Microwave and RF Vacuum Electronic Power Sources

Richard G. Carter

Microwave and RF Vacuum Electronic Power Sources

Do you design and build vacuum electron devices, or work with the systems that use them? Quickly develop a solid understanding of how these devices work with this authoritative guide, written by an author with over 50 years of experience in the field. Rigorous in its approach, it focuses on the theory and design of commercially significant types of gridded, linear-beam, crossed-field, and fast-wave tubes. Essential components such as waveguides, resonators, slow-wave structures, electron guns, beams, magnets, and collectors are also covered, as well as the integration and reliable operation of devices in microwave and RF systems. Complex mathematical analysis is kept to a minimum, and Mathcad worksheets supporting the book online aid understanding of key concepts and connect the theory with practice. Including coverage of primary sources and current research trends, this is essential reading for researchers, practitioners, and graduate students working on vacuum electron devices.

Richard G. Carter is a Professor Emeritus in the Department of Engineering at Lancaster University, and a Fellow of the IET.

'Vacuum electronic devices are marvels of technology that achieve amazing performance advantages by a sophisticated combination of a vast diversity of physics, materials, components, and engineering principles. This book finds a worthy value-add role in the pantheon of great vacuum electronics (VE) literature by masterfully marrying a comprehensive treatment with carefully crafted choices of appropriate—but not excessive—depth. Thus it informs, provides rigor, but is extremely accessible to many levels of expertise. The background treatment of electromagnetics is especially noteworthy—with its mapping of structure properties to equivalent circuit models—since the start for understanding VE is the science and design principles of electromagnetic fields in waveguides and resonators. The coverage of VE devices and critical materials is current and thus state-of-the-art, which is important since research discovery advances keep pushing the limits of the performance capabilities of modern VE. I wholeheartedly recommend this book to VE design engineers, and especially anyone who is new to this exciting field, as well as experts who are seeking new ways to teach its fundamental principles.'

John H. Booske, University of Wisconsin-Madison

'In spite of significant progress with solid state power amplifiers during the last decades, large zones in parameter space remain inaccessible by this technology and will continue to be the realm of vacuum electronics. Vacuum electronics however seems to have the undeserved reputation of being less "modern" and thus often is given less attention in University classes. The book *Microwave and RF Vacuum Electronic Power Sources* responds to the urgent need to keep and build expertise in this field. Carter accomplishes the difficult task to present the matter well-structured didactically and in sufficient depth to make this an excellent textbook, both for self-education and to accompany lectures. At the same time and thanks to its completeness, accomplished experts would certainly want it in their bookshelves as a modern reference.'

Erk Jensen, Head of RF group, CERN

'This book is a tour de force on the subject of high-power vacuum electronic devices and a "must have" for anyone in the field. Professor Carter is a leading authority on the subject and has spent his career in equal parts developing the technology and teaching it in clear and intuitive style. The book contains a comprehensive review of the principles and conventions of all types of vacuum tubes with detailed explanations and examples of the most common varieties. Part history lesson, part "how to" guide this book will be as useful to newcomers entering the field as it will be to the most experienced practitioners. The Mathcad Worksheets available online will be invaluable for those new to the subject and a starting off point for further development.'

Robert Rimmer, JLab

'Combining pedagogy, science and practice, Professor Carter brings up step by step deeper in the detailed understanding and design methods of all types of RF & Microwave Electron Devices.

A precious book for students, designers and system users who will acquire practical theory based on most advanced analytical formulas and supported by a huge list of references, as well as Mathcad worksheets. Enjoy this journey in electromagnetics and vacuum.'

Philippe Thouvenin, Thales – Microwave & Imaging Sub-Systems, *Technical Director – Microwave Sub-Systems*

'Vacuum Electronics is a dynamic and exciting field and Vacuum Electron Devices in particular remain vital building blocks in numerous critical electromagnetic systems. The new

book by Professor Richard Carter provides a solid mathematical and technological foundation for students in the field and a large knowledge base for seasoned professionals. The book includes relativistic formulations of many basic working equations, online Mathcad worksheets for instructional and practical use, and device performance characteristics helpful to systems engineers and other users.'

Richard True, L3 Technologies (Retired)

'In high-power and high-frequency systems for telecommunication, radar, remote sensing, broadcasting, processing of materials and scientific applications, vacuum electronic devices (tubes) are the only radiofrequency (RF), microwave, and millimeter-wave sources. This excellent and comprehensive monograph focuses on the theory, design, state-of-the-art and current research trends of commercially important types of gridded tubes (triodes, tetrodes, inductive output tubes), linear electron beam tubes (single-, multi- and sheet beam klystrons, travelling wave tubes), crossed field tubes (magnetrons, crossed-field amplifiers), and fast-wave microwave devices (gyro-oscillators and amplifiers, free electron masers). It is devoted to researchers, practitioners and graduate students working in the fields of R&D and application of vacuum electron devices. After the introduction of essential tube components such as electron guns, electron beam guiding systems, waveguide and slow-wave structures, resonators, output couplers, dielectric windows, and electron collectors, beam-wave interaction and the different types of microwave and RF vacuum electronic power sources are discussed in detail. Well-selected case studies for the various types of tubes and Mathcad worksheets supporting the book online (can be downloaded from the publisher's website) aid understanding of key concepts and link theory with practice. Concluding chapters on emission and breakdown phenomena, different types of magnets and system integration may help to achieve reliable operation of microwave and RF systems.'

Manfred Thumm, Karlsruhe Institute of Technology (KIT)

'This book is the author's life-long contribution to the microwave tubes' community; it is a comprehensive and systematic book including passive components and all kinds of RF vacuum devices with the analysis of cathodes, electron optics, reliability, and the system integration. This book gives both clear physical imaging, deep understanding, the author's new insight of the high frequency structure and beam-wave interaction as well as concise physical models and worksheet for engineering design with both circuit and field method using modern computer facility.'

Jinjun Feng, Tsinghua University

'Richard Carter's fifty years of internationally leading expertise is distilled in this essential reading for everyone working on high power microwave and RF sources. Researchers, graduate students and practitioners will find this monograph indispensable.

Not only does it provide an in-depth understanding of the many subtleties of the subject that will satisfy the experienced practitioner, its careful construction and clarity of expression offers the student entering this field an excellent introduction that they will continue to value and retain on their shelves as their own expertise develops.

To achieve a worthwhile understanding of this field some use of mathematics is essential and this text very thoughtfully employs the necessary mathematics to support the theoretical discussions in ways that are accessible to science and engineering graduates.

The coverage includes the most significant families of vacuum electronic microwave and RF power oscillators and amplifiers that both those working on research and development of the sources, or needing to use these sources in applications are most likely to encounter. This comprehensive list includes slow-wave sources, gridded tubes, triodes, tetrodes, inductive output tubes (IOTs), klystrons, travelling-wave tubes (TWTs), magnetrons, crossed-field

amplifiers as well as fast-wave sources such as electron cyclotron masers, gyrotrons, cyclotron auto-resonance masers, gyro-amplifers, peniotrons, ubitrons and free electron lasers. The treatments of the slow-wave devices tend to be more extensive than some of the fast-wave devices. For example although electron cyclotron masers and gyrotrons are covered in some detail the author intentionally leaves a detailed treatment of free electron lasers to the referencing of some existing publications. This balance of the coverage is well-justified, as it reflects the fact that the majority of the commercial applications of microwave and RF vacuum electronic power sources presently use, and are likely to continue to use, slow-wave devices. Consequently the students, researchers and practitioners in this field tend to be well-motivated to study and understand the range of sources that are covered in this book.

This is a "stand-alone" book since it does not assume that the reader is familiar with all of the specialized components from which vacuum electronic microwave and RF power sources are constructed. The main components such as waveguides, resonators, slow-wave structures, electron guns, beams, magnets and collectors are expertly covered in dedicated sections that add extra value. The coverage of multipactor discharges and system integration will be found to be especially useful to those interested in applying these sources.

This text is highly recommended for graduate students, practising scientists and engineers and for libraries.'

<div align="right">Alan D. R. Phelps, FRSE (Prof), University of Strathclyde</div>

'This book by Richard Carter is a comprehensive, complete, and excellent exploration of the microwave and RF vacuum electronic power source field. The content of the book has been refined and enhanced by invited lectures given in many relevant international institutes. The subject matter is not only important in its own right but also provides a splendid insight to comprehending the wider literature in the field. The mathematical analyses are kept to a minimum which encourages the reader to master the principles and establish the basis for advanced developments. The reader's understanding is enhanced by the accessibility to the worksheets used for the calculations throughout the book. Unusually if not uniquely, the author uses Mathcad as a powerful and flexible tool for modelling problems in vacuum electronics. These models are made available to the readers.

In summary, this work is an indispensable reference handbook in the field of microwaves and RF power, both for beginners and experienced professionals alike.'

<div align="right">Stephen Myers, Executive Chair ADAM SA, Former CERN Director of Accelerators and Technology</div>

'Destined to be a classic reference in the field of RF vacuum electronic device analysis and design. From density modulated power grid and induction output tubes, to slow wave linear and crossed field devices and fast wave high power gyrotrons, and their key components, every subject is treated with technical rigor interspersed with sage practical advice obtained over the course of Prof. Carter's distinguished career in the field. Notable features include the strong use of transmission line analysis for device modelling providing insight not always available with standard field analysis or computer simulation, a detailed list of references at the end of every chapter for advanced reading, and online worksheets for invaluable hands-on experience with the physics behind the equations bringing home the necessary design trades that always need to be made – an excellent learning tool. Enthusiastically recommend for both the intern/new hire as well as the seasoned RF vacuum electronics engineer and researcher.

Professor Carter says in the Introduction that the book can be considered to be fifty years in the making. I would add that it was well worth the wait – a monumental achievement.'

<div align="right">Carter M. Armstrong, L3 Technologies</div>

The Cambridge RF and Microwave Engineering Series

Series Editor

Steve C. Cripps, Distinguished Research Professor, Cardiff University

Editorial Advisory Board

James F. Buckwalter, UCSB
Jenshan Lin, University of Florida
John Wood, Maxim Integrated Products

Peter Aaen, Jaime Plá and John Wood, *Modeling and Characterization of RF and Microwave Power FETs*
Dominique Schreurs, Máirtín O'Droma, Anthony A. Goacher and Michael Gadringer (Eds), *RF Amplifier Behavioral Modeling*
Fan Yang and Yahya Rahmat-Samii, *Electromagnetic Band Gap Structures in Antenna Engineering*
Enrico Rubiola, *Phase Noise and Frequency Stability in Oscillators*
Earl McCune, *Practical Digital Wireless Signals*
Stepan Lucyszyn, *Advanced RF MEMS*
Patrick Roblin, *Nonlinear RF Circuits and the Large-Signal Network Analyzer*
Matthias Rudolph, Christian Fager and David E. Root (Eds), *Nonlinear Transistor Model Parameter Extraction Techniques*
John L. B. Walker (Ed.), *Handbook of RF and Microwave Solid-State Power Amplifiers*
Anh-Vu H. Pham, Morgan J. Chen and Kunia Aihara, *LCP for Microwave Packages and Modules*
Sorin Voinigescu, *High-Frequency Integrated Circuits*
Richard Collier, *Transmission Lines*
Valeria Teppati, Andrea Ferrero and Mohamed Sayed (Eds), *Modern RF and Microwave Measurement Techniques*
Nuno Borges Carvalho and Dominique Schreurs, *Microwave and Wireless Measurement Techniques*
David E. Root, Jason Horn, Jan Verspecht and Mihai Marcu, *X-Parameters*
Earl McCune, *Dynamic Power Supply Transmitters*
Hossein Hashemi and Sanjay Raman (Eds), *Silicon mm-Wave Power Amplifiers and Transmitters*
T. Mitch Wallis and Pavel Kabos, *Measurement Techniques for Radio Frequency Nanoelectronics*
Giovanni Ghione and Marco Pirola, *Microwave Electronics*
Isar Mostafanezhad, Olga Boric-Lubecke and Jenshan Lin (Eds), *Medical and Biological Microwave Sensors*
Richard G. Carter, *Microwave and RF Vacuum Electronic Power Sources*

Forthcoming

José Carlos Pedro, David Root, Jianjun Xu and Luis Cotimos Nunes, *Nonlinear Circuit Simulation and Modeling*
Michael Schröter and Martin Claus, *Carbon Nanotube Electronics for Analog Radio-Frequency Applications*

Microwave and RF Vacuum Electronic Power Sources

RICHARD G. CARTER

Lancaster University

To my dear wife Awena
without whose loving support
this book could never have been written

Contents

Preface

It could be said that it has taken me fifty years to write this book. I entered the field of vacuum electronics as a PhD student in 1965 at the University College of North Wales, Bangor. That date coincided with the end of what I think of as the heroic period of the field. The years from 1939 to 1965 saw contributions to the subject from people whose names are well-known to those who followed them. Their achievements in the exploration of a new technology were remarkable, especially when one considers that nearly everything had to be done by theory and experiment. Computer modelling was in its infancy. The period since 1965 has seen an explosion in the power of computers and what can be achieved with them. I took my first steps in computer programming on the EDSAC II vacuum tube computer at Cambridge University using a very primitive autocode. The point has now been reached where computing power is not only sufficient to model devices in great detail, but also to permit automatic optimisation of them. During my career computer modelling of tubes became increasingly important, and that was the main focus of my own work, and of those who worked with me.

From Bangor I moved to what was then the English Electric Valve Company (now Teledyne e2v) in Chelmsford. The culture of the company at that period has been well-described by Norman Pond in his book *The Tube Guys*. There was no separation between development and production so that we were always conscious that the tubes we were developing had to be made by the people alongside whom we worked. I was fortunate to be seconded to a development project at the Services Electronics Research Laboratory, where, under the able guidance of Robin King, I undertook the detailed design of most parts of a C-band coupled-cavity TWT. Most of the design calculations were carried out on an electro-mechanical calculator on my desk. But I did take over some FORTRAN programs for modelling TWTs, which had been written by Harold Curnow. Those programs had to be run via an unreliable landline, using a chattering teletype, on a mainframe computer 50 miles away.

After three years I moved to the newly formed Engineering Department at Lancaster University. The Department was based on the idea that engineering is a single discipline, which fitted the multidisciplinary nature of vacuum electronics rather well. I was encouraged to continue work in the field when I was offered an appointment as a consultant by EEV. That remained an important relationship for nearly thirty years, keeping me in touch with the problems that mattered to my

colleagues in industry. The relationship led to the creation of an innovative undergraduate course in Physical Electronic Engineering, sponsored by EEV. I developed and taught a module on vacuum electronics for this course. The students undertook industrial placements with the company and my role in setting up, and monitoring, those activities gave me contacts in nearly every department there. The creation of the course was ably supported by Dr John Whitmore (at that time Manager of the Light Conversion Division). John also encouraged me to start to build my own research group, which became a major focus of my life until I retired in 2009. In 1989 I was invited by the company to prepare a set of twenty videotaped lectures to be used for training purposes. The project was the brainchild of Maurice Esterson and I was given invaluable practical help by F.J. (Peter) Weaver. Those lectures could be said to be the true foundation of the present book.

Over time my research group flourished. I was especially grateful to Maurice Esterson and Brian Cooper of EEV who supported research projects, and made a point of involving me in the technical problems of their staff. Over the period until I retired nearly fifty research students, assistants, and fellows worked with me, and I owe a debt of gratitude to them for the way they forced me to try to think clearly about their work in order to guide them. There were collaborative projects with CEERI in India and with BVERI, and a number of other institutions, in China. It is a matter of pride to me that several of the people who were members of the group have since occupied senior positions in laboratories overseas. The material which has gone on to form the content of this book was honed through lectures I was invited to give at around sixteen institutions, in nine countries, on three continents. These included contributions to five CERN Accelerator Schools to which I was introduced by colleagues at the Daresbury Laboratory. I also contributed lectures, with colleagues at the universities of Lancaster and Strathclyde, to a joint MSc in High Power Radiofrequency Engineering.

The idea of trying to collect my thoughts on vacuum electronics into the form of a book has been around for a long time. But I was unable to find time to devote to the task until after I retired. Throughout my career my interest has been in the modelling and conceptual design of tubes. In teaching I have always tried to keep close to a physical understanding of how tubes operate. And I have been interested in developing simple models as an aid to understanding in the conceptual phase of design, following the example of the pioneers in the field. I have been influenced, especially, by the books by Pierce, Slater, Gittins, Beck, and Gewartowski and Watson. The literature in the field is now too extensive for any one person to master. The references cited in this book are the sources which I consulted while writing it. My aim was to try to bring this material together to give a comprehensive and consistent account of the subject. The topics I have chosen to include are those I have found useful. I hope that readers will find that this material is not only useful in its own right but also provides a route to understanding the wider literature in the field.

Microwave engineering has often been a happy hunting ground for those who love advanced mathematics. Although it is not possible to discuss the topics in this

book without some use of mathematics I have tried to make the treatment accessible. I am impressed by the thought that engineers generally try to find simple ways of modelling complex problems. We use 'j-notation' to allow us to apply DC circuit theory to AC problems. And we use transmission line, and equivalent circuit methods to model problems involving electromagnetic fields. This way of thinking about vacuum electronics remains as useful today as it was when I entered the field fifty years ago. It provides insight into the physics of devices which cannot readily be gained in any other way. I believe it to be an essential foundation for training new entrants to the field whose knowledge will, otherwise, come largely from computer simulation using general purpose electromagnetics software and particle-in-cell codes. After writing in FORTRAN for many years I became an enthusiast for the possibilities of modelling with software such as Mathcad. Although this approach is not suitable for all problems it does provide a remarkably powerful and flexible way of modelling problems in vacuum electronics. Over time I have developed Mathcad models for many aspects of tube design and I have used these extensively whilst writing this book. The models which I used are listed in the Appendix, and can be downloaded from the publisher's website. I would like to encourage my readers to use those models to explore aspects of tube design. The use of Mathcad makes it easy for the user to modify the models, and to add diagnostic features, as part of their own learning process.

I have been grateful to many people for help, advice, and inspiration at different stages of my career. There are too many to name them all, but I particularly want to mention: Rob Newton, my PhD supervisor, from whom I learnt the value of space-charge wave theory and coupled-mode analysis; Dudley Perring from ESTEC who introduced me to the world of space TWTs, becoming simultaneously a project sponsor and a research student until his untimely death; and Rodney Vaughan, from whom I learnt that good mathematical models of tubes were best kept close to the underlying physics. I would also like to acknowledge the colleagues at EEV, TMD technologies, Thales at Ulm and Vélizy, CERN, DESY, and the Daresbury Laboratory who generously gave time to advise on research projects. There have been many fruitful discussions with my colleagues at Lancaster, especially Graeme Burt, Amos Dexter, Claudio Paoloni, and Robin Tucker.

During the writing of this book I have been most grateful to colleagues who reviewed chapters in draft and made many helpful comments and suggestions. In alphabetical order: Carter Armstrong, Heinz Bohlen, Bruce Carlsten, Jim Dayton, Amos Dexter, Tushar Ghosh, Dan Goebel, Michel Grézaud, Neil Marks, John Ospechuk, Alan Phelps, Alan Sangster, Manfred Thumm, Mike Tracy, Todd Treado, and Dick True. I hope they will forgive me if I have not always taken their advice. The book, as it stands, is entirely my responsibility. I would like to thank Lancaster University for appointing me as an Emeritus Professor when I retired. That gave me continued access to the information and computer resources of the University. I am also grateful to Julie Lancashire and her colleagues at Cambridge University Press who have supported me consistently in a project that has taken much longer than either they or I expected.

Finally, I owe an immense debt of gratitude to my wife Dr Awena Carter, whose name means 'inspiration' in Welsh, for her unstinting love and support throughout the whole of my career. She has accompanied me to conferences, and on visits to institutions, and is known to many of my colleagues in the field. She has borne patiently (most of the time) with the long hours I have spent working on the book, and she brought her considerable expertise in writing to bear by reading the whole text in draft and pointing out where its clarity could be improved.

Principal Roman Symbols

A		Magnetic vector potential
A		Amplification (voltage gain)
		Area
a		Waveguide width
		Tunnel radius
		Grid spacing
B		Magnetic flux density $= \mu\mathbf{H}$
	B_B	Brillouin field
	B_0	Characteristic field
B		Susceptance
		Bandwidth
	B_b	Beam loading susceptance
b		Waveguide height
		Beam radius
		Pierce velocity parameter
C		Capacitance
		Pierce gain parameter
c		Velocity of light $= 0.2998 \times 10^9$ m s^{-1}
D		Electric displacement $= \varepsilon\mathbf{E}$
d		Dielectric radius
		Distance
E		Electric field
	E_F	Fermi energy
	E_m	Space harmonic amplitude
e		Elementary charge $= 1.602 \times 10^{-19}$ C
		Base of natural logarithms $= 2.718$
f		Frequency
G		Conductance $= 1/R$
		Power gain
	G_{dB}	Gain in decibels $= 10\ln G$
	G_b	Beam loading conductance
	$G_0 = I_0/V_0$	Beam conductance
g		Gap length

H		Magnetic field
h		Height
		Planck's constant $= 6.626 \times 10^{-34}$ J s
I		Electric current
		Bessel function
		Intensity
	I_a	Anode current
	I_g	Grid current
	I_0	DC current
	I_n	Harmonic current $n = 1, 2, 3 \ldots$
	$I_n(z)$	Modified Bessel function of the first kind
	$I_{\pm}$	Space-charge wave RF currents
i		Current
	i_g	Induced current in a gap
J		Current density
		Bessel function
	$J_n(z)$	Bessel function of the first kind
	J_C	Child-Langmuir current density
	J_S	Schottky current density
K		Bessel function
		Coupling factor
		Perveance
		Ratio of flux linkages $= \sqrt{1 - 1/m^2}$
	$K_n(z)$	Modified Bessel function of the second kind
k		Boltzmann constant $= 1.380 \times 10^{23}$ J K^{-1}
		Coupling factor
		Free space propagation factor $= \omega/c$
L		Length
		Self inductance
		Cold loss
l		Length
M		Mutual inductance
		Gap coupling factor
M		Magneto-motive force
m		Mass
	m	Normalised magnetic flux density $= B/B_B$
	m_0	Rest mass of the electron $= 9.109 \times 10^{-31}$ kg
n		Harmonic number
P		Power
		Pressure
		Permeance (= 1/reluctance)
	P_{in}	Input power
	P_{out}	Output power

p		Pitch
		Plasma frequency reduction factor
Q		Quality factor of a resonator
	Q_E	External Q
	Q_L	Loaded Q
	Q_U	Unloaded Q
	QC	Pierce space charge parameter
q		Induced current reduction factor
R		Resistance
		Spherical radius
		Normalised radius
	R	Reluctance
	R_c	Cavity shunt impedance
	R_s	Surface resistance
r		Radius
		Resistance
	r_a	Anode radius
	r_c	Cathode radius
s		Shield radius
T		Absolute temperature (K)
		Kinetic energy
t		Thickness
u		Velocity
	u_0	DC velocity
V		Voltage
		Volume
	V_a	Anode voltage
		Accelerating voltage
	V_b	Breakdown voltage
	V_c	Carrier amplitude
		Collector potential with respect to the tube body
	V_f	Final voltage
	V_g	Grid voltage
		Gap voltage
	V_H	Hull cut-off voltage
	V_i	Initial voltage
	V_{in}	Input voltage
	V_{out}	Output voltage
	V_R	Rest energy of the electron $= m_0 c^2 / e = 511\,\mathrm{keV}$
	V_s	Source voltage
	V_T	Threshold (Hartree) voltage
		Volt equivalent of temperature $= kT/e$
	V_0	DC voltage
		Characteristic voltage

V_1	AC voltage
	Beam kinetic voltage
$V_{\pm}$	Space-charge wave kinetic voltages
v	Voltage
	Velocity
$v_g = d\omega/d\beta$	Group velocity
$v_p = \omega/\beta$	Phase velocity
W	Stored energy
w	Width
X	Reactance
Y	Admittance $= 1/Z$
Y_e	Electronic admittance
Z	Impedance $= 1/Y$
	Normalised position
Z_C	Coupling impedance
Z_c	Characteristic impedance
Z_e	Electronic impedance
Z_g	Waveguide impedance
Z_P	Pierce impedance
Z_S	Surface impedance
Z_T	Total impedance
$Z_w = \sqrt{\mu/\varepsilon}$	Wave impedance
$Z_0 = \sqrt{\mu_0/\varepsilon_0}$	Wave impedance of free space

Principal Greek Symbols

α		Loss parameter
		Time constant
β		Propagation constant
	$\beta_C = \omega_C/c$	Waveguide cut-off
	$\beta_c = \omega_c/u_0$	Cyclotron propagation constant
	$\beta_e = \omega/u_0$	Electronic propagation constant
	$\beta_g = \sqrt{\beta_C^2 - \omega^2/c^2}$	Waveguide propagation constant
	$\beta_L = \omega_L/u_0$	Larmor propagation constant
	$\beta_m = \beta_0 + 2m\pi/p$	Space harmonic propagation constant
	$\beta_{mn} = \omega_{mn}/c$	Waveguide cut-off for the (m,n) mode
	$\beta_p = \omega_p/u_0$	Plasma propagation constant
	$\beta_q = \omega_q/u_0$	Reduced plasma propagation constant
	β_s	Synchronous propagation constant
	$\beta_w = \varepsilon_r \omega/c$	Propagation constant in a dielectric
	$\beta_0 = \omega/v_p$	Propagation constant in material
		Slow-wave structure propagation constant
	$\beta_0 = 2\pi/L$	Periodic electrostatic focusing
		Periodic permanent magnet focusing
	$\beta_{\pm}$	Space-charge wave propagation constants
γ		Radial propagation constant $= \sqrt{\beta^2 - \omega^2/c^2}$
		Relativistic factor $= 1/\sqrt{1 - u^2/c^2}$
δ		Skin depth
		Secondary electron emission coefficient
ε		Permittivity
	$\varepsilon_r = \varepsilon/\varepsilon_0$	Relative permittivity
	ε_0	Primary electric constant $= 8.854 \times 10^{-12}$ F m^{-1}
η		Efficiency
	η_A	Power added efficiency
	η_c	Circuit efficiency
	η_e	Electronic efficiency
	η_h	Harmonic efficiency
	η_O	Overall efficiency
	η_{rf}	RF efficiency

θ	Angle
λ	Wavelength
$\quad \lambda_C$	Waveguide cut-off wavelength
$\quad \lambda_g = 2\pi/\beta_g$	Guide wavelength
μ	Permeability
	Gap coupling factor
	Amplification factor
	Attenuation coefficient of X-rays
$\quad \mu_c$	Radial gap coupling factor
$\quad \mu_d$	Axial gap coupling factor
$\quad \mu_r$	Relative permeability
$\quad \mu_0$	Primary magnetic constant $4\pi \times 10^{-7}$ H m^{-1}
ρ	Charge density
	Resistivity
	Voltage reflection coefficient
	Density
σ	Conductivity
τ	Time constant
	Transit time
	Radial propagation constant
Φ	Magnetic flux
φ	Angle
	Phase
	Work function
ψ	Angle
	Helix pitch angle
ω	Angular frequency $= 2\pi f$
$\quad \omega_C$	Waveguide cut-off frequency
$\quad \omega_c = \dfrac{eB}{m_0}$	Cyclotron frequency
$\quad \omega_H$	Hull cut-off frequency
$\quad \omega_L = \omega_c/2$	Larmor frequency
$\quad \omega_{mn}$	Cut-off frequency of waveguide mode (m,n)
$\quad \omega_p = \sqrt{\dfrac{e\rho}{\varepsilon_0 m_0}}$	Electron plasma frequency
$\quad \omega_q = p\omega_p$	Reduced plasma frequency
$\quad \omega_s$	Synchronous frequency
	Synchronous angular velocity
$\quad \omega_0$	Resonant frequency

Abbreviations

AC	Alternating current
AM	Amplitude modulation
BWO	Backward wave oscillator
CFA	Crossed-field amplifier
CVD	Chemical vapour deposition
CW	Continuous wave
DC	Direct current
ECM	Electron cyclotron maser
ECM	Electronic counter-measures
FEL	Free electron laser
GTO	Gate turn-off thyristor
IGBT	Insulated gate bipolar transistor
IOT	Inductive output tube
PM	Phase modulation
PPM	Periodic permanent magnet
RF	Radio frequency
RMS	Root mean square
TWT	Travelling-wave tube

1   Overview

1.1   Introduction

Sources of power in the radiofrequency, microwave, and millimetre wave regions of the electromagnetic spectrum are essential components in a wide range of systems for telecommunications, broadcasting, remote sensing, and processing of materials[1]. Current research is extending the frequency range into the sub-millimetre region. These sources employ either vacuum electronic, or solid state, technologies. At the higher frequencies and power levels, vacuum electronic devices (tubes[2]) are the only sources available (see Section 1.2).

The purpose of this book is to provide a comprehensive introduction to the theory and conceptual design of the types of tubes which are of continuing importance. The design and operation of vacuum tubes requires knowledge and skills drawn both from electrical and electronic engineering and from physics. The treatment here is intended to be accessible to those whose training has been in either discipline. The use of advanced mathematics has been avoided as far as possible with considerable emphasis on the use of simple numerical methods. The book is designed to be a reference text for designers and users of vacuum tubes, and a textbook for people who are new to the field.

This is a mature field in which much has been published since its first beginnings in 1904 [1]. The sources cited here are those which have been used as the basis for the book. They are believed to comprise most of the most important sources in the field and the reader is invited to consult them for further information. References have been included to sources that provide additional information on many of the topics, but no attempt has been made to provide a comprehensive bibliography of the subject and longer lists of references are to be found elsewhere [2]. A further aim of this book is to provide the reader with the background necessary to read with understanding other papers in the field.

In any book it is necessary to make choices about what should be included and what excluded. The subjects covered are those the author believes to be important for the practical business of designing and using vacuum tubes. Because the focus

---

[1] For convenience, the term RF is used throughout this book to refer to all frequencies in the range 30 kHz to 300 GHz.

[2] We shall call these devices (vacuum) tubes because that terminology is familiar and concise.

is on theory and conceptual design very little has been said about the technology of tube construction, which is well treated elsewhere [3, 4]. Similarly, little has been said specifically about vacuum tubes for use at sub-millimetre wavelengths since they employ the same principles as those at lower frequencies, and most of the challenges are in the technology of their construction.

This chapter provides an overview of the subject of the book. The next section compares vacuum tubes with solid-state devices to show how the technologies are complementary. Section 1.3 provides an overview of the physical principles on which vacuum tubes are based and definitions of the key terms used to describe their performance. A tube converts the DC power in the initial electron stream into RF output power by interaction with electromagnetic structures. These structures support standing (resonant), or travelling, electromagnetic waves. Coupled-mode theory is introduced as a valuable conceptual tool for understanding the interactions between electron streams and travelling electromagnetic waves. The section concludes with a classification of the principal types of vacuum tube based on the preceding discussion. The principal applications of vacuum tubes are reviewed in Section 1.4 together with some of the factors which govern the availability of tubes of different types. That leads, in Section 1.5 to consideration of the communication between the designers and users of tubes in the form of a Statement of Requirements. This statement specifies both the electrical performance required, and the factors which constrain the design. Many tubes are required to amplify modulated carrier signals whose properties are not normally familiar to people whose primary discipline is Physics. An introduction to analogue and digital modulation, noise and multiplexing is provided in Section 1.6. Finally, Section 1.7 considers some of the principles of the engineering design of tubes including dimensional analysis and scaling, and the use of computer modelling.

The remainder of the book comprises four sections:

i) Chapters 2–4 deal with the properties of the passive electromagnetic components employed in vacuum tubes.
ii) Chapters 5–11 are concerned with aspects of electron dynamics in vacuum that are employed in tubes in a variety of ways. Chapter 10 also includes a discussion of methods of cooling.
iii) Chapters 12–17 show how the fundamental principles introduced earlier in the book are applied to specific types of tube, and their conceptual design.
iv) Chapters 18–20 provide an introduction to some technological issues which are common to most types of tube and their successful use in systems.

1.2   Vacuum Electronic and Solid-State Technologies

The characteristic size of any active RF device is determined by the distance travelled by the charge carriers in one RF cycle. Thus the size of a device decreases as the frequency increases and as the velocity of the charge carriers decreases. The application

of any electronic technology is limited, ultimately, by temperature since the operation of a device generates heat. Hence, the maximum continuous, or average, RF power which can be generated by a single device is determined by the power dissipation within it. At low power levels semiconductor devices have the advantages of small size, and low voltage operation. But these become disadvantages at high power levels because the current passing through the device is high. Thus there are large conduction losses, generating heat, within a small volume. They can be reduced, to some extent, by operating the device as a switch rather than in its active mode.

Transistors can currently deliver around 100 W of continuous power or 1 kW of pulsed power at frequencies in the region of 1 GHz [5–7]. Further developments, including transistors using diamond as a semiconducting material, may increase the power to several kW [8, 9]. High power amplifiers can be made by operating many transistors in parallel but the penalties of increased complexity set limits to this. The power combining can take place in space, as in active phased-array radar, where average powers up to tens of kilowatts can be achieved [10]. Alternatively, a power combining network may be used as in the 190 kW, 352 MHz, amplifier at Synchrotron SOLEIL [11]. This amplifier combines the power from four 50 kW towers each containing twenty 2.5 kW units. A unit comprises eight 315 W modules each having a pair of transistors operated in push-pull. This arrangement means that the loss of power from the failure of an individual transistor is small, and the degradation of the amplifier from this cause is gradual. However, the power output is still almost an order of magnitude less than that achieved by vacuum tubes at the same frequency, and the frequency is at the lower end of the range for which vacuum tubes have been developed. The use of low DC voltages reduces problems of reliability caused by voltage breakdown. But the consequent need for high currents leads to DC losses in the connecting bus-bars. A review of high power semiconductor RF power technology is given in [12].

Vacuum tubes, in contrast, operate at high voltages and low currents. The charge carriers have a much higher velocity than in a semiconducting material, and they are not subject to energy loss through collisions as they pass through it. Thus the active volume can be large with only RF losses within it. In many cases, the greater part of the heat generated is dissipated on electrodes which are separate from the active region, and whose size can be increased to reduce the power density. There is a common misconception that vacuum tubes are fragile, short-lived, unreliable, and inefficient. In fact, modern vacuum devices are mechanically robust and able to survive short-term electrical overloads without damage. They have demonstrated outstanding reliability and lifetimes in the very demanding environment of space. Vacuum tube amplifiers can have maximum conversion efficiencies of up to 70%, and 90% has been achieved in oscillators. However, the most reliable performance of any technology is achieved by allowing generous design margins and operating well within their limits. Vacuum tubes can be operated in parallel, in the same way as semiconductor devices, but the number of parallel devices is usually quite small [13, 14]. An exception to this is in phased-array radar where the output from a larger number of microwave power modules (see Section 20.1) is combined in space [15].
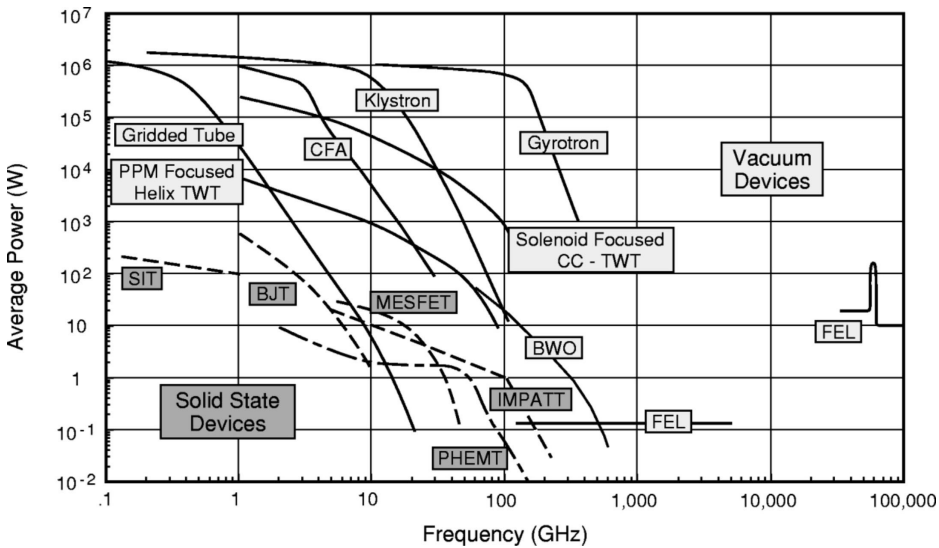
Figure 1.1: Comparison of the performance of single vacuum electronic and solid state RF power devices
(copyright 1999, IEEE, reproduced, with permission, from [16]).

Figure 1.1 shows a comparison between the performances of individual vacuum and semiconductor RF power devices [16]. There have been advances in both technologies since that figure was drawn, but the general picture remains valid. A more recent, but less detailed, figure is to be found in [17]. For further discussion of the relative merits of vacuum tube and solid-state RF power amplifiers see [18, 19]

1.3   Principles of Operation

The tubes which are the subject of this book are all power amplifiers and oscillators. For our present purpose it is sufficient to regard oscillators as amplifiers in which the RF input is provided by internal feedback. All vacuum tube amplifiers can be understood in terms of the block diagram shown in Figure 1.2. A uniform stream of electrons is emitted into the vacuum from the electron source and modulated by the RF input voltage[3]. Radiofrequency energy is extracted from the modulated stream of electrons, and their remaining energy is dissipated as heat on a collecting electrode. The arrows show the direction of motion of the electrons; the conventional current is, of course, in the opposite direction. The functions of the blocks may be combined in various ways in different devices but the overall process is essentially the same. The basic RF performance of an amplifier is defined in terms of its gain, output power, efficiency, and instantaneous or tuneable bandwidth.

---

[3] The word 'stream' is used here as being more general than the term electron beam which is used for specific types of tube.
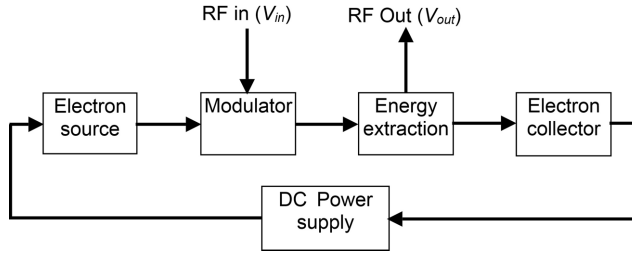
Figure 1.2: Block diagram of a vacuum tube amplifier.

1.3.1   Geometry

The majority of practical vacuum tubes have geometries which are cylindrically symmetrical, or close approximations to it. The flow of the DC current is radial or axial, and driven by a static electric field in the same direction (see Chapters 5, 8, and 9).

1.3.2   Electron Dynamics

The voltages employed in vacuum tubes are high enough, in many cases, for the velocities of the electrons to be at least mildly relativistic. The kinetic energy of a relativistic electron is [20]:

$$T = mc^2 - m_0 c^2, \tag{1.1}$$

where $m_0 c^2$ is the rest energy of the electron and the relativistic mass is

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}} = \gamma m_0, \tag{1.2}$$

where u is the velocity of the electron and c the velocity of light. If an electron starts from rest at the cathode then its velocity at a point where the potential is V, relative to the cathode,[4] is found by using the principle of conservation of energy:

$$eV = m_0 c^2 \left(\frac{1}{\sqrt{1 - u^2/c^2}} - 1 \right). \tag{1.3}$$

This equation can be rearranged as

$$u = c \left[1 - \frac{1}{\left[1 + (V/V_R) \right]^2} \right]^{\frac{1}{2}}, \tag{1.4}$$

---

[4] In this book voltages are referred to the cathode unless otherwise stated. In practice the tube body is normally at earth potential and the cathode potential is negative with respect to it.

where $V_R = m_0 c^2 / e = 511\,\text{kV}$ is the rest energy of the electron expressed in electron volts. When $V/V_R \ll 1$ the fraction can be expanded by the Binomial Theorem to give the approximate expression:

$$u = c \sqrt{\frac{2V}{V_R}}. \tag{1.5}$$

Figure 1.3 shows a comparison between the velocities calculated using the exact and approximate equations. The error in the approximate velocity is 1% at 7 kV so that the exact formula should be used for voltages higher than this.

The force acting on an electron is equal to the rate of change of momentum

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{u}) = m\frac{d\mathbf{u}}{dt} + \mathbf{u}\frac{dm}{dt}. \tag{1.6}$$

If the force acts in the direction of the motion of the electrons then

$$F = \frac{m_0}{\left(1 - u^2/c^2\right)^{3/2}} \cdot a = \gamma^3 m_0 a, \tag{1.7}$$

where a is the acceleration and $\gamma^3 m_0$ is sometimes called the *longitudinal mass*. When the force acts at right angles to the direction of motion then

$$F = \frac{m_0}{\left(1 - u^2/c^2\right)^{1/2}} \cdot a = \gamma m_0 a, \tag{1.8}$$

where γm_0 is the *transverse mass*. If the longitudinal velocity is approximately constant and much greater than the transverse velocity then it is possible to use the classical equations of motion with relativistic corrections to the mass of the electron.

The DC electron current in a tube may be collimated by a magnetic field which is either parallel to the direction of the DC current or perpendicular to it (see Chapters 7 and 8). Tubes in which the field is parallel to the current are described as 'Type O' and those in which it is perpendicular as 'Type M'. The current may also be collimated by a static electric field but this is rather rare.
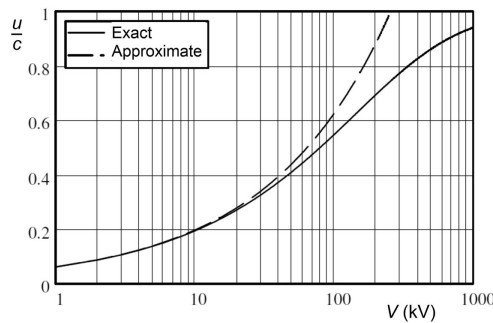
Figure 1.3: Comparison between exact and approximate electron velocities as a function of the accelerating potential.

1.3.3   Modulation of the Electron Current

For the present we will assume that the input RF voltage is purely sinusoidal with amplitude V_1 and frequency ω. Practical input signals are discussed in Section 1.6. It can be shown that the three possible methods of modulation of the current by the RF input voltage are [21]:

- Emission density modulation in which the current emitted from the source is varied.
- Deflection and concentration modulation in which the electrons are deflected sideways.
- Transit-time modulation in which the electron velocity is varied.

The effect of any of these methods of modulation, or combinations of them, is to produce a bunched stream of electrons whose current varies with time at the frequency of the input signal. Because the process is non-linear the time variation of the current is not sinusoidal but can be represented by a Fourier series. The amplitudes and phases of the harmonic components depend upon the amplitude of the input voltage, and on the distance from the source. The modulation may be either in the same direction as the DC current flow, or normal to it. Thus the current at some point z can be written

$$I(z,t) = I_0 + \sum_{n=1}^{\infty} I_n(V_{in},z)\exp(jn\omega t), \qquad (1.9)$$

where I_0 is the current in the unmodulated stream, $I_n(V_{in},z)$ are complex amplitudes, and modulation in the axial direction has been assumed. The DC and time-varying parts of the current may be in different directions in space. We note that the real current cannot be negative. Its maximum value, relative to I_0, is determined by the process of modulation, or by the maximum current which can be drawn from the source. The ratio I_0/I_0 cannot exceed 2.0 (see sections 11.8.4 and 13.3.4).

1.3.4   Amplification, Gain, and Linearity

The RF output voltage is obtained by passing the modulated stream through a region at an effective position z_2 where energy is removed from the electron bunches by an RF electric field. This can be represented by an impedance $Z_n(V_{in})$ which depends on frequency and on the magnitude of the input signal. Thus the output voltage is

$$V_{out}(t) = \mathrm{Re}\left\{\sum_{n=1}^{\infty} I_n(V_{in},z_2)Z_n(V_{in})\exp(jn\omega t)\right\}, \qquad (1.10)$$

where, for simplicity we assume that the characteristic impedances of the input and output waveguides are the same. The impedance Z_n is effectively zero above some value of the harmonic number n determined by the nature of the output section. Thus the number of harmonics at which there is appreciable output power is small and, in some cases, limited to the fundamental. If the output power is to be radiated it may be necessary to filter out the harmonics to comply with the regulations

determining the bandwidth available to the system, as specified by international agreement and national regulations [22, 23]. The transfer characteristic of the amplifier at the frequency ω is then

$$V_{out}(\omega t) = A(V_{in})\cos(\omega t + \Phi(V_{in})),\qquad(1.11)$$

where $A(V_{in})$ is the AM/AM (amplitude modulation) characteristic which is usually plotted on decibel scales and $\Phi(V_{in})$ is the AM/PM (phase modulation) characteristic of the amplifier.

The details of the transfer characteristics depend upon the type of amplifier but many of them show the same features. Figure 1.4 shows, as an example, a typical AM/AM curve of a travelling-wave tube (TWT). The output power is proportional to the input power at low drive levels (slope 1 dB/ dB) but reaches saturation as the input is increased. In the figure the input and output powers have been normalised to their values at saturation. The gain of the amplifier in decibels is given by

$$G_{dB} = 10\log\left|\frac{P_{out}}{P_{in}}\right|,\qquad(1.12)$$

where P_{in} and P_{out} are, respectively, the input and output power. Since we have assumed that the input and output waveguides have the same characteristic impedances we can write

$$G_{dB} = 20\log\left(\frac{A(V_{in})}{V_{in}}\right).\qquad(1.13)$$

The difference between the linear (small-signal) and the saturated gain is the *gain compression* which may be used as a measure of non-linearity. It is common for a tube to be described in terms of its saturated output power, but the output power available under normal operating conditions may be less than this. For example the tubes used in particle accelerators are normally operated 'backed-off' from saturation to provide a control margin for the operation of the accelerator. In a TWT there is normally some output at second, and higher, harmonic frequencies. This also depends on the input drive level as shown in Figure 1.4. At low drive levels the second harmonic output is proportional to the square of the input power, giving a slope of 2 dB/dB. The maximum second harmonic power does not necessarily occur at the drive level that saturates the fundamental. The intersection between the projections of the linear parts of the fundamental and second harmonic curves, known as the *second-order intercept point*, is a measure of the second-order distortion of the amplifier.

The corresponding AM/PM curve for a TWT is shown in Figure 1.5. The phase is plotted relative to the phase at low drive levels, and the input power is normalised to the input power at saturation. The phase of the output signal relative to the input is constant at low drive levels but changes as the drive level is increased. Both the amplitude and phase of the output depend on frequency, and on the operating
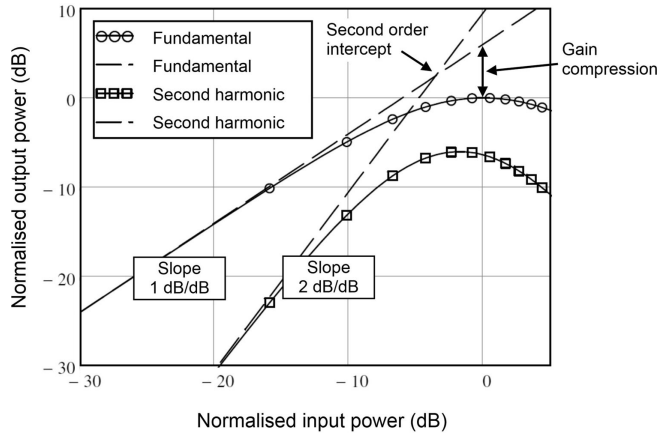
Figure 1.4: Typical curves of fundamental and second harmonic output power of an RF amplifier plotted against the input power, normalised to saturation.
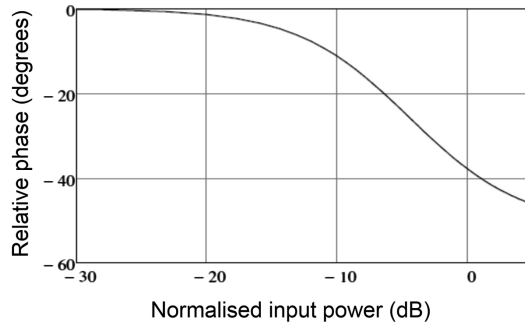
Figure 1.5: Typical curve of the phase of the output voltage of an RF amplifier against input power normalised to saturation.

conditions of the amplifier, including the voltages applied and the external RF matches. Thus any ripple in the voltages applied may result in amplitude and phase modulation of the output signal. In an oscillator, voltage ripple may also produce frequency modulation.

1.3.5   Power Output and Efficiency

An important consideration for many applications of tubes is the efficiency with which the DC input power is converted into useful RF power. The principle of conservation of energy requires that, in the steady state, the total input and output powers must balance, that is

$$P_{RF\ in} + P_{DC\ in} = P_{RF\ out} + P_{Heat}. \tag{1.14}$$

The DC input power includes the power in the electron stream, the cathode heater, and electromagnets. If the tube has a depressed collector (see Section 10.3) the

DC power is reduced by the power recovered by the collector. When a comparison is made between alternative tube types, or between vacuum tube and solid-state amplifiers, then the DC input power should be specified at the input to the power supply to allow for losses in it. Any power required for cooling fans or pumps must also be included. The RF output power comprises power at the fundamental frequency and its harmonics. Heat is generated by the impact of electrons on the collector and the tube body and by RF losses in the tube body, connecting waveguides and windows.

A number of different definitions of efficiency are in use and it is important to distinguish between them. The *overall efficiency* is defined here as the ratio of the fundamental RF output power (P_2) to the total input power to the tube

$$\eta_O = \frac{P_2}{P_{DC} + P_{RF\,in}}. \tag{1.15}$$

If the gain of the amplifier is high, the RF input power is much smaller than the DC input power so that approximately

$$\eta_O \approx \frac{P_2}{P_{DC}}. \tag{1.16}$$

The heater and electromagnet powers are typically much smaller than the stream power in continuous wave (CW) tubes but they can be comparable with the stream power in pulsed tubes. The *power added efficiency*, sometimes used when the gain is small, is

$$\eta_A = \frac{P_2 - P_{RF\,in}}{P_{DC}}. \tag{1.17}$$

This efficiency is effectively identical to that in (1.16) if the gain is 20 dB or more. The efficiency of tubes with depressed collectors is discussed in Section 10.3. The *RF efficiency* is defined here as the ratio of the useful RF output power to the power input to the electron stream (less any recovered by the collector) plus the RF input power.

$$\eta_{rf} = \frac{P_2}{P_{stream} - P_{recovered} + P_{RF\,in}}. \tag{1.18}$$

The *electronic efficiency* is the efficiency with which power is transferred from the electron stream to the RF electric field of the output circuit.

$$\eta_e = \frac{P_{RF\,out} + P_{loss}}{P_{stream}}, \tag{1.19}$$

where P_{loss} is the total RF power loss in the output circuit. If the output circuit is resonant so that only power at the fundamental frequency is included

$$\eta_e = \frac{P_2 + P_{2,loss}}{P_{stream}}, \tag{1.20}$$

where $P_{2,loss}$ is the RF loss at the fundamental frequency. Then the *circuit efficiency* can be defined by

$$\eta_c = \frac{P_2}{P_2 + P_{2,loss}}.$$ (1.21)

If, in addition, the RF input power is small enough to be neglected, and the tube does not have a depressed collector then the RF efficiency is

$$\eta_{rf} = \eta_e \eta_c.$$ (1.22)

1.3.6 Bandwidth

The AM/AM, and AM/PM, curves depend upon the frequency of the input signal. Then the transfer function of the amplifier may be written $H(\omega, V_{in})$. Figure 1.6 shows a typical graph of the RF output power of an amplifier as a function of frequency. This graph may be plotted at constant input power, or showing the saturated output power at each frequency. The bandwidth is defined in terms of the frequencies at which the power falls below the maximum by a specified amount (e.g. 1 dB bandwidth). This may be expressed in absolute terms or as a percentage of the centre frequency. The figure also shows that the output power may vary within the band and this can affect the system in which the tube is employed. These ripples are normally greatest when the tube is not saturated and are reduced at saturation by the effects of gain compression. The custom of using a decibel scale for this graph can be misleading. It is important to remember that a change of -1 dB represents a reduction in power by 20% and a change of -3 dB a reduction of 50%. Similar graphs of efficiency, small-signal gain, and saturated gain against frequency can also be plotted. The bandwidth specified for a tube may be to permit rapid changes in frequency (as in frequency-agile radar), or to enable it to be used with multiple and modulated signals (see Section 1.6).

In addition to power at harmonics of the input frequency the output of a tube may also include non-harmonic power under some operating conditions (for example during the rise and fall of the pulse during pulsed operation). These out-of-band emissions are undesirable, and limits are usually specified to avoid interference with other systems. For example it has been known for out-of-band emissions from marine radar, and from industrial microwave ovens, to interfere with microwave communication links. The tube also contributes to the noise in the system as discussed in Section 1.6.1.

1.3.7 The Electromagnetic Structure

The electromagnetic structures used to modulate the electron stream, and to extract power from it, can be divided into three groups:

- Fast-wave structures in which the phase velocity of travelling waves is greater than the velocity of light (Chapter 2).
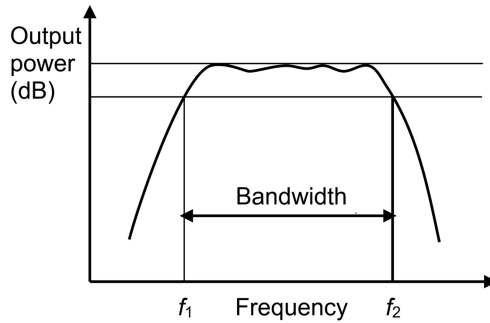
Figure 1.6: Typical graph of the output power of an RF amplifier plotted against frequency.

- Resonant (standing wave) structures in which the RF electric field profile is time-varying but fixed in space (Chapter 3).
- Slow-wave structures which employ the field of a travelling electromagnetic wave whose phase velocity is less than the velocity of light (Chapter 4).

The bandwidth of a tube using a resonant structure is typically restricted to a few percent by the loaded Q of the resonance (see Section 3.2.2). If the higher-order modes of the structure do not coincide with harmonics of the input frequency then the harmonic output of the tube is small.

The bandwidths of tubes using travelling-wave structures can be broad because the phase velocity and the impedance presented to the electron current by the structure vary slowly with frequency. The interaction is distributed so that the bunching section and the output section of the tube are effectively combined. For useful interaction to take place the mean electron velocity must be approximately synchronous with the phase velocity of the wave. There may be interaction at harmonics of the input frequency leading to appreciable output power at the second and higher harmonics as shown in Figure 1.4.

1.3.8   Coupled-Mode Theory

Useful insights into the properties of travelling-wave structures and their interactions with electron streams under small-signal (linear) conditions are given by coupled-mode theory [24–28]. For this purpose the small-signal modulation of the electron stream is described by a pair of normal modes whose impedances have opposite signs. The properties of space-charge waves on electron beams are discussed in detail in Chapter 11.

The properties of a wave propagating as $\exp j(\omega t - \beta z)$ are the solutions to the dispersion equation

$$D(\omega, \beta) = 0, \tag{1.23}$$

where the frequency (ω) and the propagation constant (β) are real if there is no attenuation.[5] The solutions of the dispersion equation can be plotted on a dispersion (ω, β) diagram. In this diagram the phase velocity of the wave is

$$v_p = \frac{\omega}{\beta} \tag{1.24}$$

and the group velocity, which is taken to be the velocity of propagation of energy, is

$$v_g = \frac{d\omega}{d\beta}. \tag{1.25}$$

The phase velocity is positive in the first quadrant of the (ω, β) diagram and negative in the second quadrant. The group velocity can be either positive or negative depending upon the properties of the mode. We shall see in Chapter 4 that waves having positive phase velocities and negative group velocities occur in periodic structures.

Where two waves propagate in the same region of space they may interact to produce coupled modes if their dispersion curves intersect as shown in Figure 1.7. The dispersion equation for the coupled modes can then be written in the form

$$D_1(\omega, \beta) D_2(\omega, \beta) = \kappa(\omega, \beta), \tag{1.26}$$

where $\kappa(\omega, \beta)$ represents the coupling between the modes. The coupled modes are the solutions to this equation. In order to investigate the properties of coupling under different conditions we will redefine ω and β so that the two uncoupled modes intersect at the origin. This transformation does not change the group velocities of the modes.

The properties of the different possible kinds of mode coupling can be illustrated by assuming that the uncoupled dispersion curves are straight lines so that

$$D(\omega, \beta) = (\omega - \beta v_g) = 0. \tag{1.27}$$

The dispersion equation for the coupled modes is then

$$(\omega - \beta v_{g1})(\omega - \beta v_{g2}) = \kappa_0, \tag{1.28}$$

where κ_0 is regarded as a constant close to the origin. This equation can be expanded to give

$$\omega^2 - (v_{g1} + v_{g2})\omega\beta + v_{g1}v_{g2}\beta^2 - \kappa_0 = 0, \tag{1.29}$$

which is quadratic in both ω and β. Solving for ω gives

---

[5] The symbol β used by engineers is employed in this book in place of the symbol k normally used by physicists.
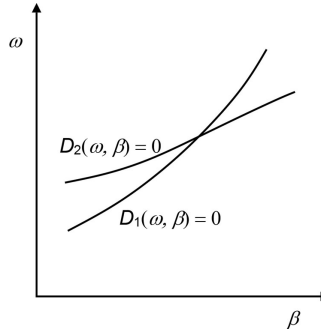
Figure 1.7: Dispersion diagram for two modes of propagation.

$$\omega = \frac{1}{2}\left\{ (v_{g1}+v_{g2})\beta \pm \sqrt{(v_{g1}-v_{g2})^2 \beta^2 + 4\kappa_0} \right\} \tag{1.30}$$

and for β gives

$$\beta = \frac{1}{2v_{g1}v_{g2}}\left\{ (v_{g1}+v_{g2})\omega \pm \sqrt{(v_{g1}-v_{g2})^2 \omega^2 + 4v_{g1}v_{g2}\kappa_0} \right\}. \tag{1.31}$$

The constant κ_0 is positive if the signs of the impedances of the uncoupled waves are the same, and negative if they are opposite. Likewise the signs of the group velocities of the waves may be the same, or opposite. Hence all possible combinations are covered by four cases, which can be explored using Worksheet 1.1. It can be shown that the properties of the coupled modes can be revealed by considering the solutions of (1.30) when β is real and the solutions of (1.31) when ω is real as follows [26, 29]:

Case A: The group velocities and the impedances have the same sign $(v_{g1}v_{g2} > 0, \ \kappa_0 > 0)$.
When β is real the solutions of (1.30) are always two real values of ω and, when ω is real, the solutions of (1.31) are two real values of β. These solutions represent travelling waves with the modified dispersion diagram shown in Figure 1.8. In Figures 1.8 to 1.11 the uncoupled modes are shown by dotted lines, the real parts of the solutions by solid lines and the imaginary parts by dashed lines.

Case B: The group velocities have opposite signs and the impedances have the same sign $(v_{g1}v_{g2} < 0, \ \kappa_0 > 0)$.
When β is real the solutions of (1.30) are always two real values of ω but, when ω is real and close to the origin, the solutions of (1.31) are complex conjugate pairs of values of β, as shown in Figure 1.9(b). Physically, the complex values of β represent travelling waves whose amplitudes decay exponentially with distance. For example see Figure 4.28(a) where the forward and backward waves in a waveguide are coupled to one another by periodic discontinuities to produce frequency stop-bands in which there are decaying (evanescent) waves.
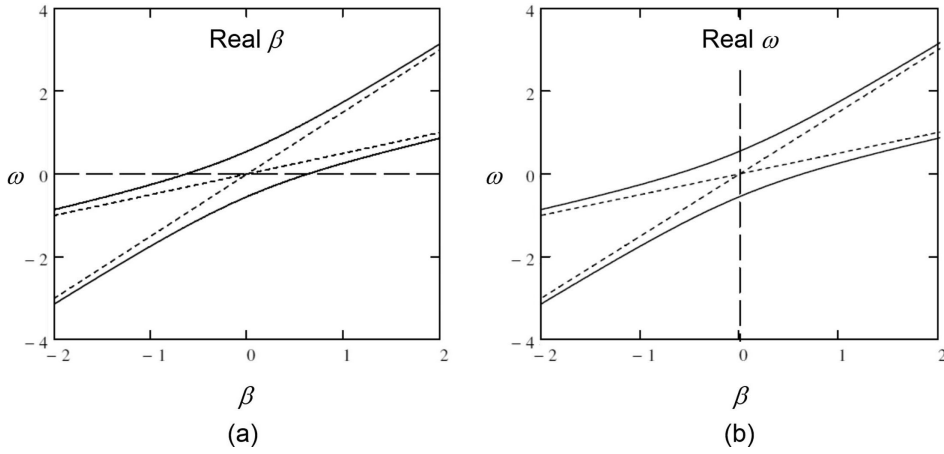
Figure 1.8: Coupling between two forward waves with impedances having the same sign: (a) ω for real β, and (b) β for real ω.
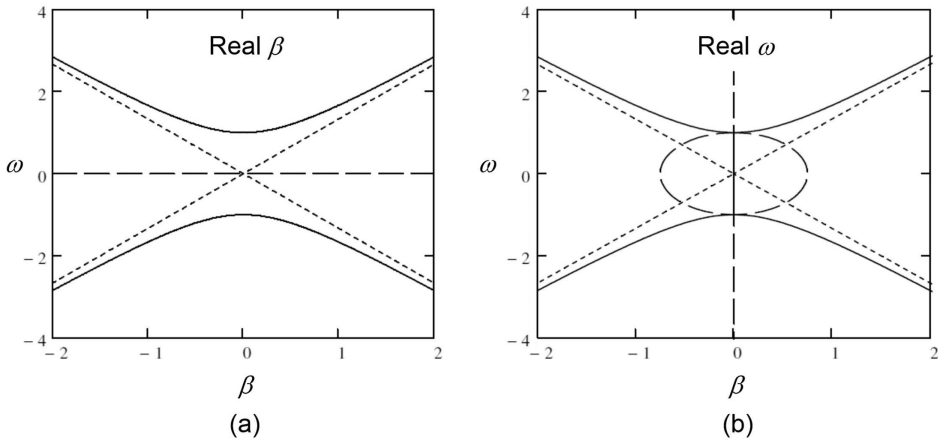
Figure 1.9: Coupling between a forward wave and a backward wave with impedances having the same sign: (a) ω for real β, and (b) β for real ω.

Case C: The group velocities have the same sign and the impedances have opposite signs $\left(v_{g1}v_{g2} > 0, \kappa_0 < 0\right)$.

The solutions for both ω and β, close to the origin, are complex conjugate pairs, as shown in Figure 1.10. It can be shown that the complex solution for β corresponds to a *convective instability* in which the waves grow and decay exponentially in space (see Figure 1.10(b)) [26, 29]. An example of this type of coupling is the growth of waves in a travelling-wave tube as shown in Figure 11.19.

Case D: The group velocities and the impedances have opposite signs $\left(v_{g1}v_{g2} < 0, \kappa_0 < 0\right)$.
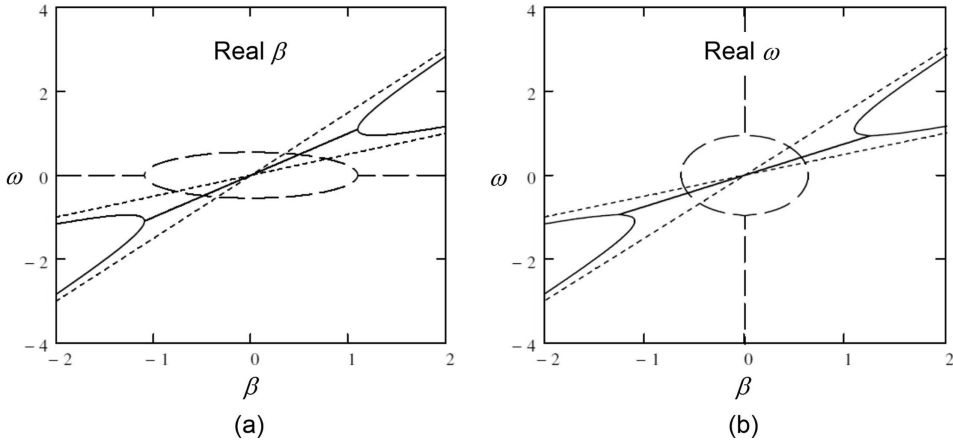
Figure 1.10: Coupling between two forward waves with impedances having opposite signs: (a) ω for real β, and (b) β for real ω.

When β is real and close to the origin the solutions of (1.30) are a complex conjugate pair of values of ω and, when ω is real, the solutions of (1.31) are always two real values of β, as shown in Figure 1.11. The solution with complex ω is a *non-convective,* or *absolute, instability* in which the two solutions grow and decay exponentially in time. The condition for this instability to exist is that there is at least one value $\omega = \omega_s$ in the upper half of the complex plane for which there are coincident roots in β_s [29]. For the case considered here the condition for the existence of coincident roots is found from (1.31)

$$\omega_s = \pm 2j \frac{\sqrt{\left|v_{g1}v_{g2}\right|\left|\kappa_0\right|}}{\left|v_{g1}\right| + \left|v_{g2}\right|}. \tag{1.32}$$

The condition is satisfied in this case when $\kappa_0 = 0$ and $\omega = 0$ so that an absolute instability exists for all negative values of κ_0. The corresponding values of β can be found from (1.31):

$$\beta_s = \frac{\left|v_{g1}\right| - \left|v_{g2}\right|}{2\left|v_{g1}v_{g2}\right|} \omega_s. \tag{1.33}$$

Thus β_s is complex when ω_s is complex. In general the start-oscillation condition can be found by finding the least magnitude of the coupling for which coincident roots of β exist when ω is real (see Figures 17.5 and 17.6) [30]. If the coupling is increased beyond this point both the frequency and the propagation constant are complex (see Section 11.7). Examples of this kind of instability are backward-wave oscillations in a travelling-wave tube (Section 11.7) and gyrotron oscillators (Section 17.3).
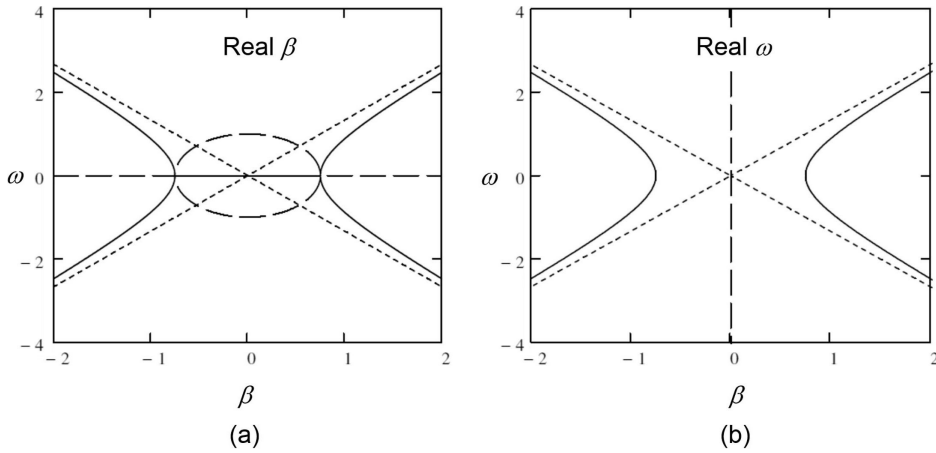
Figure 1.11: Coupling between forward and backward waves with impedances having opposite signs: (a) ω for real β, and (b) β for real ω.

The properties of coupled modes remain the same in more general cases where the uncoupled dispersion diagrams are not straight lines, κ is a function of ω and β, and more than two modes are involved in the coupling. Thus a qualitative description of the properties of the coupled system can be derived from an examination of the dispersion diagram of the uncoupled modes and the form of the dispersion equation for the coupled system [29].

1.3.9   Classification of Vacuum Tubes

The preceding sections have reviewed the different ways in which the functional blocks in Figure 1.2 can be realised, and that can be used as a basis for classifying tubes. The discussion in this book is restricted to the types of tube which are of continuing practical importance. A summary of their principal features is given in Table 1.1. A few variations on the main types of tube exist, for example: triodes and tetrodes with DC and RF current flow in the axial direction, and hybrid tubes in which the bunching section resembles a klystron and the output section a travelling-wave tube. It is possible to think of the inductive output tube as a hybrid tube with a triode bunching section and a klystron output section. Many other possible types of tube can be envisaged [31–33].

1.4   Applications of Vacuum Tubes

High-power vacuum tubes have found applications in many fields as summarised in Table 1.2 [17]. A comprehensive summary of the historical development of tubes for the microwave and millimetre-wave region of the spectrum can be found in [34, 35]. For the history of power gridded tubes and their applications see [36].

Table 1.1: Classification of the principal types of vacuum tube

Type (Chapter)	Class	DC current	RF current	Magnetic field	Electro-magnetic structure	Bunching
Triode and tetrode (12)	Gridded	Radial	Radial	None	Resonant	Radial
Inductive output tube (12)	Gridded	Axial	Axial	Axial	Resonant	Axial
Klystron (13)	Linear-beam (O)	Axial	Axial	Axial	Resonant	Axial
Travelling-wave tube (14)	Linear-beam (O)	Axial	Axial	Axial	Slow-wave	Axial
Magnetron (15)	Crossed-field (M)	Radial	Tangential	Axial	Resonant	Tangential
Crossed-field amplifier (16)	Crossed-field (M)	Radial	Tangential	Axial	Slow-wave	Tangential
Gyrotron (17)	Fast-wave	Axial	Tangential	Axial	Fast-wave	Tangential

In general, the development of high power vacuum tubes has been driven by the requirements of particular applications. The number of tubes of each type needed has not been great enough to allow the use of high volume manufacturing methods. The exception to this rule is the domestic microwave oven magnetron where very large numbers have been made. The cost per tube is then dramatically reduced when compared with magnetrons of other types. In some cases tubes, which have been developed for one purpose, have found applications elsewhere. For example high power tetrodes for radio transmitters and klystrons for television and radar transmitters have been used in particle accelerators. However, if the primary market declines, the manufacturers may not continue to produce tubes for low-volume secondary applications. This was a major motivation for the development of high power solid state amplifiers for the SOLEIL synchrotron [11]. Thus the nature of the market for high power vacuum tubes dictates the need for close communication between the engineers developing a tube and those working on the system in which it will be used. This pattern, which is typical of other markets for low-volume, high-technology, products is quite different from that for high-volume products. The successful incorporation of complex components, such as tubes, into a system requires the systems engineers to have a detailed understanding of how they work. It is not sufficient to regard tubes as 'black boxes' whose performance can be described completely by characteristics defined at their terminals. Tube engineers must, likewise, have a detailed understanding of the issues which are important for the systems in which their products are used. Some of these are reviewed in this chapter and in Chapter 20. For further information see [37].

Table 1.2: The principal applications of vacuum tubes

Field of application	Principal applications	Principal tube types
Broadcasting	Radio	Triodes Tetrodes
	Television	Inductive output tubes Klystrons
	Direct broadcasting from satellites	Helix travelling-wave tubes
Telecommunications	Point to point links	Helix travelling-wave tubes
	Satellite communications	Helix travelling-wave tubes Klystrons
	Deep space communications	Coupled-cavity travelling-wave tubes
Civilian radar	Marine radar	Magnetrons
	Airborne radar	Magnetrons
	Weather radar	Magnetrons
	Air traffic control	Magnetrons Coupled-cavity travelling-wave tubes
Military	Early warning radar	Klystrons
	Target identification radar	Magnetrons Klystrons Travelling-wave tubes Crossed-field amplifiers
	Electronic countermeasures	Helix travelling-wave tubes
Industrial	Industrial heating	Triodes Magnetrons
	Domestic ovens	Magnetrons
Scientific	Scientific particle accelerators	Tetrodes, Inductive output tubes Klystrons
	Medical accelerators	Klystrons Magnetrons
	Thermonuclear fusion reactors	Tetrodes, Klystrons Gyrotrons

1.5   The Statement of Requirements

A fundamental document for communication between the customers and the contractors for any engineering product is the Statement of Requirements (also known as a Requirements Specification or a Statement of Work). It is the source from which two further documents are derived: the Manufacturing Specification (or Manufacturing Data Package), which provides all the information required to manufacture the product; and the Test Specification, which gives details of

the tests that the product must undergo to demonstrate its compliance with the requirements. It is self-evident that any omissions or ambiguities in these documents will potentially lead to failure to achieve the outcome desired. Thus it is essential that the Statement of Requirements and the Test Specification are produced by an open dialogue between representatives of the customer and the contractor, both of whom fully understand the issues involved. For this purpose it is helpful to have a checklist of the issues to be discussed [37, 38]. These issues fall into two main groups: the performance requirements; and the design constraints. Typical top-level requirements are shown in Tables 1.3 and 1.4. The following sections discuss some of the issues which are important in the specification of performance requirements.

1.6   Signals and Noise

In order to transmit information for telecommunications (one to one) or broadcasting (one to many) it is necessary for the properties of the sinusoidal carrier wave to be modulated by varying them with time. The properties of the input signal which may be varied are amplitude, phase, and frequency. Many different methods of modulation are in use and only a brief summary is given here of the factors that are important in the specification of vacuum tube amplifiers [39–41]. The methods fall into two classes: analogue modulation in which the properties of the carrier vary continuously with time; and digital modulation in which they are switched between a number of states.

The baseband input signal is typically defined as a time-varying voltage. This could represent a continuously varying, analogue, source such as a speech or music waveform. Alternatively it could represent a digital source such as a computer data file, or the digitised form of an analogue waveform. It is common, for convenience of analysis, to assume that the input signal is random when observed over an extended period of time, though this is frequently not the case. The frequency spectrum of the signal is obtained by taking the Fourier transform of the waveform. Theoretically this spectrum extends to infinity but, in practice, it is assumed to be restricted to a bandwidth that extends from DC up to some maximum frequency. The upper limit may be somewhat arbitrary and be defined as the point at which the power density in the spectrum falls below a certain limit. This still applies if the bandwidth of the signal has been limited by a filter, since no practical filter has an infinitely sharp cut-off.

The basic digital signal is binary in which the voltage switches between two states at a fixed clock rate. However, it is often useful to convert this into a form in which greater number of states is employed. This can be illustrated by considering Table 1.5, which shows different possible representations of the symbols in a sixteen character set. Let us suppose that we wish to transmit a message made up of the 16 characters in the first column at a rate of 1 symbol/sec. It is helpful to think of these characters

Table 1.3: Typical performance requirements for vacuum tube amplifiers

Frequency
Bandwidth (instantaneous, tuneable)
Power output (peak, mean)
Gain
Linearity (amplitude, phase)
Noise and spurious emissions
Efficiency
RF matches

Table 1.4: Typical design constraints for vacuum tube amplifiers

High-voltage power supply
    (voltage, current, pulse shape)
Auxiliary power supplies
Size and weight
Cooling
Temperature range (storage, operation)
Shock and vibration
Atmosphere (pressure, humidity)
Reliability
Repairability
Lifetime cost

as representing 16 different voltage levels in a stepwise approximation to an ana-logue waveform. A little thought shows that the highest frequency is obtained when any pair of symbols is repeated alternately (e.g. ABABABA). The Fourier trans-form of all other sequences has a lower maximum frequency.

    If we now wish to transmit the same information in binary form we must send four binary digits (bits) per second. The highest frequency is obtained when the symbols 0 and 1 are repeated alternately (e.g. FFFF, or KKKK). Finally we can represent the 16 symbols by a four-symbol set by grouping the binary digits in pairs as shown in the third column. Now the highest frequency is obtained when any pair of symbols is repeated alternately (e.g. WXWX = BB). From this it can be seen that the bandwidth required to transmit the same message depends upon the number of symbols in the set used to represent it. Alternatively, we see that the maximum data rate for a fixed bandwidth is increased by a factor of two if a four-symbol set is used, and by a factor of four with a 16-character set. Now an ideal low-pass chan-nel of bandwidth B can transmit $2B$ pulses per second. Thus the maximum rate at which data can be transmitted over this channel is given by Hartley's law

$$R = 2B \log_2 m \quad \text{bits s}^{-1}, \tag{1.34}$$

where m is the number of symbols used.

1.6.1   Noise

Any communications system adds white noise to the input signal so that the signal received is corrupted to some extent [39]. The thermal noise power in watts caused by the random motion of charge carriers in resistive materials that is delivered to a matched load is

$$N_p = kT\,B, \tag{1.35}$$

where k is Boltzmann's constant $\left(k = 1.380 \times 10^{-23}\ \text{J}\,\text{K}^{-1}\right)$, T is the absolute temperature and B is the bandwidth of the system in Hz. The effectiveness of an analogue communication system is measured by the signal to noise ratio $\left(S_p/N_p\right)$ where S_p is the average signal power. This is commonly expressed in decibels as

$$SNR = 10\log_{10}\left(\frac{S_p}{N_p}\right). \tag{1.36}$$

The minimum acceptable SNR for reliable communication is generally taken to be about 10 dB.

For digital communications the corruption of the signal may result in some of the bits being received incorrectly. Shannon's theorem states that a noisy channel will theoretically support error-free data transmission at a channel capacity given by

$$C = B\log_2\left(1 + \frac{S_p}{N_p}\right). \tag{1.37}$$

Table 1.5: Representation of digital signals

16 symbols	2 symbols (binary)	4 symbols
A	0000	00 00 (WW)
B	0001	00 01 (WX)
C	0010	00 10 (WY)
D	0011	00 11 (WZ)
E	0100	01 00 (XW)
F	0101	01 01 (XX)
G	0110	01 10 (XY)
H	0111	01 11 (XZ)
I	1000	10 00 (YW)
J	1001	10 01 (YX)
K	1010	10 10 (YY)
L	1011	10 11 (YZ)
M	1100	11 00 (ZW)
N	1101	11 01 (ZX)
O	1110	11 10 (ZY)
P	1111	11 11 (ZZ)

This can be expressed alternatively as

$$C = B \log_2 \left(1 + \frac{E_b}{N_0} \cdot \frac{R}{B}\right), \tag{1.38}$$

where E_b is the energy per bit, R is the mean bit rate, and N_0 is the spectral energy density of the noise. Shannon's law represents an ideal which cannot be achieved in practice, but it remains useful for comparison with practical systems. In a digital transmission system the energy per bit, which is related to the bit error rate, is commonly used in place of the signal to noise ratio [40]. A second performance measure is the bandwidth efficiency (R/B) in bits s^{-1} Hz^{-1}.

   Each part of a communications channel contributes some noise to the system. Here we are concerned specifically with the contribution from the final power amplifier. The noise generated by an amplifier can be described by its noise figure, which is the ratio of the signal to noise ratio at the input to that at the output

$$F = \frac{SNR_i}{SNR_o}. \tag{1.39}$$

If an amplifier has a power gain A_p then the signal to noise ratio at the output is

$$SNR_o = \frac{A_p S_i}{A_p N_i + N_a}, \tag{1.40}$$

where S_i and N_i are the signal and noise powers at the input and N_a is the noise power added by the amplifier. Thus

$$F = \frac{A_p N_i + N_a}{A_p N_i}. \tag{1.41}$$

The noise figure is standardised by fixing the input noise power as that given by (1.35) when $T = 290$ K. The noise added by the amplifier can be described by an equivalent noise source at its input such that

$$N_a = k T_e B A_p, \tag{1.42}$$

where T_e is the effective noise temperature of the amplifier. The noise performance can also be described by the total noise power at the output, or by the noise power density (noise power per Hz).

1.6.2   Analogue Modulation

In order to transmit the baseband signal over a wireless channel it must be super-imposed upon an RF carrier wave by a modulator. The properties of the carrier wave which can be varied by the modulator are its amplitude, frequency, and phase, or some combination of them. In general the modulated signal at the input to the

power amplifier can be described by a sinusoidal carrier whose amplitude and phase are functions of time [42, 43]. Thus

$$V_1(t) = r(t)\cos(\omega_0 t + \psi(t)), \tag{1.43}$$

where ω_0 is the carrier frequency. The rates of change with time of the amplitude $(r(t))$ and phase $(\psi(t))$ are normally much slower than that of the carrier wave so that it is permissible to regard the instantaneous response of the amplifier as being that for a single unmodulated carrier wave. The response to more complex signals can therefore be deduced from the steady-state, single-carrier, transfer character-istics. Figure 1.12 shows examples of simple analogue modulation schemes. For purposes of illustration the modulation frequency has been taken to be one eighth of the carrier frequency although the ratio is normally much smaller than that.

The simplest example to understand is the amplitude modulation shown in Figure 1.12(a). The envelope of the carrier carries information such as an audio waveform. In the time domain the modulated waveform is given by

$$f(t) = [K + S(t)]\cos(\omega_0 t), \tag{1.44}$$

where K is the amplitude of the unmodulated carrier and $S(t)$ is the signal wave-form. When the signal is sinusoidal with frequency ω_m the Fourier transform of $f(t)$ yields the spectrum in the frequency domain shown in Figure 1.12(b). This representation shows the amplitudes, but not the phases, of the components of the signal. The central line at the carrier frequency is flanked by two sidebands at fre-quencies $(\omega_0 \pm \omega_m)$. This is known as double sideband amplitude modulation. For a more general modulating waveform the sidebands are the frequency-shifted Fourier transforms of the baseband waveform. Faithful transmission of the modulated waveform requires the bandwidth of the amplifier to be at least $2B$ (the Nyquist limit) where B is the baseband bandwidth. The instantaneous power in the modu-lated signal varies with time, as can be seen from Figure 1.12(a). Faithful transmis-sion of the signal requires the amplifier to be linear up to the highest instantaneous power. Any non-linearity of the amplifier produces a distortion of the signal in the time domain which is reflected in changes in its spectrum (see Section 1.6.4).

Figures 1.12(c) and (d) show double sideband suppressed carrier modulation. The modulated waveform is

$$f(t) = S(t)\cos(\omega_0 t). \tag{1.45}$$

The spectrum differs from that in Figure 1.12(b) in the absence of the carrier fre-quency. The bandwidth requirements for the amplifier are the same as in the previ-ous case and the dynamic range is greater. However, this modulation scheme is less susceptible to interference from noise than the previous case because all the power is in the sidebands.

Figures 1.12(e) and (f) show phase modulation. The amplitude envelope of the signal is constant in the time domain so that the instantaneous power is con-stant. This has the advantage that the working point of the amplifier is virtually
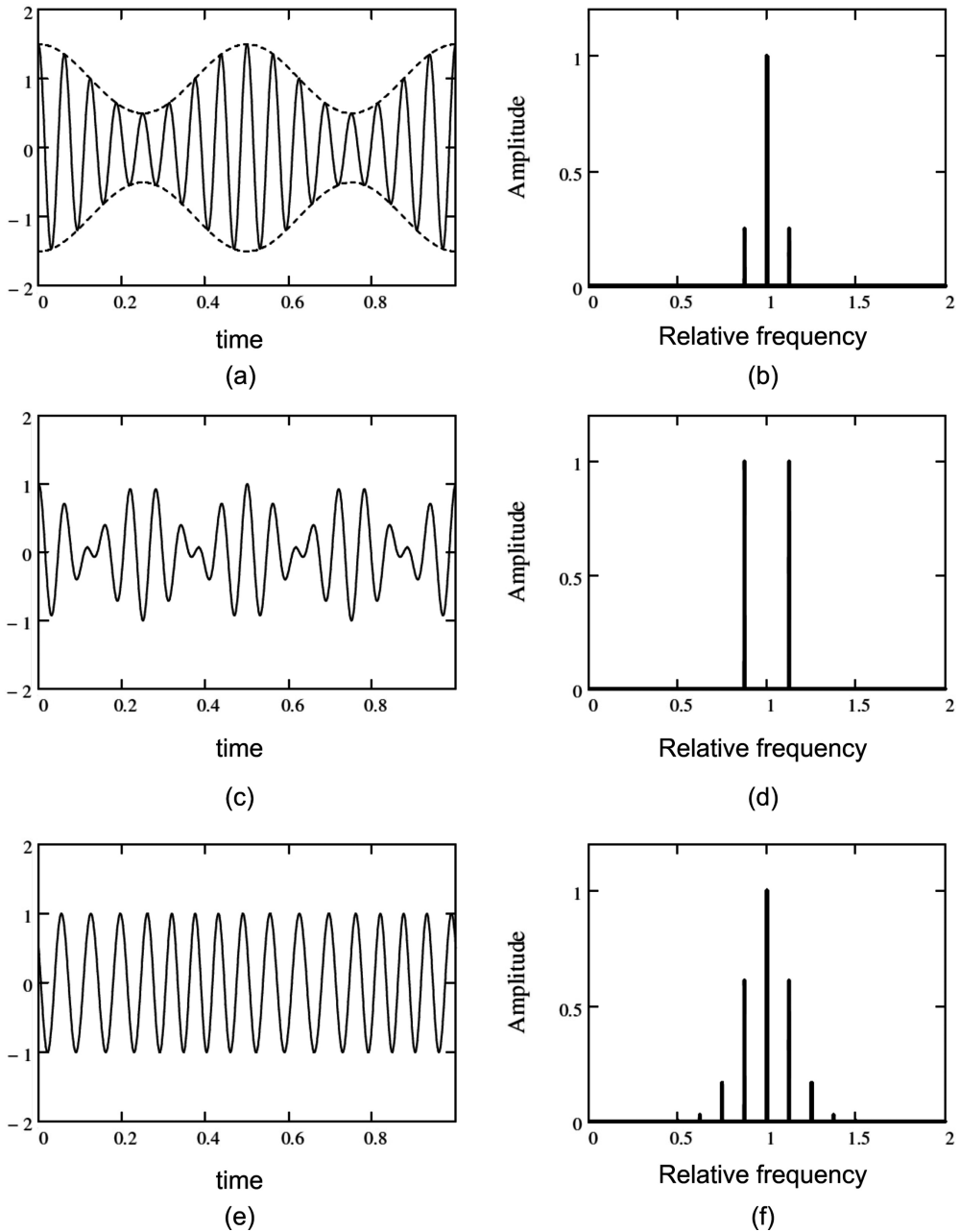
Figure 1.12: Time and frequency domain representations of analogue modulated carriers: (a) and (b) double sideband amplitude modulation; (c) and (d) double sideband suppressed carrier amplitude modulation; (e) and (f) phase modulation.

constant, so that the output signal is only affected by the frequency dependence of the gain and phase of the amplifier. The modulated signal, assuming unity amplitude, is

$$f(t) = \cos(\omega_0 t + S(t)),\qquad(1.46)$$

so that the dependence of the modulated signal on the modulation is non-linear. For sinusoidal modulation at frequency ω_m (1.46) can be written

$$f(t) = \cos(\omega_0 t + \beta \sin \omega_m t),\qquad(1.47)$$

where β is the modulation index. The spectrum for this waveform, shown in Figure 1.12(f) for $\beta = 1$, has additional lines at frequencies $(\omega_0 \pm n\omega_m)$ where $n = 1, 2, 3, \ldots$. Thus the bandwidth required for faithful amplification is greater than for amplitude modulation schemes. A rule of thumb for the transmission bandwidth is Carson's rule [40]

$$B_T = 2(\beta + 1)\omega_m.\qquad(1.48)$$

Frequency modulation can also be represented by (1.47) so that its properties are similar to those of phase modulation. The properties of analogue modulation schemes can be explored using Worksheet 1.2.

1.6.3   Digital Modulation

The simplest digital modulation schemes are amplitude shift keyed (ASK), phase shift keyed (PSK), and frequency shift keyed (FSK) modulation [39]. Figure 1.13 shows time and frequency domain representations of these methods of modulation, using the same carrier and modulation frequencies as before. It may be observed that the bandwidths of these signals are greater than those for the equivalent analogue modulation, with the exception of binary FSK modulation, which may be considered a special case of frequency modulation. The properties of binary digital modulation schemes can be explored using Worksheet 1.2.

We saw in Section 1.6.1 that the bit rate in a digital system having fixed bandwidth can be increased by using a greater number of symbols. Two common implementations of this principle are illustrated in Figure 1.14 which shows the positions of the symbols on phasor diagrams known as constellation plots. Figure 1.14(a) shows binary phase-shift keyed (BPSK) modulation with a phase shift of 180° corresponding to Figure 1.13(c) and (d). Quadrature phase-shift keyed (QPSK) modulation is shown in Figure 1.14(b), and quadrature amplitude modulation with 16 symbols (16-QAM) is shown in Figure 1.14(c). The effect of noise in the transmission channel is to produce a scatter of points around the desired constellation points at the receiver. If the scatter is too great then the receiver will fail to identify the symbols correctly leading to errors in transmission. Table 1.6 shows the bandwidth efficiencies and E_b/N_0 figures for these three channels at a bit error rate of 10^{-6}. These figures assume coherent modulation so that the clock signals at the transmitter and
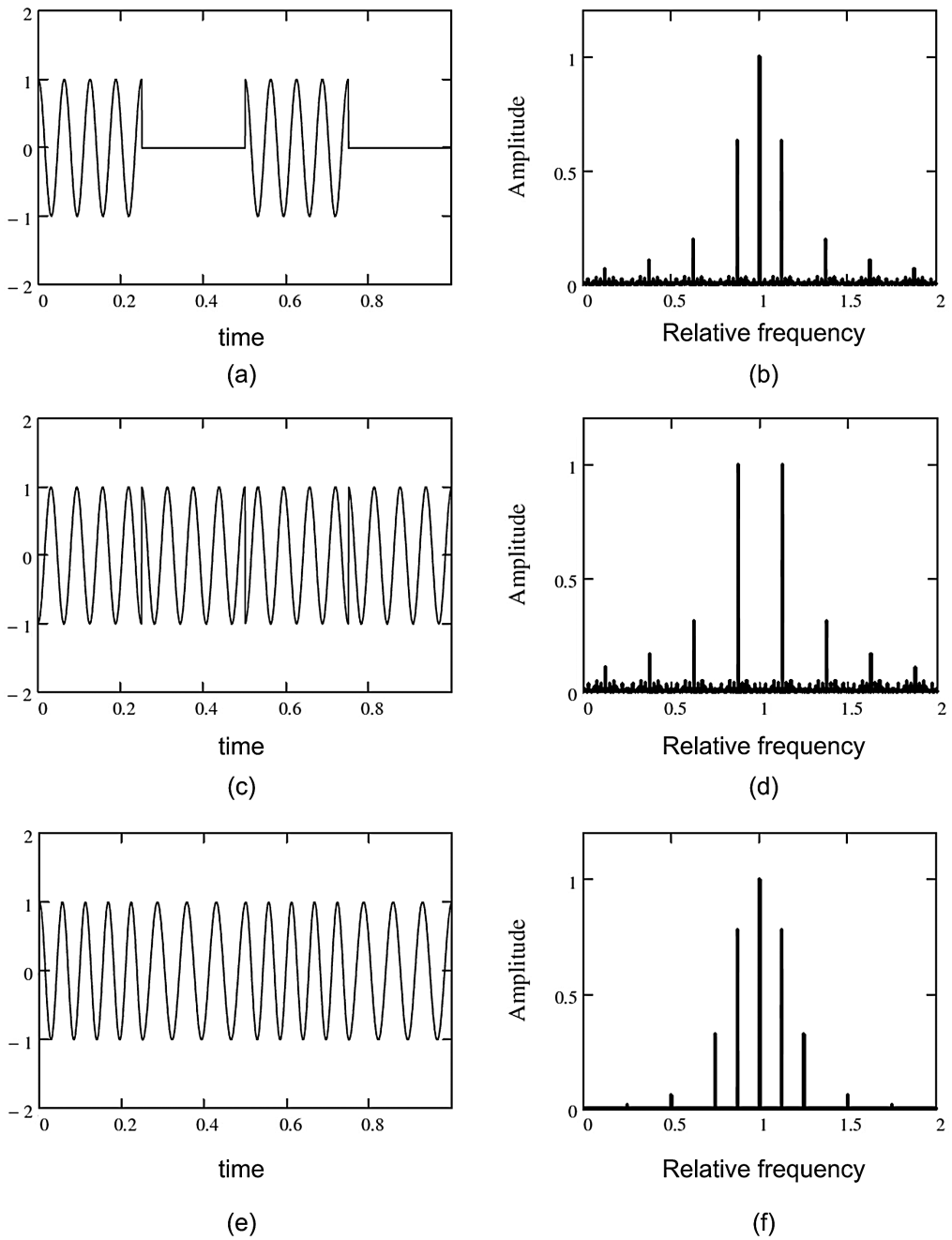
Figure 1.13: Time and frequency domain representations of simple digitally-modulated carriers: (a) and (b) binary ASK; (c) and (d) binary PSK; (e) and (f) binary FSK.

Table 1.6: Comparison between examples of coherent digital modulation schemes [40]

Modulation type	Bandwidth efficiency (bits s^{-1} Hz^{-1})	E_b/N_0
BPSK	0.5	10.6 dB
QPSK	1.0	10.6 dB
16-QAM	2.0	14.3 dB

(a) (b) (c)
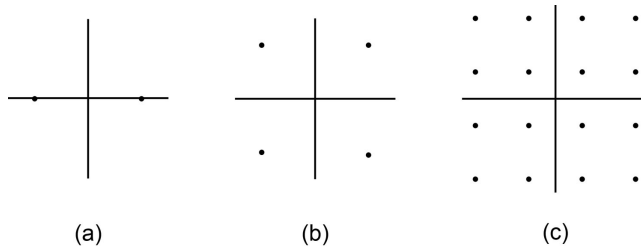
Figure 1.14: Examples of constellation diagrams for digitally-modulated carriers: (a) binary phase-shift keyed (BPSK), (b) quadrature phase-shift keyed (QPSK), and (c) quadrature amplitude modulation with 16 symbols (16-QAM).

the receiver are locked to one another. It can be seen that the bandwidth efficiency increases with the number of symbols according to (1.34) as expected.

From the point of view of the design of the power amplifier it is important to note that BPSK and QPSK are constant envelope schemes which allow the efficiency of the power amplifier to be high, subject only to the need to maintain adequate linearity. On the other hand the amplitude variation in 16-QAM means that the amplifier is only working at the highest power conversion efficiency at four of the constellation points. In addition, the overall transmitter power must be high enough to achieve an acceptable bit error rate for the innermost constellation points where the transmitted power is least. If all constellation points are equally probable then the average efficiency is about half of that at the outermost points. This makes 16-QAM unsuitable for applications such as satellite downlinks where high efficiency is important.

1.6.4   Multiplexing

It is commonly the case that one telecommunications channel carries data from multiple sources. The interleaving of the different data streams, known as multiplexing, can be achieved in many different ways [44]. Of these, the method which has implications for the design of a vacuum tube amplifier is frequency domain multiplexing (FDM) in which the input to the tube comprises a number of carriers each of which is modulated by a separate data stream [42]. In a typical scheme the available bandwidth is occupied by N carriers at frequency intervals $2\Delta\omega$. If the centre frequency is ω_0 then the channel frequencies are

$$\omega_n = \omega_0 + n\Delta\omega \quad \text{for } n = \pm 1, \pm 3, \pm 5..., \tag{1.49}$$

where each of these sub-carriers is frequency-modulated. The non-linearity of the final power amplifier produces mixing between the carriers leading to co-channel interference. It is therefore necessary to have some way of calculating this interference and specifying the maximum acceptable level.

The simplest approach is to consider the central pair of carriers at frequencies $\omega = \omega_0 \pm \Delta\omega$ with equal amplitudes. The two carriers each contribute half of the RF input power so that their amplitudes are related to that of a single carrier having the same total power by $V_c = V_1/\sqrt{2}$. The input to the final power amplifier can be written in the form given in (1.43)

$$V_{in}(t) = 2V_c \cos(\Delta\omega t)\cos(\omega_0 t) = r(t)\cos(\omega_0 t), \tag{1.50}$$

where $r(t)$ is the time-varying amplitude. Then, from (1.11), the output of the amplifier is

$$V_2(t) = A[r(t)]\cos(\omega_0 t + \Phi[r(t)]). \tag{1.51}$$

The non-linearity of the amplifier produces signals at frequencies given by

$$\omega_{m,n} = m(\omega_0 + \Delta\omega) \pm n(\omega_0 - \Delta\omega), \tag{1.52}$$

where $m,n = 0,1,2\ldots$. The amplitudes of the signals at these frequencies can be found by Fourier analysis [45]. Thus, setting $m = 1$ and $n = 0$ gives the carrier frequency $\omega_{1,0} = \omega_0 + \Delta\omega$ at which the amplitude is

$$V_{1,0}(V_1) = \frac{1}{\pi}\int_0^\pi A[r(\phi)]\cos\{\Phi[r(\phi)]\}\sin\phi\,d\phi, \tag{1.53}$$

where $\phi = \Delta\omega t$. The analysis is conveniently carried out in the base band. Since the AM/AM and AM/PM characteristics of the amplifier normally vary only slowly with frequency it is possible to assume that they are constant, so that the amplitude of the other carrier $(V_{0,1})$ is the same. Hence the total RF output power at the frequencies of the two carriers is given by replacing V_c in (1.50) by the equivalent single carrier amplitude V_1. This power can be plotted against the RF input power for comparison with the single-carrier AM/AM characteristic as shown in Figure 1.15. At small input powers the amplifier is linear and the power gain is identical to that with one carrier. However, as the input power approaches the single-carrier saturation level the output power is less with two carriers than with one. The reason for this is illustrated in Figure 1.16 which shows typical two-tone input and output waveforms for an amplifier, normalised to the saturated output voltage of the amplifier. For clarity of explanation the carrier frequency shown is much lower relative to the frequency of the envelope than would normally be the case. The output voltage is reduced when the instantaneous amplitude of the signal exceeds the saturation level of the amplifier. Thus the saturated output power under multi-carrier operation is less than that for a single carrier at the same input power. Figure 1.15 also compares the effects of AM/AM conversion alone (typical of klystrons) with those in which AM/PM conversion is also included (typical of TWTs) [45].
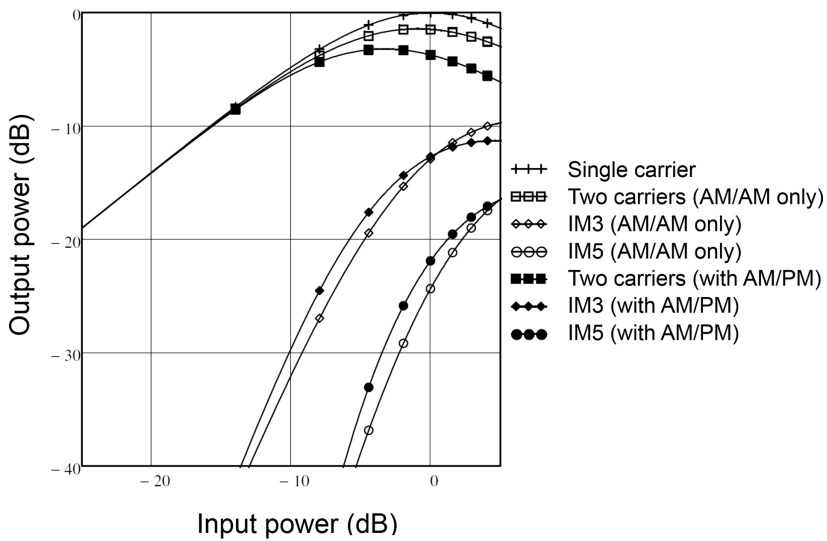
Figure 1.15: Typical AM/AM transfer characteristics for two-carrier operation showing third-order (IM3) and fifth-order (IM5) intermodulation products.
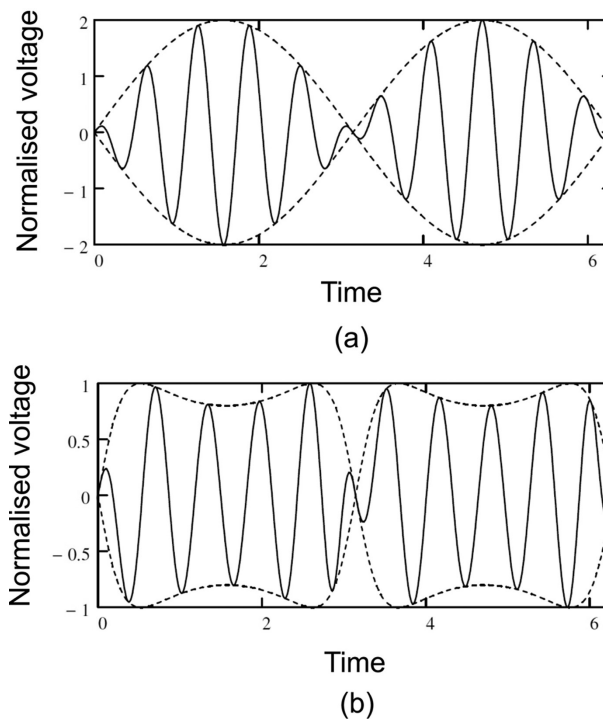
Figure 1.16: Typical waveforms for two-carrier operation of a non-linear amplifier: (a) input, and (b) output.

When we set $m = 2$ and $n = 1$ in (1.52) we obtain a third-order intermodulation product at frequency $\omega_{2,1} = \omega_0 + 3\Delta\omega$ whose amplitude is

$$V_{2,1}(V_1) = \frac{1}{\pi}\int_0^\pi A(2V_c \cos\phi)\cos\left[\Phi(2V_c \cos\phi)\right]\sin 3\phi\, d\phi. \tag{1.54}$$

In the same way, setting $m = 3$ and $n = 2$ gives a fifth-order intermodulation product at $\omega_{3,2} = \omega_0 + 5\Delta\omega$ whose amplitude is

$$V_{3,2}(V_1) = \frac{1}{\pi}\int_0^\pi A(2V_c \cos\phi)\cos\left[\Phi(2V_c \cos\phi)\right]\sin 5\phi\, d\phi. \tag{1.55}$$

The amplitudes of the other intermodulation products at $\omega_{1,2}$ and $\omega_{2,3}$ are the same. Then the total third-order (IM3) and fifth-order (IM5) intermodulation powers can be plotted as shown in Figure 1.15. The frequencies of these signals coincide with those of possible adjacent carriers resulting in co-channel interference. The acceptable level of interference is commonly expressed as the ratio $(C/I3)$ of the fundamental power to that of the third-order intermodulation product, expressed in decibels. At low drive levels the slope of the graph of the third-order product is 3 dB/dB and that of the fifth-order product is 5 dB/dB. Thus the ratio $C/I3$ can be increased by operating the amplifier at a reduced drive level. This is commonly expressed in terms of the output back-off from saturation (in dB) necessary to achieve an acceptable value of $C/I3$. The linear parts of the curves for the inter-modulation products can be extrapolated in the same manner as in Figure 1.4 to meet the extrapolation of the fundamental curve at the third-order and fifth-order intercept points. These may also be used to specify the linearity of the amplifier.

The functions A and Φ required for the evaluation of the integrals in (1.53) to (1.55) may be specified numerically using interpolation on data points which have been determined experimentally or by large-signal modelling [45]. Alternatively a suitable function may be fitted to the data points [42, 43, 46]. Polynomial expansions do not give a good fit to the data unless a large number of terms is used. In particular these models fail for drive levels approaching saturation. They do, however, reveal that the fundamental, and intermodulation, amplitudes depend only on the odd terms of the series. These can be specified independently of the terms describing the dependence of the even harmonics on the input RF power [46]. Simple functions which are a good fit to experimental data for TWTs have been proposed by Saleh [43]:

$$A(r) = \frac{\alpha_a r}{1 + \beta_a r^2} \tag{1.56}$$

and

$$\Phi(r) = \frac{\alpha_\phi r^2}{1 + \beta_\phi r^2}, \tag{1.57}$$

where α_a, β_a, α_ϕ and β_ϕ are empirical constants. If the amplifier has unity gain at saturation, so that $A(1) = 1$, it can be shown that $\alpha_a = 2$ and $\beta_a = 1$. These values are

close to the empirical figures so that it is possible to use them as a useful approxima-
tion. In the same way it may be tentatively suggested that the function required to
generate the even harmonics might be written

$$B(r) = \frac{\alpha_h r^2}{\left(1 + \beta_h r^2\right)^2}.$$

(1.58)

The curves shown in Figures 1.4 and 1.15 were generated using this non-linear
model (see Worksheet 1.3).

The ratio $C/I3$ is commonly used for specifying the intermodulation distortion
in an amplifier, not least because it is straightforward to measure it experimentally.
However, it is at best, a proxy for the intermodulation effects because, when there
are more than two carriers, the overall performance depends upon the signal levels
of all the carriers. Since the phases of the phase-modulated carriers vary in a quasi-
random manner it may be assumed that all phase combinations are equally prob-
able. Where only two carriers are used it is sufficient to consider unmodulated sine
waves because all possible phases arise in one period of the difference frequency.
Then the peak power is twice the mean power. However, with three or more carriers
this is not the case, and it is necessary to use Monte Carlo methods to examine all
possible phases [47]. As the number of carriers increases, the probability that all the
signals will be in phase simultaneously decreases. In the limit, a large number of
uncorrelated phase-modulated carriers can be represented by band-limited white
noise [45, 48]. This leads to an alternative measure of intermodulation distortion in
which the input to the amplifier is band-limited white noise from which the portion
in a small frequency range has been removed by a notch filter. The non-linearity of
the amplifier causes some signals to be generated in this frequency range. The ratio
of the amplitude of these to that of the white noise at the output of the amplifier is
the *Noise Power Ratio* (in dB). This is a measure of the level of interference which
would be experienced by a signal at the centre of the notch [49, 50]. An approxima-
tion to this can be achieved experimentally, or in simulations, by using a comb of
carriers at equal frequency intervals such that one carrier has been omitted from
the centre of the set as shown in Figure 1.17(a) [47, 51]. Figure 17(b) shows a typi-
cal output spectrum and the definition of the noise power ratio. Intermodulation
products are also generated on the edges of the input spectrum so increasing its
bandwidth. If the power in these is great enough then it may cause interference with
other systems operating in adjacent frequency bands. For this reason the permitted
bandwidth may be expressed as a spectrum envelope [52].

