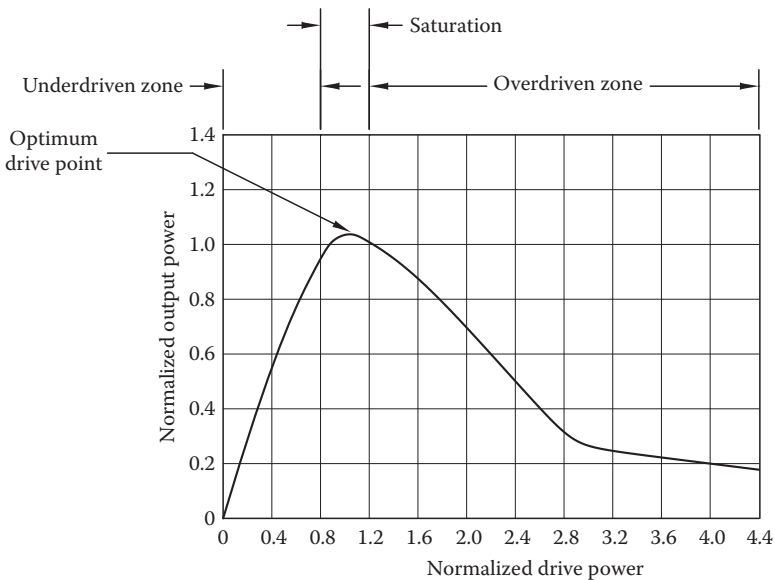


FIGURE 11.27
RF output power of a klystron as a function of RF drive power.

the RF output power is low. As the level of RF drive power increases, RF output power increases until an optimum operating point is reached. Beyond this point, further increases in RF drive power result in reduced RF output power.

Klystrons are said to be *saturated* at the point where a further increase in RF drive only decreases the RF output power. To obtain maximum RF output from a klystron, sufficient RF drive must be applied to the tube to reach the point of saturation. Operating at RF drive levels beyond saturation will only overdrive the klystron, decrease RF output power, and increase the amount of beam interception (body current) at the drift tube.

Tuning of any cavity of a multicavity klystron will affect the overall shape of the amplitude-response curve. There are two common methods of tuning klystrons:

- High-efficiency narrowband tuning
- Broadband tuning, yielding lower gain and efficiency

Figure 11.28 shows how RF output power changes with various levels of RF drive applied to a klystron under different tuning conditions. The RF power at point *A* represents the drive saturation point for a synchronously tuned tube. Point *B* shows a new point of saturation that is reached by tuning *the penultimate* (next to the last) cavity to a somewhat higher frequency. By tuning the penultimate cavity still further, point *C* is reached. At point *D*, increasing the penultimate cavity frequency no longer increases RF output power; instead, it reduces the power delivered by the tube (curve *E*).

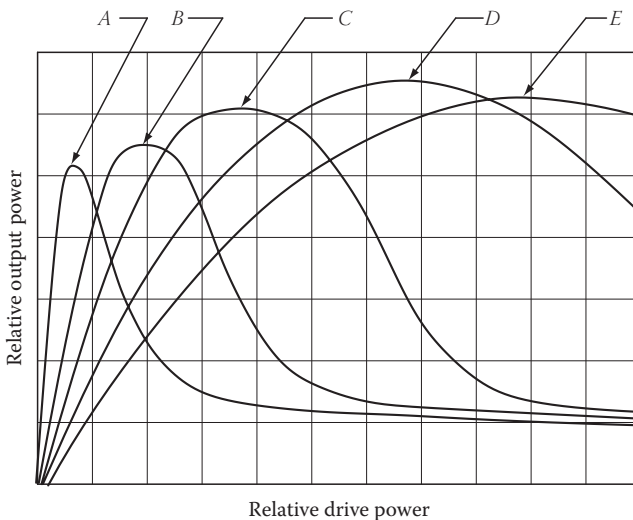


FIGURE 11.28

Output power variation of a klystron as a function of drive power under different tuning conditions.

11.3.6 Intermodulation Distortion*

IMD products in linear power amplifier circuits can be caused by either amplitude gain nonlinearity or phase shift with change in input signal level. IMD products appear when the RF signal has a varying envelope amplitude. A single continuous-frequency wave will be amplified by a fixed amount and shifted in phase by a fixed amount. The nonlinearity of the amplifier will produce only harmonics of the input wave. As the input RF wave changes, the nonlinearity of the amplifier will cause undesirable IMD.

When an RF signal with varying amplitude is passed through a nonlinear device, many new products are produced. The frequency and amplitude of each component can be determined mathematically. The nonlinear device can be represented by a power series expanded about the zero-signal operating point. IMD energy is wasted and serves no purpose other than to cause interference to adjacent channels. IMD in a power amplifier tube is principally the result of its transfer characteristics. An ideal transfer characteristic curve is shown in Figure 11.29.

Even-order products do not contribute to the IMD problem because they usually fall outside the amplifier passband. Therefore, if the transfer characteristic produces an even-order curvature at the small-signal end of the curve (from point A to point B) and the remaining portion of the curve (point B to point C) is linear, the tube is considered to have an ideal transfer characteristic. If the operating point of the amplifier is set at point 0 (located midway horizontally between point A and point B), there will be no distortion in a class

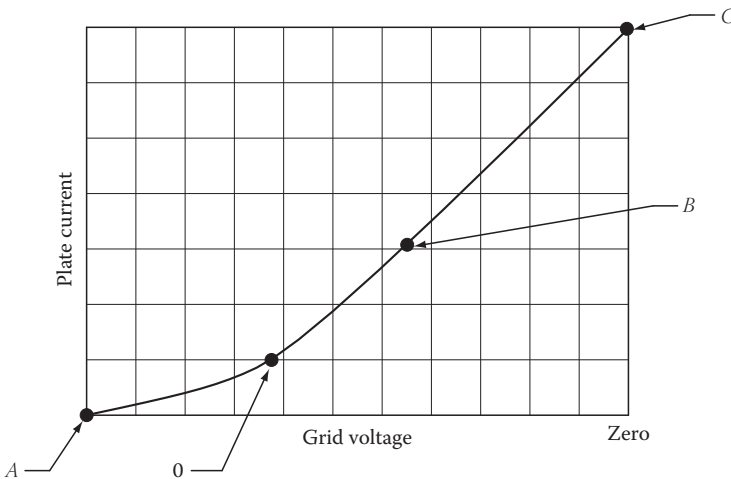


FIGURE 11.29
Ideal grid-plate transfer curve for class AB operation.

* Portions of this section were contributed by Geoffrey N. Mendenhall, P. E., Cincinnati, OH.

AB amplifier. However, no tube has this idealized transfer characteristic. It is possible, by manipulation of the electron ballistics within a given tube structure, to alter the transfer characteristic and minimize the distortion products.

11.3.6.1 Generation of Radio Frequency Intermodulation Products

When two or more transmitters are coupled to each other, new spectral components are produced by mixing of the fundamental and harmonic terms of each of the desired output frequencies.* The most common form of coupling results from transmitting antennas that are physically closely spaced. If sufficient filtering is not provided at the output stages and/or on one or both of the transmitter feedlines, energy from one transmitter may travel back to the final power amplifier of the other. In the PA stage, the desired and undesired signals mix, producing spurious IM products. Consider the following example involving two transmitters. The third-order IM products (IM3) are generated in the following manner:

- The output of transmitter 1 (f_1) is coupled into the nonlinear output stage of transmitter 2 (f_2). This condition is the result of insufficient isolation between the two output signals.
- Signal f_1 mixes with the second harmonic of f_2 , producing an in-band third-order term with a frequency of $2f_2 - f_1$ at the output of transmitter 2.
- In a similar manner, another third-order term will be produced at a frequency of $2f_1 - f_2$ at the output of transmitter 1.

This implies that the second-harmonic content within the output stage of each transmitter, along with the specific nonlinear characteristics of the output stage, will have an effect on the value of the mixing loss.

It is also possible to generate these same third-order terms in another way. If the difference frequency between the two transmitters ($f_2 - f_1$), which is an out-of-band frequency, remixes with either f_1 or f_2 , the same third-order IM frequencies will be produced.

Empirical measurements [5] indicate that the $2f_2 - f_1$ type of mechanism is the dominant mode generating third-order IM products in RF generators using a tuned cavity for the output network. Figure 11.30 illustrates the mechanism of third-order IM for the two-signal condition.

11.3.6.2 Intermodulation as a Function of Turnaround Loss*

Turnaround or *mixing loss* describes the phenomenon whereby the interfering signal mixes with the fundamental and its harmonics within the nonlinear

* Mendenhall (1983).

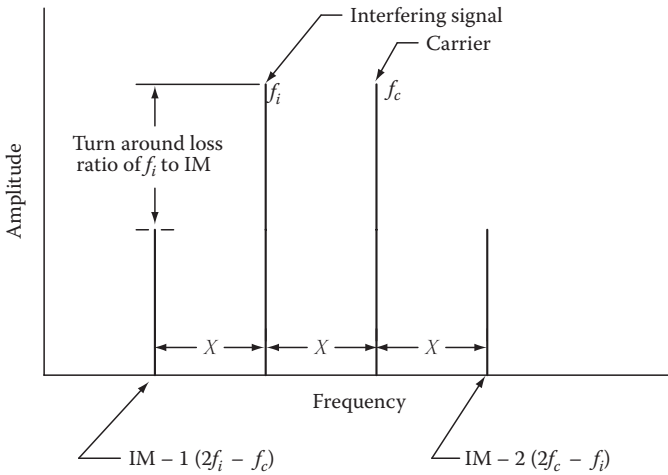


FIGURE 11.30 Frequency spectrum of third-order IM with the interfering level equal to the carrier level.

output device. This mixing occurs with a net conversion loss; hence, the term *turnaround loss* has become widely used to quantify the ratio of the interfering level to the resulting IM3 level. A turnaround loss of 10 dB means that the IM3 product fed back to the antenna system will be 10 dB below the interfering signal fed into the transmitter output stage.

Turnaround loss will increase if the interfering signal falls outside the passband of the transmitter output circuit, varying with frequency separation of the desired signal and the interfering signal. This is because the interfering signal is first attenuated by the output circuit selectivity going into the nonlinear device, then the IM3 product is further attenuated as it comes back through the frequency-selective circuit. Turnaround loss, thus, can be divided into three principal components:

- The basic in-band conversion loss of the nonlinear device
- Attenuation of the out-of-band interfering signal resulting from the selectivity of the output stage
- Attenuation of the resulting out-of-band IM3 products resulting from the selectivity of the output circuit

As the turnaround loss increases, the level of undesirable IM products is reduced, and the amount of isolation required between the transmitters also is reduced.

Transmitter output circuit loading directly affects the power amplifier source impedance and, therefore, also affects the efficiency of coupling the interfering signal into the output circuit where it mixes with the other frequencies present to yield IM3 products. Light loading reduces the amount of interference that

enters the output circuit with a resulting increase in turnaround loss. In addition, output loading will change the output circuit bandwidth (loaded Q) and, therefore, also affect the amount of attenuation that out-of-band signals will encounter in passing into and out of the output circuit.

Second-harmonic traps or low-pass filters in the transmission line of either transmitter have little effect on the generation of IM products. This is because the harmonic content of the interfering signal entering the output circuit of the transmitter has much less effect on IM3 generation than the harmonic content within the nonlinear device itself. The interfering signal and the resulting IM3 products thus fall within the passband of the low-pass filters and outside the reject band of the second-harmonic traps; these devices typically offer little, if any, attenuation to RF IM products.

Site-related IM problems and typical solutions are discussed later in this chapter.

11.4 Voltage Standing Wave Ratio

Among all the possible load measurements, VSWR is the most common. An in-line directional wattmeter commonly is used to measure both the forward and reflected power in a feedline. It is the comparison of reflected power (P_r) to forward power (P_f) that provides a qualitative measure of the match between a transmitter and its feedline/load.

The standing wave ratio (SWR) on a transmission line is related to the ratio of the forward to reflected power by the following:

$$\text{SWR} = \frac{1 + \sqrt{P_r/P_f}}{1 - \sqrt{P_r/P_f}} \quad (11.14)$$

The standing wave is the result of the presence of two components of power: one traveling toward the load and the other reflected by a load mismatch, traveling back toward the generator. These components are defined by the following:

$$P_f = \frac{E_f^2}{Z_o} = I_f^2(Z_o) \quad (11.15)$$

$$P_r = \frac{E_r^2}{Z_o} = I_r^2(Z_o) \quad (11.16)$$

$$P_n = (P_f - P_r) \quad (11.17)$$

where

P_f is the forward power

P_r is the reflected power

E_f is the forward voltage

E_r is the reflected voltage

I_f is the forward current

I_r is the reflected current

Z_o is the characteristic impedance of the line

P_n is the net power absorbed by the load (transmission line loss and antenna radiation)

Because the forward and reflected voltages and currents are traveling in opposite directions, they will add in phase at some point along the line to produce a voltage maximum. At $1/4$ wavelength along the line in either direction from this maximum, the forward and reflected components will be out of phase, and will produce a voltage minimum. The forward and reflected components of current also add vectorally to produce a current standing wave. The magnitude of this wave is defined by

$$SWR = \frac{E_{max}}{E_{min}} = \frac{I_{max}}{I_{min}} \quad (11.18)$$

The ratio of the highest voltage point on the line to the lowest voltage point on the line commonly is used to evaluate system performance. This property is known as the *voltage standing wave ratio*. A VSWR of 1.0:1.0 means that a perfect match has been achieved. A VSWR of 2.0:1.0 means that a mismatch is causing power to be reflected back from the load, approximately 11% in this particular case.

At the point of reflection (the load mismatch), the phase of the reflected current is reversed 180° from the forward current. The reflected voltage does not exhibit this phase reversal. This displaces the voltage and current standing waves by 90° along the line so that the E_{max} and I_{min} occur at the same points, while E_{min} and I_{max} occur 90° in either direction from E_{max} and I_{min} .

The fact that the reflected current is reversed in phase makes it possible to measure forward and reflected power separately using a directional coupler device. A small voltage is obtained by inductive coupling that represents the current in the transmission line. To this signal is added a sample of the voltage across the line, simultaneously obtained by capacitive coupling. These two samples are adjusted to be exactly equal when the line is terminated with its characteristic impedance. The two RF samples, when added, yield a resultant RF voltage proportional to the forward components of voltage and current. Reflected power is measured by installing a second coupling section physically turned in the opposite direction. The phase of the current sample is reversed, and the reflected components add while the forward components balance out.

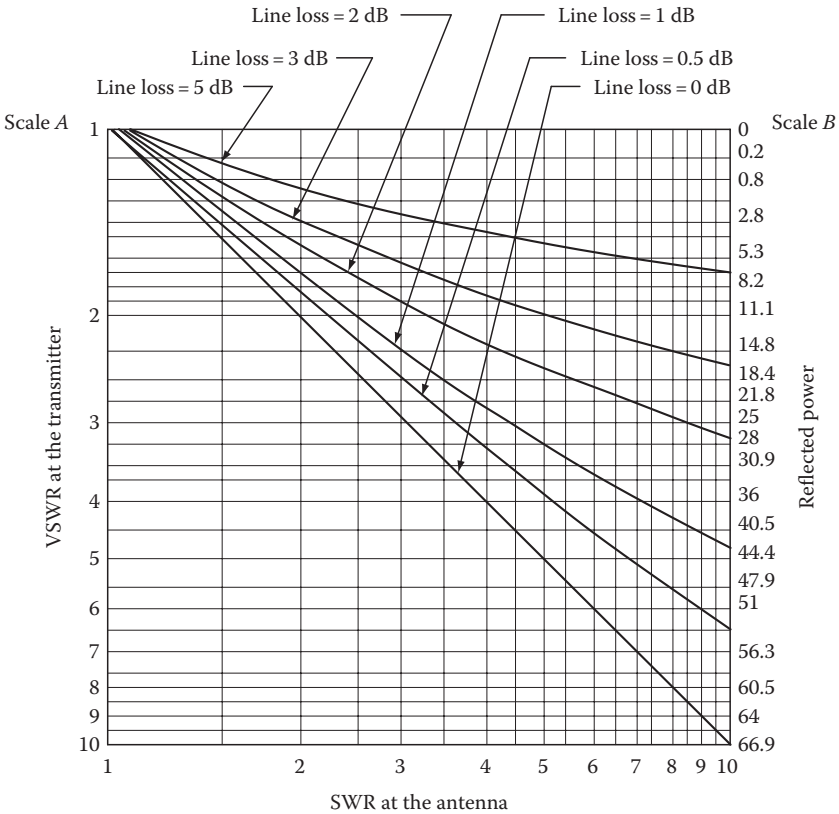


FIGURE 11.31 Illustration of how the VSWR on a transmission line affects the attenuation performance of the line. Scale A represents the VSWR at the transmitter; scale B represents the corresponding reflected power in watts for an example system.

On a theoretical, lossless feedline, VSWR is the same at all points along the line. Actually, all feedlines have some attenuation distributed uniformly along their entire lengths. Any attenuation between the reflection point and the source reduces VSWR at the source. This concept is illustrated in Figure 11.31. Standing waves on the feedline increase feedline attenuation to a level greater than normal attenuation under matched-line conditions. The amount of feedline attenuation increase depends on the VSWR and the extent of losses under matched-line conditions. The higher the matched-line attenuation, the greater the additional attenuation caused by a given VSWR. Figure 11.32 correlates additional attenuation with matched-line attenuation for various VSWR levels.

When a feedline is terminated by an impedance equal to the feedline characteristic impedance, the feedline impedance at any point is equal to the feedline characteristic impedance. When the feedline is terminated by an impedance not equal to the feedline characteristic impedance, the feedline input impedance will vary with changes in feedline length. The greater

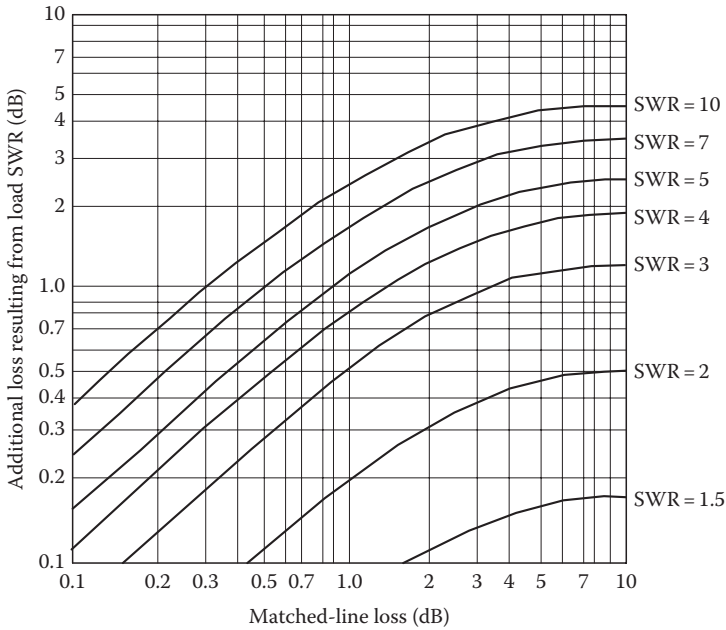


FIGURE 11.32 Illustration of how VSWR causes additional loss in a length of feedline.

the impedance mismatch at the load, the wider the feedline input impedance variation with changes in line length. The largest feedline impedance change occurs with the worst mismatches: open circuit or short circuit.

A short circuit at the end of a feedline appears as an open circuit at a point $1/4$ wavelength down the feedline. It is the greatest impedance transformation that can occur on a feedline. Less severe mismatches cause smaller impedance transformations. At a point $1/2$ wavelength down the line, the impedance equals the load impedance. Feedlines with high VSWR sometimes are called *resonant* or *tuned feedlines*. Matched lines—those with little VSWR—are referred to as *nonresonant* or *untuned feedlines*. This phenomenon means that high feedline VSWR can cause the feedline input to present a load impedance to the output tube that differs greatly from the feedline characteristic impedance. Some RF generators incorporate an output matching network that provides a *conjugate match* to a broad range of impedances.

11.5 Radio Frequency System Performance

Assessing the performance of an RF generator has been greatly simplified by recent advancements in system design. The overall performance of most transmitters can be assessed by measuring the audio, video, or data performance of

the system as a whole. If these “intelligence signal” tests indicate good performance, chances are good that all of the RF parameters also meet specifications. Practical transmission systems have certain limitations that prevent them from ever being a completely transparent medium. Through proper system design and regular equipment maintenance, however, overall performance can be maintained at acceptable levels in most operational environments.

11.5.1 Key System Measurements

The procedures for measuring transmitter performance vary widely, depending on the type of equipment being used. Certain basic measurements, however, can be applied to characterize the overall performance of most systems:

- Frequency response: the actual deviation from a constant amplitude across a given span of frequencies.
- THD: the creation, by a nonlinear device, of spurious signals harmonically related to the applied waveform. THD is sensitive to the noise floor of the system under test. For example, if the system has a S/N of 60 dB, the distortion analyzer’s best possible reading will be higher than 0.1% ($60 \text{ dB} = 0.001 = 0.1\%$).
- IMD: the creation, by a nonlinear device, of spurious signals not harmonically related to the input waveform. These distortion components are *sum-and-difference* (beat note) mixing products. The IM measurement is relatively impervious to the noise floor of the system under test.
- Signal-to-noise ratio (S/N): the amplitude difference, expressed in decibels, between a reference-level signal and the residual noise of the overall system.
- Cross talk: the amplitude difference, expressed in decibels, between two or more channels of a multichannel transmission system.

11.5.2 Synchronous AM in FM Systems*

The output spectrum of an FM transmitter contains many sideband frequency components, theoretically an infinite number. The practical considerations of transmitter design and frequency allocation make it necessary to restrict the bandwidth of FM signals. Bandwidth restriction brings with it a number of undesirable side effects, including

- Phase shifts through the transmission chain
- Generation of *synchronous AM* components
- Distortion in the demodulated output of certain receivers

* Mendenhall (1988).

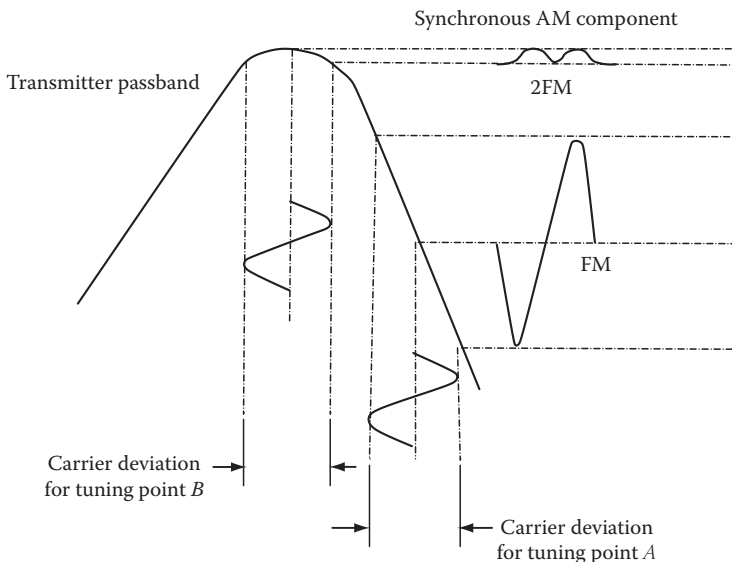


FIGURE 11.33 Generation of synchronous AM in a bandwidth-limited FM system. Note that minimum synchronous AM occurs when the system operates in the center of its passband.

In most medium- and high-power FM transmitters, the primary out-of-band filtering is performed in the output cavity of the final stage. Previous stages in the transmitter (the exciter and IPA) are designed to be broadband, or at least more broadband than the PA. As the bandwidth of an FM transmission system is reduced, synchronous AM increases for a given carrier deviation (modulation). Synchronous AM is produced as tuned circuits with finite bandwidth are swept by the frequency of modulation. The amount of synchronous AM generated is dependent on tuning, which determines (to a large extent) the bandwidth of the system. Figure 11.33 illustrates how synchronous AM is generated through modulation of the FM carrier.

11.5.2.1 Bandwidth Limiting

The design goal of most medium- and high-power FM transmission equipment is to limit the bandwidth of the transmitted RF signal at one stage only. In this way, control over system bandwidth can be tightly maintained, and the trade-offs required in any practical FM system can be optimized. If, on the other hand, more than one narrowband stage exists within the transmitter, adjustment for peak efficiency and performance can be a difficult proposition. The following factors can affect the bandwidth of an FM system:

- Total number of tuned circuits in the system
- Amplitude and phase response of the total combination of all tuned circuits in the RF path

- Amount of drive to each RF stage (saturation effects)
- Nonlinear transfer function within each RF stage

The following techniques can be used to improve the bandwidth of a given system:

- Maintain wide bandwidth until the final RF stage. This can be accomplished through the use of a broadband exciter and broadband (usually solid-state) IPA.
- Minimize the number of interactive-tuned networks within the system. Through proper matching network design, greater bandwidth and simplified tuning can be accomplished.

Bandwidth affects the gain and efficiency of any FM RF amplifier. The bandwidth is determined by the load resistance across the tuned circuit and the output or input capacitance of the RF stage. Because grid-driven PA tubes typically exhibit high input capacitance, the PA input circuit generally has the greatest effect on bandwidth limiting in an FM transmitter. The grid circuit can be broadened through resistive swamping in the input circuit, or through the use of a broadband input impedance matching network. One approach involves a combination of series inductor and shunt capacitor elements implemented on a printed wiring board, as illustrated in Figure 11.34. In this design, the inductors and capacitors are etched into the copper-clad laminate of the PWB. This approach provides the necessary impedance transformation from a 50 Ω solid-state driver to a high-impedance PA grid in a series of small steps, rather than more conventional *L*, *Pi*, or *Tee* matching networks.

11.5.2.2 Effects of Synchronous Amplitude Modulation

The effect of bandwidth on synchronous AM performance for an example FM RF amplifier is plotted in Figure 11.35. Notice that as the -3 dB points of the RF system passband are narrowed, a dramatic increase in synchronous AM occurs. The effect shown in the figure is applicable to any bandwidth-limited FM system.

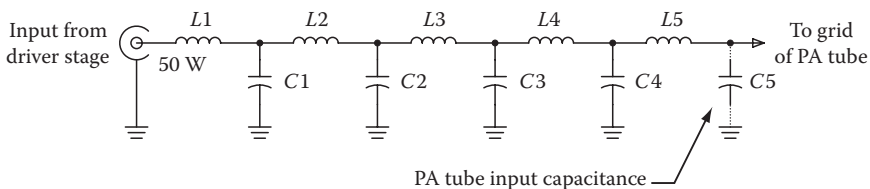


FIGURE 11.34

IPA plate-to-PA-grid matching network designed for wide bandwidth operation.

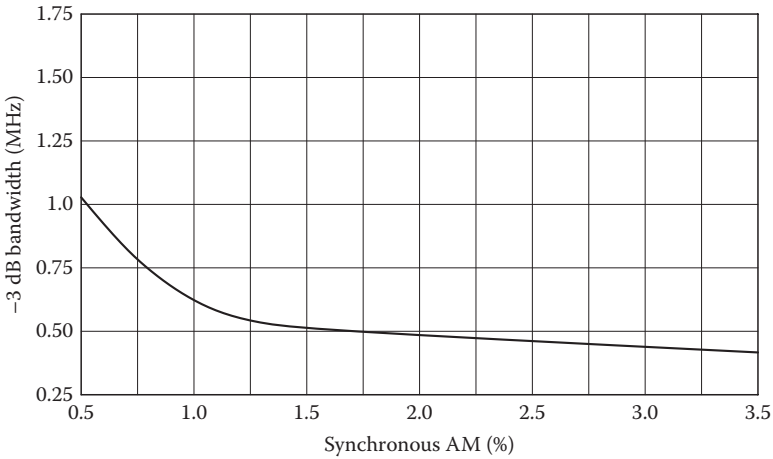


FIGURE 11.35 Synchronous AM content of an example FM system as a function of system bandwidth.

Bandwidth also affects the distortion floor of the demodulated signal of an FM system, as plotted in Figure 11.36. The data shown apply to a test setup involving a single-tuned circuit fed by an FM generator with a deviation of ± 75 kHz at a modulating frequency of 15 kHz (no de-emphasis is applied to the output signal). Remember that FM is a nonlinear process and that interpretation of distortion numbers for multiple-carrier signals is not accurate. The example shown, however, illustrates how the bandwidth of the RF channel can set a minimum performance limit on system THD. By understanding

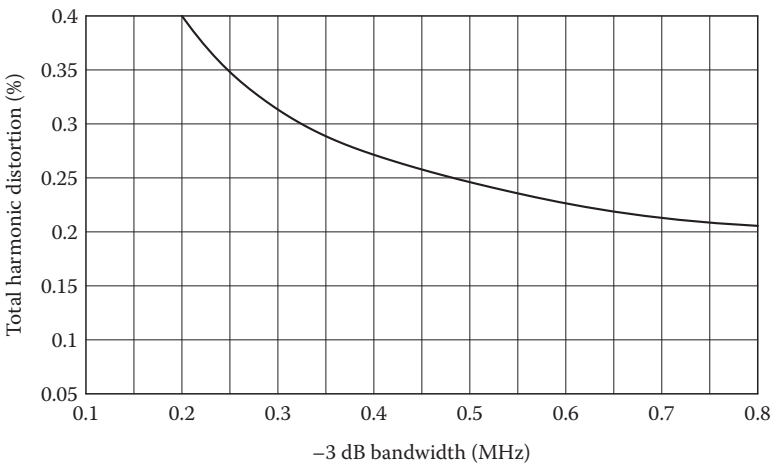


FIGURE 11.36 THD content of an example demodulated FM signal as a function of transmission-channel bandwidth.

the mechanics of synchronous AM, adjustment for optimum performance of an FM system can be accomplished.

11.5.3 Incidental Phase Modulation

Incidental phase modulation (IPM) is phase modulation produced by an AM transmitter as a result of AM. In other words, as an AM transmitter develops the amplitude-modulated signal, it also produces a phase-modulated, or PM version of the input intelligence. In theory, IPM is of little consequence for classic AM applications because FM and PM do not affect the carrier amplitude. An envelope detector will, in theory at least, ignore IPM. However, a number of applications utilize manipulation of the carrier phase to convey information in addition to AM of the carrier. Receivers designed to decode such AM/PM waveforms will not perform properly if excessive IPM is produced by the transmitter. Although the effects of IPM vary from one system to another, optimum performance is realized when IPM is minimized.

As a general rule, because IPM is a direct result of the modulation process, it can be generated in any stage that is influenced by modulation. The most common cause of IPM in plate-modulated and pulse-modulated transmitters is improper neutralization of the final RF amplifier. Adjusting the transmitter for minimum IPM is an accurate way of achieving proper neutralization. The reverse is not always true, however, because some neutralization methods will not necessarily result in the lowest IPM.

Improper tuning of the IPA stage is the second most common cause of IPM. As modulation changes the driver loading to the PA grid, the driver output also may vary. The circuits that feed the driver stage usually are sufficiently isolated from the PA that they do not produce IPM. An exception is if the power supply for the IPA is influenced by modulation. Such a problem could be caused by a lack of adequate capacitance in the filters of the high-voltage power supply.

11.5.3.1 Bandwidth Considerations

When a perfect AM transmitter is modulated with a tone, two sidebands are created, one above the carrier and one below it. PM of the same carrier, however, produces an infinite number of sidebands at intervals above and below the carrier. The number of significant sidebands depends on the magnitude of the modulation. With moderate amounts of IPM, the transmitted spectrum can be increased to the point at which the permissible channel limits are exceeded.

11.5.3.2 Corrective Procedures

The spectrum analyzer is generally the best instrument for identifying IPM components. The transmitter is first modulated with a pure sine wave, and a

THD measurement is taken with a distortion analyzer. These data will provide an idea of what the spectrum analyzer display should look like when it is connected to the transmitter. If the transmitter has an excessive amount of harmonic distortion, it will be impossible to tell the difference between the normal distortion sidebands and the IPM sidebands.

Figure 11.37 shows the trace of an ideal transmitted spectrum on a spectrum analyzer when modulated with a 1kHz tone at 100% modulation. Notice that there are just three components: one carrier and two sidebands. This transmitter would measure 0% distortion, because no harmonic sidebands are present. Also, no IPM is present, because it would show up as extra sidebands in the spectrum.

If the transmitter had an IPM problem, the spectrum analyzer display might look like Figure 11.38. The distortion analyzer still would read 0%, but

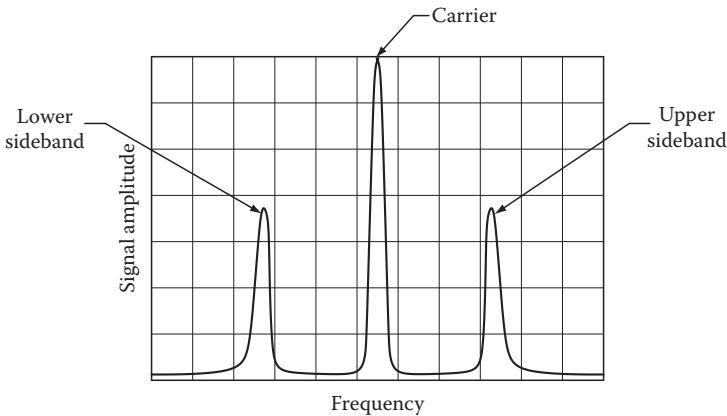


FIGURE 11.37 Spectrum of an ideal AM transmitter. THD measures 0%; there are no harmonic sidebands.

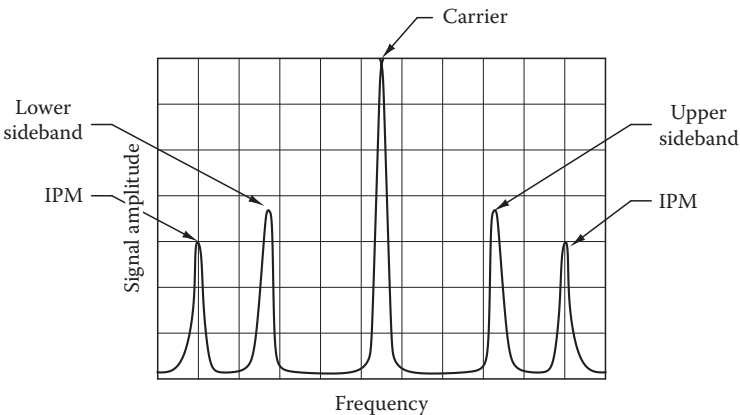


FIGURE 11.38 Spectrum of an ideal transmitter with the addition of IPM.

the sidebands shown on the spectrum analyzer would not agree. The key is to start with a low harmonic distortion reading on the transmitter so that any sidebands will be recognizable immediately as the result of IPM.

11.5.4 Carrier Amplitude Regulation

Carrier amplitude regulation (*carrier level shift*) is the effective shift in apparent carrier level as a result of the AM process [6]. Carrier level shift can be caused by the following:

- Poor power supply regulation
- Modulation-related even-order harmonic distortion that generates a dc offset component in the modulated RF envelope
- A combination of the previous two mechanisms

Carrier level shift can be either positive or negative, although is usually negative because power supply regulation is most often the major source of carrier level shift, and power supply regulation is generally negative in sign (i.e., lower voltage output at higher current loads). Strictly speaking, carrier level shift is defined as the shift in effective carrier voltage or current resulting from modulation (see Figure 11.39).

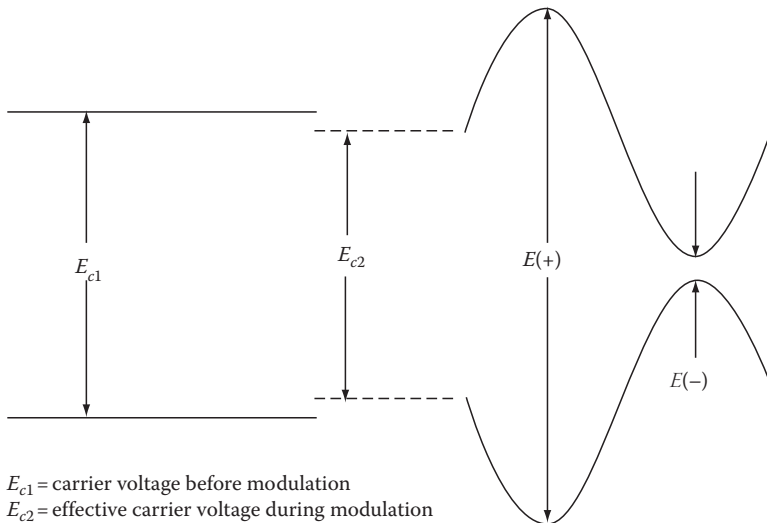


FIGURE 11.39

Carrier amplitude regulation (carrier level shift) defined. (After George, W., *AM transmitters*, in *NAB Engineering Handbook*, 9th edn., J.C. Whitaker (ed.), National Association of Broadcasters, Washington, DC, 1998. Used with permission.)

As illustrated in the figure,

$$CS = \frac{(E(+)+E(-))/2}{E_{c1}} \quad (11.19)$$

where CS equals carrier level shift. Percentage modulation is given by

$$\frac{E(+)-E(-)}{E(+)+E(-)} \times 100 \quad (11.20)$$

11.5.5 Site-Related Intermodulation Products

As discussed previously in this chapter, an IM product (intermod) is a signal created when two or more signals are unintentionally mixed to produce a new signal or group of signals. Mixing can occur in transmitter circuitry, in receiver circuitry, or elsewhere. The apparent effect is the same: one or more low-level products strong enough to interfere with desired signals and prevent proper reception. Sometimes the intermod falls on unused frequencies. Sometimes it falls on a channel in use or near enough to the channel to cause interference.

It is possible to calculate the intermod that may result from frequencies in use at a given transmission site. Normally produced by a computerized analysis, this *intermod study* is helpful in several ways. If the study is conducted as one or more new frequencies are under consideration for future use at an existing site, it may reveal which frequencies are more likely to produce a harmful intermod and which are not. If the study is conducted as part of an investigation into an interference problem, it may help in identifying transmitters that play a role in generating the intermod.

Intermod can be generated by any nonlinear device or circuit. The most common place for intermod to be generated is in the power amplifier of an FM transmitter. Such an amplifier does not need to preserve waveform integrity; the waveform is restored by the tuned output circuit. Thus, the amplifier is operated in an efficient but nonlinear mode, class C. In contrast, many amplifiers for AM service operate in a linear mode. Such amplifiers normally do not act as mixers.

Generally, solid-state amplifiers are better mixers than their electron tube counterparts, a fact that may explain why intermod may appear suddenly when older equipment is replaced. A mix occurs when RF energy from one transmitter (Tx 1) finds its way down the transmission line and into the final amplifier of another transmitter (Tx 2), as illustrated in Figure 11.40. The energy mixes in the final amplifier with the prefiltered second harmonic of Tx 2 to produce another in-band signal, a third-order intermod product. The reverse mix also may occur in the final amplifier of Tx 1, producing another intermod signal.

After the cause is discovered, a mix generated in a transmitter final amplifier is relatively easy to eliminate: increase the isolation between the

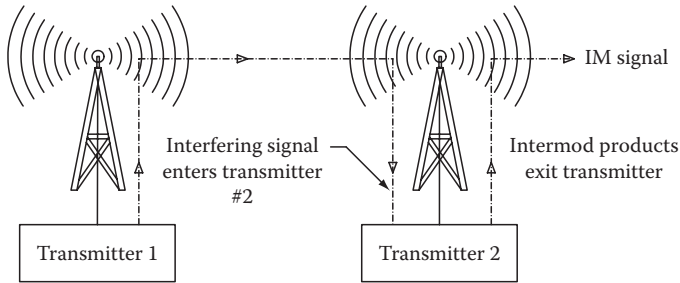


FIGURE 11.40
Mechanics of intermod generation at a two-transmitter installation.

offending transmitters. This may be accomplished by increasing the space between the antennas, or by placing an isolator at the output of each transmitter.

11.5.5.1 Receiver Intermod

Intermod also may be generated in the front end of a receiver. High-level RF energy from transmitters at a site can generate a bias voltage in the receiver sufficient to place the receiver RF amplifier into a nonlinear state. In such a mode, the stage acts as a mixer and creates unwanted frequencies. Filters may be used to solve this problem. The filters should notch out the offending transmitted signals. If another transmitter is added to the site, another notch filter usually must be added to the receiver transmission line.

11.5.5.2 Hardware Intermod

At a large transmission site, the entire area—including towers, guy wires, antennas, clamps, and obstruction lighting hardware—is saturated by high-level RF. Any loose or moving hardware can cause broadband noise. Corroded connections, tower joints, or clamps can form semiconductor junctions that may generate harmful intermod. Because the antenna system is exposed to a changing external environment, the degree of semiconductivity can vary with moisture or temperature. Changing semiconductivity causes intermittent problems, at changing levels of severity.

References

1. Cabot, R. C., Audio test and measurements, in *Audio Engineering Handbook*, K. Blair Benson (ed.), McGraw-Hill, New York, 1988.
2. Termaine, H. W., *Audio Cyclopedia*, Howard W. Sams, Indianapolis, IN, 1975.
3. Hulick, T. P., Very simple out-of-band IMD correctors for adjacent channel NTSC/DTV transmitters, *Digital Television '98*, Intertec Publishing, Overland Park, KS, 1998.

4. Laboratory Staff, *Care and Feeding of Power Grid Tubes*, Varian/Eimac, San Carlos, CA, 1984.
5. Mendenhall, G. N., M. Shrestha, and E. Anthony, FM broadcast transmitters, *NAB Engineering Handbook*, 8th edn., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, pp. 463–464, 1992.
6. George, W., AM transmitters, in *NAB Engineering Handbook*, 9th edn., J. C. Whitaker (ed.), National Association of Broadcasters, Washington, DC, 1998.

Bibliography

- Benson, K. B. and J. C. Whitaker, *Television and Audio Handbook for Engineers and Technicians*, McGraw-Hill, New York, 1989.
- Bordonaro, D., Reducing IPM in AM transmitters, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1988.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th edn., National Association of Broadcasters, Washington, DC, 1992.
- Dolgosh, Al. and R. Jakubowski, Controlling interference at communications sites, *Mobile Radio Technology*, Intertec Publishing, Overland Park, KS, February 1990.
- Hershberger, D. and R. Weirather, Amplitude bandwidth, phase bandwidth, incidental AM, and saturation characteristics of power tube cavity amplifiers for FM, Application note, Harris Corporation, Quincy, IL, 1982.
- Kinley, H., Base station feedline performance testing, *Mobile Radio Technology*, Intertec Publishing, Overland Park, KS, April 1989.
- Mendenhall, G. N., A study of RF intermodulation between FM broadcast transmitters sharing filterplexed or co-located antenna systems, *Broadcast Electronics*, Quincy, IL, 1983.
- Mendenhall, G. N., Testing television transmission systems for multichannel sound compatibility, *Broadcast Electronics*, Quincy, IL, 1985.
- Mendenhall, G. N., A systems approach to improving subcarrier performance, *Proceedings of the 1986 NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1986.
- Mendenhall, G. N., Fine tuning FM final stages, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1987.
- Mendenhall, G. N., Techniques for measuring synchronous AM noise in FM transmitters, *Broadcast Electronics*, Quincy IL, 1988.
- Mendenhall, G. N., Improving FM modulation performance by tuning for symmetrical group delay, *Broadcast Electronics*, Quincy IL, 1991.
- Terman, F. E., *Radio Engineering*, 3rd edn., McGraw-Hill, New York, 1947.
- Whitaker, J. C., *Maintaining Electronic Systems*, CRC Press, Boca Raton, FL, 1992.
- Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1990.
- Witte, R. A., Distortion measurements using a spectrum analyzer, *RF Design*, Cardiff Publishing, Denver, CO, pp. 75–84, September 1992.

12

RF System Maintenance and Troubleshooting

12.1 Introduction

Radio frequency (RF) equipment is unfamiliar to many persons entering the electronics industry. Colleges and universities do not routinely teach RF principles, favoring instead digital technology. Or, where RF principles are taught, they usually focus on low-power devices and systems. Unlike other types of products, however, RF equipment often must receive preventive maintenance to achieve its expected reliability. Maintaining RF gear is a predictable, necessary expense that facilities must include in their operating budgets. Tubes will have to be replaced from time to time no matter what the engineer does; components fail every now and then; and time must be allocated for cleaning and adjustments. By planning for these expenses each month, unnecessary downtime can be avoided.

Although the reason generally given for minimum RF maintenance is a lack of time and/or money, the cost of such a policy can be deceptively high. Problems that could be solved for a few dollars may, if left unattended, result in considerable damage to the system and a large repair bill. A standby system often can be a lifesaver; however, its usefulness sometimes is overrated. The best standby RF system is a *main system* in good working order.

12.2 Routine Maintenance

Most RF system failures can be prevented through regular cleaning and inspection and close observation. The history of the unit also is important in a thorough maintenance program so that trends can be identified and analyzed.

12.2.1 Maintenance Log

The control system front panel can tell a great deal about what is going on inside an RF generator. Develop a maintenance log and record all front-panel

meter readings, as well as the positions of critical tuning controls, on a regular basis. This information provides a history of the system and can be a valuable tool in noting problems at an early stage. The most obvious application of such logging is to spot failing power tubes, but other changes occurring in components can often be identified.

Creating a history of the line and tank pressure for a pressurized transmission line can identify developing line or antenna problems. After the regulator is set for the desired line pressure, record the tank and line readings each week and chart the data. If possible, make the observations at the same time of day each week. Ambient temperature can have a significant effect on line pressure; note any temperature extremes in the transmission line log when the pressure is recorded. The transmission line pressure usually will change slightly between carrier-on and carrier-off conditions (depending on the power level). The presence of RF can heat the inner conductor of the line, causing the pressure to increase. After a few months of charting the gradual loss of tank pressure, a pattern should become obvious. Investigate any deviation from the normal amount of tank pressure loss over a given period.

Whenever a problem occurs with the RF system, make a complete entry describing the failure in the maintenance log. Include a description of all maintenance activities required to return the system to operational condition. Make all entries complete and clear. Include the following data:

- Description of the nature of the malfunction, including all observable symptoms and performance characteristics
- Description of the actions taken to return the system to a serviceable condition
- Complete list of the components replaced or repaired, including the device schematic number and part number
- Total system downtime as a result of the failure
- Name of the engineer who made the repairs

The importance of regular, accurate logging can best be emphasized through the following examples.

12.2.1.1 Case Study #1

Improper neutralization is detected on an AM (amplitude modulation) broadcast transmitter IPA (intermediate power amplifier), shown in Figure 12.1. The neutralization adjustment is made by moving taps on a coil, and none have been changed. The history of the transmitter (as recorded in the maintenance record) reveals, however, that the power amplifier (PA) grid tuning adjustment has, over the past 2 years, been moving slowly into the higher readings. An examination of the schematic diagram leads to the conclusion that C-601 is the problem.

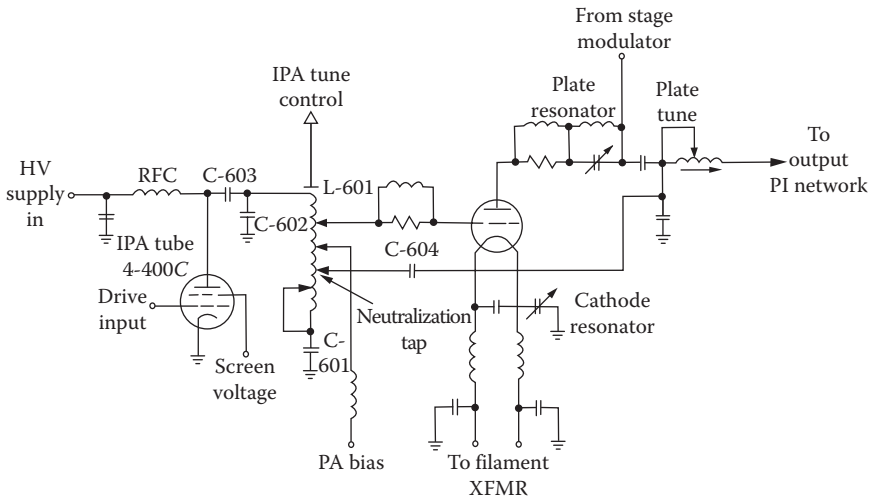


FIGURE 12.1

AM transmitter IPA/PA stage exhibiting neutralization problems. A history of IPA returning (through adjustment of L-601) helped determine that loss of neutralization was the result of C-601 changing in value.

The tuning change of the stage was so gradual that it was not thought significant until an examination of the transmitter history revealed that continual retuning in one direction only was necessary to achieve maximum PA grid drive. Without a record of the history of the unit, time could have been wasted in substituting expensive capacitors in the circuit, one at a time. Worse yet, the engineer might have changed the tap on coil L-601 to achieve neutralization, further hiding the real cause of the problem.