1.7   Engineering Design

The design of any engineering product can be divided into three phases:

i) Conceptual design, in which the principal dimensions and operating param-
    eters are established.
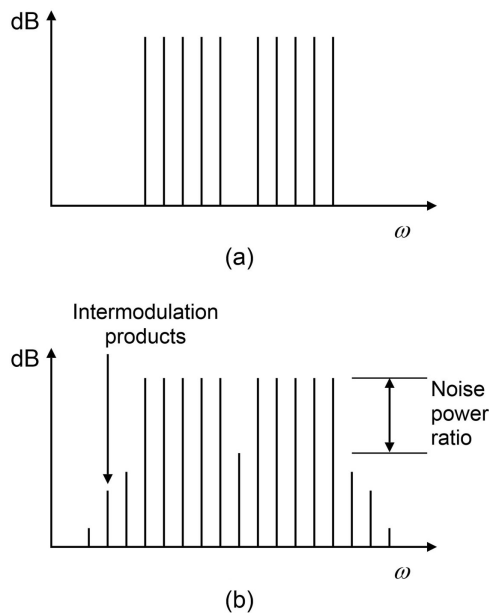
Figure 1.17: Amplification of multiple carriers: (a) input spectrum, and (b) output spectrum.

ii) Detailed design, involving refinement of the conceptual design by prototyping and advanced modelling, leading to the preparation of the manufacturing data package and the test specification.

iii) Design consolidation, incorporating further refinements suggested by product testing and field experience.

The costs that are committed, and incurred, increase at each stage. It is therefore desirable to produce a conceptual design which does not require further modification at the detailed design stage; and a detailed design which does not have to be modified as a result of unsatisfactory performance in the field. It should be noted that performance and environmental testing is normally carried out under laboratory conditions, and care must be taken to ensure that tubes which pass these tests do also perform well in the field. Differences between the testing and the real environments, such as the impedances of power supplies and matches of RF components, can affect how a tube performs.

1.7.1   Dimensionless Parameters and Scaling

Nearly all tubes are developed in organisations in which the design process can be based upon accumulated knowledge and expertise. Many tubes are designed by scaling from existing designs, or by other modifications to them, rather than starting from a clean sheet of paper. It is, however, important that this process should be underpinned by fundamental understanding.

The basis for scaling is dimensional analysis [53, 54]. Let us suppose that some aspect of the performance of a tube is governed by an equation which can be written in the general form

$$F\left(Q_1, Q_2 \ldots Q_n, r', r'' \ldots\right) = 0, \tag{1.59}$$

where $Q_1, Q_2 \ldots$ etc. are dimensional variables and r', r'' etc. are dimensionless ratios. Very commonly these are the ratios of the leading dimensions of the tube to a single dimension or to a dimensional variable such as the free-space wavelength. If two tubes are described by the same set of dimensionless ratios, so that all dimensions are changed by the same multiplying factor, they are said to be *geometrically similar* to one another. If geometrical similarity is assumed then the ratios are constants and (1.59) can be written

$$F\left(Q_1, Q_2 \ldots Q_n\right) = 0. \tag{1.60}$$

This equation is a complete description of the problem, provided that none of the variables required has been overlooked. The choice of variable for inclusion is based upon an understanding of the physics of the problem. When it involves electromagnetics then the primary electric and magnetic constants ε_0 and μ_0 may have to be included in the list. The ratio of the charge to the rest mass of an electron $\left(e/m_0\right)$ is also required when the dynamics of electrons are part of the problem. Equation (1.60) remains valid even if the algebraic form of the function is unknown. The dimensions of the variables Q can be expressed in terms of the fundamental dimensions mass [M], length [L], and time [T], together with voltage [V] or charge for electrical problems. Thus there are, at most, four fundamental dimensions. Table 1.7 shows the dimensions of some quantities relevant to vacuum tubes.

The variables Q may be combined together to form dimensionless groups denoted by Π so that (1.60) can be written

$$\psi\left(\Pi_1, \Pi_2, \ldots \Pi_i\right) = 0. \tag{1.61}$$

The number of groups required is related to the number of variables by

$$i = n - k, \tag{1.62}$$

where k is the number of dimensions required. This is known as Buckingham's theorem. An immediate effect of using dimensionless variables (or dimensionless groups) is that the number of variables required to define the problem has been reduced. Now let us choose one of the groups to be a dependent variable whose value is determined by the values of the others. Then (1.61) becomes

$$\Pi_1 = \Psi\left(\Pi_2, \Pi_3, \ldots \Pi_i\right). \tag{1.63}$$

Table 1.7: Typical independent and dependent parameters for vacuum tubes

	Parameter	Dimensions
Constants	Velocity of light in free space (c) 0.299792×10^9 m s^{-1}	$[LT^{-1}]$
	Primary electric constant (ε_0) 8.854187×10^{-12} F m^{-1}	$[MLT^{-2}V^{-2}]$
	Primary magnetic constant (μ_0) 1.256637×10^{-6} H m^{-1}	$[M^{-1}L^{-3}T^4V^2]$
	Elementary charge (e) 1.602176×10^{-19} C	$[ML^2T^{-2}V^{-1}]$
	Rest mass of the electron (m_0) 9.109382×10^{-31} kg 0.510998 MeV	$[M]$
	Electron charge to mass ratio (e/m_0) 1.758820×10^{11} C kg^{-1}	$[L^2T^{-2}V^{-1}]$
Parameters	Length	$[L]$
	Voltage	$[V]$
	Magnetic flux density	$[L^{-2}TV]$
	Frequency	$[T^{-1}]$
	R.F. power	$[ML^2T^{-3}]$
	Current	$[ML^2T^{-3}V^{-1}]$
	Electron velocity	$[LT^{-1}]$
	Charge density	$[ML^{-1}T^{-2}V^{-1}]$

If the values of the independent variables $(\Pi_2, \Pi_3, \ldots \Pi_i)$ are fixed then so is the value of the dependent variable Π_1. Two tubes which are described by the same values of the dimensionless variables are said to be *dynamically similar* to one another. Thus the basis of scaling is the maintenance of both geometrical and dynamic similarity.

The process of forming dimensionless groups is aided by choosing normalisations which are based on the underlying physics of the problem. Some examples are shown in Table 1.8.

A simple illustration is provided by the calculation of the velocity of an electron (u_0) from an applied voltage (V_a) and the dimensional constants. The dependent dimensionless group is

$$\Pi_1 = \frac{u_0}{c} \tag{1.64}$$

and the independent dimensionless group is

$$\Pi_2 = \frac{eV_a}{m_0 c^2}. \tag{1.65}$$

Table 1.8: Examples of the formation of dimensionless groups

Variable	Normalising quantity	Dimensionless group
Dimension (d)	Wavelength (λ) Propagation constant (β)	$\dfrac{d}{\lambda}$ or βd
Velocity (u)	Velocity of light (c) Phase velocity (v_p)	$\dfrac{u}{c}$ or $\dfrac{u}{v_p}$
Voltage (V)	Rest energy of the electron in eV $\left(V_0 = m_0 c^2 / e\right)$	$\dfrac{eV}{m_0 c^2}$
Electron plasma frequency $\left(\omega_p\right)$ Electron cyclotron frequency $\left(\omega_c\right)$	Frequency (ω)	$\dfrac{\omega_p}{\omega}$ and $\dfrac{\omega_c}{\omega}$

In this case there are five parameters and three independent dimensions: $[\text{M}]$, $[\text{V}]$ and $[\text{LT}^{-1}]$ so that the problem is completely described by the two dimensionless groups which have been formed. Note that in this case the dimensions $[\text{L}]$ and $[\text{T}]$ do not occur separately so that the number of independent dimensions is reduced by one. It follows that we can write

$$\Pi_1 = \Psi\left(\Pi_2\right), \tag{1.66}$$

where Ψ is an unknown function which can be determined empirically, even if its form cannot be found by mathematical analysis. The value of the reduction of the description of a problem to the form in (1.63) is greater than just that of scaling. If the form of the function Ψ is known, even over a limited range of variation of the variables then it can be used to explore the effects of design choices on Π_1. In cases where geometrical similarity is not maintained then one, or more, of the ratios r are added as dimensionless variables. Examples of the application of this method to particular problems in tube design are given later in this book.

1.7.2   Modelling

The trend in tube design is towards achieving near-ideal RF performance coupled with high reliability and low cost of ownership. For this reason, the design of a state-of-the-art vacuum tube is a complex business which involves considerable use of computer simulations to achieve right-first-time design [55, 56]. The advanced computer codes used for detailed design can take many hours to run, even on the most powerful computers available. They are invaluable tools for accurate modelling, but are not well-suited to the conceptual design stage which is the focus of this book. For that, the engineers must have a sound understanding of the theory of vacuum tubes, together with simple computer models which can be used for rapid design iterations.

The availability of advanced computer modelling tools such as Particle in Cell (PIC) codes can mislead the unwary into supposing that the best way of modelling a tube is to create a complete model in such a code. This approach is wrong on two counts:

 i) A PIC code finds a self-consistent numerical solution of Maxwell's equations and the equations of motion in a series of time steps. However, the solution is linear in the part of the problem space which is not occupied by electrons. It is therefore more efficient to compute that part of the solution once and store it for future use. The problem is then reduced to the solution of the non-linear model in the space containing electrons, subject to matching the stored solution on the boundary. This is the method used in the large-signal models of tubes which are described later in this book.

ii) Even the best computer codes cannot be wholly accurate models of reality because the complexity of the problem is too great, so that some parts of it may not be known precisely. Assumptions and approximations are always required. For this reason the mathematical modelling of tubes is as much an art as it is a science [57]. The best results are obtained when the physics of the problem are well understood so that the assumptions and approximations are valid. Any model is only useful when it has been validated by comparison with experimental results. Even then it can only be used with caution outside the range of parameters for which it has been validated.

A large number of simple models were created in Mathcad14® during the writing of the book and used to generate many of the figures. These models can be used to explore the behaviour of different aspects of tubes by changing the parameters. They are available in electronic form (see the list of worksheets in the Appendix, available online at www.cambridge.org/9780521198622). The comments included in them should be sufficient for users to create their own models using other software if they wish. The models are intended as educational tools and have not been validated for use in tube design.

References

[1]   J. A. Fleming, 'Improvements in instruments for detecting and measuring alternating electric currents', UK Patent GB190424850 (A) 1904.

[2]   R. J. Barker *et al.*, eds, *Modern Microwave and Millimeter-Wave Power Electronics*. Piscataway, NJ: IEEE Press; Hoboken, NJ: Wiley-Interscience, 2005.

[3]   W. H. Kohl, *Handbook of Materials and Techniques for Vacuum Devices*. New York: American Institute of Physics, 1995.

[4]   F. Rosebury, *Handbook of Electron Tube and Vacuum Techniques*. New York: American Institute of Physics, 1993.

[5]   R. J. Trew, 'SiC and GaN transistors – is there one winner for microwave power applications?', *Proceedings of the IEEE,* vol. 90, pp. 1032–1047, 2002.

[6]   M. Meneghini *et al.*, *Power GaN Devices: Materials, Applications and Reliability*. Springer, 2016.

[7] R. S. Pengelly *et al.*, 'A review of GaN on SiC high electron-mobility power transistors and MMICs', *IEEE Transactions on Microwave Theory and Techniques,* vol. 60, pp. 1764–1783, 2012.

[8] M. Kasu *et al.*, 'Diamond-based RF power transistors: fundamentals and applications', *Diamond & Related Materials,* vol. 16, pp. 1010–1015, 2007.

[9] V. Camarchia *et al.*, 'An overview on recent developments in RF and microwave power H-terminated diamond MESFET technology', presented at the 2014 International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC), 2014.

[10] M. A. Richards *et al.*, *Principles of Modern Radar*. Edison, NJ: SciTech, 2010.

[11] P. Marchand *et al.*, 'High power 352 MHz solid state amplifiers developed at the Synchrotron SOLEIL', *Physical Review Special Topics – Accelerators and Beams,* vol. 10, p. 112001, 2007.

[12] R. Caverly *et al.*, 'Advancements at the lower end: advances in HF, VHF, and UHF systems and technology', *IEEE Microwave Magazine,* vol. 16, pp. 28–49, 2015.

[13] H. P. Kindermann *et al.*, 'The RF power plant of the SPS', *IEEE Transactions on Nuclear Science, Vol. NS-30, No. 4,* vol. NS-30, pp. 3414–3416, August 1983.

[14] M. Jensen, 'Inductive output tube based 300 kW RF amplifier for the Diamond Light Source', in *EPAC 2004*, Lucerne, Switzerland, pp. 962–964, 2004.

[15] J. H. Booske *et al.*, 'Traveling-wave tubes', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE, pp. 171–245, 2005.

[16] V. Granatstein *et al.*, 'Vacuum electronics at the dawn of the twenty-first century', *Proceedings of the IEEE,* vol. 87, pp. 702–716, 1999.

[17] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices.* Berlin: Springer-Verlag, pp. 1–84, 2008.

[18] J. H. Booske and R. J. Barker, 'Introduction and overview', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 1–33, 2005.

[19] N. Ayllon, 'Microwave high power amplifier technologies for space-borne applications', presented at the 2015 IEEE 16th Annual Wireless and Microwave Technology Conference (WAMICON), 2015.

[20] W. G. V. Rosser, *An Introduction to the Theory of Relativity*. London: Butterworths, 1964.

[21] D. Gabor, 'Energy conversion in electronic devices', *Journal of the Institution of Electrical Engineers – Part III: Communication Engineering, including the Proceedings of the Wireless Section of the Institution,* vol. 91, pp. 128–141, 1944.

[22] National Frequency Planning Group, 'United Kingdom Frequency Allocation Table', Committee on UK Spectrum Strategy, 2013.

[23] National Telecommunications and Information Administration, 'Frequency Allocations', in *Manual of Regulations and Procedures for Federal Radio Frequency Management*, US Department of Commerce, 2014.

[24] J. R. Pierce, 'Coupling of modes of propagation', *Journal of Applied Physics,* vol. 25, pp. 179–183, 1954.

[25] W. H. Louisell, *Coupled Mode and Parametric Electronics*. New York: Wiley, 1960.

[26] R. J. Briggs, *Electron-stream Interaction with Plasmas*. Cambridge, MA: MIT Press, 1964.

[27] C. C. Johnson, *Field and Wave Electrodynamics*. New York: McGraw-Hill, 1965.

[28] J. R. Pierce, 'The wave picture of microwave tubes', *The Bell System Technical Journal,* vol. 33, pp. 1343–1372, 1954.

[29] R. N. Sudan, 'Classification of instabilities from their dispersion relations', *Physics of Fluids,* vol. 8, pp. 1899–1904, 1965.

[30] K. R. Chu and A. T. Lin, 'Gain and bandwidth of the gyro-TWT and CARM amplifiers', *IEEE Transactions on Plasma Science,* vol. 16, pp. 90–104, 1988.

[31] J. Feinstein and K. Felch, 'Status review of research on millimeter-wave tubes', *IEEE Transactions on Electron Devices,* vol. 34, pp. 461–467, 1987.

[32] G. Bekefi, 'Survey of physics research in microwave devices', in *International Electron Devices Meeting*, pp. 822–825, 1984.

[33] A. F. Harvey, 'Microwave tubes – an introductory review with bibliography', *Proceedings of the IEE – Part C: Monographs,* vol. 107, pp. 29–59, 1960.

[34] N. C. Luhmann, Jr. *et al.*, 'Historical highlights', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimeter-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 35–106, 2005.

[35] V. Granatstein *et al.*, 'Vacuum electronics at the dawn of the twenty-first century', *Proceedings of the IEEE,* vol. 87, pp. 702–716, 2002.

[36] T. E. Yingst *et al.*, 'High-power gridded tubes – 1972', *Proceedings of the IEEE,* vol. 61, pp. 357–381, 1973.

[37] L. Sivan, *Microwave Tube Transmitters*. London: Chapman & Hall, 1994.

[38] NATO, AC243, Panel 3, RSG-19, *DRG Handbook: Microwave and Millimetre Wave Tubes and Power Supplies*. NATO, 1998.

[39] J. Dunlop and D. G. Smith, *Telecommunications Engineering*. Wokingham, UK: Van Nostrand Reinhold (UK) Co. Ltd, 1984.

[40] R. Ziemer, 'Modulation', in R. A. Meyers, ed., *Encyclopedia of Telecommunications*. San Diego, CA: Academic Press, pp. 201–213, 1989.

[41] J. B. Anderson, *Digital Transmission Engineering [electronic resource]*, 2nd ed. Piscataway, NJ: IEEE Press; Hoboken, NJ: Wiley-Interscience, 2005.

[42] A. Berman and C. E. Mahle, 'Nonlinear phase shift in traveling-wave tubes as applied to multiple access communications satellites', *IEEE Transactions on Communication Technology,* vol. 18, pp. 37–48, 1970.

[43] A. A. M. Saleh, 'Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers', *IEEE Transactions on Communications,* vol. 29, pp. 1715–1720, 1981.

[44] R. Ziemer, 'Multiplexing', in R. A. Meyers, ed., *Encyclopedia of Telecommunications*. San Diego, CA: Academic Press, pp. 215–220, 1989.

[45] G. R. Stette, 'Calculation of intermodulation from a single carrier amplitude characteristic', *IEEE Transactions on Communications,* vol. 22, pp. 319–323, 1974.

[46] W. E. Kunz *et al.*, 'Traveling-wave tube amplifier characteristics for communications', *Microwave Journal,* vol. 10, pp. 41–46, 1967.

[47] R. G. Carter *et al.*, 'Computer simulation of intermodulation distortion in traveling wave tube amplifiers', *IEEE Transactions on Electron Devices,* vol. 48, pp. 178–180, January 2001.

[48] A. Guida, 'Accurate calculation of TWT intermodulation in the many-carrier case', *IEEE Transactions on Communications,* vol. 35, pp. 685–687, 1987.

[49] T. Reveyrand *et al.*, 'A novel experimental noise power ratio characterization method for multicarrier microwave power amplifiers', in *55th ARFTG Conference Digest-Spring*, pp. 1–5, 2000.

[50] A. Katz and R. Gray, 'Noise power ratio measurement tutorial', *Linearizer Technology, Inc.*, www.lintech.com, 2012.

[51] C. Loo, 'Calculation of the suppression of signals and intermodulation noise when multiple unequal carriers are amplified by a TWT', *Canadian Electrical Engineering Journal,* vol. 2, pp. 29–32, 1977.

[52] F. Amoroso, 'The bandwidth of digital data signals', *IEEE Communications Magazine,* vol. 18, pp. 13–24, 1980.

[53] E. Buckingham, 'On physically similar systems; illustrations of the use of dimensional equations', *Physical Review,* vol. 4, pp. 345–376, 1914.

[54] E. B. Wilson Jr., *An Introduction to Scientific Research*. New York: McGraw-Hill, 1952.

[55] T. M. Antonsen, Jr. *et al.*, 'Advances in modeling and simulation of vacuum electronic devices', *Proceedings of the IEEE,* vol. 87, pp. 804–839, 1999.

[56] R. G. Carter, 'Computer modelling of microwave tubes – a review', in *2nd IEEE International Vacuum Electronics Conference 2001*, Noordwijk, Netherlands, pp. 393–396, 2001.

[57] P. B. Johns, 'The art of modelling', *Electronics and Power,* vol. 25, pp. 565–569, 1979.

2   Waveguides

2.1   Introduction

Modern vacuum tubes are power amplifiers and oscillators which require the use of waveguides or coaxial lines to convey RF power into and out of them. The purpose of this chapter is to provide a summary of those topics which are important for the design of vacuum tubes. It also provides a foundation for the discussions of resonators in Chapter 3, and of slow-wave structures in Chapter 4. Section 2.2 summarises the theory of hollow metal waveguides, and two-conductor transmission lines, having uniform cross-sections. This leads to a discussion of practical coaxial lines, and rectangular, ridged and circular waveguides in Section 2.3. The properties of simple discontinuities in rectangular waveguides are considered in Section 2.4 followed by a discussion of matching techniques in Section 2.5. Sections 2.6 and 2.7 examine methods of coupling between waveguides of different cross-sections without, and with, changes in the mode of propagation. The final section reviews the different kinds of vacuum windows which are used in coaxial lines and rectangular waveguides. The theory and practice of waveguides and waveguide components is covered by many books and the reader is referred to them for detailed information [1–5].

2.2   Waveguide Theory

The propagation of electromagnetic waves in a source-free region, filled with a uniform material of permittivity ε and permeability μ, is governed by the wave equations

$$\nabla^2 \mathbf{E} - \varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \qquad (2.1)$$

and

$$\nabla^2 \mathbf{H} - \varepsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} = 0, \qquad (2.2)$$

where $\mathbf{E}$ is the electric field and $\mathbf{H}$ is the magnetic field [1]. We shall assume that the waves are guided in the z direction by conducting boundaries whose shape

does not vary with z. The vector operator can be decomposed into transverse and longitudinal parts

$$\nabla = \nabla_T + \hat{\mathbf{z}}\frac{\partial}{\partial z},$$

(2.3)

where $\hat{\mathbf{z}}$ is the unit vector in the z direction. We also assume the propagation of harmonic waves with frequency ω in the z direction so that

$$\mathbf{E} = \mathbf{E}_0 \exp j\left(\omega t - \beta_g z\right)$$

(2.4)

and

$$\mathbf{H} = \mathbf{H}_0 \exp j\left(\omega t - \beta_g z\right),$$

(2.5)

where the vector amplitudes depend only upon the spatial variables normal to the z axis. Substituting for $\mathbf{E}$ and $\mathbf{H}$ in (2.1) and (2.2) yields

$$\nabla_T^2 \mathbf{E}_0 + \left(\beta_w^2 - \beta_g^2\right)\mathbf{E}_0 = 0$$

(2.6)

and

$$\nabla_T^2 \mathbf{H}_0 + \left(\beta_w^2 - \beta_g^2\right)\mathbf{H}_0 = 0,$$

(2.7)

where

$$\beta_w^2 = \varepsilon\mu\omega^2$$

(2.8)

is the propagation constant of plane electromagnetic waves in the material filling the waveguide. We are only interested in the cases where the waveguide is empty, or filled with a low-loss dielectric material so that, effectively, $\mu = \mu_0$. The phase velocity of plane waves is

$$v_{p0} = \frac{\omega}{\beta_w} = \frac{1}{\sqrt{\varepsilon\mu}}.$$

(2.9)

If the waveguide is empty then $v_{p0} = c$, the velocity of light in free space, and the propagation constant is $\beta_w = \omega/c$. The solutions of (2.6) and (2.7) are eigenfunctions which satisfy the boundary conditions on the conducting surfaces. There is an infinity of such solutions and each has a corresponding eigenvalue β_C such that

$$\beta_C^2 = \beta_w^2 - \beta_g^2$$

(2.10)

or

$$\beta_g^2 = \beta_w^2 - \beta_C^2.$$

(2.11)

It is evident, from (2.11), that waves can only propagate when β_g is real, that is, when $\beta_w^2 > \beta_C^2$. But, from (2.8), β_w is proportional to the frequency of the waves

and, therefore, waves can propagate only at frequencies above the cut-off frequency given by

$$\omega_C = \beta_C v_{p0}. \tag{2.12}$$

Equation (2.11) can also be written

$$\omega = \sqrt{\omega_C^2 + \beta_g^2 v_{p0}^2}. \tag{2.13}$$

Every mode in a waveguide obeys (2.13) and has its own cut-off frequency. The plot of this curve, shown in normalised form in Figure 2.1, is the dispersion $(\omega - \beta)$ diagram of the waveguide mode. We shall see that dispersion diagrams are valuable tools for understanding the properties of wave guiding systems in general.

The phase velocity of the wave in the waveguide is

$$v_p = \frac{\omega}{\beta_g} \tag{2.14}$$

and this is represented by the slope of the line joining a point on the dispersion curve to the origin. The dashed line shows the dispersion curve for plane waves in unbounded material for which the phase velocity is v_{p0}. It can be seen that for guided waves $v_p > v_{p0}$ at all frequencies. The group velocity, which is the velocity of propagation of information along the waveguide, is the slope of the dispersion curve, i.e.

$$v_g = \frac{d\omega}{d\beta_g} \tag{2.15}$$
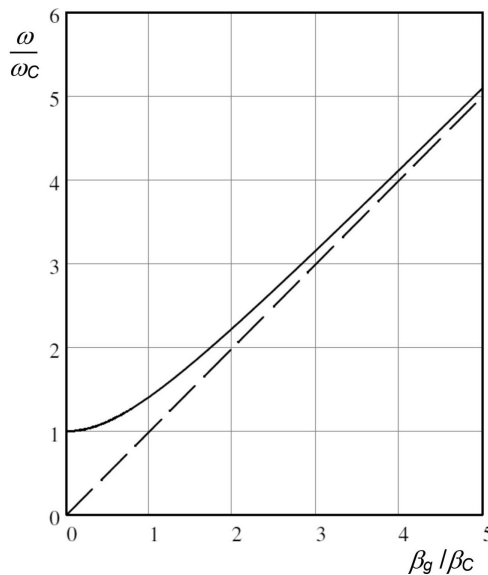
Figure 2.1: Normalised dispersion diagram for waves in a uniform waveguide.

and we see that $v_g < v_{p0}$ and that $v_g \rightarrow 0$ at cut-off. An important special case, which is discussed below, occurs when $\beta_C = 0$ and the dispersion curve is the straight dashed line through the origin.

When $\omega < \omega_C$ the waveguide is cut off and (2.11) can be written

$$\beta_g = \pm j\beta_w \sqrt{(\omega_C / \omega)^2 - 1}, \tag{2.16}$$

where the negative sign is taken for waves travelling in the positive z direction. The mode is then an *evanescent mode* and the amplitude of the wave then decays exponentially with z. The attenuation of the cut-off mode in decibels per free-space wavelength is

$$20 \log\left(\exp\left(-|\beta_g| \frac{2\pi}{\beta_w} \right) \right) = -40\pi \log(e) \sqrt{(\omega_C / \omega)^2 - 1}. \tag{2.17}$$

Thus, for example, if $\omega_C / \omega = 1.05$ the attenuation in one free-space wavelength is 17.5 dB. We shall explore the significance of this result when discussing higher-order mode effects in waveguides.

It is convenient to express the vector amplitudes of the waves as the sum of longitudinal and transverse vectors [3] so that

$$\begin{aligned} \mathbf{E}_0 &= \mathbf{E}_T + \hat{\mathbf{z}} E_z \\ \mathbf{H}_0 &= \mathbf{H}_T + \hat{\mathbf{z}} H_z, \end{aligned} \tag{2.18}$$

where $\mathbf{E}_T$ and $\mathbf{H}_T$ only have transverse components. By substituting these expressions into Maxwell's equations and making use of vector identities it can be shown that

$$\mathbf{E}_T = -\frac{j\beta_g}{\beta_C^2} (\nabla_T E_z) + \frac{j\omega\mu}{\beta_C^2} (\hat{\mathbf{z}} \wedge \nabla_T H_z) \tag{2.19}$$

and, similarly,

$$\mathbf{H}_T = -\frac{j\omega\varepsilon}{\beta_C^2} (\hat{\mathbf{z}} \wedge \nabla_T E_z) - \frac{j\beta_g}{\beta_C^2} (\nabla_T H_z). \tag{2.20}$$

This pair of equations shows that the transverse field components can be determined if the axial components are known. Furthermore they are expressed as the sum of two terms, one of which depends only on the axial electric field, and the other only on the axial magnetic field. It can be shown that a waveguide supports an infinite set of normal modes of propagation which are orthogonal to one another so that any general configuration of the fields can be expressed as a sum of normal modes. For every waveguide mode it can be shown that the electric and magnetic fields are mutually orthogonal. The time-average power flow can be found by integrating the longitudinal component of the Poynting vector over a plane normal to the z axis

$$P = \frac{1}{2} \iint (\mathbf{E} \wedge \mathbf{H})_z \, dS. \tag{2.21}$$

Three distinct cases are considered in the sections which follow.

2.2.1 The Transverse Electric and Magnetic Mode

If $\beta_C = 0$ then (2.19) and (2.20) permit the existence of non-zero solutions for the transverse field components when the longitudinal field components are both zero. The wave is then a plane electromagnetic wave which satisfies the boundary conditions and whose phase velocity is v_{p0}. This is the transverse electric and magnetic (*TEM*) mode shown by the dashed line in Figure 2.1. It can be shown that the transverse electric field can be found from the gradient of a scalar electric potential which is a solution of Laplace's equation in two dimensions. Since Laplace's equation can only have a non-zero solution in a closed region in which there are at least two independent boundaries it follows that *TEM* waves can only be supported by waveguides comprising two or more conductors, for example coaxial lines. The fact that the cut-off frequency is zero means that the waveguide must be able to support an electric field under DC conditions. This confirms the requirement that it must have at least two independent conductors. Figure 2.2(a) shows the field pattern of the *TEM* wave in a coaxial line.

The properties of each mode of a uniform waveguide can be represented by a lumped-element equivalent circuit constructed by representing displacement currents by capacitors, and conduction currents by inductors. Equivalent circuits, constructed in this way, are also valuable aids to understanding the properties of cavity resonators and slow-wave structures. The displacement current in a *TEM* line is in the transverse direction. The conduction current is normal to the direction of the tangential magnetic field at the wall and is, therefore, in the longitudinal direction. Thus the equivalent Tee circuit of a short length dz of the line can be constructed as shown in Figure 2.2(b) where C and L are, respectively, the capacitance and inductance per unit length. Analysis of this circuit shows that the propagation constant is

$$\beta_g = \pm \omega\sqrt{LC}. \tag{2.22}$$

But, from (2.11), we see that $\beta_g = \beta_w$, the propagation constant for plane waves in the material filling the space between the conductors, so the phase velocity is

$$v_p = \frac{1}{\sqrt{\varepsilon\mu}} = \frac{1}{\sqrt{LC}} = v_{p0}. \tag{2.23}$$
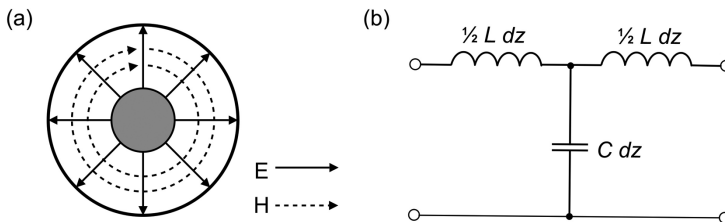
Figure 2.2: Coaxial line *TEM* mode: (a) field pattern, and (b) equivalent circuit for the mode.

The line is non-dispersive because the phase velocity is constant. The characteristic impedance is defined as

$$Z_c = \frac{|V|^2}{2P} \tag{2.24}$$

where V is the magnitude of the potential difference between the conductors, and P is the power flow. The value of Z_c can be calculated from any pair of: the magnitude of the voltage; the magnitude of the current; and the power flow, yielding identical results.

In terms of the equivalent circuit

$$Z_c = \sqrt{\frac{L}{C}}. \tag{2.25}$$

The capacitance per unit length of any *TEM* line can be found from the solution of the two-dimensional Laplace equation, and the inductance per unit length can then be calculated from (2.22).

2.2.2 Transverse Electric Modes

Modes in which the longitudinal component of the electric field is zero are known as Transverse Electric (or *TE*) modes. They are also sometimes known as *H* modes because they have a longitudinal component of the magnetic field. The longitudinal magnetic field is a solution of

$$\nabla_T^2 H_z + \beta_C^2 H_z = 0, \tag{2.26}$$

subject to the condition that $\nabla_T H_z = 0$ on the conducting boundaries. Each solution comprises an eigenfunction H_z of the transverse coordinates, which is independent of frequency, with an associated eigenvalue β_C. Once these are known the corresponding transverse field components can be found directly from (2.19) and (2.20). Figure 2.3 shows the arrangement of a rectangular waveguide. The field patterns of some *TE* modes are shown in Figure 2.4. The transverse components of the
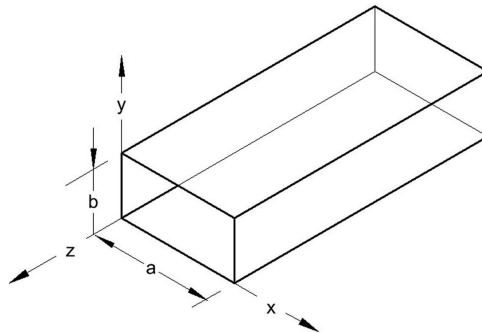
Figure 2.3: Arrangement of a rectangular waveguide.

TE MODES IN RECTANGULAR WAVEGUIDE

$$\lambda_c = 2a$$

TE$_{10}$

$$\lambda_c = \frac{2a}{\sqrt{1 + (a/b)^2}}$$

TE$_{11}$

$$\lambda_c = \frac{a}{\sqrt{1 + (a/2b)^2}}$$

TE$_{21}$

1. Cross-sectional view
2. Longitudinal view
3. Surface view

a. Inside Broad Dimension
b. Inside Narrow Dimension

I
E
H
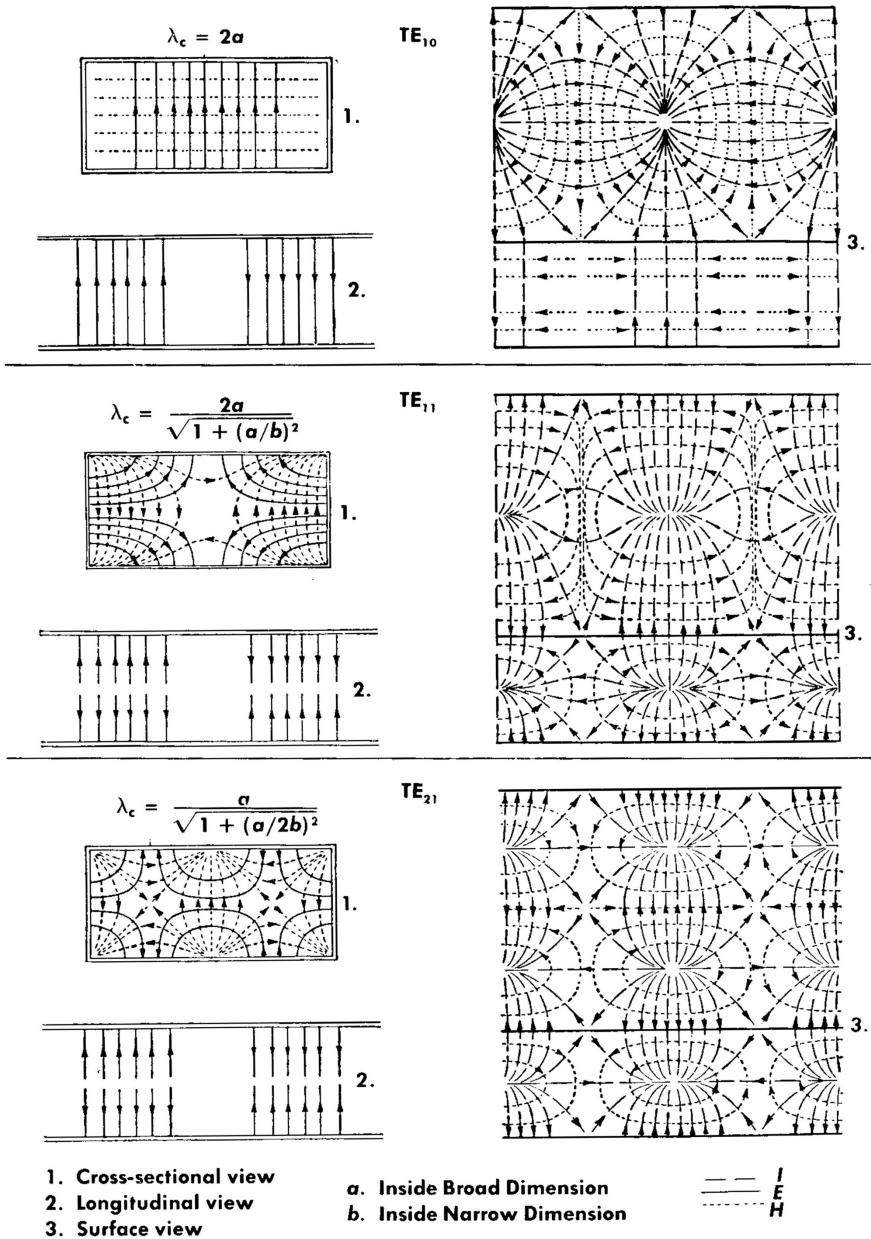
Figure 2.4: Field patterns of *TE* modes in a rectangular waveguide (reproduced, with permission from [5]).

electric and magnetic fields are always at right angles to one another. The ratio of their amplitudes is the wave impedance which is found from (2.19) and (2.20) to be

$$Z_{TE} = \left| \frac{E_T}{H_T} \right| = \frac{\omega\mu}{\beta_g} = \frac{\beta_w}{\beta_g} Z_w, \tag{2.27}$$

where $Z_w = \sqrt{\mu/\varepsilon}$ is the wave impedance of plane waves in the material filling the waveguide. Equation (2.27) is valid for all TE modes and it has been assumed that the material filling the guide, if any, is non-magnetic. It can be seen that $Z_{TE} \to Z_w$ as $\omega \to \infty$.

It is common to use transmission line methods to solve waveguide problems with the guide wavelength defined by

$$\lambda_g = \frac{2\pi}{\beta_g}. \tag{2.28}$$

The characteristic impedance is not unique, as it is for TEM waves. This arises because the electric and magnetic fields are not uniform over the surface of the waveguide so that the voltage and current can be defined in different ways [1]. Typically:

- The voltage V is defined as the maximum voltage between opposite sides of the waveguide.
- The current I is defined as the amplitude of the total axial current in the walls of the waveguide.

However, the power calculated from the voltage and current defined in this way is not equal the power flow (P) calculated from the Poynting vector. Thus the characteristic impedance can be defined by

$$Z_{PV} = \frac{V^2}{2P}, \tag{2.29}$$

$$Z_{PI} = \frac{2P}{I^2}, \tag{2.30}$$

or

$$Z_{VI} = \frac{V}{I} \tag{2.31}$$

and these equations give different results. To make matters more complicated, the relationship between them is not fixed but depends upon the shape of the waveguide. For problems where only one propagating mode is concerned (e.g. a change in the height of a rectangular waveguide) it does not matter which definition is used because the scaling factors are the same for waveguides having the same shape. There is no consensus in the literature about which definition of the characteristic impedance should be used for any given problem.

From Figure 2.4 it can be seen that the simplest TE mode (TE_{10}) has displacement current in the vertical direction. Axial and tangential conduction currents flow in the walls, corresponding to the transverse and axial components of the magnetic
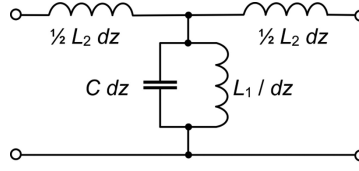
Figure 2.5: Equivalent circuit of the TE_{10} mode in a rectangular waveguide.

field. (The notation used to describe the modes is explained in Section 2.3.2.) The equivalent circuit is therefore as shown in Figure 2.5. From analysis of the circuit for a section of waveguide which is short compared with the wavelength we find that the dispersion equation is

$$\beta_g^2 = \frac{L_2}{L_1}\left(\omega^2 L_1 C - 1\right). \tag{2.32}$$

Comparing (2.32) with (2.11) shows that

$$\beta_w^2 = \omega^2 L_2 C \tag{2.33}$$

and

$$\beta_C^2 = \frac{L_2}{L_1} \tag{2.34}$$

so that the cut-off frequency is

$$\omega_C = \frac{1}{\sqrt{L_1 C}}. \tag{2.35}$$

The iterative impedance of the network is

$$Z_g = \frac{\beta_w}{\beta_g}\sqrt{\frac{L_2}{C}}. \tag{2.36}$$

It is important to understand that the component values in this equivalent circuit are not the same as those which would be found from static analysis; though, in some cases, they can be calculated from the dimensions of the waveguide in a similar manner. Moreover, they are not unique but depend upon the way in which the impedance is defined. The properties of the waveguide are completely determined by any two of the parameters since the third is then fixed.

2.2.3 Transverse Magnetic Modes

In transverse magnetic (TM) modes, also known as E modes, the longitudinal component of the magnetic field is zero. The longitudinal electric field is a solution of

$$\nabla_T^2 E_z + \beta_C^2 E_z = 0, \tag{2.37}$$

subject to the condition that $E_z = 0$ on the conducting boundaries. The full solution for each mode can then be obtained as before. Figure 2.6 shows examples of the field patterns of *TM* modes in a rectangular waveguide. The tangential electric field and the normal magnetic field are zero at the walls to satisfy the boundary conditions. The tangential magnetic field at the wall only has transverse components and the conduction current in the walls is longitudinal. The wave impedance for *TM* modes can be derived in the same way as for the *TE* modes with the result

$$Z_{TM} = \left| \frac{E_T}{H_T} \right| = \frac{\beta_g}{\omega \varepsilon} = \frac{\beta_g}{\beta_w} Z_w. \qquad (2.38)$$

As in the case of *TE* waves, $Z_{TM} \to Z_w$ as $\omega \to \infty$. We also note from (2.27) and (2.38) that

$$Z_{TE} Z_{TM} = Z_w^2. \qquad (2.39)$$

The equivalent circuit of the TM_{11} mode is shown in Figure 2.7. It should be noted that the lower line represents the wall of the waveguide and the upper line the centre of the waveguide. From analysis of this circuit the dispersion equation is

$$\beta_g^2 = \frac{C_1}{C_2} \left(\omega^2 L C_2 - 1 \right). \qquad (2.40)$$

Comparing this equation with (2.11) we see that

$$\beta_w^2 = \omega^2 L C_1 \qquad (2.41)$$

and

$$\beta_C^2 = \frac{C_1}{C_2} \qquad (2.42)$$

so that the cut-off frequency is

$$\omega_C = \frac{1}{\sqrt{L C_2}}. \qquad (2.43)$$

The iterative impedance is

$$Z_g = \frac{\beta_g}{\beta_w} \sqrt{\frac{L}{C_1}}. \qquad (2.44)$$

In this case the potential difference between opposite sides of the guide is zero and it is necessary to define the potential between the wall and the axis. But, because the axial current flow in the walls is non-zero, it is more natural to choose Z_{PI} as the definition of characteristic impedance.

TM MODES IN RECTANGULAR WAVEGUIDE

$$\lambda_c = \frac{2a}{\sqrt{1 + (a/b)^2}}$$

TM$_{11}$

$$\lambda_c = \frac{a}{\sqrt{1 + (a/2b)^2}}$$

TM$_{21}$

$$\lambda_c = \frac{a}{\sqrt{1 + (a/b)^2}}$$

TM$_{22}$

1. **Cross-sectional view**
2. **Longitudinal view**
3. **Surface view**

a. **Inside Broad Dimension**
b. **Inside Narrow Dimension**

$\begin{matrix} -- & I \\ = & E \\ \cdots & H \end{matrix}$
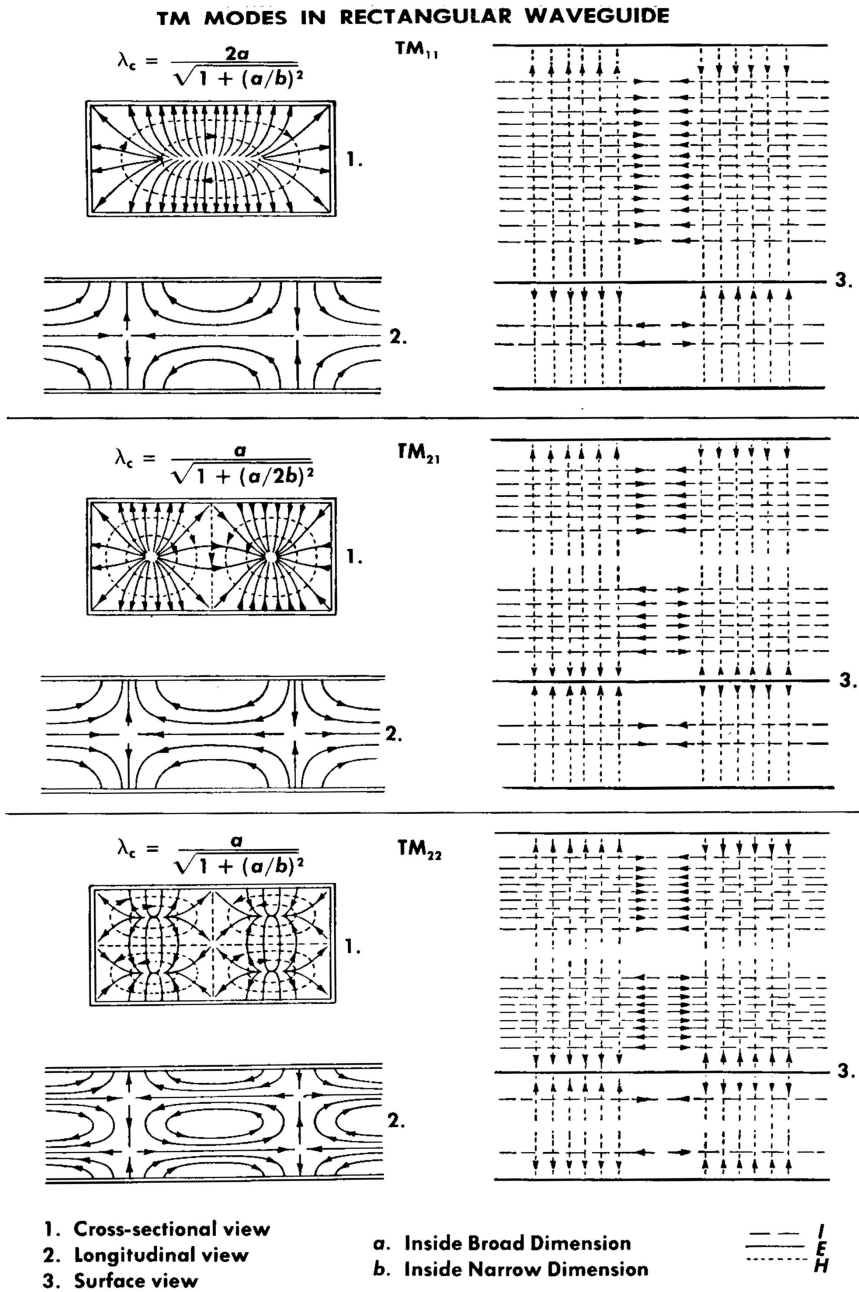
Figure 2.6: Field patterns of TM modes in a rectangular waveguide (reproduced, with permission, from [5]).
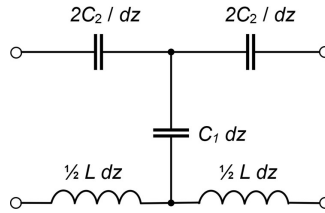
Figure 2.7: Equivalent circuit of the TM_{11} mode in a rectangular waveguide.

2.3   Practical Waveguides

Waveguides of all kinds are normally operated in the frequency band between the lowest two cut-off frequencies so that only one mode can propagate. When, exceptionally, the frequency is such that there are two, or more, modes which are not cut-off, then the waveguide is said to be *overmoded*. We shall see that the effect of any discontinuity in a waveguide is to scatter power into modes other than the one which has been excited. It follows that, when an overmoded guide is used it is necessary to take great care to avoid this possibility.

The maximum power which can be handled by a waveguide is determined by electric breakdown in air, or dielectric materials, and by multipactor discharges in vacuum (see Section 18.8). Air-filled waveguides are normally operated under a small positive pressure of dry air, or dry nitrogen, to prevent condensation and to exclude dust, both of which lead to breakdown at lower powers [6, 7]. The power handling capability outside the vacuum envelope can be increased by pressurisation of the waveguide (see Section 18.7.2). Increasing the pressure requires the waveguide to be gas-tight and reinforced to avoid distortion by the gas pressure. It may also require pressure vessel rules to be followed. The breakdown voltage can also be increased by filling the waveguide with an electron attaching gas such as sulphur hexafluoride (SF_6). This gas is chemically inert but when it is decomposed by an electrical discharge the products are corrosive and highly toxic [8]. The type and size of the waveguide external to the tube is generally specified by the system in which it is installed. Dielectric waveguide windows are required to separate sections of waveguide at different pressures, especially between the vacuum inside a tube and the atmospheric, or higher, pressure outside and these are discussed in Section 2.8. Knowledge of the properties of waveguides is important for understanding the properties of cavity resonators and slow-wave structures (see Chapters 3 and 4). It is also necessary for the design of the input and output coupling structures of tubes. We shall assume that waveguides are evacuated or air-filled unless otherwise stated.

2.3.1   Coaxial Lines

The geometry of a coaxial line is shown in Figure 2.8(a). The space between the conductors may be filled with dielectric material to locate and support the inner
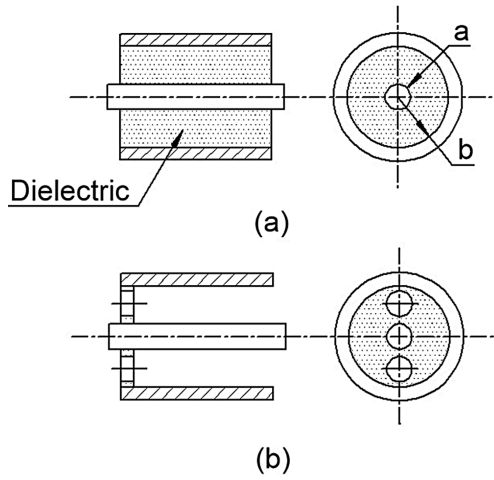
Figure 2.8: Arrangement of a coaxial line with (a) a uniform dielectric, and (b) dielectric spacers.

conductor. For high powers it is usual to use air- or vacuum-spaced lines to mini-mise the dielectric losses. The inner conductor is then supported by dielectric spacers as shown in Figure 2.8(b) [9]. Standard coaxial cables in small sizes are commonly used for the RF input connections of high power amplifiers. The high power output connection is frequently a coaxial line, either directly, or as a means of coupling RF power to some other kind of waveguide.

The lowest mode of a coaxial transmission line is the *TEM* mode in which the electric field is radial and the magnetic field azimuthal, as shown in Figure 2.2(a). Since this field pattern must be the same at all frequencies it can be derived by static analysis [10]. The results are

$$E_r = \frac{V_0}{\ln(b/a)} \cdot \frac{1}{r} \cdot \exp\left[j\left(\omega t - \beta_g z\right)\right], \tag{2.45}$$

where V_0 is the magnitude of the RF potential difference between the inner and outer conductors, and

$$H_\theta = \sqrt{\frac{\varepsilon}{\mu}} E_r. \tag{2.46}$$

The capacitance per unit length is

$$C = \frac{2\pi\varepsilon}{\ln(b/a)} \tag{2.47}$$

and the inductance per unit length is

$$L = \frac{\mu}{2\pi} \cdot \ln(b/a). \tag{2.48}$$

Then the phase velocity is

$$v_p = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\varepsilon\mu}} \qquad (2.49)$$

and the characteristic impedance is

$$Z_c = \sqrt{\frac{L}{C}} = \frac{1}{2\pi}\sqrt{\frac{\mu}{\varepsilon}}\ln(b/a). \qquad (2.50)$$

Coaxial lines for high-power transmission are available with 50 Ω and 75 Ω characteristic impedances. The former represents a compromise between breakdown field strength and power handling capacity, and the latter is chosen for minimum attenuation. The ratio b/a is fixed by the characteristic impedance of the line and the relative permittivity of the medium between the conductors. It is easy to see from the equivalent circuit and equations (2.49) and (2.50) what the effects of changes in the dimensions and dielectric material will be.

The wave on an air-spaced coaxial line loses power through conduction losses in the conductors so that its amplitude decays as $\exp(-\alpha z)$ where α is the loss parameter is given by

$$\alpha = \frac{1}{2Z_0\ln(b/a)}\left(\frac{R_a}{a} + \frac{R_b}{b}\right). \qquad (2.51)$$

Here R_a and R_b are the surface resistances of the inner and outer conductors respectively (see Section 3.3) and $Z_0 = \sqrt{\mu_0/\varepsilon_0}$ is the wave impedance of free space [6]. It can be seen that the greater part of the loss is contributed by the inner conductor because the current density is higher there than on the outer conductor. One of the problems with the use of coaxial lines at high power levels is the need to remove heat from the inner conductor. Large coaxial lines for high-power RF transmission may use copper for the inner conductor and aluminium for the outer conductor to save weight. The use of aluminium as the outer conductor increases the transmission losses by about 10%. The average power being carried on a coaxial line is given by

$$P = \frac{1}{2}\frac{V_0^2}{Z_c}. \qquad (2.52)$$

Now from (2.45) the magnitude of the RF electric field is

$$|E_r| = \frac{V_0}{\ln(b/a)}\cdot\frac{1}{r}, \qquad (2.53)$$

which is maximum when $r = a$. Substituting in (2.52) from (2.50) and (2.53) gives the relationship between the power and the maximum electric field

$$P_{max} = \frac{\pi a^2 E_a^2}{Z_w}\ln(b/a). \qquad (2.54)$$

It is customary to assume that the peak electric field in an air-filled line is the breakdown electric field strength in dry air at standard pressure and temperature (3 MV/m). It must be remembered that this figure provides no margin of safety, and that allowances must be made for changes with altitude, humidity, and the presence of dust particles in the air. For this reason the breakdown voltage is often an order of magnitude, and the maximum power two orders of magnitude, below the theoretical limit [11].

When the line is evacuated, breakdown is caused by a multipactor discharge. It should be noted that this does not apply only within the vacuum envelope of the tube but is also the cause of breakdown in waveguides used in space [12]. The maximum power can be increased by the use of anti-multipactor coatings and by the application of a DC bias between the inner and outer conductors [13, 14]. Multipactor discharges are discussed in Section 18.8.

Higher-order *TE* and *TM* modes can propagate in a coaxial line at higher frequencies. In general they are to be avoided because mode conversion from the *TEM* mode to *TE* or *TM* modes represents a source of power loss. The cut-off frequencies for these modes are found from (2.26) and (2.37). In polar co-ordinates the transverse vector operator is

$$\nabla_T^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}. \tag{2.55}$$

For TE modes the variables can be separated by writing

$$H_z(r,\theta) = R(r)\cos(m\theta), \tag{2.56}$$

where $m = 0,1,2\dots$. Substituting in (2.26) we find that

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial R}{\partial r}\right) + \left(\beta_C^2 - \frac{m^2}{r^2}\right)R = 0. \tag{2.57}$$

This form of Bessel's equation has the general solution

$$R = AJ_m(\beta_C r) + BY_m(\beta_C r), \tag{2.58}$$

where J_m and Y_m are Bessel functions of the first and second kinds [15]. The value of β_C is found by applying the boundary conditions $\partial R/\partial r = 0$ when $r = a$ and $r = b$ so that

$$J_m{}'(\beta_C a)Y_m{}'(\beta_C b) - J_m{}'(\beta_C b)Y_m{}'(\beta_C a) = 0, \tag{2.59}$$

where the prime represents differentiation with respect to the argument. This equation has a doubly infinite set of solutions $\beta_C = \beta_{m,n}$ where $n = 1,2,\dots$ is the number of zeroes of R between the conductors. The corresponding modes are denoted by TE_{mn}. Figures 2.9 and 2.10 show examples of the field patterns of *TE* and *TM* modes in a coaxial line. The lowest mode is the TE_{11} mode whose cut-off wavelength is given very nearly by

$$\lambda_C = \frac{2\pi}{\beta_C} = \pi(a+b) \tag{2.60}$$

TE MODES IN COAXIAL WAVEGUIDE

TE$_{11}$

1.

$$\lambda_c = 1.873 \frac{\pi}{2} (b + a); \; a = 3b$$

3.

TE$_{21}$

1.

$$\lambda_c = 1.023 \frac{\pi}{2} (b + a); \; a = 3b$$

3.

TE$_{31}$

1.

$$\lambda_c = 0.721 \frac{\pi}{2} (b + a); \; a = 3b$$

3.

1. Cross-sectional view
2. Longitudinal view through plane *I-I*
3. Surface view from s-s

a. Inside Radius of Outer Conductor
b. Outside Radius of Inner Conductor

— — — I
——— E
------ -- H
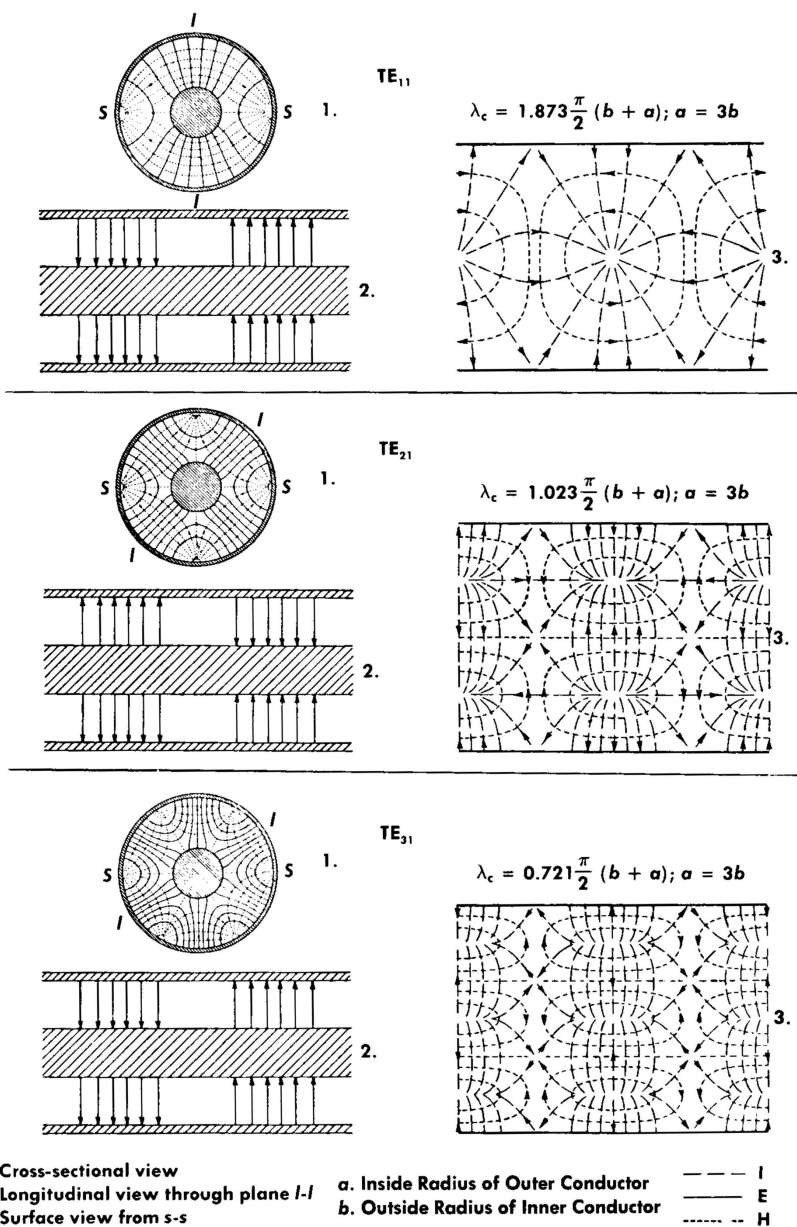
Figure 2.9: Field patterns of *TE* modes is coaxial waveguide (reproduced, with permission, from [5]).

TM MODES IN COAXIAL WAVEGUIDE

TM_{01}

1.

$\lambda_c = 2.029\,(a - b),\ a = 3b$

2.

3.

TM_{11}

1.

$\lambda_c = 1.920\,(a - b);\ a = 3b$

2.

3.

TM_{21}

1.

$\lambda_c = 1.680\,(a - b);\ a = 3b$

2.

3.

1. Cross-sectional view
2. Longitudinal view through plane *l-l*
3. Surface view from s-s

a. Inside Radius of Outer Conductor
b. Outside Radius of Inner Conductor

— — I
——— E
------ H
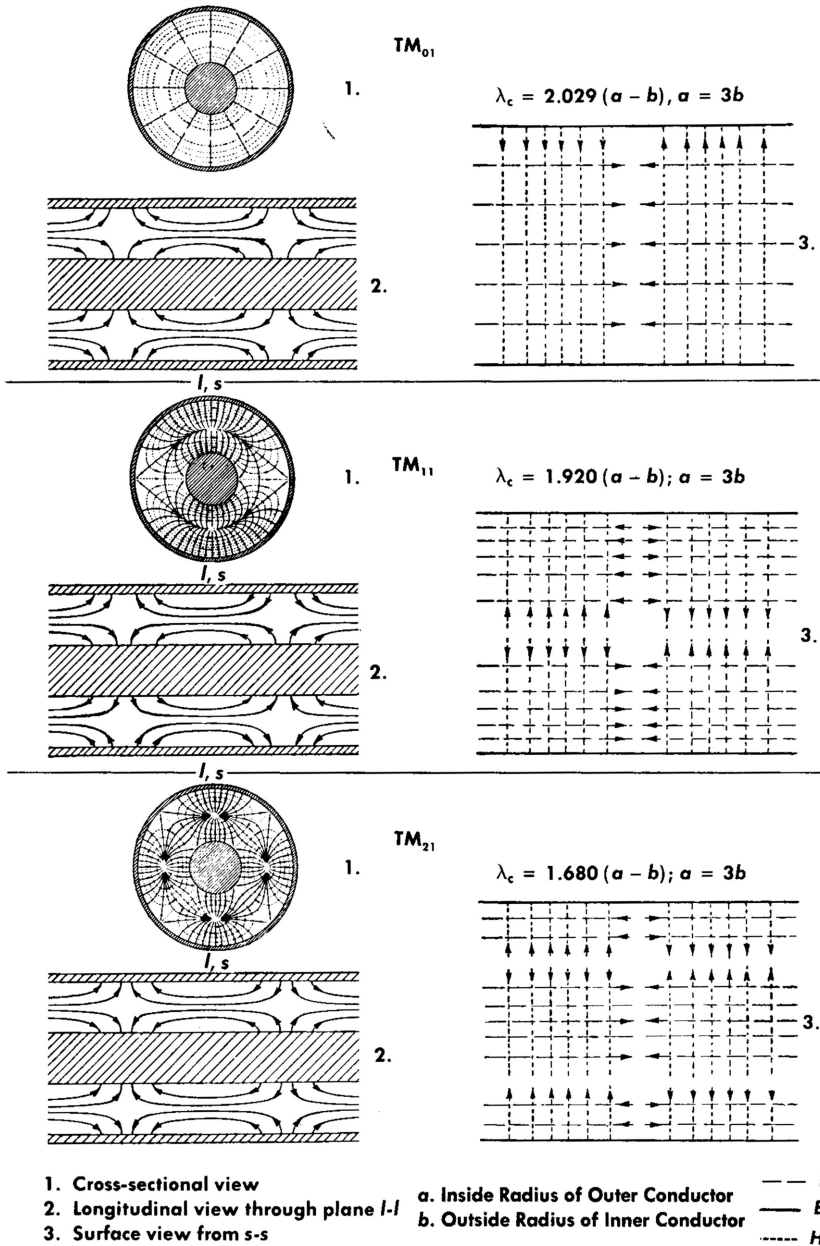
Figure 2.10: Field patterns of *TM* modes is coaxial waveguide
(reproduced, with permission, from [5]).

over the range $1 < a/b < 4$. Thus cut-off occurs when there is a standing wave in the azimuthal direction. The cut-off wavelengths for other *TE* modes and for *TM* modes can be found from (2.59). Detailed information about the higher-order modes in a coaxial line, including diagrams of the field patterns and the equations for the field components, can be found in [1, 5].

2.3.2   Rectangular Waveguides

Rectangular waveguides are the usually used for transmission of RF power at 300 MHz and above where their power handling and attenuation are better than those of coaxial lines [6]. Figure 2.3 shows the arrangement of a rectangular waveguide in which a and b are the internal width and height. Standard rectangular waveguides have aspect ratios that are close to 2:1 but reduced height waveguides are sometimes used for special purposes.

The cut-off frequencies for the modes are found from (2.26) and (2.37). In rectangular Cartesian co-ordinates the transverse vector operator is

$$\nabla_T^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{2.61}$$

For TE modes the variables can be separated by writing

$$H_z(x, y) = H_{0z} \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right), \tag{2.62}$$

where $m, n = 0, 1, 2 \ldots$. It should be noted that this choice automatically satisfies the boundary conditions on the walls of the waveguide. Substituting in (2.26) we find that

$$\beta_C^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2. \tag{2.63}$$

The modes are denoted by TE_{mn} and the lowest mode when $a > b$ is the TE_{10} mode in which the electric field is normal to the broad wall and the magnetic field is parallel to the broad wall as shown in Figure 2.4. This field pattern can be generated by superimposing plane *TEM* waves whose directions of propagation are at equal positive and negative angles with respect to the z axis [16]. At cut-off this angle is $\pi/2$ so that there is a standing wave across the guide and no power propagates along it. The equations for the components of the electric and magnetic fields can be found by substituting (2.62) in (2.19) and (2.20). Table 2.1 shows the cut-off frequencies of the lowest modes of a waveguide for which $a = 2b$, normalised to the frequency of the TE_{10} mode.

The power flow in the guide is found by integrating the z component of the Poynting vector across a transverse plane with the result

$$P = \frac{ab}{4} \cdot \frac{\beta_g}{\beta_w} \cdot \frac{|E_x|^2}{Z_w}, \tag{2.64}$$

Table 2.1: Relative cut-off frequencies of *TE* modes in a rectangular waveguide ($a = 2b$)

m	n	f_c
1	0	1.000
0	1	2.000
2	0	2.000
1	1	2.236
2	1	2.828
3	0	3.000
3	1	3.606
0	2	4.000

where $|E_x|$ is the magnitude of the electric field at the centre of the guide. The characteristic impedance is conveniently defined by (2.29) where $V = bE_x$. Substituting for P from (2.64) gives

$$Z_{PV} = \frac{2b}{a} \cdot \frac{\beta_w}{\beta_g} Z_w, \tag{2.65}$$

which can be seen to have the same form as (2.36). The impedance of an empty waveguide having a 2:1 aspect ratio ranges from around 630Ω at the bottom of the useful frequency band to 440Ω at the top. Assuming that $b = a/2$ the attenuation constant of an empty rectangular waveguide is found to be [1]

$$\alpha = \frac{2R_s \left[1 + (\omega_C / \omega)^2 \right]}{a\sqrt{\mu_0 / \varepsilon_0} \left[1 - (\omega_C / \omega)^2 \right]^{0.5}}, \tag{2.66}$$

where R_s is the surface resistance (see Section 3.3). The theoretical maximum power in an air, or gas, filled waveguide is

$$P_{\max} = \frac{E_{\max}^2 ab}{4\sqrt{\mu_0 / \varepsilon_0}} \sqrt{1 - (\omega_C / \omega)^2}. \tag{2.67}$$

where $E_{\max}$ is the breakdown field.

The values of the components in the equivalent circuit of the TE_{10} mode of an empty waveguide which are consistent with the definition of impedance in terms of power and voltage (Z_{PV}) can be found by comparing (2.65) with (2.36). The result is

$$\sqrt{\frac{L_2}{C}} = \frac{2b}{a} \sqrt{\frac{\mu_0}{\varepsilon_0}}. \tag{2.68}$$

Then from (2.33), (2.34) and (2.68)

$$C = \frac{a}{2b} \varepsilon_0 \tag{2.69}$$

$$L_1 = \frac{2ab}{\pi^2}\mu_0 \qquad\qquad (2.70)$$

$$L_2 = \frac{2b}{a}\mu_0. \qquad\qquad (2.71)$$

We recall that it is not possible to determine these component values using static field analysis as in the case of the *TEM* mode. Instead they are chosen so that the network has the correct values of the cut-off frequency, the high-frequency asymptote for the phase velocity and the characteristic impedance [16].

Higher-order waveguide modes determine the useful bandwidth of rectangular waveguide. For the conventional choice of a 2 to 1 aspect ratio in the transverse dimensions, the first higher-order mode is the TE_{20} mode (H_{20} mode), for which the cut-off frequency is just twice the cut-off frequency of the TE_{10} mode (see Table 2.1). The useable frequency band of a rectangular waveguide is usually taken to be from $1.25\,f_c$ to $1.90\,f_c$. Reference to Figure 2.1 shows that below the minimum frequency the dispersion of the waveguide increases rapidly. At the maximum frequency the attenuation of the next higher-order mode is found from (2.17) to be 18 dB in one free-space wavelength. The wavelength of the propagating TE_{10} mode at a frequency $1.90\,f_c$ is 1.18 times the free-space wavelength so that the attenuation of the TE_{20} mode in one guide wavelength of the TE_{10} mode is 21 dB. Thus any higher-order mode fields generated by discontinuities in the waveguide are of short range. The equations for the field components of the *TE* modes are given in [1, 5]. Figure 2.4 shows examples of the field patterns *TE* modes.

The cut-off frequencies for the *TM* modes can be determined from (2.37) in a similar manner by assuming that

$$E_z(x,y) = E_0 \sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right). \qquad\qquad (2.72)$$

This choice ensures that the boundary conditions are satisfied. The cut-off condition is again given by (2.63) but we note that modes in which either m or n is zero cannot exist because then $E_z = 0$ everywhere. Thus the lowest *TM* mode is the TM_{11} mode. The equations for the field components of these modes are to be found in [1, 5]. Figure 2.6 shows examples of the field patterns of *TM* modes. The dimensions and properties of standard rectangular waveguides can be found in reference books [4] and are also available from manufacturers [17].

2.3.3 Ridged Waveguides

In the previous section we saw that the useful bandwidth of a rectangular waveguide is limited to around 1.5:1. For some purposes this is inconvenient. In order to increase the useful bandwidth it is necessary to increase the separation between the cut-off frequencies of the TE_{10} and TE_{20} modes. The equivalent circuit shown in Figure 2.5

suggests a possible way. If a ridge is added to the centre of one, or both, of the broad walls of the waveguide, as shown in Figure 2.11, then the effect on the TE_{10} mode is to increase the capacitance C. As a result, the cut-off frequency of this mode is depressed. The TE_{20} mode, however, is little affected because the ridges are in regions of weak electric field for that mode. The price paid for the increase in bandwidth is a reduction in power handling capability because of the concentration of electric field produced by the ridges. At low power levels coaxial lines provide a more compact means of broad-band signal transmission, but ridged waveguides are useful where a combination of broad-bandwidth and moderately high power is required.

Because the shunt capacitance is increased by the addition of the ridges the characteristic impedance is reduced, as can be seen from (2.36). This is useful for matching between coaxial lines and waveguides because the characteristic impedance is intermediate between those of the other two types of waveguide. The notation for modes in ridged waveguide is the same as that used for rectangular waveguides [2, 4, 18, 19]. Detailed information about ridged waveguides is to be found in [18]. The dimensions and other properties of standard WRD series double-ridged waveguides are available from manufacturers [20].
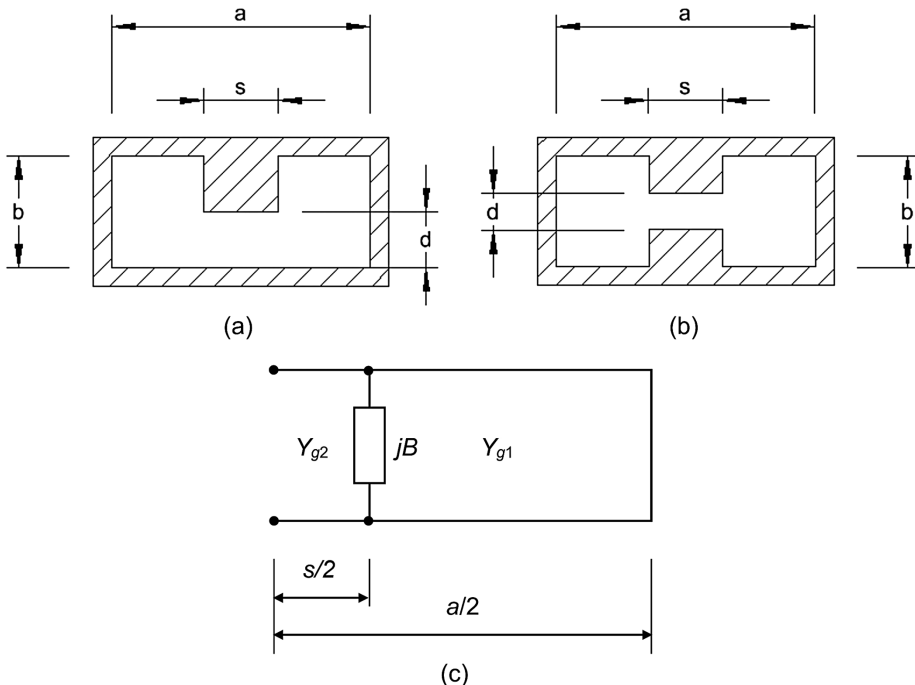
Figure 2.11: (a) Single ridged waveguide, (b) double ridged waveguide and (c) the equivalent circuit.

The cut-off frequency of the TE_{10} mode of a ridged waveguide can be calculated by recalling that, at cut-off, there is a standing *TEM* wave across the guide [19]. The problem may therefore be considered as comprising two parallel plate transmission lines having differing characteristic impedances as shown in Figure 2.11(c). It is easy to show that the capacitance and inductance per unit length of a parallel plate transmission line having unit width and separation b are

$$C = \frac{\varepsilon_0}{b} \tag{2.73}$$

and

$$L = \mu_0 b, \tag{2.74}$$

where fringing fields have been ignored. Thus the phase velocity and characteristic admittance are

$$v_p = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} = c \tag{2.75}$$

and

$$Y_g = \frac{1}{b}\sqrt{\frac{\varepsilon_0}{\mu_0}}. \tag{2.76}$$

The admittance presented at the step by the outer region of the guide is found using the equation for the transformation of admittance on a transmission line

$$Y_1 = \frac{Y_L + jY_{g1}\tan\beta_w z}{Y_{g1} + jY_L \tan\beta_w z} \cdot Y_{g1}, \tag{2.77}$$

where $\beta_w = \omega/c$. Setting $z = (a-s)/2$ and letting the terminating admittance Y_L tend to infinity gives

$$Y_1 = -j\frac{w}{b}\sqrt{\frac{\varepsilon_0}{\mu_0}}\cot\left(\beta_w(a-s)/2\right). \tag{2.78}$$

The admittance presented at the step by the inner region of the guide is found, similarly, by setting $z = s/2$ and $Y_L = 0$

$$Y_2 = j\frac{w}{d}\sqrt{\frac{\varepsilon_0}{\mu_0}}\tan\left(\beta_w s/2\right). \tag{2.79}$$

The effect of the fringing electric field at the step is represented in the equivalent circuit by the lumped susceptance B. If the height of the waveguide is small compared with the free-space wavelength, and the fringing field of the step does not extend to the outer wall of the waveguide, the determination of B may be treated as a quasi-static problem. The field between a pair of parallel plates

with a step can be determined by conformal mapping, with the result shown in Figure 2.12. The fringing field can be represented by a lumped capacitance. This is the difference between the capacitance calculated from the field pattern shown and that calculated if it is assumed that all the field lines go straight across from one plate to the other. The resulting shunt susceptance for a double-ridged guide is [21]:

$$B = \frac{\beta_w}{2\pi}\sqrt{\frac{\varepsilon_0}{\mu_0}} \cdot \left[\left(\frac{d}{b}+\frac{b}{d}\right)\ln\left(\frac{1+d/b}{1-d/b}\right) \quad - \quad 2\ln\left(\frac{4d/b}{1-d^2/b^2}\right)\right]. \tag{2.80}$$

At cut-off β_C is the value of β_w for which

$$Y_1 + Y_2 + jB = 0. \tag{2.81}$$

Figure 2.13 shows a plot of λ_c/a against s/a with d/b as a parameter when $b/a = 0.5$. A single-ridged guide can be considered as the part of a double-ridged guide above the mid plane. The normalised cut-off wavelength is therefore the same as that of a double-ridged guide of double the height. The same result can be obtained by doubling the susceptance given by (2.80).

Hopfer calculated the characteristic impedance of a ridged waveguide by first finding the transverse voltage distribution at cut-off, using transmission line theory, and then computing the stored electric energy including that in the fringing field [18]. The power flow in the waveguide was computed by multiplying the stored energy by the group velocity, and the characteristic impedance found from the power flow and the voltage across the centre of the waveguide. More accurate solutions have been
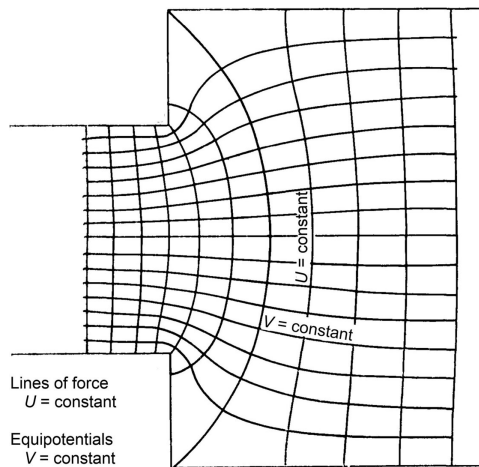
Figure 2.12: Field pattern around a step in a parallel plate transmission line ($d/b = 0.5$) (copyright 1946 IEE; reproduced, with permission, from [21]).
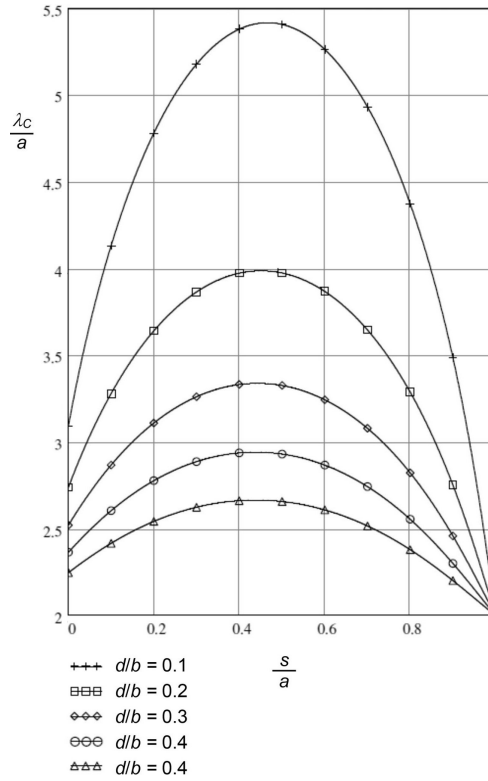
Figure 2.13: Normalised cut-off wavelength of a double-ridged waveguide ($b/a = 0.5$) (copyright 1955 IEEE; reproduced, with permission, from [18]).

obtained using variational methods [22, 23] and finite element methods [24, 25]. Montgomery found that, for the dimensions he examined, Hopfer's value of the cut-off wavelength was accurate to better than 1%, and his figure for the impedance accurate to better than 5%. Approximate closed form expressions for the parameters of ridged waveguides are given in [26]. The properties of ridged waveguides can be explored using Worksheet 2.1.

2.3.4   Circular Waveguides

The fourth type of waveguide which is important in vacuum tube engineering is the circular waveguide. The field solutions are obtained using cylindrical polar coordinates so that, for *TE* modes, the transverse variation of the axial magnetic field is given by (2.56) in the same manner as for a coaxial line. The radial variation is given by (2.58) but with the difference that the coefficient B must be zero because the Bessel function of the second kind tends to infinity on the axis. Thus

$$R = AJ_m\left(\beta_C r\right). \tag{2.82}$$

The radial boundary condition requires that

$$J_m'\left(\beta_C a\right) = 0, \tag{2.83}$$

where a is the internal radius of the waveguide. The solutions to this equation are given in [5]. The notation used for the modes is the same as that in coaxial lines. Figure 2.14 shows the field patterns for examples of TE modes [5]. The mode with the lowest cut-off frequency in a circular waveguide is the $TE_{11}\left(H_{11}\right)$ mode. This mode resembles the TE_{10} mode in a rectangular waveguide, and it is used in pill-box windows (see Section 2.8.2). The TE modes, especially the higher-order ones, are important because of their use in gyrotrons (see Chapter 17). Modes in which the plane of polarisation is rotating can be formed by combining modes with planes of polarisation at right angles to one another with a phase difference of 90° between them. For example the azimuthal variation of a rotating TE_{11} mode is given by

$$\cos\phi\cos\omega t + \sin\phi\sin\omega t = \cos\left(\phi - \omega t\right), \tag{2.84}$$

so that the direction of the electric field is rotating with angular velocity ω.

The TM modes in circular waveguides are important because the longitudinal component of the electric field is non-zero and can be used to interact with a longitudinal stream of electrons. These modes are the foundation for many of the cavity resonators and slow-wave structures which are discussed in Chapters 3 and 4. These are used in slow-wave tubes such as klystrons and TWTs. The TM modes satisfy (2.37), subject to the boundary condition $E_z = 0$ when $r = a$. This requires that

$$J_m\left(\beta_C a\right) = 0. \tag{2.85}$$

The solutions to this equation are tabulated in [15] and examples of the field patterns are shown in Figure 2.15 [5]. The characteristic impedances of circular waveguide modes can be calculated in the same way as those for rectangular waveguides (see Table 2.1).

2.3.5   Summary of Waveguide Impedances

Table 2.2 shows the characteristic impedances for the lowest modes in some common waveguides.

2.4   Waveguide Discontinuities

Coupling between waveguides involves some kind of discontinuity in the waveguiding structure leading to reflection of microwave power. This section examines the cause of this reflection and considers ways of modelling a few examples of discontinuities in rectangular waveguides. A general discussion of the problem of the junctions between two waveguides propagating TE modes is given by Farmer [27]. It is shown that, if the field patterns in the two guides are similar to one another, it

TE MODES IN CIRCULAR WAVEGUIDE

1. **Cross-sectional view**
2. **Longitudinal view through plane /-/**
3. **Surface view from s-s**

a. **Inside Radius of Outer Conductor**

— — *I*
— — *E*
..... *H*
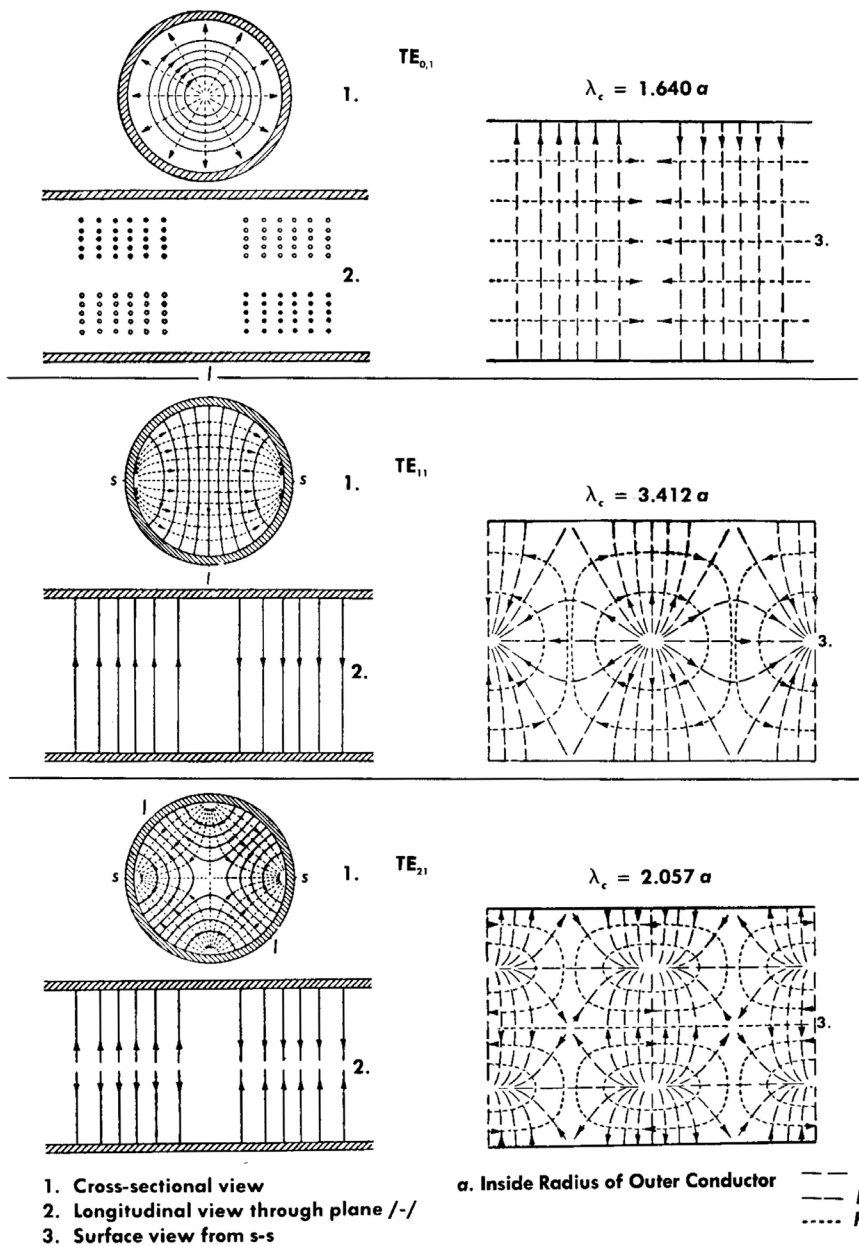
Figure 2.14: Field patterns for *TE* modes in circular waveguide (reproduced, with permission, from [5]).

TM MODES IN CIRCULAR WAVEGUIDE
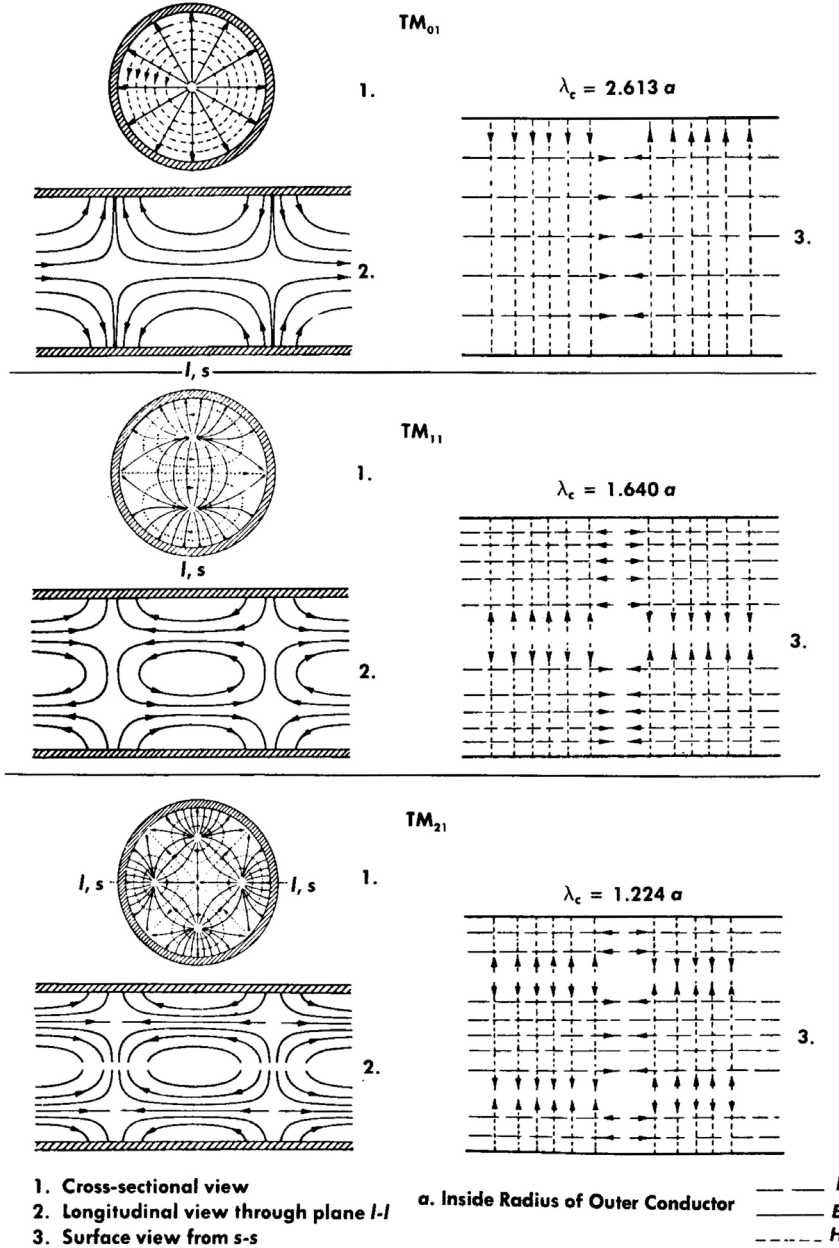
Figure 2.15: Field patterns for *TM* modes in circular waveguide (reproduced, with permission, from [5]).

Table 2.2: Characteristic impedances of waveguide modes

Waveguide	Mode	Z_{PV}	Z_{VI}	Z_{PI}
Coaxial	TEM	$\dfrac{\ln(b/a)}{2\pi}Z_w$	$\dfrac{\ln(b/a)}{2\pi}Z_w$	$\dfrac{\ln(b/a)}{2\pi}Z_w$
Rectangular	TE_{01}	$2\dfrac{b}{a}Z_{TE}$	$\dfrac{\pi}{2}\cdot\dfrac{b}{a}Z_{TE}$	$\dfrac{\pi^2}{8}\cdot\dfrac{b}{a}Z_{TE}$
	TM_{11}			$\dfrac{\pi^2}{64}\left(\dfrac{ab}{a^2+b^2}\right)Z_{TM}$
Circular	TE_{11}	$2.018\,Z_{TE}$	$1.376\,Z_{TE}$	$0.938\,Z_{TE}$
	TM_{01}			$0.0796\,Z_{TM}$

is possible to represent the waveguide junction by an abrupt transition between two transmission lines, together with a lumped susceptance at the transition. In order for this to be possible it is necessary for the characteristic impedances of the two transmission lines to be in the ratio

$$\frac{Z_g{}'}{Z_g}=\frac{\lambda_g{}'}{\lambda_g}\cdot\frac{\iint E_d'^2 dS'}{\iint E_d^2 dS}\cdot\frac{\iint E_d^2 dA}{\iint E_d'^2 dA},\tag{2.86}$$

where the unprimed and primed variables refer to the two waveguides, E_d is the transverse electric field of the dominant mode in a waveguide, S is the cross-section of a waveguide and A is the area of the aperture coupling the two guides. This definition is not necessarily consistent with the ratio calculated using any of the usual definitions of impedance, though it may be so in certain cases.

Figure 2.16 shows a thin iris which partially obstructs the width of a waveguide. At the plane of the iris the fields must satisfy the boundary conditions. The electric field distribution at the plane of the iris is approximately a compressed TE_{10} mode, as shown by the solid line in Figure 2.17(a) with $E_y=0$ outside the gap. The field in the waveguide adjacent to the diaphragm must match that at the diaphragm at every point. The field of the incident TE_{10} wave shown by the solid line in Figure 2.17(b) plainly does not obey this requirement. In order to match the boundary conditions correctly it is necessary for higher-order modes to be excited. Because the diaphragm is symmetrical the next mode which can be excited is the TE_{30} mode shown by the dashed line in Figure 2.17(b). When these two modes are superimposed with amplitudes 0.65 and 0.35 the result is the dashed line in Figure 2.17(b). This is much closer to satisfying the boundary conditions than the original TE_{10} mode. The addition of further higher-order modes (TE_{50}, TE_{70}, etc.) produces a better match, but an infinite series is needed to match the fields exactly at the plane of the diaphragm [28]. Within the normal working band of the waveguide all these higher-order modes are cut off and their fields are confined to a short distance
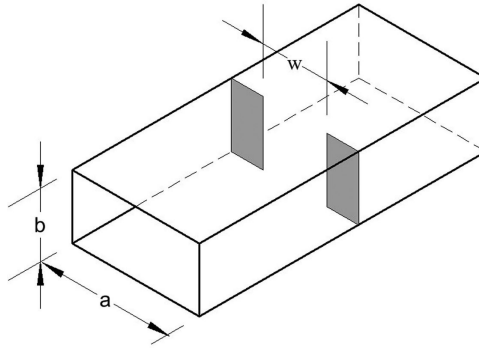
Figure 2.16: Symmetrical inductive iris in a rectangular waveguide.
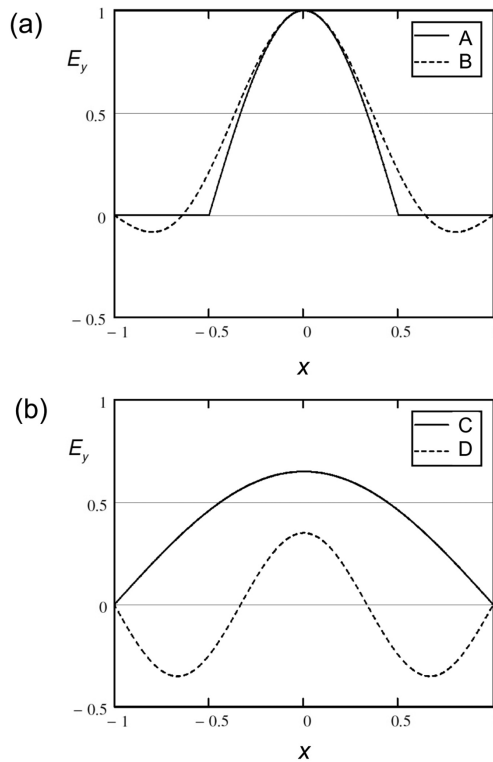
Figure 2.17: Electric field amplitudes at an inductive iris in rectangular waveguide: (a) E_y in the plane of the iris (A) and a two mode approximation (B); (b) transverse variation of E_y for the TE_{10} mode (C) and the TE_{30} mode (D).

either side of the diaphragm. These evanescent modes store energy and their effect can therefore be represented by a lumped reactance. Bearing in mind the earlier discussion of equivalent circuits, we note that the iris has a greater effect on the conduction current than on the displacement current. The discontinuity can therefore be represented by a lumped inductance. It is also possible to create an inductive discontinuity by placing one or two posts between the broad walls of the waveguide.

If the iris changes the height of the guide, instead of its width, as shown in Figure 2.18, a different set of higher-order modes is excited. Once again these are normally cut off. The discontinuity mainly affects the electric field distribution so the shunt reactance is capacitive. An alternative way of introducing a shunt capacitance is by inserting a screw through the broad wall of the waveguide. This arrangement is useful for matching purposes because the capacitance is easily adjusted, but is not suitable for use at high power levels. In fact capacitive obstacles are best avoided in high-power waveguides because they introduce concentrations of electric field and reduce the peak power which can be carried. Formulae and graphs for the susceptances of a wide range of capacitive and inductive discontinuities in rectangular, circular, and coaxial waveguides are given in [5, 21]. These susceptances depend on both the dimensions and the guide wavelength. Although this data is useful for design purposes it has been superseded to some extent by the availability of computer codes for modelling problems in electromagnetics. Additional insight is provided by equivalent circuit models as illustrated in the sections which follow.

2.4.1   Height Step in a Rectangular Waveguide

Figure 2.19 shows a symmetrical height step in a rectangular waveguide. This can be modelled by the junction between two transmission lines having the same propagation constant but differing characteristic impedances, together with a lumped susceptance at the junction. The related problem of a step change in the height of a parallel plate transmission line was considered in Section 2.3.3. The electrostatic field in the vicinity of the junction is shown in Figure 2.12. It can be seen that the field only differs from that in the unperturbed transmission lines over a small region close to the junction. If the higher-order modes excited by a waveguide step are strongly cut off, their fields are, similarly, confined to a small region close to the junction. Therefore, in the first approximation, they can be regarded as quasi-static fields which are solutions to Laplace's equation on the mid-plane of the waveguide. This enables the susceptance of the waveguide step to be deduced directly from (2.80). If the width of the transmission line is a then, from (2.76), the characteristic admittance of the transmission line is

$$Y_g = \frac{a}{b} Y_0 \tag{2.87}$$

and the susceptance of a symmetrical transmission line step can be written

$$B = \frac{\beta_w b}{2\pi} Y_g \cdot \left[\left(\frac{d}{b} + \frac{b}{d} \right) \ln \left(\frac{1 + d/b}{1 - d/b} \right) \quad - \quad 2 \ln \left(\frac{4d/b}{1 - d^2/b^2} \right) \right]. \tag{2.88}$$
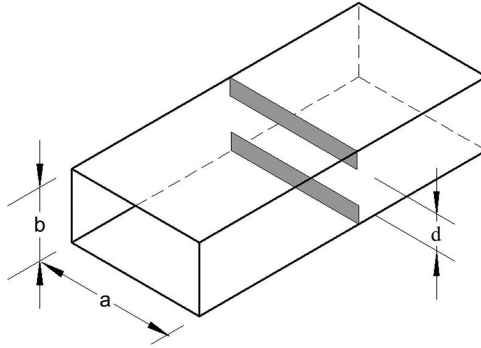
Figure 2.18: Symmetrical capacitive iris in a rectangular waveguide.
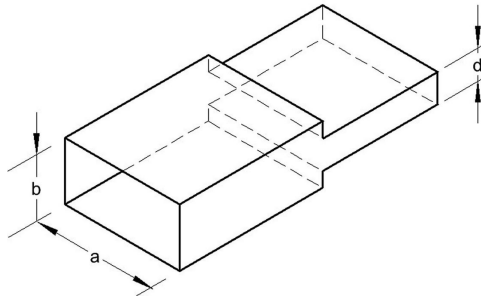
Figure 2.19: Symmetrical height step in a rectangular waveguide.

The susceptance of the symmetrical waveguide step is found by substituting β_g for β_w and Y_{PV} for Y_g [21]. The result, in normalised form, is

$$\frac{B}{Y_{PV}} \cdot \frac{\lambda_g}{b} = \left(\frac{d}{b} + \frac{b}{d}\right) \ln\left(\frac{1+d/b}{1-d/b}\right) \quad - \quad 2\ln\left(\frac{4d/b}{1-d^2/b^2}\right). \qquad (2.89)$$

We note that the expression on the right-hand side of this equation depends only on the dimensions of the step and is independent of frequency. Close to cut-off this expression is exact. A more detailed analysis includes additional terms on the right-hand side which are functions of b/λ_g [5]. However, $b/\lambda_g < 0.4$ throughout the working band of a standard rectangular waveguide, and the error introduced by ignoring this correction is less than 7.5% for all values of d/b. If the change in height is achieved by a step in one broad wall (an asymmetrical step) then the susceptance is twice that given by (2.89).

To develop a simple equivalent circuit model let us suppose that the reactive effects of the higher-order modes excited by the step can be represented by a short length δ of a waveguide having height d where δ depends only upon the geometry of the problem. The step is therefore represented by the equivalent circuit in Figure 2.5 with shunt capacitance C_s and shunt inductance and L_s. Since we expect δ to be

small we will neglect the effect of the series inductance L_2. The susceptance of this network is

$$B = \omega C_s \delta \left(1 - \frac{1}{\omega^2 L_s C_s} \right). \tag{2.90}$$

Both sections of the waveguide have the same width and, therefore, both have the same cut-off frequency. Hence

$$B = \frac{C_s \delta}{\omega} (\omega^2 - \omega_c^2) = c C_s \delta \frac{\beta_g^2}{\beta_w}, \tag{2.91}$$

where $\beta_w = \omega/c$ and c is the velocity of light. The susceptance added by the step is found by subtracting the susceptance of a section of waveguide having the same length for which $d = b$ and shunt capacitance is C so that

$$B_s = c C \delta \left(\frac{C_s}{C} - 1 \right) \frac{\beta_g^2}{\beta_w}. \tag{2.92}$$

It is reasonable to suppose that the capacitances are inversely proportional to the heights of the guides, so that

$$B_s = c C \delta \left(\frac{b}{d} - 1 \right) \frac{\beta_g^2}{\beta_w} \tag{2.93}$$

and, from (2.36), the characteristic admittance of the waveguide can be written

$$Y_g = c C \frac{\beta_g}{\beta_w}. \tag{2.94}$$

Thus, the normalised susceptance of the step is

$$\frac{B_s}{Y_g} \cdot \frac{\lambda_g}{b} = 2\pi \left(\frac{b}{d} - 1 \right) \frac{\delta}{b}. \tag{2.95}$$

This equation has the same form as (2.89) so that δ/b can be calculated as a function of d/b, as shown in Figure 2.20. It can be seen that $\delta/b \leq 0.125$ confirming the initial assumption that this quantity is small. We note that the shunt capacitance and inductance in this model depend only upon the dimensions, so that the dependence of the susceptance on frequency arises from the topology of the network.

2.4.2   Capacitive Iris in a Rectangular Waveguide

The susceptance of the symmetric capacitive iris shown in Figure 2.18 can be deduced in a similar manner. The electrostatic field around a capacitive iris in a parallel plate transmission line, obtained by conformal mapping, is shown in
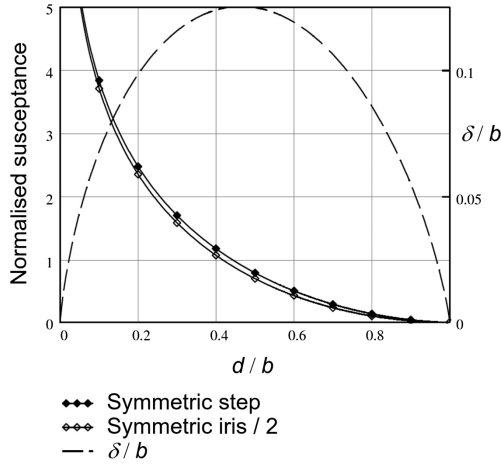
Figure 2.20: Normalised susceptance of a symmetrical capacitive step in a rectangular waveguide (with the susceptance of a symmetrical capacitive iris for comparison), and the normalised length of an equivalent section of reduced height waveguide.
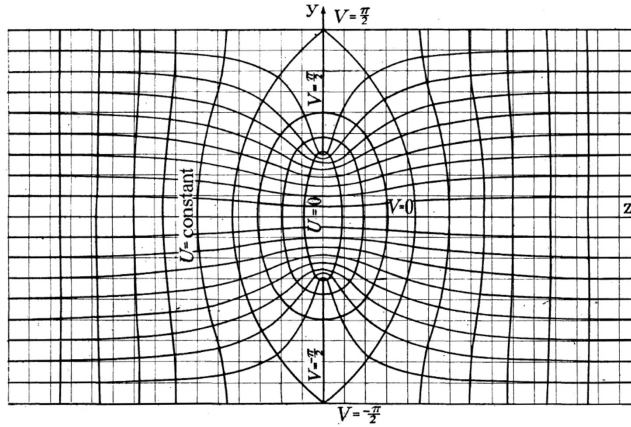
Figure 2.21: Field pattern around a thin iris in a parallel-plate transmission line $(d/b = 0.5)$ (copyright IEE 1946; reproduced, with permission, from [21]).

Figure 2.21. The normalised susceptance of the waveguide iris is given, in the first approximation, by [21]

$$\frac{B_i}{Y_g} \cdot \frac{\lambda_g}{b} = -4\ln\left(\sin\left(\frac{\pi d}{2b}\right)\right). \tag{2.96}$$

Comparison between Figures 2.12 and 2.21 shows that the fringing fields of the step, and of the iris, are very similar to one another. Thus the susceptance of a symmetrical iris is approximately twice that of a symmetrical step having the same height. The comparison between the two results in Figure 2.20 shows that result

for the iris can be used as an approximation to that of a step. More exact formulae for the susceptance of a capacitive iris in a waveguide are given in [5]. The susceptance of an iris having finite thickness can be estimated using the equivalent circuit method in the previous section by adding the thickness of the iris to δ. The properties of a capacitive iris can be explored using Worksheet 2.2.

2.4.3   Inductive Iris in a Rectangular Waveguide

The normalised susceptance of the symmetrical inductive iris shown in Figure 2.16 has been found, to a first approximation, to be [21, 30]

$$\frac{B}{Y_g} \cdot \frac{a}{\lambda_g} = -\cot^2\left(\frac{\pi w}{2a}\right). \tag{2.97}$$

A more accurate expression is given by Marcuvitz [5]. This iris can also be modelled using an equivalent circuit. In this case the shunt susceptance is

$$B = cC_i\delta\frac{\beta_i^2}{\beta_w}, \tag{2.98}$$

where C_i is the capacitance per unit length of a waveguide of height b and width w and $\beta_i = \pi/w$ is the cut-off propagation constant of the iris. As before, we subtract the shunt susceptance of the length δ of the waveguide to give the susceptance added by the iris

$$B_i = cC\delta\frac{\beta_g^2}{\beta_w}\left(\frac{C_i\beta_i^2}{C\beta_g^2} - 1\right). \tag{2.99}$$

Normalising to the characteristic admittance of the waveguide, and assuming that $C_i/C = w/a$, gives

$$\frac{B_i}{Y_g} = \frac{\delta}{\beta_g a}\left[w\left(\beta_w^2 - \frac{\pi^2}{w^2}\right) - a\left(\beta_w^2 - \frac{\pi^2}{a^2}\right)\right]. \tag{2.100}$$

By neglecting the series inductance we have implicitly assumed that the waveguide is close to cut-off so that $\beta_w \to \pi/a$ and

$$\frac{B_i}{Y_g} \cdot \frac{a}{\lambda_g} = \frac{\pi}{2} \cdot \frac{\delta}{a}\left(\frac{w}{a} - \frac{a}{w}\right). \tag{2.101}$$

We note that the susceptance is negative and, therefore, inductive. Now the fraction of the longitudinal current in the waveguide that is intercepted by the aperture is

$$k_a = \sin\left(\frac{\pi w}{2a}\right). \tag{2.102}$$

But the representation of the iris by a shunt susceptance across an equivalent transmission line means that susceptance in (2.101) must be increased by dividing it by k_a. Hence, finally,

$$\frac{B_i}{Y_g} \cdot \frac{a}{\lambda_g} = \frac{\pi}{2} \cdot \frac{\delta}{a}\left(\frac{w}{a} - \frac{a}{w}\right)\operatorname{cosec}\left(\frac{\pi w}{2a}\right). \tag{2.103}$$

The normalised susceptance depends only on the dimensions of the iris and the value of δ can be found by comparison with (2.97). It is found that, to a good approximation

$$\frac{\delta}{a} = \frac{4}{\pi^2}\left(1 - \left(\frac{w}{a}\right)^2\right). \tag{2.104}$$

The properties of an inductive iris can be explored using Worksheet 2.3. The related problem of a step change in the width of a waveguide can be treated in the same way but it is no longer the case that the dominant electric field patterns in the two waveguides are similar and the use of a two-port equivalent circuit model is only approximate [5, 27].

2.5   Matching Techniques

The input and output connections of a vacuum tube amplifier must be well-matched at all frequencies within the operating band of the tube. This is important because reflected power:

- Reduces the useful power output of the tube.
- Leads the periodic variation with frequency in the gain, and output power, because of the effects of multiple reflections.
- Increases the risk of voltage breakdown, and failure of windows, through the increased peak electric field in a standing wave.
- Can cause a tube to become unstable.

It may also be important to consider the effects of mismatches at harmonic frequencies since these can also affect the behaviour of the tube. These issues are discussed further in the chapters dealing with particular types of tube. There are many well-established matching techniques in microwave engineering [3]. The basic methods which are used singly, or in combination, are reviewed briefly below.

2.5.1   Stub Matching

Reactive elements (stubs) can be arranged so that the signal reflected by them more or less cancels out the reflection from the component to be matched. We recall that, when a waveguide having characteristic admittance Y_g is terminated by a load Y_L

the impedance perceived at a plane which is closer to the source by a distance l is given by

$$\frac{Y_L'}{Y_g} = \frac{Y_L + jY_g \tan\left(\beta_g l\right)}{Y_g + jY_L \tan\left(\beta_g l\right)}, \tag{2.105}$$

where β_g is the propagation constant in the waveguide. This equation has been quoted in terms of admittances because the matching stubs normally used in a waveguide present shunt reactances, as in the examples in the previous section. Because the losses in waveguides are small it is usual to treat them as lossless when carrying out matching calculations so that the characteristic impedance is real. It is often useful to visualise the transformation of impedances using a Smith chart, and Vector Network Analysers offer this as a display option.

The simplest matching technique uses (2.105) to find a plane in the waveguide at which the real part of Y_L' is equal to Y_g. The insertion of a stub at that plane, whose susceptance is equal and opposite to the imaginary part of Y_L', cancels the susceptance of the load. This ensures that the combination of load and stub appears matched at all planes which are closer to the source. However, because β_g depends on frequency, the electrical length $\left(\beta_g l\right)$ of the section of waveguide between the load and the stub is not constant. Thus, this technique can only give a good match at a single frequency (or, strictly speaking, the set of frequencies at which $\tan\left(\beta_g l\right)$ has the same value).

2.5.2 Broad-band Matching

To increase the band of frequencies over which there is a good match it is necessary to use a number of stubs at intervals along the waveguide. If the stubs are at least a quarter of a guide wavelength apart the evanescent higher-order modes excited by one stub have negligible effects at a neighbouring stub. It is therefore possible to treat the stubs as lumped reactances. The properties of a system comprising a number of matching elements can be illustrated simply by considering the arrangement shown in Figure 2.22. Four reactive elements are equally spaced on a transmission line with a matched termination. The reflection coefficients are ρ_1 and ρ_2 as shown, and the electrical length of the line between them is ϕ. We will assume that the reflection coefficients are small so that the amplitude of the incident wave is hardly affected by the reflections. Now suppose that D is the obstacle to be matched and A, B and C are the matching elements. The apparent reflection coefficient at A is

$$\begin{aligned}
\rho &= \rho_1 + \rho_2 e^{-2j\phi} + \rho_2 e^{-4j\phi} + \rho_1 e^{-6j\phi} \\
&= 2\rho_1 e^{-3j\phi} \cos 3\phi + 2\rho_2 e^{-3j\phi} \cos 3\phi.
\end{aligned} \tag{2.106}$$

In practical terms it is usually the magnitude of the reflection which is significant, that is

$$|\rho| = \left|8\rho_1 \cos^3 \phi + 2\left(\rho_2 - 3\rho_1\right)\cos \phi\right|. \tag{2.107}$$
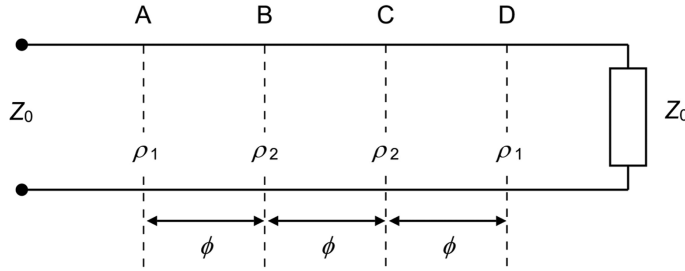
Figure 2.22: Arrangement of four regularly-spaced discontinuities on a transmission line.

which may be written

$$|\rho| = |8\rho_1 x^3 + 2(\rho_2 - 3\rho_1)x|, \tag{2.108}$$

where, for convenience, $x = \cos\phi$. Now ϕ is a function of frequency and, therefore, so is x. Equation (2.108) describes the variation of the reflection coefficient with frequency. Clearly the shape of this cubic curve can be changed by making different choices of ρ_2. One possibility is to set $\rho_2 = 3\rho_1$. The result is

$$|\rho| = |8\rho_1 x^3|, \tag{2.109}$$

shown as the continuous curve in Figure 2.23. The curve is normalised to the reflection coefficient when $x = 1$. A polynomial, in which only the coefficient of the highest power of x is non-zero, has the property that the highest possible number of derivatives of the function is zero at the origin. The curve is therefore as flat as possible close to the origin. A match having this frequency dependence is called a *maximally flat* or Butterworth response. It is the best possible match over a narrow band using a given number of matching elements.

For a given maximum reflection coefficient ρ_m the bandwidth (Δx) is the distance between the two points on the curve at which $\rho = \rho_m$ so that, from (2.109)

$$\Delta x = 2x = 2\left(\frac{\rho_m}{8\rho_1}\right)^{\frac{1}{3}} = \left(\frac{\rho_m}{\rho_1}\right)^{\frac{1}{3}}. \tag{2.110}$$

This shows that, as one would expect, it is possible to achieve a greater bandwidth by accepting a poorer match. For broad-band matching it is better to sacrifice some of the excellence of the match close to the origin in order to gain greater bandwidth. The broken curve in Figure 2.23 shows such a possibility. There is, of course, an infinite number of possible polynomials but of special interest are the Tchebychev[1] polynomials [3, 31]. The first six polynomials are listed in Table 2.3.

---

[1] Other spellings of this name in the Latin alphabet include Chebyshev, Tchebycheff, Teschebyscheff, and Tchebichef.

Table 2.3: Tchebychev polynomials

n	$T_n(x)$
0	1
1	x
2	$2x^2 - 1$
3	$4x^3 - 3x$
4	$8x^4 - 8x^2 + 1$
5	$16x^5 - 20x^3 + 5x$
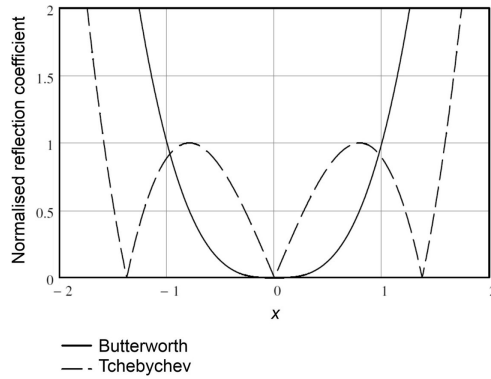
Figure 2.23: Third-order Butterworth (maximally flat) and Tchebychev (equal-ripple) reflection characteristics.

These polynomials have the property that, over the range $-1 < x < 1$, the magnitude is always less than or equal to 1. The Tchebychev polynomial has the broadest bandwidth of all possible polynomials of a given order and specified in-band ripple. To design a broad band match using Tchebychev polynomials we introduce the scaling factors ρ_m and x_m. The third-order polynomial can then be written

$$\left(\frac{\rho}{\rho_m}\right) = 4\left(\frac{x}{x_m}\right)^3 - 3\left(\frac{x}{x_m}\right). \tag{2.111}$$

The response curve has an in-band ripple equal to ρ_m as shown in Figure 2.23. Equating coefficients of x^3 in (2.108) and (2.111) we obtain

$$4\rho_m/x_m^3 = 8\rho_1 \tag{2.112}$$

so that

$$x_m^3 = \frac{\rho_m}{2\rho_1}. \tag{2.113}$$

Now ρ_1 is the reflection coefficient of the termination to be matched and ρ_m is the maximum acceptable in-band match. Therefore, from (2.113), the bandwidth which can be achieved in this way using three matching elements is

$$\Delta x = 2x_m = 2\left(\frac{\rho_m}{2\rho_1}\right)^{\frac{1}{3}} = 2^{\frac{2}{3}}\left(\frac{\rho_m}{\rho_1}\right)^{\frac{1}{3}}. \tag{2.114}$$

Comparing (2.114) with (2.110) we see that, when three matching elements are used, the bandwidth of the Tchebychev match is wider than that of the Butterworth match by a factor of $2^{2/3} = 1.59$.

To complete the design of the Tchebychev matching network the coefficients of x in (2.108) and (2.111) are equated to give

$$2\left(\rho_2 - 3\rho_1\right) = -3\rho_m/x_m \tag{2.115}$$

so that

$$\rho_2 = 3\rho_1 - 3\rho_m/2x_m . \tag{2.116}$$

Since ρ_1, ρ_m and x_m are all known ρ_2 can be calculated.

2.5.3   Stepped Impedance Transformers

Another basic matching technique employs sections of waveguide having different characteristic impedances. If the length of a section of line is one quarter of a guide wavelength $\left(\beta_g l = \pi/2\right)$ then from (2.105) the admittance at the entrance to the line is

$$Y_L' = \frac{Y_g^2}{Y_L}. \tag{2.117}$$

Thus, if the admittance of the load is real, it can be matched to a source whose admittance is also real by using a section of waveguide (a *quarter-wave transformer*) whose characteristic impedance is $Y_g = \sqrt{Y_L Y_L'}$. The method can be extended to match loads which have non-zero susceptance by first using a section of line to transform the admittance to a plane where it is real. The match achieved by a quarter-wave transformer is only exact at frequencies where the length is an odd number of quarter wavelengths. Broad-band matching is achieved by using a number of sections in series to give either a Butterworth or a Tchebychev response. This simple theory ignores the susceptances of the steps. It is possible to compensate for these by making slight changes to the lengths of each section of waveguide. Further information and tables for the design of multi-step transformers can be found in [32–34].

2.6   Coupling without Change of Mode

Microwave circuits involve bends and junctions in the waveguides that introduce discontinuities into the system. Some of these junctions do not involve any change

in the mode of propagation. These are discussed here, and those in which a change of mode is involved are discussed in the next section.

One of the simplest waveguide components is the bend. Provided that the bend is gradual the mismatch is slight and no problems arise. This is normally the case with coaxial lines. Sometimes, however, it is necessary to arrange a change of direction in a small space and the design of the bend then becomes important.

Waveguide bends are often made in the form of circular arcs, as shown in Figure 2.24a. The direction of the bend can either be in the E plane or in the H plane. It has been shown that the minimum radius for a satisfactory match is 1.5 times the width of the waveguide in the plane of the bend [4]. Thus for the E-plane bend shown in Figure 2.24(a) the minimum value of R is $1.5b$. Because very accurate manufacturing is necessary to avoid increased reflections, tight bends are usually electro-formed. More gentle bends can be made by bending a straight waveguide using special equipment.

Radiused waveguide bends are expensive so it is sometimes better to use fabricated mitred bends instead. Figure 2.24(b) shows a simple right-angle bend. Clearly the field patterns within the bend do not match those in the connecting waveguides so an appreciable reactive mismatch can be expected. Equivalent circuits can be developed for right-angled bends in a manner similar to that used above for waveguide discontinuities [5]. An improvement on the simple right-angled bend, shown in Figure 2.24(c), chamfers the outside of the bend to make the effective height of the waveguide the same as that of the connecting waveguides. Better still, if space allows, is to use a double-mitred bend as shown in Figure 2.24(d). If the separation (L) between the two joints is chosen correctly, then the mismatches caused by the junctions cancel each other out at one frequency and produce near cancellation over a useful band of frequencies.
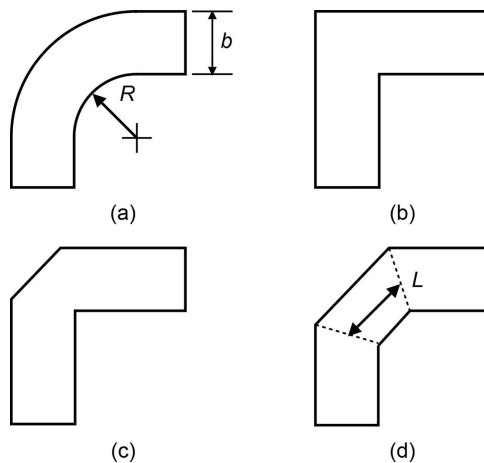
Figure 2.24: E-plane waveguide bends: (a) radiused bend, (b) mitred bend, (c) chamfered bend, and (d) double mitred bend.

It is common in microwave tube engineering to use reduced-height waveguides to make connections into cavity resonators and coupled-cavity slow-wave structures. The most straightforward way to make a transition to a reduced-height waveguide is to use a linear taper [35]. However, the taper must be around three guide wavelengths long to achieve a broad-band match having a return loss greater than 30 dB. An alternative method is to use one, or more, quarter-wave transformers [32, 33, 36, 37]. A very good broad-band match can be achieved in this way, using three or four steps, and such a transformer is considerably shorter than a linear taper having the same performance. To get the best results it is necessary to adjust the lengths of the steps slightly to compensate for their shunt capacitances. Good results can also be obtained using non-uniform transitions [38, 39]. For further information see [3, 4].

2.7   Coupling with Change of Mode

In the previous section we considered examples of coupling between waveguides which were propagating the same mode. Very often it is necessary to couple power from one mode to another. Examples of this are transitions from a coaxial line to a waveguide, and from one waveguide mode to another. The general principle is to match the field patterns of the modes in the two waveguides to one another. In many cases this cannot be done exactly and it is necessary to find some way of matching the field patterns as closely as possible. In that way a high proportion of the power is coupled into the desired mode, and relatively little into other modes. An example of this is the transition from the TE_{10} rectangular waveguide mode to the TM_{01} circular waveguide mode. Figure 2.25(a) shows the field patterns in the two waveguides. It is immediately evident that the magnetic field patterns are similar if the waveguides are arranged at right angles to each other as shown. The electric fields are not so well matched because the removal of part of the broad wall of the rectangular guide means that the field lines have to be radically redistributed. One solution is to put a post at the centre of the transition as shown in Figure 2.25(b). Additional matching elements may be needed to match the transition completely over a band of frequencies. This transition is known as a 'Door-Knob' transition. Further information about mode transducers can be found in [4, 40].

As a second example we will consider the problems of making a good broadband match from a coaxial line to a waveguide. A standard coaxial line has a characteristic impedance of 50 ohms. The characteristic impedance of a rectangular waveguide in terms of voltage and power is given by (2.65). Standard waveguides have an aspect ratio (a/b) of about 2:1 and are used over a range of (λ_g/λ_0) of about 1.7 to 1.2. The impedance to be matched therefore ranges from 630 ohms to 440 ohms, roughly ten times that of the coaxial line. A component providing a well matched connection is called a *coaxial to waveguide transformer*.
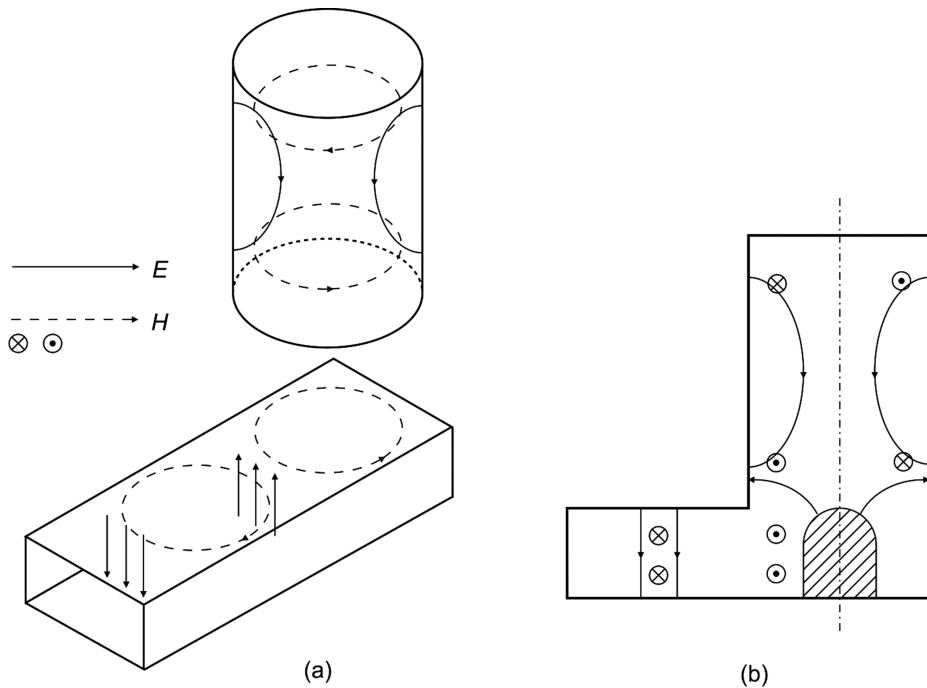
Figure 2.25: TE_{10} rectangular waveguide to TM_{01} circular waveguide junction: (a) arrangements of the fields in the waveguides, and (b) a cross-sectional view of a 'door-knob' transition showing the field patterns.

Consideration of the field patterns in the two waveguides suggests that a door-knob transition, similar to that shown in Figure 2.25(b), might be used. The field patterns are shown in Figure 2.26(a). A quarter wavelength of waveguide behind the transition is used to transform the short circuit at the end of the waveguide into an open circuit at the plane of the transition. Provided that the diameter of the door-knob is not less than $0.15a$ it is possible to neglect its series inductance so that the impedance of the coaxial line is presented directly to the waveguide [41]. Cohn found that this type of junction is best modelled by using Z_{VI} as the characteristic impedance of the waveguide but other authors have used Z_{PI} [42]. One of the difficulties of the door-knob transition is that it has to cope with the change from a low-voltage, high-current wave in the coaxial line to a high-voltage, low-current wave in the waveguide. The existence of a direct current path makes this difficult. Two solutions to this problem suggest themselves: one is to make use of a lower impedance waveguide; the other is to avoid a direct current path.

To reduce the waveguide impedance we must reduce its height. But the reduction to, perhaps, one tenth of the height of a standard waveguide presents new problems for the transformation between the two. A better solution is to use a ridged waveguide which can have an impedance of 50 ohms without needing such narrow gaps. Figure 2.26(b) shows a transition to a coaxial line using a ridged waveguide, with a
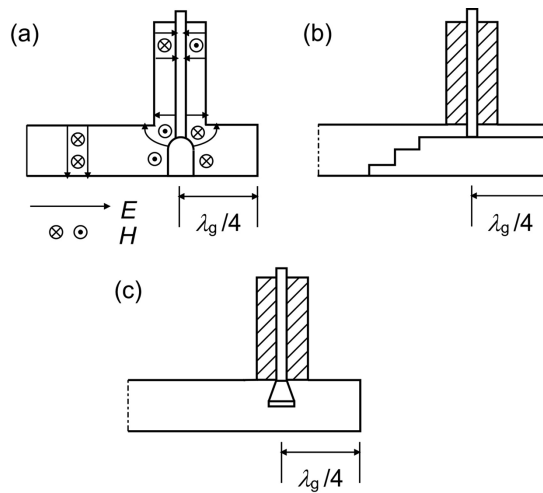
Figure 2.26: Coaxial line to waveguide transformers using: (a) a door-knob, (b) a ridged waveguide, and (c) a probe.

quarter wave section of short circuited waveguide behind it, and a stepped impedance transformer to a standard waveguide [41–43].

The third approach is to couple into the waveguide using displacement current rather than conduction current. Figure 2.26(c) shows how this is achieved. The free end of the centre conductor is, in effect, an antenna radiating into the waveguide. The use of an expanded end to the centre conductor provides the necessary impedance transformation. Many other kinds of transition have been invented, some of which are in common use [4, 40]. The three described here have been included to illustrate the principles involved in coupling between different modes.

2.8   Windows

Dielectric windows form an essential part of any microwave tube because of the need for RF power to pass into, and out of, the vacuum envelope. It is obviously important that they should have low insertion loss and be well-matched at all frequencies of importance. In addition the output window is commonly one of the most critical components in a tube because it is subject to high thermal and electrical stresses. If a window fails then the tube loses its vacuum and is useless [44–47].

Electrical stresses are caused by high RF electric fields in the region of the window, and these may be increased by the presence of a standing wave in the output waveguide. The RF electric fields can also cause single-surface multipactor discharges on the vacuum side of the window (see Section 18.8) [48]. The bombardment by multipacting electrons causes heat dissipation and increased thermal

stresses. A multipactor discharge also generates RF noise and changes the match of the window. It is therefore important to reduce the risk of multipactor when designing the window and it is common to use anti-multipactor coatings such as titanium nitride. Window failure can be caused by a waveguide arc on the air side of the window, either because of the increased standing wave, or because of the direct impact of the arc on the window surface [49]. In some tubes the window may be subject to bombardment by stray electrons from the interaction region, and by X-rays generated within the tube. Even in the absence of any other electrical causes the dielectric material of the window is heated directly by dielectric losses. These dielectric losses commonly increase with temperature, leading to the possibility of thermal runaway. The output window of a high power tube can be protected by monitoring its temperature, or by the use of an arc detector, so that the tube can be switched off rapidly in the event of abnormal heating, or the presence of an arc [47].

The mechanical failure of a window may be caused directly by mechanical stresses, including those caused by thermal expansion, or by dielectric breakdown caused by the build-up of surface charge [44]. The window is subject to mechanical stresses in any case because of the pressure differential across it and these are increased if the external waveguide is pressurised. In addition, the tube must withstand the temperature changes which occur during processing and during use. The coefficients of thermal expansion of the dielectric materials commonly used in windows differ from those of the metals used in tube construction. Thus appreciable thermal stresses can be caused by changes in temperature and the mechanical design of a window is as important as the electrical design [44].

Various dielectric materials have been used in microwave windows. Glass was used in the earliest tubes but has now generally been supplanted by ceramics. The commonest of these is high purity alumina which has low dielectric losses and is readily metallised for brazing to metal components. Beryllia has sometimes been used for windows to carry high average powers because its thermal conductivity is an order of magnitude higher than that of alumina. However, it is a highly toxic material when in powder form and is now generally avoided. A much better material is chemical-vapour-deposited (CVD) diamond which has excellent electrical and mechanical properties. Table 2.4 shows a comparison between the properties of a number of dielectric materials for use in windows [46].

The simplest form of window is a sheet of dielectric material inserted in a waveguide having uniform cross-section, as shown in Figure 2.27. The figure shows a rectangular waveguide for convenience, but the discussion which follows applies to a waveguide of any cross-section. The same mode can propagate in both the empty waveguide and the window, though the cut-off frequency in the window is less than that of the waveguide by a factor of $\sqrt{\varepsilon_r}$ where ε_r is the relative permittivity of the dielectric. Because there is no change of mode, the match of the window can be calculated by using the wave impedances of the waveguide and the window. The window always introduces a mismatch except when it is an integral number of half-wavelengths thick. This condition only holds at specific frequencies, and a half-wavelength window typically has a bandwidth of around 10 to 15% [47,

Table 2.4: Comparison of typical dielectric material properties at 25°C [46, 50]

Property	99.5% Alumina	Beryllia	CVD Diamond
Density (g cm^{-3})	3.90	2.90	3.51
Specific heat (J g^{-1} K^{-1})	0.88	1.02	0.517
Thermal conductivity (W m^{-1} K^{-1})	35	250	1800
Coefficient of thermal expansion	6.9×10^{-6}	6.4×10^{-6}	1.5×10^{-6}
Young's modulus (GPa)	370	344	1220
Poisson ratio	0.22	0.21	0.2
Tensile strength (MPa)	262	75−150	200−500
Relative permittivity	9.4	6.7	5.67
Loss tangent	9×10^{-5}	4×10^{-4}	2×10^{-5}
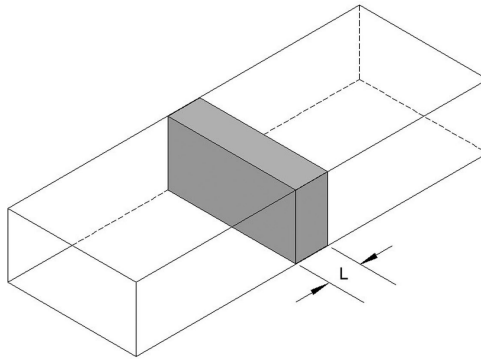
Figure 2.27: A dielectric window in a uniform waveguide.

51]. The bandwidth decreases as the relative permittivity of the dielectric increases. Although this type of window can handle high mean powers it has not been widely used except in gyrotrons. It is useful at millimetre wavelengths where other types of window become too thin to have good mechanical strength [52].

If the window is less than a half-wavelength thick it will always cause a mismatch which becomes smaller as the thickness of the window is reduced. It is therefore desirable to make the window as thin as possible consistent with sufficient mechanical strength [53]. A good match can be achieved by adding reactive elements to provide either a maximally-flat or equal ripple response. The bandwidth of half-wavelength windows can be increased in the same way [52]. Other techniques which have been suggested include the use of a self-matched pair of windows, tapered or stepped dielectrics, and multi-layer dielectrics [50, 54]. The design of broad-band windows is discussed further in the following sections.

Because the phase velocity of electromagnetic waves is reduced in the dielectric it is possible for the cut-off frequencies of higher-order modes to fall within the

working band of the window. These modes are cut-off in the adjacent waveguide and their fields are therefore confined to the region close to the window and are hard to detect. They are therefore known as ghost or trapped modes [55]. In an ideal window they would not be coupled to the propagating mode, but slight asymmetries in the construction of the window, or of the properties of the dielectric, are sufficient to cause coupling of power into them. Ghost modes have high Q factors and it is easy for the fields associated with them to become strong enough to cause failure of the window. Thus it is important to ensure that no ghost modes fall within the working frequency band of the window. They can be divided into even and odd modes depending upon whether their fields are symmetrical or anti-symmetrical about the mid-plane of the window. For windows having the arrangement shown in Figure 2.27 the ghost mode resonances are the solutions of

$$\tan\left(\beta'_{gn}L/2\right) = \left\{ \begin{array}{c} \beta'_{gn}/j\beta_{gn} \\ \varepsilon_r j\beta_{gn}/\beta'_{gn} \end{array} \right\} \tag{2.118}$$

for modes whose tangential electric field is zero at the mid-plane, and

$$\cot\left(\beta'_{gn}L/2\right) = \left\{ \begin{array}{c} \beta'_{gn}/j\beta_{gn} \\ -\varepsilon_r j\beta_{gn}/\beta'_{gn} \end{array} \right\} \tag{2.119}$$

for modes whose tangential magnetic field is zero at the mid-plane. In these equations the upper and lower expressions within the braces refer to *TE* and *TM* modes respectively, the primed variables refer to the region within the dielectric, and the propagation constants for the nth mode are determined from (2.11) [55]. Ghost modes can be a problem in any kind of window but they are most troublesome in half-wavelength windows.

2.8.1   Windows in Coaxial Lines

The characteristic impedances of the *TEM* mode of empty, and dielectric-filled, sections of a coaxial line can be calculated from (2.50). If the impedance of the dielectric-filled section is made equal to that of the air-filled section, the ratio of the diameters of the inner and outer conductors must be increased. Figure 2.28(a) shows a typical arrangement in which the diameter of the inner conductor has been decreased [45, 56]. The window can be modelled using the equivalent circuit, shown in Figure 2.28(b), comprising sections of transmission line in series, together with lumped capacitances to represent the reactances of step changes in the diameter of the inner conductor [57]. In an alternative design the diameters of both conductors differ from those of the connecting coaxial line and inductive chokes are used to improve the match [58].

Windows of the kind illustrated in Figure 2.28(a) have a partial standing wave within the dielectric window because the design has focused on the overall match, rather than the detailed internal fields. The electric field in the window can be reduced by using external reactive elements to match the impedances at the surfaces of the dielectric [59]. It has been shown that the use of a travelling-wave design can reduce the maximum field in the dielectric by a factor of four, compared with

(a)

(b)

Z_1 Z_2 Z_3 Z_2 Z_1
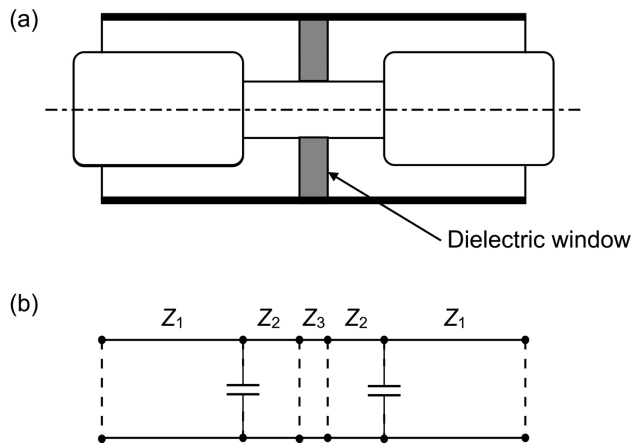
Figure 2.28: A dielectric window in coaxial line: (a) typical arrangement, and (b) equivalent circuit.

a half-wave window. A travelling-wave coaxial window in which the electric field in the dielectric was only 25% of that in the external coaxial line, successfully carried a peak power of 1 MW at 1.3 GHz for several months [60]. After initial conditioning it showed no signs of multipactor discharges.

2.8.2   Windows in Rectangular Waveguide

Although rectangular windows have been used in rectangular waveguide, they suffer from the disadvantage of being difficult to seal and of having stress concentrations at the corners. There is, therefore, a general preference for using circular windows in connection with rectangular waveguides.

The simplest technique is to use a circular window set in a transverse metal wall, as shown in Figure 2.29. The TE_{11} circular waveguide mode is excited in the window and its diameter is chosen so that it is resonant at the desired centre frequency. The diameter of the window can be less than the height of the waveguide if the relative permittivity of the dielectric is great enough. Otherwise the window may take the form of a disk whose top and bottom have been cropped to fit the waveguide. Because this kind of window is resonant it is inherently narrow band, but the bandwidth can be increased by adding reactive elements [61, 62]. When the resonant diameter of the window is greater than the height of the waveguide it is possible to use a circular disk if the matching elements take the form of inductive chokes similar to those used in waveguide choke flanges [44, 45]. It is also possible to design self-resonant windows with a pure travelling wave in the dielectric [63–65].

Self-resonant waveguide windows suffer from the disadvantage that the diameter of the window is set by the resonant condition. Thus it is not possible to reduce the power density in the window by increasing its diameter. This problem is overcome by setting the window in a section of circular waveguide carrying the TE_{11} mode, with mode transformers at either end to connect it to the TE_{10} mode in the rectangular
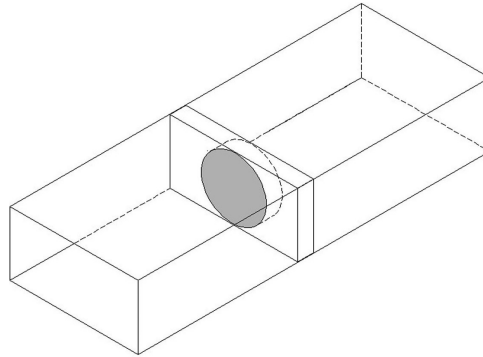
Figure 2.29: Arrangement of a resonant waveguide window.
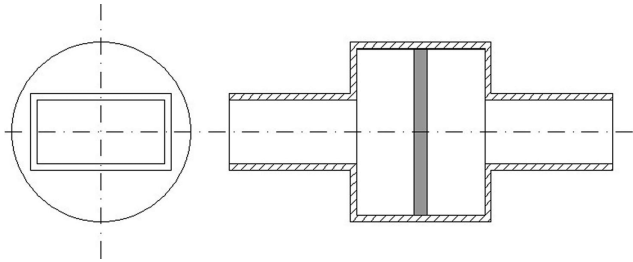
Figure 2.30: Arrangement of a pill-box window.

waveguide. The simplest form is the widely used pill-box window shown in Figure 2.30 which has a bandwidth of around 30% [47, 66]. The design of pill-box windows has been described in [45, 67, 68]. The electrical length of the circular section is chosen to be a half-wavelength when the effects of the dielectric are included. The diameter of the circular waveguide must be greater than the diagonal of the rectangular waveguide, and it is usual to make its cut-off frequency close to that of the rectangular waveguide. This type of window has ghost mode resonances associated with reflections between the transitions between the waveguides. The diameter of the circular guide is adjusted to ensure that none of these resonances falls within the working frequency band. The equivalent circuit of this window is shown in Figure 2.28(b), but some care is needed in defining the characteristic impedances in the circular section. The use of this equivalent circuit is justified by the close similarity between the electric field distributions over the aperture connecting the two waveguides. The correct relationship between their characteristic impedances can be found using the method described by Farmer [27]. An electromagnetic analysis of a pill-box window that includes full dimensions of the window is given in [69]. Pill-box windows can be designed to have travelling-wave fields in the dielectric by the use of inductive irises at the waveguide transitions [70]. Alternatively the transitions can be made using tapers, or other wide-band mode converters [50, 71]. Windows have also been described which use other modes in the circular waveguide [72, 73].

References

[1] S. Ramo *et al.*, *Fields and Waves in Communication Electronics*. New York: Wiley, 1965.

[2] J. Feinstein, 'Passive microwave components', in D. G. Fink and D. Christiansen, eds, *Electronic Engineers' Handbook*. New York: McGraw-Hill, pp. 9.2–9.27, 1982.

[3] R. E. Collin, *Foundations for Microwave Engineering*. New York: McGraw-Hill, 1966.

[4] A. F. Harvey, *Microwave Engineering*. New York: Academic Press, 1963.

[5] N. Marcuvitz, *Waveguide Handbook*. New York: McGraw-Hill, 1951.

[6] R. Cooper and R. G. Carter, 'High power RF transmission', in J. Miles, ed., *CERN Accelerator School: Radio Frequency Engineering*. Geneva: CERN, pp. 210–235, 2005.

[7] 'Selecting a transmission line for your broadcast system', vol. SP50115. Orland Park, IL: Andrew Corporation, 1998.

[8] S. Choroba, 'RF power transport', in R. Bailey, ed., *CERN Accelerator School: RF for accelerators.* Geneva: CERN, pp. 117–123, 2011.

[9] C. Kubota *et al.*, 'Parts of the 203D coaxial waveguide improved at KEK', presented at the LINAC 2002, Gyeongju, Korea, 2002.

[10] R. G. Carter, *Electromagnetism for Electronic Engineers*. Available at: http://book-boon.com/en/electromagnetism-for-electronic-engineers-ebook (accessed 3 October 2017) 2009.

[11] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer-Verlag, pp. 1–84, 2008.

[12] A. M. Perez *et al.*, 'Prediction of multipactor breakdown thresholds in coaxial transmission lines for traveling, standing, and mixed waves', *IEEE Transactions on Plasma Science,* vol. 37, pp. 2031–2040, 2009.

[13] J. Lorkiewicz *et al.*, 'Surface TiN coating of TESLA couplers at DESY as an antimultipactor remedy', in *10th Workshop on RF Superconductivity*, Tsukuba, Japan, pp. 448–452, 2001.

[14] G. Burt *et al.*, 'Multipactor simulations of the SPL power coupler', in *LINAC 2010*, Tsukuba, Japan, pp. 878–880, 2010.

[15] N. W. McLachlan, *Bessel Functions for Engineers*. Oxford University Press, 1954.

[16] R. G. Carter, *Electromagnetic Waves: Microwave Components and Devices.* London: Chapman & Hall, 1990.

[17] 'Rectangular waveguide specifications', *WAVEGUIDE COMPONENT: Specifications and Design Handbook* (7th edn) section DS-501 Exeter, NH: Cobham Defense Electronic Systems, 2006.

[18] S. Hopfer, 'The design of ridged waveguides', *IRE Transactions on Microwave Theory and Techniques,* vol. 3, pp. 20–29, 1955.

[19] S. B. Cohn, 'Properties of ridge wave guide', *Proceedings of the IRE,* vol. 35, pp. 783–788, 1947.

[20] 'Double ridge waveguide specifications', *WAVEGUIDE COMPONENT: Specifications and Design Handbook* (7th edn) section DS-503 Exeter, NH: Cobham Defense Electronic Systems, 2006.

[21] G. G. MacFarlane, 'Quasi-stationary field theory and its application to diaphragms and junctions in transmission lines and wave guides', *Journal of the Institution of Electrical Engineers – Part IIIA: Radiolocation,* vol. 93, pp. 703–719, 1946.

[22] J. P. Montgomery, 'On the complete eigenvalue solution of ridged waveguide', *IEEE Transactions on Microwave Theory and Techniques,* vol. 19, pp. 547–555, 1971.

[23] Y. Utsumi, 'Variational analysis of ridged waveguide modes', *IEEE Transactions on Microwave Theory and Techniques,* vol. 33, pp. 111–120, 1985.

[24] J. Helszajn and M. McKay, 'Voltage-current definition of impedance of double ridge waveguide using the finite element method', *Proceedings of the IEE: Microwaves, Antennas and Propagation,* vol. 145, pp. 39–44, 1998.

[25] M. McKay and J. Helszajn, 'Voltage-current definition of impedance of single-ridge waveguide', *IEEE Microwave and Guided Wave Letters,* vol. 9, pp. 66–68, 1999.

[26] W. J. R. Hoefer and M. N. Burton, 'Closed-form expressions for the parameters of finned and ridged waveguides', *IEEE Transactions on Microwave Theory and Techniques,* vol. 30, pp. 2190–2194, 1982.

[27] E. D. Farmer, 'Junction admittance between waveguides of arbitrary cross-sections', *Proceedings of the IEE – Part C: Monographs,* vol. 103, pp. 145–152, 1956.

[28] M. Guglielmi and C. Newport, 'Rigorous, multimode equivalent network representation of inductive discontinuities', *IEEE Transactions on Microwave Theory and Techniques,* vol. 38, pp. 1651–1659, 1990.

[29] S. B. Cohn, 'Microwave coupling by large apertures', *Proceedings of the IRE,* vol. 40, pp. 696–699, 1952.

[30] W. H. Watson, *The Physical Principles of Wave Guide Transmission and Antenna Systems*. Oxford: Clarendon Press, 1947.

[31] F. Lösch, *Jahnke-Emde-Lösch: Tables of Higher Functions*. Stuttgart: B.G. Teubner Verlagsgesellschaft, 1960.

[32] G. L. Matthaei *et al.*, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*. Dedham, MA: Artech House Books, 1980.

[33] L. Young, 'Tables for cascaded homogeneous quarter-wave transformers', *IRE Transactions on Microwave Theory and Techniques,* vol. 7, pp. 233–237, 1959.

[34] L. Young, 'Optimum quarter-wave transformers', *IRE Transactions on Microwave Theory and Techniques,* vol. 8, pp. 478–482, 1960.

[35] R. C. Johnson, 'Design of linear double tapers in rectangular waveguides', *IRE Transactions on Microwave Theory and Techniques,* vol. 7, pp. 374–378, 1959.

[36] S. B. Cohn, 'Optimum design of stepped transmission-line transformers', *IRE Transactions on Microwave Theory and Techniques,* vol. 3, pp. 16–20, 1955.

[37] L. Young, 'Tables for cascaded homogeneous quarter-wave transformers (Correction)', *IRE Transactions on Microwave Theory and Techniques,* vol. 8, pp. 243–244, 1960.

[38] R. E. Collin, 'The optimum tapered transmission line matching section', *Proceedings of the IRE,* vol. 44, pp. 539–548, 1956.

[39] R. W. Klopfenstein, 'A transmission line taper of improved design', *Proceedings of the IRE,* vol. 44, pp. 31–35, 1956.

[40] G. L. Ragan, *Microwave Transmission Circuits*. New York: McGraw-Hill, 1948.

[41] S. B. Cohn, 'Design of simple broad-band waveguide to coaxial line junctions', *Proceedings of the IRE,* vol. 35, pp. 920–926, 1947.

[42] Y. Zhou *et al.*, 'Design of millimeter wave wideband transition from double-ridge waveguide to coaxial line', *Journal of Infrared, Millimeter and Terahertz Waves,* vol. 32, pp. 26–33, 2011.

[43] R. Nie *et al.*, 'Simulation and design of 18–40 GHz ridge waveguide to coaxial transition', in *2011 IEEE International Conference on Microwave Technology & Computational Electromagnetics (ICMTCE)*, pp. 183–185, 2011.

[44] J. F. Gittins, *Power Travelling-Wave Tubes*. London: English Universities Press, 1965.

[45] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, UK: Research Studies Press, 1995.

[46] D. K. Abe and J. P. Calame, 'Advanced material technologies', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 649–689, 2005.

[47] A. Staprans *et al.*, 'High-power linear-beam tubes', *Proceedings of the IEEE,* vol. 61, pp. 299–330, 1973.

[48] A. Neuber *et al.*, 'Window breakdown caused by high-power microwaves', *IEEE Transactions on Plasma Science,* vol. 26, pp. 296–303, 1998.

[49] Varian, 'Technical Manual: Installation, Operation, Maintenance, Care and Handling Instructions, General: Microwave Tubes, Magnetron Tubes, Electron Tubes', 1 October 1979.

[50] D. B. Churchill, 'Problems associated with waveguide pressure windows for microwave tubes', in *7th National Conference on Tube Technology*, pp. 371–389, 1964.

[51] S. Chel *et al.*, 'Coaxial disc windows for a high power superconducting cavity input coupler', in *Proceedings of the 1999 Particle Accelerator Conference (PAC 99)*, vol. 2, pp. 916–918, 1999.

[52] S. Liu, 'A fast computational technique for RF window in millimetre wave tubes', *International Journal of Infrared and Millimetre Waves,* vol. 15, pp. 857–860, 1994.

[53] A. S. Gilmour Jr., *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-Field Amplifiers and Gyrotrons*. Norwood, MA: Artech House, 2011.

[54] M. C. Lin and D. S. Chuu, 'A novel wide-band high-transmission window for high-frequency microwave tubes', in *Third IEEE International Vacuum Electronics Conference (IVEC 2002)*, pp. 216–217, 2002.

[55] M. P. Forrer and E. T. Jaynes, 'Resonant modes in waveguide windows', *IRE Transactions on Microwave Theory and Techniques,* vol. 8, pp. 147–150, 1960.

[56] K. Cummings *et al.*, 'Results and lessons learned from conditioning 1 MW CW 350 MHz coaxial vacuum windows', in *LINAC 98*, Chicago, IL, pp. 938–940, 1998.

[57] J. Wang and R. G. Carter, 'Design of a coaxial window for high power CW magnetron', in *International Conference on Microwave and Millimeter Wave Technology 1998*, pp. 744–747, 1998.

[58] Y. Kang *et al.*, 'Electromagnetic simulations and properties of the fundamental power couplers for the SNS superconducting cavities', in *Proceedings of the 2001 Particle Accelerator Conference (PAC 2001)*, vol. 2, pp. 1122–1124, 2001.

[59] X. Hanus and A. Mosnier, 'Coaxial TW window for power couplers and multipactor considerations', in *1995 Workshop on RF Superconductivity*, Gif-sur-Yvette, France, pp. 701–705, 1995.

[60] C. Travier *et al.*, 'Design and test of a 1.3 GHz travelling wave window', in *1999 Workshop on RF Superconductivity*, Santa Fe, NM, pp. 427–436, 1999.

[61] R. Z. Gerlack, 'Small ceramic windows for large waveguides', *Microwave Journal,* vol. 37, pp. 110–114, 1994.

[62] E. Chojnacki *et al.*, 'Design of a high average power waveguide window', in *Proceedings of the 1997 Particle Accelerator Conference (PAC 1997)*, vol. 3., pp. 3177–3179, 1997.

[63] W. R. Fowkes *et al.*, 'Large diameter reduced field TE_{01} traveling wave window for X-band', in *Proceedings of the 1999 Particle Accelerator Conference (PAC 1999)*, vol. 2, pp. 783–785, 1999.

[64] N. M. Kroll *et al.*, 'Design of traveling wave windows for the PEP-II RF coupling network', in *Proceedings of the 1995 Particle Accelerator Conference (PAC 95)*, vol. 3, pp. 1741–1743, 1995.

[65] W. R. Fowkes *et al.*, 'Reduced field TE_{01} X-Band travelling wave window', in *Proceedings of the 1995 Particle Accelerator Conference (PAC 1995)*, pp. 1587–1589, 1995.

[66] R. S. Symons, 'Sealed wave guide window', US Patent 2,958,834, 1960.

[67] O. S. Lamba *et al.*, 'Design and optimization of asymmetric RF window for 6 MW pulse power S-band klystron', in *International Conference on Recent Advances in Microwave Theory and Applications (MICROWAVE 2008)*, pp. 44–46, 2008.

[68] V. V. P. Singh *et al.*, 'On the electrical design of pill-box type high power microwave window', *Journal of the IETE*, vol. 39, pp. 351–359, 1993.

[69] H. Arai *et al.*, 'An analysis of a vacuum window for lower hybrid heating', *IEEE Transactions on Plasma Science,* vol. 14, pp. 947–954, 1986.

[70] S. Michizono *et al.*, 'High-power test of pill-box and TW-in-ceramic type S-band rf windows', in *LINAC 94*, Tsukuba, Japan, pp. 457–459, 1994.

[71] Y. Otake *et al.*, 'Design and high-power test of a TE_{11} mode X-band RF window with taper transitions', in *Particle Accelerator Conference (PAC 95)*, Dallas, TX, vol. 3, pp. 1590–1592, 1995.

[72] A. E. Barrington and J. T. Hyman, 'A non-resonant waveguide window', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 104, pp. 35–38, 1957.

[73] R. L. Ives and J. Neilson, 'TE_{01} windows for high power klystron applications', in *International Electron Devices Meeting (IEDM '93) Technical Digest*, pp. 161–164, 1993.

3   Resonators

3.1   Introduction

Resonant cavities are important components in many microwave tubes because of their frequency selective properties and because they store electromagnetic energy. The purpose of this chapter is to describe the features which are common to all types of microwave resonator and to examine the properties of a few important types in more detail.

Any closed metal cavity supports an infinite number of electromagnetic resonances. In each of these the fields satisfy Maxwell's equations and the boundary conditions on the metal surface. If resistive losses associated with currents in the walls of the cavity are neglected then the electric and magnetic fields are in phase quadrature. The losses in metallic cavity resonators are usually small so that the fields differ very little from those in a loss-less cavity. The losses can then be calculated to a good approximation from the currents flowing in the walls when the losses are neglected.

The cavity resonators used in microwave tubes have simple shapes and are commonly cylindrically symmetrical around the axis of the electron beam. The modes which are chiefly of interest are those having a strong axial component of the electric field in the region of the beam. The simplest example is the pill-box cavity whose TM_{010} resonance is illustrated in Figure 3.1. This mode can be derived from the TM_{01} mode of a circular waveguide by inserting a pair of conducting planes normal to the axis. The resonant frequency for the lowest mode is the cut-off frequency of the waveguide where there is no axial variation of the electric field. Figure 3.1 shows, schematically, the fields, charges and currents in the cavity at intervals of a quarter of the resonant period T_0. The properties of this cavity are discussed in detail in Section 3.3.

For many purposes it is convenient to represent the properties of a cavity resonator by its equivalent circuit. This can be derived by representing conduction current paths by inductors, and displacement current paths by capacitors, as described in Chapter 2. Figure 3.2a shows the derivation of the equivalent circuit for the TM_{010} mode of a pill-box cavity. The series resistor r represents the conduction losses in the cavity walls. The analysis of this equivalent circuit is discussed in the next section.
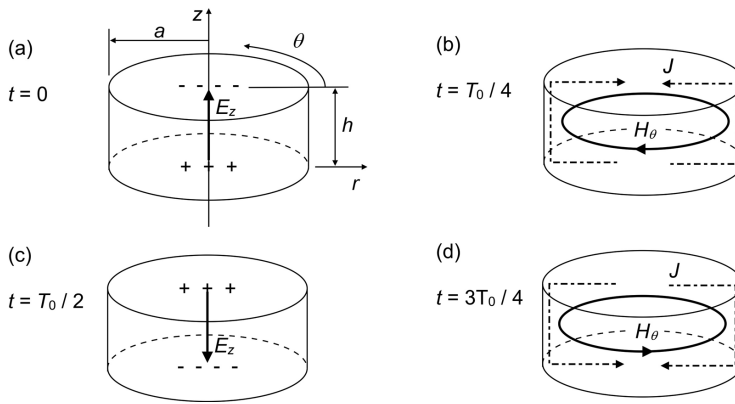
Figure 3.1: Fields, charges and currents in a pill-box cavity resonator.
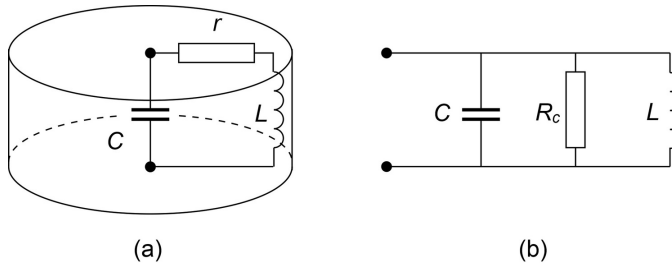
Figure 3.2: Equivalent circuits of a cavity resonator with conduction losses represented by (a) a series resistor, and (b) a parallel resistor.

The properties of pill-box cavity resonators and rectangular cavity resonators are discussed in Sections 3.3 and 3.4. Cylindrical re-entrant cavities, which are commonly used in vacuum tubes, are considered in Section 3.5. Section 3.6 discusses external coupling to cavity resonators from coaxial lines and waveguides. The last section of the chapter describes methods for measuring the properties of cavities.

3.2   Resonant Circuits

The properties of parallel resonant circuits are well-known and they are only summarised here [1]. We shall concentrate on those issues which are important when using equivalent circuits to model cavity resonators in microwave tubes.

3.2.1 The Properties of Resonant Circuits

The equivalent circuit shown in Figure 3.2(a) is derived from the physics of the resonator. Analysis of this circuit shows that the input impedance is given by

$$Z = \frac{rQ_U\left(1 - \dfrac{j}{Q_U}\dfrac{\omega_0}{\omega}\right)}{\dfrac{1}{Q_U} + j\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)}, \tag{3.1}$$

where the resonant frequency is

$$\omega_0 = \frac{1}{\sqrt{LC}}, \tag{3.2}$$

the unloaded Q factor is

$$Q_U = \frac{1}{\omega_0 rC} = \frac{\omega_0 L}{r}, \tag{3.3}$$

and

$$rQ_U = \sqrt{\frac{L}{C}}. \tag{3.4}$$

Microwave resonators commonly have unloaded Q factors of at least 1000. They can be made with values as high as 30,000 by careful design and manufacture. The second term in the numerator of (3.1) is therefore small compared with unity and can be neglected without serious error. The equation which results is then identical to that for the parallel resonant circuit shown in Figure 3.2(b).

$$Z = \frac{R_c/Q_U}{\dfrac{1}{Q_U} + j\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)}, \tag{3.5}$$

where the 'R over Q' (R/Q) is defined by

$$\frac{R_c}{Q_U} = \sqrt{\frac{L}{C}} = rQ_U \tag{3.6}$$

and

$$Q_U = \frac{R_c}{\omega_0 L} = \omega_0 R_c C. \tag{3.7}$$

The parallel resonant circuit is generally used in preference to that in Figure 3.2(a) because it is simpler to analyse. At microwave frequencies it is difficult to assign meanings to the circuit parameters C, L and R_c and it is usual to use the parameters ω_0, R_c/Q_U and Q_U to describe the resonator. Equations (3.2), (3.6) and (3.7) enable

these parameters to be calculated from the circuit parameters and vice versa. It should be noted that Q_U is the only parameter which depends upon the properties of the material from which the cavity is constructed. The resonant frequency and the R/Q of a cavity resonator therefore depend only upon its geometry.

The amplitude of the cavity impedance Z is readily derived from (3.5) as

$$\frac{|Z|}{R_c} = \frac{1}{\sqrt{1 + Q_U^2 \left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega} \right)^2}}. \tag{3.8}$$

Figure 3.3 shows the amplitude of the normalised impedance $|Z|/R_c$ in decibels as a function of normalised frequency and Q_U (see Worksheet 3.1). When the frequency of the signal applied to the circuit is equal to the resonant frequency the impedance is purely resistive and equal to R_c. It is clear from (3.8) that $|Z| < R_c$ at all other frequencies and that the width of the curve decreases as Q_U increases. This relationship is usually expressed in terms of the width of the curve at the points which are 3 dB below the peak amplitude. Then found that

$$Q_U = \frac{\omega_0}{2\Delta\omega}, \tag{3.9}$$

where $|Z| = R/\sqrt{2}$ when $\omega = \omega_0 \pm \Delta\omega$ provided that $\Delta\omega \ll \omega_0$. Modern test equipment can display the response curve with the vertical scale in decibels as shown in Figure 3.3. This makes it easy to measure the unloaded Q factor of a resonator (see also Section 3.7).

From (3.7) we see that to obtain a high unloaded Q factor the circuit must have a high shunt resistance (i.e low loss) and that leads to a useful alternative definition of Q_U. When an alternating voltage $V = V_0 \cos \omega t$ is applied to the terminals of the circuit shown in Figure 3.2(b) the maximum energy stored in the capacitor is

$$W = \frac{1}{2} C V_0^2. \tag{3.10}$$
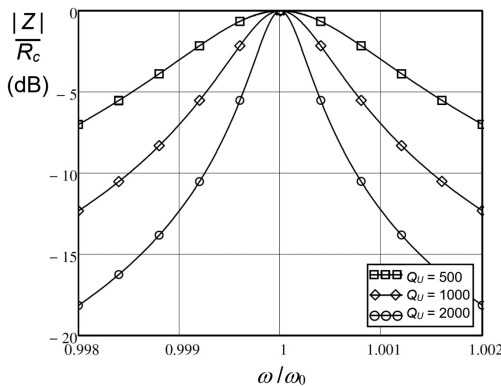
Figure 3.3: Amplitude of the impedance of a parallel resonant circuit as a function of normalised frequency and Q_U.

This energy is transferred backwards and forwards between the capacitor and the inductor during each cycle and the total stored energy remains constant, provided that the amplitude of the voltage at the terminals is held constant. The mean rate of dissipation of energy through resistive losses is

$$P_L = \frac{V_0^2}{2R_c}. \tag{3.11}$$

Eliminating V_0 between (3.10) and (3.11) and substituting for $R_c C$ in (3.7) gives

$$Q_U = \frac{\omega_0 W}{P_L} = \frac{2\pi W}{\Delta W}, \tag{3.12}$$

where ΔW is the energy dissipated per cycle. The rate of dissipation of energy is therefore

$$\frac{dW}{dt} = -\frac{\Delta W}{T_0} = -\frac{\omega_0}{Q_U} W, \tag{3.13}$$

where T_0 is the resonant period of the circuit. Equation (3.13) can readily be integrated to show that, when the external excitation is removed, the stored energy decays exponentially with a time constant $\tau = Q_U / \omega_0$. The physical significance of the parameter R/Q is revealed by substituting for P_L from (3.11) into (3.12) to give

$$\left(\frac{R_c}{Q_U} \right) = \frac{V_0^2}{2\omega_0 W}. \tag{3.14}$$

Thus R/Q is a measure of the relationship between the voltage across the terminals of the circuit and the energy stored in it.

To complete our review of the theory of parallel resonant circuits we must examine the phase of Z given by

$$\angle Z = \arctan\left[Q_U \left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0} \right) \right]. \tag{3.15}$$

Figure 3.4 shows how the phase of the cavity impedance varies with normalised frequency and Q_U. At resonance the impedance is purely resistive and the phase of Z is zero. At lower frequencies the behaviour of the circuit is dominated by the reactance of the inductor and $\angle Z \rightarrow 90°$. At higher frequencies the capacitor has the greater effect and $\angle Z \rightarrow -90°$. At the 3 dB points

$$\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0} = \pm\frac{1}{Q_U} \tag{3.16}$$

and $\angle Z = 45°$.
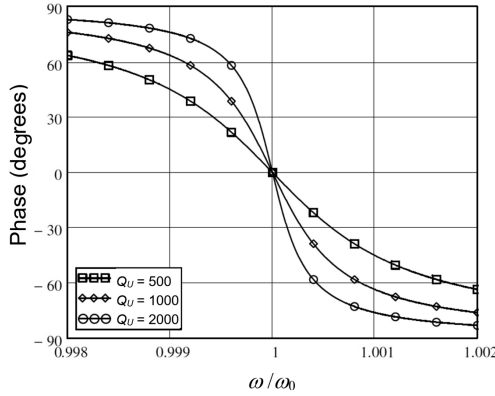
Figure 3.4: Phase of the impedance of a parallel resonant circuit as a function of normalised frequency and Q_U.

3.2.2   External Loading of Resonant Circuits

When resonant structures are used in microwave tubes their properties are usually modified by the presence of electrons within them, by connections to external wave-guides, or both. The ways in which these effects influence the behaviour of a tube differ somewhat between different types of tubes. We will therefore defer a detailed examination of electron loading and external loading effects to the discussions of particular types of tube. For our present purposes it is sufficient to consider the effect on a resonator when it is connected to an external circuit comprising a resistor and a reactor in parallel as shown in Figure 3.5. This circuit can represent the combined effects of electron loading and external loading, including any impedance transformer between the resonator and the external circuit. The admittance of this circuit, measured at its terminals, is

$$Y = \frac{1}{R_c} + j\omega C + \frac{1}{j\omega L} + \frac{1}{R_E} + jB_E. \tag{3.17}$$

At resonance the sum of the susceptances is zero. Thus the susceptance jB_E of the load modifies the resonant frequency. This effect is usually quite small and it can generally be compensated for, if necessary, by provision of an adjustable tuning element in the resonator. It manifests itself in klystrons as a small shift in the frequency of the cavities when the beam is turned on, and in magnetrons as frequency pushing and pulling (see Chapters 13 and 15). The effect of the load resistance is usually more significant because it may alter the Q of the circuit considerably. If we assume that the resonant frequency is unchanged then the Q factor is now

$$\frac{1}{Q_L} = \omega_0 L \left(\frac{1}{R_c} + \frac{1}{R_E} \right) = \frac{1}{Q_U} + \frac{1}{Q_E}, \tag{3.18}$$

where Q_L is the *loaded Q* of the resonator, and Q_U is the *unloaded Q* given by (3.7). Q_E, known as the *external Q*, is the Q factor which would be measured if the only
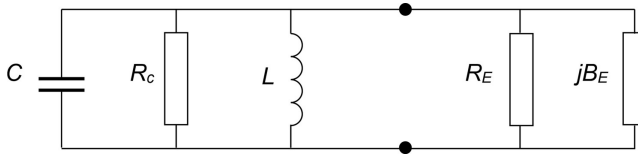
Figure 3.5: Parallel resonant circuit with an external load.

loss were that provided by the external resistance. If the external resistance is equal the resistance of the circuit at resonance then $Q_L = Q_U/2$.

We may define a coupling factor K by

$$K = \frac{R_c}{R_E} = \frac{Q_U}{Q_E} \tag{3.19}$$

so that

$$Q_U = Q_L(1 + K). \tag{3.20}$$

This effect is important in klystrons because both the beam loading and the external loading increase the bandwidths of the cavities and, therefore, of the tube. In addition we shall see that variation of the external match presented to the output of a klystron has an important effect upon the output gap voltage and on the efficiency of the tube. In magnetrons, changes in the match of the load lead to changes in the frequency of oscillation, and of the output power, through frequency pulling.

3.2.3   Excitation of Resonant Circuits

Figure 3.6 shows the equivalent circuit of a resonator connected to a source whose resistance is R_S. The rate of change with time of the current drawn from the source is

$$\frac{dI}{dt} = C \frac{d^2V}{dt^2} + \frac{1}{R'} \frac{dV}{dt} + \frac{V}{L}, \tag{3.21}$$

where R' represents the parallel combination of R_c with R_S so that

$$\frac{1}{R'} = \frac{1}{R_c} + \frac{1}{R_S} = \frac{1 + K}{K R_S}, \tag{3.22}$$

where K is defined by (3.19). The voltage across the terminals of the resonant circuit therefore obeys the differential equation

$$\frac{d^2V}{dt^2} + \frac{\omega_0}{Q_L} \frac{dV}{dt} + \omega_0^2 V = -\frac{\omega_0 R'}{Q_L} I_0 \omega \sin(\omega t) \tag{3.23}$$

in which the equivalent circuit parameters have been replaced by their microwave equivalents. The general solution of this equation for the case where the source
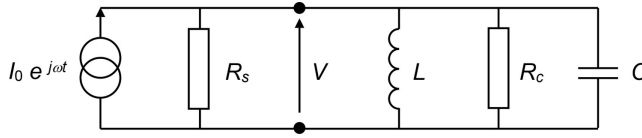
Figure 3.6: Equivalent circuit of a resonator connected to a source.

frequency is the resonant frequency of the circuit, and the source is connected to the circuit at $t = 0$, is

$$V = \left(-I_0 R' \cos(\omega_0 t) + \frac{I_0 R'}{2Q_L} \sin(\omega_0 t)\right) \exp\left(-\frac{t}{\tau}\right) + I_0 R' \cos(\omega_0 t), \quad (3.24)$$

where

$$\tau = \frac{2Q_U}{\omega_0 (1 + K)} = \frac{2Q_L}{\omega_0}. \quad (3.25)$$

The derivation of this equation has assumed that $Q_L^2 \gg 1$. The loaded Q factors of microwave resonators are typically greater than 100 so that (3.24) can be approximated as

$$V = I_0 R' \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] \cos(\omega_0 t). \quad (3.26)$$

We note that, if $Q_L \gg 1$, the voltage is in phase with the current.

The energy stored in the circuit is equal to the maximum energy stored in the capacitor, namely

$$W(t) = \frac{1}{2} C |V|^2 = \frac{1}{2} C (I_0 R')^2 \left[1 - \exp\left(-\frac{t}{\tau}\right)\right]^2. \quad (3.27)$$

When the source is delivering power into a matched load, the forward power is

$$P_+ = \frac{1}{8} I_0^2 R_S, \quad (3.28)$$

so that, making use of (3.7), (3.19) and (3.28), we can write (3.27) as

$$W(t) = P_+ \frac{Q_U}{\omega_0} \frac{4K}{(1 + K)^2} \left[1 - \exp\left(-\frac{t}{\tau}\right)\right]^2. \quad (3.29)$$

This expression is maximum when $K = 1$ (i.e. $R_c = R_S$) so that the resonator is matched to the source. It is then said to be critically coupled to the source and the final stored energy is

$$W_0 = P_+ \frac{Q_U}{\omega_0}. \quad (3.30)$$

Then (3.29) can be written

$$W(t) = W_0 \frac{4K}{(1+K)^2} \left[1 - \exp\left(-\frac{t}{\tau}\right)\right]^2.$$ (3.31)

When $K < 1$ the resonator is under-coupled and when $K > 1$ it is over-coupled. The time taken to accumulate the energy in the resonator, known as the filling time, can be estimated by noting that when $t = Q_L T_0$ the stored energy is 91.5% of its final value. When K is either much greater than, or much less than, unity the stored energy is low because most of the power supplied by the source is reflected back into it. Figure 3.7 shows the normalised stored energy $W(t)/W_0$ as a function of time for three values of K (see Worksheet 3.1). These results are relevant both when a resonator is excited by an external waveguide and when it is excited by a modulated electron beam.

As the stored energy builds up in the resonator the impedance presented to the source changes. When the source is connected to a matched load the magnitude of the load voltage in the steady state is

$$V_0 = \frac{1}{2} I_0 R_S.$$ (3.32)

The magnitude of the voltage in the general case is found from (3.22), (3.26) and (3.32) to be

$$\frac{|V|}{V_0} = \frac{2K}{1+K} \left[1 - \exp\left(-\frac{t}{\tau}\right)\right].$$ (3.33)

Since the voltage and current are in phase with each other at resonance, the voltage reflection coefficient of the resonator is given by

$$S_{11} = \frac{|V| - V_0}{V_0} = \frac{2K}{1+K} \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] - 1.$$ (3.34)
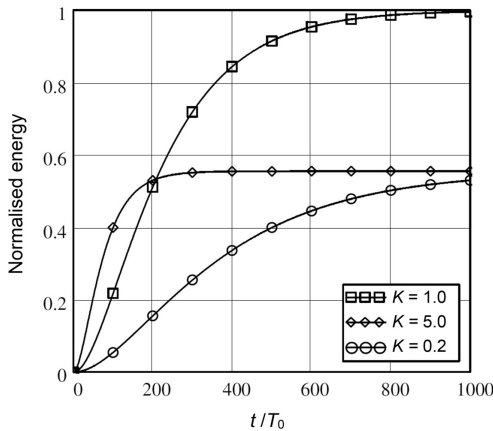
Figure 3.7: Normalised energy stored in a cavity resonator as a function of time for three coupling factors.
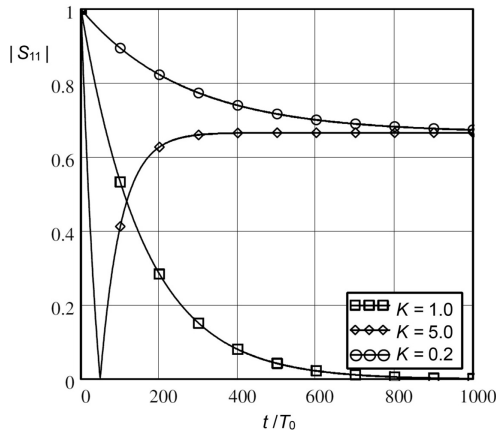
Figure 3.8: Input match of a resonant circuit as a function of time for three coupling factors.

Figure 3.8 shows how S_{11} varies with time for three values of K. When the output power from a pulsed microwave tube is supplied to a resonant cavity or structure then a time-varying mismatch is seen by the tube. As the reflected power may modify the behaviour of the tube it is necessary to place a circulator between the tube and the resonator to prevent this. When the source frequency differs from the resonant frequency the voltage and current are no longer in phase with each other and the analysis above must be modified accordingly.

3.2.4 Coupled Resonators

The bandwidth of a cavity resonator can be increased by coating its surfaces with a lossy material, by connecting it to an external resistive load, or by increasing the electron loading. These methods of reducing Q are often unsuitable because they increase the RF losses. An alternative technique is to couple two resonators together. This method is used, for example, to achieve the bandwidth required in the inductive output tubes (IOTs) used for television broadcasting (see Section 12.6). The properties of coupled-cavity slow-wave structures are considered in Section 4.6. Here we restrict our attention to a pair of identical parallel resonant circuits which are coupled by a mutual inductance, as shown in Figure 3.9. The impedance matrix for this network is

$$
\begin{bmatrix} V_1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/j\omega C & -1/j\omega C & 0 \\ -1/j\omega C & Z & -j\omega k L \\ 0 & -j\omega k L & Z \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix}, \tag{3.35}
$$

where

$$
Z = \frac{1}{j\omega C} + r + j\omega L. \tag{3.36}
$$

From the last line of this equation

$$j\omega k L\, i_2 = Z\, i_3.$$

(3.37)

Substituting for i_3 in (3.35) gives

$$\begin{bmatrix} V_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/j\omega C & -1/j\omega C \\ -1/j\omega C & Z' \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix},$$

(3.38)

where

$$Z' = Z + \frac{\omega^2 k^2 L^2}{Z}.$$

(3.39)

When the frequency is equal to the resonant frequency of either resonator

$$Z' = r\left(1 + \frac{\omega_0^2 k^2 L^2}{r^2}\right) = r\left(1 + k^2 Q_U^2\right).$$

(3.40)

If $kQ_U = 1$ the second circuit is matched to the first and the circuits are said to be critically coupled. The input impedance and phase of this circuit can be computed as a function of frequency and of the coupling factor k (see Worksheet 3.1).

Figure 3.10 shows the magnitude of the input impedance of coupled resonators whose uncoupled Q is 1000 as a function of frequency and coupling factor. When
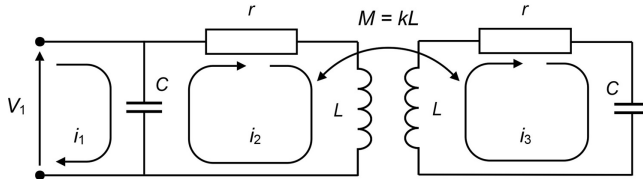
Figure 3.9: Parallel resonant circuits coupled by a mutual inductance.
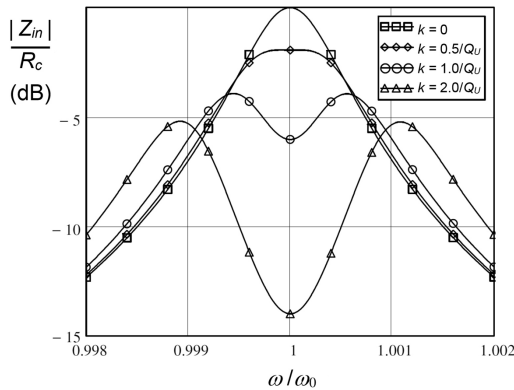
Figure 3.10: Amplitude response of two identical resonators coupled by a mutual inductance.
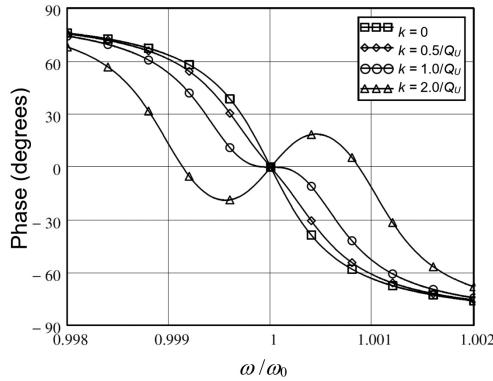
Figure 3.11: Phase response of two identical resonators coupled by a mutual inductance.

$k = 0$ the response is identical to that shown in Figure 3.3, as would be expected. As k is increased the peak of the curve first flattens and then divides into two peaks. When the resonators are critically coupled the curve resembles a Tchebychev response. The bandwidth is then increased in comparison with a single resonator having the same maximum impedance. The phase of the input impedance of the coupled resonators is shown in Figure 3.11. When k is close to zero the curve is very similar to that for a single resonator with a single frequency at which the phase is zero. This continues as k increases until there is a point of inflection when $k = 1/Q_U$. For greater values of k there are three points at which the phase is zero. The behaviour of resonators which are coupled in other ways is similar to that shown here.

3.3   Pill-Box Cavity Resonators

The resonant circuits used at frequencies greater than about 100 MHz are invariably cavity resonators. These resonators can take many forms but all support an infinite number of resonant modes. The simplest resonant cavities are formed by placing a pair of metallic walls across a metal waveguide having a uniform cross-section. These reflect the propagating modes in the waveguide and create resonances whenever there is a standing wave. This is illustrated in this section for cavities based on circular waveguides and in the following section for those based on rectangular waveguides.

Figure 3.1 shows schematically the TM_{010} mode of a resonator based on a circular waveguide. The possible resonant modes are derived from the TE and TM modes of the waveguide by applying the appropriate boundary conditions in the z direction. The $TM_{0,n}$ modes are of greatest interest because they have a longitudinal component of the electric field on the axis that can interact with an electron beam. The tangential component of the electric field must be zero on all the metal surfaces and the resonant condition is therefore

$$\beta_{m,n} h = p\pi, \tag{3.41}$$

where $p = 0, 1, 2 \ldots$ and $\beta_{m,n}$ is the propagation constant of the $TM_{m,n}$ mode of the circular waveguide found using (2.11) and (2.85). The resonant mode which satisfies (3.41) is the $TM_{m,n,p}$ mode of the cavity. The z component of the magnetic field is zero for all these modes and the field patterns can be deduced from Figure 2.15 by placing the end walls of the cavity at planes where the transverse component of the electric field is zero.

The lowest mode is the TM_{010} mode where the frequency is equal to the cut-off frequency of the TM_{01} circular waveguide mode, and the electric field has only a z component that does not vary with z. The field components are

$$E_z = E_0 J_0 \left(\beta_C r \right) \exp \left(j \omega t \right) \tag{3.42}$$

and, from (2.20),

$$H_\theta = j \sqrt{\frac{\varepsilon_0}{\mu_0}} E_0 J_1 \left(\beta_C r \right) \exp \left(j \omega t \right), \tag{3.43}$$

where $\beta_C a = 2.405$. Thus the resonant frequency is

$$\omega_0 = 2.405 \frac{c}{a}. \tag{3.44}$$

The stored energy in the cavity can be calculated from the maximum values of either the electric or magnetic field. Thus

$$W = \frac{\varepsilon_0}{2} \int_0^a 2 \pi r h E_z^2 \, dr. \tag{3.45}$$

Substituting for E_z from (3.42) gives

$$W = \pi h \varepsilon_0 E_0^2 \int_0^a J_0 \left(\beta_C r \right)^2 r \, dr. \tag{3.46}$$

Evaluating the integral gives [2]

$$W = \left(\varepsilon_0 h a^2 E_0^2 \right) \cdot \frac{\pi}{2} J_1^2 \left(2.405 \right) = 0.423 \varepsilon_0 h a^2 E_0^2. \tag{3.47}$$

If the terminals of the equivalent circuit are chosen to be at the centres of the two flat faces then the voltage across the cavity is

$$V_0 = E_0 h, \tag{3.48}$$

so that, from (3.14), (3.44), (3.47), and (3.48)

$$\left(\frac{R_c}{Q_U} \right) = 0.491 \frac{h}{a} \sqrt{\frac{\mu_0}{\varepsilon_0}} = 185 \frac{h}{a} \, \Omega. \tag{3.49}$$

The theoretical value of the shunt resistance of a high Q cavity can be calculated by assuming that the current density in the walls of the cavity is the same as in a lossless cavity, and that the surface resistance is given by

$$R_s = \frac{1}{\sigma\delta} = \sqrt{\frac{\omega\mu_0}{2\sigma}}, \tag{3.50}$$

where σ is the conductivity of the walls of the cavity and δ is the skin depth [3]. Some investigators have reported measurements of anomalous surface resistance at millimetre wavelengths that cause the surface resistance to be appreciably higher than that obtained from (3.50). However, a careful review of the experimental evidence concluded that these results are probably erroneous [4]. At frequencies up to several THz the surface resistance can be calculated correctly from the classical relaxation-effect (or Drude dispersion) model [5]

$$R_s = \sqrt{\frac{\omega\mu_0}{2\sigma}} \left[\sqrt{\sqrt{1+(\omega\tau)^2} - \omega\tau} \right] \tag{3.51}$$

where τ is the relaxation time of the metal. Using the material constants for copper ($\sigma = 5.959 \times 10^7$ S m^{-1} and $\tau = 25.018$ fs [4]) it is found that the error in using (3.50) is less than 1% at 100 GHz.

The current density is equal to the tangential magnetic field on the cavity wall and at right angles to it. Thus the current flow is radial in the flat ends of the cavity and axial in the curved walls. The power dissipated in the walls is

$$P_L = \pi ah R_s \left| H_\theta(a) \right|^2 + 2\pi R_s \int_0^a \left| H_\theta(r) \right|^2 r\, dr. \tag{3.52}$$

Substituting for H_θ from (3.43) gives

$$P_L = \pi R_s E_0^2 \frac{\varepsilon_0}{\mu_0} \left[ah J_1(2.405)^2 + 2\int_0^a J_1(\beta_C r)^2 r\, dr \right] \tag{3.53}$$

and, evaluating the integral,

$$P_L = \pi R_s E_0^2 \frac{\varepsilon_0}{\mu_0} a(a+h) J_1(2.405)^2. \tag{3.54}$$

Then, from (3.12), (3.44), (3.47), and (3.54)

$$Q_U = \frac{2.405}{2} \sqrt{\frac{\mu_0}{\varepsilon_0}} \frac{h}{R_s(a+h)}. \tag{3.55}$$

It is helpful to illustrate these equations by numerical examples. Table 3.1 shows the theoretical parameters of copper cavities for which $a = h$. It can be seen that the skin depth is small compared with the radius of the cavity in every case, so the effect of the finite conductivity on the resonant frequency is negligible compared with

Table 3.1: Theoretical parameters of copper pill-box cavity resonators (Worksheet 3.2) ($a = h$, $\sigma = 5.959 \times 10^7$ S m^{-1})

Frequency (GHz)	1.0	3.0	10.0	30.0
a (mm)	114.7	38.25	11.47	3.82
δ (µm)	2.06	1.19	0.65	0.38
R_c/Q_U (Ω)	185	185	185	185
Q_U	27,800	16,100	8,800	5,100
R_c (MΩ)	5.15	2.97	1.63	0.94

manufacturing tolerances and the effects of thermal expansion. The values of Q_U are high enough to permit the representation of the cavity losses by a shunt resistor.

3.3.1 Effects of Surface Roughness

The Q factors achieved in practice are commonly less than the theoretical values in Table 3.1 because the effective lengths of the current paths are increased by surface roughness. This problem has been studied theoretically by considering surfaces in which there are periodic grooves normal to the current flow [6, 7]. If the skin depth is small compared with the surface roughness then the resistance is increased by the ratio of the actual path length to the ideal path length [7]. When the skin depth is comparable to, or greater than, the surface roughness the resistance is less than that calculated from the path length and tends to the theoretical resistance as the skin depth increases. The resistance depends on the shape and the spacing of the grooves. Similar results have been obtained by modelling surfaces with random roughness [8]. It was found that the additional loss depends upon the RMS roughness, the correlation length and the correlation function. Grooves parallel to the current flow have a similar, but somewhat smaller, effect to those normal to the current flow. The resistance rises rapidly as the skin depth decreases and then reaches a value which is approximately constant [7].

An empirical formula for the effect of surface roughness is

$$\frac{R_r}{R_s} = 1 + \frac{2}{\pi}\arctan\left[1.4\left(\frac{\Delta}{\delta}\right)^2\right], \tag{3.56}$$

where R_s is given by (3.50), δ is the skin depth, and R_r is the surface resistance of a surface having RMS roughness Δ[8–10]. Figure 3.12 shows the variation of normalised surface resistance with Δ/δ given by (3.56). While this figure indicates in general how the surface resistance varies with the surface roughness it is clear from the discussion above that it cannot represent all possible surface conditions.

Experimental measurements of the losses in copper waveguides at 24 GHz showed normalised loss ranging from 1.09 for a machined surface to 1.8 for an electroplated surface [11]. It was shown that the loss could be increased by increasing
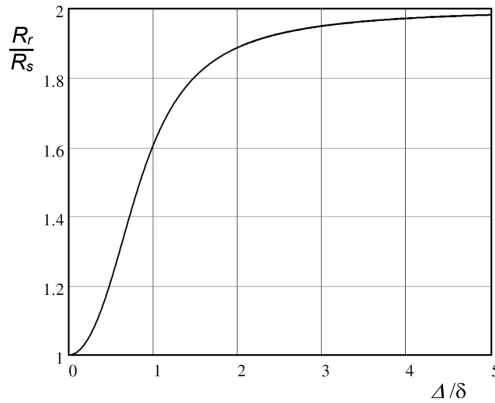
Figure 3.12: Empirical relationship between surface resistance and surface roughness.

the roughness of the surface. Waveguide measurements at 9.4 GHz were made by Benson and his co-workers, who also made measurements of the surface roughness of the experimental samples [12–14]. It was found that the effects of surface roughness could be reduced by electro-polishing or chemical polishing. Table 3.2 shows the normalised surface resistance of copper waveguides. Formulae were derived for waveguide losses which included the effects of surface roughness defined by the ratio of the actual length of the surface to the ideal length. Three different factors were included to account for the differences in the roughness of the walls of the waveguide normal to each component of the current flow. The values of these factors were measured for the waveguides studied and it was concluded that the increased loss could be explained solely by the increased path length when the skin depth was small compared with the roughness. For example a 3 GHz brass waveguide in which the surface roughness was about an order of magnitude greater than the skin depth was found to have normalised path lengths in the range 1.10 to 1.73 with an average value of 1.34.

Measurements at different frequencies made using copper waveguides are summarised in Table 3.3 [13, 15]. The theoretical surface resistance was calculated from measured values of the bulk DC conductivity. The effects of surface roughness were calculated and it was concluded that the additional loss could be explained in this way at 10 GHz but not at 35 GHz. It seems likely that the difference is to be explained by the neglect of factors other than the additional path length in the calculated values. It was found that the surface resistance could be reduced by annealing and this was explained in terms of the modification of the conductivity of a surface layer of material which had been work-hardened during manufacture [15].

Measurements of conduction losses have been made using resonators with artificially roughened surfaces at 35 GHz [16], and between 0.4 and 0.85 THz [17]. The variation of surface resistance with RMS surface roughness can be described by (3.56) if $\Delta/\delta \geq 1$. The asymptotic value corresponds to the increase in the surface area as expected. When $\Delta/\delta < 1$ the agreement with theory is poor. It was suggested

Table 3.2: Ratio of measured to calculated surface resistance for copper waveguides at about 10 GHz [12]

Copper	R_r/R_s
Bright electroplate	1.001
Electro-polished	1.002
Chemically polished	1.003
Precision drawn	1.012

Table 3.3: Ratio of measured to calculated surface resistance for drawn copper waveguides [11, 13, 15]

Frequency (GHz)	R_r/R_s
9.375	1.034
24	1.37
35	1.55–1.57
70	1.7–2.5
140	2.1–2.5

in [16] that the difference could be explained by an anomalous skin effect but this conclusion has been challenged by Lucyszyn [4].

From the preceding discussion it can be concluded that the effects of surface roughness may be represented by a modified surface resistance. The nature of the surface roughness and the direction of scratches and machining marks are important and it is not easy to calculate the effective surface resistance in general. When the surface roughness is appreciably greater than the skin depth the effective resistance can be calculated if the effective path length is known. The unloaded Q factor of a cavity resonator therefore depends on the manufacturing process as well as on the material from which it is made. In the case of the pill-box cavity azimuthal machining marks caused by turning the cavity parts on a lathe would have the greatest effect. The reduction in Q_U caused by surface roughness is not easy to quantify on theoretical grounds and estimates must be made from previous experience. In one case where the surface roughness was comparable with the skin depth at 3.2 GHz it was found that the measured Q was 72% of the theoretical value [18].

3.3.2   Higher-Order Modes

In addition to the fundamental resonance a pill-box cavity has an infinite set of higher-order resonances. The circularly symmetrical TM_{0n0} modes correspond to the solutions of $J_0\left(\beta_C a\right) = 0$ which occur at $\beta_C a = 2.405, 5.520, 8.654, \ldots$ [2]. When $m > 0$ the radial variation of the electric field is as $J_m\left(\beta_C r\right)$ and the resonances of

the TM_{mn0} modes are the zeroes of $J_m(\beta_C a)$. The electric field of these modes is weak close to the axis and therefore they do not interact strongly with an electron beam located there. Other higher-order modes (TM_{mnp} and TE_{mnp} with $p \geq 1$) exist in which **E** has components other than E_z but these modes are not usually excited. However, because a modulated electron beam has currents which are harmonics of the signal frequency it is necessary to check that the cavities do not have higher-order resonances at these frequencies.

3.4   Rectangular Cavity Resonators

A rectangular cavity resonator can be developed from a rectangular waveguide as shown in Figure 3.13. The lowest TM mode in a rectangular waveguide is the TM_{11} mode shown in Figure 2.6. The z component of the electric field of the TM_{110} resonance of the cavity is

$$E_z = E_0 \sin\left(\frac{\pi x}{a}\right)\sin\left(\frac{\pi y}{b}\right)\exp(j\omega t). \tag{3.57}$$

Substituting this expression in (2.36) we obtain

$$\beta_C^2 = \left(\frac{\pi}{a}\right)^2 + \left(\frac{\pi}{b}\right)^2. \tag{3.58}$$

The components of the magnetic field in an empty waveguide are found from (2.20) to be

$$H_x = j\sqrt{\frac{\varepsilon_0}{\mu_0}}\frac{a}{\sqrt{a^2+b^2}}E_0 \sin\left(\frac{\pi x}{a}\right)\cos\left(\frac{\pi y}{b}\right)\exp(j\omega t) \tag{3.59}$$

$$H_y = -j\sqrt{\frac{\varepsilon_0}{\mu_0}}\frac{b}{\sqrt{a^2+b^2}}E_0 \cos\left(\frac{\pi x}{a}\right)\sin\left(\frac{\pi y}{b}\right)\exp(j\omega t). \tag{3.60}$$
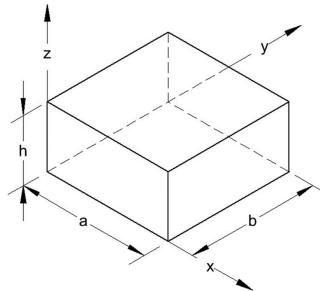
Figure 3.13: Arrangement of a rectangular cavity resonator.

The stored energy is

$$W = \frac{h}{2} \varepsilon_0 E_0^2 \int\limits_0^a \left(\int\limits_0^b \sin^2\left(\frac{\pi x}{a}\right) \sin^2\left(\frac{\pi y}{b}\right) dy \right) dx,$$ (3.61)

which can be integrated to give

$$W = \frac{\varepsilon_0}{8} abhE_0^2.$$ (3.62)

Then, since $V_0 = E_0 h$

$$\frac{R}{Q} = \frac{4}{\pi} \cdot \frac{h}{\sqrt{a^2 + b^2}} \sqrt{\frac{\mu_0}{\varepsilon_0}}.$$ (3.63)

The Q of the cavity can be found in the same way as that for the pill-box cavity [3]. The result in our notation is

$$Q = \frac{\pi}{4R_s} \sqrt{\frac{\mu_0}{\varepsilon_0}} \left[\frac{2h(a^2 + b^2)^{3/2}}{ab(a^2 + b^2) + 2h(a^3 + b^3)} \right].$$ (3.64)

Rectangular cavities support higher-order modes similar to those of the pill-box cavity.

3.5   Re-entrant Cavities

Pill-box cavities are not commonly used in microwave tubes. The reason for this can be understood by considering the time $(t = g/u_0)$ taken for an electron travelling with velocity u_0 to cross a cavity gap of length g. The change in the phase of the electric field in the cavity during this time is ωt and it is desirable that this should have a value of less than, say, $\pi/3$ to ensure that the field seen by each electron is approximately constant as it crosses the cavity. Thus

$$\frac{\omega}{u_0} \leq \frac{\pi}{3g}.$$ (3.65)

Combining this equation with (3.44) for a pill-box cavity where $g = h$, we find that, in order to keep the transit angle within the limits specified,

$$\frac{h}{a} \leq 0.435 \frac{u_0}{c}.$$ (3.66)

Most linear-beam tubes operate in the region from 5 to 100 kV and the corresponding range of u_0/c is 0.1 to 0.5. Table 3.4 shows the theoretical parameters of a copper pill-box cavity at 3 GHz for a range of beam voltages. The cavity radius is 38.2 mm and the skin depth is 1.22 μm in each case. It can be seen that the low

Table 3.4: Theoretical parameters of copper pill-box cavities at 3 GHz

u_0/c	0.1	0.3	0.5
V_0 (kV)	2.6	25	80
h/a	0.044	0.131	0.218
a (mm)	38.3	38.3	38.3
h (mm)	1.68	5.01	8.34
Q_U	1354	3722	5751
R_e/Q_U (Ω)	8.14	24.2	40.3
R_e (kΩ)	11.0	90.2	232.0

values of h/a required by (3.66) leads to very low unloaded Q factors and shunt impedances. For this reason simple cavities are only used in high-power (forward-fundamental) coupled-cavity TWTs, where the cavity performance is acceptable, and in high frequency (millimetre wave) tubes where constructional difficulties rule out the use of cavities having other shapes.

The cavities used in microwave tubes usually have the re-entrant cylindrical shape shown in Figure 3.14. Consider the comparison between a pill-box cavity and a re-entrant cavity which have the same resonant frequency and interaction gap length. The capacitance across the interaction gap is smaller in the re-entrant cavity than it is in the pill-box cavity. Therefore the inductance of the re-entrant cavity must be greater than that of the pill-box cavity to keep the resonant frequency constant. As a result the R/Q of the re-entrant cavity is greater than that of the pill-box cavity.

3.5.1   Method of Moments Model of Re-entrant Cavities

The properties of a re-entrant cavity having the general shape shown in Figure 3.14 can be calculated with great accuracy using the method of moments. The outline of the method is described in [19] and further information is given in [20, 21]. The cavity is divided into three concentric regions $(I, II$ and $III)$, each having constant axial length, whose outer radii are a, a' and A respectively. The axial component of the electric field and the azimuthal component of the magnetic field are expanded in Fourier series of basis functions on $r = a$ and $r = a'$ so that

$$\begin{Bmatrix} E_z \\ H_\theta \end{Bmatrix} = \sum_{m=0}^{\infty} \begin{Bmatrix} e_m^{II} \\ h_m^{II} \end{Bmatrix} \cos\left(\frac{m\pi z}{g}\right) \tag{3.67}$$

in region II and similarly for the other regions, with the number of terms in the summations chosen independently for each region. The (unknown) amplitudes are written as e, h on $r = a$ and ee, hh on $r = a'$. The requirement that the fields should satisfy Maxwell's equations and the boundary conditions in region I can be expressed as

$$\left[h^I\right] = \left[G^I(k)\right]\left[e^I\right], \tag{3.68}$$
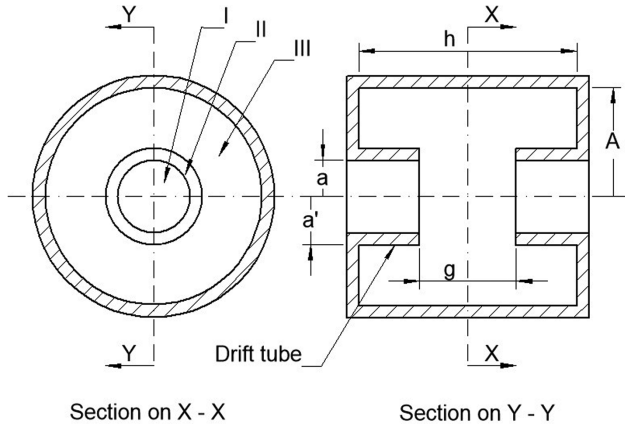
Figure 3.14: Arrangement of a cylindrical re-entrant cavity resonator.

where $k = \omega/c$ and the diagonal matrix G^I is defined in [20]. Similarly in region *III*

$$[hh^{III}] = [G^{III}(k)][ee^{III}].\tag{3.69}$$

The constants at the inner and outer boundaries of region *II* are related to one another because the fields are solutions of Maxwell's equations and this can be expressed as

$$\begin{bmatrix} h_m^{II} \\ hh_m^{II} \end{bmatrix} = [U_m(k)] \begin{bmatrix} e_m^{II} \\ ee_m^{II} \end{bmatrix}.\tag{3.70}$$

Equation (3.70) can be rearranged as a partitioned matrix

$$\begin{bmatrix} h^{II} \\ \cdots \\ hh^{II} \end{bmatrix} = \begin{bmatrix} U_{1,1}(k) & \vdots & U_{1,2}(k) \\ \cdots & \cdots & \cdots \\ U_{2,1}(k) & \vdots & U_{2,2}(k) \end{bmatrix} \begin{bmatrix} e^{II} \\ \cdots \\ ee^{II} \end{bmatrix}.\tag{3.71}$$

Expansion of E_z in region *II* in terms of the basis functions in regions *I* and *III* yields

$$[e^I] = [P_1][e^{II}]\tag{3.72}$$

and

$$[ee^{III}] = [P_2][ee^{II}].\tag{3.73}$$

Expansion of H_θ in regions *I* and *III* in terms of the basis functions in region *II* yields

$$[h^{II}] = [Q_1][h^I]\tag{3.74}$$

and

$$\left[hh^{II}\right] = \left[Q_2\right]\left[hh^{III}\right].$$ (3.75)

The forms of the connection matrices $\left[P\right]$ and $\left[Q\right]$ are given in [20]. Using (3.68), (3.69), (3.71), and (3.72) to (3.75) all the coefficients except e^{II} and ee^{II} can be eliminated to give

$$\left[W(k)\right]\begin{bmatrix} e^{II} \\ \cdots \\ ee^{II} \end{bmatrix} = 0,$$ (3.76)

so that the resonant frequency is given by the solution of

$$\left|W(k)\right| = 0.$$ (3.77)

The eigenvector containing the electric field amplitudes in region II can also be found and, from them, the electric and magnetic fields throughout the cavity. Once these are known the stored energy and power dissipation can be calculated for any given gap voltage and, hence, the Q_U and R_c/Q_U can be found. This method is implemented in Worksheet 3.3. The termination of the beam hole by a conducting boundary has no effect on the results if the depth of the beam hole is large enough.

It can be shown that, if the number of terms in the series in region II is finite, while the number of terms in regions I and III tend to infinity, then the solution is an upper bound on the frequency. The reverse procedure leads to a lower bound so that the accuracy of the solution is known. It is also found that, if the numbers of terms in the series are chosen so that the smallest wavelength in each region is approximately the same, then the result converges very rapidly with increasing numbers of terms to a very accurate figure. When this method was used with 8 terms in the series representation of the fields within the gap it was found that the resonant frequency, R_c/Q_U and Q_U could be computed to accuracy better than 0.01% for a wide range of cavity shapes [20]. When the same cavities were modelled using a commercial electromagnetics code it was found that considerable care, and much greater computational time, were needed to achieve the same accuracy. Thus the method of moments is valuable for rapid computation of the properties of cavities with the shape shown in Figure 3.14, and for benchmarking calculations using computational electromagnetics.

Figure 3.15 shows the results obtained when Worksheet 3.3 was used to investigate the properties of re-entrant cavities having a resonant frequency of 3 GHz and an interaction gap of 5 mm. These can be compared with the results for a pill-box cavity with $u_0/c = 0.3$ in the third column of Table 3.4. The inner and outer radii of the drift tube were chosen to be 5 mm and 7 mm respectively. These values are typical of those encountered in practice. The height of the cavity (h) was increased in steps and the outer radius (A) adjusted so that the correct frequency was obtained. Because the increase in height increases the inductance more than it decreases the capacitance it is found that the outer radius decreases steadily. From Figure 3.15
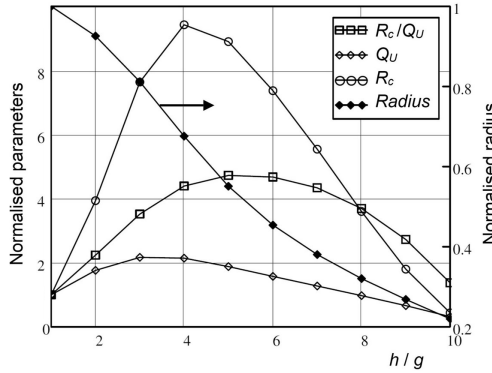
Figure 3.15: Variation in the theoretical properties of a re-entrant cavity resonator with h/g normalised to those of a pill-box cavity $(h = g)$.

we see that, as the height of the cavity increases, R_c/Q_U increases initially as the capacitance is reduced, and the inductance increased, to maintain the correct frequency. However, when the normalised height of the cavity is greater than 5, R_c/Q_U decreases because the side wall is close enough to the drift tube to increase the capacitance. This is not desirable because the increased radial component of the electric field reduces the axial component available for interaction with electrons passing through the cavity. The unloaded Q and the shunt resistance of the cavity show similar behaviour with an initial increase followed by a decrease. Figure 3.16 shows the electric field distribution in one quarter of a transverse plane of a cavity for which $h = 4g$.

Table 3.5 shows the comparison between the properties of a pill-box cavity with, and without, a beam hole, and a re-entrant cavity whose height is four times the interaction gap. These results need to be interpreted with a little caution. In the first place they only illustrate the behaviour of cavities having the dimensions chosen. Results similar to these but with differences in detail may be expected for cavities designed for other electron velocities and other frequencies. Secondly, no account has been taken of the effects of surface roughness which will reduce both Q_U and R_c in practice.

3.5.2   Fujisawa's Model of Re-entrant Cavities

The design of klystrons and inductive output tubes requires the design of cavity resonators which have a given frequency, interaction gap length, and drift tube radius. As can be seen from Figure 3.15 the height, or the radius of the cavity, may be chosen freely. The remaining parameter is then fixed by the resonant frequency. The method of moments model described in the preceding section provides a quick way of calculating the properties of a cavity. However, for parametric studies it may be faster to use an approximate equivalent circuit model [22]. In this method a doubly re-entrant cavity without a beam hole is modelled by the circuit shown in
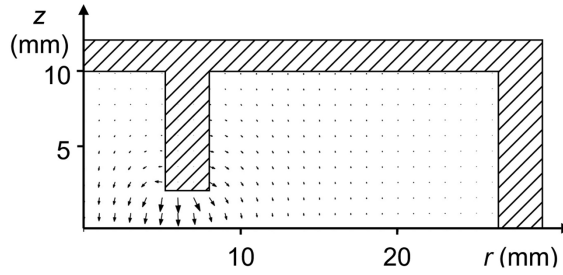
Figure 3.16: Electric field in a re-entrant cavity resonator for which $h = 4g$.

Table 3.5: Theoretical properties of copper cavities at 3 GHz ($u_0/c = 0.3$)

	Pill-box	**Pill-box with beam hole**	**Re-entrant**
a (mm)	–	5.00	5.00
a' (mm)	–	–	7.00
A (mm)	38.25	38.61	26.11
g (mm)	5.00	5.00	5.00
h (mm)	5.00	5.00	20.00
Q_U	3715	3712	7959
R_c/Q_U (Ω)	24.2	23.5	103.3
R_c (kΩ)	89.9	87.2	822.1

Figure 3.2b. The capacitor is made up of three capacitors in parallel (C_I, C_{II}, C_{III}) which represent the contribution to the capacitance from charge on the inside, the end, and the outside of the drift tube, respectively.

The capacitance associated with charge on the inside of the drift tube can be calculated to sufficient accuracy by quasi-static analysis. To do this we assume that the axial component of the electric field is constant in the gap when $r = a$. This field can be expressed in terms of an infinite set of components having sinusoidal variation $\exp(j\beta z)$ in the axial direction by taking the Fourier transform of the field at $r = a$

$$\Gamma(\beta) = \int_{-\frac{g}{2}}^{\frac{g}{2}} \frac{V_g}{g} \exp(-j\beta z)\, dz = V_g\left(\frac{\sin(\beta g/2)}{\beta g/2}\right), \tag{3.78}$$

where V_g is the gap voltage. Since we have assumed that the electric field inside the drift tube satisfies Laplace's equation we know that the radial variation of E_z is as $I_0(\beta r)$ so that, taking the inverse Fourier Transform, we find that, within the drift tube

$$E_z(r,z) = \frac{V_g}{2\pi} \int_{-\infty}^{\infty} \frac{I_0(\beta r)}{I_0(\beta a)} \cdot \frac{\sin(\beta g/2)}{\beta g/2} \exp(-j\beta z)\, d\beta. \tag{3.79}$$

The total charge within one drift tube can be found by finding the total charge on a conducting sheet at $z = 0$ for $r \leq a$

$$Q = \varepsilon_0 \int_0^a 2\pi r \, E_z(r,0) \, dr. \qquad (3.80)$$

Substituting for E_z from (3.79) and performing the integration we find that the capacitance in region I is

$$C_I = \varepsilon_0 \int_{-\infty}^{\infty} \frac{a I_1(\beta a)}{\beta I_0(\beta a)} \cdot \frac{\sin(\beta g/2)}{\beta g/2} \, d\beta. \qquad (3.81)$$

It can be shown that this capacitance is the same as that given in Figure 19 of [22]. Since we have assumed that the field in the gap is uniform the capacitance in region II is simply

$$C_{II} = \varepsilon_0 \frac{\pi(a'^2 - a^2)}{g}. \qquad (3.82)$$

The capacitance in region III is [22]

$$C_{III} = 2\varepsilon_0 a' \ln\left(\frac{e\sqrt{(A-a')^2 + (h/2)^2}}{g} \right). \qquad (3.83)$$

Thus the total capacitance is

$$C = C_I + C_{II} + C_{III}. \qquad (3.84)$$

The inductance is calculated by assuming that the azimuthal magnetic field is generated by a uniform axial current I at $r = a'$, and is negligible when $r < a'$, so that

$$H_\theta(r) = \frac{I}{2\pi r}. \qquad (3.85)$$

Then the self-inductance is

$$L = \frac{\mu_0 h}{2\pi} \ln\left(\frac{A}{a'} \right). \qquad (3.86)$$

The resonant frequency and the R/Q of the cavity can then be calculated using (3.2) and (3.6).

The stored energy is estimated from the magnetic field

$$W = \frac{\mu_0 h}{4\pi} I^2 \ln\left(\frac{A}{a'} \right). \qquad (3.87)$$

The power dissipated on the surface of the cavity is also calculated from the magnetic field

$$P_L = \frac{1}{4\pi} I^2 R_s \left\{ \frac{h}{A} + \frac{(h-g)}{a'} + 2\ln\left(\frac{A}{a'}\right) \right\}, \tag{3.88}$$

so that the unloaded Q can be calculated using (3.12).

Table 3.6 shows a comparison between the theoretical properties of re-entrant copper cavities with, and without, a beam hole calculated using the method of moments and Fujisawa's method (see Worksheet 3.4). It can be seen from these examples that, using Fujisawa's method, the accuracy of the frequency is of the order of 1% and, of the other parameters, approximately 5%. A detailed study of a wide range of cavities without beam holes showed that the accuracy of the frequency computed using Fujisawa's method is better than 5% provided that $A/a' < 5$ and $z_3/z_2 < 12$. For cavities whose shapes are typical of those used in microwave tubes the accuracy can be expected to be better than this as indicated by Table 3.6.

3.5.3   The Interaction Field

An electron beam passing along the axis of a circularly symmetrical cavity interacts with the fringing electric field of the gap in the drift tube. In the previous section it was assumed, for simplicity, that the axial component of the electric field at $r = a$ was constant. However, it is clear from Figure 3.16 that this is not true because there are field concentrations close to the drift tube noses. The field distribution depends upon the shapes of the drift tube noses as well as upon the gap length and drift tube radius. However, in order to avoid interception of electrons on the drift tube, the radius of the electron beam is normally not greater than $2a/3$. Within that radius it is found that the variation of the axial component of the electric field does

Table 3.6: Theoretical properties of re-entrant copper cavities at 3 GHz ($v_0/c = 0.3$)

	Without beam hole		With beam hole	
	Method of moments	**Fujisawa (error %)**	**Method of moments**	**Fujisawa (error %)**
a (mm)	5.0	0	5.0	5.0
a' (mm)	7.0	7.0	7.0	7.0
A (mm)	24.39	24.39	26.11	26.11
g (mm)	5.0	5.0	5.0	5.0
h (mm)	20.0	20.0	20.0	20.0
f (GHz)	3.000	2.987 (−0.4)	3.000	2.964 (−1.2)
R_c/Q_U (Ω)	99.7	93.7 (−6.4)	103.3	98.1 (−5.3)
Q_U	7499	7666 (+2.2)	7959	7934 (−0.3)
R_c (kΩ)	747.7	718.3 (−4.1)	822.1	778.1 (−5.7)

not depend strongly on the field profile at $r = a$. It is therefore possible to draw useful conclusions about the interaction field by comparing a number of approximate field profiles. In addition to the uniform field discussed above two other profiles may be considered [23]. If the drift tube noses are knife edged then

$$E_z\left(a,z\right) = \begin{cases} \dfrac{V_g}{\pi\sqrt{\left(g/2\right)^2 - z^2}} & |z| < g/2 \\ 0 & |z| \geq g/2 \end{cases}, \tag{3.89}$$

where V_g is the gap voltage. A useful approximate profile is

$$E_z\left(a,z\right) = \begin{cases} \dfrac{k\cosh\left(kz\right)}{2\sinh\left(kg/2\right)}V_g & |z| < g/2 \\ 0 & |z| \geq g/2 \end{cases}, \tag{3.90}$$

where k is chosen to match the field to that determined for the actual gap by computation or measurement. This field tends to the uniform field when $k \to 0$ and it leads to an interaction field which is very close to that given by knife-edged drift tubes when $k = 4/g$. The dependence of the interaction field on the field profile and the normalised gap length can be investigated using Worksheet 3.5. Because quasi-static fields have been assumed it is possible to normalise all dimensions to the drift tube radius. Figure 3.17 compares some typical results when $g/a = 1$ for a uniform field at $r = a$, a knife-edge field, and one which is intermediate between them. It can be seen that, in the region of space occupied by the electron beam, there is very little difference between the fields resulting from the different choices of profile. At $r = a/2$ the axial field at the centre of the gap is slightly greater than that on the axis, and it falls off a little more rapidly with increasing z. When the normalised gap length is $g/a = 0.5$ it is found that the differences between the results using the three profiles are negligible, but that there is a greater difference between the fields on the axis and at $r = a/2$. When the normalised gap length is 2.0 there is a greater
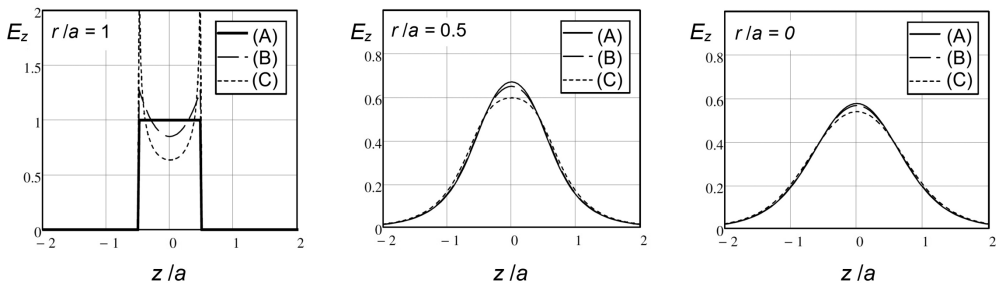

Figure 3.17: Axial variation of E_z in a gap in a drift tube when $g = a$ for three different field profiles at $r = a$: (A) uniform field, (B) hyperbolic cosine field $\left(k = 2/g\right)$, and (C) knife-edge field.

difference between the fields for the different profiles but smaller radial variation. Overall, setting $g/a = 1$ gives a good compromise between field uniformity in the radial direction and insensitivity to the field profile in the gap.

3.5.4   Practical Re-entrant Cavities

In the cavities used in practical tubes it is common for all, or part, of the outside of the drift tube to be conical, and for the tips of the drift tubes to be rounded to reduce the peak electric field on the surface. The gap length is often approximately equal to the inner radius of the drift tube but the actual dimensions are chosen to achieve a compromise between:

- the electrical properties of the cavity;
- the axial space available;
- avoiding voltage breakdown in the gap and multipactor discharges;
- thermal conduction and mechanical strength of the cavity.

The multipactor discharge is discussed in Section 18.8. The risk of gap breakdown may be further reduced by changes to the shape of the drift tube noses and by the application of coatings having low secondary electron emission coefficients [24].