12.2.1.2 Case Study #2

An ultrahigh frequency (UHF) broadcast transmitter is found to exhibit decreasing klystron body current. The typical reading with average picture content is 50 mA, but over a 4 week period, the reading dropped to 30 mA. No other parameters show deviation from normal. Yet, the decrease in the reading indicates an alternate path (besides the normal body-current circuitry) by which electrons return to the beam power supply. A schematic diagram of the system is shown in Figure 12.2. Several factors could cause the body-current variation, including water leakage into the body-to-collector insulation of the klystron. In time, this water can corrode the klystron envelope, possibly leading to a loss of vacuum and klystron failure.

Water leakage can also cause partial bypassing of the body-current circuitry, an important protection system in the transmitter. It is essential that the circuit functions normally at all times and at full sensitivity in order to detect change when a fault condition occurs. Regular logging of transmitter parameters ensures that developing problems are caught early.

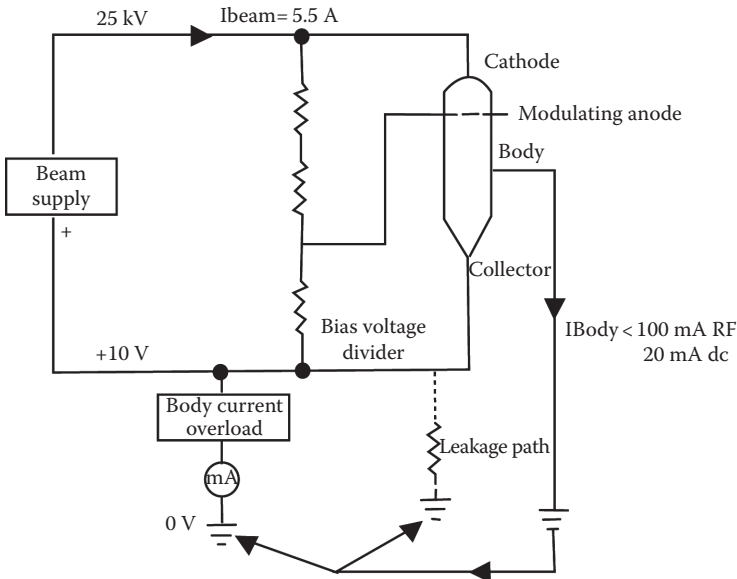


FIGURE 12.2

Simplified high-voltage schematic of a klystron amplifier showing the parallel leakage path that can cause a reduction in protection sensitivity of the body-current circuit.

12.2.2 Preventive Maintenance Routine

A thorough inspection of the RF system on a regular basis is the key to minimizing equipment downtime. Component problems often can be spotted at an early stage by regular inspection of the system. Remember to discharge all capacitors in the circuit with a grounding stick before touching any component in the high-voltage sections. Confirm that all primary power has been removed before any maintenance work begins.

Special precautions must be taken with systems that receive ac power from two independent feeds. Typically, one ac line provides 208 V three-phase service for the high-voltage section of the system, and a separate ac line provides 120 V power for low-voltage circuits. Older transmitters or high-power transmitters often utilize this arrangement. Check to see that all ac is removed before any maintenance work begins.

Consider the following preventive maintenance procedures:

12.2.2.1 Resistors and Capacitors

- Inspect resistors and RF capacitors for signs of overheating (see Figure 12.3).
- Inspect electrolytic or oil-filled capacitors for signs of leakage.
- Inspect feedthrough capacitors and other high-voltage components for signs of arcing (see Figure 12.4).

**FIGURE 12.3**

Check resistors, particularly high-power units, regularly for signs of premature wear caused by excessive heating.

**FIGURE 12.4**

Inspect high-voltage capacitors for signs of leakage around the case feedthrough terminals.

Transmitting capacitors—mica vacuum and *doorknob* types—should not run hot. They may run warm, but usually as the result of thermal radiation from nearby components (such as power tubes) in the circuit. An overheated transmitting capacitor is often a sign of incorrect tuning. Vacuum capacitors present special requirements for the maintenance technician. Care in

handling is a prime requisite for maximum service life. Because the vacuum capacitor is evacuated to a higher degree than most vacuum tubes, it is particularly susceptible to shock and rough handling. Provide adequate protection to vacuum capacitors whenever maintenance is performed. The most vulnerable parts of the capacitor are the glass or ceramic-to-metal seals on each end of the unit. Exercise particular care during removal or installation.

The current ratings of vacuum capacitors are limited by the glass-to-metal seal temperature and the temperature of the solder used to secure the capacitor plates. Seal temperature is increased by poor connecting clip pressure, excessive ambient temperatures, corrosion of the end caps and/or connecting clip, excessive dust and dirt accumulation, or excessive currents. Dust accumulation on sharp points in high-voltage circuitry near the vacuum capacitor can cause arcs or corona that may actually burn a hole through the glass or ceramic envelope.

12.2.2.2 Power Supply Components

- Inspect the mechanical operation of circuit breakers. Confirm that they provide a definite *snap* to the off position (remove all ac power for this test) and that they firmly reseal when restored. Replace any circuit breaker that is difficult to reset. This check can be done on an occasional basis. It usually does not need to be part of a weekly or monthly inspection.
- Inspect power transformers and reactors for signs of overheating or arcing (see Figure 12.5).
- Inspect oil-filled transformers for signs of leakage.
- Inspect transformers for dirt buildup, loose mounting brackets and rivets, and loose terminal connections.
- Inspect high-voltage rectifiers and transient suppression devices for overheating and mechanical problems (see Figure 12.6).

Power transformers and reactors normally run hot to the touch. Check both the transformer frame and the individual windings. On a three-phase transformer, each winding should produce about the same amount of heat. If one winding is found to run hotter than the other two, further investigation is warranted. Dust, dirt, or moisture between terminals of a high-voltage transformer may cause flashover failures. Insulating compound or oil around the base of a transformer indicates overheating or leakage.

12.2.2.3 Coils and RF Transformers

- Inspect coils and RF transformers for indications of overheating (see Figure 12.7).
- Inspect connection points for arcing or loose terminals.

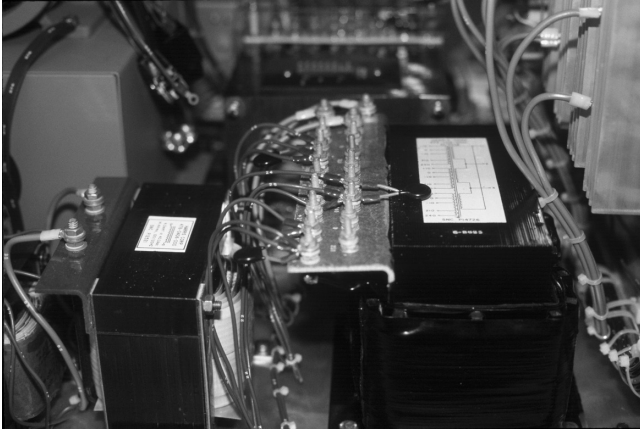


FIGURE 12.5
Inspect power transformers just after shutdown for indications of overheating, or leakage in the case of oil-filled transformers.

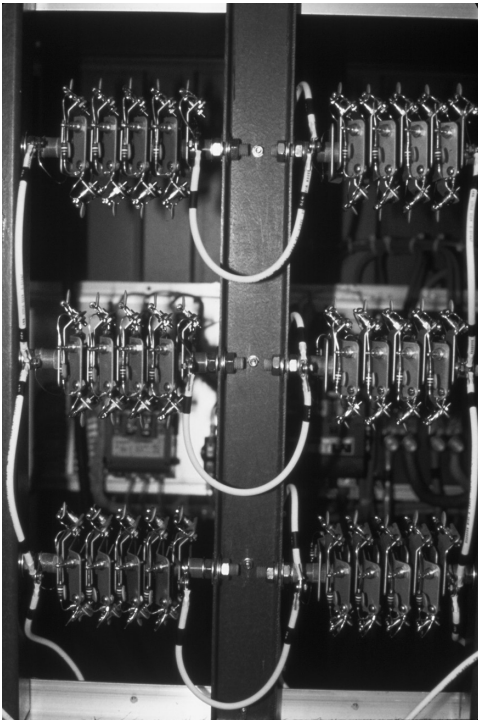


FIGURE 12.6
Check the heating of individual rectifiers in a stack assembly. All devices should generate approximately the same amount of heat.

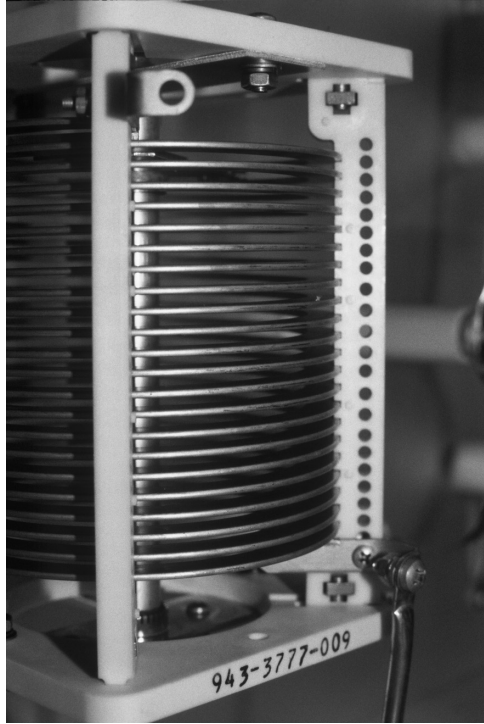


FIGURE 12.7

Clean RF coils and inductors as often as needed to keep contaminants from building up on the device loops.

Coils and RF transformers operating in a well-tuned system will rarely heat appreciably. If discoloration is noticed on several loops of a coil, consult the factory service department to see if the condition is normal. Pay particular attention to variable tap inductors. Closely inspect the roller element and coil loops for overheating or signs of arcing.

12.2.2.4 Relay Mechanisms

- Inspect relay contacts, including high-voltage or high-power RF relays, for signs of pitting or discoloration.
- Inspect the mechanical linkage to confirm proper operation. The contactor arm (if used) should move freely, without undue mechanical resistance.
- Inspect vacuum contactors for free operation of the mechanical linkage (if appropriate) and for indications of excessive dissipation at the contact points and metal-to-glass (or metal-to-ceramic) seals.

Unless problems are experienced with an enclosed relay, do not attempt to clean it. More harm than good can be done by disassembling properly working components for detailed inspection.

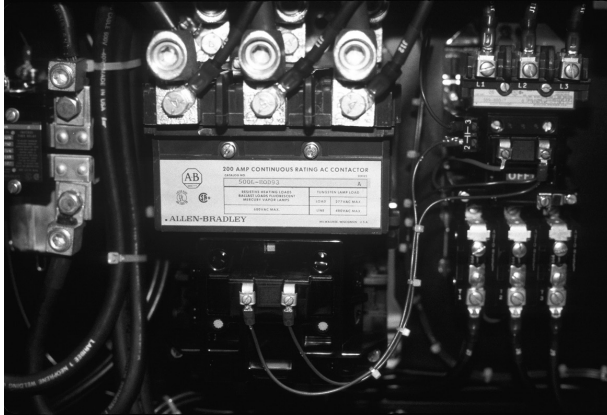


FIGURE 12.8
Check the tightness of connectors on an occasional basis, but do not stress the connection points.

12.2.2.5 Connection Points

- Inspect connections and terminals that are subject to vibration. Tightness of connections is critical to the proper operation of high-voltage and RF circuits (see Figure 12.8).
- Inspect barrier strip and printed circuit board contacts for proper termination.

Although it is important that all connections are tight, be careful not to over-tighten. The connection points on some components, such doorknob capacitors, can be damaged by excessive force. There is no section of an RF system where it is more important to keep connections tight than in the PA stage. Loose connections can result in arcing between components and conductors that can lead to system failure. For cavity PA systems, the access door is a part of the outer conductor of the coaxial transmission line circuit. High-potential RF circulating currents flow along the inner surface of the door, which must be fastened securely to prevent arcing.

12.2.3 Cleaning the System

Cleaning is a large part of a proper maintenance routine. A shop vacuum and clean brush are generally all that are required. Use isopropyl alcohol and a soft cloth for cleaning insulators on high-voltage components (see Figure 12.9). Cleaning affords the opportunity to inspect each component in the system and observe any changes.

Regular maintenance of insulators is important to the proper operation of RF final amplifier stages because of the high voltages usually present. Pay particular attention to the insulators used in the PA tube socket (see Figure 12.10). Because the supply of cooling air is passed through the socket,



FIGURE 12.9

Keep all high-voltage components, such as this rectifier bank, free of dust and contamination that might cause short-circuit paths.

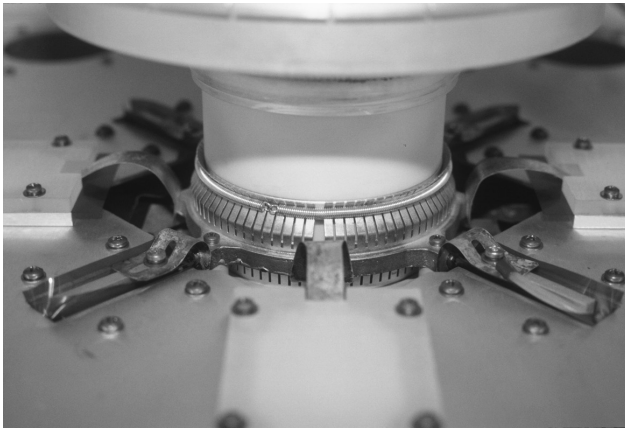


FIGURE 12.10

Carefully inspect the PA tube socket assembly. Do not remove the PA tube unless necessary.

airborne contaminants can be deposited on various sections of the assembly. These can create a high-voltage arc path across the socket insulators. Perform any cleaning work around the PA socket with extreme care. Do not use compressed air to clean out a power tube socket. Blowing compressed air into the PA or IPA stage of a transmitter will merely move the dirt from places where you can see it to places where you cannot see it. Use a vacuum instead.

When cleaning the socket assembly, do not disturb any components in the PA circuit (see Figure 12.11). Visually check the tube anode to see if dirt is clogging any of the heat-radiating fins.



FIGURE 12.11
Clean the PA cavity assembly to prevent an accumulation of dust and dirt. Check hardware for tightness.

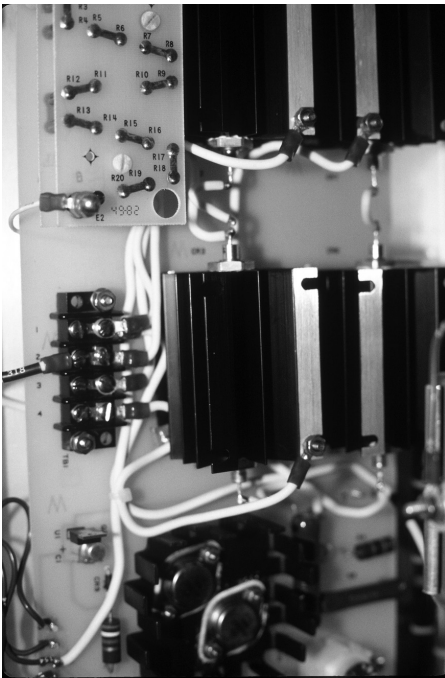


FIGURE 12.12
Clean power semiconductor assemblies to ensure efficient dissipation of heat from the devices.

Cleaning is also important to proper cooling of solid-state components in the transmitter. A layer of dust and dirt can create a thermal insulator effect and prevent proper heat exchange from a device into the cabinet (see Figure 12.12).

12.3 Transmission Line and Antenna

The voltage standing wave ratio (VSWR) of an antenna and its transmission line is a vital parameter that has a considerable effect on the performance and reliability of a transmission system. VSWR is a measure of the amount of power reflected back to the transmitter because of an antenna and/or transmission line mismatch. Figure 12.13 provides a chart for VSWR calculation. A mismatched or defective transmission system will result in a high degree of reflected power or a higher VSWR.

The amount of reflected power that a given system can accept is a function of the application. For example, it is common practice in FM (frequency modulation) broadcasting to maintain a VSWR of 1.1:1 as the maximum level within the transmission channel that can be tolerated without degrading the quality of the on-air signal. For analog TV broadcasting, a VSWR into the antenna feeder of more than 1.04:1 will start to degrade picture quality, particularly on systems that use a long transmission line. Reflections down the line from a mismatch at the antenna disrupt the performance of the transmitter output stage. When power is reflected back to the transmitter, it causes the RF output stage to look into a mismatched load with unpredictable phase and impedance characteristics. Because of the reflective nature of VSWR on a transmission system, the longer the transmission line (assuming the reflection is originating at the antenna), the more severe the problem may be for a given VSWR. A longer line means that reflected power seen at the RF output stage has greater time (phase) delays, increasing the reactive nature of the load.

VSWR is affected not only by the rating of the antenna and transmission line as individual units, but also by the combination of the two as a system. The worst-case system VSWR is equal to the antenna VSWR multiplied by the transmission line VSWR. For example, if an antenna with a VSWR of 1.05:1 is connected to a line with a VSWR of 1.05:1, the resulting worst-case system VSWR would be 1.1025:1. Given the right set of conditions, an interesting phenomenon can occur in which the VSWR of the antenna cancels the transmission line VSWR, resulting in a perfect 1:1 match. The determining factors for this condition are the point of origin of the antenna VSWR, the length of transmission line, and the observation point.

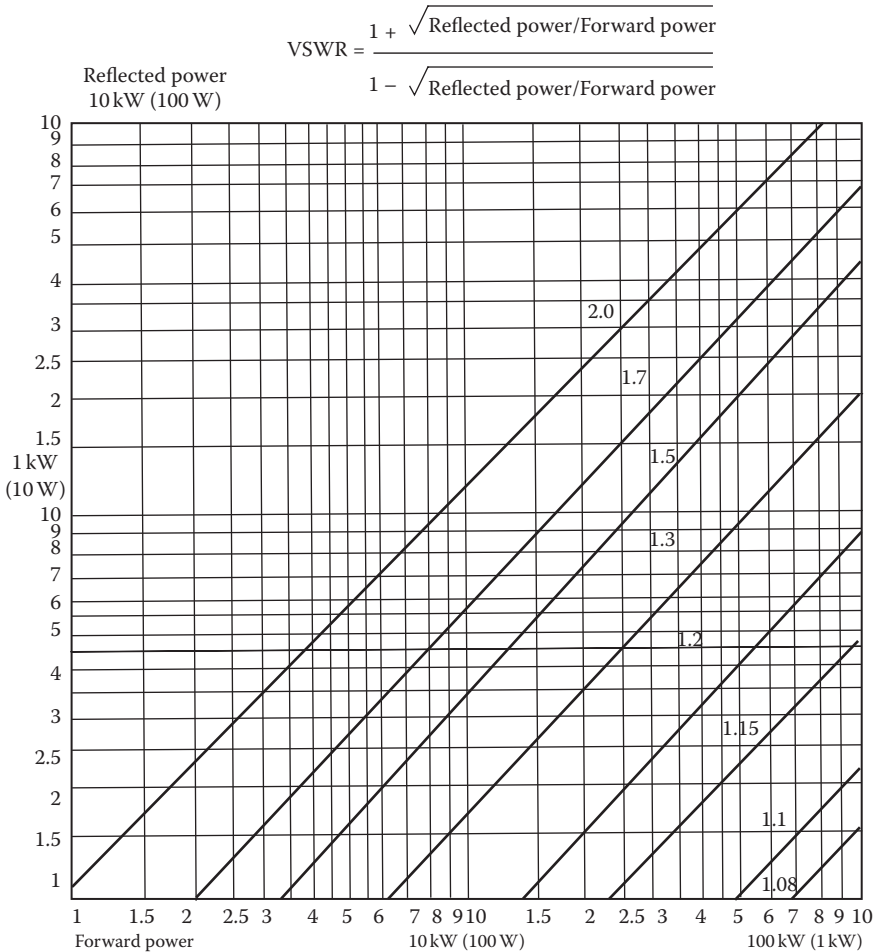


FIGURE 12.13 Transmission line VSWR graph. For low-power operation, use the values in parenthesis.

12.3.1 Effects of Modulation

The VSWR of a transmission system is a function of frequency and changes with carrier modulation. This change may be large or small, but it will occur to some extent. The cause can be traced to the frequency dependence of the VSWR of the antenna (and to a lesser extent, the transmission line.) The effects of frequency on VSWR can be observed in Figure 12.14. Although the plot of VSWR vs. frequency for a common FM antenna is good, notice that with no modulation the system VSWR is one figure. VSWR measurements are different when modulation is applied.

VSWR is further complicated because power reflected back to the transmitter from the antenna may not come from a single point, but from a number

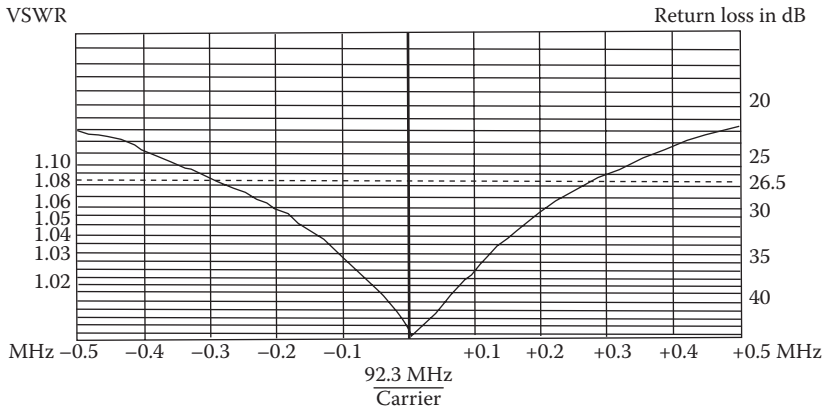


FIGURE 12.14

Measured performance of a single-channel FM antenna (tuned to 92.3 MHz). The antenna provides a VSWR of below 1.1:1 over a frequency range of nearly ± 300 kHz. (Courtesy of Jampro Antenna Company, Sacramento, CA.)

of different points. One reflection might be caused by the antenna-matching unit, another by various flanges in the line, and a third by a damaged part of the antenna system. Because these reflection points are different lengths from the transmitter PA plate, a variety of standing waves can be generated along the line, varying with the modulating frequency.

Energy reflected back to the transmitter from the antenna is not lost. A small percentage of the energy is turned into heat, but the majority of it is radiated by the antenna, delayed in time by length of the transmission line.

12.3.2 Maintenance Considerations

In order to maintain low VSWR, the transmission line and antenna system should be serviced regularly. Use the following guidelines:

- Inspect the antenna elements, interconnecting cables, impedance transformers and support braces at least once each year. Falling ice can damage antenna elements if proper precautions are not taken. Icing on the elements of an antenna will degrade the antenna VSWR because ice lowers the frequency of the electrical resonance of the antenna. Two methods are commonly used to prevent a buildup of ice on high-power transmitting antennas: electrical de-icers and *radomes*.
- Check the antenna tower regularly for structural integrity. In the case of an AM station, the tower itself is the radiator and so each section must be bonded for good electrical contact.
- Clean base insulators and guy insulators (if used) as often as required.
- Keep lighting ball gaps or other protective devices clean and properly adjusted.

- Inspect the transmission line for signs of damage. Check supporting hardware and investigate any indication of abnormal heating of the line.
- Keep a detailed record of VSWR in the facility maintenance log and investigate any increase above the norm.
- Regularly check the RF system test load. Inspect the coolant filters and flow rate, as well as the resistance of the load element.

12.4 High-Voltage Power Supply

The high-voltage plate supply is the first line of defense between external ac line disturbances and the PA stage(s). Next to the output circuit itself, the plate supply is the second section of a transmitter most vulnerable to damage because of outside influences.

12.4.1 Power Supply Maintenance

Figure 12.15 shows a high-reliability power supply of the type common in transmission equipment. Many transmitters use simpler designs, without some of the protection devices shown, but the principles of preventive maintenance are the same:

- Thoroughly examine every component in the high-voltage power supply. Look for signs of leakage on the main filter capacitor (C2).
- Check all current-carrying meter/overload shunt resistors (R1–R3) for signs of overheating.
- Carefully examine the wiring throughout the power supply for loose connections.
- Examine the condition of the filter capacitor series resistors (R4 and R5), if used, for indications of overheating. Excessive current through these resistors could point to a pending failure in the associated filter capacitor.
- Examine the condition of the bleeder resistors (R6–R8). A failure in one of the bleeder resistors could result in a potentially dangerous situation for maintenance personnel by leaving the main power supply filter capacitor (C2) charged after the removal of ac input power.
- Examine the plate voltage meter multiplier assembly (A1) for signs of resistor overheating. Replace any resistors that are discolored with the factory-specified type.