If it is necessary to be able to change the frequency of a cavity this can be done by adding a tuner. This typically takes the form of a metal plate parallel to the drift tube whose position can be changed so that the capacitance between it and the drift tubes can be varied as shown in Figure 3.18(a) [25]. In some UHF klystrons and inductive output tubes the drift tube is surrounded by a ceramic cylinder, which forms the vacuum envelope of the tube as shown in Figure 3.18(b). The outer part of the cavity is then formed in two halves, which are bolted around the tube. Because the outer part of the cavity is at atmospheric pressure it is possible to make parts of the cavity wall moveable with spring fingers to maintain good electrical contact. This type of cavity is called an external cavity.

3.6   External Coupling to Cavities

The external connections to cavity resonators used in tubes can be made using either coaxial lines or metallic waveguides. The strength of the coupling is adjusted to achieve the desired external Q.

A coaxial line may be terminated in either an electric, or a magnetic, antenna. An electric antenna is made by removing a short section of the outer conductor and the insulation so that the exposed central conductor couples to the electric field in the cavity. A magnetic antenna is made in a similar manner by forming the exposed central conductor into a loop which may be connected to the outer conductor (see Figure 3.19(a) and (b)). The loop is inductively coupled to the component of the magnetic field in the cavity that is normal to the plane of the loop. Both of these techniques are used to make probes for experimental measurements
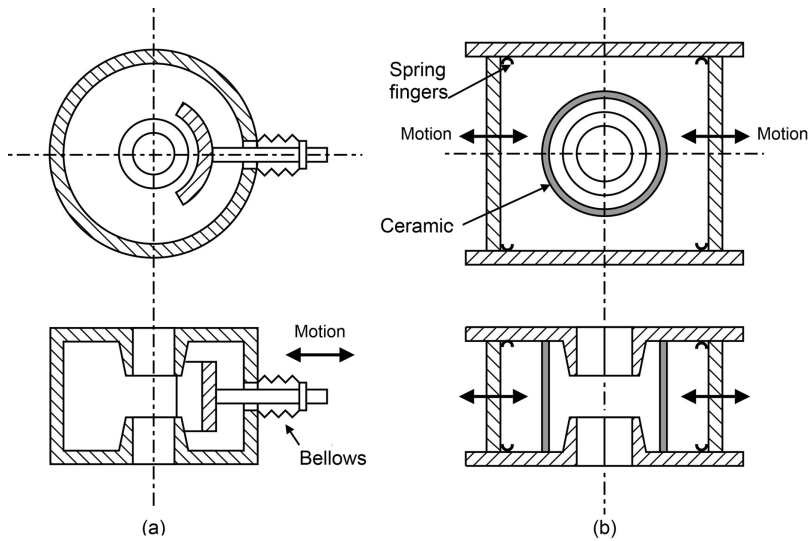
Figure 3.18: The arrangement of cavity tuners in: (a) internal cavities and (b) external cavities.
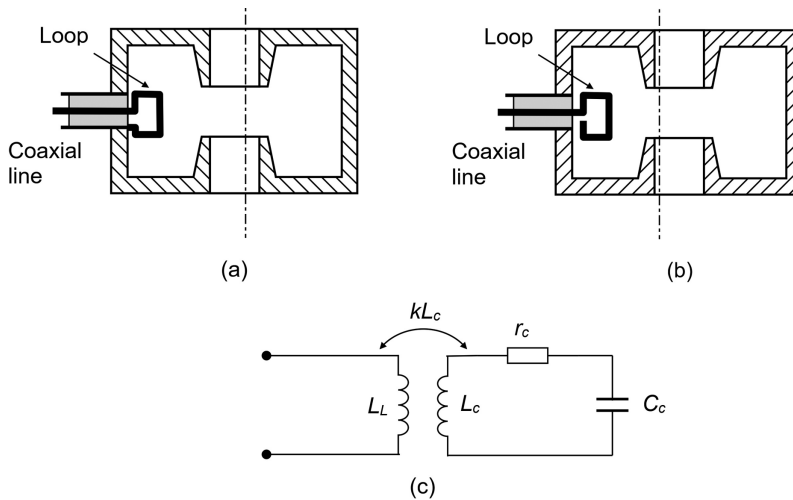
Figure 3.19: Loop-coupled cavity: (a) arrangement with a closed loop, (b) arrangement with an open loop, and (c) equivalent circuit.

on cavity resonators. The strength of the coupling between the antenna and the mode is affected by

- the size of the antenna;
- the depth of insertion of the antenna into the cavity;
- the angle between the plane of the loop of a magnetic antenna, and the direction of the local magnetic field.

Electric antennas are unsuitable for coupling into the fundamental TM_{010} mode of a cavity in a tube because the region of greatest electric field is inaccessible, and because the electric field concentration on the antenna could cause electric breakdown. Coupling using a loop (magnetic antenna) is discussed in Section 3.6.1.

When the external connection to a cavity is made using a waveguide it is arranged so that the field of the fundamental mode in the waveguide couples strongly to the cavity mode which is to be excited. Thus the broad walls of a rectangular waveguide are normally perpendicular to the axis of the cavity (see Figure 3.22). The strength of the coupling is adjusted by changing the size of the aperture (iris) where the waveguide meets the cavity. Iris coupling is discussed in Section 3.6.2.

3.6.1   Loop Coupling

Figure 3.19 shows the general arrangement of a cavity with a coupling loop. The loop may either be closed as shown in Figure 3.19(a) or open as shown in Figure 3.19(b). The loop is coupled inductively to the magnetic field in the cavity and the coupling may be adjusted by rotating the loop around the axis of the coaxial line. The equivalent circuit of a loop-coupled cavity is shown in Figure 3.19(c) where the parameters of the cavity have suffix c. The mutual inductance between the cavity and the loop is kL_c where k is the fraction of the magnetic flux within the cavity which is coupled to the loop. The self inductance of the loop is L_L. Analysis of the equivalent circuit shows that the input impedance is

$$Z_{in} = j\omega L_L + \frac{k^2 R_c}{1 + jQ_U\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)}, \tag{3.91}$$

where, from (3.6), $R_c = Q_U^2 r_c$. Comparison between this result and (3.5) shows that, apart from the effect of the loop inductance, the coupling can be represented by an ideal transformer having ratio $k:1$.

The coupling factor k can be calculated to good accuracy, if the distribution of the magnetic field in the cavity is known, by finding the fraction of the flux which is linked to the loop [26]. The self inductance of the loop can be calculated by static analysis if the dimensions of the loop are small compared with the free-space wavelength. Formulae for the self inductance of a loop in free space have been given for a rectangular loop [26] and for a circular loop [3]. The self-inductance can also be calculated by treating the loop as a short-circuited two-wire line. The results obtained in these ways agree with one another to around 20%. The self inductance of a loop can be measured by noting that the second term in (3.91) is negligible at a frequency remote from the resonant frequency of the cavity. It is found that the inductance measured in this way is up to an order of magnitude smaller than the free-space inductance. This is because the magnetic field of the loop is confined within the cavity which increases the reluctance of the flux path and reduces the flux generated by unit current. A very crude estimate of this effect can be made by

supposing that the flux line passing through the centre of a circular loop forms a circle having the same radius as the loop. The flux density at the centre of the loop is then reduced by a factor of π compared with that of the same loop in free space.

The model described above is implemented in Worksheet 3.4 which can be used to explore the properties of re-entrant cavities with loop coupling. Figure 3.20 shows the reflection coefficient S_{11} for a typical cavity if the loop inductance is neglected, and when it is reduced by a factor of π from the value calculated by treating the loop as a short-circuited two-wire line. This estimate of the loop inductance gave good agreement with experimental results in a particular case. It can be seen that the effect of the loop inductance is to shift the resonant frequency of the cavity slightly and also to change the magnitude of Z_{in} at resonance. To investigate this effect we assume that the unloaded Q of the cavity is high so that the bandwidth is narrow, and the loop reactance can be assumed to be constant. Figure 3.21 shows the equivalent circuit redrawn with an ideal transformer and constant loop reactance $X = \omega_0 L_L$ connected to a source with impedance R_S. This circuit can be put in the form shown in Figure 3.5 by writing

$$\frac{1}{R_E} + jB_E = \frac{k^2}{(R_S + jX)} = \frac{R_S - jX}{(R_S^2 + X^2)}k^2. \tag{3.92}$$
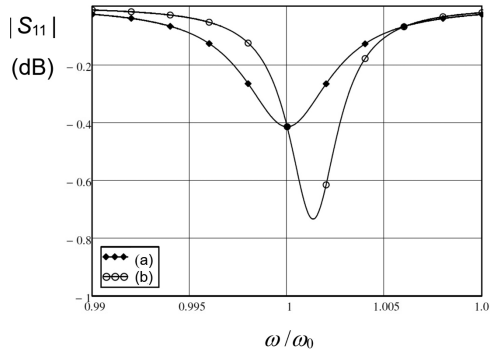

Figure 3.20: Reflection coefficient of a loop-coupled cavity $(Q_U = 7800, Q_E = 330)$: (a) ignoring the loop inductance, and (b) including the loop inductance.
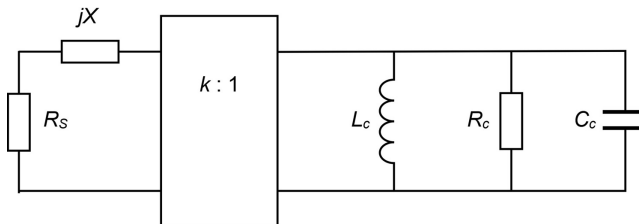
Figure 3.21: Equivalent circuit of a loop-coupled cavity with an external resistance.

so that

$$R_E = \frac{\left(R_S^2 + X^2\right)}{k^2 R_S} \tag{3.93}$$

and

$$B_E = -\frac{k^2 X}{\left(R_S^2 + X^2\right)}. \tag{3.94}$$

The external Q of the cavity and the coupling factor K can then be found from (3.18) and (3.19). The condition for resonance is that the sum of the reactances should be zero. Thus the resonant frequency of the cavity with external loading is the solution of

$$\omega C_c - \frac{1}{\omega L_c} - \frac{k^2 X}{\left(R_S^2 + X^2\right)} = 0. \tag{3.95}$$

3.6.2 Iris Coupling

Figure 3.22 shows the arrangement of a cavity coupled to a rectangular waveguide through an iris. The width of the waveguide is normally that of a standard waveguide, but the height may be reduced. The coupling factor is generally adjusted by changing the width of the aperture to obtain the desired external Q.

The coupling between a waveguide and a cavity can be modelled using the theory of coupling by small holes [27–30]. In this theory the effect of a small hole is represented by an electric dipole whose magnitude is proportional to the normal component of the electric field on the iris, and by two magnetic dipoles whose magnitudes are proportional to the transverse components of the magnetic field. It is

Figure 3.22: Arrangement of an iris-coupled cavity.

convenient to consider the case in which power is transferred from the cavity to the waveguide through the aperture. In the case of the TM_{010} mode in the cavity shown in Figure 3.22 the only non-zero field component at the iris is the magnetic field parallel to the broad wall of the waveguide. The magnetic dipole moment is then
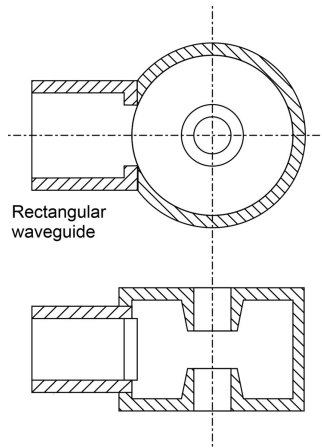
$$M = \alpha_m H_1, \tag{3.96}$$

where H_1 is the tangential magnetic field of the cavity at the centre of the aperture and α_m is the magnetic polarisability of the aperture that depends only upon its shape and size. Since the hole has been assumed to be small the polarisability can be determined by quasi-static analysis. Analytical expressions for α_m are available for circular and elliptical apertures [30], and graphs for apertures of other shapes have been found experimentally [31]. The amplitude of the transverse magnetic field induced in the waveguide is given by

$$H_0 = \frac{2\beta_g}{ab} \alpha_m H_1 \tag{3.97}$$

if the coupling hole is aligned with the centre of the waveguide [30]. The amplitude of the voltage between the centres of the broad walls of the waveguide is

$$V_0 = E_0 b = \frac{\beta_w}{\beta_g} Z_w H_0 b \tag{3.98}$$

from (2.38) where $\beta_w = \omega/c$ and $Z_w = \sqrt{\mu_0/\varepsilon_0}$. Now the coupling factor can be written

$$K = \frac{P_E}{P_L} = \frac{V_0^2}{2Z_g P_L}, \tag{3.99}$$

where P_E and P_L are the powers dissipated in the external resistance and in the cavity walls. Then, substituting from (3.97) and (3.98) we obtain

$$K = \frac{2}{Z_g} \left(\frac{Z_w \beta_w}{a} \right)^2 \frac{(\alpha_m H_1)^2}{P_L}. \tag{3.100}$$

Since we are working in terms of power and voltage the appropriate waveguide impedance is Z_{PV} given by

$$Z_{PV} = 2\frac{b}{a} \cdot \frac{\beta_w}{\beta_g} Z_w \tag{3.101}$$

so that

$$K = \frac{\beta_g \beta_w}{ab} \cdot \frac{H_1^2 Z_w}{P_L} \alpha_m^2. \tag{3.102}$$

Substituting for the power loss in the cavity from (3.12)

$$K = \frac{2\beta_g}{ab} \cdot \frac{\mu_0 H_1^2 Q_U}{2W} \cdot \alpha_m^2.$$

(3.103)

The coupling factor is seen to be the product of three terms: the first depends only on the properties of the waveguide; the second only on the properties of the cavity; and the third depends only on the properties of the aperture. The first term can be written as $4\pi/ab\lambda_g$ which is just 4π divided by the volume of the waveguide in one guide wavelength. The second term relates the magnitude of the tangential magnetic field at the aperture to the stored energy in the cavity. For a pill-box cavity H_1 and W are given by (3.43) and (3.47) so that

$$\left(\frac{\mu_0 H_1^2}{2W} \right)_{\text{Pill box}} = \frac{1}{\pi r_C^2 h},$$

(3.104)

where r_C is the radius of the cavity. We note that the right-hand side of (3.104) is the reciprocal of the volume of the cavity. Similarly, for a rectangular cavity whose width is d_x and length is d_y, H_1 and W are given by (3.59) and (3.62), and

$$\left(\frac{\mu_0 H_1^2}{2W} \right)_{\text{Rectangular}} = \frac{4}{d_x d_y h} \left(\frac{d_x^2}{d_x^2 + d_y^2} \right).$$

(3.105)

When the cavity is square the right-hand side of (3.105) reduces to 2 divided by the volume of the cavity. In general we note that the ratio $2W/\mu_0 H_1^2$ is the volume in which the stored magnetic energy of a uniform field H_1 is equal to the stored energy in the cavity [32].

For a small circular aperture of radius r_A

$$\alpha_m = \frac{4}{3} r_A^3,$$

(3.106)

which is the volume of a sphere which fills the aperture divided by π [28]. Thus the coupling factor can be expressed as

$$K = \left(\frac{4\pi}{ab\lambda_g} \cdot \frac{4}{3} r_A^3 \right) \cdot \left(\frac{\mu_0 H_1^2}{2W} \cdot \frac{4}{3} r_A^3 \right) Q_U.$$

(3.107)

The first bracket can be interpreted as the shunt reactance of a circular iris in a rectangular waveguide and the second as the coupling coefficient when two identical resonators are coupled though a circular iris. Both brackets are expressed as ratios of effective volumes [32].

It is helpful to consider a numerical example. Suppose that a pill-box cavity whose resonant frequency is 3 GHz, having the properties given in Table 3.1, is coupled to standard WR284 waveguide by a circular hole. The internal dimensions of the waveguide are 72 mm × 34 mm. The hole is taken to have radius 7.2 mm, which

is about as large is it can be while satisfying the assumption that it is small enough for the tangential magnetic field to be constant over it. The guide wavelength is 138 mm if the slight detuning of the cavity by the coupling is neglected. Hence

$$Q_E = \frac{\lambda_g ab}{2\pi} \cdot \frac{\pi r_C^2 h}{2} \cdot \left(\frac{3}{4 r_A^3}\right)^2 = 18700. \tag{3.108}$$

Now $Q_U = 16100$ so the coupling factor is 0.86. The use of this method to calculate the coupling coefficient between a cavity and a waveguide in the range 0 to 1.1 is described in [29]. The input and output couplings used in microwave tubes are generally strongly over-coupled so that the external Q is much less than the unloaded Q. For output coupling this is necessary to minimise the proportion of the output power which is dissipated in the cavity losses. Thus the coupling apertures which are required cannot be regarded as small. In particular their properties are modified from those of ideal, isolated, apertures by the proximity of the walls of the cavity and the waveguide and by the finite time taken for waves to propagate across them.

Useful insight can be gained into the properties of a rectangular cavity coupled to a waveguide by a large inductive iris by studying the arrangement shown in Figure 3.23(a). Inductive irises are commonly used because the risk of voltage breakdown is less than with capacitive irises. This problem can be modelled by the normalised transmission line circuit shown in Figure 3.23(b). The losses in the cavity are represented by a normalised lumped conductance g at the centre of the cavity where

$$g = \frac{Z_g}{(R/Q)Q_U}. \tag{3.109}$$

It can be shown that this corresponds to the usual definition of R/Q if the waveguide impedance is taken to be Z_{PV}. The susceptance of the iris is modelled to sufficient accuracy for illustration by (2.97)

$$b = -\frac{2\pi}{\beta_g a} \cot^2\left(\frac{\pi w}{2a}\right). \tag{3.110}$$
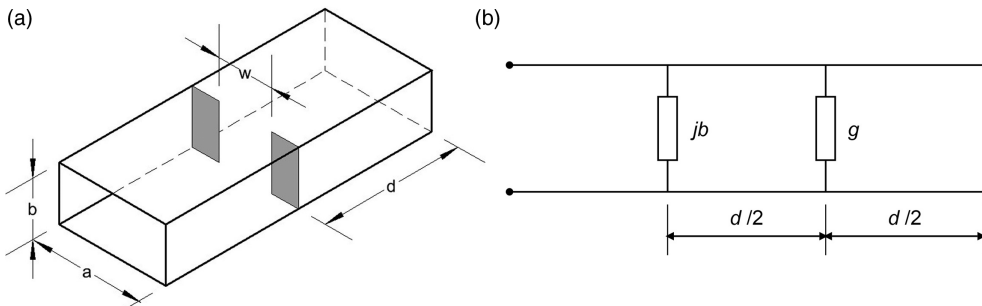

Figure 3.23: (a) Arrangement of a rectangular cavity coupled to a waveguide by an inductive iris, and (b) transmission line equivalent circuit.

The input impedance of this circuit can be found using standard transmission line theory and the perturbed resonant frequency is given by the minimum of S_{11}. The plot of the input impedance on a Smith chart is a perfect circle and the coupling parameter and loaded Q can be found using the method described by Kajfez [33]. When the unloaded Q is computed from these results it is found to be very close to that set in (3.109) so confirming the self-consistency of the model (see Worksheet 3.6). Figure 3.24 shows a typical plot with markers set to the resonant frequency and to the 3dB points of the loaded resonator.

We saw in Section 2.4.3 that an inductive iris of any width can be modelled by an equivalent circuit. The normalised shunt reactance of the iris, in ratio of volumes form, is found from (2.103) and (2.104) to be

$$x = \frac{1}{4} \cdot \frac{4\pi}{ab\lambda_g} \cdot \frac{\pi w^2 b}{4\left(1 - w^2/a^2\right)^2} \cdot \frac{\sin\left(\pi w/2a\right)}{\pi w/2a}. \tag{3.111}$$

Comparing this expression with the first bracket in (3.107) we see that the polarisability of a wide aperture which occupies the full height of the waveguide is given by

$$\alpha_m = \frac{1}{4} \cdot \frac{\pi w^2 b}{4\left(1 - w^2/a^2\right)^2} \cdot \frac{\sin\left(\pi w/2a\right)}{\pi w/2a}. \tag{3.112}$$

When $w/a \ll 1$ this reduces to

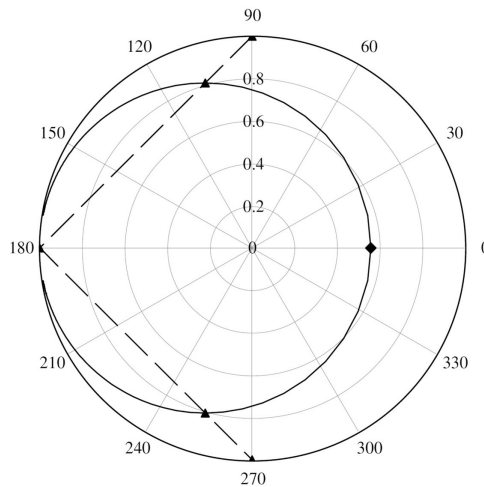$$\alpha_m = \frac{1}{4} \cdot \frac{\pi w^2 b}{4}, \tag{3.113}$$

Figure 3.24: Polar plot (Smith chart) of the reflection coefficient of a rectangular cavity coupled to a waveguide by an inductive iris showing the resonant frequency (♦) and the 3 dB points (▲).
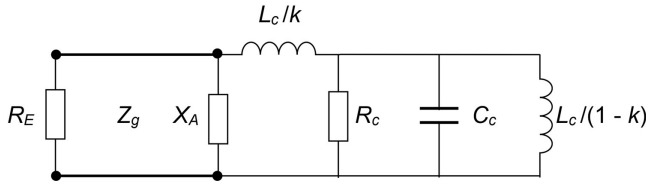
Figure 3.25: Equivalent circuit of an iris-coupled cavity.

agreeing with the formula for a narrow aperture in [32]. Substituting in (3.107) from (3.105) and (3.112) gives a closed-form expression for the coupling factor.

$$ K = \frac{1}{2}\beta_g a \left(\frac{w/a}{2\left(1 - w^2/a^2\right)^2} \cdot \sin\left(\pi w/2a\right) \right)^2 \cdot \frac{a}{d}\left(\frac{a^2}{a^2 + d^2} \right) Q_U. \qquad (3.114) $$

The coupling factor computed from the distributed circuit model is almost exactly proportional to $\left(w/a\right)^4$ up to $w/a = 0.6$. The results computed from (3.114) agree with this very closely up to $w/a = 0.4$ (see Model 1 in Worksheet 3.6). In principle this method can be used to find the loaded Q of any cavity but formulae are not available, in general, for the polarizability of rectangular apertures taking account of adjacent boundaries.

   The lumped-element equivalent circuit of an iris-coupled cavity is shown in Figure 3.25 where the components with suffix c represent the properties of the cavity. The waveguide has characteristic impedance Z_g and is terminated by an external load which we shall assume is matched to the waveguide. The iris is represented by the reactance X_A, computed using (3.110). By comparing the input impedance of this network when $\omega = \omega_0$ with that of the distributed circuit in Figure 3.23(b) it can be shown that, when the cavity is square, $k = 1/\pi$. This result is independent of the reactance of the aperture, or of the choice of characteristic impedance of the waveguide. It is tempting to interpret k as the fraction of the circulating current which is intercepted by the aperture. But considerations of symmetry would then require $k = 1/4$ (since the reactance of the aperture has already been adjusted for the fraction of the longitudinal current in the waveguide intercepted by it). It is found that there is good agreement between this model with $k = 1/\pi$ (Model 2 in Worksheet 3.6) and the distributed circuit model described above. The equivalent circuit model is useful because it is valid for any cavity if the values of k and X_A can be found.

3.7   Measurement of Cavity Parameters

Although it is possible to compute the properties of any microwave cavity to good accuracy using computational electromagnetics, it is desirable to confirm the results by cold test measurements. The parameters to be measured include the resonant frequency, R/Q, unloaded Q and external Q. In addition it may be useful to map the electric field intensity and to identify higher-order modes. The measurements

may be made on cavities at full scale or scaled up to a larger size for convenience. The cavities may be made of copper or brass. If the cavities are bolted together, rather than being brazed, it is necessary to ensure that there is a good electrical contact on all joints that intercept the current circulating within the cavity. It is therefore important to ensure that the mating surfaces are a close fit to one another. The author has obtained good results using inexpensive aluminium cavities with conducting grease (obtainable from car accessory shops) on the mating surfaces.

The resonant frequency is measured by coupling into the cavity with small electric or magnetic probes and examining either the reflection (S_{11}) or transmission (S_{21}) using a vector, or scalar, analyser. The transmission measurement is to be preferred because it is easier to detect the frequency of the resonant peak accurately. Probes are readily made from prefabricated lengths of semi-rigid coaxial cable with SMA connectors at each end. The cable is cut in half and a short length of the outer conductor and the insulation removed to make an electric probe. A magnetic probe is made in a similar manner by exposing a greater length of the central conductor, bending it to form a loop and soldering the free end to the outer conductor of the line. We saw in Section 3.6.1 that the resonant frequency of the cavity is perturbed by external coupling. It is therefore very important to make the coupling as weak as possible to avoid this effect. The positions of the probes are adjusted to give the greatest response for which there is no measurable perturbation of the frequency. If necessary the frequency can be measured as the probes are gradually inserted and the results extrapolated to zero insertion depth.

The R/Q can be determined by measuring the change in frequency when a thin dielectric rod is inserted on the axis of the cavity [34]. This method was originally proposed by Slater [35]. It is assumed that, if the perturbation is small, then the fields in the cavity outside the perturbing rod are unchanged. If a pill-box cavity of height h is perturbed by a rod having radius r and relative permittivity ε_r then the change in the capacitance of the cavity caused by the rod is

$$\Delta C = \frac{\varepsilon_0 \left(\varepsilon_r - 1\right) \pi r^2}{h}. \tag{3.115}$$

The perturbed resonant frequency is then

$$\omega = \frac{1}{\sqrt{L\left(C + \Delta C\right)}} = \omega_0 \left(1 - \frac{\Delta C}{2C}\right). \tag{3.116}$$

Now the stored energy in the cavity can be written

$$W = \frac{1}{2}CV^2, \tag{3.117}$$

where V is the effective voltage across the cavity. Hence the frequency perturbation can be written

$$\omega - \omega_0 = -\frac{\varepsilon_0 \left(\varepsilon_r - 1\right) \pi r^2 \omega_0}{h} \frac{V^2}{4W}. \tag{3.118}$$

Making use of the definition of R/Q in (3.14)

$$\omega - \omega_0 = -\frac{\varepsilon_0 \left(\varepsilon_r - 1\right) \pi r^2 \omega_0^2}{2h} \left(\frac{R}{Q}\right),$$ (3.119)

enabling R/Q to be found from a measurement of the frequency perturbation. This measurement requires accurate knowledge of the relative permittivity of the perturbing rod. That can be determined by using the rod to perturb a cavity whose properties are known theoretically. It is obviously important for the frequency shift to be large enough to be measured accurately. There is, therefore, a conflict between the requirements for the perturbation to be small enough for the theory to be valid and for it to be large enough to be measured accurately. Equation (3.118) can be rewritten as

$$\omega - \omega_0 = -\frac{\left(\varepsilon_r - 1\right)\omega_0}{2W} \cdot \frac{\pi r^2 h \, \varepsilon_0 E_z^2}{2}$$ (3.120)

in which the second term can be recognised as the stored electric energy in the volume occupied by the rod. More generally, Slater's perturbation theorem states that the change in the resonant frequency is proportional to the change in the stored energy and, thus, to the square of the magnitude of the local electric field. That makes it possible for the electric field to be mapped within a cavity by measuring the perturbation caused by a small bead. The electric field components can be measured individually by using a bead in the shape of a thin rod. A study of the accuracy of perturbation methods showed that the assumptions on which they are based cease to be valid for quite small perturbing objects [36]. However, the errors from different sources are of opposite signs so that the method gives useful accuracy even when the rod is not very small. The frequency of a pill-box cavity perturbed by a dielectric rod was calculated exactly and compared with the result from (3.119). It was found that the error in the method was about 0.3% when the ratio of the radius of the rod to that of the cavity was 0.05 and $\varepsilon_r = 10$. The same paper also examined the perturbation of the frequency of a pill-box cavity by a small dielectric sphere.

Higher-order modes can be identified by probing the field of the cavity with a small dielectric rod inserted through holes drilled in the cavity wall. The strength of the perturbation produced by a given depth of insertion of the rod provides an indication of the local strength of the electric field. It is therefore possible to detect the zeroes and the maxima of the magnitude of the electric field, and hence to identify the mode for each resonant frequency.

The loaded and unloaded Q of a cavity resonator can be measured accurately using the methods described by Kajfez [33]. The input impedance of the cavity is determined using a vector analyser with a result similar to that shown in Figure 3.24. The coupling factor is given by

$$K = \frac{D}{2 - D},$$ (3.121)

where $D\left(0 \le D \le 2\right)$ is the diameter of the circle which is the locus of the impedance on the Smith chart. The frequencies $\left(\omega_{1,2}\right)$ are found at the two points indicated by triangular markers in Figure 3.24 which subtend an angle of 90° at the anti-resonant point and then

$$Q_L = \frac{\omega_0}{\omega_1 - \omega_2}. \tag{3.122}$$

The unloaded Q of the cavity can then be determined using (3.20).

References

[1] R. G. Carter, *Electromagnetic Waves: Microwave Components and Devices.* London: Chapman and Hall, 1990.

[2] N. W. McLachlan, *Bessel Functions for Engineers*. Oxford University Press, 1954.

[3] S. Ramo *et al.*, *Fields and Waves in Communication Electronics*. New York: Wiley, 1965.

[4] S. Lucyszyn, 'Investigation of anomalous room temperature conduction losses in normal metals at terahertz frequencies', *IEE Proceedings: Microwaves, Antennas and Propagation,* vol. 151, pp. 321–329, 2004.

[5] E. Episkopou *et al.*, 'Defining material parameters in commercial EM solvers for arbitrary metal-based THz structures', *IEEE Transactions on Terahertz Science and Technology,* vol. 2, pp. 513–524, 2012.

[6] S. P. Morgan, 'Effect of surface roughness on eddy current losses at microwave frequencies', *Journal of Applied Physics,* vol. 20, pp. 352–362, 1949.

[7] C. L. Holloway and E. F. Kuester, 'Power loss associated with conducting and superconducting rough interfaces', *IEEE Transactions on Microwave Theory and Techniques,* vol. 48, pp. 1601–1610, 2000.

[8] L. Tsang *et al.*, 'Effects of random rough surface on absorption by conductors at microwave frequencies', *IEEE Microwave and Wireless Components Letters,* vol. 16, pp. 221–223, 2006.

[9] E. O. Hammerstad and F. Bekkedal, *Microstrip Handbook*. Trondheim, Norway: University of Trondheim, 1975.

[10] T. Edwards, *Foundations for Microstrip Circuit Design*, 2nd ed. Chichester, UK: Wiley, 1992.

[11] E. Maxwell, 'Conductivity of metallic surfaces at microwave frequencies', *Journal of Applied Physics,* vol. 18, pp. 629–638, 1947.

[12] J. Allison and F. Benson, 'Surface roughness and attenuation of precision-drawn, chemically polished, electropolished, electroplated and electroformed waveguides', *Proceedings of the IEE-Part B: Radio and Electronic Engineering,* vol. 102, pp. 251–259, 1955.

[13] F. Benson, 'Waveguide attenuation and its correlation with surface roughness', *Proceedings of the IEE-Part III: Radio and Communication Engineering,* vol. 100, pp. 85–90, 1953.

[14] F. Benson, 'Attenuation and surface roughness of electroplated waveguides', *Proceedings of the IEE-Part III: Radio and Communication Engineering,* vol. 100, pp. 213–216, 1953.

[15] F. Benson and D. Steven, 'Rectangular-waveguide attenuation at millimetre wavelengths', *Proceedings of the Institution of Electrical Engineers,* vol. 110, pp. 1008–1014, 1963.

[16] F. J. Tischer, 'Excess conduction losses at millimeter wavelengths', *IEEE Transactions on Microwave Theory and Techniques,* vol. 24, pp. 853–858, 1976.

[17] M. P. Kirley and J. H. Booske, 'Terahertz conductivity of copper surfaces', *IEEE Transactions on Terahertz Science and Technology,* vol. 5, pp. 1012–1020, 2015.

[18] F. Horner *et al.*, 'Resonance methods of dielectric measurement at centimetre wavelengths', *Journal of the Institution of Electrical Engineers – Part III: Communication Engineering, including the Proceedings of the Wireless Section of the Institution,* vol. 93, pp. 53–68, 1946.

[19] R. G. Carter *et al.*, 'Rapid calculation of the properties of klystron cavities', in *2008 IEEE International Vacuum Electronics Conference*, pp. 142–143, 2008.

[20] R. G. Carter *et al.*, 'Calculation of the properties of reentrant cylindrical cavity resonators', *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 2531–2538, December 2007.

[21] R. Taylor, 'Calculation of resonant frequencies of re-entrant cylindrical electromagnetic cavities', *Journal of Nuclear Energy, Part C: Plasma Physics,* vol. 3, pp. 129–134, 1961.

[22] K. Fujisawa, 'General treatment of klystron resonant cavities', *IRE Transactions on Microwave Theory and Techniques,* vol. MTT-6, pp. 344–358, 1958.

[23] H. G. Kosmahl and G. M. Branch, Jr., 'Generalized representation of electric fields in interaction gaps of klystrons and traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 20, pp. 621–629, 1973.

[24] C. Hill and R. G. Carter, 'Investigation of possible multipactor discharge in a klystron input cavity', in *2006 IEEE International Vacuum Electronics Conference held Jointly with 2006 IEEE International Vacuum Electron Sources*, Monterey, CA, pp. 81–82, 2006.

[25] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, UK: Research Studies Press, 1995.

[26] K. R. Engala *et al.*, 'Simple computation of the coupling coefficient for loop-coupled resonant cavities', *Microwave and Optical Technology Letters,* vol. 27, pp. 400–404, 2000.

[27] H. A. Bethe, 'Theory of diffraction by small holes', *Physical Review,* vol. 66, pp. 163–182, 1944.

[28] R. E. Collin, *Foundations for Microwave Engineering*. New York: McGraw-Hill, 1966.

[29] J. Gao, 'Analytical formula for the coupling coefficient β of a cavity-waveguide coupling system', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 309, pp. 5–10, 1991.

[30] J. Gao, 'Analytical formulae for the coupling coefficient β between a waveguide and a travelling wave structure', in *Proceedings of the 1993 Particle Accelerator Conference (PAC 1993)*, pp. 868–870, 1993.

[31] S. B. Cohn, 'Determination of aperture parameters by electrolytic-tank measurements', *Proceedings of the IRE,* vol. 39, pp. 1416–1421, 1951.

[32] H. A. Wheeler, 'Coupling holes between resonant cavities or waveguides evaluated in terms of volume ratios', *IEEE Transactions on Microwave Theory and Techniques,* vol. 12, pp. 231–244, 1964.

[33] D. Kajfez and E. J. Hwan, 'Q-factor measurement with network analyzer', *IEEE Transactions on Microwave Theory and Techniques,* vol. 32, pp. 666–670, 1984.

[34] J. F. Gittins, *Power Travelling-Wave Tubes*. London: English Universities Press, 1965.

[35] L. C. Maier, Jr. and J. C. Slater, 'Field strength measurement in resonant cavities', *Journal of Applied Physics,* vol. 31, pp. 68–77, 1952.

[36] R. G. Carter, 'Accuracy of microwave cavity perturbation measurements', *IEEE Transactions on Microwave Theory and Techniques,* vol. 49, pp. 918–923, May 2001.

4   Slow-Wave Structures

4.1   Introduction

The useful bandwidth of tubes based on the resonators described in the previous chapter is limited to a few percent. The electric field which interacts with the electron beam is essentially a standing wave restricted to a small region of space. The R/Q of the resonators is typically of the order of 100 Ω and the shunt impedance is proportional to Q. Thus if Q is reduced to increase the bandwidth, the gap voltage induced by a given RF beam current decreases. It eventually falls to a point where the interaction with the beam is of no practical value. When wide bandwidth is required it is necessary to adopt a different approach and arrange for the interacting electric field to be in the form of a travelling wave whose phase velocity is approximately synchronous with the electron velocity. The interaction field is smaller than is the case with a resonator but a strong interaction can be achieved by maintaining synchronism over a much longer distance that the interaction gap of a cavity resonator. This chapter deals with electromagnetic structures that support electromagnetic waves having the properties required for interaction with an electron beam. These are known as *slow-wave structures*. Many types of slow-wave structure are used in microwave tubes and many more types have been considered for potential use [1, 2]. There is a very extensive literature on the subject: these two papers cite over 300 references. It is therefore only possible to consider the general principles, and some of the more important types of structure, in the space available in this book. The basic principles of uniform and periodic structures are discussed in the subsections of this section. Practical slow-wave structures of various types are considered in Sections 4.2 to 4.6. The final section deals with methods for measuring the properties of slow-wave structures.

4.1.1   Uniform Slow-Wave Structures

The essential property of any useful slow-wave structure is that it should support an electromagnetic wave having a longitudinal component of the electric field and a phase velocity rather smaller than the velocity of light. Figure 4.1 shows the general arrangement of a cylindrical slow-wave structure. The diagram is divided into two concentric regions: the inner region $r < a$ is empty and allows for the passage of an electron beam; the outer region $r \geq a$ supports the propagation of a slow
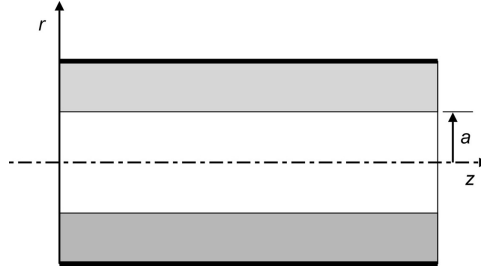
Figure 4.1: Generic slow-wave structure.

electromagnetic wave having a z-component of electric field that is non-zero on the axis. These requirements are satisfied if the structure supports a TM_{01} mode. The z component of the electric field in the inner region must then satisfy the wave equation

$$\nabla^2 E_z - \frac{1}{c^2}\frac{\partial^2 E_z}{\partial t^2} = 0. \tag{4.1}$$

In cylindrical polar coordinates this becomes

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial E_z}{\partial r}\right) + \frac{\partial^2 E_z}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E_z}{\partial t^2} = 0 \tag{4.2}$$

when azimuthal symmetry is assumed. Now let the electric field on the axis be given by

$$E_z = E_0 \exp j\left(\omega t - \beta_0 z\right), \tag{4.3}$$

where $\beta_0 = \omega/v_p$ and v_p is the phase velocity of the wave. Substituting for E_z in (4.2) gives

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial E_z}{\partial r}\right) - \left(\beta_0^2 - \frac{\omega^2}{c^2}\right)E_z = 0. \tag{4.4}$$

This is a form of Bessel's equation [3] whose solutions are modified Bessel functions. Thus the electric field within the inner region is given by

$$E_z = E_0 I_0\left(\gamma_0 r\right)\exp j\left(\omega t - \beta_0 z\right), \tag{4.5}$$

where $\gamma_0 = \sqrt{\beta_0^2 - \omega^2/c^2}$ is a real quantity because $\beta_0 > \omega/c$ for a slow wave. The Bessel function I_0 increases rapidly as its argument increases, as shown in Figure 4.2. It is therefore necessary for the electron beam to be as close as possible to the surface of the slow-wave structure to obtain a strong interaction. When the electron beam is a solid cylinder which is coaxial with the structure it is usual to choose $\gamma_0 a \leq 1$ to minimise the variation of the electric field over the cross-sectional area of the beam. It can be seen from (4.5) that E_z is non-zero when $r = a$ and, therefore, the structure in the outer region must be able to support a longitudinal electric field at its surface.
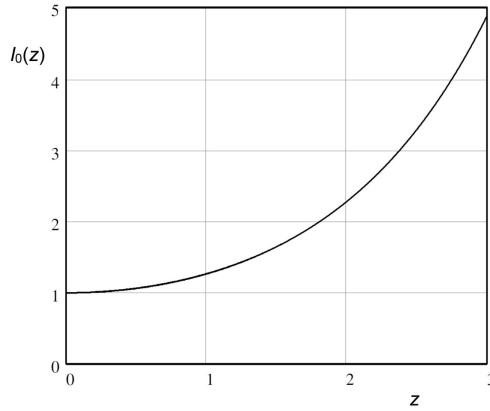
Figure 4.2: The Bessel function $I_0(z)$.

Similar conclusions are obtained when the structure has a planar geometry and the transverse variation of the electric field is a hyperbolic cosine, hyperbolic sine, or exponential function, depending upon the boundary conditions.

Further insight into the general properties of slow-wave structures can be obtained from an equivalent uniform transmission line having series inductance, and shunt capacitance, as shown in Figure 2.2(b). At a given frequency the inductance and capacitance can be chosen so that the phase velocity and the characteristic impedance have the desired values. This is the equivalent circuit of a two-wire line propagating a *TEM* wave whose phase velocity is given by

$$v_p = \frac{1}{\sqrt{LC}} = \frac{c}{\sqrt{\varepsilon_r}}, \tag{4.6}$$

where L and C are the inductance and capacitance per unit length and ε_r is the relative permittivity of the dielectric separating the conductors. Although (4.6) shows that this line can have a phase velocity less than the velocity of light it is not useful as a slow-wave structure because the longitudinal electric field is zero everywhere. The phase velocity can be reduced from that given by (4.6) by increasing the inductance per unit length. This is achieved most simply by making the inner conductor of a coaxial line in the form of a helical wire. The properties of this important type of slow-wave structure are considered in detail in Section 4.3. For our present purposes it is sufficient to note that all methods of increasing the series inductance lead to structures in which the current flow has other components besides the axial component. The magnetic field is then no longer purely transverse and the mode is not a *TEM* mode.

The characteristic impedance of the modified line is

$$Z_c = \frac{|V|^2}{2P}, \tag{4.7}$$

where V is the voltage between the conductors and P is the power flow along the line. In general, the inductance and capacitance per unit length of the modified line depend on frequency so that the line is dispersive. In that case

$$Z_c = \frac{|V|^2}{2Wv_g},\tag{4.8}$$

where W is the stored energy per unit length and v_g is the group velocity. The characteristic impedance can then be written

$$Z_c = \frac{v_p}{v_g} \cdot \frac{V^2}{2Wv_p} = \frac{v_p}{v_g}\sqrt{\frac{L}{C}} = \frac{v_p}{v_g}Z_{c0},\tag{4.9}$$

where v_p is the phase velocity and Z_{c0} is the characteristic impedance of a non-dispersive line having the same inductance and capacitance per unit length. From (4.9) it can be seen that it is desirable for the characteristic impedance to be as high as possible because that gives a larger voltage for a given power flow. Hence it is best to modify the line by increasing the inductance while keeping the capacitance small. That, in turn, means that the relative permittivity of the dielectric should also be small.

It is commonly stated that the axial electric field on the surface of the slow-wave structure is given by [4, 5]

$$E_z(a) = -\frac{\partial V}{\partial z} = j\beta_0 V.\tag{4.10}$$

However, this equation cannot be correct as it stands, because the axial component of the electric field is zero in the unmodified structure. The reason is that (4.10) does not include the contribution to the electric field from the time-varying magnetic field. When this is taken in to account we find that

$$E_z(a) = -\frac{\partial V}{\partial z} - \frac{\partial A_z}{\partial t},\tag{4.11}$$

where A_z is the magnetic vector potential which is related to the axial component of the current [6]. The magnetic vector potential also satisfies the Lorentz condition

$$\nabla \cdot \mathbf{A} = -\varepsilon\mu\frac{\partial V}{\partial t}.\tag{4.12}$$

Now the components of $\mathbf{A}$ are associated with components of the current density in the same directions. If the radial component of the current density is zero and the mode of propagation is azimuthally symmetric then (4.12) becomes

$$\frac{\partial A_z}{\partial z} = -\varepsilon\mu\frac{\partial V}{\partial t}.\tag{4.13}$$

Making use of the variation with z and t defined in (4.3)

$$A_z = \frac{\omega}{\beta_0 v_{p0}^2} V,$$ (4.14)

where v_{p0} is the phase velocity of the *TEM* wave on the unmodified structure. Substituting for A_z in (4.11) gives

$$E_z(a) = j\beta_0 \left(1 - \frac{v_p^2}{v_{p0}^2} \right) V.$$ (4.15)

It is usually the case that $v_p \ll v_{p0}$ and then (4.10) is a satisfactory approximation. We can define an impedance which relates the strength of the electric field to the power flow in the structure (P) by

$$Z_S = \frac{\left| E_z(a) \right|^2}{2\beta^2 P} = \left(1 - \frac{v_p^2}{v_{p0}^2} \right)^2 Z_c.$$ (4.16)

This impedance, which we will call the *surface impedance*, is the impedance presented to the current that is induced in the slow-wave structure by the beam. When $v_p \ll v_{p0}$ the surface impedance is equal to the characteristic impedance of the structure. An indication of the strength of the interaction between the wave on the structure and an electron beam is provided by the *Pierce Impedance* defined in terms of the electric field on the axis by

$$Z_P = \frac{\left| E_z(0) \right|^2}{2\beta_0^2 P} = \frac{Z_S}{I_0(\gamma_0 a)^2}.$$ (4.17)

A simple slow-wave structure supporting the TM_{01} mode can be made by lining a circular metallic waveguide with dielectric. Then, referring to Figure 4.1, it is necessary for the outer (dielectric) region to support a travelling electromagnetic wave. Because the velocity of this wave is less than that of light the fields in the inner (vacuum) region are evanescent and decay towards the axis as shown in Figure 4.2. Experimental devices incorporating this kind of slow-wave structure have been constructed [7]. The threshold velocity for a growing wave (Cerenkov) interaction is $c/\sqrt{\varepsilon_r}$ so that a device with a relative permittivity of the order of 10 gave maximum gain at a beam voltage of 160 kV. A device using a multi-layer photonic band-gap structure with a beam voltage of 110 kV is described in [8, 9]. This type of slow-wave structure has poor thermal dissipation making it unsuitable for continuous high power operation. There is also the possibility that static charge may accumulate on the dielectric surface leading to electric breakdown or disruption of the electron beam.

Another type of uniform slow-wave structure is the sheath helix in which dielectric material in the outer region supports a helically conducting surface at $r = a$. Such a material could be manufactured by arranging insulated metal wires side by side in a strip and then winding the strip helically around a cylindrical mandrel.

This structure, which is a useful model for the helix slow-wave structure, is discussed in detail in Section 4.3.1.

4.1.2   Periodic Slow-Wave Structures

The previous section has shown that uniform slow-wave structures are of limited use. All the types of slow-wave structure which are used in microwave tubes are periodic and the interaction space is bounded by sections of metal separated by gaps as shown in Figure 4.3. The gaps have length g, the period of the structure is p, and the electric fields in the gaps are determined by the wave propagating in the structure to which they are coupled. The details of practical slow-wave structures are discussed in later sections. Here we restrict our attention to the general properties of structures of this kind. Thus no assumptions are made about the region outside the conducting rings other than that it is capable of supporting a travelling electromagnetic wave with a propagation constant β_0 at frequency ω.

Any periodic structure can be represented at a single frequency by a uniform transmission line, having propagation constant β_0 and characteristic impedance Z_c, which is tapped at regular intervals p as shown in Figure 4.4(a). A general external connection to the line can be represented by an ideal transformer whose ratio is $1 : N$ together with a lumped susceptance Z_3 which represents the storage of electromagnetic energy in the junction [10]. Figure 4.4(b) shows a single junction in this structure in which the signs of the voltages and currents on the transmission line are chosen so that the cells can be cascaded. It is convenient to represent this by the hybrid transfer matrix

$$\begin{bmatrix} V_2 \\ I_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -Y_3 & 1 & N \\ N & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \\ I_3 \end{bmatrix} = [G] \begin{bmatrix} V_1 \\ I_1 \\ I_3 \end{bmatrix}, \tag{4.18}$$

where $Y_3 = 1/Z_3$. In this representation the circuit is driven by the voltage and current at port 1 and the current induced by the electron beam at port 3. The outputs are the voltage and current at port 2 and the voltage acting on the beam at port 3.
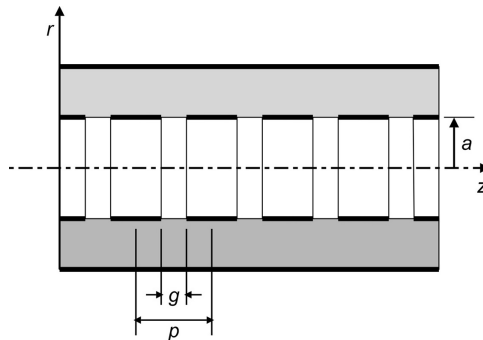

Figure 4.3: Arrangement of a periodic slow-wave structure.

Figure 4.4: Periodic slow-wave structure: (a) general equivalent circuit, and (b) a single junction (after Figure 4.24 in [10]).

The properties of a single cell of the structure, such as that between the planes 1 and 2 in Figure 4.4(a), can be found by noting that the voltage and current on the line are related to the amplitudes of the forward and backward waves by

$$\begin{bmatrix} V \\ I \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ Y_c & -Y_c \end{bmatrix} \begin{bmatrix} V_f \\ V_b \end{bmatrix}. \tag{4.19}$$

The inverse of this equation is

$$\begin{bmatrix} V_f \\ V_b \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & Z_c \\ 1 & -Z_c \end{bmatrix} \begin{bmatrix} V \\ I \end{bmatrix}. \tag{4.20}$$

Then, from (4.18), (4.19), and (4.20)

$$\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & Z_c & 0 \\ 1 & -Z_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -Y_3 & 1 & N \\ N & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ Y_c & -Y_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \\ I_3 \end{bmatrix}, \tag{4.21}$$

which can be evaluated as

$$\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_3 \end{bmatrix} = \begin{bmatrix} 1 - Y_3 Z_c/2 & -Y_3 Z_c/2 & NZ_c/2 \\ Y_3 Z_c/2 & 1 + Y_3 Z_c/2 & -NZ_c/2 \\ N & N & 0 \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \\ I_3 \end{bmatrix} = [T] \begin{bmatrix} V_{1f} \\ V_{1b} \\ I_3 \end{bmatrix}. \tag{4.22}$$

The matrix representing the propagation of waves on a length $p/2$ of the transmission line is

$$\begin{bmatrix} V_{2f} \\ V_{2b} \end{bmatrix} = \begin{bmatrix} \exp(-j\beta_0 p/2) & 0 \\ 0 & \exp(j\beta_0 p/2) \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \end{bmatrix}. \tag{4.23}$$

Thus the hybrid wave matrix for a single cell of the structure is

$$\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_3 \end{bmatrix} = \begin{bmatrix} e^{-j\beta_0 p/2} & 0 & 0 \\ 0 & e^{j\beta_0 p/2} & 0 \\ 0 & 0 & 1 \end{bmatrix} [T] \begin{bmatrix} e^{-j\beta_0 p/2} & 0 & 0 \\ 0 & e^{j\beta_0 p/2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \\ I_3 \end{bmatrix}. \tag{4.24}$$

This can be evaluated as

$$
\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_3 \end{bmatrix} = \begin{bmatrix} \left(1 - Y_3 Z_c/2\right)e^{-j\beta_0 p} & -Y_3 Z_c/2 & \left(NZ_c/2\right)e^{-j\beta_0 p/2} \\ Y_3 Z_c/2 & \left(1 + Y_3 Z_c/2\right)e^{j\beta_0 p} & -\left(NZ_c/2\right)e^{j\beta_0 p/2} \\ Ne^{-j\beta_0 p/2} & Ne^{j\beta_0 p/2} & 0 \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \\ I_3 \end{bmatrix}. \quad (4.25)
$$

We note that, if port 3 is open circuit $\left(I_3 = 0\right)$ then (4.25) shows that the forward and backward waves are coupled by the shunt admittance Y_3. Furthermore the off-diagonal terms $(1,2)$ and $(2,1)$ have opposite signs so, from the coupled-mode theory reviewed in Section 1.3.8, we recognise this as Case B coupling between the forward and the backward wave. Thus the presence of a periodic discontinuity in the line leads to the presence of a stop band, as is well-known from the theory of filters [10]. In cases where $Y_3 = 0$ there is no stop band but the periodic coupling to the transmission line remains.

By analogy with the surface impedance of a uniform structure we can define the *total impedance* of a periodic structure by

$$
Z_T = \frac{|V_3|^2}{2P}, \quad (4.26)
$$

where P is the power flow on the line. But for a forward wave

$$
P = \frac{|V_f|^2}{2Z_c} \quad (4.27)
$$

and therefore the turns ratio N is found from the element at the bottom of the first column of the matrix in (4.25) to be

$$
N = \sqrt{Z_T/Z_c}. \quad (4.28)
$$

4.1.3   Space Harmonics

Let us assume, for simplicity, that the field in each gap is uniform (which is permissible if the opposing edges are blunt) and varies sinusoidally with time with frequency ω. The phases of the fields in adjacent gaps are not independent but are related by

$$
\phi_{i+1} = \phi_i \exp{-j\left(\beta_0 p + 2m\pi\right)}, \quad (4.29)
$$

where ϕ_i is the phase in the ith gap and $m = 0, \pm 1, \pm 2, \ldots$ The additional term in the brackets in (4.29) arises because the exact phase relationship is indeterminate to a multiple of 2π. We also note that, at a fixed time, the field along the wall is a periodic function in space whose wavelength is

$$
\lambda = 2\pi/\beta_0. \quad (4.30)
$$

Figure 4.5 shows an illustration of the way in which the field at $r = a$ varies with time. It can be seen that the gap fields are samples of a wave which is travelling in the positive z direction.
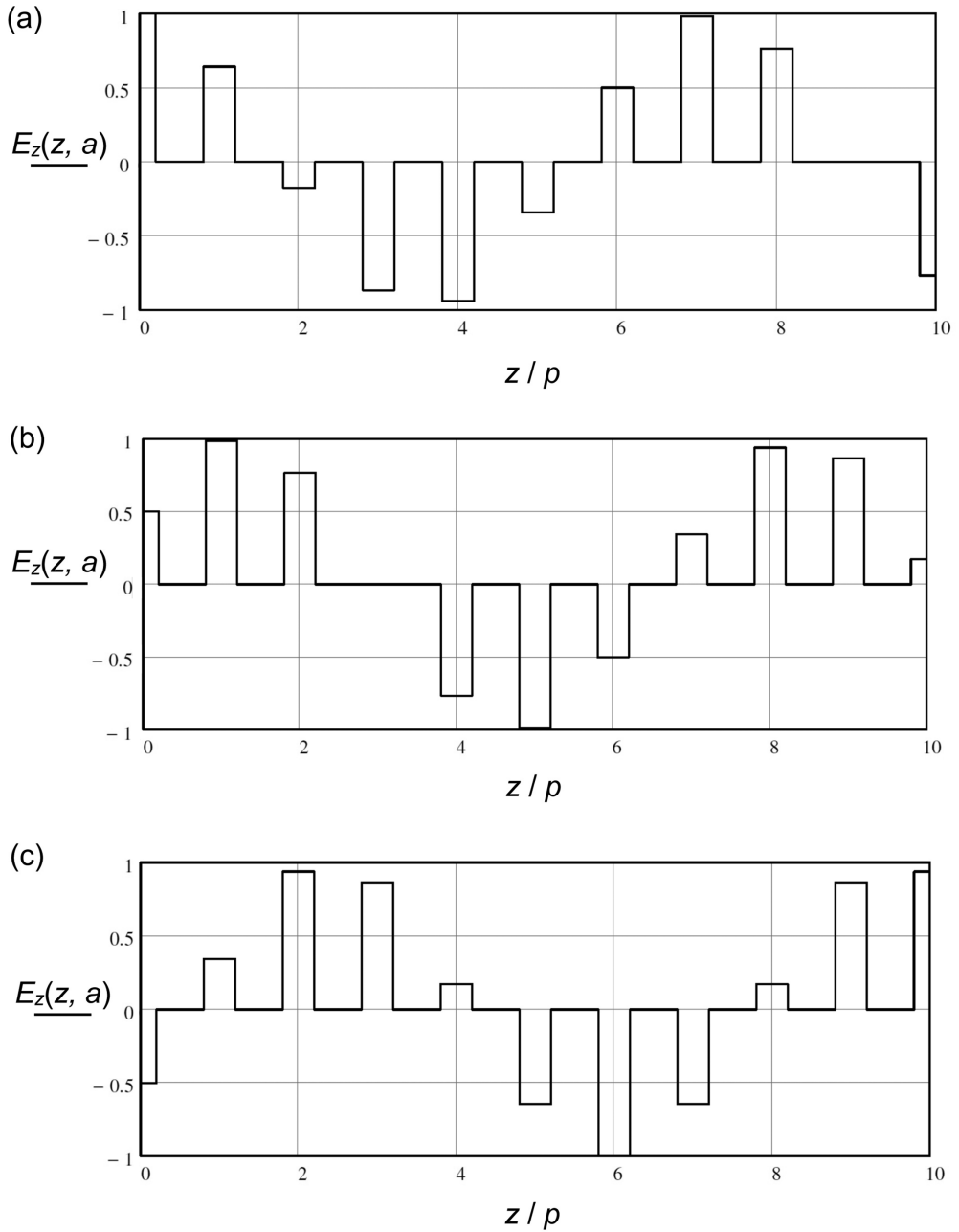
(a)

$E_z(z, a)$

z / p

(b)

$E_z(z, a)$

z / p

(c)

$E_z(z, a)$

z / p

Figure 4.5: Electric field of a periodic slow-wave structure at $r = a$ for $\beta_0 p = 50°$: (a) $\omega t = 0$, (b) $\omega t = 60°$, and (c) $\omega t = 120°$.

If we define propagation constants

$$\beta_m = \pm\left(\beta_0 + \frac{2m\pi}{p}\right)$$ (4.31)

which have periodicity p then, by virtue of Floquet's Theorem [11], the electric field along the wall can be expanded as a Fourier series in space

$$E_z(z,a) = \sum_{m=-\infty}^{\infty} E_m \exp(-j\beta_m z), \tag{4.32}$$

where E_m is the amplitude of the mth space harmonic of the field and sinusoidal time dependency as $\exp(j\omega t)$ is assumed. To find the amplitudes of the individual harmonics we multiply both sides of (4.32) by $\exp(j\beta_n z)$ and integrate from $-p/2$ to $p/2$. The result is

$$\int_{-p/2}^{p/2} E_z(z,a)\exp(j\beta_n z)\,dz = \sum_{m=-\infty}^{\infty} \left\{ \int_{-p/2}^{p/2} E_m \exp(-j(\beta_m-\beta_n)z)\,dz \right\}. \tag{4.33}$$

Now

$$\exp(-j(\beta_m-\beta_n)z) = \exp\left(-j\frac{2\pi}{p}(m-n)\right) \tag{4.34}$$

is zero except when $m=n$. Then, if $E_z(z,a) = E_0$ when $-g/2 \leq z \leq g/2$ and $E_z(z,a) = 0$ elsewhere, (4.33) becomes

$$\int_{-g/2}^{g/2} E_0 \exp(j\beta_m z)\,dz = \int_{-p/2}^{p/2} E_m\,dz, \tag{4.35}$$

which may be integrated to give

$$E_m = \frac{E_0}{j\beta_m p}\left(\exp\left(j\frac{\beta_m g}{2}\right) - \exp\left(-j\frac{\beta_m g}{2}\right)\right) \tag{4.36}$$

and rearranged as

$$E_m = E_0\left(\frac{\sin(\beta_m g/2)}{\beta_m g/2}\right)\cdot\frac{g}{p}. \tag{4.37}$$

If we suppose that the gaps are connected to a dispersionless transmission line, then the dispersion diagram of the wave in the gaps is as shown in Figure 4.6. The notation in this figure is consistent with (4.31) but differs from that used by some other authors. The advantage of the notation used here is that the space harmonics of the forward and backward waves having the same index (m) have the same normalised amplitudes. From Figure 4.6 it is evident that the group velocities $(\partial\omega/\partial\beta)$ of all the space harmonics are the same at a given frequency but that their phase velocities (ω/β) differ. The significance of this will become apparent later on. It is important to recognise that the amplitudes of the space harmonics are linked by (4.37) so that they cannot be excited independently of each other. An alternative way of thinking about space harmonics is to note that an electron passing along
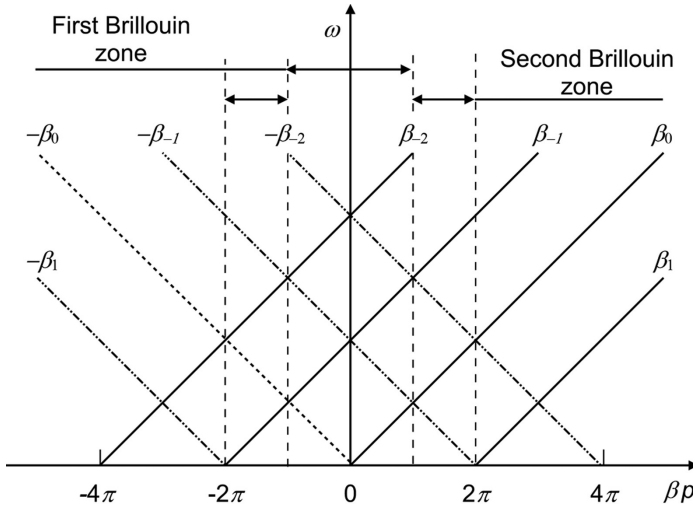
Figure 4.6: Dispersion diagram for space harmonic waves on a dispersionless periodic slow-wave structure.

the wall experiences the instantaneous field in each gap as it passes it. The phase relationship between the gaps is therefore perceived as that of the space harmonic whose phase velocity is equal to the velocity of the electron. This is an example of the aliasing that occurs when waveforms which are periodic in time are sampled periodically [12].

In Figure 4.6 the range $-\pi/p \leq \beta \leq \pi/p$ is known as the *First Brillouin Zone*, while $\pi/p \leq |\beta| \leq 2\pi/p$ is the *Second Brillouin Zone*, and so on. It can be seen that the dispersion diagram in the range $0 \leq \beta \leq \pi/p$ is repeated periodically in both directions. Thus it is often sufficient to plot the dispersion diagram only in this range. A diagram of this kind is called a *reduced zone diagram*. We have seen that if the transmission line has periodic discontinuities with spacing p then the forward and backward waves are coupled together and stop bands occur. It is easy to see that these are caused by the cumulative addition of waves reflected from the discontinuities when the electrical length from one discontinuity to the next, and back, is equal to a whole number of wavelengths. Thus coupling between the modes can occur when $\beta = \pm n\pi/p$ where n is an integer. Examples of this will be shown in later sections.

It is common in the literature to find references to 'forbidden regions' in the dispersion diagrams. These occur when a slow-wave structure is not surrounded by a conducting wall. It then radiates energy into space when the phase velocity of a space harmonic exceeds the velocity of light, as it may at some frequencies. When the structure is contained within a conducting wall, waves that propagate with phase velocities greater than c do not radiate energy and there are no forbidden regions. All the structures to be discussed in this chapter are contained within metal walls.

4.2   Planar Slow-Wave Structures

Planar slow-wave structures find some use in crossed-field amplifiers, and are of interest for their potential use in millimetre wave travelling-wave tubes. We shall see below that the concepts are helpful for the understanding of helix and helix-derived structures. All the planar structures to be discussed are based on a microstrip array, as shown in the cross-section in Figure 4.7. At a height h_2 above the ground plane there may be another earthed conducting plane or a symmetry boundary. The strips have finite thickness, as shown, and the array is assumed to be of infinite extent in the (x, z) plane. If there is a dielectric present it is normally in contact with the strips to provide mechanical support and a path for the conduction of heat. However, the theory to be described can also be applied to cases where the dielectric layer does not make contact with the strips. In many practical structures the dielectric is not continuous but comprises an array of rods at right angles to the strips. It is found that such an array of dielectric rods may be replaced, for modelling purposes, by an equivalent uniform dielectric without serious loss of accuracy.

We wish to investigate the case where the lines propagate a TEM wave in the x direction such that there is a phase difference ϕ between adjacent strips. Strictly speaking, a pure TEM wave cannot exist because of the shear caused by the difference between the phase velocities in the dielectric and in free space. However, in many cases, this effect is small and it is possible to assume that there is a quasi-TEM wave [13]. Figure 4.8 shows the self capacitance of one line, the mutual capacitances coupling it to the adjacent lines, and the voltage on each line. Consider, first, the case where there is no dielectric substrate [14]. The charge per unit length on the line is

$$q = C_0 V + \sum_{n=1}^{N} C_n V \left(2 - \exp(-jn\phi) - \exp(jn\phi) \right) \tag{4.38}$$

so that the effective capacitance per unit length of the line is

$$C(\phi) = C_0 + \sum_{n=1}^{N} 2C_n \left(1 - \cos(n\phi) \right), \tag{4.39}$$
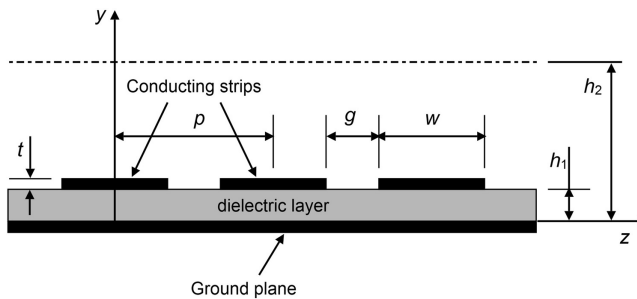
Figure 4.7: Cross-section of a microstrip array.

where the summation is taken to include as many terms as are needed to ensure good accuracy. The coefficients of capacitance can readily be found from finite difference solutions of Laplace's equation (see Worksheet 4.1). The use of quasi-static analysis is justified because the phase velocity of propagation in the z direction is normally much less than the velocity of light. In practice it is found that satisfactory results can normally be obtained with $N = 2$.

The effective inductance per unit length can be found in a similar way. The magnetic flux linked to one line is

$$\Lambda = L_s i + \sum_{n=1}^{N} M_n i \left(\exp\left(jn\phi \right) + \exp\left(-jn\phi \right) \right), \tag{4.40}$$

where i is the magnitude of the current in each line, L_s is the self-inductance per unit length of the line, and M_n are the mutual inductances coupling it to the other lines in the array. Thus the effective inductance per unit length is

$$L\left(\phi \right) = L_s + \sum_{n=1}^{N} 2M_n \cos\left(n\phi \right). \tag{4.41}$$

Now let L_0 be the inductance per unit length when $\phi = 0$. Then (4.41) can be written

$$L\left(\phi \right) = L_0 - \sum_{n=1}^{N} 2M_n \left(1 - \cos\left(n\phi \right) \right). \tag{4.42}$$

It is not necessary to find explicit expressions for the coefficients of inductance because the line supports a *TEM* wave with phase velocity c so that

$$L\left(\phi \right) = \frac{1}{c^2 C\left(\phi \right)}. \tag{4.43}$$

However we shall find (4.42) useful in understanding the difference between the properties of different types of slow-wave structure.

When there is a dielectric substrate present with relative permittivity ε_2 the effective capacitance per unit length is

$$C_K\left(\phi \right) = C_{K0} + \sum_{n=1}^{N} 2C_{Kn} \left(1 - \cos\left(n\phi \right) \right), \tag{4.44}$$

where the coefficients of capacitance can be found using finite difference calculations as before. This method can be used when the strip conductors have finite thickness, whereas analytical methods are restricted to strips having zero thickness [13, 14]. We note that, approximately

$$C_{K0} = \varepsilon_2 C_0 \tag{4.45}$$

and

$$C_{Kn} = \frac{1}{2}\left(1 + \varepsilon_2 \right) C_n. \tag{4.46}$$
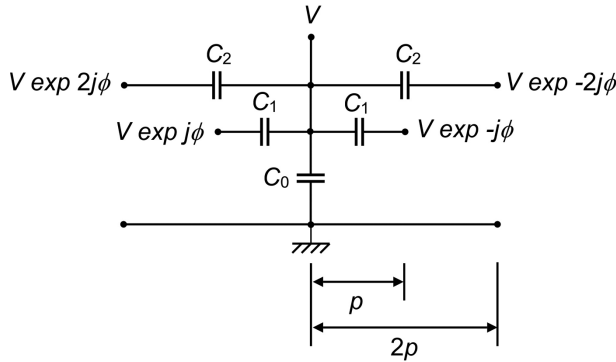
Figure 4.8: Equivalent circuit for the capacitance of one line in the microstrip array.

It is convenient to define the effective relative permittivity by

$$\varepsilon_{eff}(\phi) = \frac{C_K(\phi)}{C(\phi)}. \tag{4.47}$$

The effective permittivity calculated in this way with $N = 2$ agrees closely with that computed by Weiss [15] if it is assumed that there is a symmetry boundary at $y = h_2$ (see Worksheet 4.2).

The inductance per unit length is not changed by the presence of the dielectric so the phase velocity of waves on the line with dielectric loading is

$$v_{pe}(\phi) = \frac{1}{\sqrt{L(\phi)C_K(\phi)}} = \frac{c}{\sqrt{\varepsilon_{eff}(\phi)}}, \tag{4.48}$$

the propagation constant in the x direction is

$$\beta_x(\phi) = \frac{\omega}{v_{pe}(\phi)}, \tag{4.49}$$

and the characteristic impedance is

$$Z_c(\phi) = \sqrt{\frac{L(\phi)}{C_K(\phi)}}. \tag{4.50}$$

The results of this calculation also agree closely with those given by Weiss [15].

This basic formulation can be used to find the properties of several of different types of planar slow-wave structures [14]. The procedure is to consider a length l of the lines in the x direction and to apply boundary conditions as illustrated in Figure 4.9. In all practical cases the lines must be enclosed by a metal boundary. The region above the line may be terminated by a conducting plane, or two identical lines may be placed opposite one another. The electrons which interact
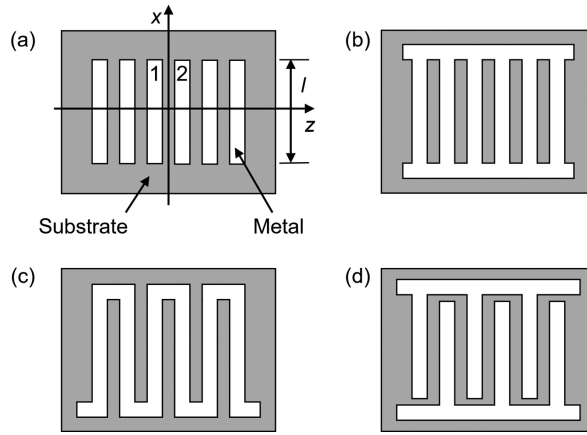
Figure 4.9: Planar slow-wave structures: (a) Finite array of parallel lines, (b) Ladder line, (c) Meander line, and (d) Inter-digital line.

with the fields of the line move in the z direction. They may either pass close to the surface of the line or, if the strips are sufficiently thick, through holes bored through them.

4.2.1   Ladder Line

A ladder line is created by connecting all the strips together at each end as shown in Figure 4.9(b). The boundary conditions are

$$\begin{cases} V_1(l/2) = V_2(l/2) = 0 \\ V_1(-l/2) = V_2(-l/2) = 0 \end{cases},$$

(4.51)

where V_1 and V_2 are the voltages on two adjacent lines. These conditions require that there is a standing wave on each line such that

$$\beta_x(\phi)l = \pi$$

(4.52)

for all ϕ. Then the dispersion equation is

$$\omega = \frac{\pi}{l}v_p(\phi) = \frac{\pi c}{l} \cdot \frac{1}{\sqrt{\varepsilon_{eff}(\phi)}}.$$

(4.53)

The short circuits at the ends of the lines can be provided by the side walls of the enclosing waveguide so no dielectric is necessary to support them. However, the circuit is resonant at $\omega = \pi c/l$ when there is no dielectric loading, so it is then of little use as a slow-wave structure. The structure becomes useful if the ladder line is enclosed by a ridged waveguide (the Karp structure) or by a cruciform waveguide (the Anti-Karp structure) [16, 17]. Figure 4.10 shows examples of ladder line slow-wave structures and their dispersion curves [1].
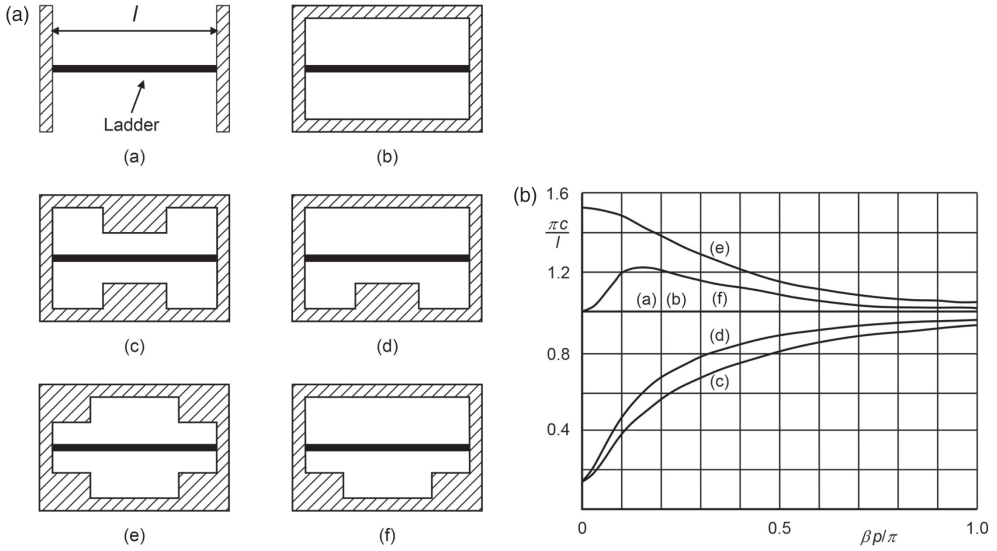
Figure 4.10: Propagation along ladder lines, (a) Cross sections of lines, (b) dispersion curves with ladder detail
(copyright IEEE 1960, reproduced, with permission, from [1]).

4.2.2   Meander Line

The boundary conditions for the meander line structure shown in Figure 4.9(c), are [15]

$$
\begin{cases}
V_1(l/2) = V_2(l) \\
I_1(l/2) = -I_2(l/2) \\
V_1(-l/2) = V_2(-l/2)\exp(j2\phi) \\
I_1(-l/2) = -I_2(-l/2)\exp(j2\phi)
\end{cases}.
\tag{4.54}
$$

In order to satisfy these conditions four propagating modes are required. Two are provided by the even mode propagating in the $\pm x$ directions with a propagation constant given by (4.49). The other two waves are those of the odd mode for which the phase is $(\pi + \phi)$ [14, 18]. The voltages on two adjacent lines are given by

$$
V_{1,2} = \left[2A\cos(\beta_e x) \mp 2jB\sin(\beta_o x) \right] \exp(\pm j\,\phi/2),
\tag{4.55}
$$

where $\beta_e = \beta_x(\phi)$ and $\beta_o = \beta_x(\phi + \pi)$ are the propagation constants in the x direction of the even and odd modes, and A and B are the amplitudes of the even and odd modes whose ratio is

$$
\frac{B}{A} = \left\{ \frac{Z_o}{Z_e} \cdot \frac{\sin(\beta_e l)}{\sin(\beta_o l)} \right\}^{\frac{1}{2}},
\tag{4.56}
$$

where $Z_e = Z_c(\phi)$ and $Z_o = Z_c(\phi + \pi)$. The dispersion equation obtained is

$$\tan^2(\phi/2) = \frac{Z_o}{Z_e} \begin{Bmatrix} \tan(\beta_e l/2)\tan(\beta_o l/2) \\ \cot(\beta_e l/2)\cot(\beta_o l/2) \end{Bmatrix}. \tag{4.57}$$

The upper line in (4.57) applies to the first pass-band of the structure and the lower line to the second pass-band. The dispersion diagram for the case presented in [15, 19] is shown in Figure 4.11. In this figure the lines are calculated from (4.57) assuming that l is the full length of the line in the x direction. The dashed lines assume TEM propagation along a folded line whose length is l with phase velocity determined by the approximate effective relative permittivity $(1 + \varepsilon_2)/2$ [13]. The experimental points are taken from [19] (see Worksheet 4.2). It can be seen that the periodic discontinuities in the line produce a stop band when $\phi = \beta_0 p = \pi$. Note that in [15] the phase φ is that for a complete period $(2p)$ of the structure so that $\varphi = 2\phi$.

To find the coupling impedance of the structure we note that the energy stored per line in the even modes is

$$W_e(\phi) = \frac{1}{2}C_K(\phi) \int_{-l/2}^{l/2} 4A^2 \cos^2(\beta_e x)\, dx \tag{4.58}$$

and in the odd modes is

$$W_o(\phi) = \frac{1}{2}C_K(\phi + \pi) \int_{-l/2}^{l/2} 4B^2 \sin^2(\beta_o x)\, dx. \tag{4.59}$$

Hence the power flow in the z direction is

$$P_z(\phi) = \frac{2l|A|^2}{p} v_g \left[\frac{1}{v_e Z_e}\left(1 + \frac{\sin\beta_e l}{\beta_e l}\right) + \frac{1}{v_o Z_o}\left(\frac{B}{A}\right)^2\left(1 - \frac{\sin\beta_o l}{\beta_o l}\right) \right] \tag{4.60}$$

where the group velocity $v_g = d\omega/d\beta_0$ is obtained from the dispersion curve. The magnitude of the electric field in the gaps can be shown to be

$$E_g(x,\phi) = \frac{4A}{g}\left\{ \cos(\beta_e x)\sin(\phi/2) + \frac{B}{A}\sin(\beta_o x)\cos(\phi/2) \right\}. \tag{4.61}$$

Figure 4.12 shows how the magnitude of the field on the meander line, whose dimensions are given in [15], varies with x for various values of ϕ. The phase of the field is constant in each gap and the phase difference between the gaps is ϕ. The folding of the line means that the curves in Figure 4.12 are reflected about the plane $x = 0$ in alternate gaps. This property means that the calculation of the space harmonics and of the coupling impedance is complicated. For the special case where $x = 0$ the magnitude of the fundamental space harmonic is obtained from (4.37) as

$$E_0(0,\phi) = E_g(0,\phi)\left(\frac{\sin(\beta_0 g/2)}{\beta_0 g/2} \right) \cdot \frac{g}{p}. \tag{4.62}$$
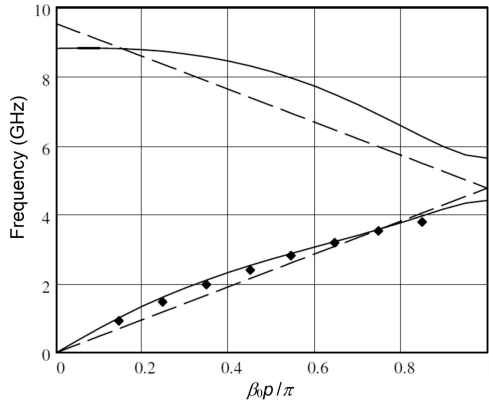
Figure 4.11: Dispersion diagram of a meander line slow-wave structure showing propagation along uncoupled lines (dashed) and experimental points for comparison (copyright 1974 IEEE, reproduced, with permission, from [15]).
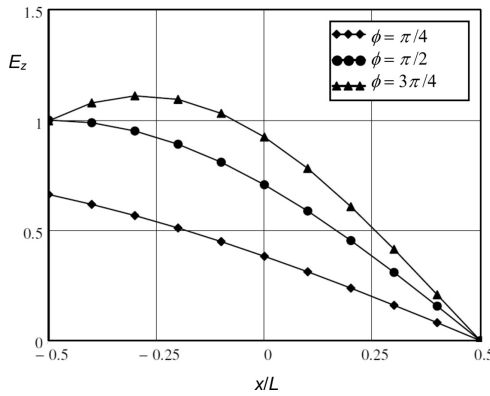

Figure 4.12: Distribution of the z component of the electric field in a meander line slow-wave structure.

where $\beta_0 p = \phi$. The surface impedance of the line at $x = 0$ is given by

$$Z_S = \frac{E_0\left(0, \phi\right)^2}{2\beta_0^2 P_z}. \tag{4.63}$$

The results obtained from this equation are of the same magnitude as those in [15] but the curve given there is probably wrong. The results of experimental measurements made on the same structure by covering it with a thin dielectric sheet are rather different but difficult to reproduce theoretically (see Worksheet 4.2) [19]. Practical examples of meander line slow-wave structures are given in [20–23].

4.2.3   Inter-digital Line

The boundary conditions for the inter-digital line shown in Figure 4.9(d) are [24]

$$\begin{cases} V_1\left(l/2\right) = V_2\left(-l/2\right) = 0 \\ I_1\left(-l/2\right) = I_2\left(l/2\right) = 0 \end{cases}.$$ (4.64)

The period of the inter-digital line is $2p$ like that of the meander line. The dispersion equation is the same as that for a meander line if the widths of the strips and the gaps are interchanged. This is because the two structures are duals of one another so that the electric field lines in one case correspond to the magnetic field lines in the other and vice-versa. This is an application of Babinet's principle [6]. The interdigital line does not require dielectric support [25, 26].

4.3   Helix Slow-Wave Structures

The planar structures discussed in the previous section are useful for interaction with sheet beams, and in crossed-field devices, but their arrangement is not ideal for interaction with the cylindrical beams used in most linear-beam tubes. Also, because they are periodic structures, they are dispersive and have limited bandwidth. These problems are avoided in the helix slow-wave structure shown in Figure 4.13. In this structure a helical conducting tape is held within a metal shield by three dielectric support rods. The support rods also provide a path for heat to flow from the helix to the shield. This structure can be regarded as a coaxial line with a helical inner conductor. Thus the series inductance is increased and the phase velocity decreased as discussed in Section 4.1.1. The dielectric loading is kept as small as possible, consistent with adequate thermal conduction, to avoid increasing the shunt capacitance and thereby reducing the characteristic impedance. The support rods are usually made of materials such as anisotropic boron nitride and CVD diamond [27]. Beryllia, which has high thermal conductivity, is toxic and its use is now generally avoided. The helix is made of tungsten or molybdenum to ensure dimensional
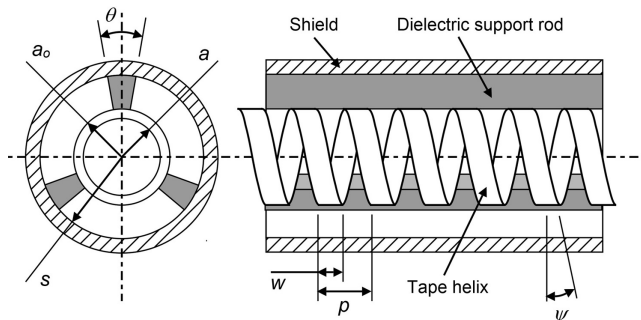
Figure 4.13: Arrangement of a helix slow-wave structure.

stability at high temperatures. It is important to maintain good thermal contact between the support rods and the helix. This can be achieved by using assembly methods which ensure that the rods exert a compressive force on the helix, or by brazing the helix to the rods [4, 28]. The thermal analysis of rod supported helix slow-wave structures is discussed in [29, 30].

The basic properties of the helix slow-wave structure can be understood using the methods of the previous section. Figure 4.14 shows two adjacent strips from a developed (opened out) helix. From the geometry of the structure the pitch angle ψ is given by

$$\tan \psi = \frac{p}{2\pi a}. \tag{4.65}$$

The boundary conditions are now just

$$\begin{cases} V_1\left(l/2\right) = V_2\left(-l/2\right) \\ I_1\left(l/2\right) = I_2\left(-l/2\right) \end{cases} \tag{4.66}$$

and the structure supports an azimuthal *TEM* wave whose phase velocity is determined by the effective permittivity. If the helix is supported by a material whose relative permittivity is unity then the axial phase velocity is expected to be

$$v_p = c\sin\psi. \tag{4.67}$$

It is found that this is not quite accurate for reasons which are explored below.
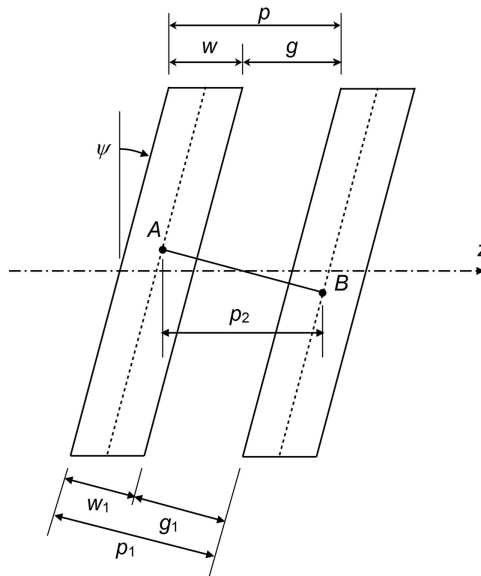
Figure 4.14: The geometry of the helix slow-wave structure.

4.3.1   The Sheath Helix Model

Although the helix is geometrically simple it is very difficult to analyse mathematically. The simplest analytical model treats the helix as a thin cylinder in which conduction is only possible at an angle ψ to transverse planes, where ψ is the pitch angle of the helix shown, in Figure 4.13 [10]. The region between the cylinder and the shield is assumed to be filled with uniform dielectric material having relative permittivity ε_2. The component of the electric field in the direction in which the cylinder conducts must be zero and therefore, in region 1 $(r \le a)$,

$$E_{\theta 1} \cos \psi + E_{z1} \sin \psi = 0 \tag{4.68}$$

and, in region 2 $(a \le r \le s)$,

$$E_{\theta 2} \cos \psi + E_{z2} \sin \psi = 0. \tag{4.69}$$

The tangential electric field normal to the direction of conduction must be continuous so that

$$E_{z1} \cos \psi - E_{\theta 1} \sin \psi = E_{z2} \cos \psi - E_{\theta 2} \sin \psi. \tag{4.70}$$

The current flow in the cylinder is in the direction of conduction and therefore the tangential magnetic field must be continuous in this direction

$$H_{\theta 1} \cos \psi + H_{z1} \sin \psi = H_{\theta 2} \cos \psi + H_{z2} \sin \psi. \tag{4.71}$$

These four equations define the boundary conditions at $r = a$. The boundary condition at $r = s$ is $E_\theta = E_z = 0$. The axial components of both the electric and magnetic field are non-zero. Therefore the wave propagating on the structure cannot be a pure TM or TE mode but is a hybrid of the two linked by the conduction angle of the cylinder. A cylindrically-symmetrical solution to this problem can be found by assuming that all quantities vary in the z direction as $\exp j\left(\omega t - \beta_0 z\right)$. The radial propagation constants are then

$$\begin{aligned} \gamma_0 &= \sqrt{\beta_0^2 - \left(\omega/c\right)^2} \quad (0 \le r \le a) \\ \gamma_2 &= \sqrt{\beta_0^2 - \varepsilon_2 \left(\omega/c\right)^2} \quad (a \le r \le s). \end{aligned} \tag{4.72}$$

In practice β_0^2 is normally much greater than $\left(\omega/c\right)^2$ and it is possible to use the approximations $\gamma_0 = \gamma_2 = \beta_0$ without serious error. For the TM fields

$$\begin{aligned} E_{z1} &= A_1 I_0\left(\gamma_0 r\right) \\ E_{z2} &= \frac{A_2}{K_0\left(\gamma_2 s\right)}\left[I_0\left(\gamma_2 r\right) K_0\left(\gamma_2 s\right) - I_0\left(\gamma_2 s\right) K_0\left(\gamma_2 r\right)\right], \end{aligned} \tag{4.73}$$

where A_1 and A_2 are constants, since the z component of tangential electric field must be zero when $r = s$. Similarly for the TE fields

$$\begin{aligned} H_{z1} &= B_1 I_1\left(\gamma_0 r\right) \\ H_{z2} &= B_2 I_1\left(\gamma_2 r\right) + C_2 K_1\left(\gamma_2 r\right), \end{aligned} \tag{4.74}$$

where B_1, B_2 and C_2 are constants. The transverse field components H_θ and E_θ can be found from (2.19) and (2.20) expressed in cylindrical polar co-ordinates. Applying the boundary condition on E_θ at $r = s$ gives C_2 in terms of B_2 so that all the field components can be expressed as functions of r and the four constants A_1, A_2, B_1 and B_2. Substituting these expressions into the equations for the four boundary conditions at $r = a$((4.68) to (4.71)) leads to the dispersion equation for the structure. The equations obtained in this way are rather complicated so we will simply summarise the results.

If the sheath helix is in free space and $s \to \infty$ the dispersion equation is [10]

$$\frac{(\gamma_0 a)^2}{(ka)^2}\tan^2 \psi = \frac{I_1(\gamma_0 a)K_1(\gamma_0 a)}{I_0(\gamma_0 a)K_0(\gamma_0 a)}. \tag{4.75}$$

The effects of the finite shield radius and the supporting dielectric material can conveniently be expressed using the methods described in [31]. Equation (4.75) is first expressed in terms of equivalent circuit parameters by

$$L_0 = \left(\frac{\mu_0}{2\pi}\right)\left(I_1(\gamma_0 a)K_1(\gamma_0 a)\right)\left(\frac{\beta_0}{\gamma_0}\right)^2 \cot^2 \psi \tag{4.76}$$

and

$$C_0 = \frac{2\pi\varepsilon_0}{I_0(\gamma_0 a)K_0(\gamma_0 a)}. \tag{4.77}$$

Substitution of L_0 and C_0 into the equation for the propagation constant of a transmission line (2.21) gives the dispersion equation (4.75). It should be noted that the inductance and capacitance are functions of frequency.

If the sheath helix is surrounded by a shield then the inductance and capacitance become [31]

$$L = L_0\left[1 - \frac{I_1(\gamma_0 a)K_1(\gamma_0 s)}{I_1(\gamma_0 s)K_1(\gamma_0 a)}\right] \tag{4.78}$$

and

$$C = C_0\left[1 - \frac{I_0(\gamma_0 a)K_0(\gamma_0 s)}{I_0(\gamma_0 s)K_0(\gamma_0 a)}\right]^{-1} \tag{4.79}$$

The diameter of the shield is normally chosen so that higher-order modes, especially the TE_{11} coaxial waveguide mode, are cut off. If this is not so then the higher-order modes can provide an unwanted feedback path when the structure is used in a travelling-wave tube [32].

The addition of a uniform dielectric supporting medium between the helix and the shield does not affect the inductance, but it multiplies the capacitance by the effective relative permittivity [33]

$$\varepsilon_{eff}\left(\gamma_0 a, \gamma_0 s\right) = 1 + \left(\varepsilon_2 - 1\right)\left(\gamma_0 a\right) I_0\left(\gamma_0 a\right) K_1\left(\gamma_0 a\right)\left\{1 + \frac{I_1\left(\gamma_0 a\right) K_0\left(\gamma_0 s\right)}{I_0\left(\gamma_0 b\right) K_1\left(\gamma_0 a\right)}\right\}. \quad (4.80)$$

where it has been assumed that $\gamma_2 = \gamma_0$. Note that this expression differs from that given in [31] which does not take account of the radial boundary conditions correctly. The effect of the additional term (in braces) is most important if s/a is close to unity. When the dielectric loading is included (4.79) becomes

$$C = C_0 \varepsilon_{eff}\left(\gamma_0 a, \gamma_0 s\right)\left[1 - \frac{I_0\left(\gamma_0 a\right) K_0\left(\gamma_0 s\right)}{I_0\left(\gamma_0 s\right) K_0\left(\gamma_0 a\right)}\right]^{-1} \quad (4.81)$$

In the derivation of these equations it was assumed that the phase velocity was small enough for the quasi-static approximation to be valid so that $\gamma_0 \simeq \beta_0$ everywhere.

The characteristic impedance of the slow-wave structure is found by substituting the values of L and C into (4.9). An alternative, and much more complicated, expression for the characteristic impedance of a helix in free space can be derived by integrating the Poynting vector to obtain the power flow [34]. The two methods give identical numerical results.

In practical slow-wave structures the helix is supported within the shield by dielectric support rods, as shown in Figure 4.13. If these are wedge-shaped then relative permittivity of the equivalent uniform dielectric is given by [31, 35]

$$\varepsilon_2 = 1 + \frac{N\theta}{2\pi}\left(\varepsilon_r - 1\right), \quad (4.82)$$

where there are N rods having relative permittivity ε_r each subtending an angle θ on the axis. This equation can be derived by considering a coaxial line in which the central conductor is supported by wedge-shaped rods. The problem is a quasi-static one in which the equipotentials are concentric circles. The capacitance between the two conductors is therefore the parallel combination of the capacitances of the vacuum and the dielectric regions. That leads directly to (4.82). When the rods have some other shape, such as the 'T' shaped rods shown in Figure 4.15(a), the effective relative permittivity can still be determined from the capacitance. It is commonly assumed that this can be done by separating the region between the helix and the shield into two or more layers, and then finding the effective relative permittivity for each, using (4.82) [36]. This method, which assumes that the equipotentials remain concentric circles, actually leads to an upper bound of the capacitance and, therefore, of the effective relative permittivity. It can be shown that a lower bound can be found by assuming that the flux lines remain radial. This is an example of the method of Tubes and Slices in which upper and lower bounds to the capacitance are calculated from the stored energy using approximations to the field [37]. An alternative method for calculating the effective relative permittivity from the stored energy is described in [38].
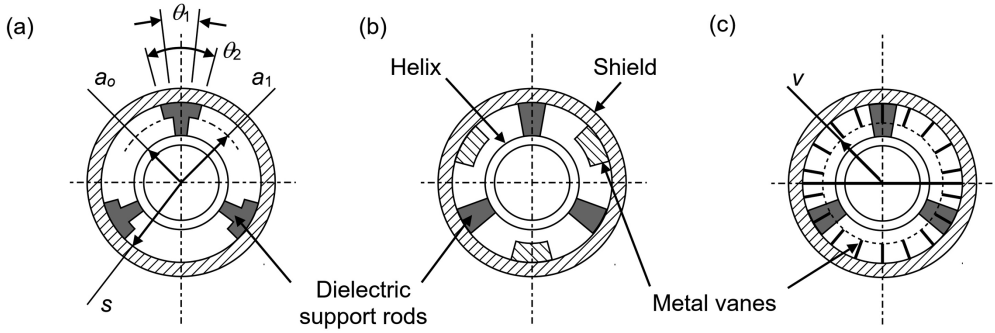
Figure 4.15: Different arrangements for helix slow-wave structures: (a) T-shaped support rods, (b) vane loading for dispersion shaping, and (c) theoretical model of vane loading.

Figure 4.16(a) shows typical dispersion curves obtained from the sheath helix model. It is useful to consider the asymptotic behaviour of these in the limits $\omega \to 0$ and $\omega \to \infty$. In the low frequency limit it can be shown that the inductance and capacitance per unit length tend to

$$L_{00} = \frac{\mu_0}{4\pi}\left(1 - \frac{a^2}{s^2}\right)\cot^2 \psi \tag{4.83}$$

and

$$C_{00}(\varepsilon_2) = \frac{2\pi\varepsilon_0\varepsilon_2}{\ln(s/a)}, \tag{4.84}$$

which is the capacitance per unit length of a coaxial line (see (2.47)). Thus the low-frequency limit of the phase velocity is

$$v_{p0} = \frac{1}{\sqrt{L_{00}C_{00}}} = \frac{c\tan\psi}{\sqrt{\varepsilon_2}}\cdot\sqrt{\frac{2\ln(s/a)}{1 - a^2/s^2}}. \tag{4.85}$$

In the high frequency limit we note that the modified Bessel functions are given approximately by [3]

$$I_n(z) = \frac{e^z}{\sqrt{2\pi z}} \tag{4.86}$$

and

$$K_n(z) = \sqrt{\frac{\pi}{2z}}\,e^{-z}. \tag{4.87}$$

Then from (4.78) and (4.79) we find that $L \to L_0$ and $C \to C_0$ because the fields are confined to the region close to the helix and are not dependent on the dimensions of the shield. Hence the high frequency limits are

$$L_0 \to \left(\frac{\mu_0}{4\pi}\right)\frac{1}{\gamma_0 a}\cot^2 \psi \tag{4.88}$$

since $\beta_0 \simeq \gamma_0$ for a slow wave, and

$$C_0 \to 4\pi\varepsilon_0\left(\gamma_0 a\right). \tag{4.89}$$

From (4.80) we find that the effective relative permittivity tends to

$$\varepsilon_{eff}\left(\gamma_0 a, \gamma_0 s\right) \to \frac{1}{2}\left(1+\varepsilon_2\right), \tag{4.90}$$

which is the average of the values of relative permittivity inside and outside the helix. Thus the asymptotic value of the phase velocity is

$$v_p \to \frac{c\tan\psi}{\sqrt{\left(1+\varepsilon_2\right)/2}}. \tag{4.91}$$

We note that this differs from the result previously obtained in (4.67). The asymptote of the phase velocity is shown in Figure 4.16(a) The curves of phase velocity do indeed tend to the asymptote if the plot is extended to higher values of $\beta_0 p$. Figure 4.16(b) shows the Pierce impedance, calculated using (4.17), as a function of the phase shift per turn.

4.3.2   Dispersion Shaping

The dispersion characteristics of a helix can be modified by placing longitudinal vanes between the support rods, as shown in Figure 4.15(b). The effect of the vanes is represented by the model shown in Figure 4.15(c). The tips of the vanes define the radial boundary as far as the electric field is concerned, whilst the magnetic fields are bounded by the outer shield. Thus s is replaced by v in (4.81). Figure 4.17 shows the effect of the vane loading on the dispersion and the Pierce impedance of a typical sheath helix slow-wave structure. The addition of the vanes changes the shape
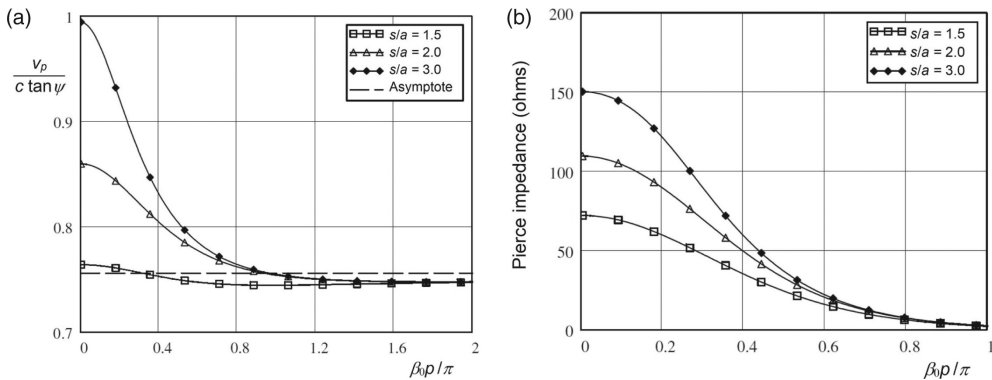
Figure 4.16: Typical properties of a sheath helix slow-wave structure: (a) normalised phase velocity, and (b) Pierce impedance, as functions of the phase shift per turn $\left(\psi = 10°;\, \varepsilon_2 = 2.5\right)$.
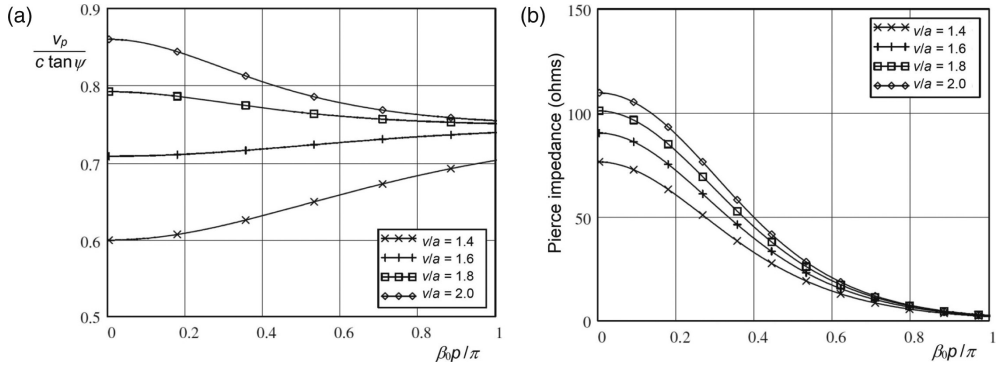
Figure 4.17: The effect of vane loading on a typical sheath helix slow-wave structure: (a) normalised phase velocity, and (b) Pierce impedance $(s/a = 2; \; \psi = 10°; \; \varepsilon_2 = 2.5)$.

of the dispersion curve so that the phase velocity is constant over a wider range of phase shift per turn. However, the Pierce impedance falls as the useful bandwidth is increased. The properties of the sheath helix model can be explored further using Worksheet 4.3.

4.3.3   Tape Helix Slow-Wave Structures

When the sheath helix model is used to calculate the properties of a real tape helix slow-wave structure it is found that there are appreciable errors. In principle a tape helix slow-wave structure could be modelled using the method of moments in a manner similar to that described in Section 3.5.1. However the tape helix problem is much more difficult than the re-entrant cavity, and the efforts of many people over a period of more than sixty years have not yet produced good agreement with experimental results. A great many papers have been published on this subject, and only the more significant ones are reviewed here.

In addition to the fundamental mode described above the sheath helix supports higher-order modes which satisfy the free space dispersion relation [39]

$$\frac{I_n'\left(\gamma a\right) K_n'\left(\gamma a\right)}{I_n\left(\gamma a\right) K_n\left(\gamma a\right)} = -\frac{\left(\gamma^2 a^2 + n\beta a \cot \psi\right)^2}{\left(ka\right)^2 \left(\gamma a\right)^2 \cot^2 \psi}. \tag{4.92}$$

It is therefore possible to model the fields of a tape helix by superimposing space-harmonic waves which satisfy (4.92) and the boundary conditions. It is important to understand that, unlike the space-harmonics discussed earlier, these harmonics have azimuthal variation. In order to make the problem tractable one or more of the following assumptions are made:

• The tape has negligible thickness, so that it is only necessary to consider the fields in the inner and outer regions.
• The region between the helix and the shield is filled with a uniform dielectric.

- The current in the helix flows in the direction of the tape and is distributed in a prescribed fashion.
- The differences between β_0 and γ_0, and between the values of γ in dielectric and in vacuum, are neglected (this is the quasi-static approximation).

The earliest work on this subject was that of Sensiper [39, 40] who studied the properties of a thin helix in free space. Two approximations were considered: narrow tapes on which the current density was taken to be constant; and narrow gaps in which the electric field was taken to be uniform. The dispersion equation for the case of a narrow tape was

$$\sum_{m=-\infty}^{\infty} \left(M_m + N_m \right) \left(\frac{\sin\left(\beta_m w/2\right)}{\beta_m w/2} \right) = 0, \tag{4.93}$$

where w is the width of the tape,

$$M_m = \left(\gamma_m a - \left(m\frac{\beta_m}{\gamma_m} \cot \psi \right) \right)^2 I_m\left(\gamma_m a \right) K_m\left(\gamma_m a \right), \tag{4.94}$$

where $\gamma_m^2 = \beta_m^2 - k^2$, and

$$N_m = \left(k^2 a^2 \cot^2 \psi \right) I_m{}'\left(\gamma_m a \right) K_m{}'\left(\gamma_m a \right). \tag{4.95}$$

The dispersion equation for the sheath helix is obtained if only the terms in $m = 0$ are considered. The same equations were obtained for the case of a wide helix with narrow gaps except that the last term in (4.93) was replaced by $J_0\left(\beta_m w/2 \right)$.

The propagation of waves on a narrow tape helix supported within a metal cylinder by a uniform dielectric was considered by Stark [41]. A dispersion relation similar to that for the sheath helix (4.75) was obtained with the numerator and denominator on the right-hand side multiplied, respectively, by the term in square brackets in (4.78) and the reciprocal of the term in square brackets in (4.79). The effects of the tape width were represented by the addition of the term $(\sin \psi)\ln(2p/\pi w)$ to both the numerator and the denominator. This approach does not seem to have been taken up by subsequent authors. The same problem was considered by Tsutaki [32] who assumed that the current density distribution on the tape could be described by a hyperbolic cosine function. It was found that when $w/p > 0.4$ the results converged if only the $n = 0, \pm 1$ space harmonics were considered. For tapes with $w/p \sim 0.1$ it was necessary to include harmonics up to $n = \pm 3$. It was concluded that in the range $0.4 < w/p < 0.8$ the solutions were not sensitive to the current distribution assumed. The phase velocities calculated were said to be accurate to with $\pm 2\,\%$ but there was no detailed comparison with experiment. The use of a helix as a backward-wave structure was discussed in [42]. It should be noted that as a consequence of (4.92) the axial component of the electric field of the higher space-harmonics, which have azimuthal variation, is zero on the axis.

The effect of the finite thickness of the tape was modelled in [43] using a thin tape at the mean radius of the helix, separated from the dielectric by an air gap equal to half the thickness of the tape. This followed the work of Swift-Hook [44] who carried out a careful theoretical and experimental study of round wire helices inside glass tubes having different wall thicknesses, but no outer metal boundary. It was found that the use of an air gap having a thickness equal to one tenth of the helix radius gave the best results in this case. The justification for the use of the same technique with tape helices, and with a gap equal to half the tape thickness is not clear. Nevertheless this method of representing the effects of tape thickness has been widely used by subsequent authors.

A simplified approach to the modelling of tape helix slow-wave structures was described in [45]. This was based upon Sensiper's dispersion equation for a helix in free space (4.93). Following the method in [31] the effects of the radial boundary and dielectric loading were introduced by multiplying the individual terms of the summation by correction factors similar to (4.78) and (4.79). The effects of dielectric loading were included by assuming that the helix was surrounded by a number of concentric dielectric layers. The effective permittivity of each layer was determined from the relative permittivity of the support rods and the fraction of each layer which was occupied by the rods. This approach was first proposed in [33, 43]. Although there are doubts about the validity of this method, since the boundary conditions are not correctly maintained, it has been widely used by other authors. The results of the simplified model of the tape helix were compared with a small number of experimental results. The agreement was marginally better than that given by the sheath helix model.

Further developments in the modelling of tape helix slow-wave structures included: removal of the assumption that the radial and axial propagation constants were equal [46]; calculation of the constants in the field expansions by successive approximations [47]; removal of the need to assume the current distribution on the tape [48–51]; use of a perturbation technique to evaluate the effects of the azimuthal distribution of the dielectric loading [52]; and an exact theory which included the finite thickness of the helix and the shapes of support rods and vanes by matching radial admittances (but which was only implemented in a simplified form) [53]. All these methods involve very complicated equations and do not add much to understanding of these structures. Where comparisons have been made with experimental results the agreement is generally of the order of 1% for the phase velocity and of the order of 10% or more for the interaction impedance. On the whole the accuracy obtained does not justify the complexity of the analysis involved. Unfortunately only a limited range of experimental data is available in the literature and some of the measurements are old and of doubtful accuracy. The problems of making accurate experimental measurements on helix slow-wave structures are discussed in Section 4.7.

The alternative to the complicated mathematical analysis discussed above is to compute the properties of structures using computational electromagnetics [54–58]. These methods also present some difficulties because the structure does not

normally conform to the boundaries of a regular mesh, and because a very large number of mesh elements is required to ensure good accuracy. The accuracies reported are comparable with those of the best analytical methods.

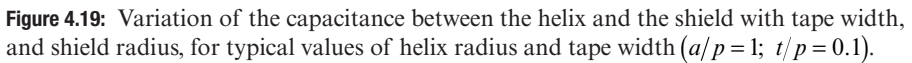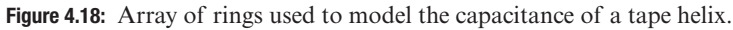### 4.3.4   Equivalent Circuit of Helix Slow-Wave Structures

In Section 4.3 we briefly considered a model comprising parallel strips. In its simple form this model does not predict the dispersion correctly [14]. The reason for this is that in the limit of low frequencies the phase velocity does not correspond to an azimuthal *TEM* wave between the helix and the shield. The presence of an axial magnetic flux inside the helix causes the inductance per turn to be less than it would be for a *TEM* wave and the phase velocity is therefore increased. However, a variation of this approach does give insight, and useful accuracy, as we shall see. A developed model of the helix as a set of parallel metal strips supported over a ground plane by dielectric strips has been used to study the effects of the loading of the structure by vanes similar to those illustrated in Figure 4.15(b) [59]. Empirical approximations to the capacitances were used. Comparisons with four sets of experimental data showed agreement to ±1 % for the phase velocity. The use of empirical approximations can be avoided by using a method similar to that described for planar structures in Section 4.2.

The electrical properties of helix slow-wave structures can be modelled using the same equivalent circuit as that for planar structures [60, 61]. The effective capacitance per unit axial distance is given by (4.39) without dielectric loading and (4.44) with dielectric loading, The direct computation of the capacitances and inductances from the dimensions of the circuit is difficult but this problem can be avoided by making use of the properties of the sheath helix model. Comparison with the results of experimental measurements shows that the properties of a tape helix structure can be calculated to good accuracy if the capacitance and inductance per unit length for a sheath helix are multiplied, respectively, by constants α_C and α_L whose values depend only upon the dimensions, and dielectric loading, of the structure. The best results are obtained when the radius of the sheath helix is taken to be the mean radius of the tape helix.

We consider, first, a structure without dielectric loading. The capacitance per unit length given by (4.39) can be calculated from a two-dimensional finite difference solution for the field between an array of conducting rings and a concentric conducting cylinder as shown in Figure 4.18. The solution is computed between the planes A–A and B–B (see Worksheet 4.4) for $\beta_0 p = 0$, $\pi/2$ and π to give the first three terms of the series. The ratio of the capacitance per unit length of the array of rings, when the phase difference between them is zero, to that of an air-spaced coaxial line having the same dimensions is

$$\alpha_{C0} = \frac{C(0)}{C_{00}(1)},\tag{4.96}$$

where C_{00} is given by (4.84). Figure 4.19 shows how α_{C0} varies with w/p for three values of s/p and typical values of a/p and t/p when C_{00} is calculated using the

Figure 4.18: Array of rings used to model the capacitance of a tape helix.

Figure 4.19: Variation of the capacitance between the helix and the shield with tape width, and shield radius, for typical values of helix radius and tape width $(a/p = 1; \ t/p = 0.1)$.

outer radius of the tape helix. As the tape width is increased the capacitance tends to that of the coaxial line as would be expected.

The value of α_L could be calculated from a numerical solution for the low-frequency inductance per unit length of the tape helix surrounded by a shield. However, this can be avoided by noting that when $\beta_0 p = \pi$ a helical TEM wave propagates with phase velocity c. The value of α_L can then be calculated directly since the phase velocity is given by

$$v_p = \frac{1}{\sqrt{\alpha_L \alpha_{C0} LC}} = c \sin \psi, \tag{4.97}$$

where L and C are given by (4.78) and (4.79) using the mean radius of the helix. It can be shown that the values of C are close to those computed from (4.39) except close to $\beta_0 p = \pi$.

The Pierce impedance is calculated by assuming that the potential difference between adjacent turns is the potential difference between the points A and B in Figure 4.14. Thus

$$
\begin{aligned}
\left|V_{AB}\right| &= V_g \left| \exp\left(j\beta_0 p_2/2\right) - \exp\left(-j\beta_0 p_2/2\right) \right|, \\
&= 2V_g \sin\left(\beta_0 p_2/2\right)
\end{aligned}
\tag{4.98}
$$

where V_g is the magnitude of the voltage between the helix and the shield. It is assumed that the magnitude of the electric field in the gap is constant so that

$$
\begin{aligned}
\left|E_z\right| &= \frac{V_{AB}}{g_1} \cos\psi \\
&= \frac{2V_g \sin\left(\beta_0 p_2/2\right)}{g_1} \cos\psi.
\end{aligned}
\tag{4.99}
$$

Then the magnitude of the fundamental space-harmonic is from (4.37)

$$
E_0 = \frac{2V_g \sin\left(\beta_0 p_2/2\right)}{g_1}\left(\frac{\sin\left(\beta_0 g/2\right)}{\beta_0 g/2}\right)\cdot\frac{g}{p}\cdot\cos\psi.
\tag{4.100}
$$

This can be written as

$$
E_0 = \beta_0 V_g\left(\frac{\sin\left(\beta_0 p_2/2\right)}{\beta_0 p_2/2}\right)\left(\frac{\sin\left(\beta_0 g/2\right)}{\beta_0 g/2}\right)\cos^2\psi
\tag{4.101}
$$

and, from (4.17), the Pierce impedance is

$$
Z_P = \left[\left(\frac{\sin\left(\beta_0 p_2/2\right)}{\beta_0 p_2/2}\right)\left(\frac{\sin\left(\beta_0 g/2\right)}{\beta_0 g/2}\right)\frac{\cos^2\psi}{I_0\left(\gamma a_1\right)}\right]^2 Z_c,
\tag{4.102}
$$

where a_1 is the inner radius of the helix and Z_c is the characteristic impedance of the tape helix slow-wave structure given by (4.9) so that

$$
Z_c = \frac{v_p}{v_g}\sqrt{\frac{\alpha_L L}{\alpha_{C0} C}}.
\tag{4.103}
$$

This model is found to agree with experimental data within the limits of experimental error. Figure 4.20 shows a comparison between the calculated and measured values of the phase velocity and the Pierce impedance (measured by three different methods) for a typical structure (see Worksheet 4.5).

Tape helix slow-wave structures with dielectric loading can be modelled by considering the effect of the dielectric rods on the shunt capacitance. The structure is first modelled without the dielectric as described above. It is assumed that the value of α_L is unchanged by the addition of the dielectric. The effect of dielectric loading on the shunt capacitance cannot be computed by multiplying $C(0)$ by ε_2 because the presence of the dielectric causes the charges on the surfaces to be redistributed.
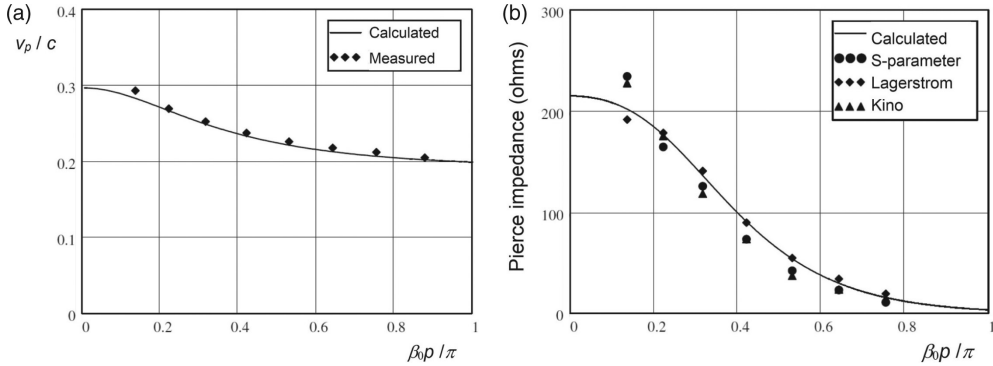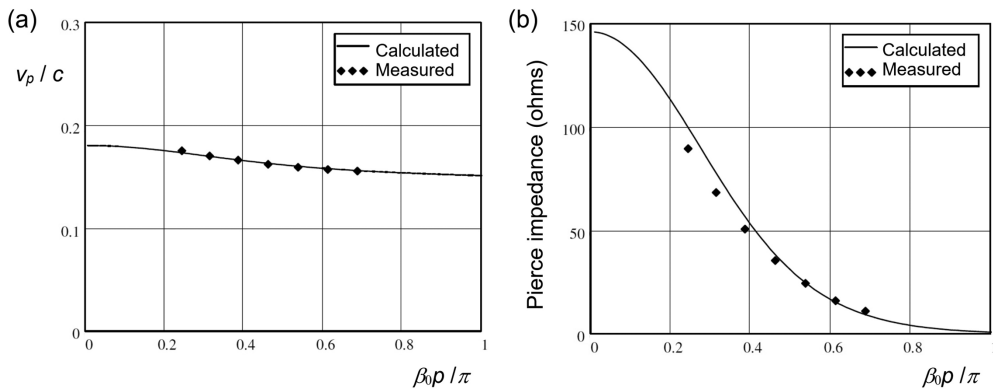
Figure 4.20: Comparisons between calculated and measured properties of a typical tape helix slow-wave structure without dielectric loading: (a) phase velocity, and (b) Pierce impedance. (a = 4.7 mm, p = 6.0 mm, s = 14.0 mm, w = 2.4 mm, t = 1 mm).

Instead, the shunt capacitance per unit length $\left(C_K\left(\beta_0 p\right)\right)$, given by (4.44), can be calculated using finite difference solutions at $\beta_0 p = 0$, $\pi/2$ and π to find the first three terms of the series (see Worksheet 4.4). Wedge shaped support rods are treated as a uniform dielectric of effective permittivity ε_2 given by (4.82). Values of capacitance calculated in this way are unsatisfactory close to $\beta_0 p = \pi$ where the quasi-static approximation starts to break down. Thus it is better to use the capacitance of the sheath helix given by (4.81) with a correcting factor α_C given by

$$\alpha_C = \frac{C_K\left(0\right)}{C_{00}\left(\varepsilon_2\right)}. \tag{4.104}$$

The effective permittivity is given by (4.47) and the phase velocity and Pierce impedance are calculated as before. Figure 4.21 shows the comparison between the results of calculations and experimental data for SWS-2 in [47]. This structure had rectangular support rods that may only been in contact with the helix along narrow lines. The best agreement between calculations and measurements was obtained by assuming wedge-shaped rods occupying the same fraction of the perimeter of the shield as the rectangular rods. The accuracy obtained using this model is comparable with that given by the much more complicated analytical method in [47]. Thus the equivalent circuit gives insight into the properties of helix slow-wave structure, and a simple method for calculating them with useful accuracy. For example, it can be shown that, if the support rods are not in good contact with the surface of the helix, there can be a marked change in the phase velocity and the impedance. In one case an air gap between the helix and the dielectric equal to 5% of the helix radius was found to reduce the shunt capacitance by a factor of 0.75.

4.3.5   Couplers and Attenuators

When a helix slow-wave structure is used in a travelling-wave tube it is important that the terminations of the structure are correctly matched. The transition between

Figure 4.21: Comparisons between calculated and measured properties of a tape helix slow-wave structure with dielectric loading: (a) phase velocity, and (b) Pierce impedance.

the helix and an external waveguide is normally made through the direct connection of the helix to a coaxial line. At low power levels this is connected directly to an external coaxial line. At high power levels it may be coupled outside the vacuum envelope to a rectangular, or ridge, waveguide. The characteristic impedance of the helix slow-wave structure can be computed from (4.103), if the effects of the support rods are included, and this can be used to design a broad-band transition to the coaxial line [60, 61]. A method for the detailed design and optimisation of coaxial couplers is discussed in [62].

Where the helix is divided (severed) into two or more sections (see Chapter 14) it is necessary to achieve a good match at each sever. This is achieved by placing a graduated attenuating coating on the support rods [63–67]. The helix may be severed physically but this is not necessary if sufficient attenuation can be introduced to prevent the passage of signals through the attenuator.

4.4   Ring-Bar and Ring-Loop Structures

The helix slow-wave structure is not suitable for use at high beam voltages because the interaction impedance falls rapidly as the pitch angle is increased. There is also an increased tendency to backward-wave oscillation. Better impedance can be obtained, at the expense of reduced bandwidth, using helix-derived structures such as the ring-bar structure shown in Figure 4.22. This structure evolved from a contra-wound bifilar helix [68, 69]. It comprises a series of rings which are connected by bars placed alternately at diametrically opposed positions, and can readily be constructed by making transverse cuts in a metal tube. The structure is supported within a concentric metal tube by two, three or four dielectric support rods [18, 70, 71]. The interaction impedance of the ring-bar structure in the fundamental mode is greater than that of a single helix, while that in the space harmonics is less.
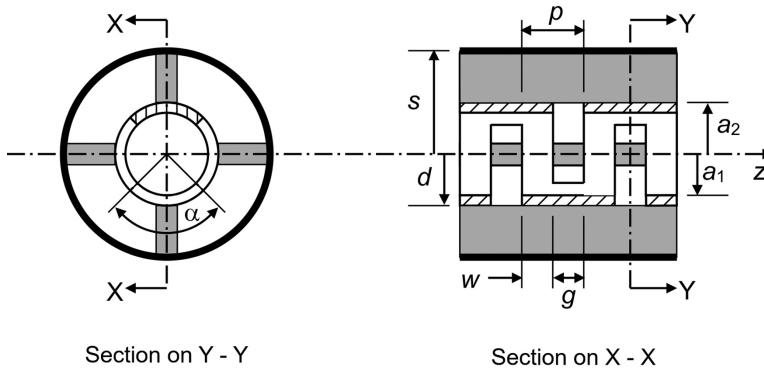
Figure 4.22: The ring-bar structure.

Thus the tendency to backward-wave oscillations in the space-harmonic modes is reduced. The thermal performance of this structure is similar to that of the helix structure so its chief use is in pulsed tubes.

Most analytical models of the ring-bar structure have been based on the methods for tape helix slow-wave structures described above [70–72]. Like the helix models from which they are derived these models involve considerable mathematical complexity but do not achieve close agreement with experimental results. The dielectric support rods are replaced by an equivalent uniform dielectric as before. A different approach, described by Ash et al. [18], is based on an analysis of a developed structure in which parallel metal strips are supported on dielectric bars, at right angles to them, in a manner similar to that used for the tape helix by Onodera and Raub [59]. This model showed useful agreement with experimental results but it does not seem to have been pursued by other authors. The ring-bar structure has also been modelled successfully using electromagnetics software [70, 73].

The main mode of propagation of the ring-bar structure is symmetrical about the longitudinal plane passing through the centres of the bars. The current passing along a bar is divided equally between the two halves of the structure which are therefore in parallel with one another [68]. A footnote in [18] observes that each half of the structure can be regarded as a meander line, but this approach does not seem to have been used elsewhere. It does, however, provide a simple method for modelling this structure using the theory of Section 4.2.2. The coefficients of capacitance in (4.39) and (4.44) can be found using a finite-difference model of isolated rings, both with and without uniform dielectric loading, in the same manner as for the tape helix (see Worksheet 4.4). The capacitance per unit axial length is converted to capacitance per unit azimuthal distance by multiplying the results of the finite-difference calculations by $p/2\pi a$ where p is the pitch of the rings and a their mean radius. The dispersion diagram of the structure can then be computed directly from the equations in Section 4.2.2 with results similar to those in Figure 4.11 (see Worksheet 4.6). Figure 4.23(a) shows a comparison between the

phase velocity computed in this way and the experimental results reported in [70]. In this calculation it has been assumed that the effective azimuthal length of each half ring is πa. The agreement between calculation and measurement is comparable with the analytical results in [70, 71].

To compute the Pierce impedance of this structure we regard it as a pair of meander lines connected in parallel. The power flow in each line is given by (4.60) where the lengths of the lines are given by $l = \pi a$ and the ratio B/A is given by (4.56). The electric field in the gaps between the lines is given by (4.61). This may be expressed as a set of azimuthal harmonics of which only the fundamental has non-zero field on the axis. The mean amplitude of the field in the gaps is then

$$E_{g0}(\phi) = \frac{1}{\pi a} \int_{-\frac{\pi a}{2}}^{\frac{\pi a}{2}} E_g(x, \phi)\, dx. \tag{4.105}$$

The magnitude of the fundamental space harmonic field on the axis is obtained using (4.37) as

$$E_0(\phi) = \frac{E_{g0}(\phi)}{I_0(\gamma_0 a)} \cdot \frac{\sin(\beta_0 g/2)}{\beta_0 g/2} \cdot \frac{g}{p}, \tag{4.106}$$

where $\gamma_0 = \sqrt{\beta_0^2 - \omega^2/c^2}$. Finally, from (4.17) the Pierce impedance is

$$Z_P = \frac{E_0^2}{2\beta_0^2 P_z}, \tag{4.107}$$

where P_z is the total power flow in the structure. Figure 4.23(b) shows a comparison between the Pierce impedance calculated in this way with the experimental results given in [70]. Further experimental results are given in [18, 74].

Although this simple model gives good agreement with experimental results, it does not take account of the widths of the bars and the capacitance between the bars and the shield. That capacitance is affected by the positions of the support rods relative to the bars resulting in changes in the Pierce impedance of the structure [18]. The properties of the structure can also be changed by changing the shape of the bars as in the ring-loop structure, shown in Figure 4.24. This structure has increased interaction impedance in the fundamental mode, while that of the backward wave mode decreases or even vanishes [75, 76]. It can be manufactured by making the rings and connecting loops from sheet and then folding them using a mandrel to ensure alignment [77]. The method outlined above can be used to model this structure if the short-circuiting bars are replaced by series inductances [75]. Other, related, structures are the stub-supported ring-bar structure [68] and the ring-plane structure [78]. In both cases the rings are supported by metal stubs or planes which are connected to the shield. When stubs are used they are chosen to have an electrical length of 90° so that they present an open circuit to the rings. Compared with the ring-bar structure these structures have improved thermal performance at the expense of greater dispersion.

(a)

v_p / c

(b)

Pierce impedance (ohms)

$\beta_0 p / \pi$ $\beta_0 p / \pi$
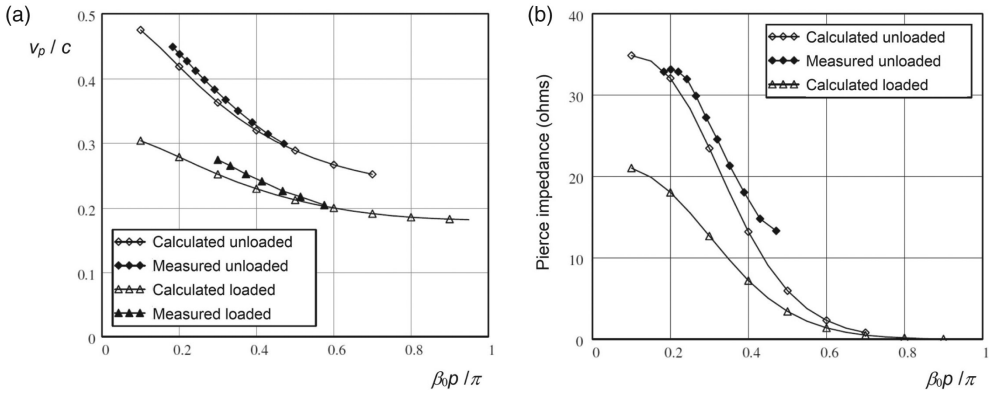
Figure 4.23: Computed properties of a ring-bar slow-wave structure both with and without dielectric loading compared with the experimental results in [70]: (a) dispersion curves, and (b) Pierce impedance.

Section on Y - Y   Section on X - X
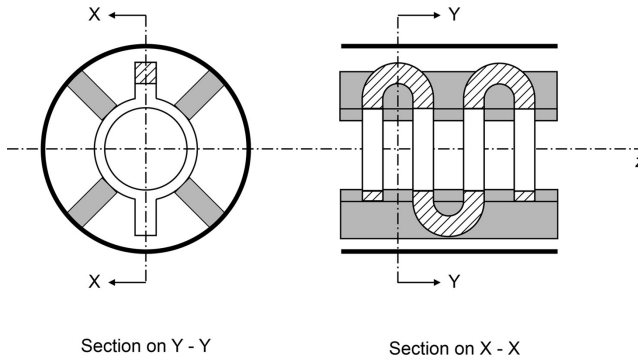
Figure 4.24: The ring-loop slow-wave structure.

4.5   Waveguide Slow-Wave Structures

The structures which have been described so far are not suitable for use at high average power levels because of their poor thermal performance. Structures which are based on metallic waveguides are much better from that point of view, even though they have greater dispersion than helix and helix-derived structures. The ladder line structure discussed in Section 4.2.1 can be considered to fall into this category if the bars are thick and the electron beam passes through holes in them.

4.5.1   The Folded Waveguide Structure

A slow-wave structure can be made using a serpentine rectangular waveguide, as shown in Figure 4.25 [79]. A simple analysis of this structure supposes that a TE_{01} rectangular waveguide mode propagates along the mid-line of the waveguide
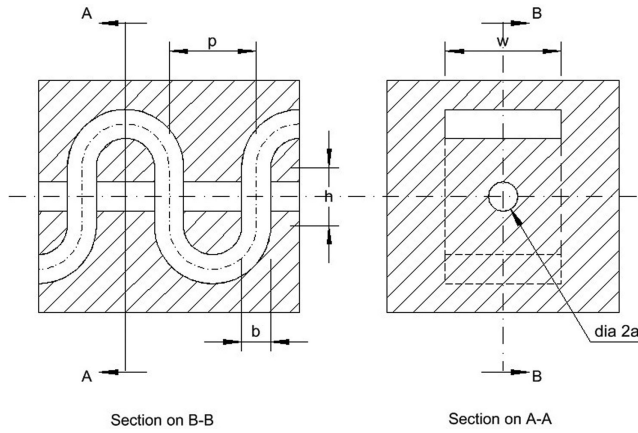
Figure 4.25: Arrangement of a serpentine (folded) waveguide slow-wave structure with swept bends.

(shown by the chain dotted line). At the cut-off frequency of the waveguide the phase is the same everywhere. However, the folding of the waveguide means that the phases in successive gaps are in anti-phase with one another. Thus the phase shift between successive gaps is π plus the phase shift along the serpentine path at the same frequency. The length of the serpentine path between adjacent gaps is $(h + \pi p/2)$ so that the phase shift perceived by an electron moving along the axis is

$$\beta_n p = \beta_g \left(h + \pi p/2\right) + \left(2n+1\right)\pi, \tag{4.108}$$

where β_g is the propagation constant for the guide and n is an integer. Making use of (2.11) we have

$$\beta_n p = \beta_C \left(\frac{\omega^2}{\omega_C^2} - 1\right)^{\frac{1}{2}} \left(h + \pi p/2\right) + \left(2n+1\right)\pi, \tag{4.109}$$

where $\beta_C = \omega_C/c = \pi/w$. Figure 4.26 shows typical dispersion curves of the space harmonics of this structure in the first four Brillouin zones. It can be seen that the $n = 0$ curve represents a backward wave in the first Brillouin zone. Thus the electron velocity must be chosen to be synchronous with the wave on the structure in the second Brillouin zone for forward-wave interaction. The total impedance of this structure is equal to the impedance Z_{PV} of the waveguide which can be calculated using (2.64), noting that a in that equation is the width of the waveguide denoted here by w. The ratio of the height of the waveguide to the pitch of the structure (b/p) can be chosen to maximise the impedance of the first space harmonic.

The derivation of (4.109) assumed that the bends do not present mismatches to the wave. In practice both the bends, and the beam holes, introduce small mismatches that cause stop bands to occur wherever the dispersion curves cross [79].
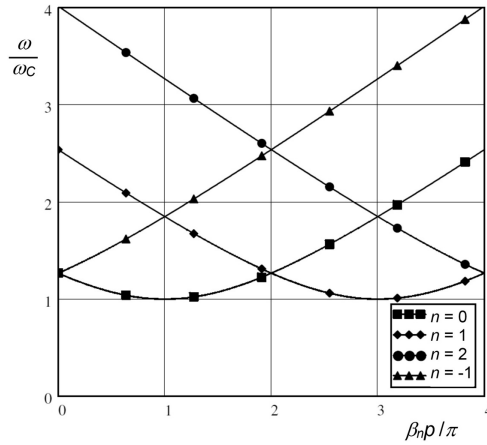
Figure 4.26: Dispersion curves of the serpentine (folded) waveguide structure shown in Figure 4.25.

At the edges of a stop band there is a standing wave in the serpentine waveguide and the group velocity is zero. At one edge of the stop band the electric field is maximum on the axis, and at the other edge it is zero. Thus the total impedance is infinite at one band edge and zero at the other.

An alternative form of the folded waveguide structure in which the height of the waveguide can be different in different sections is shown in Figure 4.27(a). This can be modelled using the equivalent circuit in Figure 4.27(b) which represents the structure between the centres of adjacent gaps. The straight lengths of waveguide are represented by transmission lines having characteristic impedances $Z_{g1,g2}$ and electrical lengths $\theta_{1,2}$. The bends are represented by an equivalent network B [80, 81]. The parameters of this network can be computed simply using quasi-static analysis. The basic properties of the structure can be illustrated using a model that ignores the impedances of the bends. Then the electrical lengths of the two sections of waveguide are

$$\theta_1 = \beta_g h/2 \quad \text{and} \quad \theta_2 = \beta_g p. \tag{4.110}$$

The value of β_g is the same in each case because we have assumed that both sections have the same width (w). The transfer matrix for a section of waveguide is

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & jZ_g\sin\theta \\ jY_g\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [T] \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}. \tag{4.111}$$

Using this definition we may write the transfer matrix for the section of the structure between the centres of adjacent gaps as

$$[T] = [T_1][T_2][T_1]. \tag{4.112}$$
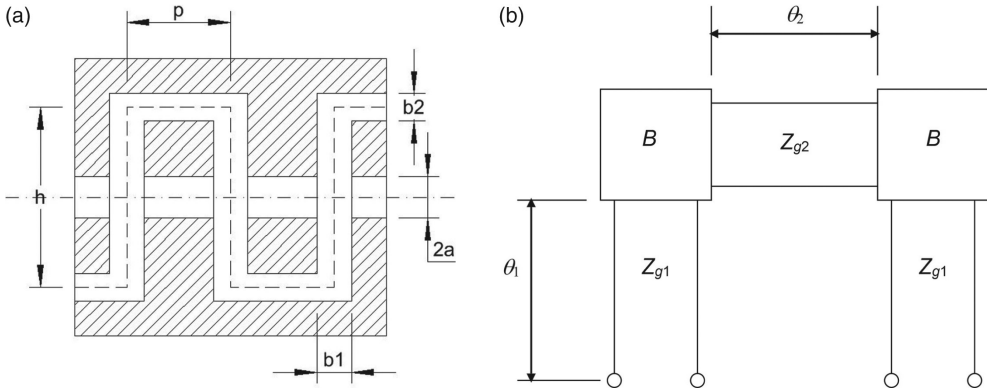
Figure 4.27: Folded waveguide structure with right-angle bends: (a) cross-sectional view, and (b) equivalent circuit.
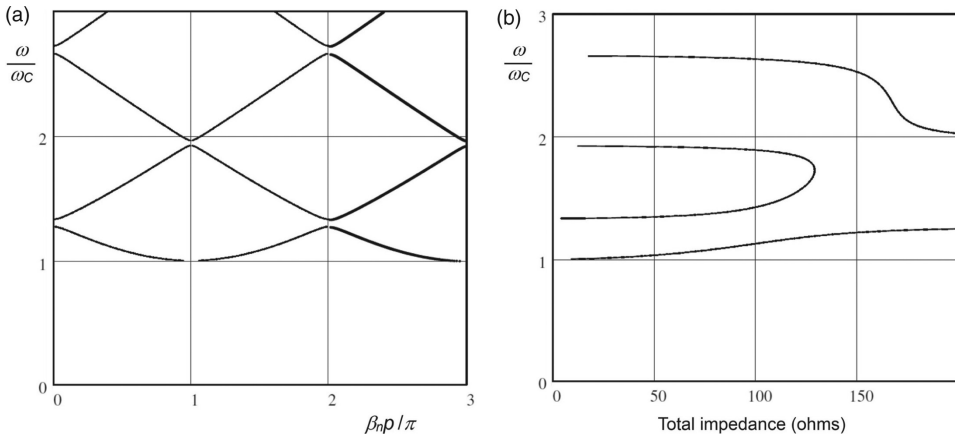
Figure 4.28: Characteristics of a typical folded waveguide structure $(b_2/b_1 = 1.2)$: (a) dispersion, and (b) total impedance.

The numerical values of the elements of $[T]$ can be computed from these equations and the properties of the structure are given by

$$\phi = \arccos T_{11} \qquad (4.113)$$

and

$$Z_T = \sqrt{\frac{T_{12}}{T_{21}}}. \qquad (4.114)$$

Figures 4.28(a) and (b) show the dispersion and impedance curves for a structure in which the impedance of the longitudinal waveguides is higher than that of the transverse waveguides $(b_2 > b_1)$. The impedance of the lowest pass-band tends to infinity at the upper cut-off frequency whilst that of the next higher band tends to zero at its lower cut-off. Thus the standing wave maximum is at the centre of

the beam hole at the top of the first pass-band and at the centre of the longitudinal waveguide at the bottom of the second pass-band. When the impedance of the slots is made less than that of the cavities then the situation is reversed. The properties of folded waveguide slow-wave structures can be explored using Worksheet 4.7.

Folded waveguide structures are usually designed so that $Z_{g2} > Z_{g1}$ because this gives the best interaction impedance in the lowest pass-band. Such structures are sometimes referred to as *normal* structures. When $Z_{g2} < Z_{g1}$ the structure is said to be *inverted*. In both cases the cold bandwidth of the lowest mode is similar, and the usable hot bandwidth is always somewhat less than this. If $Z_{g1} = Z_{g2}$ then, at least in theory, the stop bands between the modes vanish [82]. Such a structure, called a *coalesced mode structure*, would appear to offer the possibility of making a tube with a much greater bandwidth. But, because of the discontinuities caused by the bends and the beam holes, and because of the effects of manufacturing tolerances, it is very difficult to make a structure which does not have any stop bands. Even a very narrow stop band would tend to cause a dip in the gain of the tube at that frequency with the risk of oscillation because of the very high interaction impedance at one of the band edges.

4.5.2 Helical Waveguides

By analogy with the tape helix slow-wave structure it would appear that a circular waveguide having a helical groove on its inner surface should be a wide-band high-power slow-wave structure. Structures of this type have been studied by a number of authors (see for example [83–86] and the references therein). Helical waveguides are employed in gyro-TWTs but do not appear to have been used in conventional TWTs.

4.6 Coupled-Cavity Slow-Wave Structures

The rectangular folded waveguide structure shown in Figure 4.27(a) can, alternatively, be regarded as a structure in which rectangular cavity resonators are coupled together by slots in the walls which separate them. The folded waveguide structure is chiefly of interest for use at millimetre wavelengths where the dimensions are very small. At microwave frequencies improved performance can be achieved using structures having more complex geometries. These can be divided into two classes: space harmonic structures, and forward fundamental structures. Space harmonic structures that resemble the folded waveguide structure cannot achieve the highest power levels because operation in the first space harmonic limits the beam voltage that can be used. To achieve higher power levels it is necessary to use structures in which there is an additional phase reversal between the cavities so that the fundamental space harmonic is forward wave.

The properties of coupled-cavity structures can be found by measurement as described in Section 4.7.1, or by modelling using computational electromagnetics software [87–90]. Modern software allows the use of periodic boundary conditions
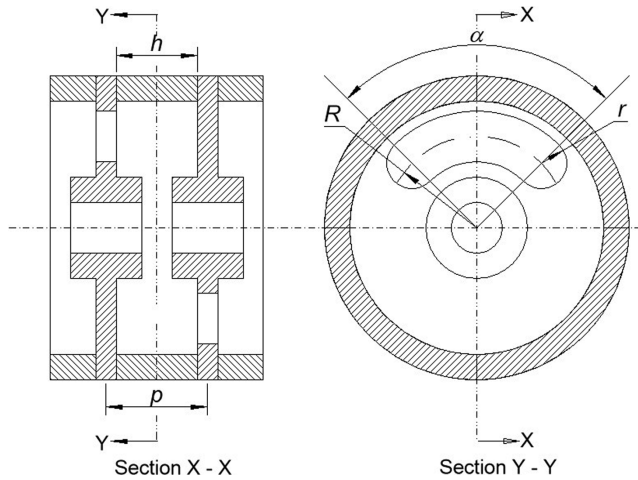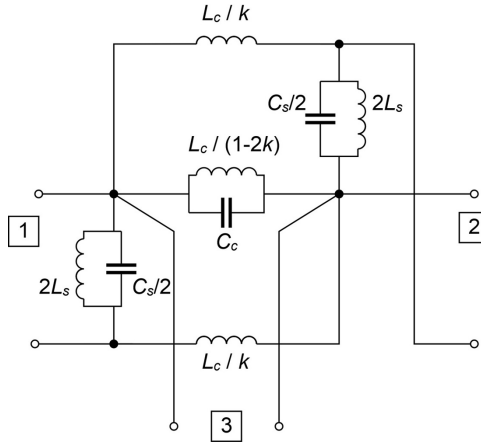
Figure 4.29: Arrangement of a space-harmonic slow-wave structure (copyright1983 IEE, reproduced, with permission, from [91]).

so that it is sufficient to model a single period of the structure, and its properties can be determined for any value of the phase change per cavity. Computer modelling also provides detailed information about the fields within the structure so that the coupling impedance can be calculated. Further insight into the properties of some of the more important types of coupled-cavity structure can be obtained from equivalent circuit models as described below. These circuits may be used for interpolation on computed, or measured data. They are also useful for rapid parametric calculations.

4.6.1   Space Harmonic Structures

Figure 4.29 shows the arrangement of a typical space harmonic ('Hughes') structure for use at microwave frequencies. The cavities are circular and re-entrant in shape and are coupled together by large kidney-shaped slots. The slots are arranged alternately on either side of the axis in successive coupling plates. Although this structure can be considered as a somewhat distorted folded waveguide, the usual conceptual approach is to consider the effects of coupling on the properties of a chain of identical resonant cavities.

The space-harmonic slow-wave structure can be represented by the equivalent circuit shown in Figure 4.30 [64, 92]. Ports 1 and 2 are connected to the adjacent cavities whilst port 3 is connected to the electron beam. The cavity inductance L_c is divided into three parts so that a fraction k of the current circulating in the cavity is intercepted by each slot. When the fields in adjacent cavities are in phase with one another the net current intercepted by the slots is zero and the frequency is equal to the resonant frequency of an isolated cavity. When the fields in adjacent cavities are in antiphase a net current is intercepted by the slots. If the frequency is below the resonant frequency of the slots they present an inductive reactance to the current and so reduce the frequency of the mode. This is consistent with a backward-wave characteristic in the first Brillouin zone.

Figure 4.30: Equivalent circuit of a space-harmonic slow-wave structure (copyright 1983 IEE, reproduced, with permission, from [91]).

The equivalent circuit in Figure 4.30 models the lowest mode quite accurately [91] but fails for the higher modes. The reason is easily understood if the figure is redrawn in the form of sections of the equivalent circuit of the TE_{01} waveguide mode in series with each other. These sections are valid only when the phase change per section is small, and that assumption is less valid as the frequency rises. Analysis of the equivalent circuit results in the dispersion equation

$$\cos\phi = 1 - \frac{1}{ka_k}\left\{1-\left(\frac{\omega}{\omega_c}\right)^2\right\}\left\{1+a_k-\left(\frac{\omega}{\omega_s}\right)^2\right\}, \tag{4.115}$$

where

$$\omega_c = 1/\sqrt{L_c C_c} \tag{4.116}$$

$$\omega_s = 1/\sqrt{L_s C_s} \tag{4.117}$$

$$a_k = 2k\left(L_s/L_c\right) \tag{4.118}$$

and the subscripts c and s refer to the cavity and the slots, respectively. The frequencies at the edges of the stop band are the solutions of (4.115) when $\phi = 0$

$$\omega = \omega_c \quad \text{and} \quad \omega = \left(1+a_k\right)^{\frac{1}{2}}\omega_s. \tag{4.119}$$

The total impedance of the slow-wave structure is given by

$$Z_T = \frac{2\left(R/Q\right)_c}{ka_k\sin\phi}\cdot\frac{\omega}{\omega_c}\left[1+a_k-\left(\omega/\omega_s\right)^2\right] \tag{4.120}$$

where $\left(R/Q\right)_c$ is the R/Q of the cavity, p is the structure pitch and v_g is the group velocity. Equations (4.115) and (4.120) can be fitted to experimental data.

The parameters of the circuit can be calculated with useful accuracy from the dimensions of the structure. The resonant frequency and R/Q of the cavity are obtained using the method of moments (see Section 3.5.1 and Worksheet 3.3) or, with less accuracy, using Fujisawa's formulae [93] (see Section 3.5.2 and Worksheet 3.4). The effective slot length is defined as the area of the slot divided by its width (see Figure 4.29)

$$l_s = \alpha R + \frac{\pi r}{2}$$ (4.121)

and the slot is assumed to be resonant when it is a half-wavelength long so that

$$f_s = c/2l_s.$$ (4.122)

The slot is treated as a short length of waveguide carrying the TE_{01} mode. In order to find the R/Q of the slot it is necessary to define its effective thickness. This is increased beyond the physical thickness by the fringing of the fields around the slot. It is found that a useful approximation to the effective slot thickness is

$$t_s = p - h + 2r$$ (4.123)

and then

$$(R/Q)_s = \sqrt{\frac{\mu_0}{\varepsilon_0}} \cdot \frac{2r}{\pi t_s}.$$ (4.124)

The coupling factor is defined by

$$k = \frac{l_s}{2\pi R}.$$ (4.125)

The dispersion curve and the total impedance can then be calculated using (4.115) and (4.120). Figure 4.31 shows typical dispersion curves, and impedance plots, calculated for two different slot angles with experimental results for comparison (see Worksheet 4.8). This model is a simplification and an improvement of the one previously described by the author, but alternative assumptions about the slot length and thickness are possible [94]. The bandwidth of the structure increases as the slot angle (α) is increased. However, if α is too great the structure becomes inverted and the coupling impedance is reduced. It might be thought that the coupling between the cavities could be increased, without inverting the structure, by using two slots per cavity in place of one. The result, known as the Chodorow-Nalos structure when the slots are in line with each other, actually has a narrow bandwidth because of direct coupling between the slots [95]. This is a form of ladder-line which is sometimes useful at millimetre wave frequencies. An alternative in which the slots are staggered by rotating successive coupling plates by 90° has better bandwidth [96].

4.6.2   The Cloverleaf Structure

At very high powers it is desirable to use a structure in which the fundamental mode is forward wave. One such is the cloverleaf structure in which the cavities have the shape
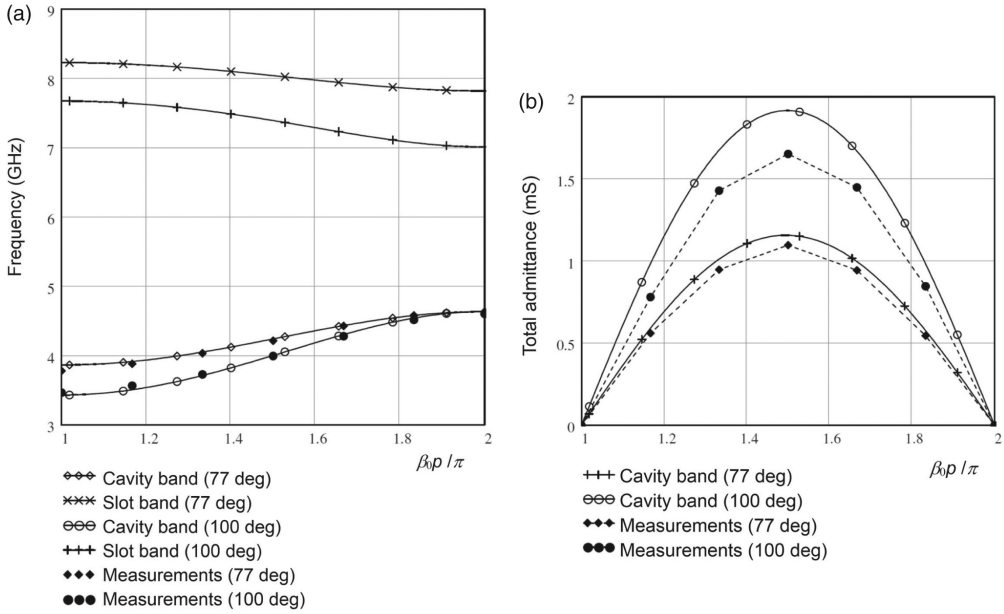
Figure 4.31: Typical results for a space-harmonic slow-wave structure: (a) dispersion, and (b) total admittance.
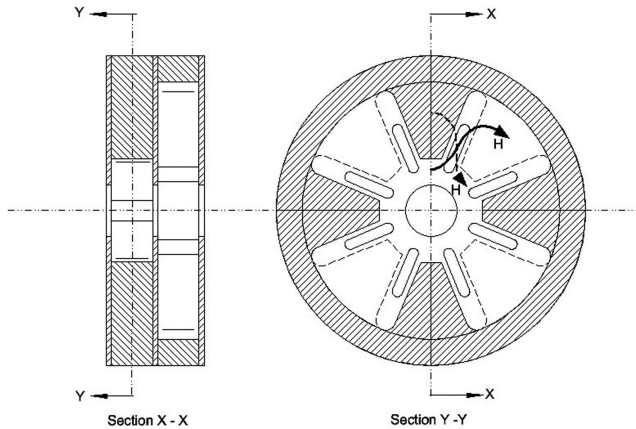
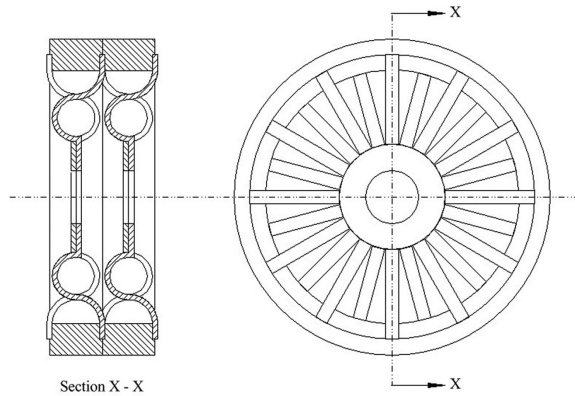**Figure 4.32:** General arrangement of a cloverleaf slow-wave structure (copyright 1983 IEE, reproduced, with permission, from [91]).

with four or more noses, shown in Figure 4.32 [97]. Successive cavities are rotated with respect to each other so that the noses in one cavity correspond to the gaps in the next. The cavities are coupled by radial slots that lie close to the noses, as shown. If the cavities are excited in phase with each other, then the magnetic field lines circulate in the same direction in adjacent cavities, as shown by the solid and dashed arrows. The associated azimuthal current components are in the same direction, and the

current intercepted by each slot is non zero. The slot inductances are then effectively in series with the inductance of the cavity, and the resonant frequency is reduced. When the fields in adjacent cavities are in antiphase, the intercepted currents cancel, and the resonant frequency is close to that of the cavity. Hence the fundamental mode of the structure is a forward wave in the first Brillouin zone.

The cloverleaf structure can be represented by the equivalent circuit in Figure 4.33 which is very like that in Figure 4.30 [92]. This time a fraction k of the current in each cavity is intercepted by all the slots because they are in line with one another. The crossing of the connections on the right-hand side of the figure represents the phase reversal brought about by the geometry. Analysis of this circuit produces

$$\cos(\phi) = -1 - \frac{\left[1-(\omega/\omega_c)^2\right]\left[1-(\omega/\omega_s)^2\right]}{a_k\left[1-k-(\omega/\omega_c)^2\right]} \qquad (4.126)$$

and

$$Z_T = \frac{2k(R/Q)_c}{a_k\sin\phi}\cdot\frac{\omega}{\omega_c}\cdot\frac{1-(\omega/\omega_s)^2}{\left[1-k-(\omega/\omega_c)^2\right]^2}\,. \qquad (4.127)$$

The phase change per cavity is π when $\omega = \omega_c$ and $\omega = \omega_s$. The cavities have properties similar to those of the pill-box resonator shown in Figure 2.2. The cavity resonant frequency is approximately that of a pill-box whose diameter is 1.1 times that of the diameter across the noses. The R/Q is slightly less than that of a pill-box cavity because of the additional energy storage in the lobes of the cavity between the noses. The slots resonate when their lengths are close to half of a free-space wavelength. They are weakly coupled to each other by the cavities so that the slot band is narrow. This structure has higher-order modes corresponding to the higher-order TM modes of the cavity. These are weakly coupled to the slots and therefore have narrow bandwidth, and low group velocity. The TM_{02} mode is potentially troublesome because E_z is non zero on the axis, and the mode has high interaction impedance. Because of the shape
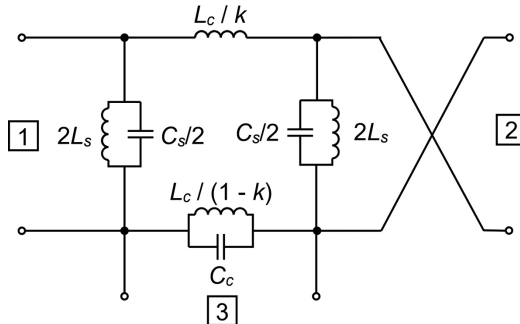
Figure 4.33: Equivalent circuit of a cloverleaf slow-wave structure (copyright 1983 IEE, reproduced, with permission, from [91]).

of the cavity the outer maximum of this mode may break up into a set of individual maxima at the centres of the lobes between the cloverleaf noses [64].

4.6.3 The Centipede Structure

Another forward-fundamental structure is the centipede structure shown in Figure 4.34 [97–100]. In this structure there is inductive coupling between the loops and the cavities.

The structure can be represented by the equivalent circuit shown in Figure 4.35 [91]. Note that it is necessary to include direct loop to loop coupling in the model. The connections on the right-hand side of the diagram are reversed to account for the phase reversal produced by the shape of the loops. An alternative equivalent circuit is described in [101].

Analysis of the circuit in Figure 4.35 results in the dispersion equation

$$\cos\phi = -1 + \frac{1}{2} \cdot \frac{\left[(\omega_c/\omega)^2 - 1\right]\left[(\omega_l/\omega)^2 + 2k_2 - 1\right]}{k_1^2 + k_2\left[(\omega_c/\omega)^2 - 1\right]}, \tag{4.128}$$

Figure 4.34: General arrangement of a centipede slow-wave structure (copyright 1983 IEE, reproduced, with permission, from [91]).
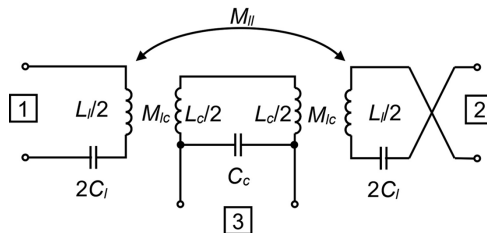
Figure 4.35: Equivalent circuit of a centipede slow-wave structure (copyright 1983 IEE, reproduced, with permission, from [91]).

where the subscripts c and l refer to the cavity and the loops, respectively,

$$k_1 = \sqrt{L_c L_l}/M_{lc} \qquad (4.129)$$

and

$$k_2 = L_l/M_{ll}. \qquad (4.130)$$

The band edges at $\phi = \pi$ are related to the cavity and loop resonant frequencies by

$$\omega = \omega_c \quad \text{and} \quad \omega = \omega_l/\sqrt{1 - 2k_2}. \qquad (4.131)$$

The cavity parameters are close to those for a pill-box cavity having the same dimensions as the cavities of the structure. The parameters of the loops must generally be determined by fitting the model to experimental data. The dispersion diagram of the centipede structure is rich in higher-order modes with varying rotational symmetries. The uniform distribution of the loops around the edge of the cavity allows modes such as the TM_{11} loop mode to exist. Pairs of cavity and loop modes having the same rotational symmetry are normal (inverted) depending on whether the lower cut-off frequency of the loop mode lies above (below) the corresponding resonant frequency of the cavities. The total impedance of the centipede structure is given by

$$Z_T = -\frac{k_1^2 \left(R/Q\right)_c}{\sin\phi} \cdot \left(\frac{\omega_c}{\omega}\right)^3 \frac{\left(\omega_l/\omega\right)^2 + 2k_2 - 1}{\left[k_1^2 + k_2\left(\left(\omega_c/\omega\right)^2 - 1\right)\right]^2}. \qquad (4.132)$$

Figure 4.36 shows a typical dispersion diagram for a centipede structure calculated using (4.128). It has been assumed that the dispersion diagrams for the higher-order modes obey the same equation with the appropriate cavity resonant frequency. The loop frequencies of the TM_{11} and TM_{21} modes were increased from that for the TM_{01} mode by factors of 1.2 and 1.3, respectively, to achieve a better match with experimental results. The centipede structure is smaller than a cloverleaf structure designed to operate at the same frequency but its heat transfer properties are inferior.

4.6.4   Termination of Coupled-Cavity Slow-Wave Structures

The input and output connections to coupled-cavity slow-wave structures are commonly made using coupling loops, or inductive irises, as described in Section 3.6. Additional elements may be used to achieve a good match over the intended working band of the tube [4, 64, 102]. The use of transition cavities to obtain a good match is discussed in [103]. Direct coupling, from reduced height waveguide, is an obvious choice for folded waveguide structures, space harmonic structures and cloverleaf structures [104]. An alternative is to make use of a door-knob transition (Figure 2.25) since a circular TM_{01} mode is to be launched in the structure [98].
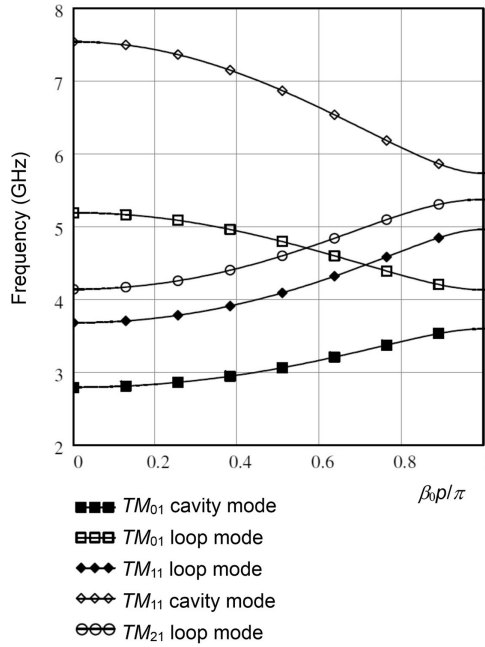
Figure 4.36: Typical mode diagram of a centipede slow-wave structure $(k_1 = 0.3, k_2 = 0.1)$.

In coupled-cavity travelling-wave tubes it is usual for there to be a break in the structure at each sever. There may be a transition to an external waveguide, allowing the sever power to be dissipated in waveguide loads either inside, or outside, the vacuum envelope. This arrangement adds to the size and complexity of the tube. Alternatively the structure may be terminated internally by adding lossy ceramics to one or more cavities [4].

4.7   Measurement of the Properties of Slow-Wave Structures

Although the properties of slow-wave structures can usually be calculated to good accuracy using computational electromagnetics, it is still desirable to be able to check the results by comparison with cold test measurements on actual, or scaled, slow-wave structures. It is normally possible to measure the phase velocity with good accuracy, but accurate measurement of the coupling impedance presents greater difficulties.

4.7.1   Measurements on Coupled-Cavity Slow-Wave Structures

Measurements on coupled-cavity slow-wave structures are made using resonant sections [64, 105]. The short-circuit planes must coincide with planes of symmetry. For some structures (e.g. cloverleaf, centipede) the mid-plane of the cavity

can be used, giving a stack with half-height end cavities. For space harmonic structures the symmetry planes are at the centres of the slots. The structure may be at full-scale, or a scaled-up version may be made to make the measurements easier. Experimental measurements only yield figures for particular points on the dispersion curve. Five or six cavities are usually sufficient, and the properties at intermediate points are then found by numerical interpolation. One way to achieve this is to fit an equivalent circuit model to the experimental data [91]. The method of measurement is the same as that used for single cavities (see Section 3.7) with the difference that many more resonant modes exist. These can usually be identified by careful probing of the structure with a dielectric or metal rod through small holes drilled in the ends, or the sides, of the structure. If the resonances are displayed using a swept frequency source, the extent of the perturbation of a resonance by probing at different positions can be observed. Since the change in frequency is proportional to the stored electric energy it is possible to find the positions of field maxima and zeroes. The axial variation of modes for which $E_z \neq 0$ on the axis can also be determined by observing the frequency shifts as a small dielectric bead is moved along the axis of the structure. In some cases it is not possible to measure the band-edge frequencies directly and they must be deduced by extrapolation from the data points available. A suitable formula is

$$f_0 \simeq 0.1\left(15 f_1 - 6 f_2 + f_3\right), \tag{4.133}$$

where f_0 is the frequency at the band edge and $f_{1,2,3}$ are the frequencies of three resonances such that all four frequencies are determined at equal intervals in ϕ. This equation fits an even, fourth-order, polynomial to the data points.

The interaction impedance is measured experimentally by perturbing the resonant stack with a dielectric rod on its axis. A simple formula can be obtained for the calculation of the Pierce impedance [64] based upon the change in the capacitance resulting from the presence of the rod,

$$Z_P = \frac{4}{\varepsilon_0\left(\varepsilon_r - 1\right)\pi^2 d^2} \cdot \frac{\Delta f}{f} \cdot \frac{p}{\phi^2} \cdot \frac{\partial \phi}{\partial f}, \tag{4.134}$$

where d is the rod diameter, ε_r its relative permittivity and Δf the change in the resonant frequency when the rod is inserted. The last term is inversely proportional to the group velocity. This equation is only strictly valid in the limit of a vanishingly small rod. The measurement accuracy is limited by the need to use a rod which is large enough for the change in frequency to be measured accurately. Equation (4.134) also ignores the effects of space-harmonic fields which become more important as the rod diameter is increased. A more accurate formula which includes these effects is given in [106, 107]. Similar methods can be used for all other structures, such as ladder lines and ring-bar structures, which possess transverse symmetry planes.

4.7.2   Measurements on Helix Slow-Wave Structures

Helix slow-wave structures do not possess any planes of symmetry and it is therefore necessary to use non-resonant methods. The one end of the structure under test is connected to an external termination, and the other end is similarly terminated or fitted with an internal load. It is very important that there should be good matches at both terminations [108]. The input to the helix is connected to a vector network analyser. A perturbing object is passed through the helix and the rate of change of the phase of the reflected signal with the position of the object is just twice the phase shift per unit length of the structure. The perturbing object may be a bead or rod made from metal or dielectric material, or it may be small helix having a small number of turns [109–112]. The measurements are normally made on a full-scale section of helix. The helix diameter is typically a few millimetres or less, and that makes it more difficult to obtain accurate results. The equipment is normally automated so that a large number of measurements can be made to increase the accuracy of the results.

To measure the coupling impedance of the helix it is necessary to measure the change in the propagation constant produced by perturbation with a dielectric or metal rod. One way of doing this is to mount a bead, or a small helix, on a dielectric rod which passes through the whole length of the structure, and to measure the phase of the reflected signal [109–112]. The impedance can be calculated from the difference between the wavelengths measured with, and without, the rod. To get reasonable accuracy with this method it is necessary to measure the wavelengths very accurately. Lagerstrom [113] measured the Pierce impedance by observing the change in the phase of the signal transmitted through the helix when a dielectric rod was inserted in it. The measurement is subject to some uncertainty because end effects mean that it is not possible to determine the precise length of the rod in the helix. This difficulty can be overcome by progressively advancing the rod into the helix, and observing the rate of change of the phase of the transmitted signal with the position of the rod. It is necessary to correct the results for the effects of space-harmonic fields and transverse field components [108]. Since the diameter of the rod is normally smaller than the internal diameter of the helix it is very important to ensure that it is accurately aligned with the axis of the helix. A metal rod was used in place of the dielectric rod by Kino [114]. The perturbations are larger with this method, but the end effects are more serious. The phase velocity and interaction impedance can be measured simultaneously, using this method, by observing the phases of both the reflected and transmitted signals as functions of the position of the rod. The phase velocity, and interaction impedance, can also be determined from the amplitude and phase of the signal reflected by a spherical dielectric bead placed within the helix [110]. This (S-parameter) method makes use of the fact that the amplitude of the reflected signal is proportional to the local change in stored electric energy produced by the bead. The accuracy achieved is comparable with that obtained by other methods (see Figure 4.20).

References

[1] A. F. Harvey, 'Periodic and guiding structures at microwave frequencies', *IRE Transactions on Microwave Theory and Techniques,* vol. 8, pp. 30–61, 1960.

[2] W. Wang *et al.*, 'Review of the novel slow-wave structures for high-power traveling-wave tube', *International Journal of Infrared and Millimeter Waves,* vol. 24, pp. 1469–1484, 2003.

[3] N. W. McLachlan, *Bessel Functions for Engineers*. Oxford: Oxford University Press, 1954.

[4] A. S. Gilmour, Jr., *Principles of Traveling Wave Tubes*. Norwood, MA: Artech House, 1994.

[5] J. H. Booske *et al.*, 'Traveling-wave tubes', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE, pp. 171–245, 2005.

[6] S. Ramo *et al.*, *Fields and Waves in Communication Electronics*. New York: Wiley, 1965.

[7] H. Kosai *et al.*, 'Plasma-filled dielectric Cherenkov maser', *IEEE Transactions on Plasma Science,* vol. 18, pp. 1002–1007, 1990.

[8] E. I. Smirnova *et al.*, 'Design, fabrication, and low-power tests of a W-band omniguide traveling-wave tube structure', *IEEE Transactions on Plasma Science,* vol. 36, pp. 763–767, 2008.

[9] D. Y. Shchegolkov *et al.*, 'Millimeter-wave gain experiments with a wide-band omniguide traveling-wave tube', in *IEEE Thirteenth International Vacuum Electronics Conference (IVEC 2012)*, pp. 297–298, 2012.

[10] R. E. Collin, *Foundations for Microwave Engineering*. New York: McGraw-Hill, 1966.

[11] L. Brillouin, *Wave Propagation in Periodic Structures*. New York: Dover, 1953.

[12] R. W. Hamming, *Numerical Methods for Scientists and Engineers*. New York: McGraw-Hill, 1973.

[13] T. Edwards, *Foundations for Microstrip Circuit Design*, 2nd ed. Chichester, UK: Wiley, 1992.

[14] P. N. Butcher, 'The coupling impedance of tape structures', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 104, pp. 177–187, 1957.

[15] J. A. Weiss, 'Dispersion and field analysis of a microstrip meander-line slow-wave structure', *IEEE Transactions on Microwave Theory and Techniques,* vol. 22, pp. 1194–1201, 1974.

[16] A. Karp, 'Traveling-wave tube experiments at millimeter wavelengths with a new, easily built, space harmonic circuit', *Proceedings of the IRE,* vol. 43, pp. 41–46, 1955.

[17] J. Froom *et al.*, 'Ridge-loaded ladder lines', *IEEE Transactions on Electron Devices,* vol. 12, pp. 411–421, 1965.

[18] E. A. Ash *et al.*, 'Dispersion and impedance of dielectric-supported ring-and-bar slow-wave circuits', *Proceedings of the Institution of Electrical Engineers,* vol. 111, pp. 629–641, 1964.

[19] W. E. Courtney, *Printed-Circuit RF-Keyed Crossed-Field Amplifier*. Lexington, MA: Lincoln Laboratory, 1975.

[20] C. Chua and S. Aditya, 'A 3-D U-shaped meander-line slow-wave structure for traveling-wave-tube applications', *IEEE Transactions on Electron Devices,* vol. 60, pp. 1251–1256, 2013.

[21] B. R. Potter *et al.*, 'High-power printed circuit traveling wave tubes', in *International Electron Devices Meeting*, pp. 521–524, 1973.

[22] S. Sengele *et al.*, 'Microfabrication and characterization of a selectively metallized W-band meander-line TWT circuit', *IEEE Transactions on Electron Devices,* vol. 56, pp. 730–737, 2009.

[23] M. Sumathy *et al.*, 'Design and RF characterization of W-band meander-line and folded-waveguide slow-wave structures for TWTs', *IEEE Transactions on Electron Devices,* vol. 60, pp. 1769–1775, 2013.

[24] A. H. W. Beck, *Space-Charge Waves and Slow Electromagnetic Waves*. London: Pergamon Press, 1958.

[25] R. C. Fletcher, 'A broad-band interdigital circuit for use in traveling-wave-type amplifiers', *Proceedings of the IRE,* vol. 40, pp. 951–958, 1952.

[26] J. Arnaud, 'Circuits of travelling wave crossed field tubes', in E. Okress *et al.*, eds, *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 47–67, 1961.

[27] K. Xie, 'CVD diamond rod and its application in the high power TWT', in *Conference Digest: 25th International Conference on Infrared and Millimeter Waves*, pp. 343–344, 2000.

[28] G. Fleury *et al.*, 'Average power limits of brazed-helix TWT's', in *International Electron Devices Meeting*, pp. 806–809, 1980.

[29] R. Crivello and R. W. Grow, 'Thermal analysis of PPM-focused rod-supported TWT helix structures', *IEEE Transactions on Electron Devices,* vol. 35, pp. 1701–1720, 1988.

[30] Y. Han *et al.*, 'Thermal analysis of a helix TWT slow-wave structure', *IEEE Transactions on Electron Devices,* vol. 55, pp. 1269–1272, 2008.

[31] S. F. Paik, 'Design formulas for helix dispersion shaping', *IEEE Transactions on Electron Devices,* vol. 16, pp. 1010–1014, 1969.

[32] K. Tsutaki *et al.*, 'Numerical analysis and design for high-performance helix traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 32, pp. 1842–1849, 1985.

[33] B. Basu, 'Equivalent circuit analysis of a dielectric-supported helix in a metal shell', *International Journal of Electronics,* vol. 47, pp. 311–314, 1979.

[34] J. R. Pierce, *Traveling-Wave Tubes*. Princeton, NJ: D. van Nostrand, 1950.

[35] L. Loshakov and E. Ol'Derogge, 'Propagation of slow electromagnetic waves along a helix with dielectric supports', *Radio Engineering and Electronic Physics,* vol. 13, pp. 45–51, 1968.

[36] P. K. Jain and B. N. Basu, 'The inhomogeneous dielectric loading effects of practical helix supports on the interaction impedance of the slow-wave structure of a TWT', *IEEE Transactions on Electron Devices,* vol. 39, pp. 727–733, 1992.

[37] P. Hammond and J. K. Sykulski, 'Tubes and slices: a new way of teaching the principles of electric and magnetic fields', *IEEE Transactions on Education,* vol. 35, pp. 300–306, 1992.

[38] Q. Chen, 'Characteristic of vane-loaded helix slow-wave structure', in *Vakuumelektronik und Displays*, Garmisch-Partenkirchen, Germany, pp. 27–31, 1983.

[39] S. Sensiper, 'Electromagnetic wave propagation on helical structures (a review and survey of recent progress)', *Proceedings of the IRE,* vol. 43, pp. 149–161, 1955.

[40] S. Sensiper, 'Electromagnetic wave propagation in helical structures', ScD, Department of Electrical Engineering, MIT, 1951.

[41] L. Stark, 'Lower modes of a concentric line having a helical inner conductor', *Journal of Applied Physics,* vol. 25, pp. 1155–1162, 1954.

[42] D. A. Watkins and E. A. Ash, 'The helix as a backward-wave circuit structure', *Journal of Applied Physics,* vol. 25, pp. 782–790, 1954.

[43] P. K. Jain *et al.*, 'Effect of the finite thickness of the helix wire on the characteristics of the helical slow-wave structure of a traveling-wave tube', *IEEE Transactions on Electron Devices,* vol. 34, pp. 1209–1213, 1987.

[44] D. T. Swift-Hook, 'Dispersion curves for a helix in a glass tube', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 105, pp. 747–755, 1958.

[45] A. K. Sinha *et al.*, 'Simplified tape model of arbitrarily-loaded helical slow-wave structures of a travelling-wave tube', *IEE Proceedings H: Microwaves, Antennas and Propagation,* vol. 139, pp. 347–350, 1992.

[46] S. Ghosh *et al.*, 'Rigorous tape analysis of inhomogeneously-loaded helical slow-wave structures', *IEEE Transactions on Electron Devices,* vol. 44, pp. 1158–1168, 1997.

[47] S. D'Agostino *et al.*, 'Accurate analysis of helix slow-wave structures', *IEEE Transactions on Electron Devices,* vol. 45, pp. 1605–1613, 1998.

[48] D. Chernin *et al.*, 'Exact treatment of the dispersion and beam interaction impedance of a thin tape helix surrounded by a radially stratified dielectric', *IEEE Transactions on Electron Devices,* vol. 46, pp. 1472–1483, 1999.

[49] N. Kalyanasundaram and G. N. Babu, 'Dispersion of electromagnetic waves guided by an open tape helix I', *Progress In Electromagnetics Research B,* vol. 16, pp. 311–331, 2009.

[50] N. Kalyanasundaram and G. N. Babu, 'Dispersion of electromagnetic waves guided by an open tape helix II', *Progress In Electromagnetics Research B,* vol. 19, pp. 133–150, 2010.

[51] N. Kalyanasundaram *et al.*, 'On the distribution of current on an open tape helix', *Progress In Electromagnetics Research M,* vol. 12, pp. 81–93, 2010.

[52] P. Greninger, 'Tape helix perturbation including 3-D dielectrics for TWTs', *IEEE Transactions on Electron Devices,* vol. 48, pp. 12–23, 2001.

[53] T. Wessel-Berg, 'CAD of tape helix circuits for traveling wave tubes by convolution methods', *IEEE Transactions on Electron Devices,* vol. 51, pp. 1918–1927, 2004.

[54] M. Aloisio and G. Sorbello, 'One-third-of-pitch reduction technique for the analysis of ternary azimuthally periodic helical slow-wave structures', *IEEE Transactions on Electron Devices,* vol. 53, pp. 1467–1473, 2006.

[55] C. L. Kory, 'Three-dimensional simulation of helix traveling-wave tube cold-test characteristics using MAFIA', *IEEE Transactions on Electron Devices,* vol. 43, pp. 1317–1319, 1996.

[56] K. Ward and J. Wlodarczyk, 'Transmission line modelling of helix slow wave structures', in *Technical Digest: International Electron Devices Meeting (IEDM'93)*, pp. 157–160, 1993.

[57] C. L. Kory and J. A. Dayton, Jr, 'Accurate cold-test model of helical TWT slow-wave circuits', *IEEE Transactions on Electron Devices,* vol. 45, pp. 966–971, 1998.

[58] M. Aloisio and P. Waller, 'Analysis of helical slow-wave structures for space TWTs using 3-D electromagnetic simulators', *IEEE Transactions on Electron Devices,* vol. 52, pp. 749–754, 2005.

[59] T. Onodera and W. Raub, 'Phase velocity dispersion of a generalized metal-segment-loaded helix as used in broad-band traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 35, pp. 533–538, 1988.

[60] R. G. Carter and P. Wang, 'Equivalent circuit modelling of helix slow-wave structures', presented at the ESA/NATO 1997 Workshop on Microwave Tubes for Space, Military and Commercial Applications ESTEC, Noordwijk, The Netherlands, 1997.

[61] R. G. Carter and P. Wang, 'An equivalent circuit model of helix slow-wave structures', presented at the Microwaves 94 Workshop on CAE, Modelling and Measurement Verification, London, pp. 85–90, 1994.

[62] T. K. Ghosh *et al.*, 'Optimization of coaxial couplers', *IEEE Transactions on Electron Devices,* vol. 54, pp. 1753–1759, July 2007.

[63] V. Kumar *et al.*, 'Experimental evaluation of lossy coatings of different materials on helix support rods for a high-efficient TWT', in *14th International Crimean Conference on Microwave and Telecommunication Technology (CriMico 2004)*, 2004, pp. 648–650.

[64] J. F. Gittins, *Power Travelling-Wave Tubes*. London: English Universities Press, 1965.

[65] V. B. Naidu *et al.*, 'Three-dimensional electromagnetic analysis of attenuator-coated helix support rods of a traveling-wave tube', *IEEE Transactions on Electron Devices,* vol. 56, pp. 945–950, 2009.

[66] D. M. Goebel *et al.*, 'Gain stability of traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 46, pp. 2235–2244, 1999.

[67] M. K. Alaria *et al.*, 'Design of tip loss profile on support rod for a helix TWT', *Journal of Infrared, Millimeter, and Terahertz Waves,* vol. 30, pp. 651–656, 2009.

[68] C. K. Birdsall and T. E. Everhart, 'Modified contra-wound helix circuits for high-power traveling-wave tubes', *IRE Transactions on Electron Devices,* vol. 3, pp. 190–204, 1956.

[69] J. E. Nevins, Jr., 'An investigation and application of the contrawound helix', *IRE Transactions on Electron Devices,* vol. 6, pp. 195–202, 1959.

[70] D. T. Lopes and C. C. Motta, 'Characterization of ring-bar and contrawound helix circuits for high-power traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 55, pp. 2498–2504, 2008.

[71] S. K. Datta *et al.*, 'Equivalent circuit analysis of a ring-bar slow-wave structure for high-power traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 56, pp. 3184–3190, 2009.

[72] W. N. Cain and R. W. Grow, 'The effects of dielectric and metal loading on the dispersion characteristics for contrawound helix circuits used in high-power traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 37, pp. 1566–1578, 1990.

[73] S. Sengele *et al.*, 'Backward-wave suppression analysis, and design and fabrication of a prototype millimeter-wave ring-bar slow-wave structure', *IEEE Transactions on Plasma Science,* vol. 42, pp. 3949–3960, 2014.

[74] D. C. Buck, 'Traveling-wave tube interaction circuits', *IEEE Transactions on Electron Devices,* vol. 14, pp. 721–723, 1967.

[75] S. Liu, 'The theory of ring-loop structure', *Scientia Sinica,* vol. 20, pp. 679–695, 1977.

[76] S. Liu, 'Calculation of the parameters for ring-loop traveling wave tube in MMW', *International Journal of Infrared and Millimeter Waves,* vol. 21, pp. 1097–1101, 2000.

[77] B. K. Vancil and E. G. Wintucky, 'A low cost electrostatically focused TWT', in *Third IEEE International Vacuum Electronics Conference (IVEC 2002)*, pp. 180–181, 2002.

[78] R. M. White *et al.*, 'Properties of ring-plane slow-wave circuits', *IEEE Transactions on Electron Devices,* vol. 11, pp. 247–261, 1964.

[79] G. Dohler *et al.*, 'Serpentine waveguide TWT', in *International Electron Devices Meeting*, pp. 485–488, 1987.

[80] R. G. Carter, 'An improved equivalent circuit model of folded-waveguide slow-wave structures', in *Vacuum Electronics and Displays*, Garmisch-Partenkirchen, Germany, pp. 55–59, 2001.

[81] J. H. Booske *et al.*, 'Accurate parametric modeling of folded waveguide circuits for millimeter-wave traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 52, pp. 685–694, 2005.

[82] B. G. James and P. Kolda, 'A ladder circuit coupled-cavity TWT at 80–100 GHz', in *1986 International Electron Devices Meeting*, pp. 494–497, 1986.

[83] K. J. Bunch and R. W. Grow, 'The helically wrapped circular waveguide', *IEEE Transactions on Electron Devices,* vol. 34, pp. 1873–1884, 1987.

[84] Y. Wei *et al.*, 'An approach to the analysis of arbitrarily shaped helical groove waveguides', *IEEE Microwave and Guided Wave Letters,* vol. 10, pp. 4–6, 2000.

[85] W. Wang *et al.*, 'Study of the ridge-loaded helical-groove slow-wave structure', *IEEE Transactions on Microwave Theory and Techniques,* vol. 45, pp. 1689–1695, 1997.

[86] G. Yu *et al.*, 'Analysis of the coaxial helical-groove slow-wave structure', *IEEE Transactions on Microwave Theory and Techniques,* vol. 50, pp. 191–200, 2002.

[87] R. G. Carter *et al.*, 'Computer modelling of coupled-cavity slow-wave structures', in *Vakuumelektronik und Displays*, Garmish-Partenkirchen, Germany, pp. 69–74, 1995.

[88] F. Kantrowitz and I. Tammaru, 'Three-dimensional simulations of frequency-phase measurements of arbitrary coupled-cavity RF circuits', *IEEE Transactions on Electron Devices,* vol. 35, pp. 2018–2026, 1988.

[89] C. L. Kory *et al.*, 'Simulation of cold-test dispersion and interaction impedances for coupled-cavity tube slow-wave circuits', in *International Electron Devices Meeting 1992 (IEDM '92) Technical Digest*, pp. 763–766, 1992.

[90] P. Prince *et al.*, 'A numerical method for predicting Brillouin plots of coupled cavity circuits', in *International Electron Devices Meeting*, pp. 676–679, 1984.

[91] R. G. Carter, 'Representation of coupled-cavity slow-wave structures by equivalent circuits', *IEE Proceedings I: Solid-State and Electron Devices,* vol. 130, pp. 67–72, 1983.

[92] H. J. Curnow, 'A general equivalent circuit for coupled-cavity slow-wave structures', *IEEE Transactions on Microwave Theory and Techniques*, vol. 13, pp. 671–675, 1965.

[93] K. Fujisawa, 'General treatment of klystron resonant cavities', *IRE Transactions on Microwave Theory and Techniques,* vol. MTT-6, pp. 344–358, 1958.

[94] R. G. Carter and S. K. Liu, 'Method for calculating the properties of coupled-cavity slow-wave structures from their dimensions', *IEE Proceedings-H Microwaves Antennas and Propagation,* vol. 133, pp. 330–334, October 1986.

[95] M. Chodorow and E. J. Nalos, 'The design of high-power traveling-wave tubes', *Proceedings of the IRE,* vol. 44, pp. 649–659, 1956.

[96] A. Karp, '"Comb-Quad" millimeter-wave coupled-cavity TWT interaction structure', in *International Electron Devices Meeting*, pp. 325–327, 1981.

[97] M. Chodorow and R. A. Craig, 'Some new circuits for high-power traveling-wave tubes', *Proceedings of the IRE,* vol. 45, pp. 1106–1118, 1957.

[98] P. S. Matthews *et al.*, 'The centipede structure in practical megawatt tubes', in *8th European Microwave Conference*, Paris, France, pp. 753–760, 1978.

[99] T. Roumbanis, 'Centipede Twystron ® amplifiers and traveling wave tubes for broad-band high-efficiency super-power amplification', *IEEE Transactions on Electron Devices,* vol. 16, p. 242, 1969.

[100] A. F. Pearce, 'A structure, using resonant coupling elements, suitable for a high-power travelling-wave tube', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 105, pp. 719–726, 1958.

[101] T. M. Reeder, 'An equivalent circuit for the "Centipede" waveguide', *IEEE Transactions on Microwave Theory and Techniques,* vol. 14, pp. 200–205, 1966.

[102] W. Wiejak, 'X-band coupled cavity slow wave structure of the traveling wave tube for airborne application', in *18th International Conference on Microwave Radar and Wireless Communications (MIKON)*, pp. 1–4, 2010.

[103] T. Kageyama, 'The design of the transition region in coupled-cavity TWT', in *Third IEEE International Vacuum Electronics Conference (IVEC 2002)*, pp. 102–103, 2002.

[104] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices.* Berlin: Springer-Verlag, pp. 1–84, 2008.

[105] A. W. Horsley and A. Pearson, 'Measurement of dispersion and interaction impedance characteristics of slow-wave structures by resonance methods', *IEEE Transactions on Electron Devices,* vol. 13, pp. 962–969, 1966.

[106] D. J. Connolly, 'Correction to "Determination of the interaction impedance of coupled cavity slow wave structures"', *IEEE Transactions on Electron Devices,* vol. 23, p. 1195, 1976.

[107] D. J. Connolly, 'Determination of the interaction impedance of coupled cavity slow wave structures', *IEEE Transactions on Electron Devices,* vol. 23, pp. 491–493, 1976.

[108] P. Wang *et al.*, 'An improved technique for measuring the Pierce impedance of helix slow-wave structures', in *24th European Microwave Conference,* pp. 998–1003, 1994.

[109] S. J. Rao *et al.*, 'Nonresonant perturbation measurements on dispersion and interaction impedance characteristics of helical slow-wave structures', *IEEE Transactions on Microwave Theory and Techniques,* vol. 45, pp. 1585–1593, 1997.

[110] P. Wang *et al.*, 'A simple technique for measuring the interaction impedance of helical slow-wave structures', in *Vakuumelektronik und Displays*, Garmisch-Partenkirchen, Germany, pp. 63–69, 1995.

[111] B. T. J. Maharaj and E. W. Schumann, 'Automated measurement methods to characterise travelling-wave tube slow-wave structures', *Elektron (South Africa),* vol. 6, pp. 8–10, 1989.

[112] J. R. Legarra, 'Measurement of microwave characteristics of helix traveling wave circuits', in *International Electron Devices Meeting*, pp. 408–411, 1979.

[113] R. P. Lagerstrom, 'Interaction impedance measurements by perturbation of traveling waves', Electronics Research Lab., Stanford University, Stanford, CA, 1957.

[114] G. S. Kino, 'Normal mode theory in perturbed transmission systems', Electronics Research Labs., Stanford University, Stanford, CA, 1955.

5   Thermionic Diodes

5.1   Introduction

The simplest vacuum electron device is the plane parallel diode shown in Figure 5.1(a). The diode comprises a cathode, from which electrons are emitted, and an anode within the same vacuum envelope. If the anode is positive with respect to the cathode, and there is no electron emission, the electrostatic potential varies linearly between the cathode and the anode as shown by the solid line in Figure 5.1(b). Now suppose that electrons are emitted from the cathode with negligible velocity and a current density which can be controlled in some way. The electrons are drawn towards the anode by the electric field between the electrodes, and the negative space-charge associated with the electron current depresses the potential in the space between the electrodes, as shown by the dashed line in Figure 5.1(b). If the electric field at the surface of the cathode is negative all the electrons which are supplied by the cathode can flow to the anode. The current through the diode is then determined by the properties of the cathode. When the potential difference between the electrodes is reversed no electrons can reach the anode. Thus the device is a diode which only permits the flow of electrons from the cathode to the anode. It should be noted that the conventional current of electric circuit theory assumes that the charge carriers are positive so that the conventional current flows from the anode to the cathode.

As the emitted current density is increased the potential depression increases until the electric field is zero at the cathode surface. The current through the diode is then determined by the geometry of the device and the potential difference across it, and not by the current density available at the cathode. The diode is then said to be *space-charge limited* and the current cannot be increased any further by increasing the cathode emission.

This chapter discusses the steady-state properties of planar, cylindrical and spherical space-charge limited diodes, including the effects of the initial energies of the electrons and of relativity. When the potential difference across a diode varies with time the static relationship continues to hold until the time taken for electrons to cross the diode becomes comparable with the time in which the voltage is changing. It is shown that this sets a limit to the maximum frequency of operation of a diode. Section 5.8 discusses the injection of electrons into a planar diode and Section 5.9 reviews the properties of diodes in which the current flow is two-dimensional.
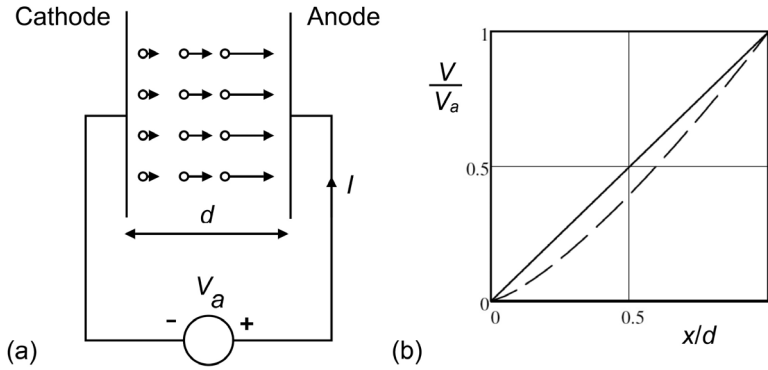
Figure 5.1: Planar vacuum diode showing: (a) general arrangement, and (b) the electrostatic potential within the diode showing the potential depression caused by space-charge (dashed line).

An important form of diode is the electron gun in which the shapes of the electrodes are chosen in such a way that the electrons do not strike the anode but pass through a hole in it and are projected into the region beyond. The electron gun is a fundamental component of many microwave tubes and the theory and design of the principal types are reviewed in Chapter 9.

5.1.1 Dimensional Analysis of Thermionic Diodes

Useful insights into the properties of thermionic diodes are provided by the use of dimensional analysis (see Section 1.7.1.). The objective is to discover how the current flowing through a diode depends upon the applied voltage and the dimensions of the diode. The shape of the diode may be quite complicated but, for the present, we will assume that it can be described by the distance between the electrodes and by the area of the cathode. Since this is a problem concerning the dynamics of charged particles we expect that it will involve the charge/mass ratio of the electron and the primary electric constant ε_0. When thermal velocities or relativistic effects are important it is necessary to introduce a reference voltage V_r related to their magnitude. The dimensional quantities required to specify the problem are summarised in Table 5.1 together with their dimensions.

The problem has seven parameters and four dimensions. The application of Buckingham's theorem shows that it can be specified by three dimensionless groups. Two dimensionless groups can readily be formed as

$$\Pi_1 = \frac{V_a}{V_r} \tag{5.1}$$

and

$$\Pi_2 = \frac{A}{d^2}. \tag{5.2}$$

Table 5.1: Parameters of a thermionic diode

Quantity	Symbol	Dimensions
Current	I	$[ML^2T^{-3}V^{-1}]$
Anode voltage	V_a	$[V]$
Reference voltage	V_r	$[V]$
Electrode separation	d	$[L]$
Area	A	$[L^2]$
Primary electric constant	ε_0	$[MLT^{-2}V^{-2}]$
Charge/mass ratio of the electron	e/m_0	$[L^2T^{-2}V^{-1}]$

The remaining group is found by assuming that

$$I = V_r^\alpha L^\beta \varepsilon_0^\gamma \left(e/m\right)^\delta \tag{5.3}$$

where the indices are to be determined. Then, examining the powers of the dimensions, we find that

$$
\begin{aligned}
M: &\quad 1 = \gamma \\
L: &\quad 2 = \beta + \gamma + 2\delta \\
T: &\quad -3 = -2\gamma - 2\delta \\
V: &\quad -1 = \alpha - 2\gamma - \delta.
\end{aligned}
\tag{5.4}
$$

The solution to these equations is $\alpha = 1$; $\beta = 0$; $\gamma = 3/2$; $\delta = \frac{1}{2}$ and therefore the third dimensionless group is

$$\Pi_3 = \left(\frac{I}{\varepsilon_0 \sqrt{(e/m)V_r^3}}\right). \tag{5.5}$$

Any one of the dimensionless groups can then be expressed as a function of the remaining groups, for example

$$\Pi_3 = f\left(\Pi_1, \Pi_2\right), \tag{5.6}$$

where f is a function, to be determined, which is the same for all diodes which are geometrically similar, that is they can be scaled from one another.

   If neither thermal velocities nor relativistic effects are important then Π_1 is not required. Then V_r can be replaced by V_a in (5.5) so that (5.6) becomes

$$\Pi_3 = f\left(\Pi_2\right). \tag{5.7}$$

For any set of geometrically similar diodes Π_2 is constant so that Π_3 is also constant. Equation (5.5) shows that the current in the diode is proportional to the 3/2 power of the applied voltage. Thus dimensional analysis shows that all diodes which do not require a reference voltage must obey the 3/2 power law regardless of

geometry. The same result was obtained in different ways by Langmuir [1] and by Tsimring [2].

Because many electron tubes obey the 3/2 power law it is convenient to describe a diode by its perveance (K) defined by

$$K = \frac{I}{V_a^{3/2}},\tag{5.8}$$

which depends only on the geometry of the diode. It follows that all diodes whose dimensions can be scaled from one another have the same perveance. Because the perveance is typically of the order of 10^{-6} A V$^{-1.5}$ it is usual to use micro-perveance $\left(\mu\mathrm{P} = \mu\mathrm{A\ V}^{-1.5}\right)$. We shall see below that when thermal velocities, or relativistic effects, are important the description of the problem requires a reference voltage and the current does not obey the 3/2 power law.

5.1.2   Current Limitation

Virtually all power vacuum tubes use thermionic cathodes whose emitted current increases with temperature (see Section 18.5). Figure 5.2 shows a typical *I-V* plot for a thermionic diode. At low voltages the diode is space-charge limited and the curve follows the 3/2 power law. As the current increases, it becomes equal to the maximum current available from the cathode, as described by the Schottky equation (18.10). This current depends on the temperature and work function of the cathode, and only increases slowly with increasing anode voltage. The diode is then said to be *temperature limited*.

Vaughan [3] suggested an empirical model for the transition from space-charge limitation to temperature limitation of the current:

$$\frac{1}{J^\alpha} = \frac{1}{J_C^\alpha} + \frac{1}{J_S^\alpha},\tag{5.9}$$

where J_C is the space-charge limited current density given by the Child-Langmuir equation (5.16), J_S is the current density given by the Schottky equation (18.10), and α is an empirical constant. It was found that the value of α was typically in the range 6–10 for well-designed diodes (electron guns). However, for diodes with uneven emission, caused by surface defects or uneven heating, the value could be in the range 2–5.

The diodes used in power vacuum tubes are usually designed to be operated in the space-charge limited region because it ensures that:

- The current depends only on the cathode voltage, and the geometry of the diode, and not on the condition of the cathode.
- The cathode surface is not damaged by the impact of positive ions caused by the presence of a strong electric field there.

If it is suspected that the cathode emission is failing, an *I-V* plot can be drawn to check that the operating point is below the knee of the curve. There is a trade-off, for
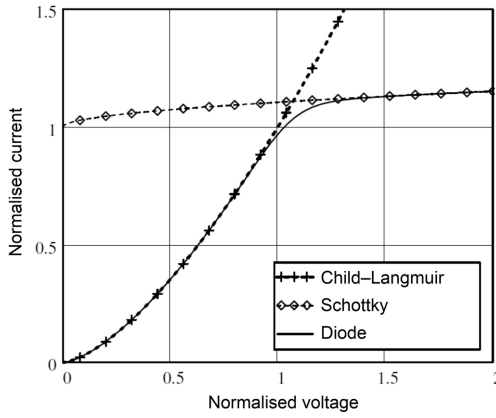
Figure 5.2: Plot of current against voltage for a thermionic diode showing the Child-Langmuir (space-charge limited) and Schottky (temperature limited) regions.

a particular cathode, between the current density and the life of the cathode. Thus tubes which are required to have very long working lives (such as those for use in satellites) are designed to have low cathode current density. However, temperature-limited diodes are used in gyrotrons so that the electron current can be controlled by the cathode temperature.

5.2   The Planar Space-Charge Limited Diode

Consider a diode comprising two plane parallel electrodes separated by distance d as shown in Figure 5.1. The electrostatic potential (V) is zero at the cathode at $x = 0$ and $V = V_a$ at the anode at $x = d$. For the moment we will assume that electrons are emitted with zero energy from the cathode and that relativistic effects can be neglected. The theory of this diode was first studied by Child [4] and Langmuir [1]. The velocity of an electron at any point is found from the principle of conservation of energy to be

$$u = \sqrt{2(e/m_0)V},$$
(5.10)

where V is the potential with respect to the cathode. In the steady state the current density is constant and given by

$$J = \rho u,$$
(5.11)

where ρ is the charge density. The potential is related to the space-charge density by Poisson's equation in one dimension,

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon_0}.$$
(5.12)

Combining equations (5.10) to (5.12) gives

$$\frac{d^2V}{dx^2} = -\frac{J}{\varepsilon_0\sqrt{2(e/m_0)V}}.$$ (5.13)

Equation (5.13) can be integrated by multiplying both sides by $2(dV/dx)$, and integrating with respect to x to give

$$\left(\frac{dV}{dx}\right)^2 = -\frac{4J}{\varepsilon_0\sqrt{2(e/m_0)}}\sqrt{V} + C,$$ (5.14)

where C is a constant. Applying the boundary conditions to (5.14) gives

$$E_a^2 - E_k^2 = -\frac{4J}{\varepsilon_0\sqrt{2(e/m_0)}}\cdot\sqrt{V_a},$$ (5.15)

where V_a is the potential of the anode with respect to the cathode and E_a and E_k are, respectively, the electric field at the anode and the cathode. The current density is maximum when the left-hand side of (5.15) is maximum. From Figure 5.1(b) we see that, as the space-charge density increases E_a increases and E_k decreases. Thus the maximum possible current density is obtained when $E_k = 0$ and this is the condition for space-charge limitation of the current. When the current is space-charge limited $C = 0$ and (5.14) may be integrated a second time to give

$$J_C = -\frac{4}{9}\cdot\frac{\varepsilon_0\sqrt{2(e/m_0)}}{x^2}\cdot V^{\frac{3}{2}} = -K\frac{V^{\frac{3}{2}}}{x^2},$$ (5.16)

the well-known Child–Langmuir law where $K = 2.334 \times 10^{-6}$ A V$^{-1.5}$. Note that the current density in (5.16) is negative because the conventional current flows from the anode to the cathode in the negative x direction. A simple physical derivation of this equation which gives insight into the reason for the power law was given by Umstattd et al. [5].

Equation (5.16) can be rearranged, using dimensionless groups, to be in the form (5.6)

$$\left(\frac{I}{\varepsilon_0\sqrt{2(e/m)V_r^3}}\right) = -\frac{4}{9}\left(\frac{V_a}{V_r}\right)^{\frac{3}{2}}\left(\frac{A}{d^2}\right),$$ (5.17)

where $I = AJ_C$, A is the cross-sectional area of the diode, and the factor $\sqrt{2}$ has been included in the term on the left-hand side of the equation. In this case it is convenient to choose a reference voltage of 1.0 V.

The transit time of electrons in a space-charge-limited diode may be calculated by noting that the potential in the diode is given by

$$V = V_a\left(\frac{x}{d}\right)^{\frac{4}{3}}$$ (5.18)

so that the velocity of an electron which starts from rest at the cathode is

$$u = \sqrt{\frac{2eV_a}{m_0}} \left(\frac{x}{d}\right)^{\frac{2}{3}}. \tag{5.19}$$

The time taken for this electron to reach the anode is

$$\tau = \frac{1}{u_a} \int_0^d \left(\frac{d}{x}\right)^{\frac{2}{3}} dx = \frac{3d}{u_a}, \tag{5.20}$$

where $u_a = \sqrt{2eV_a/m_0}$ is the velocity of the electron when it reaches the anode. The transit time calculated in (5.20) can be compared with the result for the space-charge free case where the acceleration is uniform and the transit time is $\tau_0 = 2d/u_a$. Equation (5.20) can also be written as

$$\tau = \frac{K^{\frac{1}{3}}}{3\sqrt{2e/m_0}} \left(\frac{d}{J_C}\right)^{\frac{1}{3}}. \tag{5.21}$$

5.3   The Planar Diode Including the Effects of Thermal Velocities

The analysis in the previous section is non-physical, strictly speaking, because no electrons can leave the cathode when the electric field there is zero, and because the solution implies that the charge density at the cathode surface is infinite. In practice the electrons have thermal velocities as they leave the cathode so that they can move away from it. Two approaches to this problem were discussed by Langmuir [6–7].

The first method assumes that all the electrons have the same initial velocity when they leave the cathode. This enables them to move towards the anode in the presence of an opposing electric field. There must, therefore, be a plane between the cathode and the anode at which the gradient of the potential is zero as shown in Figure 5.3. At this plane the electrons come to rest, a virtual cathode exists, and the charge density is infinite. Thus this model suffers from the same objections as before. It is assumed that the current density at the cathode is equal to the saturation (Schottky) current density at a given temperature. Then some of the electrons must be reflected by the virtual cathode, and some pass it to reach the anode. This analysis can be criticised on the ground that there is no reason why electrons, which have the same initial velocities, should behave in different ways in the steady state. There is also the difficulty of knowing how to define the initial velocity in terms of the cathode temperature [6]. This approach is therefore not valid when the initial velocities are thermal. It is, however, the basis of the discussion of space-charge limitation of the current when a mono-energetic beam of electrons is injected into a diode as discussed in Section 5.8.

The second method assumes that the electrons emitted from the cathode have a Maxwell–Boltzmann energy spectrum. Consider a source of electrons at zero
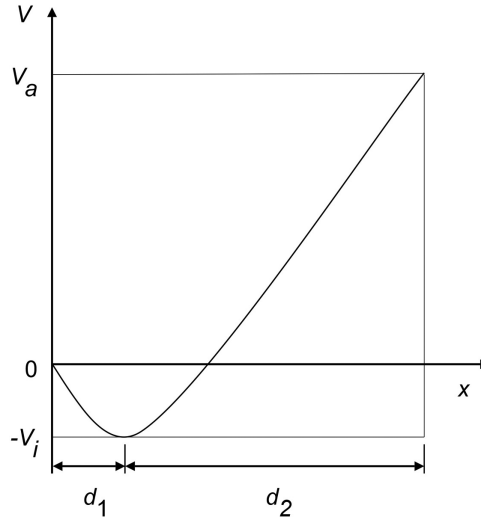
Figure 5.3: Potential variation in a planar space-charge-limited diode with non-zero emission velocities.

potential with temperature T which has saturation current density J_S. The current density collected on an electrode at a potential $-V$ with respect to the source is

$$J = J_S \exp\left(-\frac{V}{V_T}\right), \tag{5.22}$$

where $V_T = kT/e$ is the volt equivalent of temperature [6]. The current density emitted from the cathode in the energy range $(V, V + dV)$ is

$$dJ = \frac{J_S}{V_T} \exp\left(-\frac{V}{V_T}\right) \cdot dV. \tag{5.23}$$

When the current is space-charge-limited there is a virtual cathode between the cathode and the anode with potential is $-V_m$ relative to the cathode. The current collected by the anode (J_a) comprises those electrons which have sufficient energy to pass the virtual cathode. Thus

$$J_a = J_S \exp\left(-\frac{V_m}{V_T}\right). \tag{5.24}$$

The electron flow in the regions on either side of the virtual cathode must be considered separately.

5.3.1   Electron Flow between the Potential Minimum and the Anode

Consider electrons launched with initial energy V_i at the potential minimum. At a plane where the potential relative to the minimum is V they have velocity

$$u_i = \sqrt{2(e/m_0)(V + V_i)}. \tag{5.25}$$

From (5.23) the current density of these electrons is

$$dJ_i = \frac{J_a}{V_T} \exp\left(-\frac{V_i}{V_T}\right) \cdot dV_i \qquad (5.26)$$

so that the contribution to the charge density at this plane is, from (5.11)

$$d\rho_i = \frac{dJ_i}{v_i} = \frac{J_a}{V_T \sqrt{2(e/m_0)(V+V_i)}} \exp\left(-\frac{V_i}{V_T}\right) \cdot dV_i. \qquad (5.27)$$

Integrating this expression over all possible values of V_i the charge density at a plane where the potential relative to the minimum is V is given by

$$\rho(V) = \frac{J_a}{V_T} \int_0^\infty \frac{1}{\sqrt{2(e/m_0)(V+V_i)}} \cdot \exp\left(-\frac{V_i}{V_T}\right) \cdot dV_i. \qquad (5.28)$$

Equation (5.28) can be written

$$\rho(v) = \frac{J_a}{\sqrt{2(e/m_0)V_T}} \int_0^\infty \frac{1}{\sqrt{v+v_i}} \cdot \exp(-v_i) \cdot dv_i, \qquad (5.29)$$

where $v = V/V_T$ and $v_i = V_i/V_T$. Substituting for ρ in Poisson's equation (5.12) gives

$$\frac{d^2 v}{dX^2} = \int_0^\infty \frac{1}{\sqrt{v+v_i}} \cdot \exp(-v_i) \cdot dv_i, \qquad (5.30)$$

where the normalised position X is defined by

$$X = \sqrt{\frac{-J_a}{\varepsilon_0 \sqrt{2(e/m_0)V_T^3}}} \cdot x. \qquad (5.31)$$

For convenience let

$$y^2 = v + v_i. \qquad (5.32)$$

Then (5.30) may be written

$$\frac{d v^2}{dX^2} = 2 \cdot \exp(v) \cdot \int_{\sqrt{v}}^\infty \exp(-y^2) \cdot dy. \qquad (5.33)$$

Now the probability function, defined as

$$P(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) \cdot dt \qquad (5.34)$$

tends to unity as $x \to \infty$ so that (5.33) may be written

$$\frac{d^2 v}{dX^2} = \sqrt{\pi} \cdot \left(1 - P(\sqrt{v})\right) \cdot \exp(v). \qquad (5.35)$$

Equation (5.35) can be integrated by multiplying both sides by $2(dU/dX)$ to give

$$\left(\frac{d\upsilon}{dX}\right)^2 = 2\sqrt{\pi}\int\left[1 - P\left(\sqrt{\upsilon}\right)\right]\cdot\exp(\upsilon)\cdot d\upsilon. \tag{5.36}$$

Integrating by parts and noting that $(d\upsilon/dX) = 0$ when $\upsilon = 0$ we obtain

$$\left(\frac{d\upsilon}{dX}\right)^2 = 2\sqrt{\pi}\cdot\left[\exp(\upsilon) - 1 - \exp(\upsilon)P\left(\sqrt{\upsilon}\right) + \frac{2}{\sqrt{\pi}}\sqrt{\upsilon}\right]. \tag{5.37}$$

This expression agrees with the term in brackets in equation (11) in [6], taking the lower signs. The constant on the right-hand side agrees with that in equation (6) in [6] when the difference in the units is allowed for by setting $\varepsilon_0 = 1/4\pi$ (see equation (4) in [6]). Langmuir's variables are related to those used here by

$$\xi = \sqrt{2\sqrt{\pi}}\cdot X \tag{5.38}$$

and

$$\eta = \upsilon. \tag{5.39}$$

5.3.2 Electron Flow between the Cathode and the Potential Minimum

In the region between the cathode and the potential minimum we must consider the flow in two parts. Those electrons with energies greater than or equal to V_m are able to pass the minimum and reach the anode with a total current density J_a. The charge density due to these electrons at a plane where the potential is V relative to the minimum is, from (5.29)

$$\rho_1(\upsilon) = \frac{J_a}{\sqrt{2(e/m_0)V_T}}\int_0^\infty \frac{1}{\sqrt{\upsilon + \upsilon_i}}\cdot\exp(-\upsilon_i)\cdot d\upsilon_i, \tag{5.40}$$

where $\upsilon = V/V_T$.

Those electrons which start with energies less than V_m are reflected by the field at the point where their initial energy is equal to the potential. The potential at any point relative to the cathode is $(V - V_m)$ so that the velocity of an electron whose energy at the cathode is V_i is

$$u_i = \sqrt{2(e/m_0)(V_i + V - V_m)}. \tag{5.41}$$

Thus at the virtual cathode, where $V = 0$, an electron whose initial energy is $V_i = V_m$ is just brought to rest. Because electrons whose initial energies are less than V_m are reflected, they contribute twice to the charge density. The contribution to the charge density at a plane is therefore

$$d\rho_i = \frac{2dJ_i}{\upsilon_i} = \frac{2J_S}{V_T\sqrt{2(e/m_0)(V_i + V - V_m)}}\exp\left(-\frac{V_i}{V_T}\right)\cdot dV_i. \tag{5.42}$$

At any plane this must be integrated to cover all electrons whose initial energy allows them to reach that plane, but does not permit them to pass the minimum. Thus

$$\rho_2(V) = \frac{2J_S}{V_T} \int_{V_m - V}^{V_m} \frac{1}{\sqrt{2(e/m_0)(V_i + V - V_m)}} \exp\left(-\frac{V_i}{V_T}\right) \cdot dV_i. \tag{5.43}$$

Substituting for J_S from (5.24)

$$\rho_2(V) = \frac{2J_a}{V_T} \int_{-V}^{0} \frac{1}{\sqrt{2(e/m_0)(V + V_i')}} \exp\left(-\frac{V_i'}{V_T}\right) \cdot dV_i', \tag{5.44}$$

where $V_i' = V_i - V_m$. Then (5.44) can be written

$$\rho_2(\upsilon) = \frac{2J_a}{\sqrt{2(e/m_0)V_T}} \int_{-\upsilon}^{0} \frac{1}{\sqrt{\upsilon + \upsilon_i}} \cdot \exp(-\upsilon_i) \cdot d\upsilon_i, \tag{5.45}$$

where the prime on the variable of integration can be omitted without loss of generality. The total charge density at a plane where the normalised potential with respect to the minimum is υ is then the sum of ρ_1 and ρ_2. Substituting for the charge density from (5.40) and (5.45) in Poisson's equation (5.13) and making use of (5.31) we obtain

$$\frac{d^2\upsilon}{dX^2} = \int_0^\infty \frac{1}{\sqrt{\upsilon + \upsilon_i}} \cdot \exp(-\upsilon_i) \cdot d\upsilon_i + 2 \cdot \int_{-\upsilon}^{0} \frac{1}{\sqrt{\upsilon + \upsilon_i}} \cdot \exp(-\upsilon_i) \cdot d\upsilon_i. \tag{5.46}$$

Equation (5.46) can be expressed in terms of the probability function as

$$\frac{d^2\upsilon}{dX^2} = \sqrt{\pi} \cdot \exp(\upsilon)\left[1 + P\left(\sqrt{\upsilon}\right)\right]. \tag{5.47}$$

As before this may be integrated to give

$$\left(\frac{d\upsilon}{dX}\right)^2 = 2\sqrt{\pi} \cdot \left[\exp(\upsilon) - 1 + \exp(\upsilon)P\left(\sqrt{\upsilon}\right) - \frac{2}{\sqrt{\pi}}\sqrt{\upsilon}\right] \tag{5.48}$$

that agrees with (11) in [6].

The formal solution to (5.37) and (5.48) is

$$X = (4\pi)^{-\frac{1}{4}} \cdot \int_0^\upsilon \frac{d\upsilon}{\left[\exp(\upsilon) - 1 \pm \exp(\upsilon)P\left(\sqrt{\upsilon}\right) \mp 2\sqrt{\upsilon/\pi}\right]^{\frac{1}{2}}}, \tag{5.49}$$

which is, essentially, (11) in [6]. In this equation the upper signs are taken between the cathode and the minimum, and the lower signs between the minimum and the anode. The distance between the cathode and the minimum is obtained by using the

upper signs and setting $V = V_m$. The distance between the minimum and the anode is obtained by using the lower signs and setting $V = V_m + V_{ak}$. For comparison equation (5.16) can be rewritten in the form

$$\frac{-J_C\, x^2}{\varepsilon_0 \sqrt{2(e/m_0)}\, V_T^{\frac{3}{2}}} = \frac{4}{9} \cdot \upsilon^{\frac{3}{2}}$$

(5.50)

or

$$X = \frac{2}{3} \cdot \upsilon^{\frac{3}{4}}.$$

(5.51)

Figure 5.4 shows the solutions of (5.49) and (5.51). The curves are identical to those in Figure 42 in [7] apart from the change in the scale of the horizontal axis. In this figure the cathode is to the left of the origin and the anode to the right.

5.3.3   Numerical Evaluation

It is not possible to evaluate the integral in (5.49) by analytical methods. Tables of the solutions found using numerical methods are given in [6]. The presence of the exponential functions can cause numerical overflows for large values of U. In the region between the potential minimum and the anode (5.49) can be approximated without loss of accuracy, when $U > 50$, by

$$X = (4\pi)^{-\frac{1}{4}} \cdot \int_0^{\upsilon} \frac{d\upsilon}{\left[2\sqrt{\upsilon/\pi} - 1\right]^{\frac{1}{2}}}.$$

(5.52)
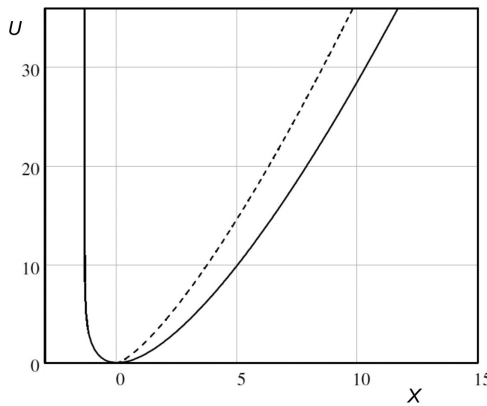
Figure 5.4: Dependence of normalised potential against normalised position in a space-charge-limited diode with thermal initial velocities showing the Child-Langmuir curve (dashed) for comparison
(copyright 1931, The American Physical Society, reproduced, with permission, from [7]).

Between the cathode and the minimum $X \to 1.3565$ when $\upsilon > 20$. The solutions shown in Figure 5.4 were obtained in this way (see Worksheet 5.1). Alternatively, equations (5.30) and (5.46) can be integrated directly using the Runge-Kutta method. Both methods give identical results.

To calculate the current-voltage characteristic of a thermionic diode we note first that the potential at the minimum is given by (5.22) as

$$V_m = V_T \, \ln\left(\frac{J_a}{J_S}\right). \tag{5.53}$$

For the purposes of illustration we will assume that the cathode temperature is 1300 °K, which is typical of the temperature of the cathodes used in microwave tubes, so that $V_T = 0.112$ V. When the ratio of the anode current to the saturation current is 10^{-6} the potential at the minimum is about 1.5 V which is small compared with typical potential differences between the electrodes. The distance of the minimum from the cathode can be computed using (5.49) for typical values of the parameters. The results are shown in Figure 5.5 where the distance between the minimum and the cathode is plotted against the normalised anode current. It can be seen that the distance is much less than 1 mm, except when the saturation current is small and when the anode current is a small fraction of the saturation current. Hence we expect that the effects of thermal velocities will be negligible except for small diode spacings, and low potential differences. Thus in Figure 5.2 d_1 is normally much smaller than d_2.