When changing components in the transmitter high-voltage power supply, be certain to use parts that meet with the approval of the manufacturer. Do not

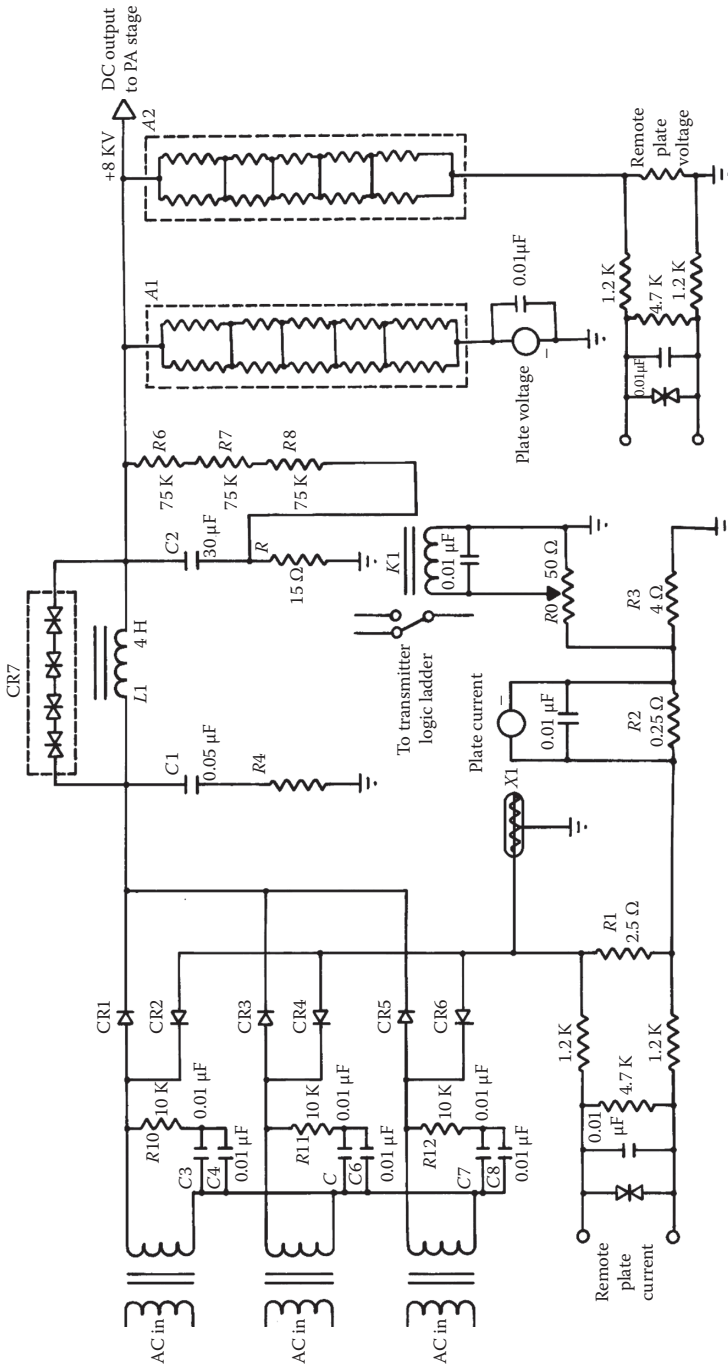


FIGURE 12.15
Common high-voltage transmitter power supply circuit design.

settle for a close match of a replacement part. Use the exact replacement part. This ensures that the component will work as intended and will fit in the space provided in the cabinet.

12.4.2 Power Supply Metering

Proper metering is one of the best ways to prevent failures in transmission equipment. Accurate readings of plate voltage and current are fundamental to RF system maintenance. Check each meter for proper mechanical and electrical operation. Replace any meter that sticks or will not zero.

With most transmitter plate current meters, accuracy of the reading can be verified by measuring the voltage drop across the shunt element ($R2$ of the previous figure) and using Ohm's law to determine the actual current in the circuit. Be certain to take into consideration the effects of the meter coil itself. Contact the transmitter manufacturer for suggestions on how best to confirm the accuracy of the plate current meter.

The plate voltage meter can be checked for accuracy by using a high-voltage probe and a high-accuracy external voltmeter. Be extremely careful when making such a measurement. Follow instructions regarding use of the high-voltage probe to the letter. Do not defeat transmitter interlocks to make this measurement. Instead, fashion a secure connection to the point of measurement and route the meter cables carefully out of the transmitter. Never use common test leads to measure a voltage of more than 600 V. Standard test lead insulation for most meters is not rated for use above 600 V.

12.4.3 Overload Sensor

The simplest type of plate supply overload sensor is arranged as shown in Figure 12.15. An adjustable resistor—either a fixed resistor with a movable tap or a potentiometer—is used to set the sensitivity of the plate overload relay. Check potentiometer-type adjustments periodically. Fixed-resistor-type adjustments rarely require additional attention. Most manufacturers have a chart or mathematical formula that may be used to determine the proper setting of the adjustment resistor ($R9$) by measuring the voltage across the overload relay coil ($K1$) and observing the operating plate current value. If mechanical problems are encountered with a relay, replace it.

Transmitter control logic for a high-power system is usually configured for two states of operation:

- An operational level, which requires all the “life-support” systems to be present before the high-voltage (HV) command is enabled
- An overload level, which removes HV when one or more fault conditions occur

Inspect the logic ladder for correct operation during any major maintenance work.

12.4.4 Transient Disturbances

Every electronic installation requires a steady supply of clean power in order to function properly. Recent advances in technology have made the question of ac power quality even more important, as microcontrollers are integrated into transmission equipment. Different types and makes of transmitters have varying degrees of transient overvoltage protection.

Figure 12.15 shows surge suppression at two points in the power supply circuit. C1 and R4 make up an R–C snubber network that is effective in shunting high-energy, fast-rise time spikes that may appear at the output of the rectifier assembly (CR1–CR6). Similar R–C snubber networks (R10–R12 and C3–C8) are placed across the secondary windings of each section of the three-phase power transformer. Any signs of resistor overheating or capacitor failure are an indication of excessive transient activity on the ac power line. Transient disturbances should be suppressed before the ac input point of the transmitter.

Assembly CR7 is a surge suppression device that should be given careful attention during each maintenance session. CR7 is typically a selenium thyrector assembly that is essentially inactive until the voltage across the device exceeds a predetermined level. At the trip point, the device will break over into a conducting state, shunting the transient overvoltage. CR7 is placed in parallel with L1 to prevent damage to other components in the transmitter in the event of a loss of RF excitation to the final stage. A sudden drop in excitation will cause the stored energy of L1 to be discharged into the power supply and PA circuits in the form of a high-potential pulse. Such a transient can damage or destroy filter, feedthrough, or bypass capacitors; damage wiring; or cause PA tube arcing. CR7 prevents these problems by dissipating the stored energy in L1 as heat. Investigate discoloration or other outward signs of damage to CR7. Such an occurrence could indicate a problem in the exciter or IPA stage of the transmitter. Immediately replace CR7 if it appears to have been stressed.

Check spark gap surge suppressor X1 periodically for signs of overheating. X1 is designed to prevent damage to circuit wiring in the event that one of the meter/overload shunt resistors (R1–R3) opens. Because the spark gap device is nearly impossible to accurately test in the field and is relatively inexpensive, it may be an advisable precautionary measure to replace the component every few years.

12.4.5 Single Phasing

Any transmitter using a three-phase ac power supply is subject to the problem of *single-phasing*—the loss of one of the three legs from the primary ac power distribution source. Single phasing is usually a utility company problem, caused by a downed line or a blown pole-mounted fuse. The loss of one leg of a three-phase line results in a particularly dangerous situation for three-phase motors, which will overheat and sometimes fail.

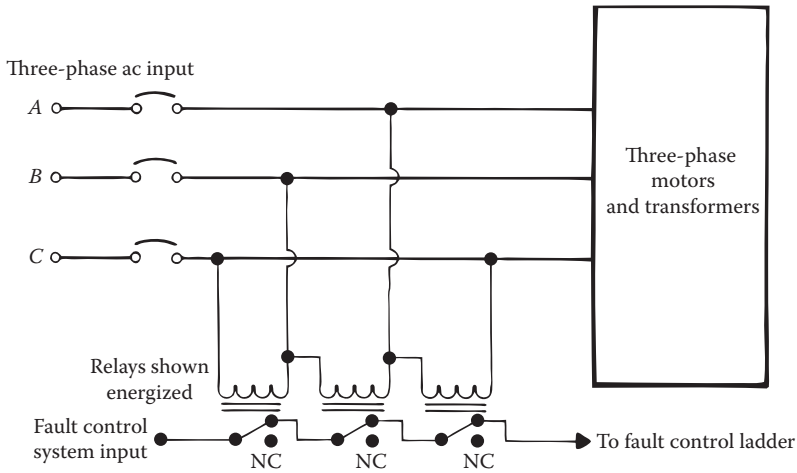


FIGURE 12.16

Protection circuit using relays for utility company ac phase loss protection.

Figure 12.16 shows a simple protection scheme that has been used to protect transmission equipment from damage caused by single phasing. Although at first glance the system looks as if it would easily handle the job, operational problems can result. The loss of one leg of a three-phase line rarely results in zero (or near-zero) voltages in the legs associated with the problem line. Instead, a combination of leakage currents caused by “regeneration” of the missing legs in inductive loads and the system load distribution usually results in voltages of some sort on the fault legs of the three-phase supply. It is possible, for example, to have phase-to-phase voltages of 220, 185, and 95 V on the legs of a three-phase, 208 V ac line experiencing a single-phasing problem. These voltages often change, depending upon the equipment turned on at the transmitter site.

Integrated circuit technology has provided a cost-effective solution to this common design problem in medium- and high-power RF equipment. Phase-loss protection modules are available from several manufacturers that provide a contact closure when voltages of proper magnitude and phase are present on the monitored line. The relay contacts can be wired into the logic control ladder of the transmitter to prevent the application of primary ac power during a single-phasing condition. Figure 12.17 shows the recommended connection method. Note that the input to the phase monitor module is taken from the final set of three-phase blower motor fuses. In this way, any failure inside the transmitter that might result in a single-phasing condition is taken into account. Because three-phase motors are particularly sensitive to single-phasing faults, the relay interlock is tied into the filament circuit logic ladder. The phase-loss protector shown in the figure includes a sensitivity adjustment for various nominal line voltages.

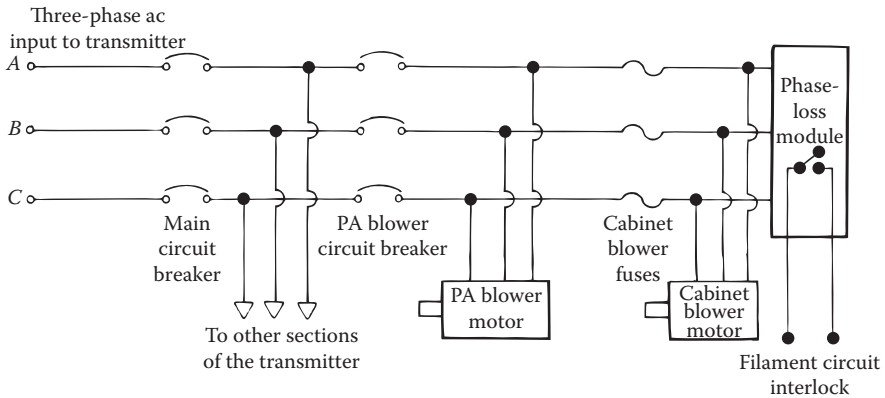


FIGURE 12.17

High-performance single-phasing protection circuit using a phase-loss module as the sensor.

12.5 Troubleshooting Procedure

When a problem occurs, the first task is to keep the transmitter on the air. If a standby transmitter is available, the solution is obvious. If the facility does not have a standby, quick thinking will be needed to minimize downtime and keep the unit running until repairs can be made. Most transmitters have sufficient protective devices so that it is impossible to operate them with serious problems. If the transmitter will not stay on the air on normal power, try lowering the power output and see if the trip-offs are eliminated. Failing this, many transmitters have driver outputs that can be connected to the antenna on a temporary basis, thereby bypassing the final amplifiers—provided, of course, the failure is in one of the PA stages. Do not allow the transmitter to operate at any power level if the meter readings are out of tolerance. Serious damage can result to the system.

When presented with a problem, proceed in an orderly manner to trace it down. Many failures are simple to repair if you stop and think about what is happening. Examine the last set of transmitter readings, and make a complete list of meter readings in the failure mode. Note which overload lamps are lit, and what other indicators are in an alarm state. With this information assembled, the cause of the failure can often be identified. Looking over the available data and the schematic diagram for 10–15 min can save hours of trial-and-error troubleshooting.

When checking inside the unit, look for changes in the physical appearance of components in the problem area. An overheated resistor or leaky capacitor may be the cause of the problem, or point to the cause. Devices never fail without a reason. Try and piece together the sequence of events that led to the problem. Then, the cause of the failure—not just the more obvious symptoms—will be corrected. When working with direct-coupled

transistors, a failure in one device will often cause a failure in another, so check all semiconductors associated with one found to be defective.

In high-power transmitters, look for signs of arcing in the RF compartments. Loose connections and clamps can cause failures that are hard to locate. Never rush through a troubleshooting job. A thorough knowledge of the theory of operation and history of the transmitter is a great aid in locating problems in the RF sections. Do not overlook the possibility of tube failure when troubleshooting a transmitter. Tubes can fail in unusual ways; substitution may be the only practical test for power tubes used in modern transmitters.

Study the control ladder of the transmitter to identify interlock or fail-safe system problems. Most newer transmitters have troubleshooting aids built in to help locate problems in the control ladder. Older transmitters, however, often require a moderate amount of investigation before repairs can be accomplished.

There is no such thing as a quick fix when it comes to transmission equipment. Think out any problem and allow ample time for repair. It makes little sense to rush through a repair job in order to get the system back on the air if another failure occurs right after the technician walks out the door. Careful analysis of the cause and effects of the failure will ensure that the original problem is solved, not just its obvious symptoms. If temporary repairs must be made in order to return the transmitter to a serviceable condition, make them and then finish the job as soon as the needed replacement parts are available.

12.5.1 Factory Service Assistance

Factory service engineers are available to aid in troubleshooting transmission equipment, but such services have their limits. No factory engineer can fix a transmitter over the phone. The factory can suggest areas of the system to investigate and relate the solutions to similar failure modes, but the facility engineer is the person who does the repair work. If the engineer knows the equipment, and has done a good job of analyzing the problem, the factory can help. When calling the factory service department, have the following basic items on hand:

- The type of transmitter and the exact failure mode. The service department will need to know the meter readings before and after the problem, and whether any unusual circumstances preceded the failure. For example, it would be important for the factory to know that the failure occurred after a brief power outage, or during an ice storm.
- A list of what has already been done in an effort to correct the problem. All too often the factory is called before any repair efforts

are made. The service engineer will need to know what happens when the high voltage is applied, and what overloads may occur.

- A copy of the transmitter diagram and component layout drawings. A thorough knowledge of the transmitter design and construction allows the maintenance engineer to intelligently converse with the factory service representative.

12.5.2 Plate Overload Fault

Of all the problems that can occur in a transmitter, probably the best known—and most feared—is the plate supply overload. Occasional plate trip-offs (one or two a month) are not generally cause for concern. Most of these occurrences can be attributed to power line transients. More frequent trip-offs require a closer inspection of the transmission system. For the purposes of our discussion, we will assume that the plate supply overload occurs frequently enough to make continued operation of the transmitter difficult.

12.5.2.1 Troubleshooting Procedure

The first step in any transmitter troubleshooting procedure is to switch the system to *local control* so that the maintenance technician, not an off-site operator, has control over the unit. This is important for safety reasons. Next, switch the transmitter automatic recycle circuit off. While troubleshooting, the transmitter should not be allowed to cycle through an overload any more times than absolutely necessary. Such action only increases the possibility of additional component damage. Use a logical, methodical approach to finding the problem. The following procedure is recommended:

- Determine the fault condition. When the maintenance engineer arrives at the transmitter site, the unit will probably be down. The carrier will be off, but the filaments may still be lit. Check all multi-meter readings on the transmitter and exciter. If they indicate a problem in a low voltage stage, troubleshoot that failure before bringing the high voltage up.
- Assuming that all low-voltage systems appear to be operating normally, switch the filaments off and make a quick visual check inside the transmitter cabinet. Determine whether there is any obvious problem. Pay particular attention to the condition of power transformers and high-voltage capacitors. Check for signs of arcing in the PA compartment. Look on the floor of the transmitter and in the RF compartments to see if there are any pieces of components lying around. Sniff inside the cabinet for hints of smoke. Check the circuit breakers and fuses to see what failures might be indicated.

- After running through these preliminary steps, restart the filaments. Then bring up the high voltage. Watch the front panel meters to see how they react. Observe what happens and listen for any sound of arcing. If the transmitter will come up, quickly run through the PA and IPA meter readings. Check the VSWR meter for excessive reflected power.

Assuming that problems persist, determine whether the plate supply overload is RF or dc based. With the plate off, switch the exciter off. Bring up the high voltage (plate supply). If the overload problem remains, the failure is based in the dc high-voltage power supply. If the problem disappears, the failure is centered in the transmitter RF chain. Proper bias must be present on all vacuum tube stages of the RF system when this test is performed. The PA tube bias supply is usually switched on with the filaments, and can generally be read from the front panel of the transmitter. Confirm proper bias before applying high voltage with no excitation. It is also important that the exciter is switched off while the high voltage is off. Removing excitation from a transmitter while it is on the air can result in a large transient overvoltage, which can cause arcing or component damage. Consult with the transmitter manufacturer for the recommended procedure.

If the overload is based in the high-voltage dc power supply, shut down the transmitter and check the schematic diagram for the location in the circuit of the plate overload sensor relay (or comparator circuit). This will indicate within what limits component checking will be required. The plate overload sensor is usually found in one of two locations: the PA cathode dc return, or the high-voltage power supply negative connection to ground. Transmitters using a cathode overload sensor generally have a separate high-voltage dc overload sensor in the plate power supply.

A sensor in the cathode circuit will substantially reduce the area of component checking required. A plate overload with no excitation in such an arrangement would almost certainly indicate a PA tube failure, because of either an interelectrode short circuit or a loss of vacuum. Do not operate the transmitter when the PA tube is out of its socket. This is not an acceptable method of determining whether a problem exists with the PA tube. Substitute a spare tube instead. Operating a transmitter with the PA tube removed can result in damage to other tubes in the transmitter when the filaments are on, and damage to the driver tubes and driver output/PA input circuit components when the high voltage is on.

If circuit analysis indicates a problem in the high-voltage power supply itself, use an ohmmeter to check for short circuits. Remove all power from the transmitter and discharge all filter capacitors before beginning any troubleshooting work inside the unit. When checking for short circuits with an ohmmeter, take into account the effects that bleeder resistors and high-voltage meter multiplier assemblies can have on resistance readings. Most access panels on transmitters use an interlock system that

will remove the high voltage and ground the high-voltage supplies when a panel is removed. For the purposes of ohmmeter tests, these interlocks may have to be temporarily defeated. Never defeat the interlocks unless all ac power has been removed from the transmitter and all filter capacitors have been discharged using the grounding stick supplied with the transmitter. Remove ac power from the transmitter by tripping the wall-mounted circuit breaker(s). Place a conspicuous note on the breaker panel stating that maintenance work is underway and that no changes should be made to the settings. This is particularly important if the breaker panel and the transmitter are in separate rooms.

Following the preliminary ohmmeter tests, check the following components in the dc plate supply:

- Oil-filled capacitors for signs of overheating or leakage.
- Feedthrough capacitors for signs of arcing or other damage.
- The dc plate blocking capacitor for indications of insulation breakdown or arcing.
- All transformers and chokes for signs of overheating or winding failure.
- Transient suppression devices for indications of overheating or failure.
- Bleeder resistors for signs of overheating.
- Any surge-limiting resistors placed in series with filter capacitors in the power supply for indications of overheating or failure. A series resistor that shows signs of overheating can be an indication that the associated filter capacitor has failed.

If the plate overload trip-off occurs only at elevated voltage levels, ohmmeter checks will not reveal the cause of the problem. It may be necessary, therefore, to troubleshoot the problem using the process of elimination.

12.5.2.2 Process of Elimination

Troubleshooting through the process of elimination involves isolating various portions of the circuit—one section at a time—until the defective component is found. Special precautions are required before performing such work:

- Never touch anything inside the transmitter without first removing all ac power and then discharging all filter capacitors with the grounding stick.
- Never perform troubleshooting work alone; another person should be present.

- Whenever a wire is disconnected, temporarily wrap it with electrical tape and secure the connector so it will not arc-over to ground or another component when power is applied.
- Analyze each planned test before it is conducted. Every test in the troubleshooting process requires time, and so steps should be arranged to provide the greatest amount of information about the problem.
- Check with the transmitter manufacturer to find out what testing procedures the company recommends. Ask what special precautions should be taken.

Figure 12.18 shows a typical transmitter high-voltage power supply. Begin the troubleshooting process by breaking the circuit at point "A." If the overload condition persists, the failure is caused by a problem in the power supply itself, not in the PA compartment. If, on the other hand, the overload condition disappears, a failure in the feedthrough capacitor (C1), decoupling capacitors (C2, C3), or blocking capacitor (C4) is indicated.

If a problem is indicated in the PA compartment, reconnect the high-voltage supply line at point "A" and break the circuit at point "B." A return of the overload problem would indicate a failure in one of the decoupling capacitors or feedthrough capacitor.

In order to avoid unnecessary effort and time in troubleshooting, use the process of elimination to identify sections of the circuit to be examined. If, for example, the test at point "A" had indicated the problem was not in the

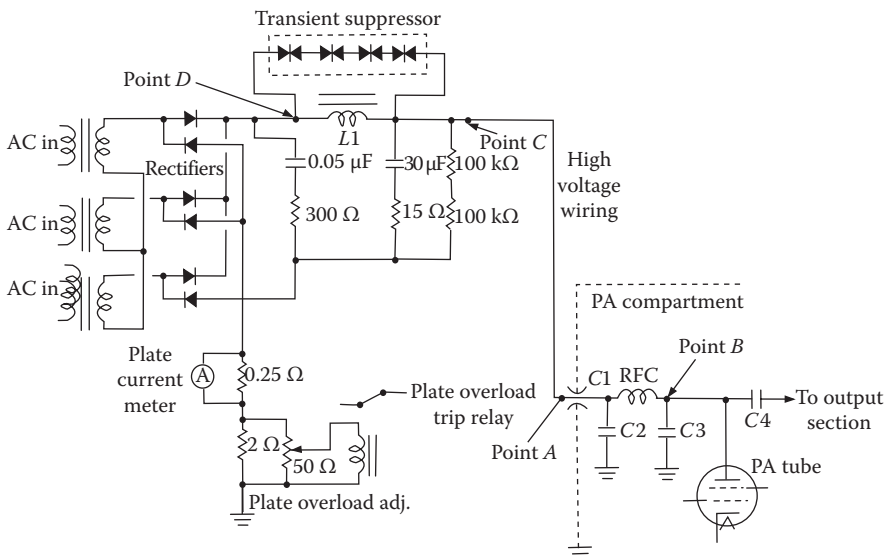


FIGURE 12.18
Typical transmitter high-voltage three-phase power supply circuit.

load, but in the power supply, a logical spot to perform the next test would be at point "C" (for long high-voltage cable runs). This test would identify or eliminate the interconnecting cable as a cause of the fault condition. If the cable run from the high-voltage supply to the PA compartment is short, point "D" might be the next logical point to check. Breaking the connection at the input to the power supply filter allows the rectifiers and interconnecting cables to be checked. Components protected by transient suppression devices (as shown above L1) should be considered a part of the component they are designed to protect. If a choke is removed from the circuit for testing, its protective device must also be removed. Failure to remove both connections may result in damage to the protective device.

It is important to remember that with no excitation, the plate current will be essentially zero. As such, there will be no appreciable voltage drop across resistors in series with the plate supply line. This may present voltage stand-off or overload concerns for ancillary circuits that feed from the plate supply.

To avoid creating a new problem while trying to correct the original failure, break the circuit in only one point at a time. Also, study the possible adverse effects of each planned step in the process. Disconnecting certain components from the circuit may cause overvoltages or power supply ripple that may damage other components in the transmitter. Consult the manufacturer to be sure.