Once d_1 has been found then d_2 is given by

$$d_2 = d - d_1 \tag{5.54}$$

and (5.49) can be solved for the potential at the anode relative to the minimum as a function of the anode current. Finally the potential difference between the anode and the cathode is found by subtracting the potential at the minimum relative to the cathode.
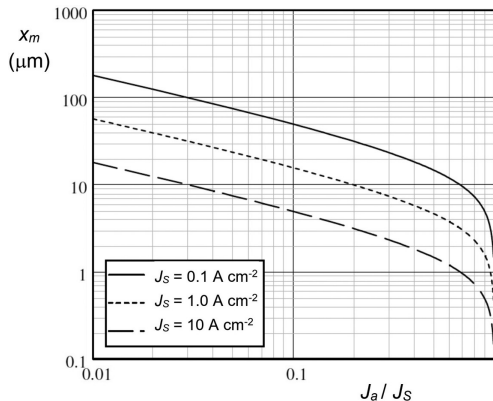
Figure 5.5: Distance between the potential minimum and the cathode, in a space-charge-limited diode with thermal velocities, plotted against normalised current density.

As the anode current drops, the potential difference between the minimum and the cathode increases and X tends to a constant, as has been noted. But, from (5.31), we see that $x \propto \sqrt{-J_a}$ and therefore the distance between the cathode and the minimum increases without limit. Thus, in any diode, the minimum will reach the anode if the anode current is small enough. When the minimum just reaches the anode the potential of the anode is negative with respect to the cathode and equal to the potential at the minimum. For smaller anode currents d_2 is negative and the relationship between the anode current and the anode-cathode voltage is obtained from (5.22). The results of these calculations for a range of parameter values are shown in Figure 5.6. The normalised current (J_a/J_S) is plotted against the anode-cathode voltage, normalised to the voltage at the saturation current for a diode without thermal velocities, given by (5.16). The graphs are plotted with the saturation current as a parameter for four different diode spacings. In each case the curve given by (5.16) is included for comparison. In microwave tubes the saturation current is typically in the range 1.0 to 10 A cm^{-2} (see Chapter 18). It is clear
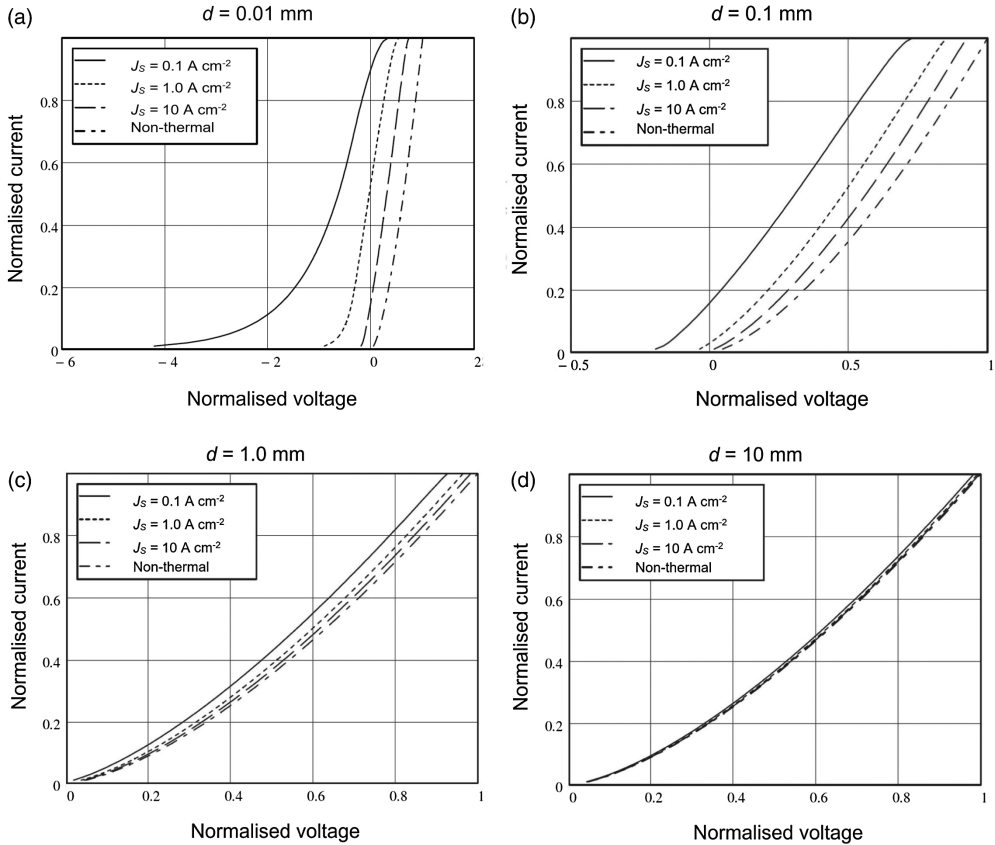
Figure 5.6: Normalised current in a planar space-charge limited diode with thermal velocities plotted against the anode voltage, normalised to the voltage at the saturation current in a diode without thermal velocities.

from Figure 5.6 that there is very little error in using (5.16) when the diode spacing is greater than 1.0 mm. For example, when $J_S = 1.0 \, \text{A cm}^2$ and $d = 1 \, \text{mm}$ the maximum anode voltage is 264 V so thermal effects can be neglected if the anode voltage is much greater than this. In general, thermal effects are not important in many power vacuum tubes. The exceptions are gridded tubes, and gridded electron guns, where the spacing between the control grid and the cathode may be much less than 1 mm, and also vacuum micro-electronic devices.

5.4   The Planar Diode Including the Effects of Relativity

In many electron tubes the potential difference between the anode and the cathode is high enough for relativistic effects to be important. In that case (5.10) is replaced by (1.4)

$$u = c \left[1 - \frac{1}{\left(1 + V/V_R\right)^2} \right]^{\frac{1}{2}}, \tag{5.55}$$

where $V_R = \left(m_0 c^2 / e\right) = 511 \text{kV}$ is the rest energy of an electron in electron volts [8–10]. Following the same method as before we find that, for space-charge limited flow, (5.14) is replaced by

$$\left(\frac{dV}{dx}\right)^2 = -\frac{2J}{\varepsilon_0 c} \cdot \sqrt{2V_R V + V^2}, \tag{5.56}$$

where the constant C is zero. It is convenient to introduce the normalised variables

$$U = \frac{V}{V_R} \tag{5.57}$$

and

$$X = \sqrt{\frac{-J}{2\varepsilon_0 c V_R}} \cdot x. \tag{5.58}$$

Note that these are not the same as the normalised variables used in the previous section. Equation (5.56) then becomes

$$\left(\frac{dU}{dX}\right)^2 = 4\sqrt{2U + U^2}, \tag{5.59}$$

whose solution can be written formally as

$$X = \frac{1}{2} \int_0^U \left(2U + U^2\right)^{-\frac{1}{4}} \cdot dU. \tag{5.60}$$

This equation cannot be integrated analytically. Acton [8] and Boers and Kelleher [9] derived solutions as power series. An alternative solution expressed in terms of elliptic integrals given in [10] was used by Zhang et al. [11] to derive an approximate equation for the space-charge limited current density in a relativistic planar diode

$$J = \frac{2\varepsilon_0 m_0 c^2}{ed^2} \cdot \frac{\left(\gamma_0^{2/3} - 1\right)^{3/2}}{\left(\sqrt{3} - 1\right)\gamma_0^{-0.392} + 1}, \tag{5.61}$$

where $\gamma_0 = \left(1 + e V_a / m_0 c^2\right)$. The error in this equation is less than 1% for any voltage.

The series given in [9] can be written

$$X = \left(U^2 + 2U\right)^{-\frac{1}{4}} \left(\frac{2}{3}U + \frac{1}{21}U^2 + \frac{1 \cdot 3}{21 \cdot 11}U^3 + \cdots (-1)^{n+1}\frac{1 \cdot 3 \cdot 5 \cdots (2n-5)(2n-3)}{21 \cdot 11 \cdot 15 \ldots (4n-5)(4n-1)}U^n\right) \tag{5.62}$$

Alternatively (5.60) can be evaluated directly using numerical methods. When voltage is small the terms in U^2 and higher powers in (5.60) can be neglected, and the integral evaluated analytically, to give

$$X = \frac{1}{3}(2U)^{\frac{3}{4}}. \tag{5.63}$$

It is easy to show that this equation is identical to (5.16). Figure 5.7 shows a comparison between curves calculated using (5.63), the numerical evaluation of (5.60), and the first three terms of (5.62). It can be seen that, for a given U, the effect of relativity is to reduce the value of X and, therefore, the current for a given diode spacing. The series solution is a good approximation to the exact result with an error of less than 1% in X when $U < 1.3$ if three terms are taken (see Worksheet 5.1).

For a fixed diode spacing the current is proportional to X^2. The ratio of the relativistic current computed from (5.60) to the non-relativistic current computed from
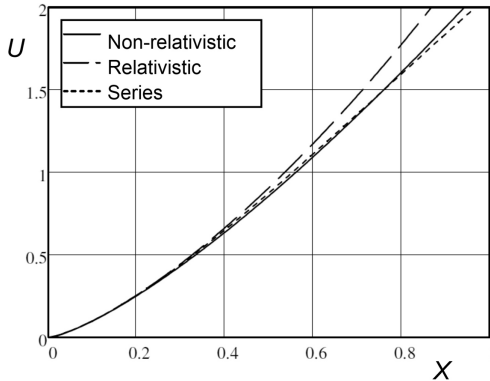
Figure 5.7: Normalised potential in a planar space-charge-limited diode plotted against normalised position both with, and without, relativistic effects and using the series approximation in (5.62).
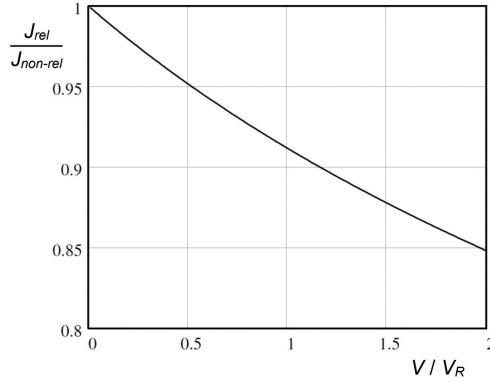
Figure 5.8: Relativistic correction to the current in a planar space-charge-limited diode plotted against normalised anode voltage.

(5.63) is shown in Figure 5.8. It can be seen that equation (5.63) over-estimates the current by around 9% when $U = 1$.

5.5　The Cylindrical Space-Charge Limited Diode

The problem of space-charge limited flow between coaxial cylinders when the initial electron velocity is zero is discussed in [12] for the non-relativistic case, and in [8] for the relativistic case. For symmetrical flow in a cylindrical diode Poisson's equation is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) = -\frac{\rho}{\varepsilon_0}. \tag{5.64}$$

The total current per unit length I_L is independent of radius so the current density is

$$J = \frac{I_L}{2\pi r}. \tag{5.65}$$

Then, from equations (5.11), (5.55) and (5.64)

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) = -\frac{I_L}{2\pi\varepsilon_0 cr}\left[1-\frac{1}{\left(1+V/V_R\right)^2}\right]^{-\frac{1}{2}}. \tag{5.66}$$

Let $R = r/r_c$, where r_c is the radius of the cathode, and $U = V/V_R$ so that (5.66) can be written

$$\frac{d^2U}{dR^2} + \frac{1}{R}\frac{dU}{dR} = \frac{k_1}{R}\left(\frac{U+1}{\sqrt{U^2+2U}}\right), \tag{5.67}$$

where

$$k_1 = \frac{-I_L r_c}{2\pi\varepsilon_0 c V_R}.$$ (5.68)

Equation (5.67) can be integrated by numerical methods with the initial conditions $U = 0$ and $dU/dR = 0$ when $R = 1$.

A power series solution for the non-relativistic case can be derived by setting

$$\gamma = \ln R$$ (5.69)

and assuming that the current obeys the 3/2 power law so that (in SI units)

$$I_L = \frac{8\pi\varepsilon_0 \sqrt{2(e/m_0)}}{9} \cdot \frac{V^{\frac{3}{2}}}{(r\beta^2)}.$$ (5.70)

Where V is the potential at radius r relative to the cathode, and β is a function of R [12]. Two series expansions for β have been found:

$$\beta = \sum_{n=0}^{\infty} A_n \gamma^n$$ (5.71)

and

$$\beta = \exp\left(-\frac{\gamma}{2}\right) \cdot \sum_{n=0}^{\infty} B_n \gamma^n,$$ (5.72)

where the coefficients are given in Table 5.2. The second series generally converges faster than the first except when $R < 0.05$ when γ is negative and the exponential function in (5.72) becomes very large.

Acton [8] obtained the following power series solution for the non-relativistic case:

$$U = a\gamma^{\frac{4}{3}} (1 + 0.1333\gamma + 0.02444\gamma^2 + 0.0039236\gamma^3 + 0.00052966\gamma^4 + \cdots),$$ (5.73)

Table 5.2: Coefficients for equations (5.71) and (5.72)

n	A_n	B_n
0	0.0	0.0
1	+1.0	+1.0
2	−0.40	+0.10
3	+0.916667	+0.016667
4	−0.014242	+0.002424
5	+0.001679	+0.000266
6	−0.000161	+0.000026

where

$$a^3 = \frac{81}{32} \left(\frac{I_L r_0}{2\pi\varepsilon_0 c V_R} \right)^2 = \frac{81}{32} k_1^2. \tag{5.74}$$

Equation (5.70) can be written in the same notation as

$$U = a \left(R^2 \beta^4 \right)^{\frac{1}{3}}. \tag{5.75}$$

It can be shown that (5.73) and (5.75) give identical results. Both equations are written in a dimensionless form similar to (5.6). We note that, when R is constant, U is proportional to a and, therefore, the current is proportional to the 3/2 power of the voltage. Langmuir showed that the solution in series was identical to that obtained by direct integration of (5.67).

For the relativistic case Acton [8] obtained the power series solution

$$U = a\gamma^{\frac{4}{3}} \left(1 + \frac{2}{15}\gamma + \frac{11}{450}\gamma^2 + \ldots \right) + a^2 \gamma^{\frac{8}{3}} \left(\frac{1}{14} + \frac{6}{175}\gamma + \ldots \right), \tag{5.76}$$

which is valid when $U < 2$. The first series on the right-hand side of (5.76) is identical to the series in (5.73) so it follows that the effects of relativity are represented by the second series on the right-hand side. It would appear from (5.76) that, if the difference between the relativistic and non-relativistic values of U is divided by a^2, the result should be a function of R only. Calculations show that this is only true for $a < 0.1$ and we conclude that (5.76) is an approximation to a double power series expansion of U in terms of a and γ. Zhang et al. [11] derived an approximate expression for this case in the same manner as that for the planar diode

$$J_c = \frac{2\varepsilon_0 m_0 c^2}{2(r_c - r_a)^2} \cdot \frac{\left(\gamma_0^{2/3} - 1 \right)^{3/2}}{\left(\sqrt{3} - 1 \right) \gamma_0^{-0.392} + 1}, \tag{5.77}$$

where $\gamma_0 = \left(1 + eV_a / m_0 c^2 \right)$, J_c is the current density on the surface of the cathode and r_c and r_a are the radii of the cathode and the anode respectively. The relativistic term is identical to that in (5.61) and the error is less than 5% for all voltages and less than 1% for mildly relativistic voltages. The accuracy is best when the ratio of the anode radius to the cathode radius is close to unity.

When the results from (5.76) are compared with those obtained by direct integration it is found that the agreement is good when $a \le 1$ and $R > 1$ (see Worksheet 5.2). When $R < 1$ the agreement is not so good. For example, when $R = 0.5$ and $a = 1$ the value of U given by (5.76) is about 5% greater than that obtained by direct integration. It appears that the series in (5.76) do not contain enough terms to give good accuracy. This equation also shows that U is not proportional to a in the relativistic case and that, therefore, the current does not obey the 3/2 power law when relativistic effects are important.

Figure 5.9 shows, as an example, comparisons between the relativistic and non-relativistic solutions when $a = 1$ and the anode is outside, and inside, the cathode. It can be seen that, for a given value of R, the effect of relativity is to increase U.
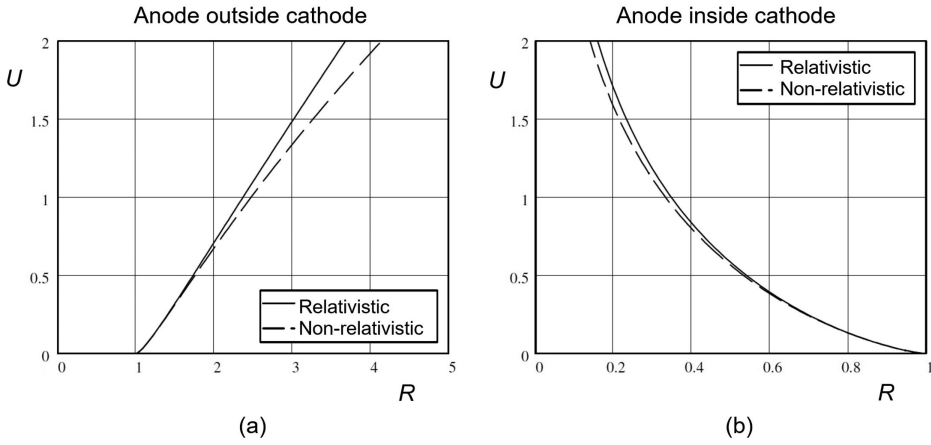
Figure 5.9: Comparisons between the relativistic and non-relativistic solutions for a cylindrical space-charge-limited diode when $a = 1$ and the anode is: (a) outside, and (b) inside, the cathode.

Thus, for a given geometry and applied voltage, the current is reduced by relativistic effects. It is found that U decreases when a decreases so that relativistic effects only become important for more extreme values of R.

5.6   The Spherical Space-Charge Limited Diode

The problem of space-charge limited flow between concentric spheres when the initial electron velocity is zero is discussed in [13] for the non-relativistic case and in [8] for the relativistic case. For symmetrical flow in a spherical diode Poisson's equation is

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r}\right) = -\frac{\rho}{\varepsilon_0}. \tag{5.78}$$

The total current I is independent of the radius so the current density is

$$J = \frac{I}{4\pi r^2}. \tag{5.79}$$

Then, from equations (5.11), (5.55) and (5.78)

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r}\right) = -\frac{I}{4\pi\varepsilon_0 c r^2}\left[1 - \frac{1}{\left(1 + V/V_R\right)^2}\right]^{-\frac{1}{2}}. \tag{5.80}$$

Again we set $R = r/r_c$, where r_c is the radius of the cathode, and $U = V/V_R$ so that (5.80) can be written

$$\frac{d^2U}{dR^2} + \frac{2}{R}\cdot\frac{dU}{dR} = \frac{k_2}{R^2}\cdot\frac{1+U}{\sqrt{2U+U^2}}, \tag{5.81}$$

where

$$k_2 = \frac{-I}{4\pi\varepsilon_0 c V_R}.$$ (5.82)

Equation (5.81) can be integrated by numerical methods with the initial conditions $U = 0$ and $dU/dR = 0$ when $R = 1$.

A power series solution for the non-relativistic case can be found by setting

$$\gamma = \ln R$$ (5.83)

and assuming that the current obeys the 3/2 power law so that (in SI units)

$$I = \frac{16\pi\varepsilon_0 \sqrt{2(e/m_0)}}{9} \cdot \frac{V^{\frac{3}{2}}}{\alpha^2},$$ (5.84)

where α is a function of R [13]. The following series expansion for α was found:

$$\alpha = \gamma - 0.3\gamma^2 + 0.075\gamma^3 - 0.0143182\gamma^4 + 0.0021609\gamma^5 - 0.00026791\gamma^6 + \cdots$$ (5.85)

Equation (5.84) can be written like (5.75) as

$$U = a\alpha^{\frac{4}{3}},$$ (5.86)

where

$$a^3 = \frac{81 k_2^2}{32}.$$ (5.87)

Comparison between (5.86) and the results of direct integration shows excellent agreement when $0.2 < R < 5$ (see Worksheet 5.3). For values of R outside this range Langmuir used direct integration.

Acton [8] obtained power series solutions in a similar manner. For the non-relativistic case

$$U = a\gamma^{\frac{4}{3}}\left(1 - \frac{2}{5}\gamma + \frac{3}{25}\gamma^2 + \ldots\right).$$ (5.88)

And, for the relativistic case when $U < 2$

$$U = a\gamma^{\frac{4}{3}}\left(1 - \frac{2}{5}\gamma + \frac{417}{25\times 144}\gamma^2 + \ldots\right) + a^2\gamma^{\frac{8}{3}}\left(-\frac{1}{14} + \frac{5}{7\times 72}\gamma + \ldots\right).$$ (5.89)

Here the first term on the right-hand side is similar, but not identical, to that in equation (5.88) so it appears that it is not permissible to regard the second term as a relativistic correction. These series do not contain enough terms to give good accuracy outside the range $0.5 < R < 2$.

Figure 5.10 shows, as an example, comparisons between the relativistic and non-relativistic solutions when $a = 1$ with the anode outside, and inside, the cathode. As
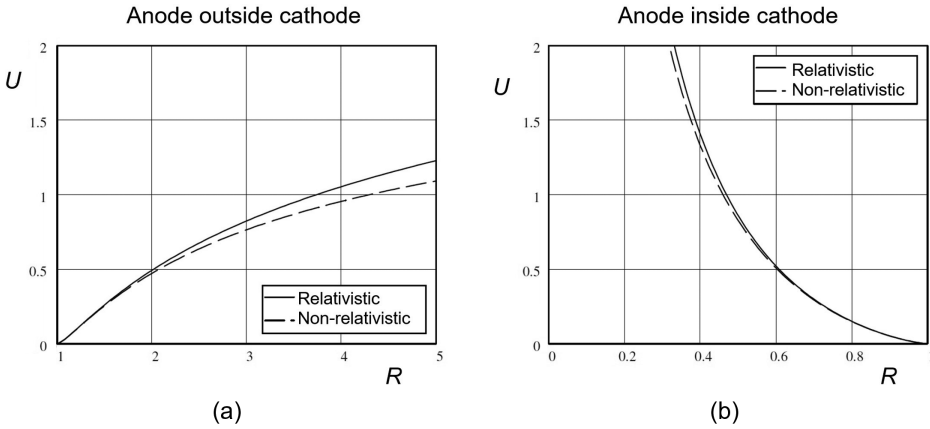
Figure 5.10: Comparisons between the relativistic and non-relativistic solutions for a spherical space-charge-limited diode when $a = 1$ and the anode is: (a) outside, and (b) inside, the cathode.

before, the effect of relativity is to increase U. Thus, for a given geometry and applied voltage, the current is reduced by relativistic effects. It is found that U decreases when a decreases so that relativistic effects only become important for more extreme values of R.

5.7   Transit-Time Effects in a Planar Diode

The analyses of the properties of space-charge-limited diodes in the preceding sections have assumed that the diodes are in the steady state. When the voltage across a diode varies with time this will no longer be the case if the transit time of an electron is comparable with the time during which the voltage changes. The high-frequency properties of a space-charge limited diode, when a small high-frequency voltage is superimposed upon a large DC voltage, are described by the Llewellyn-Peterson equations [14]. This can be important for low-power diodes but these are now of limited importance. High power diodes, however, are generally operated in the large-signal regime and it is usual to employ their static characteristics in design calculations. It is therefore important to know the conditions under which the static characteristics are a valid approximation to the dynamic behaviour.

It can be shown [15] that, when an electron current is flowing in a region with a number of electrodes,

$$\sum_{k=1}^{n} V_k i_k = \iiint \left(\rho \mathbf{v} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot \mathbf{E} \, dv, \tag{5.90}$$

where V_k is the potential of the kth electrode, i_k is the current flowing to it from the external circuit and the integral is taken over the whole volume occupied by the field

of the electrodes. This equation can be understood as a statement of the principle of conservation of energy. The left-hand side represents the instantaneous power flowing into the region and the two terms in the integral represent, respectively, the rate of change of the kinetic energy of the electrons and of the potential energy stored in the electric field. When this equation is applied to a planar diode that has one electrode earthed then

$$Vi = A \int \left(\rho v + \varepsilon_0 \frac{\partial E}{\partial t} \right) E \, dx, \tag{5.91}$$

where A is the area of each electrode. This equation shows that, under dynamic conditions, the current flowing to an electrode is not necessarily equal to the electron current landing on it. In particular, an induced current can start to flow before any electrons reach the anode. The term in the brackets in (5.91) is the total current whose divergence is zero. Thus the flux of the total current through any transverse plane is constant and depends only on time. To find the relationship between the external current and voltage under large-signal conditions it is necessary to start from the equations of motion of the electrons and apply the appropriate boundary conditions.

The large-signal high-frequency characteristics of vacuum tubes were studied by Wang [16] who investigated the behaviour of a diode when subjected to a short, parabolic, voltage pulse of amplitude V_{pk} and duration $2t_0$ given by

$$V(t) = \frac{t}{t_0} \left(2 - \frac{t}{t_0} \right) V_{pk} \quad \left(0 \le t \le 2t_0 \right). \tag{5.92}$$

This pulse approximates to a single cycle of a tube in class B or class C operation. When $t \le 0$ there are no electrons between the electrodes. When $t > 0$ electrons are emitted from the surface of the cathode at a rate sufficient to ensure that the electric field is zero there. When $t > t_0$ the voltage starts to drop and, at some point the emission ceases because the field of the electrodes is no longer sufficient to overcome the retarding field due to space-charge. The analysis assumes that the electrodes are close enough together for the propagation delay of the electromagnetic field to be negligible. Figure 5.11 shows plots of the positions of electrons as a function of time. Each curve is specified by the time τ at which the electron is emitted. A number of times which are of interest are marked:

t_1 = the time at which an electron emitted at $t = 0$ reaches the anode;
t_2 = the time at which the cathode stops emitting electrons;
t_3 = the time at which the last electron reaches the anode;
t_4 = the time at which the last electron leaves the space between the electrodes.

Four cases are shown in Figure 5.14:

(a) The transit time is negligible compared with the duration of the voltage pulse and the behaviour of the diode is quasi-static.
(b) The transit time is an appreciable fraction of the pulse duration, not all of the electrons are able to reach the anode, and the current pulse is appreciably
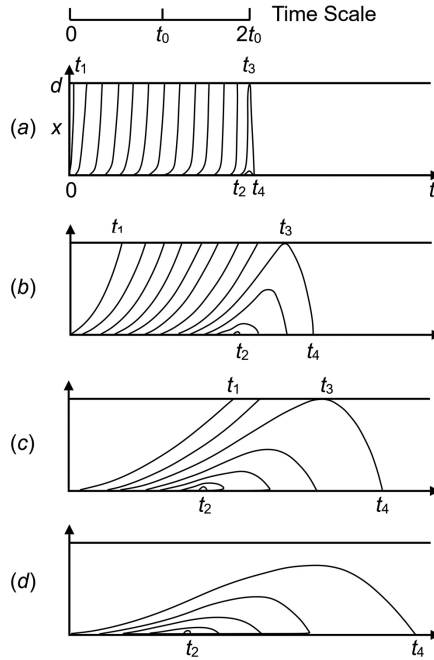
Figure 5.11. Electron trajectories in a diode when the applied voltage is given by (5.92) (copyright 1941, IEEE, reproduced, with permission, from [16]).

shorter than the voltage pulse. Electrons remain in the space between the electrodes for some time after the cathode stops emitting.

(c) The transit time of the electron emitted at $t = 0$ is a substantial fraction of the pulse duration and the cathode stops emitting electrons before the first electron reaches the anode.

(d) The reversal of the voltage is so fast that no electrons are able to reach the anode.

For our purposes it is case (a) which is of interest. When the transit time is small the diode is completely filled with electrons throughout the pulse. The transit time is greatest at the start and end of the pulse when the voltage is small. Wang provides a detailed analysis of this case and shows that, when second-order terms are included, the convection current at the anode is given by

$$J(t) = \frac{4\varepsilon_0}{9}\sqrt{\frac{2e}{m_0}} \cdot \frac{V(t)^{\frac{3}{2}}}{d^2} \quad - \quad \frac{2\varepsilon_0}{5d} \cdot \frac{dV(t)}{dt}. \tag{5.93}$$

The first term on the right hand side of this equation is identical to (5.16) and the second term is 40% of the reactive current flowing in the capacitance of the diode, but with the sign reversed. To investigate the point at which transit time effects first become important we consider a diode subjected to a sinusoidal voltage given by

$$V(t) = V_a \sin(\omega t). \tag{5.94}$$

The first term in (5.93) is non-zero only during the positive half-cycle. This waveform can be expanded as a Fourier series. If only the first harmonic term is retained then substitution in (5.93) gives the approximation

$$J(t) \simeq \frac{4\varepsilon_0 V_a u_0}{9d^2}\left[0.458\sin(\omega t) \; - \; \frac{9}{10}\beta_e d \cdot \cos(\omega t)\right]$$

$$\simeq \frac{\varepsilon_0 V_a u_0}{5d^2}\left[\sin(\omega t) \; - \; 2\beta_e d \cdot \cos(\omega t)\right], \tag{5.95}$$

where $\beta_e d = \omega d/u_0$ and $u_0 = \sqrt{2eV_a/m_0}$. The transit time effects are represented by the second term in the brackets whose magnitude depends on $\beta_e d$. This factor increases with frequency, and with the diode spacing, and decreases as the peak voltage applied to the diode is increased. Transit time effects therefore increase the magnitude of the current by a factor of approximately $\sqrt{1+(2\beta_e d)^2}$ and retard its phase by $2\beta_e d$. The condition that transit time effects should be negligible is therefore $\beta_e d \ll 1/\sqrt{2}$. This can be used as a guide to the spacings required between electrodes if transit time effects are to be avoided. It is important to note that (5.95) can only be used to calculate the convection current when transit time effects are small.

5.8   Injection of Electrons into a Planar Diode

We saw in Section 5.3 that it is not possible to use a single initial velocity to model the emission from a thermionic cathode. However, there are other cases where this problem is of interest. Consider, for example, a plane-parallel diode in which the electrodes are conducting grids that are perfectly transparent to the electrons. Let a stream of electrons having uniform current density J_i and velocity u_0 normal to the grids be injected into the space between them as shown in Figure 5.12. This problem has been studied by many authors (see [17] and the references cited therein). Here we follow the treatment in [18].

If relativistic effects are neglected, the initial velocity of the electrons is given by (5.10) as

$$u_0 = \sqrt{2(e/m_0)V_k}, \tag{5.96}$$

where the source of electrons is at potential $-V_k$ relative to the electrode at $x = 0$, then the velocity at a plane where the potential is V is

$$u = u_0 \sqrt{1+V/V_k}. \tag{5.97}$$

Making use of (5.11) and (5.12) we obtain

$$\frac{d^2V}{dx^2} = -\frac{J_i}{\varepsilon_0 u_0 \sqrt{1+V/V_k}}. \tag{5.98}$$
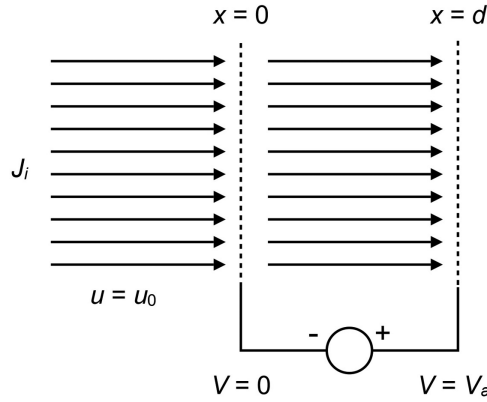
Figure 5.12: Injection of electrons into a diode formed by parallel conducting grids.

Multiplication by $2(dV/dx)$ as before and integrating yields

$$\left(\frac{dV}{dx}\right)^2 = -\frac{4J_i V_s}{\varepsilon_0 u_0} \cdot \sqrt{1+V/V_k} \quad + \quad C_1, \qquad (5.99)$$

where C_1 is a constant. It is convenient to introduce normalised variables $U = V/V_k$ and

$$X = \sqrt{\frac{-4J_i}{\varepsilon_0 u_0 V_k}} \cdot x. \qquad (5.100)$$

Substitution in (5.99) gives

$$\frac{dU}{dX} = \pm \left(\sqrt{1+U} \quad + \quad \alpha_1\right)^{\frac{1}{2}}, \qquad (5.101)$$

where α_1 is a dimensionless constant. Setting $W = \sqrt{1+U}$ and integrating a second time gives

$$\frac{4}{3}(W - 2\alpha_1)(W + \alpha_1)^{\frac{1}{2}} = \pm X + \alpha_2, \qquad (5.102)$$

where α_2 is a second constant of integration. We note that (5.101) implies that the gradient of the potential may be either positive or negative. If the gradient is positive then the potential will increase monotonically through the diode whereas, if the gradient is initially negative there will be a potential minimum followed by a monotonic increase.

a) Monotonic increase (Branch I)

Let $X = X_a$ correspond to $x = d$ and let

$$W_a = \sqrt{1+U_a} = \sqrt{1+V_a/V_k} \qquad (5.103)$$

so that $W = 1$ when $x = 0$. Substituting these boundary conditions in (5.102) and taking the plus sign gives

$$\frac{4}{3}(1-2\alpha_1)(1+\alpha_1)^{\frac{1}{2}} = \alpha_2 \qquad (5.104)$$

and

$$\frac{4}{3}(W_a - 2\alpha_1)(W_a + \alpha_1)^{\frac{1}{2}} = X_a + \alpha_2. \qquad (5.105)$$

The constants α_1 and α_2 can then be determined and, hence, the potential distribution from (5.102) (taking the positive sign).

b) Negative potential minimum (Branch II)

Let the position of the minimum be $X = X_m$. In the range $0 \le X \le X_m$

$$\frac{4}{3}(W - 2\alpha_1)(W + \alpha_1)^{\frac{1}{2}} = -X + \alpha_3 \qquad (5.106)$$

and, in the range $X \ge X_m$

$$\frac{4}{3}(W - 2\alpha_1)(W + \alpha_1)^{\frac{1}{2}} = X + \alpha_4, \qquad (5.107)$$

where α_3 and α_4 are new constants of integration. Since W is continuous we find, from (5.101), that $W + \alpha_1 = 0$ when $X = X_m$ so that

$$\alpha_3 = X_m \quad \text{and} \quad \alpha_4 = -X_m. \qquad (5.108)$$

Substituting for α_3 in (5.106) and setting $W = 1$ when $X = 0$ gives

$$\frac{4}{3}(1 - 2\alpha_1)(1 + \alpha_1)^{\frac{1}{2}} = X_m \qquad (5.109)$$

and setting $W = W_a$ when $X = X_a$ equation (5.107) becomes

$$\frac{4}{3}(W_a - 2\alpha_1)(W_a + \alpha_1)^{\frac{1}{2}} = X_a - X_m. \qquad (5.110)$$

The properties of a diode with injected current can be determined from equations (5.104), (5.105), (5.109) and (5.110) (see Worksheet 5.4). We note that $\alpha_1 \ge -1$ for both branches because α_2 and X_m are real. Equation (5.101) requires $\alpha_1 \le 0$ for branch II because otherwise it is not possible for the gradient of the potential to be zero. The solution has two branches corresponding to the two cases described above. It is convenient to plot graphs of X_a against α_1 with U_a as a parameter, as shown in Figure 5.13. Branch I solutions are plotted as solid curves and branch II solutions as dashed curves in this figure.

Solutions for given values of X_a and U_a are found from the intersections between lines of constant X_a and U_a. It can be seen from the figure that the number of intersections can be zero, one, or two. If the diode spacing is fixed and the diode voltage, and the injection voltage, are held constant then U_a is constant. When the injected current is gradually increased then X_a increases from zero. Initially there is only a
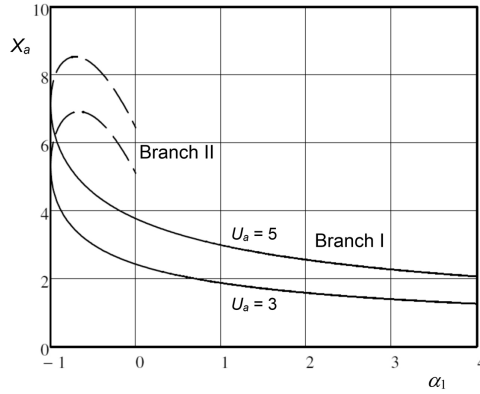
Figure 5.13: Dependence of the normalised diode spacing X_a on α_1 with the normalised potential difference U_a as a parameter
(copyright 1944, The American Physical Society, reproduced, with permission, from [18]).

single intersection with branch I for a large value of α_1, and the potential in the diode increases monotonically with x. As X_a increases a point is reached where two solutions are possible: one in branch I, and the other in branch II. It can be shown that the branch II solution has higher potential energy and is therefore unstable. Eventually at $\alpha_1 = -1$ the space charge becomes sufficient to generate a potential minimum and there are two solutions on branch II of which the one corresponding to the more negative value of α_1 is stable. Finally the two solutions coalesce at the maximum of the curve of constant U_a, the current is space-charge limited and cannot increase any further. It is straightforward to show that at the maximum

$$\alpha_1 = -\frac{W_a}{1+W_a}.\tag{5.111}$$

Substituting in (5.110) from (5.109) and (5.111) gives the maximum value of X_a for a given value of W_a

$$X_{\max} = \frac{4}{3}(1+W_a)^{\frac{3}{2}}.\tag{5.112}$$

This equation can be expressed in terms of physical variables in the form

$$J_i = -\frac{4}{9}\cdot\frac{\varepsilon_0\sqrt{2(e/m_0)}}{x^2}V_a^{3/2}\left(\sqrt{V_k/V_a}+\sqrt{1+V_k/V_a}\right)^3,\tag{5.113}$$

which can be compared directly with the Child–Langmuir Law given by (5.16). It can be seen that the two equations become identical when $V_k \to 0$. For non-zero values of the injection potential the last term in the equation is greater than unity and the maximum current which can flow through the diode is therefore increased as a result of the finite injection velocity. We note that the potential at the minimum is always greater than that at the source of electrons, and that there is no plane within the diode at which the electrons come to rest.

When the electrons are injected into the diode, the current is usually determined externally. It is therefore of interest to enquire what happens if we attempt to inject a current greater than that given by (5.113). This problem was studied by Langmuir [7]. The additional space charge depresses the potential at the minimum so that a virtual cathode is formed at which $V = -V_k$ and the electrons come to rest. We have to assume that the electrons do not all have the same injection velocity, but that there is a small velocity spread caused by their emission from a thermionic cathode. As a consequence some electrons are reflected by the virtual cathode while others pass onwards and are collected by the anode. We saw in Section 3.2.1 that, for a small spread of velocities, the behaviour of the diode is described by the Child–Langmuir Law. It is therefore justifiable to assume partial reflection of the electrons while ignoring the velocity spread. We assume that the injected current density is J_i, that the current density at the anode is J_a, and that the virtual cathode forms at $x = x_m$. In the region $0 \le x \le x_m$ the effective current density is $(2J_i - J_a)$ since the space-charge density is not affected by the direction of motion of the electrons. Then, from (5.16) we have

$$2J_i - J_a = -\frac{4}{9} \cdot \frac{\varepsilon_0 \sqrt{2(e/m_0)}}{(x_m - x)^2} \cdot (V + V_k)^{\frac{3}{2}}.$$

(5.114)

Equation (5.114) can be written

$$X = X_m - \frac{4}{3} \cdot \sqrt{\frac{J_a}{2J_i - J_a}} \cdot W^{\frac{3}{2}},$$

(5.115)

where

$$X = \sqrt{\frac{-4J_a}{\varepsilon_0 u_0 V_k}} \cdot x$$

(5.116)

and

$$W = \sqrt{1 + \frac{V}{V_k}}.$$

(5.117)

When $X = 0$, $W = 1$, and therefore

$$X_m = \frac{4}{3} \cdot \sqrt{\frac{J_a}{2J_i - J_a}}.$$

(5.118)

Substituting in (5.115) gives for the region $0 \le x \le x_m$

$$X = \frac{4}{3} \cdot \sqrt{\frac{J_a}{2J_i - J_a}} \cdot \left(1 - W^{\frac{3}{2}}\right).$$

(5.119)

Similarly in the region $x_m \le x \le d$

$$J_a = -\frac{4}{9} \cdot \frac{\varepsilon_0 \sqrt{2(e/m_0)}}{(x - x_m)^2} \cdot (V + V_k)^{\frac{3}{2}}, \qquad (5.120)$$

which becomes

$$X = \frac{4}{3} \cdot W^{\frac{3}{2}} + \frac{4}{3} \cdot \sqrt{\frac{J_a}{2J_i - J_a}}. \qquad (5.121)$$

If there is no reflection the anode current is equal to the injected current and, at the anode,

$$X_a = \frac{4}{3}\left(1 + W_a^{\frac{3}{2}}\right). \qquad (5.122)$$

This is identical to the result obtained by setting $\alpha_1 = 0$ in (5.110). We note that this is less than the space-charge limited value of X given by (5.112). If the injected current is increased, at constant electron velocity and anode voltage, then the last term in (5.121) decreases and, therefore, so does the anode current. In particular, it is observed experimentally that, if the injected current is increased slightly above the space-charge limit, then there is an abrupt reduction in the anode current [19–20]. Thus the behaviour of such a diode is chaotic. This behaviour is exploited in virtual cathode oscillators [21]. Further information about diodes with injected current is to be found in [17] and [22].

5.9   Diodes with Two-Dimensional Flow of Current

In many cases the electron flow in a diode can be regarded as one dimensional. However, the emitting surface necessarily has a finite area, and it is usual to provide focusing electrodes at the edges which are designed to compensate for edge effects to produce near-uniform current density on the cathode surface. This technique is discussed in detail in Chapter 9.

If no focusing electrodes are used then the electron flow becomes two, or three, dimensional according to the geometry. Problems of this nature have been studied by a number of authors using Particle in Cell (PIC) codes. Two-dimensional flow between planar electrodes, when the electrons are emitted from a narrow strip on the surface of one of them, was studied in [23] using two different simulations. The results from both simulations showed good agreement. The current density across the emitting surface was assumed to be uniform and increased until oscillatory behaviour was observed. This was taken to define the limiting current in two-dimensional flow. The current was limited by the formation of a virtual cathode at the centre of the emitting strip. Simulations were carried out both with, and without, a uniform magnetic field normal to the electrodes. It was found that

the limiting current was almost independent of the intensity of the magnetic field. The results obtained were generalised as an empirical geometrical correction to the Child–Langmuir law (5.16) given by

$$\frac{J_{2D}}{J_{1D}} = 0.3145\left(\frac{d}{w}\right) - 0.0004\left(\frac{d}{w}\right)^2, \tag{5.123}$$

where d is the distance between the electrodes, w is the width of the emitting strip, J_{2D} is the limiting current density and J_{1D} is the one-dimensional Child–Langmuir current density given by (5.16). Equation (5.123) was found to be valid to an accuracy of 5% in the range $0.1 < w/d < 8$, and for magnetic fields in the range 0 to 100 T.

Further studies of this problem showed that, when the current density emitted from the cathode was allowed to be non-uniform, there was a sharp rise in the current density at the edges of the strip as shown in Figure 5.14 [24]. This is to be expected because the absence of space-charge in the vacuum region immediately outside the strip reduces the space-charge potential depression that limits the current. When the strip is wide compared with the gap, the current density is given by the one-dimensional Child–Langmuir law over the greater part of the width of the strip, only rising at the extreme edges. When $w = 2d$ the current density is greater than that given by the one-dimensional law, except at the centre of the strip. For narrower normalised strip widths the current density everywhere exceeds that given by the one-dimensional law.

When a step was introduced in the gap between the electrodes, opposite the centre of the emitting strip, it was found that the emitted current density closely followed the square of the local vacuum electric field, except at the edges of the strip. It was also found that a very slight change in the electric field at the edge of the strip, such as that produced by a small non-emitting focus electrode, could have a significant effect on the emitted current density.
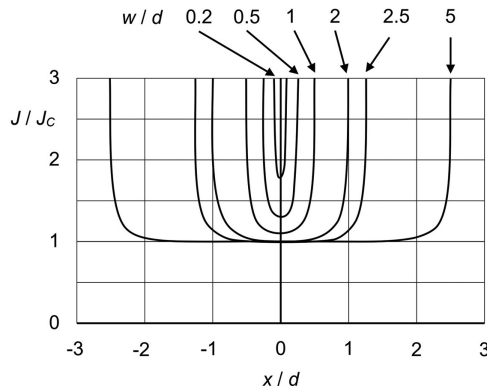
Figure 5.14: Simulated normalized current density emitted at cathode versus position (normalized to gap distance) for a variety of emission strip widths (w) and gap distances (d) in cm; each trace is labeled with its corresponding (w/d) value
(copyright 2001, The American Physical Society, reproduced, with permission, from [24]).

The scaling of the equations governing space-charge limited emission showed that the ratio J_{2D}/J_{1D} is a universal function of w/d for zero emission velocity [23]. This universal function was derived in [25] where it was shown that, to first order,

$$\frac{J_{2D}}{J_{1D}} \cong 1 + \frac{d}{\pi w} = 1 + 0.318 \frac{d}{w}. \qquad (5.124)$$

This expression is very close to the empirical expression in (5.123). The same method was also used to find the geometrical correction when the electrons are emitted from a circular patch on the cathode of radius R

$$\frac{J_{2D}}{J_{1D}} = 1 + \frac{d}{4R}. \qquad (5.125)$$

It was stressed that neither (5.124) nor (5.125) give the correct limits as W and R tend to zero. A review of these results was given in [26]. The method in [25] was extended in [27] to a number of other cases with plane parallel electrodes and finite emitting patches of varying shapes including squares, rectangles, and polygons.

The space-charge limited current flow between concentric cylinders, when electrons are emitted from a band on one of them, was studied in [28]. The simulations ignored the self magnetic field of the electron flow and assumed that the current density was uniform on the surface of the emitter. Both convergent flow (anode inside the cathode) and divergent flow were considered. It was found that both cases were described by the same empirical correction factor

$$\frac{I_{2D}}{I_{LB}} = 1 + 0.1536 \left(\frac{R}{L} \right) + 0.0183 \left(\frac{R}{L} \right)^2, \qquad (5.126)$$

where R is the radius of the outer electrode, L is the length of the emitting band and I_{LB} is the current calculated for a length L using (5.70). Equation (5.126) was found to be valid to an accuracy of 2.5% for $R/r = 3$ and 10 (where r is the radius of the inner electrode) and $0.2 < L/R < 5$. The reason why (5.126) does not depend upon R/r is that the value of β in (5.70) rapidly approaches an asymptotic value when $R/3 > 3$. When $R/r < 3$ this is no longer the case and (5.126) ceases to be valid. A theoretical expression for this case takes the form

$$\frac{I_{2D}}{I_{LB}} = 1 + \frac{4(R-r)}{\pi L} F(R/r), \qquad (5.127)$$

where $F(R/r)$ is a function (plotted in [27]) that takes different values for convergent and divergent flow. In the limit $R/r \to 1$, $F \to 1/4$, which is approximately the average of the values for convergent and divergent flow for $R/r > 1$. It is shown that (5.126) and (5.127) give similar results for $R/r = 3$ and 10 when $L/R > 1$ but that they differ for smaller values of R/r where (5.126) does not tend to the planar limit as it should. Neither method is accurate for $L/R < 1$ because the enhanced emission at the edges of the emitting band has not been included.

References

[1] I. Langmuir, 'The effect of space charge and residual gases on thermionic currents in high vacuum', *Physical Review,* vol. 2, pp. 450–486, 1913.

[2] S. E. Tsimring, *Electron Beams and Microwave Vacuum Electronics*. Hoboken, NJ: Wiley, 2007.

[3] R. Vaughan, 'A synthesis of the Longo and Eng cathode emission models', *IEEE Transactions on Electron Devices,* vol. 33, pp. 1925–1927, 1986.

[4] C. D. Child, 'Discharge from hot CaO', *Physical Review (Series I),* vol. 32, pp. 492–511, 1911.

[5] R. J. Umstattd *et al.*, 'A simple physical derivation of Child-Langmuir space-charge-limited emission using vacuum capacitance', *American Journal of Physics,* vol. 73, pp. 160–163, February 2005.

[6] I. Langmuir, 'The effect of space charge and initial velocities on the potential distribution and thermionic current between parallel plane electrodes', *Physical Review,* vol. 21, pp. 419–435, 1923.

[7] I. Langmuir and K. T. Compton, 'Electrical discharges in gases Part II: fundamental phenomena in electrical discharges', *Reviews of Modern Physics,* vol. 3, pp. 191–257, 1931.

[8] E. W. V. Acton, 'The space-charge limited flow of charged particles in planar, cylindrical and spherical diodes at relativistic velocities', *International Journal of Electronics,* vol. 3, pp. 203–210, August 1957.

[9] J. E. Boers and D. Kelleher, 'Exact solution of Poisson's equation for space-charge-limited flow in a relativistic planar diode', *Journal of Applied Physics,* vol. 40, pp. 2409–2412, 1969.

[10] H. R. Jory and A. W. Trivelpiece, 'Exact relativistic solution for the one-dimensional diode', *Journal of Applied Physics,* vol. 40, pp. 3924–3926, 1969.

[11] Y. P. Zhang *et al.*, 'Simple solutions for relativistic generalizations of the Child–Langmuir law and the Langmuir–Blodgett law', *Physics of Plasmas,* vol. 16, p. 044511, April 2009.

[12] I. Langmuir and K. B. Blodgett, 'Currents limited by space charge between coaxial cylinders', *Physical Review,* vol. 22, pp. 347–356, 1923.

[13] I. Langmuir and K. B. Blodgett, 'Currents limited by space charge between concentric spheres', *Physical Review,* vol. 24, pp. 49–59, 1924.

[14] F. B. Llewellyn, 'Vacuum tube electronics at ultra-high frequencies', *Proceedings of the Institute of Radio Engineers,* vol. 21, pp. 1532–1573, 1933.

[15] C. K. Jen, 'On the induced current and energy balance in electronics', *Proceedings of the I.R.E.,* vol. 29, pp. 345–349, 1941.

[16] C. C. Wang, 'Large-signal high-frequency electronics of thermionic vacuum tubes', *Proc. IRE,* pp. 200–214, April 1941.

[17] P. V. Akimov *et al.*, 'The true nature of space-charge-limited currents in electron vacuum diodes: a Lagrangian revision with corrections', *Physics of Plasmas,* vol. 8, pp. 3788–3798, 2001.

[18] G. Jaffe, 'On the currents carried by electrons of uniform initial velocity', *Physical Review,* vol. 65, pp. 91–98, 1944.

[19] E. W. B. Gill, 'XCIX. A space charge effect', *Philosophical Magazine Series 6,* vol. 49, pp. 993–1005, 1925.

[20] V. I. Kuznetsov and E. A. Ya., 'Time-dependent regimes of a Bursian Diode I: Stability of steady solutions', *Plasma Physics Reports,* vol. 36, pp. 226–235, 2010.

[21] J. Benford *et al.*, *High Power Microwaves*. Boca Raton, FL: CRC Press, 2007.

[22] H. Kolinsky and H. Schamel, 'Arbitrary potential drops between collector and emitter in pure electron diodes', *Journal of Plasma Physics,* vol. 57, pp. 403–423, 1997.

[23] J. W. Luginsland *et al.*, 'Two-dimensional Child–Langmuir law', *Physical Review Letters,* vol. 77, pp. 4688–4670, 1996.

[24] R. J. Umstattd and J. W. Luginsland, 'Two-dimensional space-charge-limited emission: beam-edge characteristics and applications', *Physical Review Letters,* vol. 87, p. 145002, 2001.

[25] Y. Y. Lau, 'Simple theory for the two-dimensional Child–Langmuir law', *Physical Review Letters,* vol. 87, p. 278301, 2001.

[26] J. W. Luginsland *et al.*, 'Beyond the Child–Langmuir law: a review of recent results on multidimensional space-charge-limited flow', *Physics of Plasmas,* vol. 9, pp. 2371–2376, May 2002.

[27] W. S. Koh *et al.*, 'Three-dimensional Child–Langmuir law for uniform hot electron emission', *Physics of Plasmas,* vol. 12, p. 053107, 2005.

[28] K. G. Kostov and J. J. Barroso, 'Space-charge-limited current in cylindrical diodes with finite-length emitter', *Physics of Plasmas,* vol. 9, p. 1039, 2002.

6   Triodes and Tetrodes

6.1   Introduction

A triode is constructed by placing a conducting *control grid* between the cathode and the anode of a vacuum diode, as shown schematically in Figure 6.1(a). The current drawn from the cathode depends on the potentials of the grid and the anode with respect to the cathode. The anode is known as the *plate* in the USA. The grid potential can therefore be used to control the current flowing to the anode. The electron flow between the cathode and the grid is space-charge limited, and it is found experimentally that the properties of many tubes can be described by the equation

$$I_a = K \left(V_g + V_a / \mu \right)^{\frac{3}{2}},\qquad (6.1)$$

where I_a is the anode current, V_g and V_a are the potentials of the grid and the anode with respect to the cathode, and K and μ are constants. In some tubes the power of the term in the brackets is greater than 1.5 when the grid voltage is negative, and it may be as high as 2.5 for reasons which are explained later. It is also found that K and μ are not precisely constant but may vary somewhat with the conditions of operation. The parameter K is the *perveance* by analogy with the perveance of a diode and μ is known as the *amplification factor* from its role in the small-signal theory of triode amplifiers in which [1]

$$\mu = \left(\frac{\partial V_a}{\partial V_g} \right)_{I_a}.\qquad (6.2)$$

This parameter is typically in the range from 5 to 200 depending upon the purpose for which the tube has been designed [2]. The *penetration factor*, defined by $D = 1/\mu$ is more useful when the physics of the device are under consideration. The symbol (D) comes from the German for penetration (*Durchgriff*), Other parameters used in small-signal theory are the *mutual conductance* (or *transconductance*) defined by

$$g_m = \left(\frac{\partial I_a}{\partial V_g} \right)_{V_a}\qquad (6.3)$$
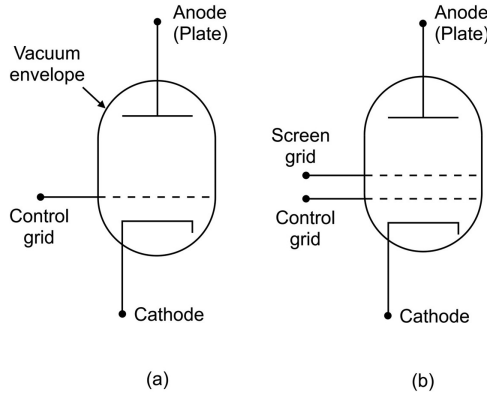
Figure 6.1: Schematic diagrams of (a) the triode, and (b) the tetrode.

and the *dynamic anode resistance* (or *anode slope resistance*) defined by

$$r_a = \left(\frac{\partial V_a}{\partial I_a} \right)_{V_g}.$$ (6.4)

From these definitions it follows that

$$\mu = g_m r_a.$$ (6.5)

It must be emphasised that the values of these parameters depend on the quiescent point chosen for operation of the tube. They are therefore of limited value in dealing with the operation of triodes under large-signal conditions. They are mentioned here for completeness because their values are sometimes given in manufacturers' catalogues.

Figure 6.1(b) shows a schematic diagram of a tetrode in which a second grid (the *screen grid*) has been placed between the control grid and the anode. The screen grid is maintained at a constant positive potential so that it screens the cathode from the varying voltage of the anode. The characteristics of a tetrode are given approximately by

$$I_a = K \left(V_g + \frac{V_s}{\mu_s} + \frac{V_a}{\mu_a} \right)^{\frac{3}{2}},$$ (6.6)

where V_s is the potential of the screen grid with respect to the cathode. Typically $\mu_s \sim 5-10$ and $\mu_a \sim 100-200$. The basic theory of triodes and tetrodes is discussed in this chapter whilst practical devices are dealt with in Chapter 12.

6.2   Electrostatic Models of Triodes

The accurate calculation of the performance of triodes and tetrodes requires detailed modelling of the electron trajectories under the influence of the electrode

potentials, space-charge, and secondary electron emission. This can be achieved with a particle in cell (PIC) code [3]. The objective, here, is to provide the reader with simple models which provide insight into the properties of gridded tubes and the means to make conceptual design calculations. This is approached through the use of electrostatic models from which the large-signal behaviour, including the effects of space-charge, can be inferred.

Figure 6.2 shows the arrangement of a planar triode whose geometry is defined by the parameters d_1, d_2, r, a. The geometrical screening factor is defined by $S = 2r/a$. In the absence of space-charge, the normal component of the electric field on the surface of the cathode is linearly dependent on the potentials of the grid and the anode with respect to the cathode. Thus we may write

$$E_x(0, y) = -\frac{V_g}{d_g(y)} - \frac{V_a}{d_a(y)}, \tag{6.7}$$

where $d_g(y)$ and $d_a(y)$ are unknown functions of y which have the dimensions of distance. Equation (6.7) can be rewritten as

$$E_x(0, y) = -\frac{V_g + D(y)V_a}{d_g(y)}, \tag{6.8}$$

where $D(y)$ is the penetration factor which, in general, varies over the surface of the cathode. Now consider the case where the grid is thin and biased to the potential which would exist at that plane if the grid were removed. Thus

$$V_g = \frac{d_1}{d_1 + d_2} V_a. \tag{6.9}$$

Then, substituting in (6.8)

$$E_x(0, y) = -\frac{d_1 + D(y)(d_1 + d_2)}{d_g(y)} \cdot \frac{V_a}{d_1 + d_2}. \tag{6.10}$$
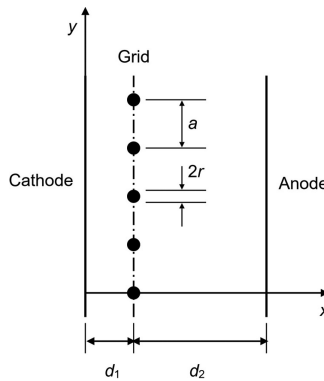
Figure 6.2: Arrangement of a planar triode.

But, in the absence of the grid, the electric field on the cathode surface is uniform and

$$E_x(0) = \frac{V_a}{d_1 + d_2}.$$ (6.11)

Thus

$$d_g(y) = d_1 + D(y)(d_1 + d_2)$$ (6.12)

and, substituting in (6.8),

$$E_x(0, y) = -\frac{V_g + D(y)V_a}{d_1 + D(y)(d_1 + d_2)}.$$ (6.13)

In many cases, as we shall see, the electric field is uniform over the surface of the cathode and then (6.13) reduces to

$$E_x(0) = -\frac{V_g + DV_a}{d_1 + D(d_1 + d_2)},$$ (6.14)

where D is a constant. The same argument can be used to derive the corresponding equations for cylindrical triodes. These are not required here, however, because the distances between the electrodes in coaxial high power triodes are small compared with their diameters so that the expressions for planar geometry may be used.

It is useful to investigate the relationship between the expressions derived, and the inter-electrode capacitances shown in Figure 6.3. It should be noted that these capacitances are not normally the same as those measured at the external terminals of the tube because of the capacitances of the connections between the internal electrodes and the terminals. The charge per unit length on the section of the cathode associated with one grid wire is given by

$$q_c = -\left(C_{cg}V_g + C_{ca}V_a\right).$$ (6.15)

If the charge density on the cathode is uniform then electric field normal to the cathode is

$$E_c = \frac{q_c}{a\varepsilon_0} = -\frac{1}{a\varepsilon_0}\left(C_{cg}V_g + C_{ca}V_a\right),$$ (6.16)
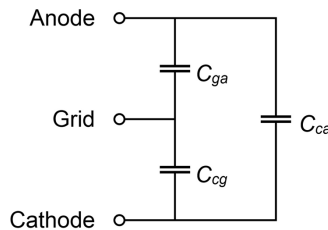
Figure 6.3: Inter-electrode capacitances in a triode.

where a is the width of the strip of cathode between two adjacent grid wires. Comparing (6.16) with (6.14) we see that

$$C_{cg} = \frac{a\varepsilon_0}{d_1 + D(d_1 + d_2)} \tag{6.17}$$

and

$$C_{ca} = \frac{Da\varepsilon_0}{d_1 + D(d_1 + d_2)} = DC_{cg} \tag{6.18}$$

so that

$$D = C_{ca}/C_{cg}. \tag{6.19}$$

Equation (6.17) can be rewritten as

$$C_{cg} = \frac{a\varepsilon_0}{d_1\left(1 + d_2 D\left(\dfrac{1}{d_1} + \dfrac{1}{d_2}\right)\right)}. \tag{6.20}$$

This can be interpreted as the capacitance of a parallel plate capacitor whose spacing is modified by the term in brackets which represents the difference between the grid and a metal sheet. The modification should be independent of the position of the grid, provided that the distances between it and the other electrodes are greater than a. Thus the quantity $d_2 D$ should depend only upon the dimensions of the grid. We shall see later that this conclusion is confirmed by field theory. Similarly, the capacitance between the grid and the anode is obtained by interchanging d_1 and d_2 to give

$$C_{ga} = \frac{a\varepsilon_0}{d_2\left(1 + d_2 D\left(\dfrac{1}{d_1} + \dfrac{1}{d_2}\right)\right)} = \frac{d_1}{d_2} C_{cg}. \tag{6.21}$$

The conclusions drawn from this analysis are valid as long as the grid wires are small compared with the other dimensions.

6.3   Penetration Factor in a Planar Triode

The potential distribution in a planar triode can be found by placing a line charge at the centre of each grid wire. This problem was first solved by Maxwell using conformal mapping [4]. The solution is derived from that of a transformed problem in which two line charges are placed within a conducting cylinder [5]. The potential of a grid of equally spaced wires is given by

$$V(x, y) = -\frac{q_g}{4\pi\varepsilon_0} \ln\left[2\left(\cosh\frac{2\pi x}{a} - \cos\frac{2\pi y}{a}\right)\right] + C, \tag{6.22}$$

where q_g is the charge per unit length on each wire and C is a constant. If the grid is located at $x = d_1$ and a second grid carrying charge $-q_g$ is located at $x = -d_1$ then the potential on the plane $x = 0$ is zero and that plane can be taken to be the cathode. Adding a linear variation of potential, for the effect of the anode, gives

$$V(x,y) = \frac{1}{a\varepsilon_0}\left(f_g(x,y)q_g + xq_a\right), \tag{6.23}$$

where

$$f_g(x,y) = \frac{a}{4\pi}\ln\left[\frac{\cosh 2\pi(x+d_1)/a - \cos 2\pi y/a}{\cosh 2\pi(x-d_1)/a - \cos 2\pi y/a}\right] \tag{6.24}$$

and q_a is the charge per unit length on a strip of the anode of width a. On the grid $V_g = V(d_1, r)$ so that

$$V_g = \frac{1}{a\varepsilon_0}\left(f_g(d_1,r)q_g + d_1 q_a\right). \tag{6.25}$$

The anode must be an equipotential surface and, therefore, when $x = d_1 + d_2$ the hyperbolic cosines must be much greater than unity. This condition is satisfied if $d_2/a \geq 1$ (when $\cosh(2\pi d_2/a) \geq 268$) and then, to a good approximation,

$$f_g(d_1+d_2, y) = \frac{a}{4\pi}\ln\left[\frac{\exp 2\pi(2d_1+d_2)/a}{\exp 2\pi d_2/a}\right] = d_1. \tag{6.26}$$

Thus

$$V_a = \frac{1}{a\varepsilon_0}\left(d_1 q_g + (d_1+d_2)q_a\right). \tag{6.27}$$

Equations (6.27) and (6.25) can be written as

$$\begin{bmatrix} V_g \\ V_a \end{bmatrix} = \frac{1}{a\varepsilon_0}\begin{bmatrix} f_g(d_1,r) & d_1 \\ d_1 & d_1+d_2 \end{bmatrix}\begin{bmatrix} q_g \\ q_a \end{bmatrix}. \tag{6.28}$$

Inverting this equation we obtain

$$\begin{bmatrix} q_g \\ q_a \end{bmatrix} = C_1\begin{bmatrix} d_1+d_2 & -d_1 \\ -d_1 & f_g(d_1,r) \end{bmatrix}\begin{bmatrix} V_g \\ V_a \end{bmatrix}. \tag{6.29}$$

where

$$C_1 = \frac{a\varepsilon_0}{(d_1+d_2)f_g(d_1,r) - d_1^2}. \tag{6.30}$$

6.3.1   A Triode with Uniform Electric Field on the Cathode

If $d_1/a \geq 1$ the potential is independent of y close to the cathode and, to a good approximation,

$$f_g(d_1,r) = \frac{a}{4\pi} \ln\left[\frac{\exp(4\pi d_1/a)}{2 - 2\cos(2\pi r/a)}\right]$$
$$= d_1 - \frac{a}{2\pi} \ln\left[2\sin\left(\frac{\pi r}{a}\right)\right] \tag{6.31}$$

and $f_g(x,y) \rightarrow x$ as $x \rightarrow 0$. The normal component of the electric field on the cathode is given by

$$E_x(0,y) = -\frac{(q_g + q_a)}{a\varepsilon_0} = -\frac{d_2 V_g + (f_g(d_1,r) - d_1)V_a}{(d_1 + d_2)f_g(d_1,r) - d_1^2}. \tag{6.32}$$

Comparing (6.32) with (6.14) we see that

$$d_2 D = f_g(d_1,r) - d_1 \tag{6.33}$$

so, from (6.31),

$$D = -\frac{a}{2\pi d_2} \ln\left[2\sin\left(\frac{\pi r}{a}\right)\right] \tag{6.34}$$

and $d_2 D$ depends only on r and a as required by (6.20).

The validity of (6.34) has been confirmed experimentally when $r/a = 0.025$, $d_1/a = 0.97$ and $0.86 < d_2/a < 8.8$ [5]. Substituting in (6.30) we find that

$$C_1 = \frac{a\varepsilon_0}{d_2(d_1 + D(d_1 + d_2))}. \tag{6.35}$$

The potential at any point within a planar triode may be calculated using (6.23), (6.24) and (6.29). The potential contours for a typical geometry are shown in Figure 6.4 for $V_a = 100\text{V}$ and a range of grid voltages (see Worksheet 6.1). The triode is close to cut off when $V_g = -12$ V. The electric field on the surface of the cathode is independent of position in every case. The equipotential surfaces close to the grid wires are approximately circles centred on the line charges representing the grid. Thus the model represents the potentials correctly if r/a is small. However, as the radius of the equipotential surface increases, it is found that it is no longer centred on the line charge. A number of authors have presented formulae for calculating the penetration factor for larger screening factors. These formulae are based on the addition of line, dipole, and multipole, charges to ensure that the surfaces of the grids wire are
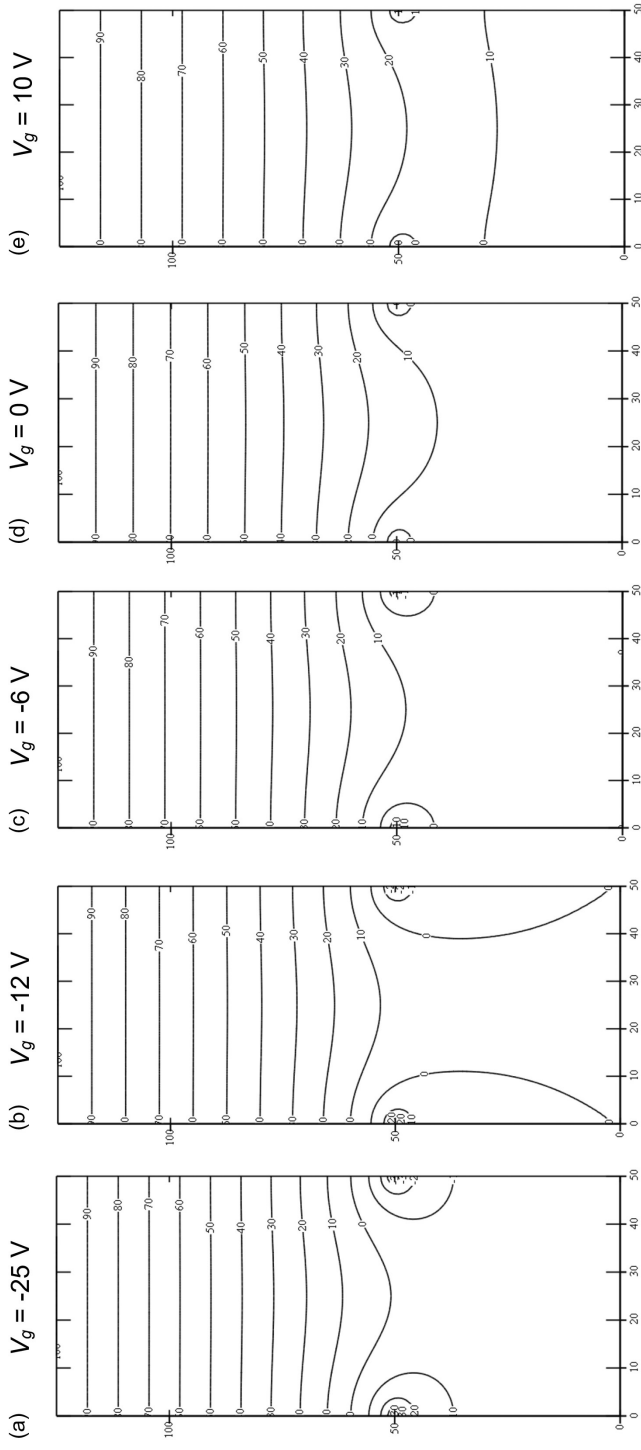
Figure 6.4: Equipotential contours in a planar triode for various grid voltages and $V_a = 100\,\text{V}$ ($d_1/a = 1$, $d_2/a = 1.5$, $r/a = 0.05$, $S = 0.1$) (copyright 1948, McGraw-Hill Education, reproduced, with permission, from [8]).

equipotential surfaces. Equation (6.34) is valid if $r/a \le 0.05$. For $r/a \le 0.08$ it has been shown [6] that

$$D = \frac{\ln\left(\coth\left(\dfrac{2\pi r}{a}\right)\right)}{\dfrac{2\pi d_2}{a} - \ln\left(\cosh\left(\dfrac{2\pi r}{a}\right)\right)}. \tag{6.36}$$

We note that $d_2 D$ is not now a function of r and a only because the diameters of the grid wires are no longer small compared with the other dimensions. Further formulae which are valid up to $r/a = 0.16$ are given in [7, 8]. These expressions are somewhat complicated and do not give additional insight into the properties of triodes. In addition, they are of limited relevance to modern high power triodes in which the grid 'wires' are usually rectangular in cross-section.

6.3.2   A Triode with Non-Uniform Electric Field on the Cathode

When $d_1/a < 1$ the x component of the electric field is found by differentiating (6.23)

$$E_z(x, y) = -\frac{\partial V}{\partial x} = -\frac{1}{a\varepsilon_0}\left(f_g{}'(x,y)q_g + q_a\right), \tag{6.37}$$

where

$$f_g{}'(x, y) = \frac{1}{2}\frac{\sinh 2\pi(x + d_1)/a}{\cosh 2\pi(x + d_1)/a - \cos 2\pi y/a} - \frac{1}{2}\frac{\sinh 2\pi(x - d_1)/a}{\cosh 2\pi(x - d_1)/a - \cos 2\pi y/a}. \tag{6.38}$$

In the limit $x \to 0$

$$f_g{}'(0, y) = \frac{\sinh\left(2\pi d_1/a\right)}{\cosh\left(2\pi d_1/a\right) - \left(\cos 2\pi y/a\right)} \tag{6.39}$$

so that, on the surface of the cathode

$$E_z(0, y) = -\frac{1}{a\varepsilon_0}\left(f_g'(0, y)q_g + q_a\right). \tag{6.40}$$

Substituting for q_g and q_a from (6.29)

$$E_z(0, y) = -\frac{C_1}{a\varepsilon_0}\begin{bmatrix} f_g'(0, y) & 1 \end{bmatrix}\begin{bmatrix} d_1 + d_2 & -d_1 \\ -d_1 & f_g(d_1, r) \end{bmatrix}\begin{bmatrix} V_g \\ V_a \end{bmatrix}. \tag{6.41}$$

The electric field depends linearly on the grid and anode potentials and therefore, following Fremlin, we define the electrostatic penetration factor by

$$D_g(y) = \left(\frac{\partial E_z}{\partial V_a} \middle/ \frac{\partial E_z}{\partial V_g}\right)_{x=0}. \tag{6.42}$$

From (6.41)

$$\frac{\partial E_z}{\partial V_a} = -\frac{C_1}{a\varepsilon_0}\left(-d_1 f_g'(0,y) + f_g(d_1 r)\right).$$ (6.43)

Similarly

$$\frac{\partial E_c}{\partial V_g} = -\frac{C_1}{a\varepsilon_0}\left((d_1 + d_2) f_g'(0,y) - d_1\right)$$ (6.44)

so that

$$D_g(y) = \frac{f_g(d_1, r) - d_1 f_g'(0,y)}{(d_1 + d_2) f_g'(0,y) - d_1},$$ (6.45)

which is valid when $r/a \le 0.05$. In the limit $d_1 \gg a$ $f_g'(0,y) \rightarrow 1$ and $D_g(y) \rightarrow D$ as expected. It can be shown, by substitution, that

$$E_z(0,y) = -\frac{V_g + D_g(y)V_a}{d_1 + D_g(y)(d_1 + d_2)},$$ (6.46)

which was previously derived as (6.13).

Figure 6.5 shows how the normalised penetration factor $\left(D_g(y)/D\right)$ varies with y for a number of grid-cathode spacings. This figure confirms that the emission from
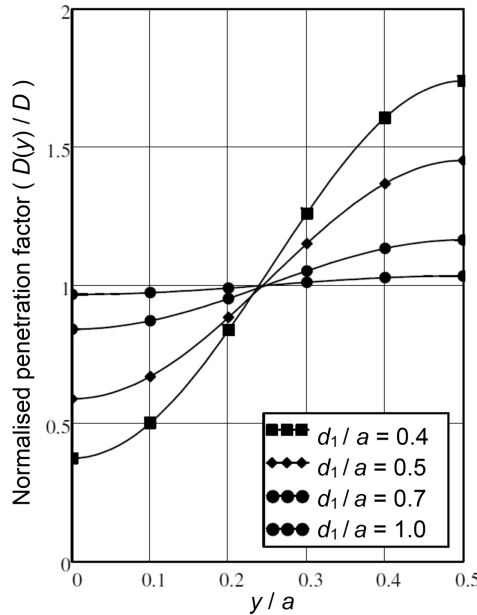
Figure 6.5: Variation of normalised penetration factor with position $(d_2/a = 5, \quad r/a = 1/14, \quad V_a = 100 \text{ V}, \quad V_g = -2 \text{ V}, \quad D = 0.026)$.

the cathode is uniform when $d_1/a \geq 1$ (see Worksheet 6.1). A different method for calculating the potential when $d_1/a \geq 0.4$ using theta functions gives essentially the same results [9].

6.3.3   Calculation of Penetration Factors Using Numerical Methods

We have seen that the analytical solutions for the penetration factor are restricted to grids composed of circular wires whose diameter is small compared with their spacing. For wires of other shapes and for general dimensions it is necessary to use numerical methods. It is sufficient to calculate the potentials in two cases:

i) $V_g = 1, \quad V_a = 0$
ii) $V_g = 0, \quad V_a = 1,$

since all other cases can be found by superposition. The penetration factor is then the ratio of the electric field at the surface of the cathode computed in (ii) to that computed in (i).

6.4   Static Characteristics of Triodes

The accurate calculation of the static characteristic curves of a triode requires the simultaneous solution of the equations of motion of the electrons, and Poisson's equation for the electrostatic potential in the presence of space charge. This is a difficult task, and a number of authors have offered expressions for the current based on equivalent diodes derived from the electrostatic solution. When the electric field at the cathode is uniform and given by equation (6.14) we can write

$$E_x(0, y) = \frac{V_e}{d_e}, \tag{6.47}$$

where the equivalent voltage is

$$V_e = A\left(V_g + DV_a\right), \tag{6.48}$$

and the spacing of the equivalent diode is

$$d_e = A\left(d_1 + D\left(d_1 + d_2\right)\right), \tag{6.49}$$

where A is a constant whose value is to be chosen so that the current density in the triode is equal to the space-charge-limited current in the equivalent diode. Thus

$$J = K\frac{V_e^{\frac{3}{2}}}{d_e^2} \tag{6.50}$$

or

$$J = \frac{K}{\sqrt{A}} \cdot \frac{\left(V_g + DV_a\right)^{\frac{3}{2}}}{\left(d_1 + D\left(d_1 + d_2\right)\right)^2}.$$ (6.51)

To find an expression for A we consider two limiting cases:

i) The distance between the anode and the grid is increased $\left(d_2/a \to \infty\right)$ keeping all other parameters constant. Then

$$J \to \frac{K}{\sqrt{A}} \cdot \frac{V_g^{\frac{3}{2}}}{\left(d_1 - \frac{a}{2\pi}\ln\left[2\sin\left(\frac{\pi r}{a}\right)\right]\right)^2}$$ (6.52)

from (6.34) since $D \to 0$ as $1/d_2$. We note that the denominator in (6.52) is constant and slightly greater then d_1 reflecting the fact that there is some penetration of the electric field through the array of grid wires. If we now reduce the spacing between the grid wires, while keeping r/a constant, then the current density should tend to that of a diode in which there is a continuous electrode at the plane of the grid. This requires $\sqrt{A} \to 1$ as $D \to 0$.

ii) The normalised grid wire radius r/a is small, so that the current intercepted by the grid is negligible. Then, from (6.34), $D \to \infty$ as $r \to 0$ and

$$J \to \frac{K}{\sqrt{AD}} \cdot \frac{V_a^{\frac{3}{2}}}{\left(d_1 + d_2\right)^2}.$$ (6.53)

This must be the current density in a diode with voltage V_a and spacing $\left(d_1 + d_2\right)$ which requires $\sqrt{AD} \to 1$ as $D \to \infty$.

These two limits are satisfied correctly if $A = 1/(1 + D)$. Then

$$V_e = \frac{V_g + DV_a}{1 + D}$$ (6.54)

and

$$d_e = \frac{d_1 + D\left(d_1 + d_2\right)}{1 + D} = d_1 + \frac{Dd_2}{1 + D}.$$ (6.55)

It can be seen from (6.55) that this definition of the equivalent diode places the equivalent anode at a plane whose distance from the grid is independent of d_1 [10]. Then the current density in the triode is given by

$$J = K\sqrt{1 + D} \cdot \frac{\left(V_g + DV_a\right)^{\frac{3}{2}}}{\left(d_1 + D\left(d_1 + d_2\right)\right)^2}.$$ (6.56)

This equation was proposed by Walker [11] but does not appear to have been generally adopted.

Two other definitions of the equivalent diode appear in the literature:

- The equivalent voltage at the plane of the grid is given by (6.14) [9, 10] so that the current density is given by

$$J = \frac{K}{d_1^2}\left(\frac{V_g + DV_a}{1 + D(1 + d_2/d_1)}\right)^{\frac{3}{2}}. \tag{6.57}$$

- The numerator of (6.14) is taken to be the equivalent voltage, and the denominator to be the equivalent spacing [8, 9] with the result that

$$J = K\frac{\left(V_g + DV_a\right)^{1.5}}{\left(d_1 + D(d_1 + d_2)\right)^2}. \tag{6.58}$$

Neither of these expressions has the correct behaviour at both the limits discussed above. It has been argued that (6.57) should be modified so that the potential gradient close to the grid corresponds to that in a space-charge-limited diode [10, 12–14] to give

$$J = \frac{K}{d_1^2}\left(\frac{V_g + DV_a}{1 + D(1 + (4d_2)/(3d_1))}\right)^{\frac{3}{2}}. \tag{6.59}$$

The adoption of this expression by industry suggests that it has been found useful in tube design [15]. However, it fails to tend to the correct limit as $D \to \infty$.

The equivalent diode approach has been criticised on the grounds that the assumptions made are somewhat arbitrary [5]. As an alternative, it was proposed that the perveance of the triode should be determined by setting V_g to the potential which would exist in the resulting space-charge-limited diode if the grid were removed [5, 16] with the result that

$$J = \frac{K}{d_1^2}\left(\frac{V_g + DV_a}{1 + D(1 + d_2/d_1)^{\frac{4}{3}}}\right)^{\frac{3}{2}}. \tag{6.60}$$

It appears, therefore, that (6.56) is the only equation that can be fully justified. However (6.59) has been widely used, presumably because it was found to give useful agreement with experimental results. It has been shown that (6.60) agrees with experiment for a limited range of cases. Figure 6.6 shows a comparison between the values of Jd_1^2/K calculated from these three equations, normalised to the value when $D \to 0$. For values of D which are of practical importance there is good agreement between (6.56) and (6.59). As a further test Figure 6.7 shows a comparison between the experimental points from Figure 7 in [5], and a curve calculated using
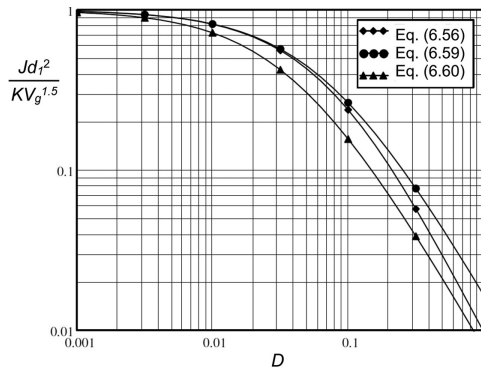
Figure 6.6: Comparison between normalised values of perveance calculated from (6.56), (6.59) and (6.60) when $d_2/d_1 = 10$.
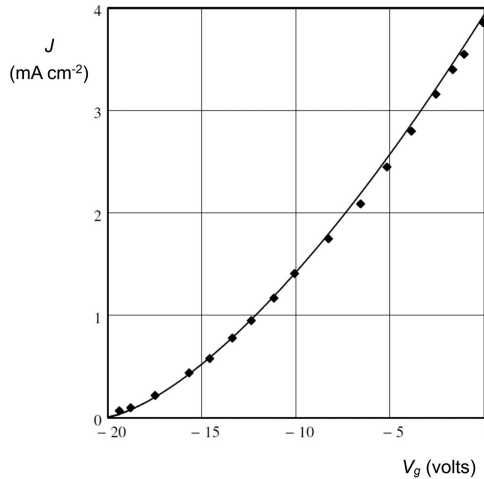
Figure 6.7: Variation of current density with grid potential at an anode potential of 200 volts. Experimental points from Figure 7 of [5] and curve calculated using (6.56). $d_1 = 1.7\,\text{mm}$, $d_2 = 5.03\,\text{mm}$, $r = 0.0445\,\text{mm}$, $a = 1.75\,\text{mm}$.

(6.56) (see Worksheet 6.2). There is excellent agreement between the theoretical and experimental curves confirming that (6.56) is correct.

The effective diode voltage may be affected by the thermal velocities of the electrons, and by the electrode contact potentials. Fremlin included a correction of about 1.2 V for thermal velocities when comparing his calculated curve with experimental results. In this case the effect of thermal velocities, calculated using the methods described in Chapter 3, can be represented by adding 0.1 V to the equivalent voltage. It therefore appears that the correction applied by Fremlin was too big. The contact potential depends on the materials from which the electrodes are made and upon surface contamination, and must therefore be determined experimentally. The potential may be positive or negative but is usually less than 0.5 V [11]. Both

these effects are small compared with the voltages used in large power tubes, and they may normally be neglected.

6.4.1   Grid Current

When the grid potential is positive with respect to the cathode then some of the primary electrons are collected by the grid, so reducing the anode current. The situation is complicated by the possibility of secondary electron emission, and thermionic emission, from the grid. Thus any calculations based on primary grid current can only be approximate. However, as the energies of secondary and thermionic electrons are small, the primary current is useful in calculating the power deposited on the grid. This problem is discussed in [8, 11, 17].

   If the effects of space-charge are ignored, then the trajectories of non-relativistic electrons remain unchanged when the ratio of the anode voltage to the grid voltage is constant. Since the trajectories depend on the velocities of the electrons, we expect that the ratio I_g/I_a is a function only of $\sqrt{V_g/V_a}$. It is found experimentally in many cases that when $V_a \geq V_g$ the ratio of the anode current to the grid current is given to useful accuracy by

$$\frac{I_a}{I_g} = \delta \sqrt{\frac{V_a}{V_g}}, \qquad (6.61)$$

where δ is the current division factor which is a constant of the tube [8]. The value of δ can be found by considering the electron trajectories when $V_a = V_g$. The trajectory which just grazes the surface of the grid provides a boundary between those electrons which are collected by the grid, and those which are collected by the anode. The approximate primary current division factor is [17]

$$\delta = \frac{a}{\dfrac{a}{\pi(1+D)} \cdot \dfrac{r}{2d_1} \ln\left(\dfrac{4ed_1}{r}\right) + 2r} - 1. \qquad (6.62)$$

An alternative approach is described in [11, 12]. An electron which just grazes the grid wire starts from the cathode at a distance x_1 from the centre line of a grid wire. The electron enters the deflecting field around the grid wire with energy given by an equivalent grid voltage obtained from (6.14) as

$$V_{eg} = -d_1 E_x(0), \qquad (6.63)$$

which can be expressed in terms of V_g/V_a for a given anode voltage. The angular momentum of the electron is conserved as it moves through the deflecting field so that

$$x_1 = r\sqrt{\frac{V_g}{V_{eg}}}. \qquad (6.64)$$

The ratio of the anode current to the grid current is then

$$\frac{I_a}{I_g} = \frac{a}{2r}\sqrt{\frac{V_{eg}}{V_g}} - 1. \tag{6.65}$$

This equation assumes that the grid-cathode spacing is large enough for the current density on the cathode surface to be uniform. The effects of space charge, and thermal velocities, are neglected. An improved expression, based on a more accurate representation of the field between the grid and the cathode, is [11]

$$\frac{I_a}{I_g} = \frac{a}{2r}\sqrt{\frac{V_{eg}}{V_g}}\left\{\frac{2V_{eg}\ln\left(a/2\pi r\right)}{2V_{eg}\ln\left(a/2\pi r\right)+\left(V_g-V_{eg}\right)}\right\} - 1. \tag{6.66}$$

A third way of calculating the grid current uses numerical integration of the equations of motion of the electrons in the electrostatic field of the triode. The effects of space charge are neglected. Figure 6.8 shows the trajectories corresponding to the potential maps in Figure 6.4. The grid current can be computed by finding the trajectory which just touches the edge of the grid (see Worksheet 6.1). Figure 6.9 shows a plot of I_g/I_a against V_g/V_a obtained in this way with curves given by (6.61), (6.65) and (6.66) for comparison. The computed grid current is zero for small positive grid voltages because the electrons are steered away from the grid by the field. It can be seen that (6.61) only gives a very rough guide to the magnitude of the grid current and that the assumption of a constant current division factor is not justified. The best agreement is given by (6.66).
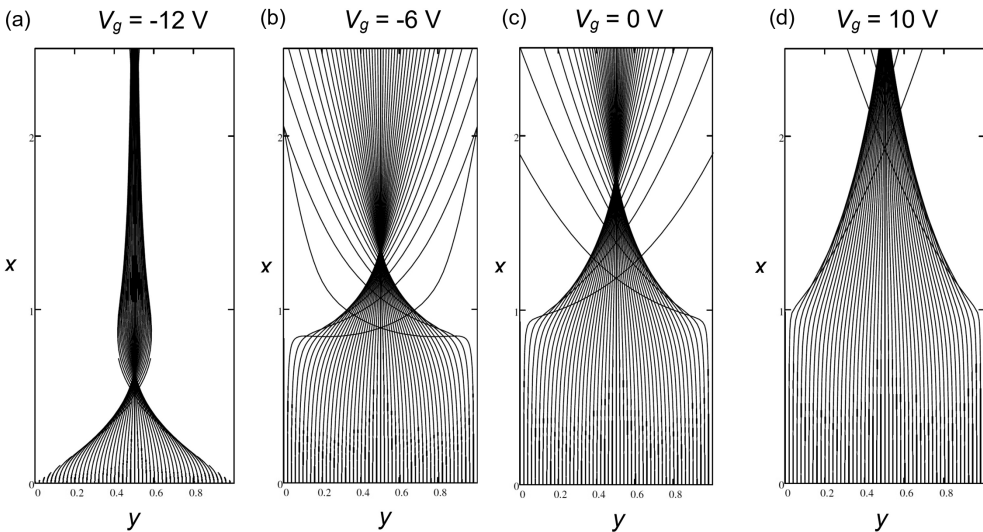
Figure 6.8: Electron trajectories without space charge corresponding to the potential distributions shown in Figure 6.4 $\left(d_1/a = 1, d_2/a = 1.5, r/a = 0.05, V_a = 100\,\text{V}\right)$.
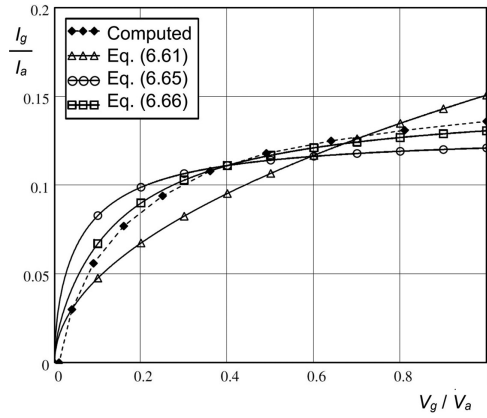
Figure 6.9: Dependence of grid current fraction on the normalised grid voltage $\left(d_1/a = 1, d_2/a = 1.5, r/a = 0.05, V_a = 100 \text{ V}\right)$.

6.4.2   Triodes with Island Formation

We saw in Section 6.3.2 that, when $r/a < 1$, the electric field at the surface of the cathode depends on the transverse position in the absence of electrons. In particular, as the grid is made more negative with respect to the cathode, the polarity of the electric field is reversed on the section of the cathode immediately opposite the grid. In that case only the part of the cathode on which the vacuum electric field is accelerating can emit electrons, the current density is non-uniform and parts of the cathode may cease to contribute to the current. This condition is known as island formation [5, 9, 18, 19].

The current in a triode, when there is island formation, can be calculated by summing the currents contributed by elementary strips of the cathode. From (6.46) and (6.56) the average current density is given by

$$J = \frac{K}{a} \int_0^a \sqrt{1 + D_g(y)} \cdot \frac{\left(V_g + D_g(y)V_a\right)^{\frac{3}{2}}}{\left(d_1 + D_g(y)(d_1 + d_2)\right)^2} \cdot dy. \tag{6.67}$$

where the integrand is zero if the effective voltage in the numerator is negative. The current is cut off when the effective voltage is less than or equal to zero for all values of y. It can be seen from Figure 6.5 that the penetration factor is greatest when $y = a/2$, thus the grid voltage at cut-off is given by

$$V_{g0} = -D_g(a/2)V_a. \tag{6.68}$$

The magnitude of this voltage increases as d_1 is reduced. As the grid voltage is increased a point is reached at which electrons are emitted from the whole cathode surface. For higher grid voltages the effects of non-uniform cathode

emission tend to average out and it may be expected that the current will obey the 3/2 power law. This can be revealed by plotting $J^{2/3}$ against the grid voltage, as shown in Figure 6.10, for a range of values of d_1 (see Worksheet 6.1). The effects of island formation are shown in the curvature when the grid voltage is negative. This curvature is not present when $d_1 = a$ and is barely perceptible, for the parameters illustrated, when $d_1 = 0.5a$. For smaller spacings, the grid voltage required to reduce the current to zero is increasingly negative and the rate of change of current with grid voltage is reduced. Some authors have suggested that, when there is island formation, the behaviour of the triode can still be represented by (6.56) if the power law is increased [9, 16]. This model is only appropriate when the grid voltage is negative because it does not represent the 3/2 power law behaviour when the grid voltage is positive. It is therefore preferable to regard the 3/2 power law as the norm and consider the effects of island formation as a modification of it for negative grid voltages and small electrode spacings. As the grid voltage increases all the curves tend to the straight lines calculated from (6.56). The agreement is exact for $d_1 \geq 0.5a$ and is a good approximation for $d_1 = 0.2a$. Figure 6.11 shows the variation of the current density across the cathode for a variety of grid voltages when $d_1 = 0.2a$. The current density is normally far from uniform. It is least under the grid wires for negative grid voltages, and greatest under the grid wires for positive grid voltages. This non-uniform emission means that the grid current is increased. There is also a possibility that the emission saturates in regions of high current density, so that the electron flow ceases to be space-charge limited everywhere. Saturation would be revealed, in a plot like that in Figure 6.10, by current densities below the asymptotic straight line at high positive grid voltages.
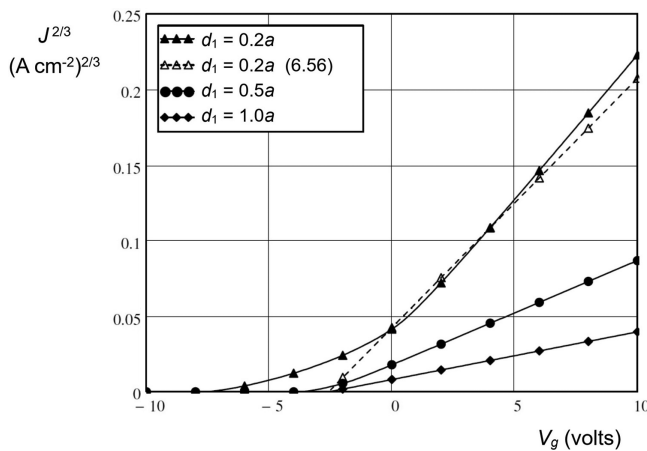
Figure 6.10: Variation of cathode current density to the power 2/3 with grid voltage showing the effects of island formation in triodes with a range of values of d_1 $(a = 1\ \text{mm}, r = 0.05\ \text{mm}, d_2 = 5\ \text{mm}, V_a = 100\ \text{V})$.
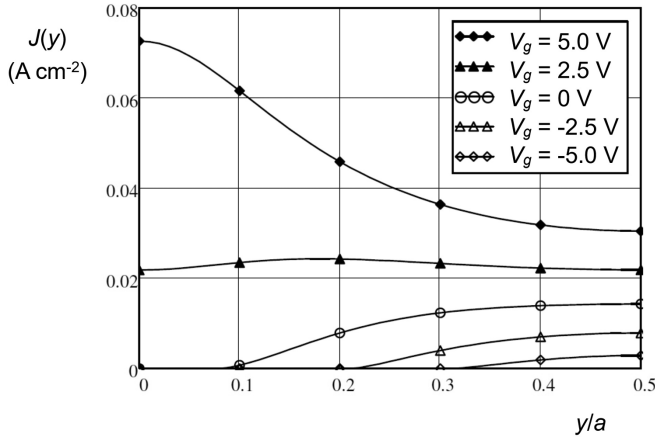
Figure 6.11: Variation of current density across the cathode surface for a range of grid voltages showing island formation $\left(a = 1\,\text{mm}, r = 0.05\,\text{mm}, d_1 = 0.2\,\text{mm}, d_2 = 5\,\text{mm}, V_a = 100\,\text{V}\right)$.

6.5   Electrostatic Models of Tetrodes

Figure 6.12 shows the arrangement of a planar tetrode. It has been assumed, for convenience, that the spacings of both grids are the same. The electrostatic analysis of the tetrode proceeds in exactly the same manner as that for the triode. By analogy with (6.7) the electric field on the surface of the cathode is linearly dependent on the potentials of the electrodes with respect to the cathode

$$E_x\left(0, y\right) = -\frac{V_g}{d_g\left(y\right)} - \frac{V_s}{d_s\left(y\right)} - \frac{V_a}{d_a\left(y\right)}. \tag{6.69}$$

This equation can be re-written as

$$E_x\left(0, y\right) = -\frac{V_g + D_1 V_s + D_2 V_a}{d_g}, \tag{6.70}$$

where the dependence on y has been omitted for simplicity, and

$$D_1\left(y\right) = \frac{d_g}{d_s} \quad \text{and} \quad D_2\left(y\right) = \frac{d_g}{d_a}. \tag{6.71}$$

Now, assume that the grids are thin and their potentials are set to

$$V_g = \frac{d_1}{d_1 + d_2 + d_3} V_a \quad \text{and} \quad V_s = \frac{d_1 + d_2}{d_1 + d_2 + d_3} V_a. \tag{6.72}$$

Then, substituting for V_g and V_s and noting that $E_x\left(0, y\right) = -V_a/\left(d_1 + d_2 + d_3\right)$, we obtain

$$d_g\left(y\right) = d_1 + D_1\left(d_1 + d_2\right) + D_2\left(d_1 + d_2 + d_3\right) \tag{6.73}$$
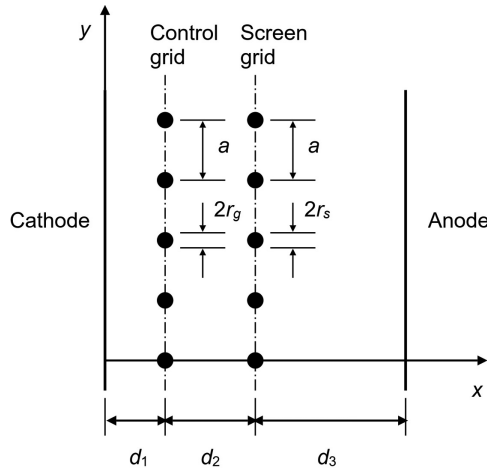
Figure 6.12: Arrangement of a planar tetrode.

so that, substituting in (6.70),

$$E_x(0, y) = -\frac{V_g + D_1 V_s + D_2 V_a}{d_1 + D_1 (d_1 + d_2) + D_2 (d_1 + d_2 + d_3)}.$$ (6.74)

6.6   Penetration Factors in Tetrodes

To find the penetration factors in a planar tetrode, we use the principle of super-position to add the potential of the screen grid to the expression for the potential previously developed for the triode [20]. The potential within the tetrode is then, from (6.23)

$$V(x, y) = \frac{1}{a\varepsilon_0} \left(f_g(x, y) q_g + f_s(x, y) q_s + x q_a \right),$$ (6.75)

where q_s is the charge per unit length on a screen grid wire, $f_g(x, y)$ is defined by equation (6.24), the equivalent expression for the screen grid is

$$f_s(x, y) = \frac{a}{4\pi} \ln \left[\frac{\cosh 2\pi (x + d_1 + d_2)/a - \cos 2\pi y/a}{\cosh 2\pi (x - d_1 - d_2)/a - \cos 2\pi y/a} \right],$$ (6.76)

and it has been assumed that both grids have the same spacing. We will also assume, for simplicity, that the distances between the electrodes are all greater than or equal to a so that the potential at the plane of each electrode is independent of y. Then

$$D_s = -\frac{a}{2\pi d_3} \ln \left[2 \sin \left(\frac{\pi r_s}{a} \right) \right]$$ (6.77)

$$V_g = \frac{1}{a\varepsilon_0}\left(f_g\left(d_1,r_g\right)q_g + d_1 q_s + d_1 q_a\right)$$

$$V_s = \frac{1}{a\varepsilon_0}\left(d_1 q_g + f_s\left(d_1+d_2,r_s\right)q_s + \left(d_1+d_2\right)q_a\right) \tag{6.78}$$

$$V_a = \frac{1}{a\varepsilon_0}\left(d_1 q_g + \left(d_1+d_2\right)q_s + \left(d_1+d_2+d_3\right)q_a\right),$$

where, from (6.33)

$$f_g\left(d_1,r_g\right) = d_1 + d_2 D_g, \tag{6.79}$$

and

$$f_s\left(d_1+d_2,r_s\right) = d_1 + d_2 + d_3 D_s, \tag{6.80}$$

with

$$D_g = -\frac{a}{2\pi d_2}\ln\left[2\sin\left(\frac{\pi r_g}{a}\right)\right]. \tag{6.81}$$

Substituting in (6.78) gives

$$\begin{bmatrix} V_g \\ V_s \\ V_a \end{bmatrix} = \frac{1}{a\varepsilon_0}\begin{bmatrix} d_1 + d_2 D_g & d_1 & d_1 \\ d_1 & d_1 + d_2 + d_3 D_s & d_1 + d_2 \\ d_1 & d_1 + d_2 & d_1 + d_2 + d_3 \end{bmatrix}\begin{bmatrix} q_g \\ q_s \\ q_a \end{bmatrix}. \tag{6.82}$$

This equation can be inverted to give

$$\begin{bmatrix} q_g \\ q_s \\ q_a \end{bmatrix} = C_2\begin{bmatrix} d_3\left(\begin{array}{c}d_1 + d_2 + \\ D_s\left(d_1+d_2+d_3\right)\end{array}\right) & -d_1 d_3 & -d_1 d_3 D_s \\ -d_1 d_3 & d_1\left(d_2+d_3\right)+d_2 D_g\left(d_1+d_2+d_3\right) & -d_2\left(d_1 + D_g\left(d_1+d_2\right)\right) \\ -d_1 d_3 D_s & -d_2\left(d_1+D_g\left(d_1+d_2\right)\right) & d_2\left[d_1+D_g\left(d_1+d_2+d_3 D_s\right)\right] \end{bmatrix}\begin{bmatrix} V_g \\ V_s \\ V_a \end{bmatrix}$$

$$\tag{6.83}$$

where

$$C_2 = \frac{a\varepsilon_0}{d_3\left[d_1\left(d_2 + D_s\left(d_2+d_3\right)\right)+d_2 D_g\left(d_1+d_2+D_s\left(d_1+d_2+d_3\right)\right)\right]}. \tag{6.84}$$

Now the electric field at the surface of the cathode is given by

$$E_x\left(0,y\right) = -\frac{\left(q_g + q_s + q_a\right)}{a\varepsilon_0} \tag{6.85}$$

so that

$$E_x(0,y) = \frac{\left(d_2 + D_s\left(d_2 + d_3\right)\right)V_g + d_2 D_g\left(V_s + D_s V_a\right)}{d_1\left(d_2 + D_s\left(d_2 + d_3\right)\right) + d_2 D_g\left(d_1 + d_2 + D_s\left(d_1 + d_2 + d_3\right)\right)},\qquad(6.86)$$

which can be rewritten as

$$E_x(0,y) = \frac{V_g + D_g'\left(V_s + D_s V_a\right)}{d_1 + D_g'\left(d_1 + d_2 + D_s\left(d_1 + d_2 + d_3\right)\right)},\qquad(6.87)$$

where

$$D_g' = \frac{D_g}{1 + D_s\left(1 + d_3/d_2\right)}.\qquad(6.88)$$

Comparing (6.87) with (6.74) we see that

$$D_1 = D_g' \quad \text{and} \quad D_2 = D_g' D_s.\qquad(6.89)$$

Equation (6.87) can also be written

$$E_x(0,y) = -\frac{V_g + D_g V_e}{d_1 + D_g d_e}\qquad(6.90)$$

in which the tetrode has been replaced by an equivalent triode whose anode potential is

$$V_e = \frac{V_s + D_s V_a}{1 + D_s\left(1 + d_3/d_2\right)}\qquad(6.91)$$

and the position of the anode is

$$d_e = \frac{d_1 + d_2 + D_s\left(d_1 + d_2 + d_3\right)}{1 + D_s\left(1 + d_3/d_2\right)}.\qquad(6.92)$$

This is similar, but not identical, to the expression for reducing a triode to an equivalent diode (see equation (6.55).

6.7   Static Characteristics of Tetrodes

The static characteristics of a tetrode may be calculated from the equivalent diode in exactly the same way as those of triodes. Thus from (6.90) we obtain

$$J = K\sqrt{1 + D_g} \cdot \frac{\left(V_g + D_g V_e\right)^{\frac{3}{2}}}{\left(d_1 + D_g d_e\right)^2}\qquad(6.93)$$

by analogy with (6.56).

Island formation occurs when $d_1/a < 1$. Since, normally, $d_2/a \geq 1$ the variation of the vacuum electric field on the cathode surface depends only upon the control grid. Then the characteristics of the tetrode can be determined from the equivalent triode using (6.45) and (6.46) with

$$D_E(y) = \frac{d_1\left(1 - f_g'(0,y)\right) + d_e D_g}{(d_1 + d_e)f_g'(0,y) - d_1},$$

(6.94)

where $f_g'(0,y)$ is given by (6.39) and

$$E_z(0,y) = -\frac{V_g + D_E(y)V_e}{d_1 + D_E(y)d_e}.$$

(6.95)

Then the characteristic curves can be computed using

$$J = \frac{K}{a}\int_0^a \sqrt{1 + D_E(y)} \cdot \frac{\left(V_g + D_E(y)V_e\right)^{\frac{3}{2}}}{\left(d_1 + D_E(y)d_e\right)^2} \cdot dy.$$

(6.96)

6.7.1   Grid Currents in Tetrodes

The current intercepted by the control grid in a tetrode may be estimated in exactly the same way as for a triode. The current is zero if the grid is negative and increases with increasingly positive voltages. Theoretically, the screen grid should not intercept any electrons because the wires are aligned with those of the control grid. Hence the fields should focus the electron trajectories so that they pass between the screen grid wires [21]. However, at high cathode current and low anode voltage, the accumulation of space charge between the screen grid and the anode causes the electron trajectories to move outwards so that they may be intercepted by the wires of the screen grid. There is no simple way of estimating the current intercepted by the screen grid but it is commonly a small fraction of the cathode current except when the anode voltage is low and the anode current is high. For a model of a planar tetrode see Worksheet (6.3).

6.7.2   Effect of Space-Charge between the Screen Grid and the Anode

The tetrode behaves approximately like an electron source whose current is controlled by the control grid voltage and whose energy is controlled by the screen grid voltage. Thus the flow of current in the space between the screen grid and the anode approximates to that in a diode with injected current as described in Section 5.8. The potential of the anode must always be greater than that of the screen grid so that any secondary electrons emitted by the anode are returned to the anode. Therefore we are only interested in conditions in which there is a monotonic increase in the potential between the screen grid and the anode, or where

there is a potential minimum which is not deep enough to form a virtual cathode. Using equations (5.96) and (5.100), and recalling that the potentials in Section 5.8 are referred to the potential at the injection plane (in this case the screen grid), the normalised spacing between the screen grid and the anode is

$$X_a = \sqrt{\frac{-4J_a}{\varepsilon_0 \sqrt{2e/m_0} V_s^{1.5}}} \cdot d_3,$$

(6.97)

where V_s is the potential of the screen grid with respect to the cathode. The normalised anode potential is

$$W_a = \sqrt{\frac{V_a}{V_s}}.$$

(6.98)

The parameter α_1 can be then found by numerical solution of the equation

$$X_a = \frac{4}{3}(W_a - 2\alpha_1)(W_a + \alpha_1)^{0.5} \mp \frac{4}{3}(1 - 2\alpha_1)(1 + \alpha_1)^{0.5},$$

(6.99)

derived from (5.104) and (5.105) by eliminating α_2, where the negative sign is taken if there is no potential minimum between the screen grid and the anode. Then, from (3.90), the normalised potential gradient at the screen grid is

$$\frac{dU}{dX} = \pm(1 + \alpha_1)^{0.5}.$$

(6.100)

We assume that this potential gradient is produced by an equivalent vacuum anode voltage V_a' so that

$$\frac{dU}{dX} = \left(\frac{V_a' - V_s}{V_s}\right) \cdot \frac{1}{X_a}.$$

(6.101)

Since the effective anode voltage has been reduced, it is necessary for the control grid voltage to be increased to maintain constant anode current by

$$\Delta V_g = D_g' D_s (V_a - V_a').$$

(6.102)

As the space charge increases, the electric field at the screen grid is reduced, V_a' becomes smaller and, eventually, becomes negative as a potential minimum is formed. Thus this model predicts that the effects of space-charge are to increase the control grid voltage required to produce a given current. The space-charge effects are greatest when the current is maximum and the anode voltage is least, that is when $W_a = 1$. Then, taking the positive sign in (6.99) we find that the maximum possible value of X_a is 3.77 when $\alpha_1 = -1/2$ [22]. Thus, once the screen grid voltage and the maximum current have been specified there is a maximum permissible value of the spacing between the screen grid and the anode beyond which the current is space-charge limited. It is shown in Chapter 12 that this simple model is in qualitative agreement with the behaviour of real tetrodes. Equation (6.97) shows that

the effect of increasing the spacing between the screen grid and the anode is similar to that of increasing the injected current. Thus the space-charge effects can be reduced by decreasing that spacing. Equation (6.98) shows that the effects of space-charge at a given anode voltage are increased as the screen grid voltage is increased. However, the range of screen grid voltages over which a tube is operated is usually small, so this is less important.

References

[1] B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*. London: Oxford University Press, 1957.

[2] EIMAC, *Care and Feeding of Power Grid Tubes*, 5th ed. San Carlos, CA: CPI Inc. Eimac Division, 2003.

[3] S. G. McNees, 'Two megawatts RF power tetrode', *IEEE Transactions on Nuclear Science,* vol. NS-20, pp. 422–423, 1973.

[4] J. C. Maxwell, *Electricity and Magnetism*, 3rd ed., vol. 1. Oxford: Clarendon Press, 1892.

[5] J. H. Fremlin, 'LXVIII. Calculation of triode constants', *Philosophical Magazine Series 7,* vol. 27, pp. 709–741, 1939.

[6] F. B. Vogdes and F. R. Elder, 'Formulas for the amplification constant for three-element tubes in which the diameter of grid wires is large compared to the spacing', *Physical Review,* vol. 24, p. 683, 1924.

[7] F. Ollendorf, 'Berechnung des Durchgriffes durch enger Steggitter', *Elektotech. und Machinenbau,* vol. 21, pp. 59–64, 1934.

[8] K. R. Spangenberg, *Vacuum Tubes*. New York: McGraw-Hill, 1948.

[9] W. Bennett and L. Peterson, 'The electrostatic field in vacuum tubes with arbitrarily spaced elements', *Bell System Technical Journal,* vol. 28, pp. 303–314, 1949.

[10] B. J. Thompson, 'Space-current flow in vacuum-tube structures', *Proceedings of the IRE,* vol. 31, pp. 485–491, 1943.

[11] A. H. W. Beck, *Thermionic Valves: Their Theory and Design*. Cambridge: Cambridge University Press, 1953.

[12] B. D. H. Tellegen, 'The calculation of the emitted current in a triode', *Physica,* vol. 5e, pp. 301–315, 1925.

[13] W. Dahlke, 'Gittereffektivpotential un Kathodenstromdichte einer ebenen Triode unter Berucksichtigung der Inselbildung', *Telefunken Zeitung,* vol. 24, pp. 213–222, 1951.

[14] S. Magdo, 'The optimum grid to cathode spacing in planar triodes', *IEEE Transactions on Electron Devices,* vol. 12, pp. 273–280, 1965.

[15] L. C. Scholz, 'Calculation of fields and currents', in R. S. Burnap, ed., *Electron Tube Design*. Harrison, NJ: Radio Corporation of America, pp. 159–201, 1962.

[16] J. W. Gewartowski and H. A. Watson, *Principles of Electron Tubes*. Princeton, NJ: D. van Nostrand Co. Inc., 1965.

[17] K. R. Spangenberg, 'Current division in plane-electrode triodes', *Proceedings of the I.R.E.,* vol. 28, pp. 226–236, 1940.

[18] W. Dahlke, 'Gittereffektivpotential un Kathodenstromdichte einer ebenen Triode unter Berucksichtigung der Inselbildung', *Telefunken Zeitung,* vol. 24, pp. 213–222, 1951.

[19] W. Dahlke, 'Statische Kennwerte der nichtidealen Triode', *Telefunken Zeitung,* vol. 27, pp. 172–186, September 1954.

[20] S. Deb, 'Amplification factors and mutual conductance of a beam power valve', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 102, pp. 469–474, 1955.

[21] C. S. Bull, 'The alignment of grids in thermionic valves', *Journal of the Institution of Electrical Engineers – Part III: Radio and Communication Engineering,* vol. 92, pp. 86–92, 1945.

[22] G. Jaffe, 'On the currents carried by electrons of uniform initial velocity', *Physical Review,* vol. 65, pp. 91–98, 1944.

7   Linear Electron Beams

7.1   Introduction

In many vacuum electron devices the interaction between the electrons and RF electric fields takes place over an extended linear region. In order for this to be possible it is necessary to control the paths of the electrons to ensure that they are held in the correct relationship with the RF fields. This means that the electrons must be formed into an electron beam having a cross-section which is approximately constant. The charge density in the electron beams used in microwave tubes is high enough to cause them to spread out rapidly because of the mutual repulsion of the electrons. If this were not controlled in some way, the electrons would be intercepted on the tube body. They would then be lost to the interaction, and would cause damaging dissipation of heat.

The greater part of this chapter deals with methods of controlling the cylindrical electron beams used in the majority of microwave tubes. The effects of external electric and magnetic fields are considered both with, and without, space-charge forces. The properties of simple electric and magnetic lenses are discussed. There is an extensive treatment of the properties of cylindrical electron beams in the presence of a uniform axial magnetic field (solenoid focusing). The conditions for stable electron flow are discussed together with periodic variations of the beam radius (scalloping), and the changes in beam radius produced by modulation of the charge density (beam stiffness). This is followed by consideration of the spreading of electron beams, in the absence of a magnetic field, caused by space-charge, and centrifugal, forces. Next it is shown that stable electron flow can be achieved using periodic permanent magnets (PPM focusing), and periodic electrostatic fields. A brief summary is given of the properties of sheet beams and annular beams. Finally, there is a review of the imperfections introduced by thermal electron velocities and trapped ions.

7.2   Cylindrical Electron Beams

The majority of linear-beam tubes employ electron beams having circular symmetry about the axis of propagation. If the beam has radius b and is located concentrically within a cylindrical conducting drift tube of radius a as shown in Figure 7.1 then the electric field produced by the electron space charge is everywhere radial.
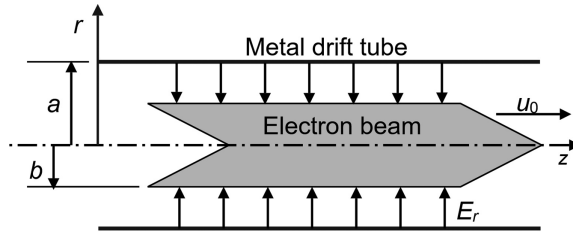
Figure 7.1: A cylindrical electron beam.

   If it is assumed that the charge density is uniform and the electrons are moving with uniform axial velocity u_0 then the current carried by the beam is

$$I_0 = \pi b^2 \rho u_0,$$ (7.1)

where ρ is the charge density, and u_0 represents the axial component of the electron velocity, including the effects of relativity, unless otherwise stated. A simple application of Gauss' Theorem shows that the radial component of the electric field, when $r \leq b$, is

$$E_r = \frac{\rho}{2\varepsilon_0} r = \frac{I_0}{2\pi b^2 \varepsilon_0 u_0} r$$ (7.2)

and, when $r > b$

$$E_r = \frac{I_0}{2\pi \varepsilon_0 u_0} \cdot \frac{1}{r}.$$ (7.3)

The electrostatic potential within these two regions is found by integrating (7.2) and (7.3) with respect to r to give

$$V = -\frac{I_0}{4\pi b^2 \varepsilon_0 u_0} r^2 + C_1$$ (7.4)

for $r \leq b$, and

$$V = -\frac{I_0}{2\pi \varepsilon_0 u_0} \ln r + C_2$$ (7.5)

for $r > b$, where C_1 and C_2 are constants. If the potential is zero when $r = a$ then (7.5) becomes

$$V = -\frac{I_0}{2\pi \varepsilon_0 u_0} \ln(r/a).$$ (7.6)

Equating the values of the potential at $r = b$ gives

$$C_1 = \frac{I_0}{4\pi \varepsilon_0 u_0} \left[1 - 2\ln(b/a) \right]$$ (7.7)

so that, within the beam

$$V = \frac{I_0}{2\pi \varepsilon_0 u_0} \left[1 - 2\ln(b/a) - \frac{r^2}{b^2} \right].$$ (7.8)

Since $b < a$ the logarithm is negative and the term within the brackets is always positive. The conventional current in the direction of electron flow is negative and so the potential within the beam is negative with respect to the surrounding drift tube. The properties of electron beams can be calculated using Worksheet 7.1.

The effect of this space-charge potential depression is best illustrated by an example. Suppose that $V_a = 10$ kV, $I_0 = -1$ A, $b = 5$ mm and $a = 10$ mm. Then the electron velocity at zero potential is 58.5×10^6 m s^{-1}. Substitution of these figures into (7.8) shows that the potential is -213 V with respect to the drift tube at the beam edge, and -367 V on the axis. Strictly speaking, the small change in the electron velocity produced by this difference in potential should be used to recalculate the potential depression iteratively until mutually consistent values are found. In practice it is either ignored, or used to make a small correction to the velocity of all the electrons. The situation is complicated by the effect of the rotation of the electrons about the axis caused by the magnetic focusing field which is discussed in Section 7.4. In practical tubes the tunnel is not continuous, but interrupted by a series of gaps. The effect of the gaps can be represented by an effective tunnel radius, if necessary [1].

7.3   Electron Optics without Space-Charge

Consider an electron which is moving in fields which are cylindrically symmetrical so that the tangential component of the electric field is zero. The equations of motion of an electron in cylindrical polar co-ordinates are from (1.6)

$$\frac{d}{dt}\left(\gamma m_0 \frac{dr}{dt}\right) = -e\left(E_r + r\dot{\theta}B_z - \dot{z}B_\theta\right) + r\dot{\theta}^2 \tag{7.9}$$

$$\frac{d}{dt}\left(\gamma m_0 r^2 \dot{\theta}\right) = er\left(\dot{r}B_z - \dot{z}B_r\right) \tag{7.10}$$

$$\frac{d}{dt}\left(\gamma m_0 \frac{dz}{dt}\right) = -eE_z + e\left(r\dot{\theta}B_r - \dot{r}B_\theta\right). \tag{7.11}$$

We note that, from (1.3) we can write

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}} = 1 + \frac{V}{V_R}, \tag{7.12}$$

where $V_R = m_0 c^2/e$ and V is the electrostatic potential relative to the cathode.[1]

---

[1] In developing the theory of vacuum tubes it is convenient take the cathode as the zero of electrostatic potential. However, it should be noted that, in practice, the body of the tube is normally at earth potential and the cathode is negative with respect to it.

7.3.1   The Paraxial Ray Equation of Electrostatic Electron Optics

Consider an electron which is moving, without rotation, in a cylindrically symmetric electric field [2]. From (7.9) the radial motion is determined by

$$\frac{d}{dt}\left(\gamma\frac{dr}{dt}\right)+\frac{eE_r}{m_0}=0. \tag{7.13}$$

Substituting for γ from (7.12) in (7.13) gives

$$\left(V+V_R\right)\frac{d^2r}{dt^2}+\frac{dV}{dt}\cdot\frac{dr}{dt}-\frac{eV_0}{m_0}\frac{\partial V}{\partial r}=0. \tag{7.14}$$

If the electron is always close to the axis, so that its radial velocity is small compared with its axial velocity, then, to a good approximation,

$$\frac{dz}{dt}=u, \tag{7.15}$$

where u is the electron velocity determined from (1.4). Then we may write

$$\frac{d}{dt}=u\frac{d}{dz} \tag{7.16}$$

and

$$\frac{d^2}{dt^2}=u^2\frac{d^2}{dz^2}+u\frac{\partial u}{\partial z}\cdot\frac{d}{dz} \tag{7.17}$$

so that (7.14) becomes

$$\left(V+V_R\right)u^2\frac{d^2r}{dz^2}+\left(V+V_R\right)u\frac{\partial u}{\partial z}\cdot\frac{dr}{dz}+u^2\frac{\partial V}{\partial z}\cdot\frac{dr}{dz}-c^2\cdot\frac{\partial V}{\partial r}=0. \tag{7.18}$$

If space charge can be neglected, then the electric field satisfies Laplace's equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right)+\frac{\partial^2 V}{\partial z^2}=0. \tag{7.19}$$

When the trajectory of an electron lies close to the axis (a paraxial trajectory) the potential does not differ much from the potential on the axis (V_A). Then (7.19) can be integrated with respect to r, to give the approximate relationship

$$\frac{\partial V}{\partial r}=-\frac{1}{2}r\frac{d^2V_A}{dz^2}. \tag{7.20}$$

Equation (1.4) can be written in the form

$$u^2=c^2\frac{V\left(V+2V_R\right)}{\left(V+V_R\right)^2} \tag{7.21}$$

and differentiated with respect to z to give

$$u\frac{du}{dz} = \frac{V_R^2 c^2}{(V+V_R)^3}.$$ (7.22)

Then, substituting into (7.18) from (7.20) and (7.22) we obtain

$$\frac{V_A(V_A+2V_R)}{(V_A+V_R)}\cdot\frac{d^2r}{dz^2} + \frac{\partial V_A}{\partial z}\cdot\frac{dr}{dz} + \frac{1}{2}\cdot\frac{\partial^2 V_A}{\partial z^2}r = 0.$$ (7.23)

This is the paraxial ray equation of electrostatic electron optics. Its application to electrostatic lenses is discussed in the next section.

7.3.2   Thin Electrostatic Lenses

Figure 7.2 shows a short transition region, defined by the planes 1 and 2, between two regions in which the axial electric field is constant. Since $\nabla\cdot\mathbf{E} = 0$ the radial component of the electric field must be non-zero in this region. When an electron enters the transition region from the left moving parallel to the axis at a radius r_1 its trajectory is perturbed. It then emerges on a trajectory that intersects the axis at a distance f_2 from the region. The electric field in the transition region therefore acts as a lens. If the lens is thin then the radial position of the trajectory can be assumed to be the same on entering and on leaving the lens $(r_1 = r_2 = r)$.

In order to derive an expression for the focal length of a thin electrostatic lens it is convenient to express (7.23) in the form

$$U\frac{d}{dz}\left(U\frac{dr}{dz}\right) = -\frac{1}{2}(V+V_R)\frac{d^2V}{dz^2}r,$$ (7.24)

where the subscript A has been dropped because the potential is understood to be defined on the axis relative to the potential of the cathode, and

$$U = \sqrt{V(V+2V_R)}.$$ (7.25)
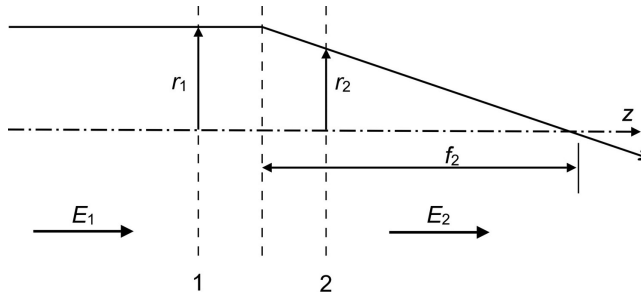
Figure 7.2: Geometry of an electrostatic electron lens.

Because the electric field is constant, except between the two planes, it follows that d^2V/dz^2 is zero outside the transition region. Integration of (7.24) between planes 1 and 2 produces

$$\left\{ U \frac{dr}{dz} \right\}_1^2 = -\frac{1}{2} \int_{z_1}^{z_2} \frac{V + V_R}{U} \cdot \frac{d^2V}{dz^2} r \, dz. \tag{7.26}$$

Now the gradient of the trajectory at plane 1 is zero and at plane 2 it is $-r/f_2$, so the focal length of the lens is given by

$$\frac{1}{f_2} = \frac{1}{2U_2} \int_{z_1}^{z_2} \frac{V + V_R}{U} \cdot \frac{d^2V}{dz^2} \, dz, \tag{7.27}$$

where U_2 is the value of U at plane 2. If $V(z)$ is known, the focal length can be calculated from (7.27). Integration of the right-hand side of this equation by parts yields

$$\frac{1}{f_2} = \frac{1}{2U_2} \left[\frac{V + V_0}{U} \cdot \frac{dV}{dz} \right]_{z_1}^{z_2} + \frac{1}{2U_2} \int_{z_1}^{z_2} \frac{V_0^2}{U^3} \left(\frac{dV}{dz} \right)^2 dz. \tag{7.28}$$

Equation (7.28) can be applied to the two simple lenses illustrated in Figure 7.3.

An Einzel lens (Figure 7.3(a)) has a central electrode at one potential within another electrode at a different potential (commonly zero). The electric field of the lens is confined between the planes 1 and 2, so that the gradient of the potential is zero at both planes, and the first term on the right-hand side of (7.28) is zero. Since the square of the electric field appears in the integral, the second term is positive for
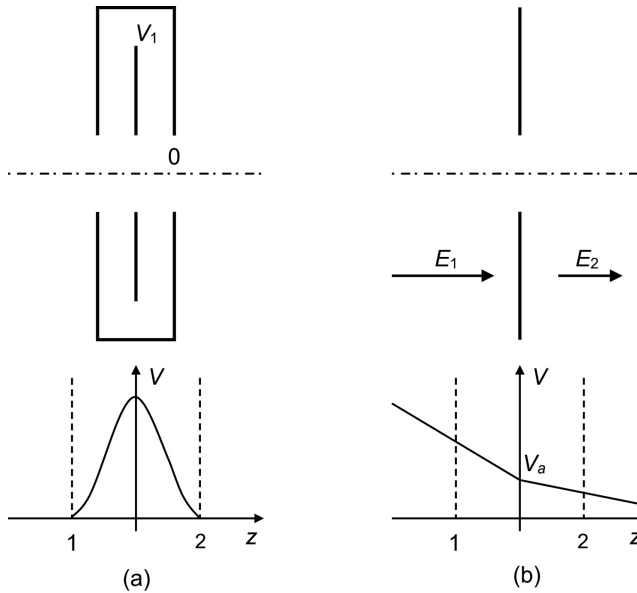
Figure 7.3: Electrostatic lenses: (a) an Einzel lens, and (b) an aperture lens.

all potentials of the inner electrode and this type of lens is always converging. The application of this to the theory of periodic electrostatic focusing of electron beams is discussed in Section 7.6.3.

In an aperture lens (Figure 7.3(b)) the initial and final fields are not zero, and it is found that the first term in (7.28) is commonly much larger than the second. The potential does not change very much through the lens so that $V_1 \approx V_2$. The focal length of the lens is given by

$$\frac{1}{f_2} \simeq \frac{V_a + V_R}{2V_a\left(V_a + 2V_R\right)}\left[\frac{dV}{dz}\right]_{z_1}^{z_2} = \frac{\left(V_a + V_R\right)\left(E_1 - E_2\right)}{2V_a\left(V_a + 2V_R\right)}, \tag{7.29}$$

where V_a is the potential of the aperture. The focal length may be positive or negative (converging or diverging) depending upon the signs and the magnitudes of the fields on the two sides of the aperture. The application of (7.29) in the theory of the Pierce electron gun is discussed in Section 9.2.1. Further information on electrostatic lenses can be found in [3].

7.3.3   Busch's Theorem

We now establish an important theorem relating to the motion of an electron through a cylindrically symmetrical magnetic field. The proof given here is based on [4] which is more general and rigorous than those published elsewhere. Consider an electron which is distant r from the axis. The magnetic flux linked to a circle of radius r is defined by

$$\Phi = 2\pi\int_0^r rB_z\,dr. \tag{7.30}$$

The rate of change of the flux linked to a circle which travels with the electron is found by differentiating of (7.30) with respect to time

$$\frac{d\Phi}{dt} = \frac{\partial\Phi}{\partial r}\dot{r} + \frac{\partial\Phi}{\partial z}\dot{z} = 2\pi\left(B_z r\dot{r} + \dot{z}\int_0^r r\frac{\partial B_z}{\partial z}\,dr\right). \tag{7.31}$$

Now the axial and radial components of the magnetic field are linked together by the requirement that $\nabla\cdot\mathbf{B} = 0$. In cylindrical polar coordinates this becomes

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rB_r\right) + \frac{\partial B_z}{\partial r} = 0 \tag{7.32}$$

so the integral in (7.31) can be written

$$\int_0^r r\frac{\partial B_z}{\partial z}\,dr = -\int_0^r \frac{\partial}{\partial r}\left(rB_r\right)dr = -rB_r \tag{7.33}$$

and (7.31) becomes

$$\frac{d\Phi}{dt} = 2\pi\left(r\dot{r}B_z - r\dot{z}B_r\right). \tag{7.34}$$

Comparing (7.34) with (7.10) we observe that the right hand sides are identical apart from constants and, therefore

$$\frac{d}{dt}\left(r^2\dot{\theta}\right) - \frac{e}{2\pi\gamma m_0}\frac{d\Phi}{dt} = 0, \tag{7.35}$$

where it is assumed that the electrostatic potential is constant, so that the relativistic factor γ is constant. When (7.35) is integrated with respect to time we obtain the important result that, for the motion of the electron,

$$r^2\dot{\theta} - \frac{e}{2\pi\gamma m_0}\Phi = \text{constant.} \tag{7.36}$$

Equation (7.36) is known as Busch's theorem. It is used to calculate the angular velocity of electrons as they move through an axially symmetric magnetic field that varies with axial position. It remains valid when the effects of space-charge are important.

7.3.4   Magnetostatic Electron Optics without Space Charge

The equation of radial motion for an electron moving in a region of constant electrostatic potential, and a magnetic field with cylindrical symmetry, is obtained from (7.9) as

$$\frac{d^2r}{dt^2} = r\dot{\theta}^2 - \frac{e}{\gamma m_0}r\dot{\theta}B_z. \tag{7.37}$$

It is assumed that the space charge forces, and the magnetic field of the electron current, are negligible. The relativistic factor γ is constant because the electrostatic potential is constant. Then, making use of (7.17) and, assuming paraxial motion, (7.37) becomes

$$u^2\frac{d^2r}{dz^2} = r\dot{\theta}^2 - \frac{e}{\gamma m_0}r\dot{\theta}B_z. \tag{7.38}$$

If the electron has originated from a cathode shielded from the magnetic field, then the constant in (7.36) is zero. Also, since the electron remains close to the axis, we can write $\Phi = \pi r^2 B_0$, where B_0 is the flux density on the axis, so that, from (7.36)

$$\dot{\theta} = \frac{eB_0}{2\gamma m_0} = \frac{\omega_L}{\gamma}, \tag{7.39}$$

where ω_L is known as the Larmor frequency [5]. The right hand side of (7.38) is

$$r\dot{\theta}^2 - \frac{e}{\gamma m_0} r\dot{\theta} B_z = -r\frac{\omega_L^2}{\gamma^2}.$$ (7.40)

Substitution in (7.38) yields the magnetostatic paraxial ray equation

$$\frac{d^2 r}{dz^2} + \frac{\beta_L^2}{\gamma^2} r = 0,$$ (7.41)

where $\beta_L = \omega_L/u$.

7.3.5   Thin Magnetic Lenses

Magnetic electron lenses commonly take the form shown in Figure 7.4 where the magnetic field generated by a coil, or a permanent magnet, is confined axially by a pair of iron pole pieces. The magnetic field is therefore zero outside the lens. If the lens is thin, r is constant within it, and the change of the slope of the trajectory is obtained by integrating (7.41) through the lens

$$\left.\frac{dr}{dz}\right|_1^2 = -r\left(\frac{e}{2\gamma m_0 u}\right)^2 \int_{z_1}^{z_2} B_0^2 \; dz.$$ (7.42)

The focal length of the lens is given by

$$\frac{1}{f_2} = \left(\frac{e}{2\gamma m_0 u}\right)^2 \int_{z_1}^{z_2} B_0^2 \; dz.$$ (7.43)

Since the integral is always positive, a magnetic lens is always converging. Further information on magnetic lenses can be found in [3].
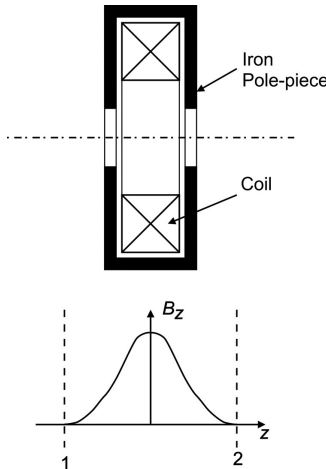
Figure 7.4: A magnetic electron lens.

7.4   Electron Optics with Space-Charge

An uncontrolled cylindrical electron beam spreads out radially under the influence of its space charge. In order to create a beam with sufficient length and power for use in a vacuum tube it is necessary to control it so that the tendency to spread is eliminated. This is usually achieved by using either a uniform axial magnetic field or a periodic array of converging magnetic lenses. An alternative which is rarely used is a periodic array of converging electrostatic lenses.

   In order to study the properties of electron beams under the influence of magnetic fields we make use of equation (7.9) for the radial motion of an electron on the edge of the beam. We will assume that the axial velocity of the electrons (u_0) is constant to a good approximation. The radial component of the electric field is given by (7.3) and the azimuthal component of the magnetic field caused by the beam current is

$$B_\theta = \frac{\mu_0 I_0}{2\pi r}. \tag{7.44}$$

Substituting in (7.9) gives

$$\frac{d^2 r}{dt^2} = -\frac{eI_0}{2\pi u_0 \varepsilon_0 \gamma^3 m_0} \cdot \frac{1}{r} + \frac{e\dot{\theta} B_z}{\gamma m_0} r + \frac{\dot{\theta}^2}{\gamma m_0} r, \tag{7.45}$$

where u_0 is the relativistically correct velocity. This can be written

$$\frac{d^2 r}{dt^2} = -\frac{\omega_p^2}{\gamma^3} \cdot \frac{b^2}{2r} + \frac{e\dot{\theta} B_z}{\gamma m_0} r + \frac{\dot{\theta}^2}{\gamma m_0} r, \tag{7.46}$$

where ω_p is the *electron plasma frequency* [6] defined when $r = b$ so that

$$\omega_p^2 = -\frac{e\rho}{\varepsilon_0 m_0} = -\frac{e}{\varepsilon_0 m_0} \cdot \frac{I_0}{\pi b^2 u_0}. \tag{7.47}$$

We will assume that the beam is launched with zero angular velocity and radius r_c from a cathode where the magnetic flux density is B_c. Then, by Busch's theorem (7.36), the angular velocity at a position where the radius is r, and the flux density is B_z, is given by

$$\dot{\theta} = \frac{e}{2\gamma m_0} \left(B_z - \frac{r_c^2}{r^2} B_c \right). \tag{7.48}$$

Substituting for the angular velocity in (7.46) gives

$$\frac{d^2 r}{dt^2} = \frac{1}{2} \frac{\omega_p^2}{\gamma^3} \frac{b^2}{r} + \left(\frac{eB_c}{2\gamma m_0} \frac{r_c^2}{r^2} \right)^2 r - \left(\frac{eB_z}{2\gamma m_0} \right)^2 r. \tag{7.49}$$

If the flux density $B_z = B_0 f(z)$ then (7.49) can be written

$$\frac{d^2 r}{dt^2} = \frac{1}{2} \frac{\omega_p^2}{\gamma^3} \frac{b^2}{r} + \frac{\omega_L^2}{\gamma^2} K \frac{b^4}{r^3} - \frac{\omega_L^2}{\gamma^2} f(z)^2 r, \tag{7.50}$$

where ω_L is the Larmor frequency defined in (7.39), and

$$K = \left(\frac{B_c r_c^2}{B_0 b^2} \right)^2.$$

(7.51)

Finally we use (7.17) to change the independent variable from t to z, and introduce the normalised radius $R = r/b$ so that

$$\frac{d^2 R}{dz^2} = \frac{\beta_p^2}{2\gamma^3} \cdot \frac{1}{R} + \frac{\beta_L^2}{\gamma^2} K \frac{1}{R^3} - \frac{\beta_L^2}{\gamma^2} f(z)^2 R,$$

(7.52)

where $\beta_p = \omega_p / u_0$ and $\beta_L = \omega_L / u_0$. This is the general differential equation governing the flow of a cylindrical electron beam through an axial magnetic field. The non-relativistic form is found by setting $\gamma = 1$ to give

$$\frac{d^2 R}{dz^2} = \frac{\beta_p^2}{2} \cdot \frac{1}{R} + \beta_L^2 K \frac{1}{R^3} - \beta_L^2 f(z)^2 R.$$

(7.53)

In the following sections the solutions of this equation are examined for a number of cases of interest. Non-relativistic analysis will be used because the effects of relativity can readily be included by adjusting the parameters β_p and β_L.

7.4.1   Solenoid Focusing

The simplest way to control a cylindrical electron beam is by using a uniform axial magnetic field (B_0). Such a field is usually produced by an electromagnet, though permanent magnets are sometimes used [7, 8]. The condition for smooth flow is that the radial electrostatic, magnetic and centrifugal forces should balance. Thus setting $R = 1$ and $f(z) = 1$ in (7.52)

$$\beta_p^2 = 2\gamma \beta_L^2 (1 - K).$$

(7.54)

Since β_p, β_L and K are all positive this equation can only be satisfied when $0 \le K \le 1$. The smallest magnetic field for which smooth flow is possible occurs when the magnetic flux linked to the cathode is zero and, therefore, $K = 0$. This condition is known as Brillouin flow [9]. The Brillouin field (B_B) is given by

$$\beta_L = \frac{1}{\sqrt{2\gamma}} \beta_p \quad \text{or} \quad B_B = \sqrt{\frac{2}{\gamma}} \cdot \frac{\omega_p}{e/m_0}.$$

(7.55)

If the magnetic field is m times the Brillouin field then the relationship between the Larmor frequency and the plasma frequency is

$$\omega_L = \frac{m}{\sqrt{2\gamma}} \omega_p.$$

(7.56)

Substituting for ω_L in (7.54) we find that the relationship between K and m can be written

$$K = 1 - \frac{1}{m^2}. \tag{7.57}$$

When $m \to \infty$ then $K \to 1$ so that the flux linked to the cathode is equal to that linked to the beam. This condition, which is known as *confined flow*, is useful as a theoretical concept but would normally require too strong a magnetic field to be useful in practice. Equation (7.57) is important for the design of linear electron beams because it is a necessary, though not sufficient, condition for achieving uniform electron flow. The injection of the beam into the focusing field is discussed in Chapter 9. A detailed discussion of the magnetic focusing of relativistic beams is given in [10, 11].

When the current in a beam is modulated, the radius is perturbed to an extent that decreases as m increases (see Section 7.4.3). This radial perturbation is undesirable and it is usual to select a value of m greater than unity to ensure adequate beam stiffness. Thus the beams in practical tubes generally operate in *space-charge balanced flow* in which $m > 1$. In order to select a suitable value of B_0 the Brillouin field is computed from (7.55), m is chosen to give the desired beam stiffness, and (7.57) then gives the cathode flux required (see Worksheet 7.1). Alternatively, if B_0 and the cathode flux are known, then an iterative calculation employing (7.55) and (7.57) can be used to find the equilibrium beam radius and the corresponding value of m. The design of the magnetic field in the beam entry region to achieve the correct initial conditions is described in Section 9.3.1, and the design of the solenoid and the magnetic circuit are described in Section 19.5.2.

The angular velocity of the electrons on the surface of a solenoid focused beam in smooth flow is found from (7.48) to be

$$\dot{\theta} = \frac{\omega_L}{\gamma}\left(1 - \sqrt{K}\right), \tag{7.58}$$

which is in the right-hand corkscrew direction when B_0 is in the positive z direction. Because a solenoid focused beam is rotating about its axis, the axial velocity of the electrons is reduced by the transfer of some of the kinetic energy into rotational motion. If the charge density and the magnetic flux density are uniform across the beam, at a given plane, then the angular velocity of all the electrons is given by (7.58). The axial velocity of an electron at radius r is given by

$$\dot{z} = \sqrt{u^2 - r^2 \dot{\theta}^2}, \tag{7.59}$$

where u is determined from (1.4) by setting V to the local potential relative to the cathode, allowing for space-charge potential depression. Now the rotational kinetic energy is zero on the axis and increases with r, whilst the potential depression is greatest on the axis and decreases with r. Thus these two effects compensate for each other to some extent. For the special case of a beam in Brillouin flow it can be

shown that the two effects cancel each other exactly [7, 12]. The axial velocity of all the electrons is then the same as that of an electron on the axis. When the beam is not in Brillouin flow $(m > 1)$ the angular velocity is smaller and the axial velocity of the electrons varies with radius. For example, if $m = 2$, then the angular velocity is only 13% of the Brillouin value, and the variation in the axial velocity is determined almost entirely by the variation of potential. These corrections to the calculation of the axial velocity are small but can be important, especially when travelling-wave tubes (TWTs) are modelled, because the centre frequency is particularly sensitive to variations in the electron velocity. In many cases it is sufficient to use the velocity of an electron on the axis as an approximation to the mean axial velocity of the electron beam.

7.4.2   Scalloping

An electron beam will only flow smoothly through a uniform axial magnetic field if it has the correct angular velocity, and if the radial velocity is zero at the equilibrium beam radius $(r = b)$. When these conditions are not satisfied the force on an electron on the surface of the beam is not zero, and the radial motion of the electron is governed by equation (7.52). In general the solution of this equation requires the use of numerical methods (see Worksheet 7.2) but useful insight is provided by examining the approximate equation obtained by assuming that the radial perturbations are small and neglecting relativistic effects. To do this we assume that $R = 1 + x$ where $x \ll 1$. The powers of R are expanded using the Binomial Theorem, and powers of x higher than the first are ignored. Then

$$\frac{d^2 x}{dz^2} = \left(\frac{1}{2} \beta_p^2 + \beta_L^2 K - \beta_L^2 \right) - \left(\frac{1}{2} \beta_p^2 + 3\beta_L^2 K + \beta_L^2 \right) x. \tag{7.60}$$

From (7.54) we find that the first bracket on the right-hand side of (7.60) is zero and that the equation can be rewritten as

$$\frac{d^2 x}{dz^2} = -2\beta_L^2 \left(1 + K \right) x. \tag{7.61}$$

An alternative form of the equation can be found by substituting for β_L and K using (7.56) and (7.57) to give

$$\frac{d^2 x}{dz^2} = -\left(2m^2 - 1 \right) \beta_p^2 x. \tag{7.62}$$

This is the equation of simple harmonic motion. Therefore the electrons at the edge of the beam execute a sinusoidal motion in the radial direction, with a wavelength which is equal to $2\pi / \beta_p$ when $m = 1$. This wavelength decreases as m is increased. This motion is known as *scalloping*. We shall see in the Section 7.4.3 that the decrease in

the scalloping wavelength as the magnetic field increases is consistent with increasing stiffness of the beam against radial perturbations. At the level of approximation we have used in this analysis, a beam which is launched into the focusing field at a radius $(b-x)$ will scallop between radii $(b-x)$ and $(b+x)$. It should be noted that a similar argument applies to electrons lying within the beam. It is necessary to ensure that the launching conditions are correct for all the electrons if the beam is to be well-controlled. The behaviour of a beam when the initial conditions are varied can be explored using Worksheet 7.2. The same worksheet can be used to study the effects of different assumptions about the axial variation of the magnetic field. If the beam is initially in smooth flow then a step change in the magnetic field produces scalloping. If, on the other hand, there is a gradual increase in the field then the beam moves smoothly to a new equilibrium beam radius with little scalloping even for quite rapid changes in the field.

7.4.3   Beam Stiffness

In a linear-beam tube the electrons in the beam become bunched axially to produce a time-varying current. This bunching upsets the equilibrium of the beam so that it tends to expand in regions of high charge density, and to contract elsewhere. In order to estimate the magnitude of this effect we consider the way in which the equilibrium beam radius varies if the beam current is increased, while keeping the magnetic field and the cathode flux constant. Since this calculation is only approximate we can safely ignore second-order effects resulting from changes in the space potential. We also assume that the wavelength on the beam is great enough for axial components of the RF electric field to be negligible.

The effect of changes in the local beam current on the behaviour of a solenoid focused beam can be explored using equation (7.52). We note that the beam current only affects the first term on the right hand side of the equation and that the square of the plasma frequency is proportional to the beam current. If we assume that the current density varies as $k(z)$ then the effect of the changes in current density can be modelled by multiplying the space-charge term in equation (7.52) by this factor. The effect of different assumptions about $k(z)$ can be examined using Worksheet 7.2. It is found that smooth changes in the beam current cause the beam to move to a new equilibrium radius with little scalloping and we are therefore justified in supposing that the effects of density modulation of the beam can be described quite accurately by the changes in the equilibrium beam radius. It should be noted that this differs from the approach used by other authors who assume that the modulation of the beam leads to scalloping [7]. We shall see that the analysis presented here is consistent with the small-signal properties of magnetically-focused beams presented in Chapter 11. The equilibrium equation is obtained from (7.52) as

$$\frac{\beta_p^2 k}{2\gamma^3} \cdot \frac{1}{R} + \frac{\beta_L^2}{\gamma^2} K \frac{1}{R^3} - \frac{\beta_L^2}{\gamma^2} R = 0, \tag{7.63}$$

where the instantaneous beam current is kI_0. Multiplying through by R^4, and substituting for β_p and K using (7.56) and (7.57), gives

$$m^2 R^4 - kR^2 - (m^2 - 1) = 0, \tag{7.64}$$

which has the solution

$$R^2 = \frac{1}{2m^2}\left(k + \sqrt{k^2 + 4m^2(m^2 - 1)}\right). \tag{7.65}$$

The solution is trivial when $m = 1$ because then $R^2 = k$ so that a doubling of the current increases the beam radius by a factor of $\sqrt{2}$. Thus, in a Brillouin focused beam the current density is constant when the beam is bunched, and the modulation takes the form of perturbation of the beam radius. For other values of m it is simple to compute the variation of R with k and the results are shown in Figure 7.5. It is evident that the radial perturbation of the beam boundary decreases sharply as m increases. Thus the stiffness of the beam increases with m, as suggested in the previous section. This is an important result because the power density in the electron beam in a microwave tube is high enough to cause damage to the surrounding metal structures if an appreciable part of the beam current is intercepted by them.

It is therefore necessary to design the focusing system with the objective of reducing the intercepted current as far as possible. The interception could, in theory, be reduced to zero by making the magnetic field very strong. This is sometimes assumed for purposes of calculation, but it is not a practical solution because of the size, weight, and cost of the focusing system that would be required. The use of Brillouin flow $(m = 1)$ is also not practical because it generally leads to an unacceptably large intercepted current. In practice it is usual to use a value of m in the range 1.3 to 3.0 where the electrons are in space-charge-balanced flow. To achieve the correct conditions it is necessary to arrange for a portion of the focusing flux to be linked to the cathode to satisfy (7.57). Sometimes, for simplicity, the beam may be launched from a shielded cathode into a magnetic field which is stronger than the Brillouin field.
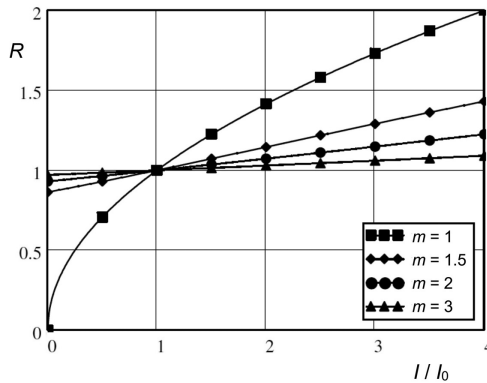
Figure 7.5: Stiffness of a solenoid-focused beam.

The beam then adopts the equilibrium radius given by (7.65) and will scallop, unless it is launched into Brillouin flow, and then compressed by a gradual increase of the field. Beams launched from non-shielded cathodes show similar behaviour [13–17]. The behaviour of such a beam can be investigated using Worksheet 7.2.

In high-power linear-beam tubes the peak beam current is of the order 2–4 times the DC current. We shall see that it is desirable to keep the diameter of the beam tunnel small, and the beam filling factor (b/a) as large as possible, in order to get a strong RF interaction between the beam and the surrounding structure. But the use of a small beam diameter requires a stronger magnetic field and, consequently, the dissipation of more power in the solenoid which supplies it. There is, thus, a trade-off between the RF interaction and the size, weight, and cost of the solenoid and its power supply. Figure 7.5 shows that, with a filling factor of 0.67, there should be no interception when m is as low as 1.5. In practice the current density profile in an electron beam does not have a sharp cut-off at the designed beam radius but commonly extends a little beyond it. Thus it may be necessary to use a higher value of m to provide a safety margin. We shall return to this issue when discussing the design principles for klystrons and TWTs. It is usual to use a computer model of the electron gun, and a portion of the focusing structure, to check that the beam is well-controlled and has the intended diameter [18].

The results shown in Figure 7.5 can be used to estimate the relative proportions of the RF beam current which are carried by modulation of the charge density (bulk current) and of the beam radius (surface current). If we assume that the axial velocity of the electrons is u_0 and that the RF current density (J_1) is uniform across the beam then we can write

$$J_1 = \frac{k I_0}{\pi r^2} - \frac{I_0}{\pi b^2}. \tag{7.66}$$

The amplitude of the bulk current (I_B) is taken to be the RF current contained within the equilibrium beam diameter, so that

$$I_B = \pi b^2 J_1 = \left(\frac{b^2}{r^2} k - 1 \right) I_0. \tag{7.67}$$

Now the amplitude of the total RF current I_1 is given by

$$I_1 = (k-1) I_0 \tag{7.68}$$

so that the normalised bulk current is

$$\frac{I_B}{I_1} = \left(\frac{1}{k-1} \right) \left(\frac{b^2}{r^2} k - 1 \right). \tag{7.69}$$

The RF surface current (I_S) is just $(I_1 - I_B)$. Figure 7.6 shows how the normalised bulk and surface currents vary with m for various values of k. Since the current density in a Brillouin focused beam is equal to the equilibrium current density, it follows that the whole of the RF current is surface current when $m = 1$. As the
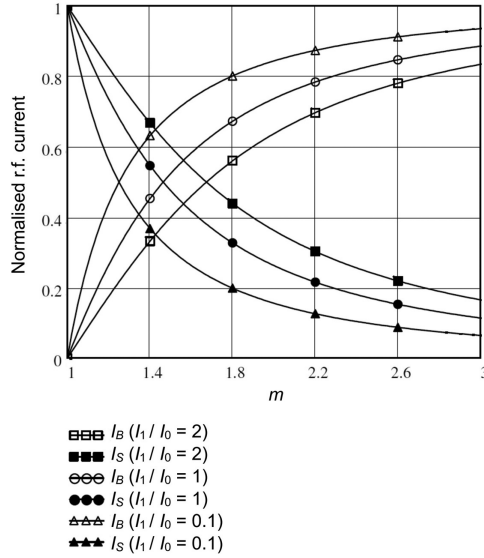
Figure 7.6: Bulk and surface currents in a modulated solenoid-focused beam.

magnetic field is increased the surface current falls and the body current increases but, even when $m = 3$ (which is the highest value normally used), there is still an appreciable surface current, especially at the higher values of k. We shall see in Section 11.3.1 that, in the limit of small modulation $(k \to 0)$, this analysis agrees with the results of space-charge wave theory.

7.5 Beam Spreading

An uncontrolled electron beam spreads out through the influence of its own space-charge. If, in addition, the electrons are rotating about the axis of the beam then the rate at which the beam spreads is increased by the centrifugal force. It is necessary to be able to compute the spreading of an electron beam under these conditions, both to model its expansion into the collector as it leaves the end of the focusing system, and to understand the behaviour of focusing systems based on the use of magnetic or electrostatic lenses.

7.5.1 The Universal Beam-Spreading Curve

If there is no magnetic field and the angular velocity of the electrons is zero, the equation of motion of an electron on the edge of a non-relativistic beam is obtained from (7.53) as [19]

$$\frac{d^2R}{dz^2} = \frac{1}{2}\beta_p^2 \frac{1}{R}.$$ (7.70)

Multiplying both sides of equation (7.70) by $2(dR/dz)$ and integrating with respect to z gives

$$\left(\frac{dR}{dz}\right)^2 = \beta_p^2 \ln(R) + C. \tag{7.71}$$

If we choose $R = 1$ when the electrons are moving parallel to the axis (the *beam waist*) the constant of integration C is zero. We recall that β_p has been defined at this plane (see (7.47)). Then

$$\frac{dR}{dz} = \beta_p \sqrt{\ln(R)} \tag{7.72}$$

so that

$$\beta_p z = \int_1^R \frac{dR}{\sqrt{\ln(R)}}. \tag{7.73}$$

The integral can be evaluated numerically after first making the substitution $x^2 = \ln(R)$ to give

$$\beta_p z = \int_0^{\sqrt{\ln(R)}} 2\exp(x^2)\,dx. \tag{7.74}$$

The result of the integration is the universal beam spreading curve shown in Figure 7.7. This curve is symmetrical about the R axis because the rate of convergence is the same as the rate of expansion. An approximate analytical equation for the curve can be obtained by assuming that $R = 1 + x$ where $x \ll 1$ and using the series expansion of the logarithm

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \tag{7.75}$$
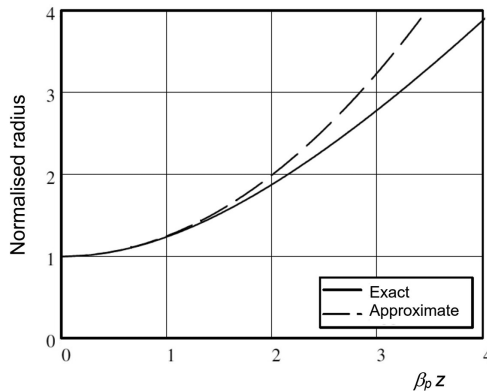
Figure 7.7: The universal beam spreading curve.

so that (7.73) can be written approximately as

$$\beta_p z = \int_0^x \frac{dx}{\sqrt{x}},$$

(7.76)

which can be integrated to give

$$x = \left(\frac{1}{2}\beta_p z\right)^2$$

(7.77)

so that

$$R \approx 1 + \frac{1}{4}\beta_p^2 z^2.$$

(7.78)

Figure 7.7 shows a comparison between the universal beam spreading curve and the approximation given by (7.78). It can be seen that the approximation is useful when $\beta_p z \leq 2$. The spreading of a relativistic beam can be found by replacing β_p with $\beta_p / \gamma^{1.5}$.

7.5.2   Spreading of Rotating Beams

Consider a beam composed of mono-energetic electrons emerging from a focusing field into an electron collector. If the electrons are launched from a cathode which is shielded from magnetic flux then, by Busch's theorem, they emerge from the magnetic field with no angular velocity and beam edge within the collector follows the universal beam spreading curve. In many cases, however, the beam is designed to be in space-charge balanced flow with magnetic flux linked to the cathode. When such a beam emerges into a region which is free from magnetic field its non-relativistic equation of motion is obtained from (7.53) as [7]

$$\frac{d^2 R}{dz^2} = \frac{1}{2}\beta_p^2 \frac{1}{R} + \beta_L^2 K \frac{1}{R^3}.$$

(7.79)

The derivation now follows the same procedure as that for the universal beam spreading curve. Equation (7.79) is multiplied by $2(dR/dz)$ and integrated with respect to z with the initial condition $R = 1$. The result is

$$\left(\frac{dR}{dz}\right)^2 = \beta_p^2 \ln(R) + \frac{1}{2}\beta_L^2 K\left(1 - \frac{1}{R^2}\right).$$

(7.80)

Substituting for β_L and K using (7.56) and (7.57) gives

$$\left(\frac{dR}{dz}\right)^2 = \beta_p^2 \left\{\ln(R) + \frac{1}{4}(m^2 - 1)\left(1 - \frac{1}{R^2}\right)\right\}.$$

(7.81)
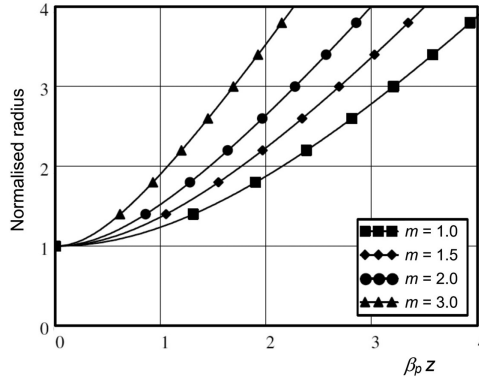
Figure 7.8: Spreading of rotating beams.

Equation (7.81) reduces to (7.72) when $m = 1$. As before we make the substitution $x^2 = \ln R$ so that

$$\beta_p z = \int_0^{\sqrt{\ln R}} \frac{2x \exp(x^2)}{\sqrt{x^2 + \frac{1}{4}(m^2 - 1)(1 - \exp(-2x^2))}}\, dx. \tag{7.82}$$

Equation (7.82) can be integrated numerically (see Worksheet 7.3) and the results are shown in Figure 7.8 for a range of values of m. The trajectories spread more rapidly as m increases as a result of the increasing centrifugal force. An approximate solution to (7.81) can be found by assuming that, by analogy with (7.78), the beam radius varies with position as

$$R \approx 1 + \alpha^2 z^2. \tag{7.83}$$

Substituting in (7.80) yields

$$\alpha^2 = \frac{1}{4}\left(\beta_p^2 + K\beta_L^2\right) \tag{7.84}$$

so that

$$R \approx 1 + \frac{1}{4}\left(\beta_p^2 + K\beta_L^2\right)z^2. \tag{7.85}$$

This reduces to (7.78) when $K = 0$. The spreading of relativistic beams can be found by making relativistic corrections to β_p and β_L.

7.6   Periodic Focusing

As an alternative to solenoid focusing it is possible to control an electron beam by passing it through a periodic array of converging lenses, as shown in Figure 7.9
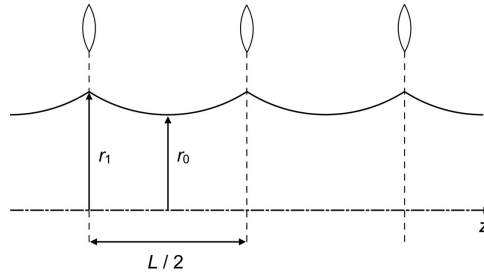
Figure 7.9: Focusing of a beam by an array of converging lenses.

[20–22]. The beam converges as it leaves a lens in such a way that the space-charge forces produce an equal divergence as it enters the next lens. The edge of the beam therefore has a ripple with the periodicity of the lenses. Systems using magnetic lenses with permanent magnets, known as *periodic permanent magnet* (PPM) systems, are commonly used for TWTs and occasionally for klystrons. Periodic electrostatic focusing is sometimes used, but is usually more difficult to implement.

7.6.1   Periodic Permanent Magnet (PPM) Focusing

Figure 7.10(a) shows the general arrangement of a periodic permanent magnet system. The permanent magnets, in the form of rings, are magnetised axially and arranged so that their polarities alternate. The flux is conveyed to the region of the electron beam by iron pole pieces so that the field on the axis varies approximately sinusoidally with position, as shown in Figure 7.10(b). This arrangement has the advantage that the size and weight of the magnets is considerably less than would be required to produce an equivalent uniform magnetic field. The properties of a PPM system can be modelled by finding the solutions of (7.53) when the magnetic field profile is defined by

$$f(z) = \cos\left(\frac{2\pi z}{L}\right) = \cos(\beta_0 z). \tag{7.86}$$

The differential equation to be solved is

$$\frac{d^2 R}{dz^2} = \frac{1}{2}\beta_p^2 \frac{1}{R} + \beta_L^2 K \frac{1}{R^3} - \beta_L^2 \cos^2(\beta_0 z) R. \tag{7.87}$$

If we let $\theta = \beta_0 z$ then (7.87) becomes

$$\frac{d^2 R}{d\theta^2} = \beta \frac{1}{R} + 2\alpha K \frac{1}{R^3} - 2\alpha \cos^2(\theta) R, \tag{7.88}$$

where the magnetic field parameter

$$\alpha = \frac{1}{2}\frac{\beta_L^2}{\beta_0^2}, \tag{7.89}$$
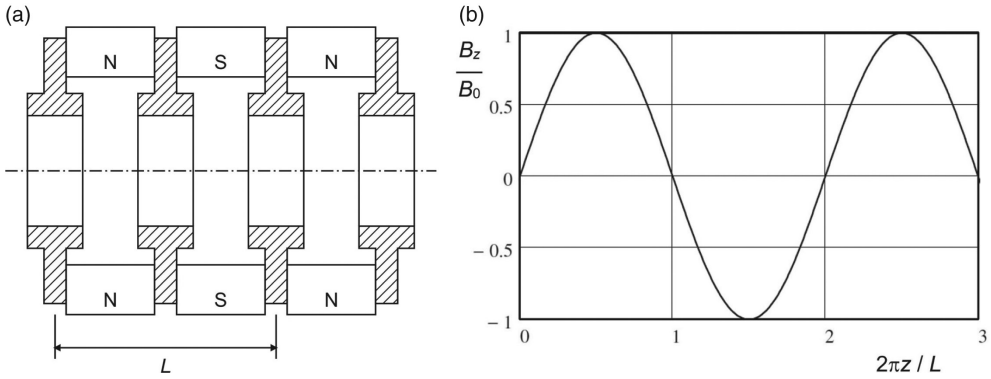
Figure 7.10: Periodic permanent magnet (PPM) focusing system: (a) sectioned elevation, and (b) the magnetic field profile.

the space-charge parameter

$$\beta = \frac{1}{2}\frac{\beta_p^2}{\beta_0^2},$$ (7.90)

and the cathode flux parameter K are defined at $z = 0$. This equation cannot be solved analytically but solutions using analogue computers are given in [22] for $K = 0$ and in [23] for $K \geq 0$. The solution of (7.88) is readily achieved using numerical methods, as shown in Worksheet 7.4, with results which agree with those in the literature. Figure 7.11 shows some typical solutions. The beam profile is approximately sinusoidal with a wavelength which is one half of the wavelength of the magnetic field. This is because the focusing effect depends upon the square of the magnetic field and, therefore, there are two converging lenses in each wavelength of the magnetic field. If the focusing conditions are correct then the electron trajectories are perfectly periodic as shown in Figure 7.11(a). If, however, the focusing is not correct then scalloping is superimposed upon the PPM ripple as shown in Figures 7.11(b) and (c) in which β has been altered by ±20%.

When the beam has no scalloping (the minimum ripple solution) it is found that the value of β is completely determined when α and K are fixed, as shown in Figure 7.12 [23]. The ripple in the beam edge defined by

$$\delta = \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}}$$ (7.91)

also depends only on α and K, as shown in Figure 7.13 [23].

If α is increased, with a fixed value of K, and β is adjusted to seek for a minimum ripple solution it is found that, beyond a certain value of α, no stable solutions can be found. For higher values of α the ripple may show a sinusoidal modulation but, for a further increase in α, the ripple increases exponentially with z. It has been shown experimentally that for a shielded cathode $(K = 0)$ $\alpha < 0.57$ [24]. Theoretical analyses have suggested other values for this limit [23–25]. In practice, however, the

$\alpha = 0.3$ $\beta = 0.187$ $K = 0.1$ $R_{max} = 1$ $R_{min}\ 0.828$ ripple = 9.4

(a)

$\alpha = 0.3$ $\beta = 0.187$ $K = 0.1$ $R_{max} = 1$ $R_{min}\ 0.828$ ripple = 9.4

(b)

$\alpha = 0.3$ $\beta = 0.15$ $K = 0.1$ $R_{max} = 1$ $R_{min}\ 0.748$ ripple = 14.4

(c)
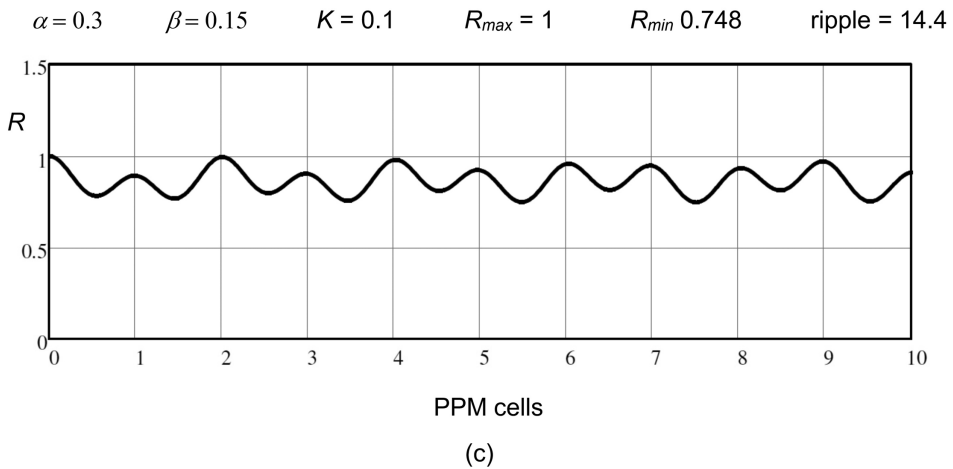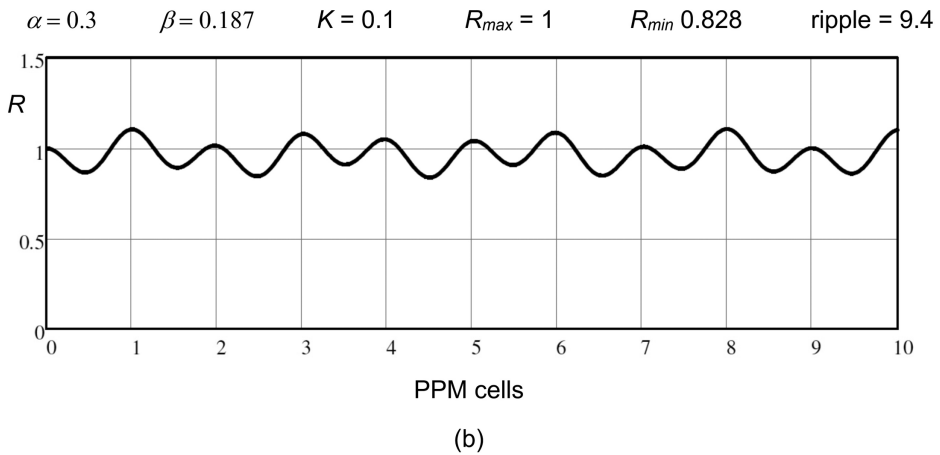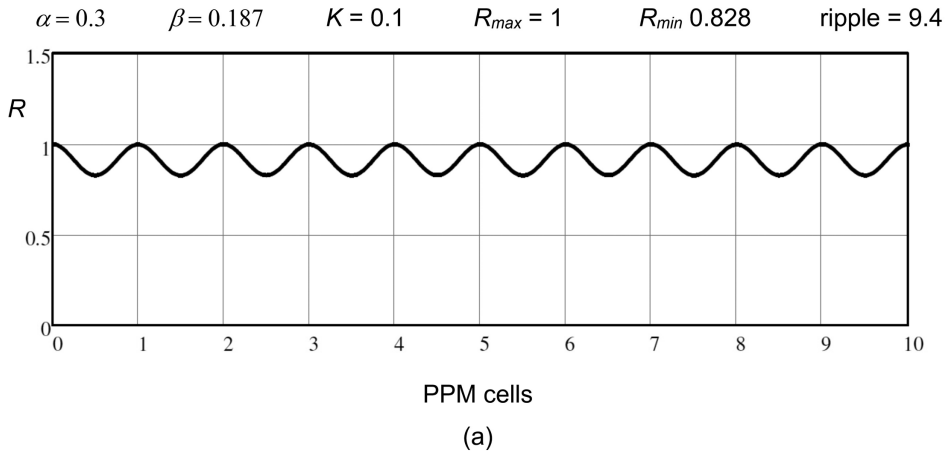
Figure 7.11: PPM focused beam profiles $(\alpha = 0.3, K = 0.1)$: (a) $\beta = 0.187$, (b) $\beta = 0.224$, and (c) $\beta = 0.15$.
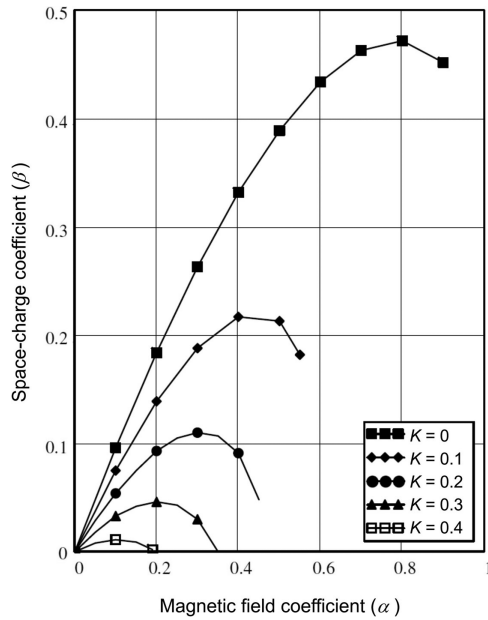
Figure 7.12: Dependence of β on α and K for minimum ripple of a PPM focused beam (copyright 1955, IEEE, reproduced, with permission, from [23]).
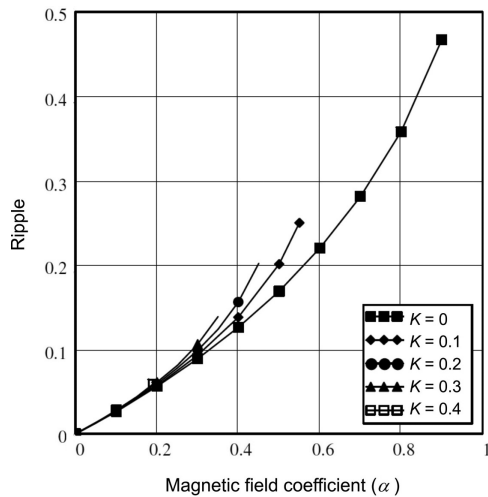
Figure 7.13: Dependence of ripple on α and K for minimum ripple of a PPM focused beam (copyright 1955, IEEE, reproduced, with permission, from [23]).

requirement for the ripple to be small in a well-focused beam means that the value of α chosen is well below the limiting value.

Approximate analytical solutions to (7.88) have been described using Fourier expansion or successive approximations [23, 25, 26], but these are mathematically

difficult. A good approximation can be obtained, however, by assuming that the beam profile is

$$R_m = 1 + \delta \cos(2\beta_0 z), \tag{7.92}$$

where R_m is the beam radius normalised to the mean radius. Following [26] we express (7.87) in terms of the mean beam radius as

$$\frac{d^2 R_m}{dz^2} = \frac{1}{2}\beta_{pm}^2 \frac{1}{R_m} + \beta_{Lm}^2 K_m \frac{1}{R_m^3} - 2\beta_{Lm}^2 \cos^2(\beta_0 z) R_m, \tag{7.93}$$

where the suffix m denotes that the parameters are defined at the plane where the beam has its mean radius, that is when $\beta_0 z = \pi/4$. We note that the magnitude of the magnetic field at that plane is $B_0/\sqrt{2}$ which is the RMS value of the magnetic field. Substitution for R_m in (7.93) gives

$$-4\beta_0^2 \delta \cos(2\beta_0 z) = \frac{\beta_{pm}^2}{2(1 + \delta \cos(2\beta_0 z))} + \frac{\beta_{Lm}^2 K}{(1 + \delta \cos(2\beta_0 z))^3} \\ - 2\beta_{Lm}^2 \cos^2(\beta_0 z)(1 + \delta \cos(2\beta_0 z)). \tag{7.94}$$

When $\beta_0 z = \pi/4$ (7.94) becomes

$$\frac{\beta_{pm}^2}{2} + \beta_{Lm}^2 K_m - \beta_{Lm}^2 = 0 \tag{7.95}$$

or

$$\beta_{pm}^2 = 2\beta_{Lm}^2(1 - K_m) = \frac{2\beta_{Lm}^2}{m^2}. \tag{7.96}$$

Comparing (7.96) with (7.54) we see that the equilibrium conditions for an electron beam, having a given mean beam radius and cathode flux, are the same for both solenoid and PPM focusing if the RMS PPM field is equal to the solenoid field. Note that that (7.96) shows that there is a linear relationship between β_{pm} and β_{Lm} for a given choice of K_m within the range of validity of this analysis.

To find the amplitude of the ripple we set $\beta_0 z = \pi/2$ in (7.94) and obtain

$$4\beta_0^2 \delta = \frac{\beta_{pm}^2}{2(1 - \delta)} + \frac{\beta_{Lm}^2 K_m}{(1 - \delta)^3}. \tag{7.97}$$

Substituting for β_{pm} and rearranging gives

$$\alpha_m = 2\delta(1 - \delta)\left[1 - K_m\left(1 - \frac{1}{(1 - \delta)^2}\right)\right]^{-1}, \tag{7.98}$$

where α_m, given by

$$\alpha_m = \frac{1}{2}\frac{\beta_{Lm}^2}{\beta_0^2} = \frac{1}{2}\alpha \tag{7.99}$$

has been defined by analogy with (7.89). Now because $\alpha = 2\alpha_m$ and it can easily be shown that

$$K_m = 2K\left(1 + \delta\right)^4 \tag{7.100}$$

so we may calculate the dependence of α on δ for different values of K for comparison with Figure 7.13. The comparison shows good agreement for $\alpha \leq 0.4$ so this approximate theory is valid for a ripple of up to 15%. Also, since

$$\beta_p^2 = \frac{\beta_{pm}^2}{\left(1 + \delta\right)^2}, \tag{7.101}$$

we can solve (7.98) numerically for δ in terms of α_m, and then calculate β in terms of α and K using (7.96). When the results are plotted and compared with Figure 7.12 there is, again, good agreement for $\alpha \leq 0.4$ (see Worksheet 7.4). This range of parameters covers the values which are generally used, so the approximate theory can be used for design calculations.

We have seen, in Figure 7.11, that when a PPM focused beam is launched incorrectly, scalloping is superimposed upon the PPM ripple. The stiffness of a PPM focused beam can be investigated in the same way as that for solenoid focusing by multiplying the space-charge coefficient β by a factor $k = I/I_0$. As before, we assume that the axial variation of charge density is slow so that it is possible to consider each part of the beam as being in equilibrium. It is then only necessary to calculate the equilibrium radius as a function of the space-charge and magnetic field parameters.

From (7.93) $d^2 R_m/d\theta^2 = 0$ when $\theta = \pi/4$ so the equilibrium radius is the solution of

$$\beta_m \frac{1}{R_m} + 2\alpha_m K_m \frac{1}{R_m^3} - 2\alpha_m R_m = 0, \tag{7.102}$$

which is identical in form to (7.63). The stiffness of a PPM focused beam is, therefore, the same as that of the equivalent solenoid-focused beam. Thus, a modulated PPM focused beam remains in equal ripple flow but with the mean radius determined locally by the instantaneous beam current. This can be confirmed by numerical integration of (7.87) in which the current increases linearly with distance (see Worksheet 7.4).

In some tubes the axial variation of the magnetic field is periodic but not sinusoidal [27]. The focusing in these cases can be investigated by the use of Worksheet 7.4, if a suitable function for the field profile is defined. A particular case is the use of double period PPM focusing in which adjacent pairs of cells have the same polarity. This is used when it is necessary to fit the periodicity of the focusing to the periodicity of the RF structure surrounding the beam [28]. Long period focusing in which the magnetic field is kept approximately constant in each of a series of sections separated by abrupt field reversals is described in [29, 30].

7.6.2   Practical PPM Focusing Systems

PPM focusing systems are sensitive to variations from the ideal field. These may be caused by:

- variations in the strengths of the magnets caused by manufacturing tolerances or accidental demagnetisation when the stack is being assembled;
- variations in the strength of the field in the cells near the ends of the stack caused by end effects [31], or by the omission of magnets to allow input and output waveguides to pass through [32];
- variations in the azimuthal symmetry of the fields caused by tilting of the pole-pieces, misalignment of the magnets, or omission of magnets as already mentioned [32].

It is standard practice to make empirical adjustments to the magnetic field of a PPM focused tube by the addition of soft iron shunts that bridge the pole-pieces on the outside of the stack. In this way both the magnitude and the symmetry of the field can be adjusted until good beam transmission is obtained. The process is time-consuming and it is desirable to design PPM focusing systems in such a way that empirical adjustments are kept to a minimum.

Problems caused by end effects, or by the need to pass waveguides through the PPM structure can be dealt with, in principle, by adjusting the design of the cells near to ends of the stack to compensate for the departures from the ideal field. Kory, who examined the magnitudes of these effects for a helix TWT, found that the greatest changes were caused by the omission of magnets to allow for the passage of waveguides [32]. The asymmetries in the field caused appreciable movement of the axis of the beam away from the axis of the stack. The effects of a simulated shunt were small, and those of misalignment of a magnet were found to be negligible.

The effects of random variations in the strengths of the magnets have been discussed by a number of authors [33–36]. It was found that such variations cause the ripple in the beam to grow with distance. The growth can be reduced by increasing the strength of the magnetic field above the theoretical value. This is in qualitative agreement with experimental observations. The most comprehensive study was carried out by MacGregor who examined the conditions for optimal beam transmission with random variations in magnet strengths of 2.5% and 5%. Figure 7.14 summarises his results using Harker's parameters [33, 34]. McGregor also showed that good beam transmission would be difficult to obtain with values of α greater than 0.4, when the field variations were 2.5%, and 0.3 when they were 5%.

The design of a PPM focusing system starts from the desired current, voltage, and mean radius of the beam. It is also necessary to fix the maximum acceptable ripple in the beam and the value of m which will give adequate stiffness. The design calculations are:

i) Calculate the Brillouin field using (7.55) and the peak field $B_0 = \sqrt{2} m B_B$. In order to ensure good beam transmission it is usual to increase the magnetic field from this theoretical value as described above (see Figure 7.14).

ii) Calculate K_m using (7.57).
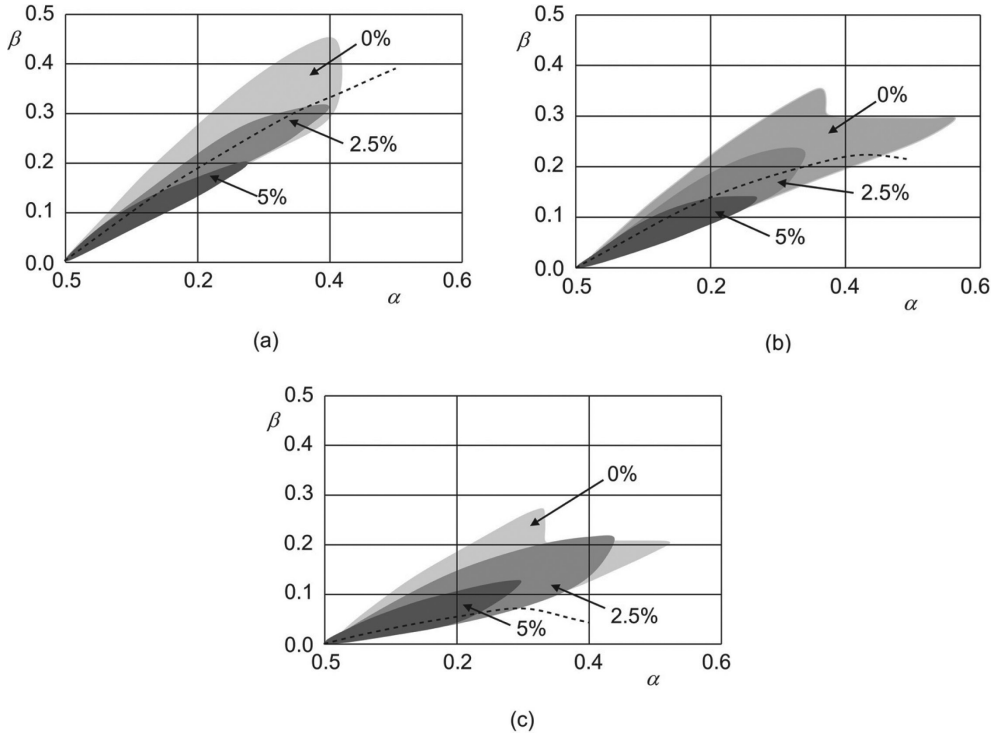
Figure 7.14: Regions of acceptable beam transmission with 2.5% and 5% random variations of magnet strength compared with the minimum ripple solution (dashed): (a) K = 0, (b) K = 0.1, and (c) K = 0.2
(copyright 1986, A. MacGregor, reproduced, with permission from [34]).

iii) Calculate α_m using (7.98) and check that the value is less than 0.2.
iv) Calculate β_0 and then L using (7.99).
v) Design the magnetic field in the beam entry region to achieve the correct initial conditions as described in Section 9.3.2.
vi) Design the magnet stack as discussed in Section 19.7.3.

Finally, in nearly every case, it is necessary to make empirical adjustments to the field in each tube by adding shunts to the outside of the PPM stack to minimise the intercepted current.

7.6.3   Periodic Electrostatic Focusing

A periodic electrostatic focusing system can be constructed as shown in Figure 7.9 using electrostatic lenses [21, 37]. If cylindrical symmetry is assumed then the non-relativistic equations of motion are, from (7.9) and (7.11)

$$\frac{d^2r}{dt^2} = -\frac{e}{m_0} E_r \tag{7.103}$$

and

$$\frac{d^2 z}{dt^2} = -\frac{e}{m_0} E_z.$$

(7.104)

The angular velocity is zero in the absence of a magnetic field. Note that it is no longer possible to assume that the axial velocity is constant. Let us suppose that the electrostatic potential on the axis relative to the cathode is given by

$$V(0, z) = V_a + V_1 \left[1 - \cos(\beta_0 z) \right].$$

(7.105)

This potential could be provided by a series of equally spaced plates, with circular holes for the passage of the beam, that are held alternately at potentials V_a and $(V_a + 2V_1)$ relative to the cathode. From the discussion of the Einzel lens in Section 7.3.2 we expect this potential variation to correspond to a series of converging lenses. The higher-order terms in the Fourier expansion of the potential at the edge of the apertures decay rapidly with decreasing radius. Thus the potential variation close to the axis can be represented by a single term from the series. Substituting in Laplace's equation (7.19) we obtain

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r} \right) - \beta_0^2 V = 0.$$

(7.106)

This is a form of Bessel's equation which has the general solution

$$V(r, z) = V_1 I_0 (\beta_0 r) \cos(\beta_0 z),$$

(7.107)

where $I_0 (\beta_0 r)$ is the modified Bessel function of the first kind [38]. If $\beta_0 r \leq 1$ the Bessel function can be written to a good approximation as

$$I_0 (\beta_0 r) \simeq 1 + \frac{(\beta_0 r)^2}{4}.$$

(7.108)

Then, in the paraxial ray approximation, the axial electric field is given by

$$E_z = -\frac{\partial V}{\partial z} \approx -\beta_0 V_1 \sin(\beta_0 z)$$

(7.109)

and the radial electric field is given by

$$E_r = -\frac{\partial V}{\partial r} \approx \frac{1}{2} V_1 \beta_0^2 r \cos(\beta_0 z).$$

(7.110)

When the radial field component is added to the radial component of the space-charge field from (7.3) and substituted in (7.103) the result can be expressed as

$$\frac{d^2 R}{d\theta^2} = \beta\frac{1}{R} - \frac{1}{4} U_1 R \cos(Z),$$

(7.111)

where V_a is the accelerating potential, $R = r/b$, $\theta = \beta_0 u_0 t$, $Z = \beta_0 z$, β is defined by (7.90) and the strength of the lenses is represented by the normalised potential $U_1 = V_1/V_a$. Similarly (7.104) can be written as

$$\frac{d^2 Z}{d\theta^2} = \frac{1}{2} U_1 \sin(Z). \tag{7.112}$$

This pair of equations corresponds to equation (7.88) for PPM focusing. The properties of this system can be investigated by the simultaneous numerical integration of (7.111) and (7.112) with the initial conditions $R = 1$, $dR/d\theta = 0$, $Z = 0$ and $dZ/d\theta = \sqrt{1 - U_1}$. Stable minimum ripple solutions can be found in the same manner as for PPM focusing, as shown in Figure 7.15 (see Worksheet 7.5). In the case shown, the potentials of alternate electrodes are positive with respect to the potential at the input and the beam radius is always less than or equal to the initial radius. If the potentials of those electrodes are made negative by reversing the sign of U_1 then minimum ripple solutions can be found in which the minimum radius is equal to the initial radius.

Figure 7.16 shows the relationship between β and U_1 for minimum ripple, and Figure 7.17 shows the dependence of the ripple on the normalised potential. Figure 7.17 is similar to the curve for $K = 0$ in Figure 7.13. However, from the definition of the normalised potential (U_1), it can be seen that for 10% ripple the variation of potential on the axis is nearly 80% of the accelerating voltage. The voltage required on off-axis electrodes would be appreciably higher. Comparison between Figure 7.16 and Figure 7.12 shows that the value of β in a periodic electrostatic system is much less than that in a PPM system with the same ripple. Thus, in the electrostatic system, the space-charge in the beam must be much less, and/or the spacing between the lenses must be smaller, than in a PPM system. It should also be noted that the electrostatic system has only one converging lens per period of the field whereas a PPM system has two. Finally we note that, when alternate electrodes are positive, the mean velocity of the electrons is appreciably greater than the initial velocity. For further information see [39, 40]. In practice, the need to use
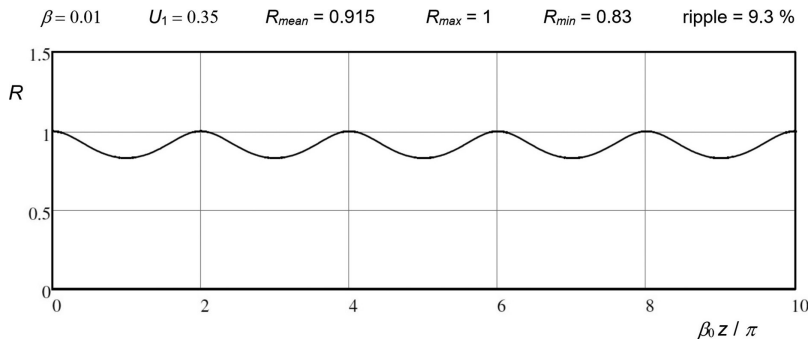
Figure 7.15: Example of a minimum ripple solution for periodic electrostatic focusing: $(\beta = 0.01, U_1 = 0.35)$.
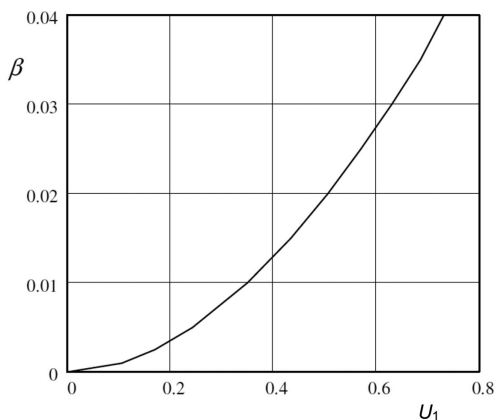
Figure 7.16: Periodic electrostatic focusing parameters for minimum ripple solutions.
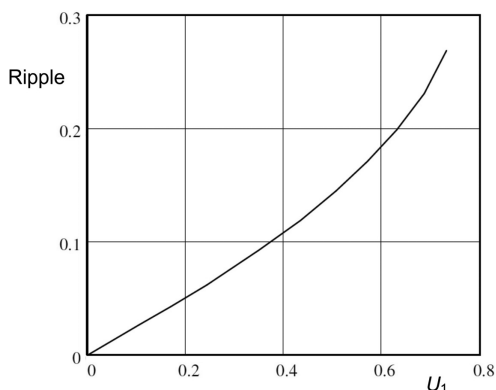
Figure 7.17: Periodic electrostatic focusing: dependence of the beam edge ripple on the normalised electrostatic potential.

high voltages, and the difficulty of combining the lens electrodes with the radiofrequency structure of a tube has, meant that the method has only been used for a few low power tubes [8, 37].

7.7   Other Forms of Linear Electron Beam

Although most linear beam tubes use solid, cylindrical, electron beams, these suffer from the disadvantage that the diameter of the beam must be kept small to minimise the variation in the field of the radio-frequency structure over the beam. That, in turn means that the beam current is limited by the need to focus the beam. Two other arrangements, described below, are sometimes used that address this issue by concentrating the electrons in a region of strong radio-frequency electric field.

7.7.1   Sheet Electron Beams

A sheet electron beam provides an alternative way of achieving a high current with an acceptable current density. This type of beam has become the subject of intense interest in recent years because of its potential application in devices at frequencies higher than 100 GHz where the size of the radiofrequency structure becomes very small. It is then advantageous to employ the interaction between a planar structure and an electron beam in the form of a thin sheet passing close to its surface. Unfortunately, this kind of electron beam can become unstable [41] because the space-charge fields caused by perturbations of the sheet normal to its plane act to increase those perturbations. It is found that the edges of the sheet tend to curl so that the uniformity of the beam is destroyed [42]. This is a particular example of the diocotron instability which occurs when a thin electron beam is controlled by crossed electric and magnetic fields [12]. However, it is possible to achieve stable flow of a sheet beam by arranging that the shape of the beam tunnel is such that the forces tending to produce instability are minimised [42, 43]. Stable flow has also been produced by using periodically cusped magnetic fields [44, 45], or with periodic permanent magnet fields [46].

7.7.2   Annular Electron Beams

An ideal annular beam has a uniform current density within an annulus centred on the axis. This type of beam has the advantage that it can concentrate the electrons in a region where the radio-frequency electric field of the surrounding structure is strong. It therefore strengthens the interaction between them. Like sheet beams, annular beams have a tendency to become unstable at high current densities [41, 47, 48]. An annular beam can be thought of as equivalent to a sheet beam in which periodic boundary conditions are applied at the edges. Although annular beams have been used in high power klystrons [49] the additional complexities involved in creating them and focusing them are not generally worthwhile. It has been shown that PPM focused annular beams are inherently unstable [50]. An analysis by Lau [51], which included relativistic effects, showed that at high beam energies the negative mass and cyclotron maser instabilities are more important than the diocotron instability. These are exploited in fast-wave devices such as the gyrotron where it is found that stable transport of annular beams is possible when the space-charge density is small.

7.8   Imperfections in Electron Beams

The theory and design of microwave linear-beam tubes is usually based on the assumption that the electron beam has uniform axial velocity and current density, and that the electrons flow without scalloping. In practice, all real electron beams depart from these ideals to some extent, and the imperfections cause the

performance of the tube to depart from the predictions of theory. It is generally possible to produce an electron beam whose properties are close to the theoretical ideal by careful design of the electron gun and the beam entry conditions, as described in Chapter 9. Here we consider briefly the effects of thermal velocities, and ion trapping, which may be significant even when the beam is well collimated.

7.8.1   Thermal Velocities

Since the electrons in the beam are normally obtained from a thermionic cathode, which is operated at a temperature of around 1000 °C, they are not actually mono-energetic but have an energy spectrum with a width of rather less than 1 eV. The variation in the axial velocity is, therefore, very small and it does not normally have much effect on the operation of the tube other than contributing to the noise signal at the output. The radial velocity components, though small, tend to cause the beam to have a 'halo' outside the theoretical beam edge. These electrons may be collected on the RF structure so that there is a small, but non-zero, body current even in a tube which should, theoretically, show no such interception [52–55] (see Section 9.2.4). Further information can be found in [56].

7.8.2   Trapped Ions

The potential well, caused by the electronic space-charge, acts as a trap for any positive ions produced by collisions between the electrons and any residual gas molecules in the tube. This is not normally a problem in pulsed tubes because there is insufficient time for ions to accumulate during the pulse. Any ions which have been formed tend to drift to the tube walls between pulses. In continuous wave (CW) tubes the accumulation of ions can lead to:

- Disruption of the focusing because the ionic space-charge tends to neutralise the electronic space charge. In extreme cases where the background gas pressure is high it has been known for a tube to continue to operate without a magnetic focusing field. The trapping, and de-trapping, of ions can lead to relaxation oscillations which modulate the output of the tube at low frequencies [57–60].
- Ion plasma oscillations which are the result of a two-stream interaction between the electron beam and the stationary cloud of ions. The ion plasma frequency is much lower than the electron plasma frequency because of the greater masses of the ions. Ion plasma oscillations typically show themselves as a modulation of the output of the tube at a frequency of a few megahertz [61].
- Bombardment of the cathode by ions that drift back into the field of the electron gun and strike it with high energies. The result is usually seen as a damaged spot in the centre of the cathode. In some tubes a hole is provided in the centre of the cathode to allow any ions to pass through and be collected harmlessly on an electrode provided for the purpose. An alternative strategy is to provide an ion trapping electrode close to the electron gun which provides a potential barrier

to the flow of ions towards the gun. That has the disadvantage that the density of ions trapped in the beam region will increase resulting in the effects noted in (a) and (b) above.

On the whole the best solution to the problems caused by ions is to try to reduce their generation, as far as possible, by maintaining a low background gas pressure in the tube. This requires careful attention to the outgassing of the tube by baking it for a sufficient time when it is evacuated. In large tubes it is normal for one, or two, ion pumps to be permanently attached to the tube to maintain the gas pressure when it is in storage, and to enable gas bursts to be monitored. In a working tube the electrons tend to sweep ions into the collector where they become embedded in the surface. For this reason the gas pressure tends to drop when a tube is operated for a long period.

References

[1] V. Srivastava and R. G. Carter, 'Effect of boundaries on the space charge potential in coupled cavity travelling wave tubes', *IEE Proceedings I: Solid-State and Electron Devices,* vol. 133, pp. 185–188, 1986.

[2] J. D. Lawson, *The Physics of Charged-Particle Beams*. Oxford: Oxford University Press, 1977.

[3] S. E. Tsimring, *Electron Beams and Microwave Vacuum Electronics*. Hoboken, NJ: John Wiley & Sons, 2006.

[4] W. J. Kleen, *Electronics of Microwave Tubes*. New York: Academic Press, 1958.

[5] B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*. London: Oxford University Press, 1957.

[6] S. Ramo *et al.*, *Fields and Waves in Communication Electronics*. New York: Wiley, 1965.

[7] J. F. Gittins, *Power Travelling-Wave Tubes*. London: English Universities Press, 1965.

[8] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, England: Research Studies Press Ltd., 1995.

[9] L. Brillouin, 'A theorem of Larmor and its importance for electrons in magnetic fields', *Physical Review,* vol. 67, pp. 260–266, 1945.

[10] W. Neugebauer, 'Equlibrium solutions for partially immersed relativistic electron beams', *IEEE Transactions on Electron Devices,* vol. ED-14, pp. 686–693, 1967.

[11] W. Neugebauer, 'Equlibrium solutions for partially immersed relativistic electron beams', *IEEE Transactions on Electron Devices,* vol. ED-15, p. 43, 1968.

[12] A. H. W. Beck, *Space-Charge Waves and Slow Electromagnetic Waves*. London: Pergamon Press, 1958.

[13] E. A. Ash, 'Compression and reflection of Brillouin beams', *Journal of Electronics and Control,* vol. 15, pp. 401–417, 1963.

[14] O. P. Gandhi and N. C. Vaidya, 'Behavior of electron beams under magnetic compression', *Proceedings of the IEEE,* vol. 52, pp. 1052–1053, 1964.

[15] N. C. Vaidya and O. P. Gandhi, 'Performance of magnetically compressed O-type electron beams emitted from nonshielded cathodes', *IEEE Transactions on Electron Devices,* vol. 13, pp. 453–458, 1966.

[16] K. Amboss, 'Studies of a magnetically compressed electron beam', *IEEE Transactions on Electron Devices,* vol. 16, pp. 897–904, 1969.

[17] J. A. Seeger, 'Magnetic compression of axially symmetric Brillouin-focused electron beams', *IEEE Transactions on Electron Devices,* vol. 16, pp. 1–5, 1969.

[18] W. B. Herrmannsfeldt, 'Numerical design of electron guns and space charge limited transport systems', *Nuclear Instruments and Methods in Physics Research,* vol. 187, pp. 245–253, 1981.

[19] K. R. Spangenberg, *Vacuum Tubes*. New York: McGraw-Hill, 1948.

[20] J. R. Pierce, 'Spatially alternating magnetic fields for focusing low-voltage electron beams', *Journal of Applied Physics,* vol. 24, p. 1247, 1953.

[21] A. M. Clogston and H. Heffner, 'Focusing of an electron beam by periodic fields', *Journal of Applied Physics,* vol. 25, pp. 436–447, 1954.

[22] J. T. Mendel *et al.*, 'Electron beam focusing with periodic permanent magnet fields', *Proceedings of the I.R.E.,* vol. 42, pp. 800–810, 1954.

[23] K. J. Harker, 'Periodic focusing of beams from partially shielded cathodes', *IRE Transactions on Electron Devices,* vol. ED-2, pp. 13–19, October 1955.

[24] J. T. Mendel *et al.*, 'Electron beam focusing with periodic permanent magnet fields', *Proceedings of the Institute of Radio Engineers,* vol. 42, pp. 800–810, 1954.

[25] K. K. N. Chang, 'Stability of periodic-field beam focusing', *Journal of Applied Physics,* vol. 27, pp. 1527–1532, 1956.

[26] K. Ura and M. Terada, 'Equivalence of periodic magnetic field to uniform magnetic field in electron beam focusing', *IEEE Transactions on Electron Devices,* vol. ED-13, pp. 930–934, 1966.

[27] D. C. Buck, 'Stability of a cylindrical electron beam in nonsinusoidal periodic magnetic-focusing fields', *IRE Transactions on Electron Devices,* vol. 4, pp. 44–49, 1957.

[28] S. K. Datta *et al.*, 'Stability analysis for electron beam transport in double-periodic permanent magnet focusing structure', in *Eighth IEEE International Vacuum Electronics Conference*, Kitakyushu, Japan, pp. 357–358, 2007.

[29] P. F. C. Burke, 'Compensated reversed field focusing of electron beams', *Proceedings of the IEEE,* vol. 51, pp. 1653–1659, 1963.

[30] W. Siekanowicz and J. Cash Jr, 'Focusing of Brillouin electron beams by use of long-period magnetic fields', in *1964 International Electron Devices Meeting*, pp. 88–90, 1964.

[31] M. J. Schindler, 'The magnetic field and flux distributions in a periodic focusing stack for TWTs', *RCA Review,* vol. 21, pp. 414–436, 1960.

[32] C. L. Kory, 'Effect of geometric azimuthal asymmetries of PPM stack on electron beam characteristics [TWTs]', *IEEE Transactions on Electron Devices,* vol. 48, pp. 38–44, 2001.

[33] R. G. Carter and A. MacGregor, 'Effect of imperfections on periodic permanent magnet focusing of electron beams', in *2nd IEEE International Vacuum Electronics Conference*, Noordwijk, The Netherlands, pp. 211–213, 2001.

[34] A. MacGregor, 'Periodic permanent magnet focusing of electron beams', PhD, Engineering Department, Lancaster University, Lancaster, UK, 1986.

[35] R. R. Moats, 'Calculations of beam trajectories under non-ideal conditions of PPM focusing', in *1976 International Electron Devices Meeting*, pp. 515–519, 1976.

[36] B. Minakovic, 'Effect on an electron beam of variations in periodic permanent-magnet focusing systems', *Electrical Communication,* vol. 38, pp. 415–424, 1963.

[37] P. K. Tien, 'Focusing of a long cylindrical electron stream by means of periodic electro-static fields', *Journal of Applied Physics,* vol. 25, pp. 1281–1288, 1954.

[38] N. W. McLachlan, *Bessel Functions for Engineers*. Oxford University Press, 1954.

[39] P. B. Priestland and H. L. Hartnagel, 'Theory of periodic electrostatic focusing of electron beams', *IEEE Transactions on Electron Devices,* vol. 15, pp. 915–935, 1968.

[40] P. B. Priestland and H. L. Hartnagel, 'A further contribution to the theory of electrostatic focusing', *IEEE Transactions on Electron Devices,* vol. 16, pp. 803–812, 1969.

[41] C. C. Cutler, 'Instability in hollow and strip electron beams', *Journal of Applied Physics,* vol. 27, pp. 1028–1029, 1956.

[42] K. T. Nguyen *et al.*, 'Intense sheet electron beam transport in a uniform solenoidal magnetic field', *IEEE Transactions on Electron Devices,* vol. 56, pp. 744–752, 2009.

[43] J. Pasour *et al.*, 'Demonstration of a 100-kW solenoidally focused sheet electron beam for millimeter-wave amplifiers', *IEEE Transactions on Electron Devices,* vol. 58, pp. 1792–1797, 2011.

[44] J. H. Booske *et al.*, 'Stability and confinement of nonrelativistic sheet electron beams with periodic cusped magnetic focusing', *Journal of Applied Physics,* vol. 73, pp. 4140–4155, 1993.

[45] J. H. Booske and M. A. Basten, 'Demonstration via simulation of stable confinement of sheet electron beams using periodic magnetic focusing', *IEEE Transactions on Plasma Science,* vol. 27, pp. 134–135, 1999.

[46] Z. L. Wang *et al.*, 'The conditions for stable sheet electron beams transport in periodic permanent magnet fields', *Journal of Infrared Millimeter and TeraHertz Waves,* vol. 31, pp. 649–658, June 2010.

[47] R. L. Kyhl and H. F. Webster, 'Breakup of hollow cylindrical electron beams', *IRE Transactions on Electron Devices,* vol. 3, pp. 172–183, 1956.

[48] J. R. Pierce, 'Instability of hollow beams', *IRE Transactions on Electron Devices,* vol. 3, pp. 183–189, 1956.

[49] E. K. Demmel, 'Some studies on a high-perveance hollow-beam klystron', *IEEE Transactions on Electron Devices,* vol. 11, pp. 66–73, 1964.

[50] G. Dohler, 'On the stability of hollow beams in PPM focused TWTs', *IEEE Transactions on Electron Devices,* vol. 28, pp. 602–604, 1981.

[51] Y. Y. Lau, 'A unified theory of the diocotron, cyclotron maser, and negative-mass instabilities', *IEEE Transactions on Electron Devices,* vol. ED-31, pp. 329–337, 1984.

[52] A. Szabo, 'Thermal velocity effects in magnetically confined beams', *IRE Transactions on Electron Devices,* vol. 5, pp. 183–185, 1958.

[53] J. L. Palmer and C. Susskind, 'Effects of transverse velocities in magnetically focused cylindrical electron beams', in *Third International Congress on Microwave Tubes*, Munich, pp. 456–460, 1960.

[54] P. T. Kirstein, 'On the effects of thermal velocities in two-dimensional and axially symmetric beams', *IEEE Transactions on Electron Devices,* vol. 10, pp. 69–80, 1963.

[55] J. R. Hechtel, 'Magnetic focusing of electron beams in the presence of transverse velocity components', *IEEE Transactions on Electron Devices,* vol. 28, pp. 473–482, 1981.

[56] G. Herrmann, 'Optical theory of thermal velocity effects in cylindrical electron beams', *Journal of Applied Physics,* vol. 29, pp. 127–136, 1958.

[57] R. L. Jepsen, 'Ion oscillations in electron beam tubes; ion motion and energy transfer', *Proceedings of the IRE,* vol. 45, pp. 1069–1080, 1957.

[58] A. D. Sutherland, 'Relaxation instabilities in high-perveance electron beams', *IRE Transactions on Electron Devices,* vol. 7, pp. 268–273, 1960.

[59] C. B. Thorington, 'Computer simulation of ion trapping and detrapping in a PPM focused traveling wave tube', *IEEE Transactions on Electron Devices,* vol. 48, pp. 56–61, 2001.

[60] W. Tighe *et al.*, 'Transient ion disturbances in traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 48, pp. 82–87, 2001.

[61] R. G. Carter and R. H. C. Newton, 'An experimental study of charge densities and radio-frequency propagation characteristics in an ion beam/plasma system', *J. Phys. D,* vol. 7, pp. 1670–1676, 1974.

8   Electron Flow in Crossed Fields

8.1   Introduction

This chapter reviews the theory of electron motion in crossed fields in both planar and cylindrical geometries. The effects of relativity are neglected because they are not normally important in commercial crossed-field devices. Relativistic magnetrons, which are outside the scope of this book, are discussed in [1].

The equation of motion for a non-relativistic electron moving in electric and magnetic fields is

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{e}{m_0}\left(\mathbf{E} + \frac{d\mathbf{r}}{dt}\times\mathbf{B}\right). \tag{8.1}$$

The force exerted by the magnetic field is always perpendicular both to the direction the magnetic field, and to the motion of the electron, as shown by the vector product. If the term in brackets on the right-hand side of the equation is zero then the electron will move in a straight line with constant velocity. This can be achieved by arranging that the electric field (the sum of externally imposed and space-charge fields), the magnetic field, and the direction of motion of the electron, are mutually perpendicular. This method of controlling the flow of electrons, using crossed fields, provides an alternative to the use of a magnetic field parallel to the electron velocity, discussed in the previous chapter.

Electron flow in crossed fields falls into four categories, as shown in Figure 8.1. It can be seen that the electron motion is essentially two-dimensional, and that the beam is a sheet beam. The beam may be injected from an electron gun, similar to that in a linear beam tube, or emitted from the whole surface of the negative electrode (cathode). The geometry may either be planar, or cylindrical. In injected beam tubes, the potential of the negative electrode (sometimes known as the *sole*) may differ from that of the cathode from which the electrons are emitted (see Section 8.2.2). The positive electrode, incorporating the RF structure with which the electrons interact, is known as the anode (or sometimes the *line*). This chapter examines the conditions necessary to achieve uniform, stable, electron flow in both planar and cylindrical geometries of both kinds. A review of the experimental evidence shows the limitations of the theoretical treatments.
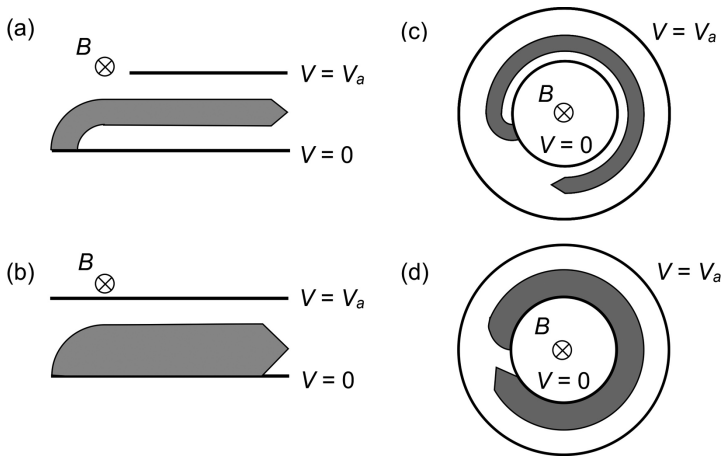
Figure 8.1: Types of crossed-field device: (a) Planar injected beam, (b) Planar emitting cathode, (c) Cylindrical injected beam, and (d) Cylindrical emitting cathode.

Crossed-field tubes include magnetron oscillators (see Chapter 15), and crossed-field amplifiers and oscillators (see Chapter 16). In the older literature they are sometimes referred to as 'M-type' tubes, from their description in French as 'Tubes a Ondes Magnetique'. All tubes which are of continuing importance are of the cylindrical, emitting cathode, type. Planar geometry has only been used in a few experimental devices because cylindrical tubes are smaller, and lighter, for the same performance. Injected beams, which tend to be unstable, were used in some low power devices but are now, largely, obsolete.

8.2   Crossed-Field Electron Flow in Planar Geometry

Electron flow in planar geometry is considered in this section, and the following one, because it provides useful insights into the properties of electrons in crossed-field flow. The cylindrical case, which is of greater practical importance is discussed in Sections 8.4 and 8.5. Figure 8.2 shows a planar magnetron diode immersed in a uniform magnetic field normal to the plane of the paper. An electron emitted from the cathode is accelerated towards the anode by the electric field but the effect of the magnetic field is to bend the trajectory as shown. As the magnetic field is increased, at constant anode voltage, a point is reached where the electron can no longer reach the anode. The current through the diode is then zero and it is said to be *cut-off* or *magnetically insulated*. Alternatively, the current ceases to flow if the anode voltage is reduced, at constant magnetic field. Then the electron just grazes the surface of the anode at the *cut-off potential* V_c.
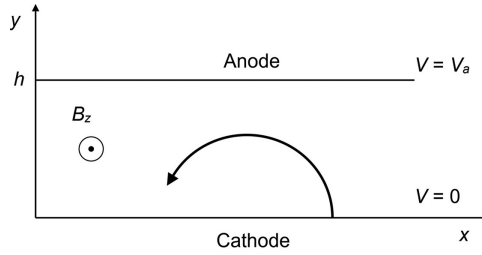
Figure 8.2: Crossed-field electron trajectory in planar geometry.

8.2.1   Electron Motion without Space-Charge

Let us suppose that the magnetic field is in the z direction and that the electron is moving in the (x, y) plane. Then from (8.1) the equations of motion in Cartesian co-ordinates are

$$\frac{d^2 x}{dt^2} = -\frac{e}{m_0}\frac{dy}{dt}B_z = -\omega_c \frac{dy}{dt} \tag{8.2}$$

$$\frac{d^2 y}{dt^2} = -\frac{e}{m_0}E_y + \frac{e}{m_0}\frac{dx}{dt}B_z = \omega_c \left(\frac{dx}{dt} - \frac{E_y}{B_z}\right), \tag{8.3}$$

where $\omega_c = eB_z/m_0$ is the *cyclotron frequency*. It is easy to show by substitution that a solution of these equations, in the absence of an electric field, is

$$\begin{aligned} x &= -r\sin(\omega_c t) \\ y &= r\cos(\omega_c t), \end{aligned} \tag{8.4}$$

when the initial position of the electron is $(x, y) = (0, r)$. Equations (8.4) are just the parametric form of a circle centred on the origin. The initial vector velocity is $(\dot{x}, \dot{y}) = (-r\omega, 0)$.

When a uniform electric field is added in the y direction equation (8.3) shows that the centre of rotation, known as the *guiding centre*, moves with uniform velocity given by

$$u_0 = \frac{E_y}{B_z}. \tag{8.5}$$

By extension of (8.4) the solution of (8.2) and (8.3) is

$$\begin{aligned} x &= -r\sin(\omega_c t) + u_0 t \\ y &= r\cos(\omega_c t), \end{aligned} \tag{8.6}$$

where the initial conditions are as before, and the initial velocity components are

$$\begin{aligned} \dot{x} &= u_0 - \omega_c r, \\ \dot{y} &= 0. \end{aligned} \tag{8.7}$$

The electron trajectories calculated from (8.6) are shown in Figure 8.3 for a number of different initial velocities. It can be seen that the choice of initial position of the electron with respect to the guiding centre (which lies on the x axis) determines the initial velocity and that the motion of the electron is cycloidal. Strictly speaking the trajectory in Figure 8.3(a) is a lengthened cycloid, 8.3(b) is a cycloid and 8.3(c) is a shortened cycloid [2]. When $r = 0$ then the electric and magnetic forces are balanced, $\dot{x} = u_0$, and the electron moves in a straight line parallel to the x axis. The solutions for other initial conditions can be derived from (8.6) by changes of origin (see Worksheet 8.1).

$$r = -1.5, \quad dx/dt = 2.5$$

(a)

$$r = -1.0, \quad dx/dt = 2.0$$

(b)

$$r = -0.5, \quad dx/dt = 1.5$$

(c)
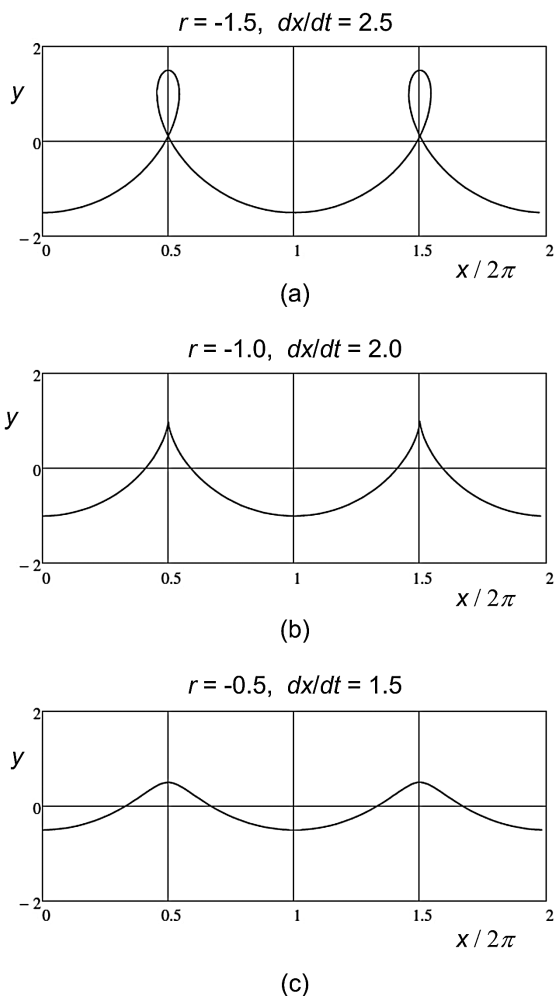
Figure 8.3: Electron trajectories in crossed fields in Cartesian geometry.