Perform any troubleshooting work on a transmitter with extreme care. Transmitter high voltages can be lethal. Work inside the transmitter only after all ac power has been removed and after all capacitors have been discharged using the grounding stick provided with the transmitter. Remove primary power from the unit by tripping the appropriate power distribution circuit breakers in the transmitter building. Do not rely on internal contractors or silicon controlled rectifiers (SCRs) to remove all dangerous ac. Do not defeat protective interlock circuits. Although defeating an access panel interlock switch may save work time, the consequences can be tragic.

12.5.3 Radio Frequency System Faults

Although RF troubleshooting may seem intimidating, there is no secret to it. Patient examination of the circuit and careful study of the schematic diagram will go a long way toward locating the problem. The first step in troubleshooting an RF problem is to determine whether the fault is RF based or dc based (as outlined in the previous section).

For RF-related failures, check the load by examining the transmitter overload indicators. Most transmitters monitor reflected power from the antenna and will trip-off if excessive VSWR is detected. If the VSWR fault indicator is not lit, the load is not likely the cause of the problem. A definitive check of the load can be made by switching the transmitter output to a dummy load and bringing up the high voltage. The PA tube may be checked by substituting one of known quality. When the tube is changed, carefully inspect the contact fingerstock for signs of overheating or arcing. Be careful to protect the socket

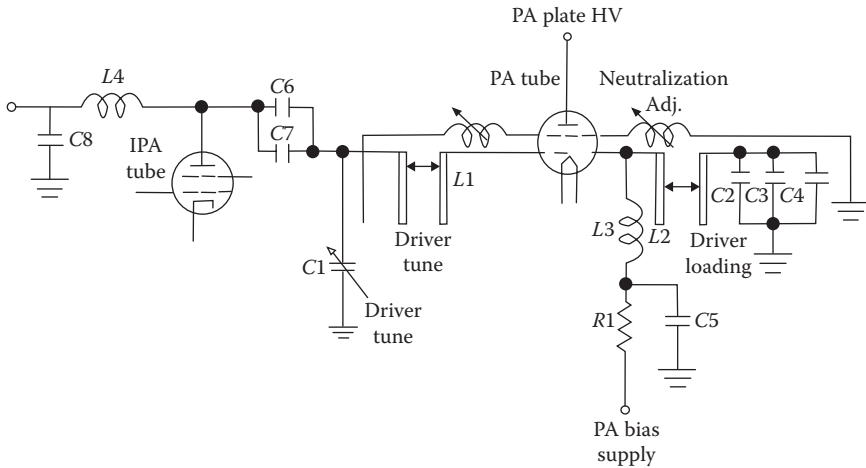


FIGURE 12.19
PA grid input circuit of a grounded screen transmitter.

from damage when removing and inserting the PA tube. Do not change the tube unless there is good reason to believe that it may be defective.

If problems with the PA stage persist, examine the grid circuit of the tube. Figure 12.19 shows the input stage of a grounded screen FM transmitter. A short circuit in any of the capacitors in the grid circuit (C1–C5) will effectively ground the PA grid. This will cause a dramatic increase in plate current, because the PA bias supply will be shorted to ground along with the RF signal from the IPA stage.

The process of finding a defective capacitor in the grid circuit begins with a visual inspection of the suspected components. Look for signs of discoloration because of overheating, loose connections, and evidence of package rupture. The voltage and current levels found in a transmitter PA stage are often sufficient to rupture a capacitor if an internal short circuit occurs. Check for component overheating right after shutting the transmitter down. (As mentioned previously, remove all ac power and discharge all capacitors first.) A defective capacitor will often overheat. Such heating can also occur, however, because of improper tuning of the PA or IPA stage, or a defective component elsewhere in the circuit.

Before replacing any components, study the transmitter schematic diagram to determine which parts in the circuit could cause the failure condition that exists. By knowing how the transmitter works, many hours can be saved in checking components that an examination of the fault condition and the transmitter design would show to be an unlikely cause of the problem.

Check blocking capacitors C6 and C7. A breakdown in either component would have serious consequences. The PA tube would be driven into full conduction, and could arc internally. The working voltages of capacitors C1–C5 could also be exceeded, damaging one or more of the components.

Because most of the wiring in the grid circuit of a PA stage consists of wide metal straps (required because of the skin effect), it is not possible to view stress points in the circuit to narrow the scope of the troubleshooting work. Areas of the system that are interconnected using components that have low power dissipation capabilities, however, should be closely examined. For example, the grid bias decoupling components shown in Figure 12.19 ($R1$, $L3$, and $C5$) include a low wattage (2W) resistor and a small RF choke. Because of the limited power dissipation ability of these two devices, a failure in decoupling capacitor $C5$ would likely cause $R1$ and possibly $L3$ to burn out. The failure of $C5$ in a short circuit would pull the PA grid to near ground potential, causing the plate current to increase and trip off the transmitter high voltage. Depending on the sensitivity and speed of the plate overload sensor, $L3$ could be damaged or destroyed by the increased current it would carry to $C5$, and therefore, to ground.

If $L3$ were able to survive the surge currents that resulted in PA plate overload, the choke would continue to keep the plate supply off until $C5$ was replaced. Bias supply resistor $R1$, however, would likely burn out because the bias power supply is generally switched on with the transmitter filament supply. Therefore, unless the PA bias power supply line fuse opened, $R1$ would overheat and probably fail.

Because of the close spacing of components in the input circuit of a PA stage, carefully check for signs of arcing between components or sections of the tube socket. Keep all components and the socket itself clean at all times. Inspect all interconnecting wiring for signs of damage, arcing to ground, or loose connections.

12.5.3.1 Component Substitution

Substituting a new component for a suspected part can save valuable time when troubleshooting. With some components, it is cost effective to replace a group of parts that may include one defective component because of the time involved in gaining access to the damaged device. For example, the grid circuit of the PA stage shown in Figure 12.19 includes three *doorknob* capacitors ($C2$ – $C4$) formed into a single assembly. If one device was found to be defective, it might be advantageous to simply replace all three capacitors. These types of components are often integrated into a single unit that may be difficult to reach. Because doorknob capacitors are relatively inexpensive, it may be best to replace the three as a group. This way, the entire assembly is eliminated as a potential cause of the fault condition.

A good supply of spare parts is a valuable troubleshooting tool. In high-power transmitting equipment, substitution is sometimes the only practical means of finding the cause of a problem. The manufacturer's factory service department can usually recommend a minimum spare parts inventory. Obvious candidates for inventory include components that are not available

locally, such as high-voltage fixed-value capacitors, vacuum variable capacitors, and specialized semiconductors.

Many failures resulting from problems within a cavity are the result of poor mechanical connections. All screws and connections must be kept tight. Every nut and bolt in a PA cavity was included for a reason. There are no insignificant screws that do not need to be tight. However, do not over tighten either. Stripped threads and broken component connection lugs will only cause additional grief.

When a problem occurs in a PA cavity, it is usually difficult to determine which individual element (neutralization inductor, plate tuning capacitor, loading capacitor, etc.) is defective from the symptoms the failure will display. A fault within the cavity is usually a catastrophic event that will take the transmitter off the air. It is often impossible to bring the transmitter up for even a few seconds to assess the fault situation. The only way to get at the problem is to shut the transmitter down and take a look inside.

Closely inspect every connection, using a trouble light and magnifying glass. Look for signs of arcing or discoloration of components or metal connections. Check the mechanical integrity of each element in the circuit. Be certain the tuning and loading adjustments are solid, without excessive mechanical play. Look for signs of change in the cavity. Check areas of the cavity that may not seem like vital parts of the output stage, such as the maintenance access door fingerstock and screws. A failure in the integrity of the cavity, whether at the base of the PA tube or on part of the access door, can cause high circulating currents to develop and may prevent proper operation of the stage. If a problem is found that involves damaged fingerstock, replace the affected sections. Failure to do so will likely result in future problems because of the currents that can develop at any discontinuity in the cavity's inner or outer conductor.

12.5.3.2 VSWR Overload

A VSWR overload in transmission equipment can result from a number of different problems. The following checklist presents some common problems and solutions:

1. VSWR overloads are usually caused by an improper impedance match external to the transmitter. The first step in the troubleshooting procedure is to substitute a dummy load for the entire antenna and transmission line system. Connect the dummy load at the transmitter output port, thereby eliminating all coax, external filters, and other RF hardware that might be present in the system.
2. If the VSWR trip fault is eliminated in step 1, the problem lies somewhere in the transmission line or antenna. The dummy load can next be moved to the point at which the transmission line leaves the

- building and heads for the tower (if different than the point checked in step 1). This test will check any RF plumbing, switches, or filter assemblies. If the VSWR overload condition is still absent, the problem is centered in the transmission line or the antenna.
3. If a standby antenna is not available, operation may still be possible at reduced power on a temporary basis. For example, if arcing occurs in the antenna or line at full power, emergency operation may be possible at half power. Inspect the antenna and line for signs of trouble. Repair work beyond this point normally requires specialized equipment and a tower crew. This discussion assumes that the problem is not caused by ice buildup on the antenna, which can be alleviated by reducing the transmitter power output until VSWR trips do not occur.
 4. If step 1 shows the VSWR overload to be internal to the transmitter, determine whether the problem is caused by an actual VSWR overload or by a failure in the VSWR control circuitry. Check for this condition by disabling the transmitter exciter and bringing up the high voltage. Under these conditions, RF energy will not be generated. (It is assumed that the transmitter has proper bias on all stages and is properly neutralized.) If a VSWR overload is indicated, the problem is centered in the VSWR control circuitry and not the RF chain. Possible explanations for control circuitry failure include loose connections, dirty switch contacts, dirty calibration potentiometers, poor PC board edge connector contacts, and defective IC amplifiers or logic gates.
 5. If step 4 shows the VSWR overload is real, and not the result of faulty control circuitry, check all connections in the output and coupling sections of the final stage. Look for signs of arcing or loose hardware, particularly on any movable tuning components. Inspect high-voltage capacitors for signs of overheating, which might indicate failure; and check coils for signs of dust buildup, which might cause a flashover. In some transmitters, VSWR overloads can be caused by improper final stage tuning or loading. Consult the equipment maintenance manual for this possibility. Also, certain transmitters include glass-enclosed spark-gap lightning protection devices (*gas-gaps*) that can be disconnected for testing.

12.5.4 Power Control Faults

A failure in the transmitter ac power control circuitry can result in problems ranging from zero RF output to a fire inside the unit. Careful, logical troubleshooting of the control system is mandatory. Two basic types of primary ac control are used in transmitters today: SCR (thyristor) system and relay contactors.

12.5.4.1 Thyristor Control System

A failure in the thyristor power control system of a transmitter is not easy to overlook. In the worst case, no high voltage at all will be produced by the transmitter. In the best case, power control may be erratic or uneven when using the continuously variable power adjustment mode. Understanding how the servo circuit works and how it is interconnected is the first step in correcting such a problem.

Figure 12.20 shows a block diagram of a typical thyristor control circuit. Three gating cards are used to drive back-to-back SCR pairs, which feed the high-voltage power transformer primary windings. Although the applied voltage is three phase, the thyristor power control configuration simulates a single-phase design for each phase-to-phase leg. This allows implementation of a control circuit that consists basically of a single-phase gating card duplicated three times (one for each load phase). This approach has advantages from the standpoint of design simplicity and also from the standpoint of field troubleshooting. In essence, each power control circuit is identical, allowing test voltages and waveforms from one gating card to be directly compared with a gating card experiencing problems.

12.5.4.2 Interlock Failures

Newer transmitters provide the engineer with built-in diagnostic readouts on the status of the transmitter interlock circuits. These may involve discrete LEDs or a microcontroller-driven visual display of some type. If the transmitter has an advanced status display, the process of locating an interlock fault is relatively simple. For an older transmitter that is not so equipped, substantially more investigation will be needed.

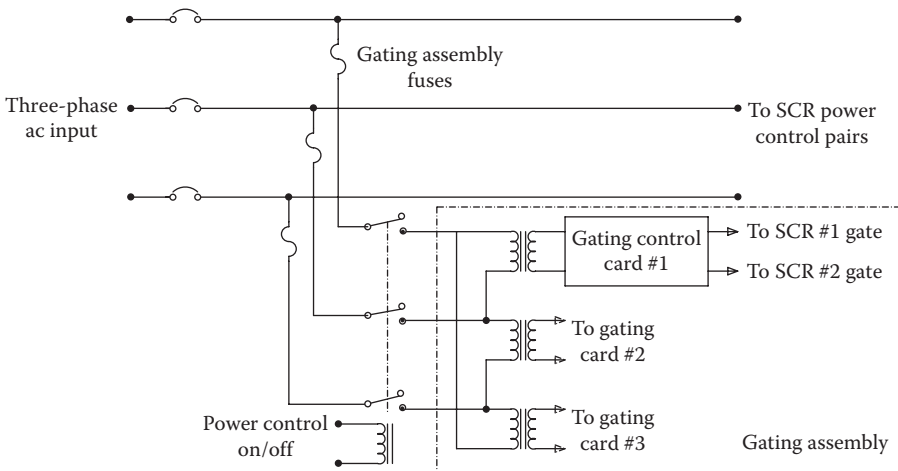


FIGURE 12.20

Block diagram of a three-phase thyristor power control system.

Make a close observation of the status of all fuses, circuit breakers, transmitter cabinet doors, and access panels. Confirm that all doors are fully closed and secured. Switch the transmitter from *remote* to *local* control (if operated remotely) to eliminate the remote control system as a possible cause of the problem. Observe the status of all control-panel indicator lamps. Some transmitters include an "interlocks open" lamp; other units provide an indication of an open interlock through the "Filament On" or "Plate Off" push-button lamps. These indicators can save valuable minutes or even hours of troubleshooting, so pay attention to them. Replace any burned-out indicator lamps as soon as they are found to be defective.

If the front-panel indicators point to an interlock problem, pull out the schematic diagram of the transmitter, get out the DMM, and shut down the transmitter. If the transmitter interlock circuit operates from a low-voltage power supply, such as 24V dc, use a voltmeter to check for the loss of continuity. Remove ac power from all sections of the transmitter except the low-voltage supply by tripping the appropriate front-panel circuit breakers. Be extremely careful when working on the transmitter to avoid any line voltage ac. If the layout of the transmitter does not permit safe troubleshooting with only the low-voltage power supply active, remove all ac from the unit by tripping the wall-mounted main breaker. Then, use an ohmmeter to check for the loss of continuity.

If the transmitter interlock circuit operates from 120V ac or 220V ac, remove all power from the transmitter by tripping the wall-mounted main breaker. Use an ohmmeter to locate the problem. Some older transmitters use line voltages in the interlock system. Do not try to troubleshoot such transmitters with ac power applied.

Finding a problem such as an open control circuit interlock is basically a simple procedure, in spite of the time involved. Do not rush through such work. When searching for a break in the interlock system, use a methodical approach to solving the problem. Consider the circuit configuration shown in Figure 12.21. The most logical approach to finding a break in the control ladder is to begin at the source of the 24V dc input and, step by step, work toward the input of the power controller. Although this approach may be logical, it can also be time consuming. Instead, eliminate stages of the interlock circuit. For example, make the first test at terminal "A." A correct voltage reading at this point in the circuit will confirm that all of the interlock door switches are operating properly.

With the knowledge that the problem is after terminal "A," move on to terminal "B." If the 24V supply voltage disappears, check the fault circuit overload relays to find where the control signal is lost. Often, such interlock problems can be attributed to dirty contacts in one of the overload relays. If a problem is found with one set of relay contacts, clean all of the other contacts in the overload interlock string for good measure. Be sure to use the proper relay-contact cleaning tools. If sealed relays are used, do not attempt to clean them. Instead, replace the defective unit.

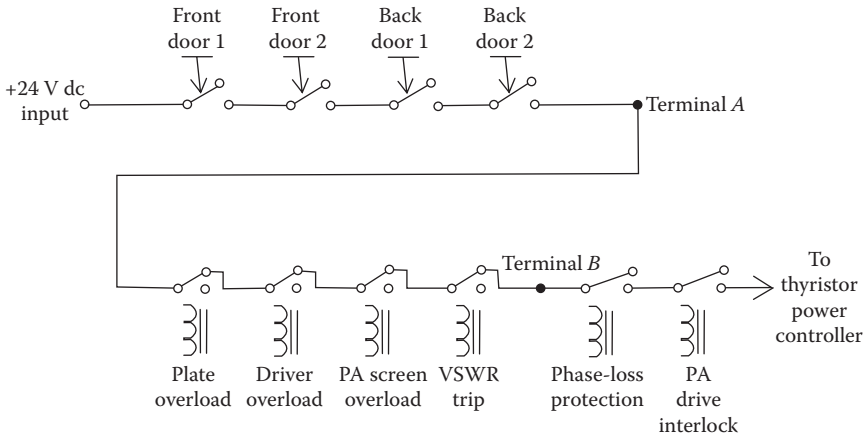


FIGURE 12.21

Typical transmitter interlock circuit. Terminals “A” and “B” are test points used for troubleshooting the system in the event of an interlock system failure.

12.5.4.3 Step-Start Faults

The high-voltage power supply of any medium- or high-power transmitter must include provisions for in-rush current limiting upon the application of a Plate-On command. The filter capacitor(s) in the power supply will appear as a virtual short circuit during a sudden increase in voltage from the rectifier stacks. To avoid excessive current surges through the rectifiers, capacitor(s), choke, and power transformer, nearly all transmitters use some form of step-start arrangement. Such circuits are designed to limit the in-rush current to a predictable level. This can be accomplished in a variety of ways.

For transmitters using a thyristor power control system, the step-start function can be easily designed into the gate firing control circuit. An R - C network at the input point of the gating cards is used to ramp the thyristor pairs from a zero conduction angle to full conduction (or a conduction angle preset by the user). This system provides an elegant solution to the step-start requirement, allowing plate voltage to be increased from zero to full value within a period of several seconds.

Transmitters employing a conventional ac power control system usually incorporate a step-start circuit consisting of two sets of contactors: the *start contactor* and the *run contactor*, as illustrated in Figure 12.22. Surge-limiting resistors provide sufficient voltage drop upon application of a Plate-On command to limit the surge current to a safe level. Auxiliary contacts on the start contactor cause the run contactor to close as soon as the start contacts have seated.

A fault in the step-start circuit of a transmitter is often evidenced—initially at least—by random tripping of the plate supply circuit breaker at high-voltage turn-on. If left uncorrected, this condition can lead to problems such as failed power rectifiers or filter capacitors.

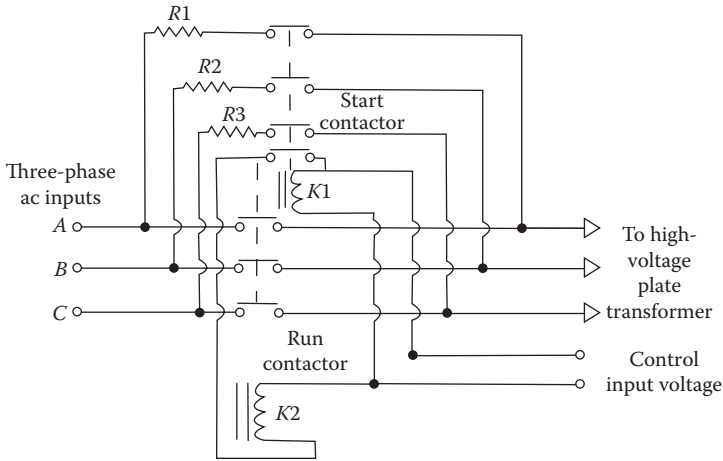


FIGURE 12.22
Typical three-phase step-start power control system.

Troubleshooting a step-start fault in a system employing thyristor power control should begin at the R-C ramp network. Check the capacitor to see if it has opened. Monitor the control voltage to the thyristor gating cards to confirm that the output voltage of the controller slowly increases to full value. If it does and the turn-on problem persists, the failure involves one or more of the gating cards.

When troubleshooting a step-start fault in a transmitter employing the dual contactor arrangement, begin with a close inspection of all contact points on both contactors. Pay careful attention to the auxiliary relay contacts of the start contactor. If the contacts fail to properly close, the full load of the high-voltage power supply will be carried through the series resistors and start contactor. These devices are normally sized only for intermittent duty; they are not intended to carry the full load current for an extended length of time. Look for signs of arcing or overheating of the contact pairs and current-carrying connector bars. Check the current-limiting resistors for excessive dissipation and continuity.

12.5.5 Protection Circuits

Many engineers enjoy a false sense of security with transmission equipment because of the protection devices included in most designs. Although conventional overload circuits provide protection against most common failure modes, they are not foolproof. The first line of defense in the transmitter, the ac power system circuit breakers, can allow potentially disastrous currents to flow under certain fault conditions.

Consider the thyristor ac power serve system shown in Figure 12.23. This common type of voltage-regulator circuit adjusts the condition angle of the SCR pairs to achieve the desired dc output from the high-voltage

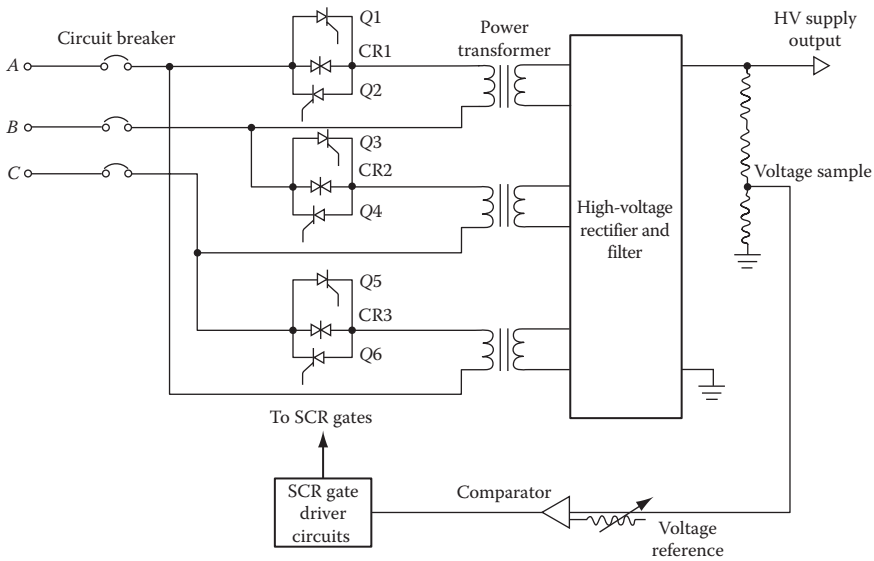


FIGURE 12.23
Common three-phase thyristor servo ac power control system.

power supply. An alternative configuration could have the output voltage sample derived from a transmission line RF pickup and amplifier/detector. In this way, the primary power control is adjusted to match the desired RF output from the transmitter. If one of the high-voltage rectifier stacks of this system failed in a short-circuit condition, the output voltage (and RF output) would fall, causing the thyristor circuit to increase the conduction period of the SCR pairs. Depending on the series resistance of the failed rectifier stack and the rating of the primary side circuit breaker, the breaker may or may not trip. Remember that the circuit breaker was chosen to allow operation at full transmitter power with the necessary headroom to prevent random tripping. The primary power system, therefore, can dissipate a significant amount of heat under reduced power conditions, such as those that would be experienced with a drop in the high-voltage supply output. The difference between the maximum designed power output of the supply (and, therefore, the transmitter) and the failure-induced power demand of the system can be dissipated as heat without tripping the main breaker.

Operation under such fault conditions, even for 20 s or less, can cause considerable damage to power supply components, such as the power transformer, rectifier stack, thyristors, or system wiring. Damage can range from additional component failures to a fire in the affected section of the transmitter. Disasters such as this are rare, but they do occur. Be prepared to respond to any emergency condition by thoroughly understanding how the transmitter works, and by identifying potential weak points in the system.

13

Safe Handling of Vacuum Tube Devices

13.1 Introduction

Electrical safety is important when working with any type of electronic hardware. Because vacuum tubes operate at high voltages and currents, safety is doubly important. The primary areas of concern, from a safety standpoint, include

- Electric shock
- Nonionizing radiation
- Beryllium oxide (BeO) ceramic dust
- Hot surfaces of vacuum tube devices
- Polychlorinated biphenyls (PCBs)

13.2 Electric Shock

Surprisingly little current is required to injure a person. Studies at Underwriters Laboratories (UL) show that the electrical resistance of the human body varies with the amount of moisture on the skin, the muscular structure of the body, and the applied voltage. The typical hand-to-hand resistance ranges between $500\ \Omega$ and $600\ \text{k}\Omega$, depending on the conditions. Higher voltages have the capability to break down the outer layers of the skin, which can reduce the overall resistance value. UL uses the lower value, $500\ \Omega$ as the standard resistance between major extremities, such as from the hand to the foot. This value is generally considered the minimum that would be encountered and, in fact, may not be unusual because wet conditions or a cut or other break in the skin significantly reduces human body resistance.

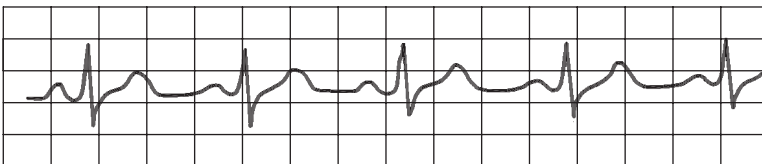
13.2.1 Effects on the Human Body

Table 13.1 lists some effects that typically result when a person is connected across a current source with a hand-to-hand resistance of $2.4\text{ k}\Omega$. The table shows that a current of approximately 50 mA will flow between the hands, if one hand is in contact with a 120 V ac source and the other hand is grounded. The table indicates that even the relatively small current of 50 mA can produce *ventricular fibrillation* of the heart, and perhaps death. Medical literature describes ventricular fibrillation as rapid, uncoordinated contractions of the ventricles of the heart, resulting in loss of synchronization between heartbeat and pulse beat. The electrocardiograms shown in Figure 13.1 compare a healthy heart rhythm with one in ventricular fibrillation.

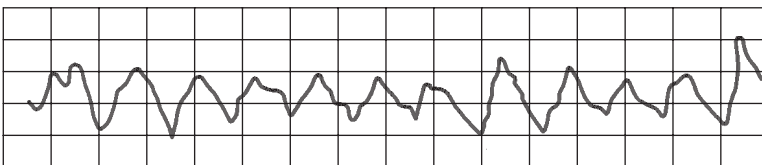
TABLE 13.1

Effects of Current on the Human Body

Current	Effect
1 mA or less	No sensation, not felt
More than 3 mA	Painful shock
More than 10 mA	Local muscle contractions, sufficient to cause “freezing” to the circuit for 2.5% of the population
More than 15 mA	Local muscle contractions, sufficient to cause “freezing” to the circuit for 50% of the population
More than 30 mA	Breathing is difficult, can cause unconsciousness
50–100 mA	Possible ventricular fibrillation
100–200 mA	Certain ventricular fibrillation
More than 200 mA	Severe burns and muscular contractions; heart more apt to stop than to go into fibrillation
More than a few amperes	Irreparable damage to body tissue



(a)



(b)

FIGURE 13.1

Electrocardiogram of a human heartbeat: (a) healthy rhythm and (b) ventricular fibrillation.

Unfortunately, once ventricular fibrillation occurs, it will continue. Barring resuscitation techniques, death will ensue within a few minutes.

The route taken by the current through the body has a significant effect on the degree of injury. Even a small current, passing from one extremity through the heart to another extremity, is dangerous and capable of causing severe injury or electrocution. There are cases where a person has contacted extremely high current levels and lived to tell about it. However, usually when this happens, the current passes only through a single limb and not through the body. In these instances, the limb is often lost, but the person survives.

Current is not the only factor in electrocution. Figure 13.2 summarizes the relationship between current and time on the human body. The graph shows that 100 mA flowing through a human adult body for 2 s will cause death by electrocution. An important factor in electrocution, the *let-go* range,

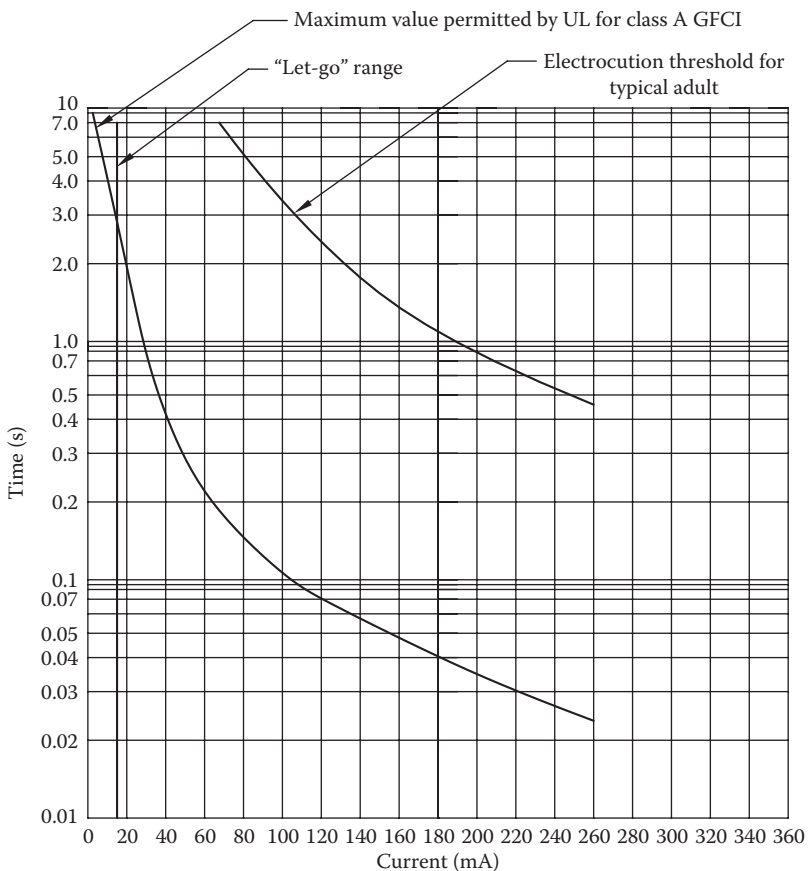


FIGURE 13.2

Effects of electric current and time on the human body. Note the "let-go" range.

TABLE 13.2**Required Safety Practices for Engineers Working around High-Voltage Equipment***High-voltage precautions*

- Remove all ac power from the equipment. Do not rely on internal contactors or SCRs to remove dangerous ac
- Trip the appropriate power distribution circuit breakers at the main breaker panel
- Place signs as needed to indicate that the circuit is being serviced
- Switch the equipment being serviced to the *local control* mode as provided
- Discharge all capacitors using the discharge stick provided by the manufacturer
- Do not remove, short-circuit, or tamper with interlock switches on access covers, doors, enclosures, gates, panels, or shields
- Keep away from live circuits

also is shown on the graph. This range is described as the amount of current that causes “freezing” or the inability to let go of the conductor. At 10 mA, 2.5% of the population will be unable to let go of a “live” conductor. At 15 mA, 50% of the population will be unable to let go of an energized conductor. It is apparent from the graph that even a small amount of current can “freeze” someone to a conductor. The objective for those who must work around electric equipment is how to protect themselves from electric shock. Table 13.2 lists required precautions for personnel working around high voltages.

13.2.2 Circuit Protection Hardware

The typical primary panel or equipment circuit breaker or fuse will not protect a person from electrocution. In the time it takes a fuse or circuit breaker to blow, someone could die. However, there are protection devices that, properly used, may help prevent electrocution. The *ground-fault current interrupter* (GFCI), shown in Figure 13.3, works by monitoring the current being

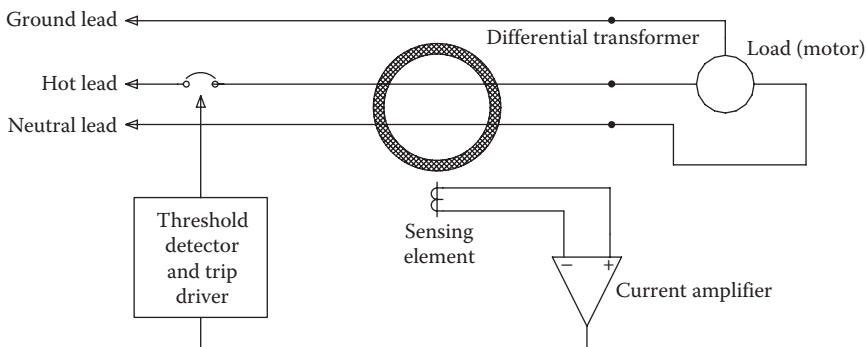


FIGURE 13.3
Basic design of a GFI.

applied to the load. The ground-fault interrupter (GFI) uses a differential transformer and looks for an imbalance in load current. If a current (5 mA, ± 1 mA) begins to flow between the neutral and ground or between the hot and ground leads, the differential transformer detects the leakage and opens up the primary circuit within 2.5 ms.

GFIs will not protect a person from every type of electrocution. If the victim becomes connected to both the neutral and the hot wire, the GFI will not detect an imbalance.

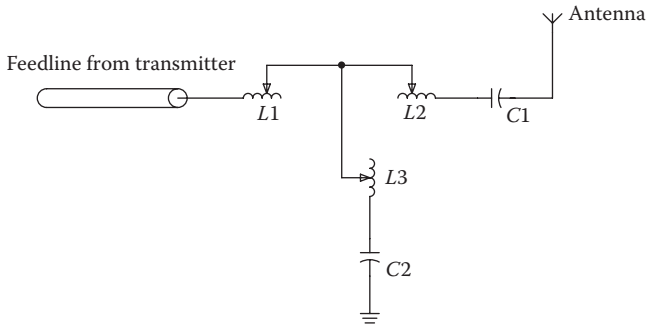
13.2.3 Working with High Voltage

Rubber gloves are commonly used by engineers working on high-voltage equipment. These gloves are designed to provide protection from hazardous voltages or radio frequency (RF) when the wearer is working on “hot” ac or RF circuits. Although the gloves may provide some protection from these hazards, placing too much reliance on them can have disastrous consequences. There are several reasons why gloves should be used with a great deal of caution and respect. A common mistake made by engineers is to assume that the gloves always provide complete protection. The gloves found in many facilities may be old or untested. Some may show signs of user repair, perhaps with electrical tape. Few tools could be more hazardous than such a pair of gloves.

Another mistake is not knowing the voltage rating of the gloves. Gloves are rated differently for both ac and dc voltages. For example, a *class 0* glove has a minimum dc breakdown voltage of 35 kV; the minimum ac breakdown voltage, however, is only 6 kV. Furthermore, high-voltage rubber gloves are not usually tested at RF frequencies, and RF can burn a hole in the best of them. It is possible to develop dangerous working habits by assuming that gloves will offer the required protection.

Gloves alone may not be enough to protect an individual in certain situations. Recall the axiom of keeping one hand in a pocket while working around a device with current flowing? That advice is actually based on simple electricity. It is not the “hot” connection that causes the problem, but the ground connection that lets the current begin to flow. Studies have shown that more than 90% of electric equipment fatalities occurred when the grounded person contacted a live conductor. Line-to-line electrocution accounted for less than 10% of the deaths.

When working around high voltages, always look for grounded surfaces. Keep hands, feet, and other parts of the body away from any grounded surface. Even concrete can act as a ground if the voltage is sufficiently high. If work must be performed in “live” cabinets, then consider using, in addition to rubber gloves, a rubber floor mat, rubber vest, and rubber sleeves. Although this may seem to be a lot of trouble, consider the consequences of making a mistake. Of course, the best troubleshooting methodology is never to work on any circuit without being certain that no hazardous voltages are present.

**FIGURE 13.4**

Example of how high voltages can be generated in an RF load matching unit.

In addition, any circuits or contactors that normally contain hazardous voltages should be firmly grounded before work begins.

Most companies and organizations have established safety procedures for working with various types of equipment. Know the local rules and obey them.

13.2.3.1 Radio Frequency Considerations

Engineers sometimes rely on electrical gloves when making adjustments to live RF circuits. This practice, however, can be extremely dangerous. Consider the typical load matching unit shown in Figure 13.4. In this configuration, disconnecting the coil from either $L2$ or $L3$ places the full RF output literally at the engineer's fingertips. Depending on the impedances involved, the voltages can become quite high, even in a circuit that normally is relatively tame.

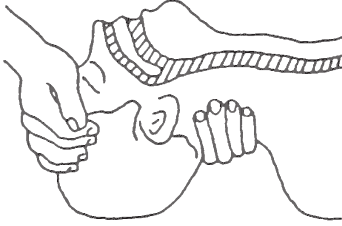
In Figure 13.4, assume that the load impedance is approximately $106 + j202 \Omega$. With 1 kW feeding into the load, the rms voltage at the matching output will be approximately 700 V. The peak voltage (which determines insulating requirements) will be close to 1 kV, and perhaps more than twice that if the carrier is being amplitude-modulated. At the instant the output coil clip is disconnected, the current in the shunt leg will increase rapidly, and the voltage easily could more than double.

13.2.4 First Aid Procedures

All engineers working around high-voltage equipment should be familiar with first aid treatment for electric shock and burns. Always keep a first aid kit on hand at the facility. Figure 13.5 illustrates the basic treatment for victims of electric shock. Copy the information, and post it in a prominent location. Better yet, obtain more detailed information from the local heart association or Red Cross chapter. Personalized instruction on first aid usually is available locally.

If the victim is not responsive, follow the a-b-c's of basic life support.

Airway: If the victim is unconscious, open airway.

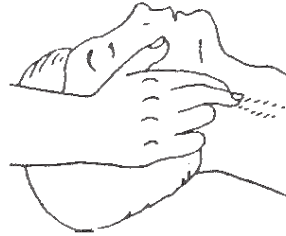


1. Lift up neck
2. Push forehead back
3. Clear out mouth if necessary
4. Observe for breathing

Breathing: If the victim is not breathing, begin artificial breathing.

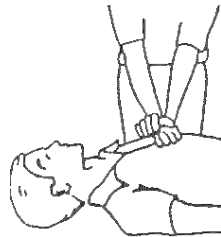
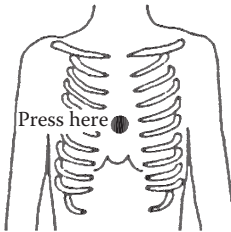


1. Tilt head
2. Pinch nostrils
3. Make airtight seal
4. Provide four quick full breaths



Check carotid pulse. If pulse is absent, begin artificial circulation. Remember that mouth-to-mouth resuscitation must be commenced as soon as possible.

Circulation: Depress the sternum 1.2–2 in.



For situations in which there is one rescuer, provide 15 compressions and then 2 quick breaths. The approximate rate of compressions should be 80/min.

For situations in which there are two rescuers, provide 5 compressions and then 1 breath. The approximate rate of compressions should be 60/min.

Do not interrupt the rhythm of compressions when a second person is giving breaths.

If the victim is responsive, keep warm and quiet, loosen clothing, and place in a reclining position. Call for medical assistance as soon as possible.

FIGURE 13.5
Basic first aid treatment for electric shock.

13.3 Operating Hazards

A number of potential hazards exist in the operation and maintenance of high-power vacuum tube RF equipment. Maintenance personnel must exercise extreme care around such hardware. Consider the following guidelines:

- Use caution around the high-voltage stages of the equipment. Many power tubes operate at voltages high enough to kill through electrocution. Always break the primary ac circuit of the power supply, and discharge all high-voltage capacitors.
- Minimize exposure to RF radiation. Do not permit personnel to be in the vicinity of open, energized RF generating circuits, RF transmission systems (waveguides, cables, or connectors), or energized antennas. High levels of radiation can result in severe bodily injury, including blindness. Cardiac pacemakers may also be affected.
- Avoid contact with BeO ceramic dust and fumes. BeO ceramic material may be used as a thermal link to carry heat from a tube to the heat sink. Do not perform any operation on any BeO ceramic that might produce dust or fumes, such as grinding, grit blasting, or acid cleaning. BeO dust and fumes are highly toxic, and breathing them can result in serious injury or death. BeO ceramics must be disposed of as prescribed by the device manufacturer.
- Avoid contact with hot surfaces within the equipment. The anode portion of most power tubes is air cooled. The external surface normally operates at a high temperature (up to 250°C). Other portions of the tube also may reach high temperatures, especially the cathode insulator and the cathode/heater surfaces. All hot surfaces may remain hot for an extended time after the tube is shut off. To prevent serious burns, avoid bodily contact with these surfaces during tube operation and for a reasonable cool-down period afterward. Table 13.3 lists basic first aid procedures for burns.

13.3.1 OSHA Safety Considerations

The U.S. government has taken a number of steps to help improve safety within the workplace under the auspices of the Occupational Safety and Health Administration (OSHA). The agency helps industries monitor and correct safety practices. OSHA has developed a number of guidelines designed to help prevent accidents. OSHA records show that electrical standards are among the most frequently violated of all safety standards. Table 13.4 lists 16 of the most common electrical violations, including exposure of live conductors, improperly labeled equipment, and faulty grounding.

TABLE 13.3**Basic First Aid Procedures for Burns***Extensively burned and broken skin*

- Cover affected area with a clean sheet or cloth
- Do not break blisters, remove tissue, remove adhered particles of clothing, or apply any salve or ointment
- Treat victim for shock as required
- Arrange for transportation to a hospital as quickly as possible
- If arms or legs are affected, keep them elevated

Less severe burns (first and second degree)

- Apply cool (not ice-cold) compresses using the cleanest available cloth article
- Do not break blisters, remove tissue, remove adhered particles of clothing, or apply salve or ointment
- Apply clean, dry dressing if necessary
- Treat victim for shock as required
- Arrange for transportation to a hospital as quickly as possible
- If arms or legs are affected, keep them elevated

More detailed information can be obtained from any Red Cross office.

TABLE 13.4**Sixteen Common OSHA Violations**

Fact Sheet	Subject	NEC^a Reference
1	Guarding of live parts	110-17
2	Identification	110-22
3	Uses allowed for flexible cord	400-7
4	Prohibited uses of flexible cord	400-8
5	Pull at joints and terminals must be prevented	400-10
6.1	Effective grounding, Part 1	250-51
6.2	Effective grounding, Part 2	250-51
7	Grounding of fixed equipment, general	250-42
8	Grounding of fixed equipment, specific	250-43
9	Grounding of equipment connected by cord and plug	250-45
10	Methods of grounding, cord and plug-connected equipment	250-59
11	AC circuits and systems to be grounded	250-5
12	Location of overcurrent devices	240-24
13	Splices in flexible cords	400-9
14	Electrical connections	110-14
15	Marking equipment	110-21
16	Working clearances about electric equipment	110-16

Source: National Electrical Code, NFPA #70.

^a National Electrical Code.

13.3.1.1 Protective Covers

Exposure of live conductors is a common safety violation. All potentially dangerous electric conductors should be covered with protective panels. The danger is that someone may come into contact with the exposed current-carrying conductors. It is also possible for metallic objects such as ladders, cable, or tools to contact a hazardous voltage, creating a life-threatening condition. Open panels also present a fire hazard.

13.3.1.2 Identification and Marking

Circuit breakers and switch panels should be properly identified and labeled. Labels on breakers and equipment switches may be many years old and may no longer reflect the equipment actually in use. This is a safety hazard. Casualties or unnecessary damage can be the result of an improperly labeled circuit panel if no one who understands the system is available in an emergency. If a number of devices are connected to a single disconnect switch or breaker, a diagram should be provided for clarification. Label with brief phrases, and use clear, permanent, and legible markings.

Equipment marking is a closely related area of concern. This is not the same thing as equipment identification. Marking equipment means labeling the equipment breaker panels and ac disconnect switches according to device rating. Breaker boxes should contain a nameplate showing the manufacturer, rating, and other pertinent electrical factors. The intent is to prevent devices from being subjected to excessive loads or voltages.

13.3.1.3 Grounding

OSHA regulations describe two types of grounding: *system grounding* and *equipment grounding*. System grounding actually connects one of the current-carrying conductors (such as the terminals of a supply transformer) to ground (see Figure 13.6). Equipment grounding connects all of the noncurrent-carrying metal surfaces together and to ground. From a grounding standpoint, the only difference between a grounded electrical system and an ungrounded electrical system is that the *main bonding jumper* from the service equipment ground to a current-carrying conductor is omitted in the ungrounded system. The system ground performs two tasks:

- It provides the final connection from equipment-grounding conductors to the grounded circuit conductor, thus completing the ground-fault loop.
- It solidly ties the electrical system and its enclosures to their surroundings (usually earth, structural steel, and plumbing). This prevents voltages at any source from rising to harmfully high voltage-to-ground levels.

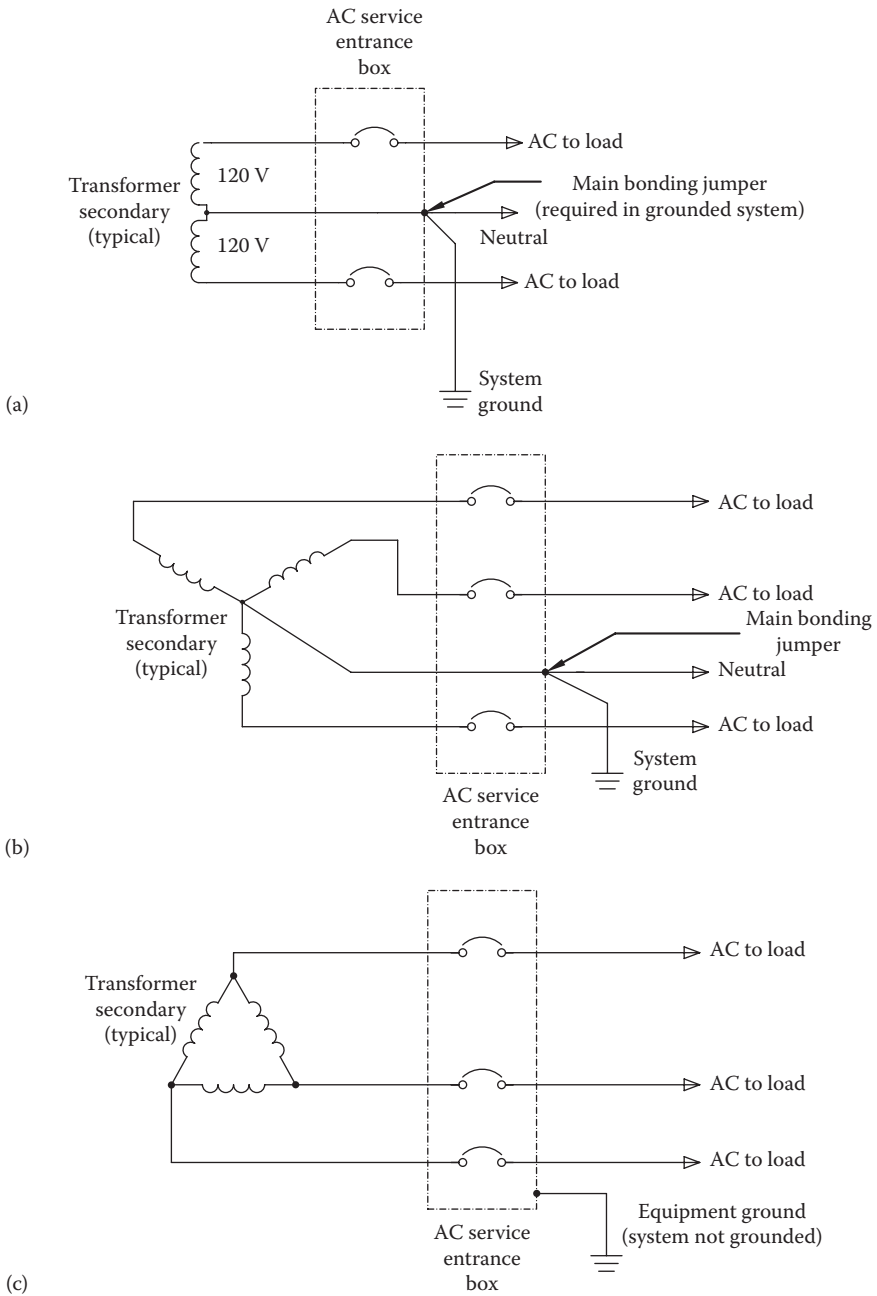


FIGURE 13.6 AC service entrance bonding: (a) 120 V phase-to-neutral (240 V phase-to-phase), (b) three-phase 208 V wye (120 V phase-to-neutral), and (c) three-phase 240 V (or 480 V) delta.

Note that equipment grounding—bonding all electric equipment to ground—is required whether or not the system is grounded. Equipment grounding serves two important tasks:

- It bonds all surfaces together so that there can be no voltage difference among them.
- It provides a ground-fault current path from a fault location back to the electrical source, so that if a fault current develops, it will rise to a level high enough to operate the breaker or fuse.

The National Electrical Code (NEC) is complex and contains numerous requirements concerning electrical safety [1]. The fact sheets listed in Table 13.4, and additional information, are available from OSHA.

13.3.2 Beryllium Oxide Ceramics

Some tubes, both power grid and microwave, contain BeO ceramics, typically at the output waveguide window or around the cathode. Never perform any operations on BeO ceramics that produce dust or fumes, such as grinding, grit blasting, or acid cleaning. BeO dust and fumes are highly toxic, and breathing them can result in serious personal injury or death. Because BeO warning labels may be obliterated or missing, maintenance personnel should contact the tube manufacturer before performing any work on a damaged device. Some tubes have BeO internal to the vacuum envelope.

13.3.3 Corrosive and Poisonous Compounds

The external output waveguides and cathode high-voltage bushings of microwave tubes are sometimes operated in systems that use a dielectric gas to impede microwave or high-voltage breakdown. If breakdown does occur, the gas may decompose and combine with impurities, such as air or water vapor, to form highly toxic and corrosive compounds. Examples include Freon gas, which may form lethal *phosgene*, and sulfur hexafluoride (SF₆) gas, which may form highly toxic and corrosive sulfur or fluorine compounds such as *beryllium fluoride*. When breakdown does occur in the presence of these gases, proceed as follows:

- Ventilate the area to outside air.
- Avoid breathing any fumes or touching any liquids that develop.
- Take precautions appropriate for beryllium compounds and for other highly toxic and corrosive substances.
- Contact the tube manufacturer for recommendations on how to proceed.

If a coolant other than pure water is used, follow the precautions supplied by the coolant manufacturer.

13.3.4 FC-75 Toxic Vapor

The decomposition products of FC-75 are highly toxic. Decomposition may occur as a result of any of the following:

- Exposure to temperatures above 200°C
- Exposure to liquid fluorine or alkali metals (lithium, potassium, or sodium)
- Exposure to ionizing radiation

Known thermal decomposition products include *perfluoroisobutylene* (PFIB; $[\text{CF}_3]_2\text{C} = \text{CF}_2$), which is highly toxic in small concentrations.

If FC-75 has been exposed to temperatures above 200°C through fire, electric heating, or prolonged electric arcs or has been exposed to alkali metals or strong ionizing radiation, take the following steps:

- Strictly avoid breathing any fumes or vapors.
- Thoroughly ventilate the area.
- Strictly avoid any contact with the FC-75.
- Promptly contact an appropriate specialist with experience in such matters.

13.4 Nonionizing Radiation

Nonionizing radio frequency radiation (RFR) resulting from high-intensity RF fields is a growing concern to engineers who must work around high-power transmission equipment. The principal medical concern regarding nonionizing radiation involves heating of various body tissues, which can have serious effects, particularly if there is no mechanism for heat removal. Recent research has also noted, in some cases, subtle psychological and physiological changes at radiation levels below the threshold for heat-induced biological effects. However, the consensus is that most effects are thermal in nature.

High levels of RFR can affect one or more body systems or organs. Areas identified as potentially sensitive include the ocular (eye) system, reproductive system, and the immune system. Nonionizing radiation also is thought to be responsible for metabolic effects on the central nervous system and cardiac system.

In spite of these studies, many of which are ongoing, there is still no clear evidence in Western literature that exposure to medium-level nonionizing

radiation results in detrimental effects. Russian findings, on the other hand, suggest that occupational exposure to RFR at power densities above 1.0 mW/cm² does result in symptoms, particularly in the central nervous system.

Clearly, the jury is still out as to the ultimate biological effects of RFR. Until the situation is better defined, however, the assumption must be made that potentially serious effects can result from excessive exposure. Compliance with existing standards should be the minimum goal, to protect members of the public as well as facility employees.

13.4.1 NEPA Mandate

The National Environmental Policy Act of 1969 required the Federal Communications Commission to place controls on nonionizing radiation. The purpose was to prevent possible harm to the public at large and to those who must work near sources of the radiation. Action was delayed because no hard and fast evidence existed that low- and medium-level RF energy is harmful to human life. Also, there was no evidence showing that radio waves from radio and TV stations did not constitute a health hazard.

During the delay, many studies were carried out in an attempt to identify those levels of radiation that might be harmful. From the research, suggested limits were developed by the American National Standards Institute (ANSI) and stated in the document known as ANSI C95.1-1982. The protection criteria outlined in the standard are shown in Figure 13.7.

The energy-level criteria were developed by representatives from a number of industries and educational institutions after performing research on the possible effects of nonionizing radiation. The projects focused on absorption

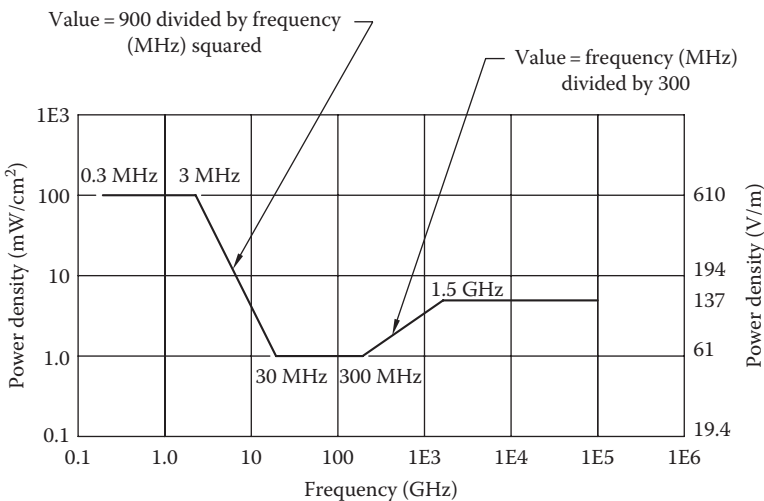


FIGURE 13.7
Power density limits for nonionizing radiation exposure for humans.

of RF energy by the human body, based upon simulated human body models. In preparing the document, ANSI attempted to determine those levels of incident radiation that would cause the body to absorb less than 0.4 W/kg of mass (averaged over the whole body) or peak absorption values of 8 W/kg over any 1 g of body tissue.

From the data, the researchers found that energy would be absorbed more readily at some frequencies than at others. The absorption rates were found to be functions of the size of a specific individual and the frequency of the signal being evaluated. It was the result of these absorption rates that culminated in the shape of the *safe curve* shown in the figure. ANSI concluded that no harm would come to individuals exposed to radio energy fields, as long as specific values were not exceeded when averaged over a period of 0.1 h. It was also concluded that higher values for a brief period would not pose difficulties if the levels shown in the standard document were not exceeded when averaged over the 0.1 h time period.

The FCC adopted ANSI C95.1-1982 as a standard that would ensure adequate protection to the public and to industry personnel who are involved in working around RF equipment and antenna structures.

13.4.2 Revised Guidelines

The ANSI C95.1-1982 standard was intended to be reviewed at 5 year intervals. Accordingly, the 1982 standard was due for reaffirmation or revision in 1987. The process was indeed begun by ANSI, but was handed off to the Institute of Electrical and Electronics Engineers (IEEE) for completion. In 1991, the revised document was completed and submitted to ANSI for acceptance as ANSI/IEEE C95.1-1992.

The IEEE standard incorporated changes from the 1982 ANSI document in four major areas:

- An additional safety factor was provided in certain situations. The most significant change was the introduction of new *uncontrolled* (public) exposure guidelines, generally established at one-fifth of the *controlled* (occupational) exposure guidelines. Figure 13.8 illustrates the concept for the microwave frequency band.
- For the first time, guidelines were included for body currents; examination of the electric and magnetic fields were determined to be insufficient to determine compliance.
- Minor adjustments were made to occupational guidelines, including relaxation of the guidelines at certain frequencies and the introduction of *breakpoints* at new frequencies.
- Measurement procedures were changed in several aspects, most notably with respect to spatial averaging and to minimum separation from reradiating objects and structures at the site.

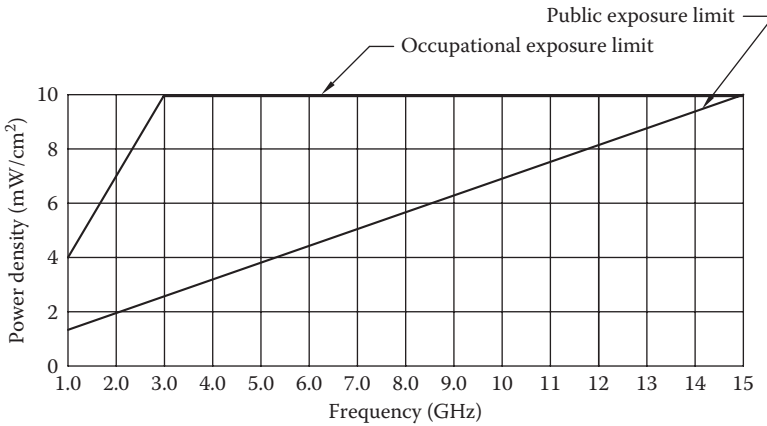


FIGURE 13.8
ANSI/IEEE exposure guidelines for microwave frequencies.

IEEE C95.1-2005, “IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Fields, 3 kHz to 300 GHz,” represents the culmination of a 14 year effort [2]. This standard, a major revision of IEEE Standard C95.1-1991, is the fourth revision of the original C95.1 standard. As with the earlier documents, the 2005 revision is designed to “protect against adverse health effects in human beings associated with exposure to electric, magnetic, and electromagnetic fields in the range of 3 kHz to 300 GHz.” The basic restrictions for frequencies between 100 kHz and 3 GHz are expressed in terms of a *specific absorption rate* (SAR)—both whole-body-averaged and peak spatial average. From these are derived the *maximum permissible exposure* (MPE), expressed in terms of more easily determined quantities, for example, incident field strength and power density. The basic rules of C95.1 are presented in graphical form in Figure 13.9.

The revised guidelines are complex and beyond the scope of this handbook. Refer to the ANSI/IEEE document for details.

13.4.3 Multiple-User Sites

At a multiple-user site, the responsibility for assessing the RFR situation—although officially triggered by either a new user or the license renewal of all site tenants—is, in reality, the joint responsibility of all the site tenants. In a multiple-user environment involving various frequencies, and various protection criteria, compliance is indicated when the fraction of the RFR limit within each pertinent frequency band is established and added to the sum of all the other fractional contributions. The sum must not be greater than 1.0. Evaluating the multiple-user environment is not a simple matter, and corrective actions, if indicated, may be quite complex.

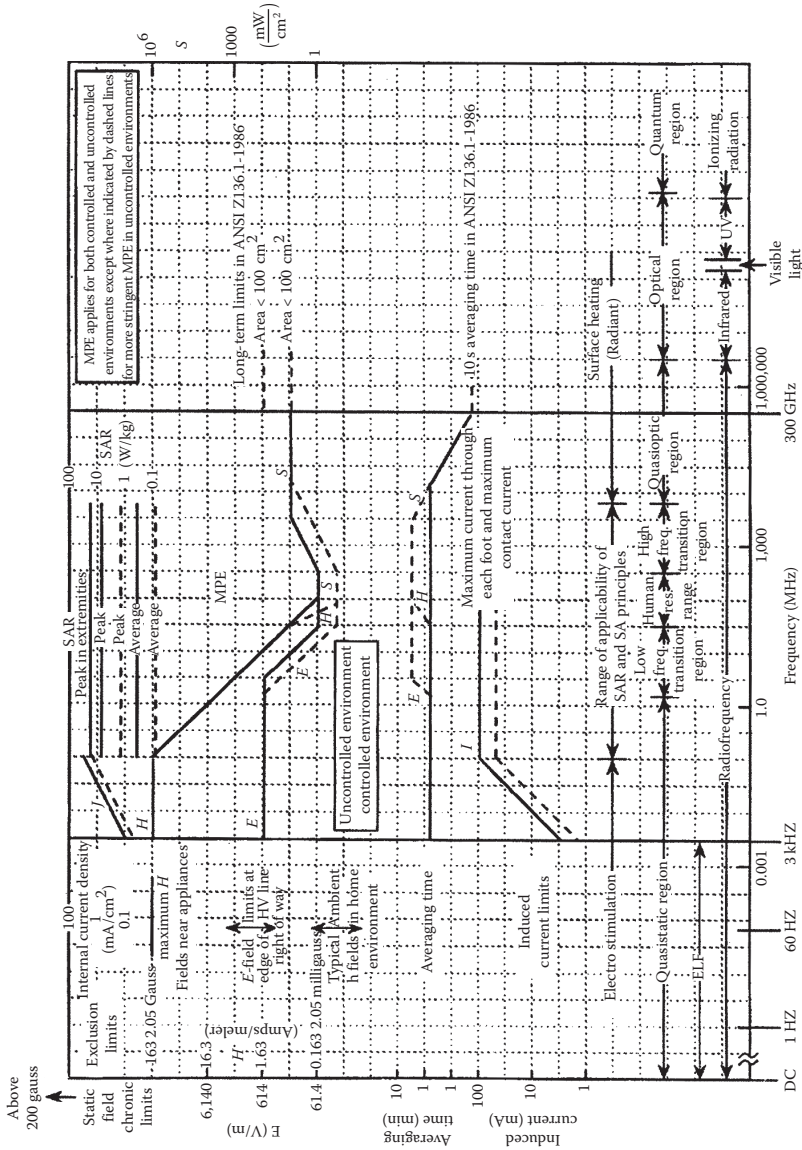


FIGURE 13.9 Capsule guide to the IEEE C95.1 standard showing the relevant features. (From Osepchuk, J.M. and R.C. Peterson, Safety and environmental issues, in *RF and Microwave Applications and Systems*, M. Golio (ed.), CRC Press, Boca Raton, FL, 2008. Used with permission.)

13.4.4 Operator Safety Considerations

RF energy must be contained properly by shielding and transmission lines. All input and output RF connections, cables, flanges, and gaskets must be RF leakproof. The following guidelines should be followed at all times:

- Never operate a power tube without a properly matched RF energy absorbing load attached.
- Never look into or expose any part of the body to an antenna or open RF generating tube, circuit, or RF transmission system that is energized.
- Monitor the RF system for radiation leakage at regular intervals and after servicing.

13.4.5 X-Ray Radiation Hazard

The voltages typically used in microwave tubes are capable of producing dangerous x-rays. As voltages increase beyond 15 kV, metal-body tubes are capable of producing progressively more dangerous radiation. Adequate x-ray shielding must be provided on all sides of such tubes, particularly at the cathode and collector ends, as well as at the modulator and pulse transformer tanks (as appropriate). High-voltage tubes never should be operated without adequate x-ray shielding in place. The x-ray radiation of the device should be checked at regular intervals and after servicing.

13.4.6 Implosion Hazard

Because of the high internal vacuum in power grid and microwave tubes, the glass or ceramic output window or envelope can shatter inward (implode) if struck with sufficient force or exposed to sufficient mechanical shock. Flying debris can result in bodily injury, including cuts and puncture wounds. If the device is made of BeO ceramic, implosion may produce highly toxic dust or fumes.

In the event of such an implosion, assume that toxic BeO ceramic is involved unless confirmed otherwise.

13.4.7 Hot Coolant and Surfaces

Extreme heat occurs in the electron collector of a microwave tube and the anode of a power grid tube during operation. Coolant channels used for water or vapor cooling also can reach high temperatures (boiling—100°C—and above), and the coolant is typically under pressure (as high as 100 psi). Some devices are cooled by boiling the coolant to form steam.

Contact with hot portions of the tube or its cooling system can scald or burn. Carefully check that all fittings and connections are secure, and monitor back pressure for changes in cooling system performance. If back pressure is increased above normal operating values, shut the system down and clear the restriction.

For a device whose anode or collector is air cooled, the external surface normally operates at a temperature of 200°C–300°C. Other parts of the tube also may reach high temperatures, particularly the cathode insulator and the cathode/heater surfaces. All hot surfaces remain hot for an extended time after the tube is shut off. To prevent serious burns, take care to avoid bodily contact with these surfaces during operation and for a reasonable cool-down period afterward.

13.4.8 Polychlorinated Biphenyls

PCBs belong to a family of organic compounds known as *chlorinated hydrocarbons*. Virtually all PCBs in existence today have been synthetically manufactured. PCBs have a heavy oil-like consistency, high boiling point, a high degree of chemical stability, low flammability, and low electrical conductivity. These characteristics resulted in the widespread use of PCBs in high-voltage capacitors and transformers. Commercial products containing PCBs were widely distributed between 1957 and 1977 under several trade names including

- Aroclor
- Pyroclor
- Sanotherm
- Pyranol
- Askarel

Askarel is also a generic name used for nonflammable dielectric fluids containing PCBs. Table 13.5 lists some common trade names used for Askarel. These trade names typically will be listed on the nameplate of a PCB transformer or capacitor.

PCBs are harmful because once they are released into the environment, they tend not to break apart into other substances. Instead, PCBs persist, taking several decades to slowly decompose. By remaining in the environment, they can be taken up and stored in the fatty tissues of all organisms, from which they are slowly released into the bloodstream. Therefore, because of the storage in fat, the concentration of PCBs in body tissues can increase with time, even though PCB exposure levels may be quite low. This process is called *bioaccumulation*. Furthermore, as PCBs accumulate in the tissues of simple organisms, and as they are consumed by progressively higher organisms, the concentration increases. This process is called

TABLE 13.5

Commonly Used Names for PCB Insulating Material

PCB Trade Names					
Apiolio	Abestol	Askarel ^a	Aroclor B	Chlorexto	Chlophen
Chlorinol	Clorphon	Diaclor	DK	Dykanol	EEC-18
Elemex	Eucarel	Fenclor	Hyvol	Inclor	Inerteen
Kanechlor	No-Flamol	Phenodlor	Pydraul	Pyralene	Pyranol
Pyroclor	Sal-T-Kuhl	Santothern FR	Santovac	Solvol	Thermin

^a Generic name used for nonflammable dielectric fluids containing PCBs.

biomagnification. These two factors are especially significant because PCBs are harmful even at low levels. Specifically, PCBs have been shown to cause chronic (long-term) toxic effects in some species of animals and aquatic life. Well-documented tests on laboratory animals show that various levels of PCBs can cause reproductive effects, gastric disorders, skin lesions, and cancerous tumors.

PCBs may enter the body through the lungs, the gastrointestinal tract, and the skin. After absorption, PCBs are circulated in the blood throughout the body and stored in fatty tissues and a variety of organs, including the liver, kidneys, lungs, adrenal glands, brain, heart, and skin.

The health risk from PCBs lies not only in the PCB itself, but also in the chemicals that develop when PCBs are heated. Laboratory studies have confirmed that PCB by-products, including *polychlorinated dibenzofurans* (PCDFs) and *polychlorinated dibenzo-p-dioxins* (PCDDs), are formed when PCBs or *chlorobenzenes* are heated to temperatures ranging from approximately 900°F to 1300°F. Unfortunately, these products are more toxic than PCBs themselves.

13.4.8.1 Governmental Action

The U.S. Congress took action to control PCBs in October 1975 by passing the Toxic Substances Control Act (TSCA). A section of this law specifically directed the EPA to regulate PCBs. Three years later the Environmental Protection Agency (EPA) issued regulations to implement the congressional ban on the manufacture, processing, distribution, and disposal of PCBs. Since that time, several revisions and updates have been issued by the EPA. One of these revisions, issued in 1982, specifically addressed the type of equipment used in industrial plants and transmitting stations. Failure to properly follow the rules regarding the use and disposal of PCBs has resulted in high fines and even jail sentences.

Although PCBs are no longer being produced for electrical products in the United States, there are thousands of PCB transformers and small PCB

capacitors still in use or in storage. The threat of widespread contamination from PCB fire-related incidents is one reason behind the EPA's efforts to reduce the number of PCB products in the environment. The users of high-power equipment are affected by the regulations primarily because of the widespread use of PCB transformers and capacitors. These components usually are located in older (pre-1979) systems, so this is the first place to look for them. However, some facilities also maintain their own primary power transformers. Table 13.6 lists the primary classifications of PCB devices.

TABLE 13.6

Definition of PCB Terms as Identified by the EPA

Term	Definition	Examples
PCB	Any chemical substance that is limited to the biphenyl molecule that has been chlorinated to varying degrees, or any combination of substances that contain such substances	PCB dielectric fluids, PCB heat-transfer fluids, PCB hydraulic fluids, and 2,2',4-trichlorobiphenyl
PCB article	Any manufactured article, other than a PCB container, that contains PCBs and whose surface has been in direct contact with PCBs	Capacitors, transformers, electric motors, pumps, and pipes
PCB container	A device used to contain PCBs or PCB articles, and whose surface has been in direct contact with PCBs	Packages, cans, bottles, bags, barrels, drums, and tanks
PCB article container	A device used to contain PCB articles or equipment, and whose surface has not been in direct contact with PCBs	Packages, cans, bottles, bags, barrels, drums, and tanks
PCB equipment	Any manufactured item, other than a PCB container or PCB article container, that contains a PCB article or other PCB equipment	Microwave systems, fluorescent light ballasts, and electronic equipment
PCB item	Any PCB article, PCB article container, PCB container, or PCB equipment that deliberately or unintentionally contains, or has as a part of it, any PCBs	
PCB transformer	Any transformer that contains PCBs in concentrations of 500 ppm or greater	
PCB contaminated	Any electric equipment that contains more than 50 but less than 500 ppm of PCBs. (Oil-filled electric equipment other than circuit breakers, reclosers, and cable whose PCB concentration is unknown must be assumed to be PCB-contaminated electric equipment)	Transformers, capacitors, contaminated circuit breakers, reclosers, voltage regulators, switches, cables, and electromagnets

References

1. National Electrical Code, NFPA #70.
2. Osepchuk, J. M. and R. C. Peterson, Safety and environmental issues, in *RF and Microwave Applications and Systems*, M. Golio (ed.), CRC Press, Boca Raton, FL, 2008.

Bibliography

Code of Federal Regulations, 40, Part 761.

Current intelligence bulletin #45, National Institute for Occupational Safety and Health, Division of Standards Development and Technology Transfer, February 24, 1986.

Electrical Standards Reference Manual, U.S. Department of Labor, Washington, DC.

Hammar, W., *Occupational Safety Management and Engineering*, Prentice Hall, New York.

Hammett, W. F., Meeting IEEE C95.1-1991 requirements, in *NAB 1993 Broadcast Engineering Conference Proceedings*, National Association of Broadcasters, Washington, DC, pp. 471–476, April 1993.

Markley, D., Complying with RF emission standards, *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1986.

Occupational injuries and illnesses in the United States by industry, OSHA Bulletin 2278, U.S. Department of Labor, Washington, DC, 1985.

OSHA, Electrical hazard fact sheets, U.S. Department of Labor, Washington, DC, January 1987.

OSHA, *Handbook for Small Business*, U.S. Department of Labor, Washington, DC.

Pfrimmer, J., Identifying and managing PCBs in broadcast facilities, *1987 NAB Engineering Conference Proceedings*, National Association of Broadcasters, Washington, DC, 1987.

Safety Precautions, Publication no. 3386A, Varian Associates, Palo Alto, CA, March 1985.

Smith, M. K., Jr., RF radiation compliance, in *Proceedings of the Broadcast Engineering Conference*, Society of Broadcast Engineers, Indianapolis, IN, 1989.

Toxics Information Series, Office of Toxic Substances, July 1983.

Whitaker, J. C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1990.

Whitaker, J. C., *Maintaining Electronic Systems*, CRC Press, Boca Raton, FL, 1991.

Whitaker, J. C., G. DeSantis, and C. Paulson, *Interconnecting Electronic Systems*, CRC Press, Boca Raton, FL, 1993.

Whitaker, J. C., *AC Power Systems*, 2nd edn., CRC Press, Boca Raton, FL, 1998.

Appendix: Mathematics, Symbols, and Physical Constants

TABLE A.1

Standard Units

Name	Symbol	Quantity
Ampere	A	Electric current
Ampere per meter	A/m	Magnetic field strength
Ampere per square meter	A/m ²	Current density
Becquerel	Bq	Activity (of a radionuclide)
Candela	Cd	Luminous intensity
Coulomb	C	Electric charge
Coulomb per kilogram	C/kg	Exposure (x- and gamma-rays)
Coulomb per square meter	C/m ²	Electric flux density
Cubic meter	m ³	Volume
Cubic meter per kilogram	m ³ /kg	Specific volume
Degree Celsius	°C	Celsius temperature
Farad	F	Capacitance
Farad per meter	F/m	Permittivity
Henry	H	Inductance
Henry per meter	H/m	Permeability
Hertz	Hz	Frequency
Joule	J	Energy, work, and quantity of heat
Joule per cubic meter	J/m ³	Energy density
Joule per Kelvin	J/K	Heat capacity
Joule per kilogram Kelvin	J/(kg K)	Specific heat capacity
Joule per mole	J/mol	Molar energy
Kelvin	K	Thermodynamic temperature
Kilogram	kg	Mass
Kilogram per cubic meter	kg/m ³	Density, mass density
Lumen	lm	Luminous flux
Lux	lx	Luminance
Meter	m	Length
Meter per second	m/s	Speed, velocity
Meter per second square	m/s ²	Acceleration
Mole	mol	Amount of substance
Pascal	Pa	Pressure, stress

(continued)

TABLE A.1 (continued)

Standard Units

Name	Symbol	Quantity
Pascal second	Pa s	Dynamic viscosity
Radian	rad	Plane angle
Radian per second	rad/s	Angular velocity
Radian per second squared	rad/s ²	Angular acceleration
Second	s	Time
Siemens	S	Electric conductance
Square meter	m ²	Area
Steradian	sr	Solid angle
Tesla	T	Magnetic flux density
Volt	V	Electrical potential
Volt per meter	V/m	Electric field strength
Watt	W	Power, radiant flux
Watt per meter Kelvin	W/(m K)	Thermal conductivity
Watt per square meter	W/m ²	Heat (power) flux density
Weber	Wb	Magnetic flux

TABLE A.2

Standard Prefixes

Multiple	Prefix	Symbol
10 ¹⁸	Exa	E
10 ¹⁵	Peta	P
10 ¹²	Tera	T
10 ⁹	Giga	G
10 ⁶	Mega	M
10 ³	Kilo	K
10 ²	Hecto	H
10	Deka	Da
10 ⁻¹	Deci	d
10 ⁻²	Centi	c
10 ⁻³	Milli	m
10 ⁻⁶	Micro	μ
10 ⁻⁹	Nano	n
10 ⁻¹²	Pico	P
10 ⁻¹⁵	Femto	f
10 ⁻¹⁸	Atto	a

TABLE A.3

Common Standard Units for Electronics Work

Unit	Symbol
Centimeter	cm
Cubic centimeter	cm ³
Cubic meter per second	m ³ /s
Gigahertz	GHz
Gram	g
Kilohertz	kHz
Kilohm	kΩ
Kilojoule	kJ
Kilometer	km
Kilovolt	kV
Kilovoltampere	kVA
Kilowatt	kW
Megahertz	MHz
Megavolt	MV
Megawatt	MW
Megohm	MΩ
Microampere	μA
Microfarad	μF
Microgram	μg
Microhenry	μH
Microsecond	μs
Microwatt	μW
Milliampere	mA
Milligram	mg
Millihenry	mH
Millimeter	mm
Millisecond	ms
Millivolt	mV
Milliwatt	mW
Nanoampere	nA
Nanofarad	nF
Nanometer	nm
Nanosecond	ns
Nanowatt	nW
Picoampere	pA
Picofarad	pF
Picoseconds	ps
Picowatt	pW

TABLE A.4

Power Conversion Factors (Decibels to Watts)

dBm	dBw	Watts	Multiple	Prefix
+150	+120	1,000,000,000,000	10^{12}	1 terawatt
+140	+110	100,000,000,000	10^{11}	100 gigawatts
+130	+100	10,000,000,000	10^{10}	10 gigawatts
+120	+90	1,000,000,000	10^9	1 gigawatt
+110	+80	100,000,000	10^8	100 megawatts
+100	+70	10,000,000	10^7	10 megawatts
+90	+60	1,000,000	10^6	1 megawatt
+80	+50	100,000	10^5	100 kilowatts
+70	+40	10,000	10^4	10 kilowatts
+60	+30	1,000	10^3	1 kilowatt
+50	+20	100	10^2	1 hectowatt
+40	+10	10	10	1 decawatt
+30	0	1	1	1 watt
+20	-10	0.1	10^{-1}	1 deciwatt
+10	-20	0.01	10^{-2}	1 centiwatt
0	-30	0.001	10^{-3}	1 milliwatt
-10	-40	0.0001	10^{-4}	100 microwatts
-20	-50	0.00001	10^{-5}	10 microwatts
-30	-60	0.000001	10^{-6}	1 microwatt
-40	-70	0.0000001	10^{-7}	100 nanowatts
-50	-80	0.00000001	10^{-8}	10 nanowatts
-60	-90	0.000000001	10^{-9}	1 nanowatt
-70	-100	0.0000000001	10^{-10}	100 picowatts
-80	-110	0.00000000001	10^{-11}	10 picowatts
-90	-120	0.000000000001	10^{-12}	1 picowatt

TABLE A.5

Voltage Standing Wave Ratio and Key Parameters

VSWR	Reflection Coefficient	Return Loss (dB)	Power Ratio	Percent Reflected
1.01:1	0.0050	46.1	0.00002	0.002
1.02:1	0.0099	40.1	0.00010	0.010
1.04:1	0.0196	34.2	0.00038	0.038
1.06:1	0.0291	30.7	0.00085	0.085
1.08:1	0.0385	28.3	0.00148	0.148
1.10:1	0.0476	26.4	0.00227	0.227
1.20:1	0.0909	20.8	0.00826	0.826
1.30:1	0.1304	17.7	0.01701	1.7
1.40:1	0.1667	15.6	0.02778	2.8
1.50:1	0.2000	14.0	0.04000	4.0
1.60:1	0.2308	12.7	0.05325	5.3
1.70:1	0.2593	11.7	0.06722	6.7
1.80:1	0.2857	10.9	0.08163	8.2
1.90:1	0.3103	10.2	0.09631	9.6
2.00:1	0.3333	9.5	0.11111	11.1
2.20:1	0.3750	8.5	0.14063	14.1
2.40:1	0.4118	7.7	0.16955	17.0
2.60:1	0.4444	7.0	0.19753	19.8
2.80:1	0.4737	6.5	0.22438	22.4
3.00:1	0.5000	6.0	0.25000	25.0
3.50:1	0.5556	5.1	0.30864	30.9
4.00:1	0.6000	4.4	0.36000	36.0
4.50:1	0.6364	3.9	0.40496	40.5
5.00:1	0.6667	3.5	0.44444	44.4
6.00:1	0.7143	2.9	0.51020	51.0
7.00:1	0.7500	2.5	0.56250	56.3
8.00:1	0.7778	2.2	0.60494	60.5
9.00:1	0.8000	1.9	0.64000	64.0
10.00:1	0.8182	1.7	0.66942	66.9
15.00:1	0.8750	1.2	0.76563	76.6
20.00:1	0.9048	0.9	0.81859	81.9
30.00:1	0.9355	0.6	0.87513	97.5
40.00:1	0.9512	0.4	0.90482	90.5
50.00:1	0.9608	0.3	0.92311	92.3

TABLE A.6

Specifications of Standard Copper Wire

Wire Size AWG	Diameter (mils)	Circular mil Area	Turns per Linear Inch ^a			Ohms per 100 ft ^b	Current- Carrying Capacity ^c	Diameter (mm)
			enam.	SCE	DCC			
1	289.3	83,810	—	—	—	0.1239	119.6	7.348
2	257.6	05,370	—	—	—	0.1563	94.8	6.544
3	229.4	62,640	—	—	—	0.1970	75.2	5.827
4	204.3	41,740	—	—	—	0.2485	59.6	5.189
5	181.9	33,100	—	—	—	0.3133	47.3	4.621
6	162.0	26,250	—	—	—	0.3951	37.5	4.115
7	144.3	20,820	—	—	—	0.4982	29.7	3.665
8	128.5	16,510	7.6	—	7.1	0.6282	23.6	3.264
9	114.4	13,090	8.6	—	7.8	0.7921	18.7	2.906
10	101.9	10,380	9.6	9.1	8.9	0.9989	14.8	2.588
11	90.7	8,234	10.7	—	9.8	1.26	11.8	2.305
12	80.8	6,530	12.0	11.3	10.9	1.588	9.33	2.063
13	72.0	5,178	13.5	—	12.8	2.003	7.40	1.828
14	64.1	4,107	15.0	14.0	13.8	2.525	5.87	1.628
15	57.1	3,257	16.8	—	14.7	3.184	4.65	1.450
16	50.8	2,583	18.9	17.3	16.4	4.016	3.69	1.291
17	45.3	2,048	21.2	—	18.1	5.064	2.93	1.150
18	40.3	1,624	23.6	21.2	19.8	6.386	2.32	1.024
19	35.9	1,288	26.4	—	21.8	8.051	1.84	0.912
20	32.0	1,022	29.4	25.8	23.8	10.15	1.46	0.812
21	28.5	810	33.1	—	26.0	12.8	1.16	0.723
22	25.3	642	37.0	31.3	30.0	16.14	0.918	0.644
23	22.6	510	41.3	—	37.6	20.36	0.728	0.573
24	20.1	404	46.3	37.6	35.6	25.67	0.577	0.511
25	17.9	320	51.7	—	38.6	32.37	0.458	0.455
26	15.9	254	58.0	46.1	41.8	40.81	0.363	0.406
27	14.2	202	64.9	—	45.0	51.47	0.288	0.361
28	12.6	160	72.7	54.6	48.5	64.9	0.228	0.321
29	11.3	127	81.6	—	51.8	81.83	0.181	0.286
30	10.0	101	90.5	64.1	55.5	103.2	0.144	0.255
31	8.9	50	101	—	59.2	130.1	0.114	0.227
32	8.0	63	113	74.1	61.6	164.1	0.090	0.202
33	7.1	50	127	—	66.3	206.9	0.072	0.180
34	6.3	40	143	86.2	70.0	260.9	0.057	0.160
35	5.6	32	158	—	73.5	329.0	0.045	0.143
36	5.0	25	175	103.1	77.0	414.8	0.036	0.127
37	4.5	20	198	—	80.3	523.1	0.028	0.113
38	4.0	16	224	116.3	83.6	659.6	0.022	0.101
39	3.5	12	248	—	86.6	831.8	0.018	0.090

^a Based on 25.4 mm.^b Ohms per 1000 ft measured at 20°C.^c Current-carrying capacity at 700 circular mils per ampere.

TABLE A.7

Celsius-to-Fahrenheit Conversion

$^{\circ}\text{Celsius}$	$^{\circ}\text{Fahrenheit}$	$^{\circ}\text{Celsius}$	$^{\circ}\text{Fahrenheit}$
-50	-58	125	257
-45	-49	130	266
-40	-40	135	275
-35	-31	140	284
-30	-22	145	293
-25	-13	150	302
-20	4	155	311
-15	5	160	320
-10	14	165	329
-5	23	170	338
0	32	175	347
5	41	180	356
10	50	185	365
15	59	190	374
20	68	195	383
25	77	200	392
30	86	205	401
35	95	210	410
40	104	215	419
45	113	220	428
50	122	225	437
55	131	230	446
60	140	235	455
65	149	240	464
70	158	245	473
75	167	250	482
80	176	255	491
85	185	260	500
90	194	265	509
95	203	270	518
100	212	275	527
105	221	280	536
110	230	285	545
115	239	290	554
120	248	295	563

TABLE A.8

Inch-to-Millimeter Conversion

Inch	0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	Inch
0	0.0	3.18	6.35	9.52	12.70	15.88	19.05	22.22	0
1	25.40	28.58	31.75	34.92	38.10	41.28	44.45	47.62	1
2	50.80	53.98	57.15	60.32	63.50	66.68	69.85	73.02	2
3	76.20	79.38	82.55	85.72	88.90	92.08	95.25	98.42	3
4	101.6	104.8	108.0	111.1	114.3	117.5	120.6	123.8	4
5	127.0	130.2	133.4	136.5	139.7	142.9	146.0	149.2	5
6	152.4	155.6	158.8	161.9	165.1	168.3	171.4	174.6	6
7	177.8	181.0	184.2	187.3	190.5	193.7	196.8	200.0	7
8	203.2	206.4	209.6	212.7	215.9	219.1	222.2	225.4	8
9	228.6	231.8	235.0	238.1	241.3	244.5	247.6	250.8	9
10	254.0	257.2	260.4	263.5	266.7	269.9	273.0	276.2	10
11	279	283	286	289	292	295	298	302	11
12	305	308	311	314	317	321	324	327	12
13	330	333	337	340	343	346	349	352	13
14	356	359	362	365	368	371	375	378	14
15	381	384	387	391	394	397	400	403	15
16	406	410	413	416	419	422	425	429	16
17	432	435	438	441	445	448	451	454	17
18	457	460	464	467	470	473	476	479	18
19	483	486	489	492	495	498	502	505	19
20	508	511	514	518	521	524	527	530	20

TABLE A.9

Millimeters-to-Decimal Inches Conversion

mm	Inches	mm	Inches	mm	Inches
1	0.039370	46	1.811020	91	3.582670
2	0.078740	47	1.850390	92	3.622040
3	0.118110	48	1.889760	93	3.661410
4	0.157480	49	1.929130	94	3.700780
5	0.196850	50	1.968500	95	3.740150
6	0.236220	51	2.007870	96	3.779520
7	0.275590	52	2.047240	97	3.818890
8	0.314960	53	2.086610	98	3.858260
9	0.354330	54	2.125980	99	3.897630
10	0.393700	55	2.165350	100	3.937000
11	0.433070	56	2.204720	105	4.133848
12	0.472440	57	2.244090	110	4.330700
13	0.511810	58	2.283460	115	4.527550
14	0.551180	59	2.322830	120	4.724400
15	0.590550	60	2.362200	125	4.921250
16	0.629920	61	2.401570	210	8.267700
17	0.669290	62	2.440940	220	8.661400
18	0.708660	63	2.480310	230	9.055100
19	0.748030	64	2.519680	240	9.448800
20	0.787400	65	2.559050	250	9.842500
21	0.826770	66	2.598420	260	10.236200
22	0.866140	67	2.637790	270	10.629900
23	0.905510	68	2.677160	280	11.032600
24	0.944880	69	2.716530	290	11.417300
25	0.984250	70	2.755900	300	11.811000
26	1.023620	71	2.795270	310	12.204700
27	1.062990	72	2.834640	320	12.598400
28	1.102360	73	2.874010	330	12.992100
29	1.141730	74	2.913380	340	13.385800
30	1.181100	75	2.952750	350	13.779500
31	1.220470	76	2.992120	360	14.173200
32	1.259840	77	3.031490	370	14.566900
33	1.299210	78	3.070860	380	14.960600
34	1.338580	79	3.110230	390	15.354300
35	1.377949	80	3.149600	400	15.748000
36	1.417319	81	3.188970	500	19.685000
37	1.456689	82	3.228340	600	23.622000
38	1.496050	83	3.267710	700	27.559000
39	1.535430	84	3.307080	800	31.496000
40	1.574800	85	3.346450	900	35.433000
41	1.614170	86	3.385820	1000	39.370000
42	1.653540	87	3.425190	2000	78.740000
43	1.692910	88	3.464560	3000	118.110000
44	1.732280	89	3.503903	4000	157.480000
45	1.771650	90	3.543300	5000	196.850000

TABLE A.10

Conversion of Common Fractions to Decimal and Millimeter Units

Common Fractions ^a	Decimal Fractions ^a	mm (Approx.)	Common Fractions	Decimal Fractions	mm (Approx.)
1/128	0.008	0.20	1/2	0.500	12.70
1/64	0.016	0.40	33/64	0.516	13.10
1/32	0.031	0.79	17/32	0.531	13.49
3/64	0.047	1.19	35/64	0.547	13.89
1/16	0.063	1.59	9/16	0.563	14.29
5/64	0.078	1.98	37/64	0.578	14.68
3/32	0.094	2.38	19/32	0.594	15.08
7/64	0.109	2.78	39/64	0.609	15.48
1/8	0.125	3.18	5/8	0.625	15.88
9/64	0.141	3.57	41/64	0.641	16.27
5/32	0.156	3.97	21/32	0.656	16.67
11/64	0.172	4.37	43/64	0.672	17.07
3/16	0.188	4.76	11/16	0.688	17.46
13/64	0.203	5.16	45/64	0.703	17.86
7/32	0.219	5.56	23/32	0.719	18.26
15/64	0.234	5.95	47/64	0.734	18.65
1/4	0.250	6.35	3/4	0.750	19.05
17/64	0.266	6.75	49/64	0.766	19.45
9/32	0.281	7.14	25/32	0.781	19.84
19/64	0.297	7.54	51/64	0.797	20.24
5/16	0.313	7.94	13/16	0.813	20.64
21/64	0.328	8.33	53/64	0.828	21.03
11/32	0.344	8.73	27/32	0.844	21.43
23/64	0.359	9.13	55/64	0.859	21.83
3/8	0.375	9.53	7/8	0.875	22.23
25/64	0.391	9.92	57/64	0.891	22.62
13/32	0.406	10.32	29/32	0.906	23.02
27/64	0.422	10.72	59/64	0.922	23.42
7/16	0.438	11.11	15/16	0.938	23.81
29/64	0.453	11.51	61/64	0.953	24.21
15/32	0.469	11.91	31/32	0.969	24.61
31/64	0.484	12.30	63/64	0.984	25.00

^a Units in inches.

TABLE A.11

Conversion Ratios for Length

Known Quantity	Multiply by	Quantity to Find
Inches (in.)	2.54	Centimeters (cm)
Feet (ft)	30	Centimeters (cm)
Yards (yd)	0.9	Meters (m)
Miles (mi)	1.6	Kilometers (km)
Millimeters (mm)	0.04	Inches (in.)
Centimeters (cm)	0.4	Inches (in.)
Meters (m)	3.3	Feet (ft)
Meters (m)	1.1	Yards (yd)
Kilometers (km)	0.6	Miles (mi)
Centimeters (cm)	10	Millimeters (mm)
Decimeters (dm)	10	Centimeters (cm)
Decimeters (dm)	100	Millimeters (mm)
Meters (m)	10	Decimeters (dm)
Meters (m)	1000	Millimeters (mm)
Dekameters (dam)	10	Meters (m)
Hectometers (hm)	10	Dekameters (dam)
Hectometers (hm)	100	Meters (m)
Kilometers (km)	10	Hectometers (hm)
Kilometers (km)	1000	Meters (m)

TABLE A.12

Conversion Ratios for Area

Known Quantity	Multiply by	Quantity to Find
Square inches (in. ²)	6.5	Square centimeters (cm ²)
Square feet (ft ²)	0.09	Square meters (m ²)
Square yards (yd ²)	0.8	Square meters (m ²)
Square miles (mi ²)	2.6	Square kilometers (km ²)
Acres	0.4	Hectares (ha)
Square centimeters (cm ²)	0.16	Square inches (in. ²)
Square meters (m ²)	1.2	Square yards (yd ²)
Square kilometers (km ²)	0.4	Square miles (mi ²)
Hectares (ha)	2.5	Acres
Square centimeters (cm ²)	100	Square millimeters (mm ²)
Square meters (m ²)	10,000	Square centimeters (cm ²)
Square meters (m ²)	1,000,000	Square millimeters (mm ²)
Ares (a)	100	Square meters (m ²)
Hectares (ha)	100	Ares (a)
Hectares (ha)	10,000	Square meters (m ²)
Square kilometers (km ²)	100	Hectares (ha)
Square kilometers (km ²)	1,000	Square meters (m ²)

TABLE A.13

Conversion Ratios for Mass

Known Quantity	Multiply by	Quantity to Find
Ounces (oz)	28	Grams (g)
Pounds (lb)	0.45	Kilograms (kg)
Tons	0.9	Tonnes (t)
Grams (g)	0.035	Ounces (oz)
Kilograms (kg)	2.2	Pounds (lb)
Tonnes (t)	100	Kilograms (kg)
Tonnes (t)	1.1	Tons
Centigrams (cg)	10	Milligrams (mg)
Decigrams (dg)	10	Centigrams (cg)
Decigrams (dg)	100	Milligrams (mg)
Grams (g)	10	Decigrams (dg)
Grams (g)	1000	Milligrams (mg)
Dekagram (dag)	10	Grams (g)
Hectogram (hg)	10	Dekagrams (dag)
Hectogram (hg)	100	Grams (g)
Kilograms (kg)	10	Hectograms (hg)
Kilograms (kg)	1000	Grams (g)
Metric tons (t)	1000	Kilograms (kg)

TABLE A.14

Conversion Ratios for Volume

Known Quantity	Multiply by	Quantity to Find
Milliliters (mL)	0.03	Fluid ounces (fl oz)
Liters (L)	2.1	Pints (pt)
Liters (L)	1.06	Quarts (qt)
Liters (L)	0.26	Gallons (gal)
Gallons (gal)	3.8	Liters (L)
Quarts (qt)	0.95	Liters (L)
Pints (pt)	0.47	Liters (L)
Cups (c)	0.24	Liters (L)
Fluid ounces (fl oz)	30	Milliliters (mL)
Teaspoons (tsp)	5	Milliliters (mL)
Tablespoons (tbsp)	15	Milliliters (mL)
Liters (L)	100	Milliliters (mL)

TABLE A.15

Conversion Ratios for Cubic Measure

Known Quantity	Multiply by	Quantity to Find
Cubic meters (m ³)	35	Cubic feet (ft ³)
Cubic meters (m ³)	1.3	Cubic yards (yd ³)
Cubic yards (yd ³)	0.76	Cubic meters (m ³)
Cubic feet (ft ³)	0.028	Cubic meters (m ³)
Cubic centimeters (cm ³)	1,000	Cubic millimeters (mm ³)
Cubic decimeters (dm ³)	1,000	Cubic centimeters (cm ³)
Cubic decimeters (dm ³)	1,000,000	Cubic millimeters (mm ³)
Cubic meters (m ³)	1,000	Cubic decimeters (dm ³)
Cubic meters (m ³)	1	Steres
Cubic feet (ft ³)	1,728	Cubic inches (in. ³)
Cubic feet (ft ³)	28.32	Liters (L)
Cubic inches (in. ³)	16.39	Cubic centimeters (cm ³)
Cubic meters (m ³)	264	Gallons (gal)
Cubic yards (yd ³)	27	Cubic feet (ft ³)
Cubic yards (yd ³)	202	Gallons (gal)
Gallons (gal)	231	Cubic inches (in. ³)

TABLE A.16

Conversion Ratios for Electrical Quantities

Known Quantity	Multiply by	Quantity to Find
Btu per minute	0.024	Horsepower (hp)
Btu per minute	17.57	Watts (W)
Horsepower (hp)	33,000	Foot-pounds per min (ft lb/min)
Horsepower (hp)	746	Watts (W)
Kilowatts (kW)	57	Btu per minute
Kilowatts (kW)	1.34	Horsepower (hp)

TABLE A.17

Basic Electrical Units

Unit	Symbol	Definition
Ampere	A	One ampere is the constant current flowing in two parallel conductors one meter apart that would produce a force of 2×10^{-7} newtons per meter of length
Coulomb	C	One coulomb is the charge that is moved in one second by a current of one ampere
Farad	F	One farad is the capacitance of a capacitor that produces a potential of one volt between the plates when charged to one coulomb
Henry	H	One henry is the inductance of a coil that has one volt induced in it when the current varies uniformly at one ampere per second
Joule	J	One joule is the work done by a force of one newton acting over a distance of one meter
Ohm	Ω	One ohm is the resistance between two points of a conductor that produces a current of one ampere when one volt is applied between these points
Siemens	S	One siemens is the conductance between two points of a conductor that produces a current of one ampere when one volt is applied between these points. Reciprocal of the ohm (formerly mho)
Volt	V	One volt is the potential difference between two points in a wire carrying one ampere when the power dissipated between these points is one watt
Watt	W	One watt is the power that produces energy at one joule per second
Weber	Wb	One weber is the magnetic flux that produces an electromotive force of one volt in a circuit of one turn as the flux changes uniformly from maximum to zero in one second

POWER VACUUM TUBES HANDBOOK

Third Edition

Providing examples of applications, **Power Vacuum Tubes Handbook, Third Edition** examines the underlying technology of each type of power vacuum tube device in common use today. The author presents basic principles, reports on new development efforts, and discusses implementation and maintenance considerations. Supporting mathematical equations and extensive technical illustrations and schematic diagrams help readers understand the material.

Translate Principles into Specific Applications

This one-stop reference is a hands-on guide for engineering personnel involved in the design, specification, installation, and maintenance of high-power equipment utilizing vacuum tubes. It offers a comprehensive look at the important area of high-frequency/high-power applications of microwave power devices, making it possible for general principles to be translated into specific applications. Coverage includes power grid tubes—triodes, tetrodes, and pentodes—as well as microwave power tubes such as klystrons, traveling wave tubes, gyrotrons, and other high-frequency devices. These vacuum tubes are used in applications from radio broadcasting to television, radar, satellite communications, and more.

Explore a Wide Variety of Methods in Power Vacuum Tube Design

This third edition includes updates on vacuum tube technology, devices, applications, design methods, and modulation methods. It also expands its scope to cover properties of materials and RF system maintenance and troubleshooting. Explaining difficult concepts and processes clearly, this handbook guides readers in the design and selection of a power vacuum tube-based system.

 **CRC Press**
Taylor & Francis Group
an **informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

K12306

ISBN: 978-1-4398-5064-0