Integrating (8.2) with respect to t gives

$$\dot{x} = -\frac{e}{m_0} B_z y = -\omega_c y, \tag{8.8}$$

where it has been assumed that $\dot{x} = 0$ when $y = 0$. It can be seen that $\dot{x}$ depends only on y and B_z. By the principle of conservation of energy the relationship between the electron velocity and the electric potential with respect to the cathode (V) is

$$V = \frac{m_0}{2e}\left(\dot{x}^2 + \dot{y}^2\right). \tag{8.9}$$

Substituting for $\dot{x}$ from (8.8) gives

$$V = \frac{m_0}{2e}\left(\dot{y}^2 + \omega_c^2 y^2\right). \tag{8.10}$$

When $\dot{y} = 0$ the electron is moving parallel to the anode and (8.10) becomes

$$V = \frac{m_0}{2e}\omega_c^2 y^2. \tag{8.11}$$

In particular, when the diode is just cut-off $y = h$, and we obtain the *Hull cut-off voltage* (V_H) for a planar diode [3]

$$V_H = \frac{e}{2m_0} B_z^2 h^2. \tag{8.12}$$

The derivation of this equation is independent of the space-charge density.

In order to use crossed fields to control the flow of electrons in a planar device we envisage an arrangement such as that shown in Figure 8.4. The electrons start from rest at a cathode at zero potential and are injected into a space in which there is a uniform electric field in the negative y direction. The field is provided by a planar anode at potential V_a and a parallel electrode at a lower potential. To achieve crossed-field flow the magnetic field must be in the negative z direction. It is immediately obvious that the electron velocity at $x = 0$ is not constant because the
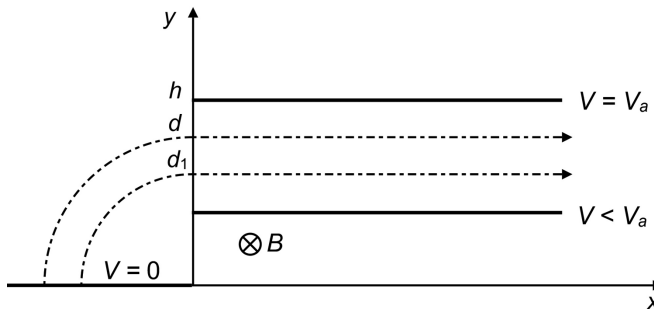
Figure 8.4: Injection of a sheet electron beam into crossed-field flow.

potential varies with y. In the absence of space-charge the electric field is uniform and, from the results of the preceding section, we see that motion in a straight line is only possible for those electrons which satisfy (8.5) and which therefore enter the focusing system at a particular value of y. All other electrons must exhibit cycloidal motion with guiding centre velocity u_0.

8.2.2   Injected Beam with Space-Charge

When space-charge is taken into consideration the electric field is no longer independent of y. We therefore wish to find whether a sheet electron beam can exist in which all quantities are independent of x [4–5]. The x component of the electron velocity is given by (8.8). If the flow is uniform, the electric and magnetic forces must balance, so that

$$E_y = \dot{x} B_z. \tag{8.13}$$

Eliminating $\dot{x}$ between (8.8) and (8.13) we obtain

$$E_y = -\frac{e}{m_0} B_z^2 y. \tag{8.14}$$

But the electric field must also satisfy the one-dimensional form of Gauss' Law

$$\frac{\partial E_y}{\partial y} = \frac{\rho}{\varepsilon_0}. \tag{8.15}$$

Differentiating (8.14) with respect to y and combining it with (8.15) gives

$$-\frac{e\rho}{\varepsilon_0 m_0} = \left(\frac{e}{m_0} B_z\right)^2, \tag{8.16}$$

which may be written using (7.47) as

$$\omega_p^2 = \omega_c^2. \tag{8.17}$$

This is the condition for smooth *Brillouin flow* in a planar crossed-field diode [6]. Note that, since the magnetic field is constant everywhere, (8.16) shows that the charge density must be constant within the electron beam. The electron velocity is not constant but depends upon y. The beam current density is

$$J_x = \rho \dot{x}, \tag{8.18}$$

where the x component of the velocity is obtained from (8.13) and (8.14) as

$$\dot{x} = -\frac{e}{m_0} B_z y, \tag{8.19}$$

hence

$$J_x = -\frac{e\rho}{m_0} B_z y. \tag{8.20}$$

Substituting for the charge density from (8.16) yields

$$J_x = \varepsilon_0 \left(\frac{e}{m_0}\right)^2 B_z^3 y. \tag{8.21}$$

Slater has pointed out that the current density can be substantial for quite modest values of the constants [4]. For example, if $B_z = 0.1\,\mathrm{T}$ and $y = 10\,\mathrm{mm}$, then $J_x = 277\,\mathrm{A\ cm^{-2}}$.

To find the electrostatic potential within the beam we integrate (8.14) with respect to y to give

$$V = \frac{e}{2m_0} B_z^2 y^2 + \text{const.} \tag{8.22}$$

This is identical to (8.11), which applies to all electrons moving parallel to the anode, and therefore the constant of integration is zero. Outside the beam the charge density is zero and, therefore, from (8.15) the electric field is constant. Now the electric field is continuous at the edges of the beam so that

$$E_y = -\frac{e}{m_0} B_z^2 d_1 \quad (0 \le y \le d_1) \tag{8.23}$$

and

$$E_y = -\frac{e}{m_0} B_z^2 d \quad (d \le y \le h). \tag{8.24}$$

Therefore the potential of the anode is given by

$$V_a = \frac{e}{2m_0} B_z^2 d^2 + \frac{e}{m_0} B_z^2 d (h - d)$$

$$= \frac{e}{2m_0} B_z^2 d^2 \left(\frac{2h}{d} - 1\right). \tag{8.25}$$

We note that this expression is independent of d_1 and that it therefore applies to all cases including $d_1 = 0$. Thus it is theoretically possible to inject an electron beam into uniform flow in crossed electric and magnetic fields. The method by which this can be accomplished is discussed in the Section 9.6.1. In order for the electron velocity to be close to uniform it is necessary for the beam to be thin. However, it is found that a beam of this kind suffers from an instability known as the diocotron instability [7–8]. Injected beam crossed-field devices are therefore not common, and tend to be of low power.

The theory in this section can also be applied to devices with an emitting cathode (see Figure 8.1(b)). The potential variation within the beam given by (8.22) with the constant set to zero so that $V = 0$ when $y = 0$. Comparing (8.22) with (8.12) we see that the potential on the upper surface of the beam is equal to the Hull cut-off voltage. Thus it is apparently possible for Brillouin flow to exist in a planar crossed-field device with a space-charge layer close to the cathode. This raises the question of how this condition can be set up, because the electrons emitted from the cathode do not experience any force in the y direction.

8.3   The Planar Magnetron Diode

In a planar magnetron diode the whole of the zero-potential electrode forms the cathode as illustrated in Figure 8.1(b). This resembles the planar space-charge limited diode, discussed in Chapter 5, with the addition of a uniform transverse magnetic field. We will assume that the problem is one-dimensional so that all quantities depend only upon y. The properties of the diode should tend to the Child–Langmuir solution as the magnetic field tends to zero. The difference is that the magnetron diode becomes cut-off for sufficiently high magnetic fields. Two cases must therefore be examined depending on whether the diode is conducting or cut off.

8.3.1   The Diode Is Conducting

Poisson's equation in one dimension is

$$\frac{d^2V}{dy^2} = -\frac{\rho}{\varepsilon_0}, \tag{8.26}$$

where the charge density ρ is related to the y component of the current density by

$$J_y = \rho \dot{y}, \tag{8.27}$$

provided that $\dot{y} \geq 0$ everywhere. Because all quantities only vary with y it follows from the continuity equation that J_y is constant. Substituting for V from (8.10), and for ρ from (8.27), in (8.26)

$$\dot{y}\frac{d}{dy}\left(\dot{y}\frac{d\dot{y}}{dy}\right) + \omega_c^2 \dot{y} = -\frac{eJ_y}{\varepsilon_0 m_0}, \tag{8.28}$$

which may be written

$$\frac{d^2\dot{y}}{dt^2} + \omega_c^2 \dot{y} = -\frac{eJ_y}{\varepsilon_0 m_0}. \tag{8.29}$$

This equation has the particular integral

$$\dot{y} = -\frac{eJ_y}{\varepsilon_0 m_0 \omega_c^2} \tag{8.30}$$

and the complementary function

$$\dot{y} = A\sin(\omega_c t) + B\cos(\omega_c t), \tag{8.31}$$

where A and B are constants. Thus the general solution of (8.29) is

$$\dot{y} = A\sin(\omega_c t) + B\cos(\omega_c t) - \frac{eJ_y}{\varepsilon_0 m_0 \omega_c^2}. \tag{8.32}$$

Now $\dot{y} = 0$ for an electron which leaves the cathode when $t = 0$ so

$$B = \frac{eJ_y}{\varepsilon_0 m_0 \omega_c^2} \tag{8.33}$$

and

$$\dot{y} = \frac{eJ_y}{\varepsilon_0 m_0 \omega_c^2}(\cos(\omega_c t) - 1) + A\sin(\omega_c t). \tag{8.34}$$

For space-charge-limited flow the normal component of the electric field is zero at the surface of the cathode so that $\ddot{y} = 0$ when $t = 0$. Then $A = 0$ and

$$\dot{y} = \frac{eJ_y}{\varepsilon_0 m_0 \omega_c^2}(\cos(\omega_c t) - 1). \tag{8.35}$$

Integrating (8.35) with respect to t we obtain

$$y = \frac{eJ_y}{\varepsilon_0 m_0 \omega_c^3}(\sin(\omega_c t) - \omega_c t), \tag{8.36}$$

where $y = 0$ when $t = 0$. In the limit $B_z \to 0$ the trigonometric functions in (8.35) and (8.36) can be expanded as power series to give

$$y \to \frac{e(-J_y)}{6\varepsilon_0 m_0} t^3 \tag{8.37}$$

and

$$\dot{y} \to \frac{e(-J_y)}{2\varepsilon_0 m_0} t^2. \tag{8.38}$$

Substituting for t in (8.38) and then for $\dot{y}$ in (8.10) we recover the Child–Langmuir Law (5.16) which can be written in the form

$$\frac{eJ_C}{\varepsilon_0 m_0 \omega_c^3 h} = -\frac{2}{9}\left(\frac{2e}{m_0} \cdot \frac{V_a}{\omega_c^2 h^2}\right)^{\frac{3}{2}}, \qquad (8.39)$$

where J_C is the Child–Langmuir current density in a planar diode in the absence of a magnetic field and h is the distance between the cathode and the anode. It is not possible to find an explicit solution of this kind for the magnetron diode. However, eliminating J_y from (8.35) using (8.36), gives

$$\frac{\dot{y}}{\omega_c y} = \frac{1-\cos\left(\omega_c t\right)}{\omega_c t - \sin\left(\omega_c t\right)} \qquad (8.40)$$

and, substituting for $\left(\dot{y}/\omega_c y\right)$ from (8.10), we find that

$$\frac{2e}{m_0} \cdot \frac{V_a}{\omega_c^2 h^2} = \left(\frac{1-\cos\left(\omega_c t_a\right)}{\omega_c t_a - \sin\left(\omega_c t_a\right)}\right)^2 + 1, \qquad (8.41)$$

where t_a is the transit time. Also, from (8.36) we find that

$$\frac{eJ_y}{\varepsilon_0 m_0 \omega_c^3 h} = -\frac{1}{\omega_c t_a - \sin\left(\omega_c t_a\right)}. \qquad (8.42)$$

Thus, from (8.42), (8.39), and (8.41) we can compute the ratio J_y/J_C as a function of the normalised transit time $\omega_c t_a$. At the Hull cut-off condition $\dot{y} = 0$ so that, from (8.40), $\omega_c t_a = 2\pi$ when $y = h$ and $J_y/J_C \to 9/4\pi = 0.716$.

The ratio of the magnetic field to the cut-off field $\left(B_H\right)$ at constant anode potential $\left(V_a\right)$ can be written, using (8.12), as

$$\frac{B^2}{B_H^2} = \left(\frac{2eV_a}{m_0 \omega_c^2 h^2}\right)^{-1}, \qquad (8.43)$$

which can be computed as a function of the normalised transit angle. We can therefore plot the normalised current against the normalised magnetic field as shown in Figure 8.5. At cut-off the curve has infinite slope as the current drops abruptly to zero [9] (see Worksheet 8.2).

It is also interesting to enquire how the current varies with anode voltage at constant magnetic field. From (8.10)

$$\frac{V_a}{V_H} = 1 + \left(\frac{\dot{y}_a}{\omega_c h}\right)^2, \qquad (8.44)$$

so that, using (8.40) we can plot the normalised current against the normalised voltage, with $\omega_c t_a$ as a parameter, as shown in Figure 8.6. The current is zero until
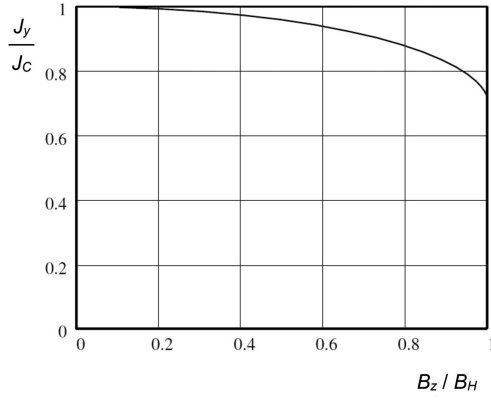
Figure 8.5: Variation of current with magnetic field in a planar crossed-field diode (copyright 1993, IEEE, reproduced, with permission, from [9]).
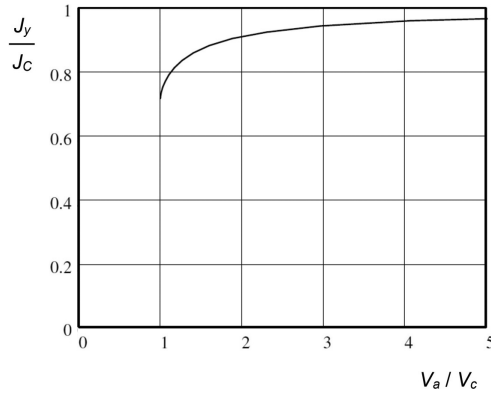
Figure 8.6: Variation of current with applied voltage in a planar crossed-field diode.

the anode voltage rises above the cut-off voltage. It then increases with increasing anode voltage and tends to the Child-Langmuir limit.

To find an electron trajectory we substitute for y in (8.8) using (8.36) to give

$$\dot{x} = -\frac{eJ_y}{\varepsilon_0 m_0 \omega_c^2}\left(\sin\left(\omega_c t\right) - \omega_c t\right). \qquad (8.45)$$

Integrating with respect to t gives

$$x = \frac{eJ_y}{\varepsilon_0 m_0 \omega_c^3}\left(\cos\left(\omega_c t\right) - 1 + \frac{1}{2}\omega_c^2 t^2\right), \qquad (8.46)$$
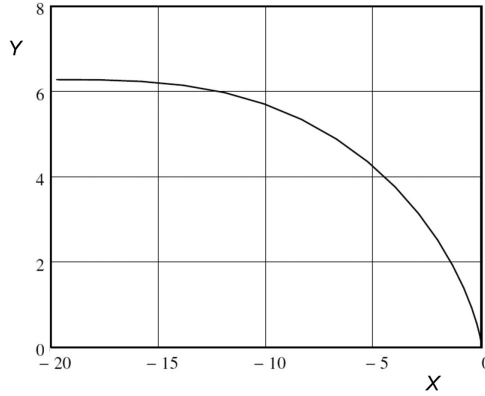
Figure 8.7: Normalised electron trajectory in a space-charge limited magnetron diode which is conducting.

where the constant of integration has been chosen so that the initial position of the electron is at the origin. Normalised x and y co-ordinates can be defined by

$$(X,Y) = \frac{\varepsilon_0 m_0 \omega_c^3}{e J_y}(x,y). \tag{8.47}$$

Figure 8.7 shows the normalised electron trajectory plotted using (8.36) and (8.46) with normalised transit time $(\omega_c t)$ as a parameter. At the maximum transit time the diode is just cut off and the trajectory is parallel to the x axis (see Worksheet 8.2).

8.3.2   The Diode Is Cut-Off

When the diode is cut-off the net current in the y direction is zero. It is commonly assumed that the electrons are in Brillouin flow so that the electrons flow parallel to the cathode, as discussed in Section 8.2.2. There is then a space-charge layer in the range $0 \le y \le d$, where $d < h$, so that no electrons are able to reach the upper electrode and the diode is cut off. Eliminating B_z between (8.25) and (8.12) produces an expression for the position of the surface of the space-charge layer within the diode in terms of the cut-off voltage, and the anode voltage

$$\frac{V_a}{V_H} = \frac{d}{h}\left(2 - \frac{d}{h}\right). \tag{8.48}$$

Equation (8.48) describes a parabola which passes through the origin and has a maximum when $d = h$. The equation can be rearranged to give the dependence of the height of the electron layer on the anode voltage

$$\frac{d}{h} = 1 - \sqrt{1 - \frac{V_a}{V_H}}. \tag{8.49}$$

From equation (8.12) we know that, for fixed h, the cut-off voltage increases with increasing magnetic field. It follows that the effect of increasing the magnetic field at constant anode voltage is to reduce d. All practical crossed field devices are operated below cut-off with a ratio d/h of the order of 0.5.

Brillouin flow is not the only possible solution to this problem. If it is supposed that the motion of the electrons is cycloidal, as in the space-charge-free case, then alternative solutions exist. The electron flow is treated as the superposition of electron currents having equal and opposite y components J_+ and J_- [4]. This is known as *two-stream flow*. Since the space-charge density is independent of the direction of motion of the electrons, each current component must be half the current in a conducting diode. Thus (8.28) becomes

$$\dot{y}\frac{d}{dy}\left(\dot{y}\frac{d\dot{y}}{dy}\right) + \omega_c^2 \dot{y} = -\frac{2eJ_+}{\varepsilon_0 m_0},\tag{8.50}$$

where J_+ is y component of the outward flow of current. The solution to the problem can then be derived directly from that in the previous section by replacing J_y by $2J_+$ in (8.34) to give

$$\dot{y} = \frac{2eJ_+}{\varepsilon_0 m_0 \omega_c^2}\left(\cos\left(\omega_c t\right)-1\right) + A\sin\left(\omega_c t\right),\tag{8.51}$$

where we have, again, assumed that the electron starts from rest on the cathode surface. This type of solution was first discussed by Slater [4] who assumed space-charge limited emission from the cathode. Then the normal component of the electric field is zero at the cathode surface and $A = 0$. The normalised (Slater) electron orbits can then be obtained from (8.47) with the substitution of $2J_+$ for J_y as shown in Figure 8.8. The points of inflexion of the curve are branching points at which the trajectory may continue to move upwards, or follow a downwards path, which is the mirror image of the initial upward path. There is an infinity of such possible Slater orbits corresponding to transit angles from the cathode to the anode of
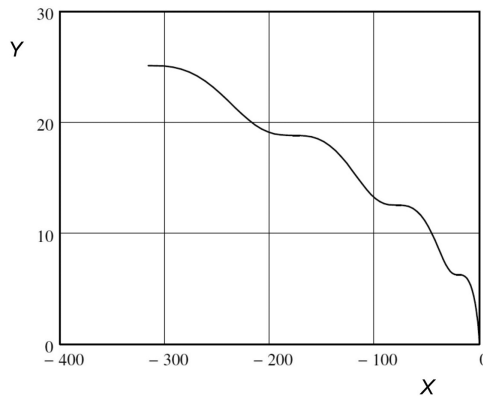
Figure 8.8: Two-stream electron orbit in a cut-off planar magnetron diode.
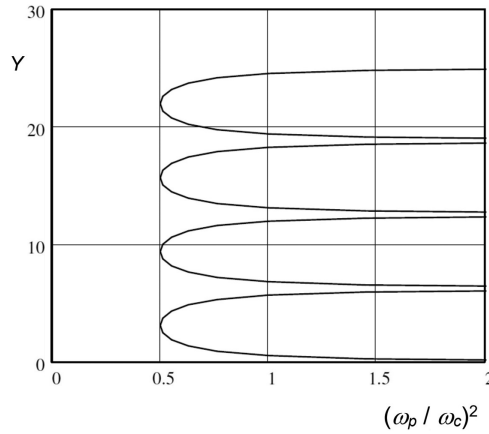
Figure 8.9: Dependence of normalised charge density $\left(\omega_p^2/\omega_c^2\right)$ on normalised position in a cut off planar magnetron diode.

2π, 4π, …$2n\pi$,…, where n is a positive integer. In each case the maximum value of y is given by (8.12) so the trajectories become flatter with increasing n.

The charge density is found by substituting for $\dot{y}$ in (8.27) using (8.35) and the result can be written as

$$\omega_p^2 = \frac{\omega_c^2}{1-\cos(\omega_c t)},\tag{8.52}$$

which can be compared with (8.17) for Brillouin flow. Using $\omega_c t$ as a parameter in (8.36) we can plot the variation in normalised charge density $\left(\omega_p^2/\omega_c^2\right)$ with Y, for two-stream flow, as shown in Figure 8.9. The charge is found to be in layers (described as *striations* by Slater [4]) with a minimum charge density half that of Brillouin flow. The average charge density, found by integrating ρ with respect to y over a whole number of layers, is identical to that for Brillouin flow. It follows that the electric field between the top of the space-charge layer and the anode is the same in both cases. The fact that the charge density is infinite close to the cathode resembles the situation in a space-charge limited diode where it is necessary to suppose that the thermal velocities of the electrons allow them to diffuse away from the cathode.

The current flowing from the cathode can be calculated in the same way as for a conducting diode. For the first Slater orbit this current is just one half of that for a diode which is just conducting (see Figure 8.5). The results for the first few orbits are given in Table 8.1. Slater suggested that Brillouin flow might be seen as the limiting case when $n \to \infty$ and the current emitted from the cathode falls to zero. Thus, an infinite set of theoretical solutions exists for the electron flow in a cut-off planar magnetron diode, and it is not apparent which of these is correct in any particular case. We shall see in Section 8.5 that the problems raised by two-stream flow are resolved when cylindrical geometry is considered.

Table 8.1: Ratio of the injected current to the
Child–Langmuir current for the first few Slater orbits

n	J_+ / J_C
1	0.358
2	0.179
3	0.119
4	0.090
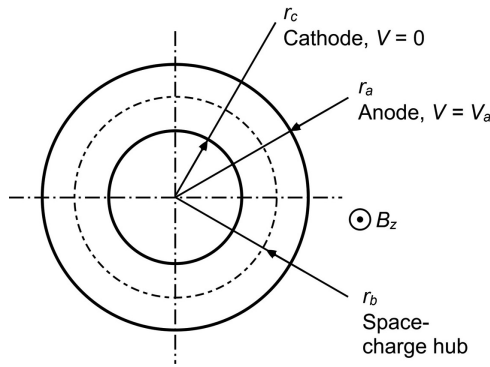
Figure 8.10: Arrangement of a cylindrical magnetron diode.

8.4   Crossed-Field Electron Flow in Cylindrical Geometry

Figure 8.10 shows the arrangement of a cylindrical crossed-field diode in which the anode lies outside the cathode. The alternative arrangement with the anode inside the cathode is not much used and will not be discussed further here. When the diode is cut off it is assumed that the space-charge is confined into a cylindrical layer (or hub) surrounding the cathode.

8.4.1   Electron Motion without Space-Charge

The analysis of a cylindrical, crossed-field diode, in which the outer electrode is the anode, proceeds in the same manner as that for a planar diode. The non-relativistic equations of motion of the electrons are (7.11) and (7.12)

$$\frac{d^2 r}{dt^2} = -\frac{e}{m_0} E_r + r\dot{\theta}^2 - \frac{e}{m_0} r\dot{\theta} B_z \tag{8.53}$$

and

$$\frac{d}{dt}\left(r^2 \dot{\theta}\right) = \frac{e}{m_0} r\dot{r} B_z. \tag{8.54}$$

This equation can be integrated with respect to time to give

$$m_0 r^2 \dot{\theta} = \frac{1}{2} e B_z r^2 + \text{const.} \tag{8.55}$$

If it is assumed that the electrons have zero angular velocity as they leave the cathode then

$$\dot{\theta} = \frac{1}{2} \omega_c \left(1 - \frac{r_c^2}{r^2}\right), \tag{8.56}$$

where r_c is the cathode radius. Using the principle of conservation of energy, the relationship between the electron velocity and the electric potential with respect to the cathode is

$$V = \frac{m_0}{2e} \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right). \tag{8.57}$$

Substituting for $\dot{\theta}$ from (8.56) gives

$$V = \frac{m_0}{2e} \left(\dot{r}^2 + \frac{1}{4} r^2 \omega_c^2 \left(1 - \frac{r_c^2}{r^2}\right)^2 \right). \tag{8.58}$$

When the diode is just cut off the electrons have zero radial velocity as they reach the anode. Then the relationship between the anode potential at cut-off (V_H) and the magnetic field is

$$V_H = \frac{e}{8 m_0} B_z^2 r_a^2 \left(1 - \frac{r_c^2}{r_a^2}\right)^2. \tag{8.59}$$

This is the Hull cut-off equation for a cylindrical diode (compare (8.12) for the planar diode). Note that the derivation of this equation is similarly unchanged by the presence of space charge.

To find the electron trajectories we note that, in the absence of space-charge,

$$E_r = -\frac{V_a}{r \ln(r_a / r_c)}. \tag{8.60}$$

Substituting for E_r and for $\dot{\theta}$ from (8.56) in (8.53) gives

$$\frac{d^2 r}{dt^2} = \frac{e}{m_0} \cdot \frac{V_a}{r \ln(r_a / r_c)} + \frac{1}{4} r \omega_c^2 \left(\frac{r_c^4}{r^4} - 1\right). \tag{8.61}$$

Now, if we define the normalised variables $R = r/r_c$, $A = r_a/r_c$, $\alpha = \omega_c t$ and $U = V_a/V_H$, then (8.61) and (8.56) can be written

$$\frac{d^2 R}{d\alpha^2} = \frac{1}{8} \frac{A^2 U}{\ln(A)} \left(1 - \frac{1}{A^2}\right)^2 \frac{1}{R} + \frac{1}{4} R \left\{\frac{1}{R^4} - 1\right\} \tag{8.62}$$
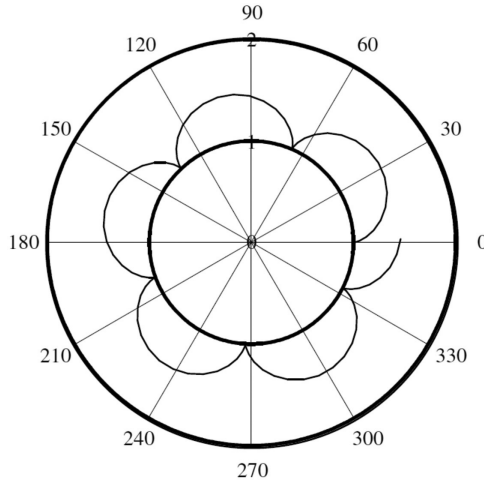
Figure 8.11: Typical electron trajectory in a cylindrical magnetron without space-charge $(A = 2, U = 0.5)$.

and

$$\frac{d\theta}{d\alpha} = \frac{1}{2}\left(1 - \frac{1}{R^2}\right).$$

(8.63)

This pair of equations can be integrated numerically to find the trajectories of the electrons. The trajectories are epicycloidal in the counter-clockwise direction as shown in Figure 8.11 (see Worksheet 8.1).

8.4.2   Injected Beam with Space-Charge

We wish to find out whether Brillouin flow, in which all quantities are independent of θ, can exist in the space between the two electrodes, as shown in Figure 8.10. If the flow is uniform the radial acceleration of the electrons must be zero. Therefore, from (8.53)

$$E_r = \frac{m_0}{e} r\dot{\theta}^2 - r\dot{\theta}B_z.$$

(8.64)

Substituting for the angular velocity from (8.56) and collecting terms gives

$$E_r(r) = -\frac{e}{4m_0}B_z^2 r\left(1 - \frac{r_c^4}{r^4}\right).$$

(8.65)

Now Gauss' Law in cylindrical co-ordinates is

$$\frac{1}{r}\frac{\partial}{\partial r}(rE_r) = \frac{\rho}{\varepsilon_0}.\tag{8.66}$$

Substitution of E_r from (8.65) into (8.66) gives

$$\omega_p^2 = \frac{1}{2}\omega_c^2\left(1+\frac{r_c^4}{r^4}\right).\tag{8.67}$$

Comparison between (8.67) and the corresponding equation for a planar magnetron (8.17) shows that they are similar but that the charge density is not uniform in the cylindrical diode.

To find the electrostatic potential within the space-charge layer (8.65) is integrated with respect to r to give

$$V = \frac{e}{8m_0}B_z^2 r^2\left(1-\frac{r_c^2}{r^2}\right)^2,\tag{8.68}$$

since $V = 0$ when $r = r_c$. We note that when $r = r_a$ this is identical to the Hull cut-off equation (8.59). The radial component of the electric field on the surface of the space-charge layer at $r = r_b$ is obtained from (8.65) as

$$E_r = -\frac{e}{4m_0}B_z^2 r_b\left(1-\frac{r_c^4}{r_b^4}\right).\tag{8.69}$$

Now, the electric field is continuous at $r = r_b$ and $\nabla.\mathbf{E} = 0$ in the space between the surface of the space-charge layer and the anode. The electric field in this region is, therefore,

$$E_r = E_r(r_b)\frac{r_b}{r} \quad (r_b \le r \le r_a),\tag{8.70}$$

where $E_r(r_b)$ is obtained from (8.65). The potential difference between the surface of the space-charge layer and the anode is

$$V_{ab} = -\int_{r_b}^{r_a} E_r(r_b)\frac{r_b}{r}\,dr = -E_r(r_b)r_b\ln\left(\frac{r_a}{r_b}\right).\tag{8.71}$$

The potential of the anode is, therefore,

$$V_a = \frac{e}{4m_0}B_z^2 r_b^2\left(1-\frac{r_c^4}{r_b^4}\right)\cdot\ln\left(\frac{r_a}{r_b}\right) + \frac{e}{8m_0}B_z^2 r_b^2\left(1-\frac{r_c^2}{r_b^2}\right)^2.\tag{8.72}$$

This equation is equivalent to equation (8.25) for the planar diode. Because that equation is so much simpler than (8.72) it is commonly used instead with the substitutions $d = r_b - r_c$ and $h = r_a - r_c$. When the results of this substitution are examined it is found that the difference between the two equations is up to 10% when

$r_a/r_c = 1.1$ and that it increases as r_a/r_c increases. However, $r_a/r_c > 1.1$ in most crossed-field devices, so the use of the equation for planar geometry, as an approximation, is hardly ever valid.

8.5   The Cylindrical Magnetron Diode

The behaviour of a cylindrical magnetron diode close to cut-off resembles that of the planar diode described in Section 8.3, but with some important differences. This problem was studied in [4, 10–11]. The equation of motion (8.53) can be written using (8.56) in terms of normalised parameters as

$$\frac{d^2 R}{d\tau^2} = \frac{1}{R^3} - R + E'_r, \tag{8.73}$$

where $R = r/r_c$, $\tau = \omega_L t$, $\omega_L = \omega_c/2$ is the Larmor frequency, and the normalised electric field E'_r is defined by

$$E'_r = \frac{e}{m_0} \frac{E_r}{r_c \omega_L^2}. \tag{8.74}$$

When the diode is cut off there are equal and opposite currents J_+ and J_- per unit surface area of the cathode. Gauss' Law gives

$$\frac{1}{r} \frac{d}{dr}(r E_r) = \frac{\rho}{\varepsilon_0}, \tag{8.75}$$

where, as in Section 8.3.2, half of the charge density is associated with each of the components of the current. Equation (8.75) can be written

$$\frac{d}{dR}(R E_r) = \frac{\rho r_c}{\varepsilon_0} R. \tag{8.76}$$

The continuity equation is

$$\frac{r_c}{r} J_+ = \frac{1}{2} \rho \frac{dr}{dt}, \tag{8.77}$$

which can be written as

$$J_+ = \frac{1}{2} \rho R r_c \omega_L \frac{dR}{d\tau}. \tag{8.78}$$

Combining (8.76) and (8.78) gives

$$\frac{d}{dR}\left(RE_r\right) = \frac{2J_+}{\varepsilon_0 \omega_L u}, \tag{8.79}$$

where $u = dR/d\tau$, or

$$\frac{d}{d\tau}\left(RE_r\right) = \frac{2J_+}{\varepsilon_0 \omega_L}. \tag{8.80}$$

Then, integrating (8.79) along a trajectory

$$RE_r - E_c = \frac{2J_+}{\varepsilon_0 \omega_L} \int_{R=1}^{R} d\tau, \tag{8.81}$$

where E_c is the radial electric field on the surface of the cathode. Now, using the normalised electric field defined in (8.74), equation (8.81) becomes

$$RE_r' - E_c' = J_+' \tau, \tag{8.82}$$

where J_+' is the normalised cathode current density given by

$$J_+' = -\frac{2(e/m_0)J_+}{\varepsilon_0 \omega_L^3 r_c}. \tag{8.83}$$

Then, substituting for E_r' from (8.82), equation (8.73) becomes

$$\frac{d^2 R}{d\tau^2} = \frac{1}{R^3} - R + \frac{E_c'}{R} + \frac{J_+' \tau}{R}. \tag{8.84}$$

This equation can be solved numerically to give J_+' as a function of E_c' and $R_a = r_a/r_c$. It is found that, if the normalised field at the cathode $E_c' = 0$, equation
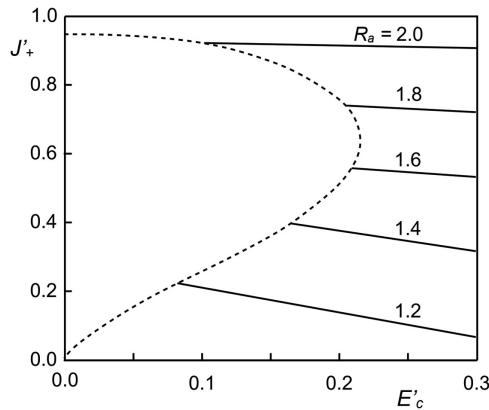
Figure 8.12: The relationship between the cathode current density and the electric field on the surface of the cathode in a cylindrical magnetron diode (copyright 1973, IEE, reproduced, with permission, from [10]).

(8.84) has no solutions when $R_a < 2.01$. This is the condition which normally exists in magnetron oscillators. Figure 8.12 shows the solutions of (8.84) for first-order orbits, that is, those where the radial component of the velocity is only zero on the surface of the anode (see Worksheet 8.3). Higher-order solutions can exist to the left of the dashed envelope curve but these exist only for values of R_a greater than those usually found in magnetron oscillators [10]. This analysis suggests that the electrons in the space-charge layer of a cut-off cylindrical diode follow first-order Slater orbits, and that the electric field on the surface of the cathode is greater than zero. Numerical modelling of magnetrons supports this conclusion [12]. We saw, in Section 8.3.2, that the average charge density is the same for both zero-order, and first-order, space-charge layers in a planar diode. A comparison between the charge per unit length in the space-charge hub of a cylindrical diode shows that they are the same to within ±5% in the range $1.2 \le r_a/r_c \le 2.0$ (see Worksheet 8.3).

The current in a smooth bore magnetron diode is provided by thermionic emission. In a magnetron oscillator there is usually a combination of thermionic and secondary electron emission. In either case, provided that the available current is great enough, we expect that the operating point lies on the dashed envelope shown in Figure 8.12. It is interesting to compare the current density injected from the cathode with that in a diode obeying the Child–Langmuir Law where the current per unit length is given by (5.69)

$$ I_C = \frac{8\pi\varepsilon_0 \sqrt{2\left(e/m_0\right)}}{9} \cdot \frac{V_a^{\frac{3}{2}}}{\left(r_a\beta^2\right)}, \tag{8.85} $$

where β is a function of R_a. Using the same normalisations as before, and noting that the anode voltage is given by (8.68), we obtain

$$ J_C' = -\frac{4}{9} \cdot \frac{R_a^2}{\beta^2}\left(1 - \frac{1}{R_a^2}\right)^3. \tag{8.86} $$
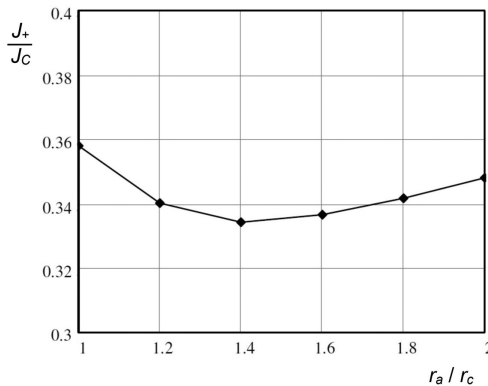
Figure 8.13: Ratio of the injected current, for first-order Slater orbits, to the Child–Langmuir current in a cylindrical magnetron diode which is just cut off.

Figure 8.13 shows the ratio J_+/J_C, plotted against the normalised anode radius, for a diode which is just cut off (see Worksheet 8.3). When $R_a \rightarrow 1$ the current ratio tends to the value 0.358 obtained for a planar diode (see Section 8.3.2). It can be seen that the current ratio varies by less than 10% over the whole range of R_a which has been plotted. If the diode is more strongly cut off then R_a is replaced by the normalised radius of the space-charge layer R_b. When the diode is just conducting then the current is doubled, as it is in a planar diode, with results very similar to those in Figure 8.6 [11].

Early numerical simulations by Hartree suggested that two-stream flow would tend to Brillouin flow within a short time [7]. More recent simulations show that two-stream flow is unstable, when subject to small perturbations, leading to a flow similar to Brillouin flow but with some turbulence [13]. Agafonov showed that Brillouin flow could be seen as the limiting case of two-stream flow [14]. These conclusions, however, cannot explain key features of the experimental behaviour of cut-off magnetron diodes, which are reviewed in the next section. We shall see that the experimental results mean that the principle of conservation of energy cannot apply to individual electrons and, therefore, that neither the Brillouin nor the two-stream model can be correct. However, it should be noted that these models are still of value in throwing light on the problem and also because the results are found to be useful for modelling magnetron oscillators and emitting sole CFAs (see Chapters 15 and 16).

8.6   Experimental Behaviour of Magnetron Diodes

When a magnetron diode is conducting the current is space-charge-limited, provided that the cathode emission is sufficient, and it decreases as cut-off is approached, as predicted by theory [9, 11]. When the diode is cut-off, however, it is found that there are major effects which cannot be explained by the theories discussed in the
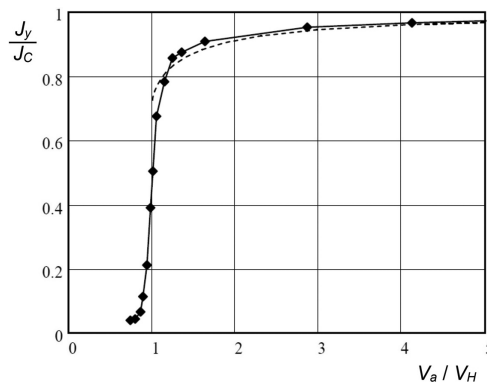
Figure 8.14: Dependence of the current in a magnetron diode on the applied voltage, with the curve from Figure 8.6 (dashed) for comparison.
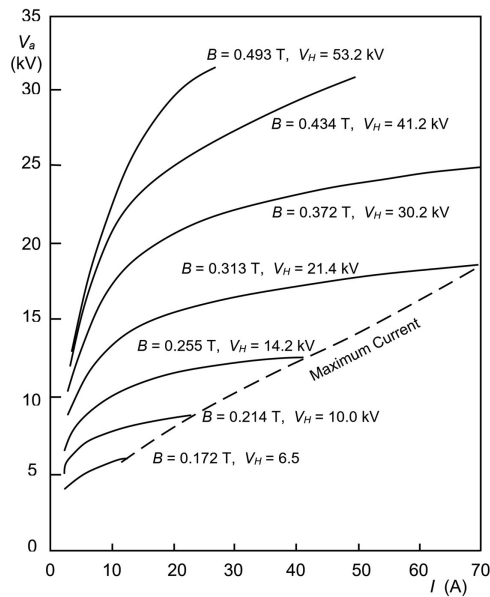
Figure 8.15: Characteristic curves of a smooth-bore magnetron diode with a platinum cathode $(r_a/r_c = 1.58)$
(copyright 1951, AIP Publishing, reproduced, with permission, from [15]).

previous sections. Very early in the study of magnetron diodes it was noticed that the cut-off characteristic was not sharp as predicted by the theories reviewed above [3]. Appreciable current continues to flow below the cut-off voltage, as shown in Figure 8.14, for a diode in which $r_a/r_c = 2.3$ [11]. It was found that this effect cannot be explained by the thermal velocities of the electrons.

All the early experiments employed diodes in which the cathode was a thin wire and the anode voltage was a few hundred volts. These conditions are not typical of the dimensions and voltages used in practical tubes. An important series of experiments was carried out by Jepsen and Muller who observed that the currents in magnetron oscillators were often greater, by up to two orders of magnitude, than could be sustained by thermionic emission from the cathode [15]. It was thought that this was the result of secondary electron emission from the cathode. They constructed a series of experimental magnetron diodes with smooth anodes whose dimensions corresponded to those of an existing magnetron oscillator. In these diodes the cathode was not heated but formed of one of a number of pure materials having different secondary electron emission properties. The temperature of the cathode was not sufficient to allow any appreciable thermionic emission. A small thermionic priming cathode was included at one end of the main cathode. The results were displayed in the form of curves of anode voltage against anode current at constant magnetic field, as shown in Figure 8.15. These curves are very similar to those for

an operating magnetron oscillator having similar dimensions. They curves reveal a number of important features:

- The anode current was tens of amps, even though the anode voltage was always less than the cut off voltage for the magnetic field used. This current was much greater than that which could be supplied by the priming cathode, except when the main cathode was made of carbon (whose secondary electron emission coefficient is less than unity). The current was always much smaller than the space-charge-limited current which would flow in the absence of the magnetic field.
- For each magnetic field it was found that, as the applied voltage was increased, the current increased until it reached a maximum. If the voltage was increased further then the current fell abruptly to approximately that available from the thermionic priming cathode. This is a particular feature of tubes with pure metal cathodes. The maximum current boundary was approximately a straight line and depended on the dimensions of the tube and on the material of the cathode.
- The flow of current to the anode was accompanied by heating of the main cathode caused by back-bombardment.

Figure 8.16 shows typical results for the dependence of the back-bombardment power, normalised to the power input to the diode, on the anode current. At low currents a high proportion of the input power is dissipated on the cathode, but that this falls as the anode current is increased. It was noted that the back-bombardment power was greater in the smooth-bore magnetron than in the equivalent magnetron oscillator by a factor of at least 2. The results shown in Figure 8.16 were obtained using a tube in which the ratio of anode diameter to cathode diameter was 1.58. It was found that the back-bombardment power decreased rapidly as this ratio was increased.

Measurements were made of the currents in the priming cathode, the main cathode, and the anode as a function of anode voltage, with the results shown in
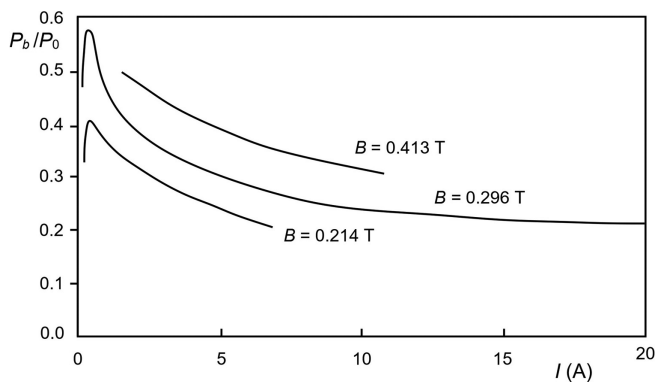
Figure 8.16: Back-bombardment power in a smooth-bore magnetron with a molybdenum cathode $(r_a/r_c = 1.58)$.
(copyright 1951, AIP Publishing, reproduced, with permission, from [15]).
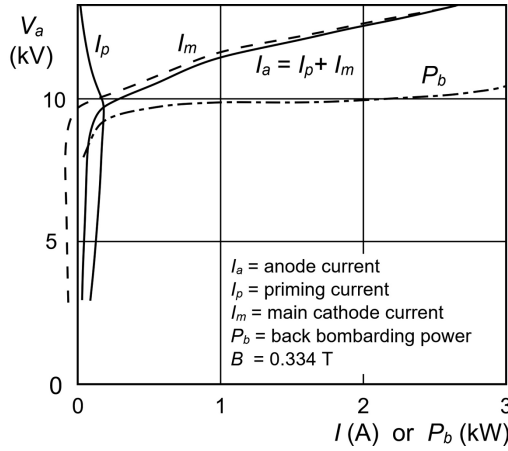
Figure 8.17: Division of current between the priming cathode and the main cathode of a smooth-bore magnetron diode.
(copyright 1951, AIP Publishing, reproduced, with permission, from [15]).

Figure 8.17. It can be seen that up to a critical anode voltage all the currents are small, and the main cathode current is negative, indicating that more electrons are landing on the cathode than are emitted by it. As the voltage is increased a point is reached at which the main cathode current changes sign and then increases rapidly. The zero of the main cathode current was interpreted as being the point at which the energy of the bombarding electrons corresponded to a secondary electron emission coefficient of unity. Thus, at higher voltages, the greater part of the anode current is from secondary electron emission.

The maximum current boundary was explained qualitatively by observing that the total current emitted from the cathodes must always be greater than, or equal to, the anode current, that is

$$I_p + I_b\left(\delta_e - 1\right) \geq I_a,\tag{8.87}$$

where I_p, I_b and I_a are, respectively, the priming cathode current, the back-bombardment current, and the anode current. The effective secondary electron emission coefficient $\left(\delta_e\right)$ takes account of the angles of incidence of the bombarding electrons. The energy of the bombarding electrons increases as the anode voltage increases. However, as shown in Section 18.3, the secondary electron emission coefficient decreases with increasing primary electron energy, and eventually falls below unity. Thus, at some point, the inequality in (8.87) can no longer be satisfied, as the anode voltage is increased, and the anode current must fall to around that emitted by the priming cathode.

The work of Jepsen and Muller showed conclusively that the simple theory of the magnetron diode was inaccurate. It was noted, on the basis of a very rough calculation, that electron scattering based on collisions failed to explain the observations
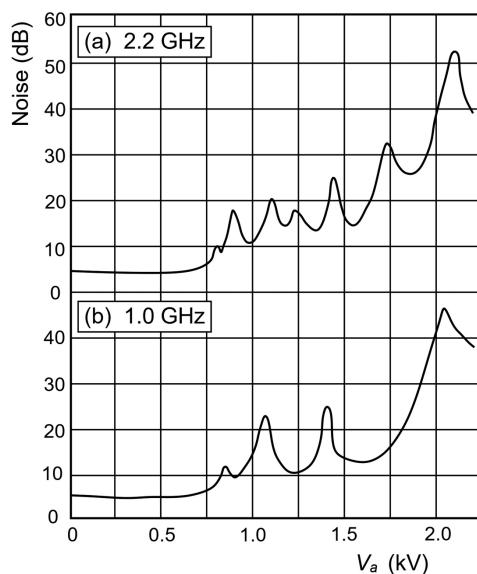
Figure 8.18: Typical noise/voltage curves measured, with uncalibrated probes in a cavity magnetron operated below oscillation
(copyright 1955, IEE, reproduced, with permission from [16]).

of current striking the cathode by a factor of around 10^6. Because the experiments employed a diode in which the thermionic current was deliberately kept small they are not representative of the situation in most practical devices where thermionic emission plays a larger role. However, the similarity between the curves in Figure 8.15, and the characteristic curves of a magnetron oscillator, suggest that there may be similarities between the electron dynamics in the two cases.

An important series of measurements of the noise in cut off magnetrons was made by Glass et al. [16]. It was found that noise could be detected in fluctuations of the anode current, or in noise power radiated by the heater leads, when a magnetron oscillator was operated in its pre-oscillating state. This noise was found to contain clear frequency peaks whose amplitude and frequency depended on the anode voltage. Frequency peaks were detected in 3 GHz tubes in the range 0.5 MHz to 4.0 GHz. Typical results are shown in Figure 8.18. The general conclusions were:

- The noise, including the peaks, increased as the anode voltage was increased.
- There was always a noise maximum near the oscillating voltage.
- A given peak moved to higher frequencies as the anode voltage was increased.

When the voltages at which the peaks occurred were plotted against frequency it was found that they lay on a series of curves diverging from the origin. Calculations suggested that these lines corresponded to azimuthal modes with differing periodicity. The conclusion of this paper was:
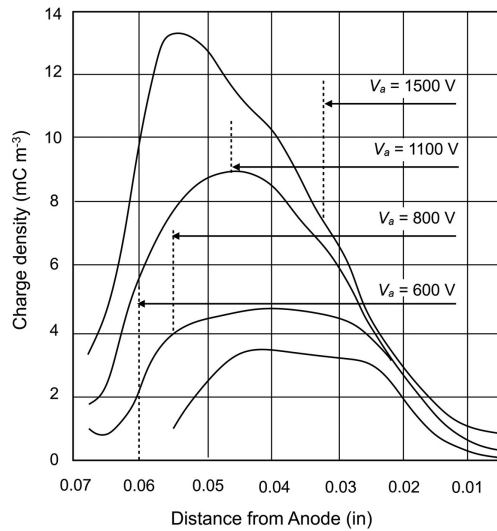
Figure 8.19: Distribution of space-charge in a cut-off magnetron diode for various voltages (copyright 1955, AIP Publishing, reproduced, with permission, from [19]).

'The experimental results described above have shown the presence of oscillations in the space-charge cloud down to extremely low voltages. Thus for all practical purposes the space-charge can be considered as being always in a state of oscillation, and a true steady state therefore never really exists. *The question of whether a single- or double-stream steady state is present is thus largely an academic one* (my emphasis)'

(copyright 1955, IEE, reproduced, with permission from [16]).

Despite this, authors have continued to argue the claims of these theoretical states to be the stable state existing in a magnetron diode. Further experimental evidence on oscillations in magnetron diodes was presented by Osepchuk [17] who noted that, although violation of cut-off is not always accompanied by detectable oscillations, it was possible that oscillations were present under most conditions.

At the same period a number of authors reported measurements of the space-charge distribution between the cathode and the anode [18]. A very careful set of results was obtained by Nedderman [19] who used the radiation from low-pressure helium within the magnetron to calculate the charge density. A representative set of results is shown in Figure 8.19 with the edge of the space-charge layer according to the Hull equation indicated by a dotted line for each anode voltage. The position of the cathode coincides with the left-hand edge of the figure. The method of measurement was unable to detect electrons having energies less than about 20 eV and, for that reason, the space-charge density appears to be very low close to the cathode. Nedderman estimated the charge which had been missed by the measurements and showed that it could mostly be accounted for by continuing the curves to the left of their peaks with constant charge density. It is evident from these results that the space-charge occupies most of the space between the edge of the theoretical space-charge layer and the anode. This is particularly noticeable

when the diode is strongly cut off. Further measurements were made at a reduced cathode temperature and at voltages too low for secondary emission to occur. It was found that, though the space-charge was reduced by this process, it was still far from zero. The conclusion was drawn that electrons were being trapped by the field and unable to reach either the anode or the cathode. Such electrons might be expected to remain there for long periods and gradually drift out of the field under the influence of large numbers of small perturbations. It was concluded that the experimental results could not be explained by theories of the magnetron diode based on self-consistent fields with unique electron trajectories.

More recent experimental results were reported by Hirsch [20–22] who studied strongly cut-off magnetrons. It was shown that the anode current distribution was not uniform but was concentrated towards the ends of the anode, suggesting that any satisfactory theory would need to be three-dimensional. The effects of background gas pressure were investigated and it was shown that the anode current increased because of scattering from gas molecules at pressures greater than 10^{-5} Torr. At low currents oscillations were observed at very well-defined frequencies with very little noise. The frequency of each mode of oscillation was directly proportional to the magnetic field, and independent of the anode voltage. Most of these measurements were made on a tube which had a filamentary cathode, and a large ratio of anode to cathode radius, at voltages of a few hundred volts and currents of a few microamps.

Direct measurements of the potential distribution in a strongly cut-off magnetron showed the presence of space-charge outside the theoretical space-charge layer [23]. It was shown that the potential curve coincided with the theoretical curve for Brillouin flow close to the cathode, but diverged at greater distances. The radius at which the divergence commenced was around 40% of the theoretical radius of the space-charge cloud. The results of the measurements were said to be in agreement with those of Nedderman [19].

8.7　　The Magnetron Problem

It is evident that the experimental results obtained with cut-off magnetron diodes cannot be explained by the elementary theory developed above. This problem is the subject of an extensive literature with divided opinions. In the discussion below, the papers cited are those which appear most important in the development of the argument. According to the elementary theory it is impossible for any electrons to reach the anode when the diode is cut off. This is a consequence of the assumption that the principle of conservation of energy applies to each electron, and to the assumption that the flow does not vary with time. The electrons on the surface of the space-charge layer have no radial acceleration because the electric and magnetic forces are in balance. If any of these electrons are to be able to reach the anode they must first lose energy so that the magnetic force is reduced. This can only happen if they exchange energy with other electrons which are speeded up, and so returned to

the cathode, which they strike with non-zero energy. This explains the simultaneous appearance of anode current and back-bombardment of the cathode. It follows that while energy and momentum must be conserved for the electron cloud, as a whole, these quantities cannot be conserved for individual electrons [24]. Thus it is impossible for any theory based on the conservation of energy by individual electrons to explain the experimental results. Three possible causes of energy exchange have been proposed in the literature:

- two- or three-body collisions between electrons [25–28];
- cumulative effects of random thermal oscillations [25–26, 29];
- oscillations of the whole electron cloud [7, 17, 21–22, 24, 27, 30–33].

The consensus appears to be that electron collisions are too rare, and random thermal effects too small, to explain observed phenomena. The importance of collective instabilities is supported by the existence of peaks in the noise spectrum which can be explained in terms of azimuthal space-charge modes (see also Section 15.2.3). It may be noted that those authors who believe that collisions are important also refer to collective effects. The explanation of the behaviour of cut-off magnetrons in terms of space-charge oscillations has been confirmed by PIC simulations [12, 32–33] which also showed the importance of non-stationary secondary electron emission from the cathode. In a 2D simulation, agreement was obtained with experimental results to within 10% when the spacing between the cathode and the anode was small compared with their length [15, 33].

References

[1] J. Benford *et al.*, *High Power Microwaves*. Boca Raton, FL: CRC Press, 2015.

[2] I. N. Bronshtein and K. A. Semendyayev, *A Guide-Book to Mathematics*. Frankfurt and Zurich: Verlag Harry Deutsch, 1971.

[3] A. W. Hull, 'The effect of a uniform magnetic field on the motion of electrons between coaxial cylinders', *Physical Review,* vol. 18, pp. 31–57, 1921.

[4] J. C. Slater, *Microwave Electronics*. New York: D. van Nostrand, 1950.

[5] R. G. E. Hutter, *Beam and Wave Electronics in Microwave Tubes*. Princeton, NJ: D. van Nostrand, 1960.

[6] L. Brillouin, 'Theory of the magnetron. I', *Physical Review,* vol. 60, pp. 385–396, 1941.

[7] O. Buneman, 'Symmetrical states and their breakup', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 209–233, 1961.

[8] R. M. Gilgenbach *et al.*, 'Crossed-field devices', in R. J. Barker *et al.*, *Modern Microwave and Millimeter-wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 289–342, 2005.

[9] Y. Y. Lau *et al.*, 'Limiting current in a crossed-field gap', *Physics of Fluids B,* vol. 5, pp. 4486–4489, 1993.

[10] A. H. Falkner, 'Double stream flow in the smooth-bore magnetron', *Proceedings of the Institution of Electrical Engineers,* vol. 120, pp. 959–961, 1973.

[11] J. A. Bradshaw, 'Cutoff characteristics of the static magnetron diode', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 261–274, 1961.

[12] H. L. McDowell, 'Smooth-bore magnetron simulations using a moving wavelength computer code', *IEEE Transactions on Plasma Science,* vol. 30, pp. 980–983, 2002.

[13] P. J. Christenson and Y. Y. Lau, 'One-dimensional modulation instability in a crossed-field gap', *Physical Review Letters,* vol. 76, pp. 3324–3327, 1996.

[14] A. V. Agafonov and V. S. Voronin, 'Brillouin flow in a magnetron diode as the kinetic model limit', in *Proceedings of the 1997 Particle Accelerator Conference*, Vancouver, BC, vol. 1, pp. 1302–1304, 1997.

[15] R. L. Jepsen and M. W. Muller, 'Enhanced emission from magnetron cathodes', *Journal of Applied Physics,* vol. 22, pp. 1196–1207, 1951.

[16] R. C. Glass *et al.*, 'Noise in cut-off magnetrons', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 102, pp. 81–86, 1955.

[17] J. M. Osepchuk, 'Magnetron diode oscillations', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 275–318, 1961.

[18] R. L. Jepsen, 'Electron interaction in the static magnetron', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 251–259, 1961.

[19] H. C. Nedderman, 'Space-charge distribution in a static magnetron', *Journal of Applied Physics,* vol. 26, pp. 1420–1430, 1955.

[20] E. H. Hirsch, 'Cut-off characteristics of a cylindrical magnetron', *International Journal of Electronics,* vol. 21, pp. 521–533, 1966.

[21] E. H. Hirsch, 'Electron transport in the strongly cut-off magnetron', *International Journal of Electronics,* vol. 22, pp. 297–305, 1967.

[22] E. H. Hirsch, 'Space charge oscillations in the cylindrical magnetron', *International Journal of Electronics,* vol. 23, pp. 497–509, 1967.

[23] V. A. Berbasov and L. M. Groshkov, 'Measurement of the potential distribution in a magnetron diode for a strongly transcritical mode', *Radiophysics and Quantum Electronics,* vol. 13, pp. 617–618, 1970.

[24] G. D. Sims, 'Preoscillation phenomena in space-charge clouds below the main oscillation threshold', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 179–208, 1961.

[25] G. Hok, 'A statistical approach to the space-charge distribution in a cut-off magnetron', *Journal of Applied Physics,* vol. 23, pp. 983–989, 1952.

[26] G. Hok, 'Statistical theory of the magnetron space charge', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 235–250, 1961.

[27] E. Q. Zhang, 'An interpretation of magnetron problem', *IEEE Transactions on Electron Devices,* vol. 27, pp. 1280–1283, 1980.

[28] E. Q. Zhang, 'An interpretation of the magnetron problem – Reply', *IEEE Transactions on Electron Devices,* vol. 28, p. 893, 1981.

[29] K. Mouthaan and C. Susskind, 'Anode current in the smooth-bore magnetron', *Proceedings of the IEEE,* vol. 54, pp. 430–432, 1966.

[30] E. H. Hirsch, 'Comment on "An interpretation of the magnetron problem"', *IEEE Transactions on Electron Devices,* vol. 28, pp. 892–893, 1981.

[31] E. Q. Zhang, 'On the magnetron cathode', *IEEE Transactions on Electron Devices,* vol. 33, pp. 1383–1384, 1986.

[32] A. V. Agafonov *et al.*, 'Dynamics of magnetic insulation violation in smooth-bore magnetrons', in *1997 Particle Accelerator Conference*, vol. 1, pp. 1299–1301, 1997.

[33] A. V. Agafonov *et al.*, 'Dynamics of magnetic insulation failure and self-organization of an electron flow in a magnetron diode', *Technical Physics,* vol. 49, pp. 93–103, 2004.

9   Electron Guns

9.1   Introduction

Many high-power vacuum tubes depend for their operation on the interaction between an extended electron beam and an electromagnetic structure. In general the current density required in the beam is greater than that which can be obtained from the cathode. Thus the area of the cathode must be greater than that of the beam. Careful design of static electric and magnetic fields is needed to ensure that the electron flow converges smoothly from the cathode into the beam. The arrangement of electrodes and magnetic materials used to achieve this is known as an electron gun. The objective of this chapter is to review the main types of electron gun and the principles of their design. The greater part of the chapter is devoted to Pierce electron guns which are used in inductive output tubes (IOTs) (Chapter 12), klystrons (Chapter 13) and travelling-wave tubes (TWTs) (Chapter 14). The discussion includes the electrostatic design of the Pierce gun, and methods for launching the beam formed by the gun into uniform and periodic magnetic focusing fields. The remainder of the chapter introduces Kino guns, used in injected beam crossed-field tubes (Chapter 16), and magnetron injection guns, used in gyrotrons (Chapter 17). Most electron beams are cylindrically symmetrical about the direction of flow being either solid or hollow cylinders. However beams whose cross-sections are either flat sheets or flattened ellipses are of increasing interest. Multiple-beam tubes incorporate several electron beams in parallel with one another and this raises special problems in the design of the electron guns.

    Many electron guns are space-charge limited diodes whose perveance is defined by (5.8)

$$ K = \frac{I_0}{V_a^{1.5}}, \tag{9.1} $$

where I_0 is the beam current and V_a the potential of the anode with respect to the cathode. This ratio is typically in the order of 10^{-6} A V$^{-1.5}$ for electron guns and it is customary to express K in microperveance. The performance of an electron gun is limited by:

- The maximum current density available at the cathode. This must always be sufficient to maintain space-charge-limited flow over the whole area of the cathode

to ensure that the current is not dependent on the properties of the cathode surface or on its temperature. (But note that the magnetron injection guns used in gyrotrons are normally operated in the temperature-limited region (see Section 9.6.2)).)

- The ratio of the cross-sectional areas of the electron flow at the cathode and in the final beam (the *Area Convergence* of the gun). If this is too great it becomes impossible to ensure that the electron trajectories flow smoothly into the beam without crossovers. Most guns have area convergences less than 100:1.
- Electric breakdown between the focus electrode and the anode (see Section 18.7.1).

9.2   The Pierce Electron Gun

The essential feature of the Pierce electron gun is rectilinear electron flow based on a simple space-charge limited diode together with the use of electrodes at cathode and anode potentials to maintain the correct boundary conditions along the edge of the beam [1]. The commonest form, used in most linear-beam tubes, is the conical convergent gun illustrated in Figure 9.1. This is based on the spherical diode, discussed in Section 5.6, and it is described in detail below. The purpose of the gun is to create a cylindrical electron beam having current I_0, velocity u_0, radius b and uniform current density, in which the electron trajectories do not cross over one another.

9.2.1   Electrostatic Theory of the Pierce Electron Gun

The electrons flow in a cone having half-angle θ from the cathode whose radius of curvature is R_c towards a concentric anode having radius of curvature R_a. For the present we will assume that the anode takes the form of a perfectly permeable
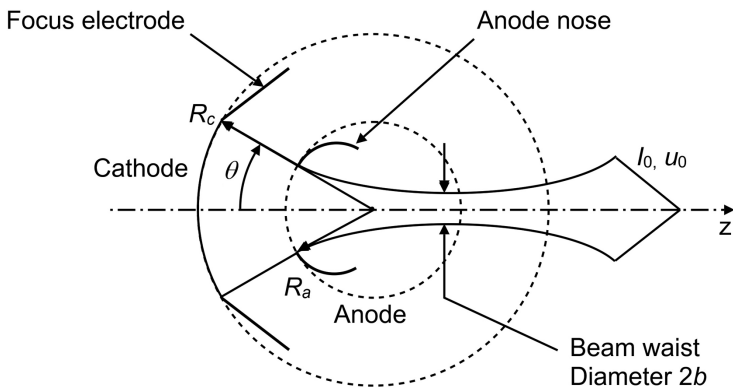
Figure 9.1: Arrangement of a Pierce electron gun.

conducting grid at potential V_a with respect to the cathode. The fraction of the surface area of the cathode sphere from which current is drawn is

$$\frac{2\pi R_c^2 \left(1 - \cos\theta\right)}{4\pi R_c^2} \tag{9.2}$$

and the space-charge-limited current is then derived from (5.84) as

$$I_0 = \frac{16\pi\varepsilon_0 \sqrt{2\left(e/m_0\right)}}{18} \cdot \frac{V_a^{1.5}}{\alpha_a^2} \left(1 - \cos\theta\right), \tag{9.3}$$

where α_a is the value of the function α of R_c/R defined by (5.85) calculated at $R = R_a$. This equation neglects relativistic effects which can be included by making use of (5.89). Equation (9.3) can be written

$$\cos\theta = 1 \;\; - \;\; \frac{9\alpha_a^2}{8\pi\varepsilon_0 \sqrt{2\left(e/m_0\right)}} \cdot \frac{I_0}{V_a^{1.5}}, \tag{9.4}$$

from which it can be seen that the angle of convergence required to produce a beam having a given perveance depends only upon α_a and therefore upon the ratio R_c/R_a. It follows from (9.4) that all guns which are geometrically similar (i.e. scaled from one another) have the same perveance.

In practical electron guns it is not possible to use a grid at the anode because it would quickly be destroyed by the power dissipated by interception of electrons. Instead there must be an aperture in the anode sufficiently large for the beam to pass through without interception. Figure 9.2 shows the modification in the field of a spherical capacitor when a hole is bored through the inner electrode. It can be seen that the electric field lines penetrate into the hole and that the transverse
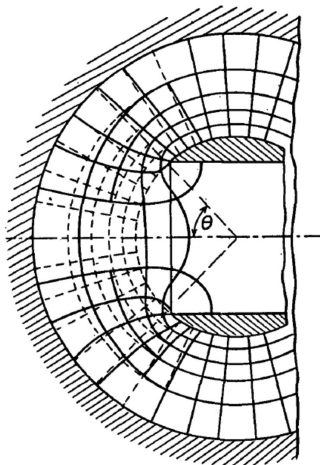
Figure 9.2: Field distortion in a spherical capacitor due to a hole in the inner electrode (copyright 1956, The IET, reproduced, with permission, from [2]).

component of the electric field is no longer zero. The transverse field acts as a diverging lens so that the slope of the beam edge trajectory, as it leaves the anode region, is less than the initial angle of convergence [1].

The strength of this lens can be estimated by the using the aperture lens theory described in Section 7.3.2. The effects of space-charge may be neglected, in the first approximation, because the local fields due to the anode are much stronger than the space-charge field. The focal length of the anode lens is given by (7.29)

$$\frac{1}{f_0} = \frac{(V_a + V_R)(E_1 - E_2)}{2V_a(V_a + 2V_R)},$$

(9.5)

where $V_R = m_0 c^2 / e$ and E_1 and E_2 are the axial electric fields before and after the anode so that $E_2 = 0$ in the drift region and f_0 is positive for a converging lens. The electrostatic potential between the cathode and the anode can be written, using (9.3), in the form[1]

$$V(R) = V_a \left(\frac{\alpha(R_c/R)}{\alpha(R_c/R_a)} \right)^{\frac{4}{3}}.$$

(9.6)

If the electric field at the surface of the ideal spherical anode is taken as an estimate of E_1 then

$$E_1 = -\left(\frac{\partial V}{\partial R} \right)_{R=R_a} = -\frac{4V_a}{3\alpha_a} \frac{R_c}{R_a^2} \cdot \left(\frac{d\alpha}{d(R_c/R)} \right)_{R=R_a} = -\frac{4V_a}{3} \frac{R_c}{R_a^2} \cdot \frac{\alpha_a'}{\alpha_a}$$

(9.7)

so that, substituting in (9.5) the focal length of the lens, normalised to the spherical radius of the anode, is given by

$$\frac{R_a}{f_0} = -\frac{2(V_a + V_R)}{3(V_a + 2V_R)} \cdot \frac{R_c}{R_a} \cdot \frac{\alpha_a'}{\alpha_a}.$$

(9.8)

In the non-relativistic limit (9.8) becomes

$$\frac{R_a}{f_0} = -\frac{1}{3} \cdot \frac{R_c}{R_a} \cdot \frac{\alpha_a'}{\alpha_a}.$$

(9.9)

We note that, at a fixed anode voltage, R_a/f_0 is a function of R_c/R_a only.

Figure 9.3 shows the geometry of the model of the Pierce electron gun including the effects of the anode aperture. From the lens equation

$$\frac{r_a}{z_1} = \frac{r_a}{R_a \cos \theta_0} + \frac{r_a}{f_0},$$

(9.10)

which can be written as

---

[1] Other authors express (9.6) and (9.7) in terms of α^2 because this function was tabulated by Langmuir in his 1924 paper. As it is easy to calculate α directly the use of α^2 is unnecessary.
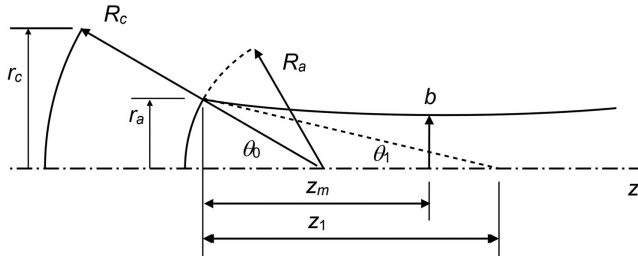
Figure 9.3: Geometry of the Pierce electron gun.

$$\tan \theta_1 = \tan \theta_0 + \frac{R_a}{f_0} \sin \theta_0. \tag{9.11}$$

We note that θ_1 is a function of R_c / R_a and the perveance. Variations of this equation appear in the literature. Some authors assume that the angles are small so that $\theta \approx \sin \theta \approx \tan \theta$ [1, 3, 4] Others have $\sin \theta_0$ in place of $\tan \theta_0$ as the first term on the right-hand side [5, 6]. This appears to be a consequence of the inconsistent choice of the position of the lens at the interception of the theoretical anode with the axis, whilst maintaining the point at which the angle of the trajectory changes, as shown in Figure 9.3. Others have sine functions throughout for reasons which are not stated [7, 8]. If the initial angle of convergence is small it is possible for the second term on the right-hand side of (9.11) to be greater than the first so that θ_1 is negative and the beam diverges after leaving the lens. Since the beam must be converging as it leaves the lens this sets a minimum value to θ_0.

After the electrons have passed through the anode aperture they move into a region where the electrostatic potential on the boundaries of the beam tunnel is constant. The beam then obeys the universal beam spreading curve, discussed in Section 7.5.1, and has minimum radius b. It is necessary for the initial electron flow to be convergent so that the beam waist falls at a convenient distance beyond the anode. The magnetic or electrostatic focusing system for the beam is then designed to maintain this radius. The slope of the edge of the beam as it leaves the anode lens can be written, using (7.72) as

$$\tan \theta_1 = \beta_p b \sqrt{\ln (r_a / b)}, \tag{9.12}$$

where β_p includes the relativistic correction if required. Rearranging this equation and making use of the geometry of the gun to eliminate r_a gives

$$\frac{r_c}{b} = \frac{R_c}{R_a} \exp \left(\frac{\tan \theta_1}{\beta_p b} \right)^2. \tag{9.13}$$

We note again that, for given parameters of the electron beam, the ratio r_c / b depends only upon R_c / R_a and the perveance. Thus, if r_c / b and the perveance are

known, then the value of R_c/R_a is fixed by (9.13) and the other parameters of the gun can be found by substitution.

9.2.2   The Focus Electrode and Anode Nose

In order to maintain the desired conical flow of the electrons it is necessary to choose the shapes of the focus electrode and of the anode in such a way that the electrostatic potential outside the beam is continuous with that defined by (9.6). It is possible to find the theoretical electrode shapes for conical flow Pierce guns [1, 9]. In practice the presence of the anode aperture distorts the field, as shown in Figure 9.2, and it is found that electrodes having the general form shown in Figure 9.4 are required to restore the equipotential surfaces close to the cathode to the desired spherical form [2, 10, 11]. Because of the concentration of equipotential surfaces close to the anode nose it is not possible to match the potential distribution along the whole of the beam edge from the cathode C to the theoretical position of the anode A. Instead the potentials are matched from C to the intersection of a surface such as B with the beam edge. Brewer noted that, if B is too close to C, then the equipotential surfaces close to the cathode will be distorted, and that if B is too close to A then the fields are strongly distorted at the edge of the beam close to the anode. He suggested that B should be approximately 0.6 of the distance between C and A [10]. For typical electron guns the potential at B is between 25% and 35% of V_a and the electron velocity is, therefore, between 50% and 60% of u_0.

Comparison between experimental results, and the simple theory of the Pierce electron gun presented in the previous section, shows that the strength of the anode lens is generally greater that that predicted by (9.9), especially as the angle of convergence increases. Danielson et al. investigated this problem and suggested that the last term in (9.11) should be multiplied by a factor which was calculated to be 1.1 for a 0.33 µPerv gun with $R_c/R_a = 2.37$[12]. The correction factor was found to be close to $\sqrt{V_a/V_x}$ where V_x was approximately equal to the potential on the axis at its intersection with the theoretical anode. Müller used ray tracing through the field of
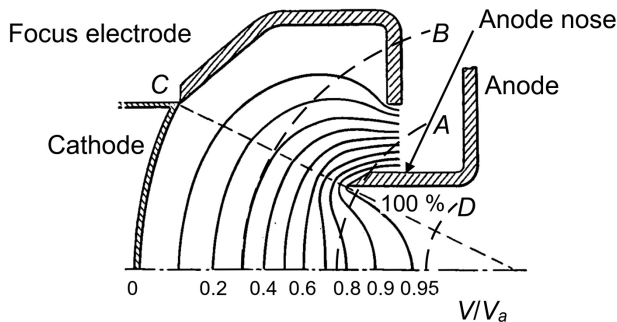
Figure 9.4: Equipotential surfaces in a realistic Pierce electron gun (copyright 1956, the IET, reproduced, with permission, from [2]).

the pierced vacuum capacitor shown in Figure 9.2 and obtained a correction factor of $1/\cos\theta_0$ for $\theta_0 \geq 20°$ [2]. Vaughan suggested an empirical correction factor of 1.25 based the analysis of a number of guns [5]. An alternative approach was used by Yang et al. who fitted the results for θ_0 of a set of 30 electron guns to experimental values by multiplying the last term in (9.4) by an empirical constant [13]. These empirical approaches are not very satisfactory and it is useful to consider the problem more closely.

In order to model the gun it is convenient to divide it into three regions as shown in Figure 9.4. Between C and B the equipotential surfaces are spherical and governed by (9.6) and the current is determined by the space-charge limited flow. The equipotential surfaces outside the beam in this region should be very similar to those determined theoretically [9]. The solution for conical flow is obtained in a plane containing the axis and expressed in polar coordinates centred on the centre of curvature of the cathode. The parameter γ defined in (5.83) is generalised as

$$\gamma = \ln\left(\frac{R}{R_c}\right) + j(\theta - \theta_0) \tag{9.14}$$

and the potential is given by

$$V(R,\theta) = k\,\mathrm{Re}\left\{\alpha(R,\theta)^{\frac{4}{3}}\right\}, \tag{9.15}$$

where k is a constant. The equipotential surfaces calculated in two dimensions remain solutions of Laplace's equation when they are rotated about the axis of symmetry of the gun. Figure 9.5 shows the shapes of the equipotential surfaces which are obtained. It should be noted that these curves are plotted for the case where $\theta_0 = 0$ with the cathode at the origin. The shapes for other angles of convergence are
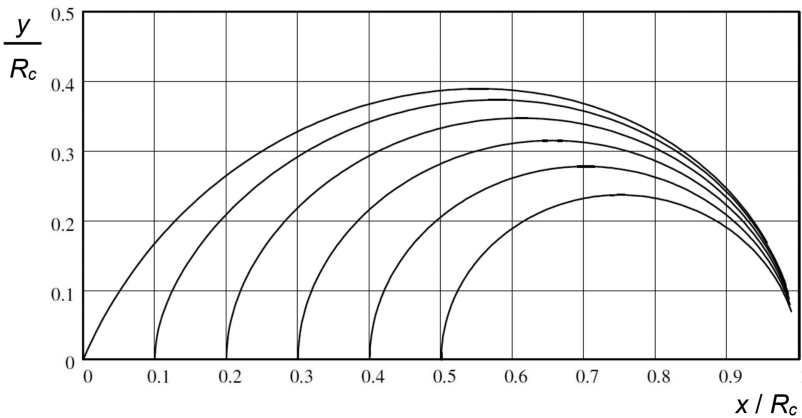
Figure 9.5: Normalised shapes of the beam-forming electrodes for a solid convergent conical beam
(copyright 1982, Elsevier, reproduced, with permission, from [9]).

obtained by rotating the figure so that the horizontal axis coincides with the edge of the beam. The curve corresponding to the focus electrode makes an angle of $67.5°$ with the edge of the beam [1]. Comparison between Figures 9.4 and 9.5 shows the similarity between the equipotential surfaces in this region.

Beyond D the external potential is effectively equal to the anode potential and the beam drifts under the influence of space-charge forces alone and follows the universal beam spreading curve.

Between B and D the electrons continue to accelerate but gradually deviate from the initial conical flow under the influence of the transverse electric fields in the region of the anode. As the electron flow is space-charge limited the electric field at the cathode is zero and, thus, from Gauss' Law we know that the flux of **D** out of the conical volume bounded by C and A must be equal to the charge enclosed. Now, the current is continuous, but the electron velocity increases as the electrons move from the cathode towards the anode. Therefore the flux of **D** is determined very largely by the charge close to the cathode. The simple model of the gun discussed in Section 9.2.1 assumes that the whole of this flux passes transversely through the beam edge as it passes through the anode. In the region between A and D the simple model assumes that the transverse electric field is determined by the local charge density in the beam. Thus the whole of the electric field which influences the transverse motion of the electrons is determined by the distribution of the space charge, and of the equipotential boundaries. Because the greater part of the charge is concentrated close to the cathode, the electric field associated with it is much greater than that due to the local space-charge in the region of the anode. An estimate of the contribution from the local space charge shows that it is typically less than 10% of that from the charge close to the cathode [14]. Thus the effects of the local space-charge on the strength of the anode lens can be neglected in the first approximation [15]. We note that the region between B and D is not thin and that the electron velocity varies considerably within it. However, the transverse electric field is strongest in the part of it which is close to A. It is therefore an acceptable approximation to model the electron motion in this region as: conical flow from C to A; a thin electrostatic lens at A; and motion governed by the universal beam spreading curve beyond A. It is desirable, however, to include the effects of large angles of convergence and variation of the electron velocity in the model of the anode lens.

9.2.3   Improved Model of the Anode Lens

We have already noted that the modification of the electron trajectories as they pass through the anode aperture is a consequence of the transverse component of the electric field. Following [12] we note that the flux of the electric field within the beam which would have terminated on the theoretical anode is terminated on the anode aperture. This flux is

$$\Phi = 2\pi R_a^2 \left(1 - \cos\theta_0\right) E_1, \tag{9.16}$$
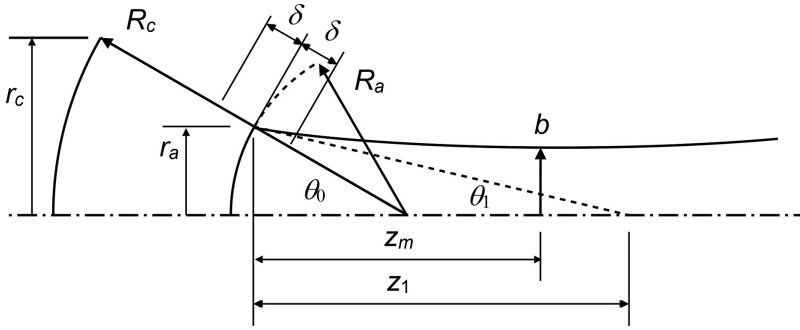
Figure 9.6: Geometry of the anode lens.

where E_1 is given by (9.7). Let us suppose that this transverse flux is confined to a short axial length 2δ of the beam edge trajectory centred on the point at which it intersects the theoretical anode as shown in Figure 9.6. Thus we are assuming that the anode lens is thin. This is not strictly correct, but the detailed distribution of the electric field depends on the shape of the anode and, therefore, it is not possible to derive a simple general model which includes the spatial variation of the electric field. The area of conical surface is

$$A = 4\pi R_a \delta \sin \theta_0. \tag{9.17}$$

Then, because the flux must be conserved, the average of the normal component of the electric field over the element is

$$E_n = \frac{\Phi}{A} = \frac{R_a}{\delta} \cdot \frac{(1-\cos\theta_0)}{2\sin\theta_0} E_1. \tag{9.18}$$

Now consider an electron which enters the lens at an angle θ_0 and passes through it with effective velocity u_e corresponding to a potential V_e. This is less than u_0 because the penetration of the field into the aperture means that the average potential in the lens is less than V_a. The impulse given to the electron is equal to the change in its momentum normal to the direction of motion. Thus

$$\frac{m_0}{\sqrt{1-u_e^2/c^2}} \Delta u_r = -e \int E_n dt. \tag{9.19}$$

But we have assumed that E_n is constant so (9.19) becomes

$$\Delta u_r = -\frac{2\delta\sqrt{1-u_e^2/c^2}}{u_e} \cdot \frac{e}{m_0} E_n \tag{9.20}$$

and, substituting for E_n from (9.18)

$$\Delta u_n = -\frac{\sqrt{1-u_e^2/c^2}}{u_e} \cdot \frac{(1-\cos\theta_0)}{\sin\theta_0} \cdot \frac{e}{m_0} R_a E_1. \tag{9.21}$$

Thus the direction of the trajectory changes by an angle given by

$$\tan\phi = \frac{\Delta u_n}{u_e} = -\frac{\sqrt{1-u_e^2/c^2}}{u_e^2}\cdot\frac{(1-\cos\theta_0)}{\sin\theta_0}\cdot\frac{e}{m_0}R_a E_1. \tag{9.22}$$

Now

$$\frac{\sqrt{1-u_e^2/c^2}}{u_e^2} = \frac{1}{c^2}\cdot\frac{V_R(V_e+V_R)}{V_e(V_e+2V_R)} \tag{9.23}$$

so that

$$\tan\phi = -\frac{(V_e+V_R)}{V_e(V_e+2V_R)}\cdot\frac{(1-\cos\theta_0)}{\sin\theta_0}\cdot R_a E_1 \tag{9.24}$$

and, using (9.5), this may be written as

$$\tan\phi = -\frac{(V_e+V_R)}{V_e(V_e+2V_R)}\cdot\frac{V_a(V_a+2V_R)}{(V_a+V_R)}\cdot\frac{2(1-\cos\theta_0)}{\sin\theta_0}\cdot\frac{R_a}{f_0}. \tag{9.25}$$

When relativistic effects are not important this reduces to

$$\tan\phi = -\frac{V_a}{V_e}\cdot\frac{2(1-\cos\theta_0)}{\sin\theta_0}\cdot\frac{R_a}{f_0}. \tag{9.26}$$

Hence the angle of the trajectory as it leaves the lens is given by

$$\theta_1 = \theta_0 - \tan^{-1}\left[\frac{V_a}{V_e}\cdot\frac{2(1-\cos\theta_0)}{\sin^2\theta_0}\cdot\frac{R_a}{f_0}\sin\theta_0\right]. \tag{9.27}$$

Comparing (9.27) with (9.11) we see that the estimate of the focal length has been adjusted by two terms. The first recognises that the electron has not yet reached its full speed as it passes through the lens, and the second takes account of the finite angle between the trajectory and the axis. We note also the lens effect is shown in the addition of angles rather than their tangents. When inner trajectories are considered the effective voltage varies with the angle of convergence as can be seen from Figure 9.4. Thus the lens shows spherical aberration in which the focal length varies with the angle of the trajectory. Spherical aberration causes the current density in the beam to be non-uniform. This can be corrected by changing the shape of the cathode surface so that the spherical aberration is compensated by changes in the current density [5, 11, 16].

Examination of Figure 9.4 shows that the effective potential V_e depends upon the shape of the anode and upon the distance between the beam edge trajectory and the edge of the aperture. It is, therefore, not possible to find a theoretical expression for V_e/V_a. However, we may expect that this ratio will have a similar value for many electron guns and that the computed dimensions of the gun will not depend strongly on it. This was investigated by applying the revised theory to the data for electron guns

in [5, 13]. This includes guns with microperveances in the range 0.29 to 3.67 and area convergences from 5 to 300. In each case the ratio V_e/V_a was adjusted so that the angle θ_0 was the same as that tabulated. In all but six cases the ratio lay in the range 0.6–0.8 with a mean value of 0.7. This is consistent with the typical potential distribution shown in Figure 9.4. Figure 9.7 shows how the normalised effective voltage depended upon the angle of convergence. The scatter may be explained by differences between the shape and position of the anode nose relative to the beam edge in different guns. The effective voltage tends to decrease as the angle of convergence increases. This can also be understood by reference to Figure 9.4 by noting that the ratio of the diameter of the anode aperture to the distance between the anode and the cathode increases as θ_0 increases. Thus the field penetrates further into the beam tunnel and the effective potential at the theoretical anode decreases.

The angle θ_0 was computed for the same data set, using the fixed ratio $V_e/V_a = 0.7$, with the results shown in Figure 9.8. It can be seen that there is a very strong correlation between the computed and experimental results. A useful empirical expression
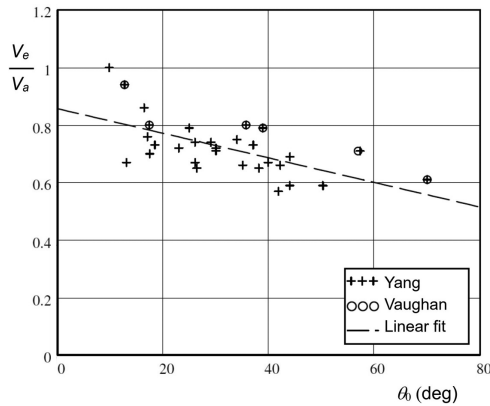
Figure 9.7: Dependence of the normalised effective voltage on the angle of convergence.
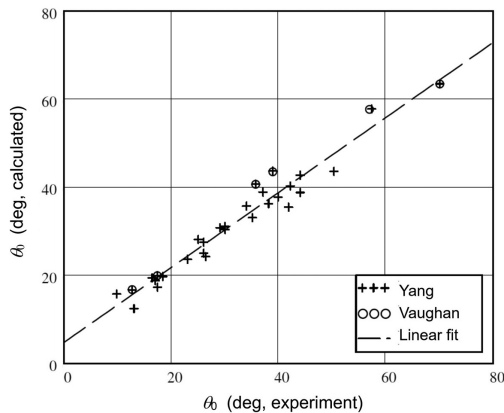
Figure 9.8: Comparison between computed and experimental angles of convergence.

for the relationship between the angle of convergence in degrees, the perveance and the area convergence has been given by True [17]

$$\theta_0 = 12.9° \times \sqrt{\mu P (r_c/b)},\qquad(9.28)$$

where μP is the microperveance. This expression is said to be valid for $\theta < 40°$, $\mu P < 3$ and $2 < r_c/b < 8$. When (9.28) was applied to the subset of the data given in [13] that satisfies these criteria, the scatter of the points was very similar to that shown in Figure 9.8. It was found that the mean value of the ratio of the estimated, and experimental, angles of convergence was unity with a standard deviation of 0.14.

9.2.4   The Effects of Thermal Velocities

Because the electrons in a Pierce gun are obtained by thermionic emission from the cathode their random thermal velocities are superimposed on the overall motion. These thermal velocities mean that the edge of the electron beam is not abrupt, as has been assumed so far, but the current density falls off more gradually. This problem was studied by Cutler and Hines who determined the effect of thermal velocities on the current density profile in the beam formed by a Pierce gun [18]. It was shown that the ratio of the current density at radius r to the idealised current density J_0 can be expressed as

$$\frac{J_r}{J_0} = \exp(-r^2/2\sigma^2) \int_0^{r_e/\sigma} \left(\frac{R}{\sigma}\right) I_0\left(\frac{rR}{\sigma^2}\right)\exp(-R^2/2\sigma^2)\, d\left(\frac{R}{\sigma}\right)\qquad(9.29)$$

where r_e is the expected beam radius, and the normalised value of σ is given by

$$\frac{\sigma}{R_c}\sqrt{\frac{2eV_a}{kT}} = \frac{r_e}{r_c} A(-\alpha_a)^{2/3} + 3\frac{r_e}{b}(B \pm C)\sqrt{\frac{\pi}{2}}(-\alpha_a)^2.\qquad(9.30)$$

Here

$$A = \int_1^{R_c/R_a} \frac{1}{(-\alpha)^{2/3}}\, d\left(\frac{R_c}{R}\right)\qquad(9.31)$$

$$B = \operatorname{erf}\left(\ln\sqrt{r_a/b}\right)\qquad(9.32)$$

and

$$C = \operatorname{erf}\left(\ln\sqrt{r_e/b}\right).\qquad(9.33)$$

Figure 9.9 shows the variation in current density with radius for a range of parameters. The fall-off in current density takes place over a range $r/\sigma \sim 5$ centred on the value of r_e/σ. Thus thermal velocity spreading of the beam is negligible when r_e/σ is greater than 100. For a particular case the value of σ can be computed from (9.30)
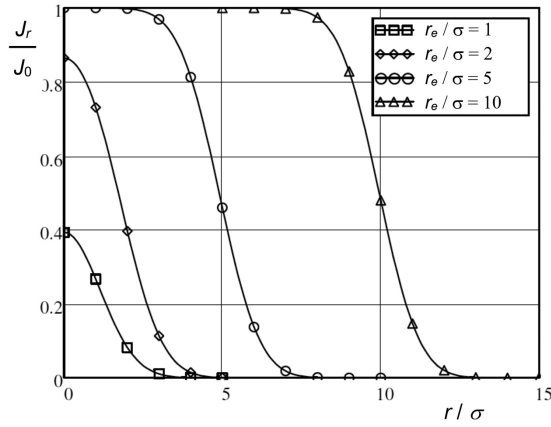
Figure 9.9: Variation in current density with radius from (9.29) (copyright 1955, IEEE, reproduced, with permission, from [18]).
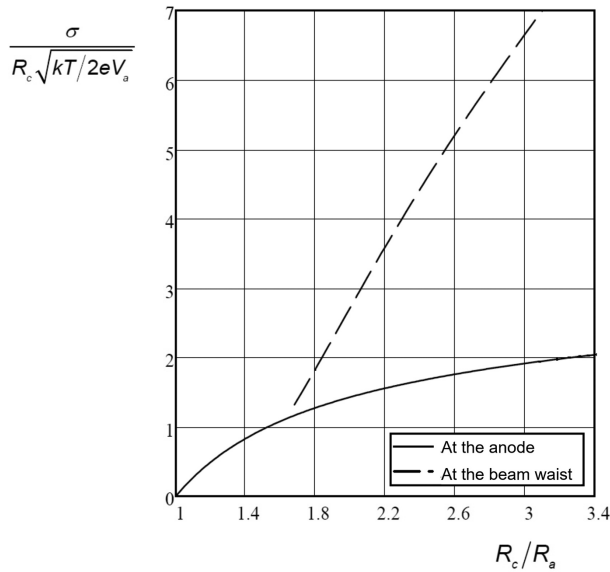
Figure 9.10: Variation in the normalised value of σ at the anode and at the beam waist with R_c/R_a from (9.30) (copyright 1955, IEEE, reproduced, with permission, from [18]).

where the negative sign is taken for points between the anode and the beam waist, and the positive sign for points beyond the beam waist.

At the anode, and at the beam waist, the right-hand side of (9.30) depends only on R_c/R_a, and the normalised value of σ on the left-hand side of (9.30) can be plotted against R_c/R_a, as shown in Figure 9.10. This figure differs slightly from Figure 5 in [18] because of the difference in the method of calculating the anode lens effect. It can be seen that the effects of thermal velocity are more pronounced at the beam

waist than at the anode. In any given case the importance of thermal velocity effects can be computed from this graph or directly from (9.30).

In practice it is important that, in linear-beam tubes, as little of the beam current should be intercepted by the structure of the tube as possible. If we require the beam transmission to exceed 99% then the envelope enclosing 99% of the beam current must be smaller than the diameter of the beam tunnel throughout its length. The fraction of the beam current within radius r is given by

$$\frac{I\left(r/\sigma, r_e/\sigma\right)}{I_0} = 2\left(\frac{\sigma}{r_e}\right)^2 \int_0^{r/\sigma}\left(\frac{r}{\sigma}\right)\cdot\frac{J_r\left(r/\sigma, r_e/\sigma\right)}{J_0} \, d\left(\frac{r}{\sigma}\right). \tag{9.34}$$

The beam radius within which any given proportion of the current is contained can be computed by solving this equation (see Worksheet 9.1). Figure 9.11 shows, as an example, the normalised beam tunnel radius (a/σ), for different transmission percentages, as a function of the normalised expected radius. It can be seen that the increase in the effective beam radius is greatest when r_e/σ is small.

Thermal velocity effects are greatest at low beam voltages, and for guns which have low perveance and high area convergence. The beam radius (r_{99}) enclosing 99% of the beam current was computed at the beam waist, with an anode voltage of 10 kV, for the 30 guns in [13]. It was found that r_{99}/b was more than 1.1 in about half the cases, more than 1.2 in a third of them, and that the greatest value was 1.6. Similar computations for the six guns in [5] all showed ratios of more than 1.1 and the greatest ratio was 2.6. Thermal effects were dominant in all cases of this
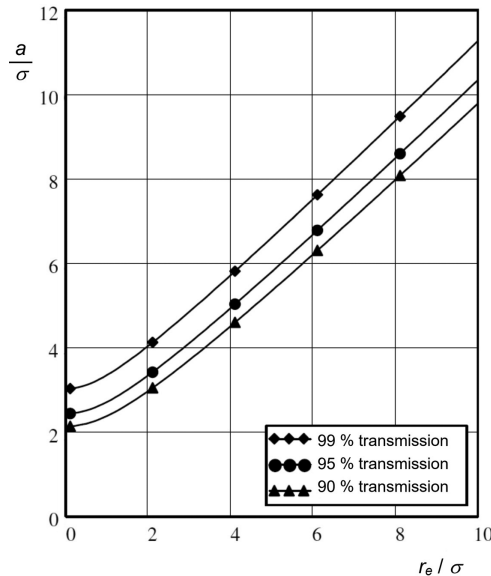
Figure 9.11: Normalised radius enclosing different percentages of the current (copyright 1955, IEEE, reproduced, with permission, from [18]).

set except No. 4 and No. 5. When these two cases were recomputed, with the target value of the cold beam radius (b) reduced in proportion to the thermal increase, then r_{99} was found to be very close to the original target radius. A small increase of up to about 1° in the angle of convergence was also required. If r_{99}/b is greater than 1.2 then the position of the beam waist is closer to the anode than is predicted by calculations which ignore thermal velocities [18]. Thus for thermal increases in the beam radius greater than about 20% it is necessary to compute the full beam profile including the effects of thermal velocities in order to find the true position and radius of the beam waist.

9.2.5   Electrostatic Design of a Pierce Electron Gun

The theory given in the preceding sections is useful for the initial design of Pierce electron guns with conical flow. The starting point is to choose the beam voltage and current, the final beam radius, and the cathode current density. The cathode loading is determined by whether the operation of the gun is pulsed or continuous, by the type of cathode chosen, and the intended lifetime of the cathode (see Section 18.5). In a typical electron gun it is found that the current density varies approximately as $1/\cos^2 \alpha$ over the surface of the cathode, where α is the angle from the centre of curvature to the point on the cathode surface [19]. The value of the normalised effective voltage at the anode lens can be chosen by reference to the results presented above, or from data for other electron guns similar to that to be designed. The scatter in Figure 9.8 suggests that setting this figure to 0.7 is likely to give a good starting point for the initial design. The problem is then one of finding the value of R_c/R_a using the equations above. Various authors have described the solution of this problem using graphical methods [1, 8], iterative methods [5] or direct methods [3, 13]. Modern computer software makes it very easy to seek the solutions of the equations directly by numerical methods. This is illustrated in Worksheet 9.1 which includes the improved model of the anode lens described above. Once the value of R_c/R_a has been found the values of the other dimensions in Figure 9.3 follow directly.

Sufficient clearance must be left between the anode tunnel and the beam to ensure that electrons are not intercepted on it. Typically the diameter of the aperture should be 20–25% greater than the beam diameter at that plane [5]. The tip of the anode nose usually lies close to the theoretical spherical anode surface. The detailed shapes of the focus electrode and of the anode nose are found empirically. Originally this was achieved using an electrolytic tank in which a strip of dielectric material was placed along the beam edge. The shapes and positions of the external electrodes were adjusted until the potential distribution measured along this strip matched that required [10]. More recently the same task has been achieved using computer programs which model the whole electron gun [20–26]. In high-voltage guns it is necessary to ensure that the surfaces are designed to avoid voltage breakdown (see Section 18.7.1) [27].

Table 9.1: Parameters of a Pierce electron gun

	Frost et al. Figure 9 [11]	Vaughan, example 2 [5]	Calculated $V_e/V_a = 0.7$	Calculated $V_e/V_a = 0.8$
b (mm)	1.35	1.35	1.35	1.35
R_c (mm)	12.8	12.5	11.6	12.9
R_a (mm)	–	–	4.39	5.24
θ_0 (deg)	38.5	40.1	43.6	38.5
θ_1 (deg)	–	–	12.2	12.7
z_m (mm)	26.3	25.9	22.1	23.5
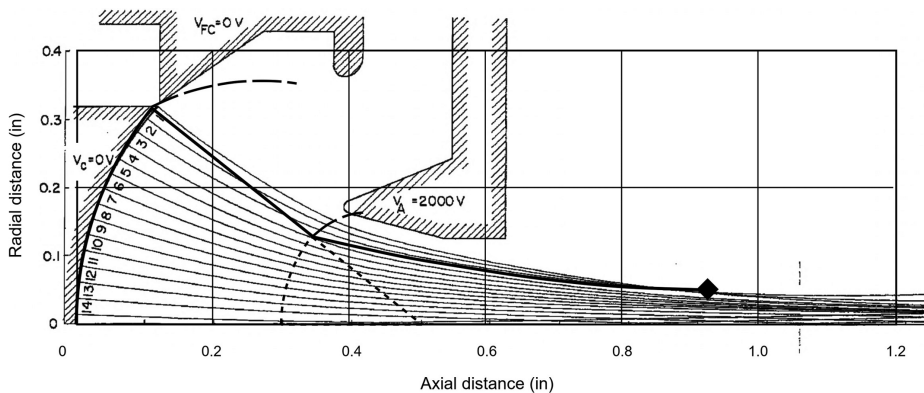
Figure 9.12: Theoretical and computed beam profiles for an electron gun (copyright 1966, IEEE, reproduced, with permission, from [24]).

As an example, let us consider the type 4A gun shown in Figure 9 of [11] which has been used as an example by other authors [5, 24]. This was designed to have a microperveance of 1.9 and cathode radius 8.0 mm (0.316 in). The spherical radius of the cathode was increased to improve the uniformity of the current density in the beam. The measured beam radius was 1.35 mm and the area convergence approximately 35:1. The leading dimensions of this gun can be calculated using the method described above using Worksheet 9.1 with the normalised effective anode voltage set to $0.7V_a$. Better agreement with the experimental data is obtained by increasing the effective anode voltage to $0.8V_a$ Table 9.1 shows a comparison between the dimensions of the gun and the results of the calculation.

Figure 9.12 shows the results of the calculations with $V_e/V_a = 0.8$ superimposed upon the trajectories computed in [24]. The dashed lines show the positions of the theoretical anode, the focus electrode, and the anode nose. The choice of the higher value of V_e/V_a can be justified by noting that the anode aperture is only about 11% greater than the computed diameter of the beam at that plane. It can be seen that

the tip of the anode nose lies close to the theoretical spherical anode. The plane of the beam waist lies slightly closer to the cathode than the position determined experimentally. Overall the comparison shows that the simple calculation provides a good starting point for subsequent improvement of the design. A gun for a 50 MW pulsed klystron is described in [28] and a detailed description of the design of a 0.52 µperv gun for a TWT is to be found in [29].

9.3   Magnetic Field Design for a Pierce Electron Gun

We have seen that a Pierce electron gun uses convergent flow of the electrons to form an electron beam in which the electrons at the beam waist are flowing parallel to the axis. In order for the beam to be useable in a linear-beam tube it must be controlled by magnetic, or electric, fields as described in Chapter 7. Here we shall only consider magnetic focusing using either a solenoid field or a PPM field. We saw in Chapter 7 that, in order to ensure smooth flow of the beam, some of the magnetic flux must be linked to the cathode (see equations (7.54) and (7.96)). This condition is necessary, but not sufficient, to achieve smooth flow. It is also necessary to ensure that the magnetic field between the cathode and the beam waist is not such as to cause the beam to scallop. The usual practice is to carry out the electrostatic design of the gun so that the beam waist radius is equal to the desired beam radius, and then to design the magnetic field in the beam entry region to ensure that that diameter is maintained without scalloping.

9.3.1   Solenoid Focusing

The magnetic field of the solenoid for a linear-beam tube is normally terminated by a soft iron sheet (pole-piece). The purpose of this is to provide a return path for the magnetic flux so ensuring that the most efficient use is made of the electro- or permanent magnets providing it. As there must be a hole in the pole-piece, which is at least big enough for the beam to pass through it, some of the magnetic flux leaks into the beam entry region. The magnetic field in the beam entry region is thus controlled primarily by the position of the pole-piece relative to the cathode and by the size of the aperture in it. A theoretical expression for the flux leakage caused by the hole in the pole-piece is obtained by assuming that the pole-piece can be represented by a thin sheet which is a perfect magnetic conductor coinciding with the inner face of the pole-piece [30]. The problem is rotationally symmetrical and described in ellipsoidal coordinates (ξ, ζ) which are related to cylindrical polar coordinates (r, z) by the equations

$$\frac{z^2}{a^2 \zeta^2} + \frac{r^2}{a^2 \left(1 + \zeta^2\right)} = 1 \tag{9.35}$$

and

$$-\frac{z^2}{a^2\xi^2}+\frac{r^2}{a^2\left(1-\xi^2\right)}=1.\tag{9.36}$$

It follows that

$$z=a\zeta\xi\tag{9.37}$$

and

$$r=a\sqrt{\left(1+\zeta^2\right)\left(1-\xi^2\right)}\tag{9.38}$$

where the sheet is in the $z=0$ plane and the edge of the hole is at $r=a$. The solution of the analogue problem in electrostatics is found from Laplace's equation in this system of co-ordinates subject to the boundary conditions $E_z\to E_0$ when $z\to\infty$ and $E_z\to0$ when $z\to-\infty$. The resulting scalar potential is

$$V=aE_0\xi\left[\zeta-\frac{1}{\pi}\left(\zeta\cot^{-1}\xi-1\right)\right].\tag{9.39}$$

This is quite a good approximation provided that the material of the pole-piece is not close to saturation around the edge of the hole. The parametric equations of the flux lines are

$$r=r_0\left\{1-\frac{1}{\pi}\left[\cot^{-1}\zeta-\frac{\zeta}{1+\zeta^2}\right]\right\}^{-\frac{1}{2}}\tag{9.40}$$

and

$$z=\zeta\left\{a^2-\frac{r^2}{1+\zeta^2}\right\}^{\frac{1}{2}},\tag{9.41}$$

where a is the radius of the aperture, z_a is the axial position referred to the aperture, and $r\to r_0$ as $\zeta\to\infty$. Figure 9.13 shows the flux lines calculated using equations (9.40) and (9.41). For a paraxial ray $r\ll a$ and $\zeta=z/a$ so that the flux density on the axis varies as

$$B_z=B_0\left\{1-\frac{1}{\pi}\left[\cot^{-1}\left(\frac{z}{a}\right)-\frac{z/a}{1+\left(z/a\right)^2}\right]\right\},\tag{9.42}$$

as shown in Figure 9.14. It is evident that the flux is very close to its asymptotic values at $z=\pm2a$.

The profile of the magnetic field in the beam entry region must be designed in such a way that the electrons flow without scalloping in a beam having the intended radius. That means that, at the electrostatic beam waist, the radial component of velocity must be zero and the angular velocity must have the correct value to
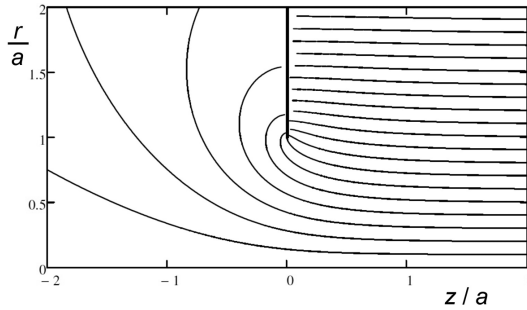
Figure 9.13: Magnetic flux lines in the region of an aperture in a pole-piece.
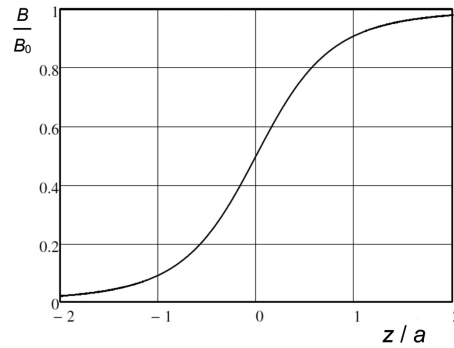
Figure 9.14: Flux density on the axis in the region of an aperture in a pole-piece.

balance the space-charge forces. Thus at least two parameters are needed to specify the magnetic field profile. This can be achieved, in principle, by choosing the diameter of the aperture in the pole-piece, and the distance of the inner face of the pole-piece from the cathode.

If the beam is focused by a uniform magnetic field whose strength is m times the Brillouin field then the ratio of the flux linked to the cathode to that linked to the beam is given by (7.57)

$$\frac{B_c r_c^2}{B_0 b^2} = \sqrt{K} = \sqrt{1 - \frac{1}{m^2}}. \tag{9.43}$$

Thus the flux line which passes through the edge of the cathode is asymptotic within the solenoid to radius r_0 given by

$$r_0 = b\left(1 - \frac{1}{m^2}\right)^{\frac{1}{4}}. \tag{9.44}$$

Using this value of r_0 we can find the value $\zeta = \zeta_c$ for which $r = r_c$ in (9.40). The distance between the pole-piece and the cathode rim is then found, as a function of a, by substituting $r = r_c$ and $\zeta = \zeta_c$ in (9.41). Hence the radius of the aperture and

its distance from the cathode rim are not independent of each other. In order to complete the specification of the magnetic field we need some way of deciding how to choose one of these dimensions so that the flux varies smoothly through the gun. Now the differential equation governing the edge of the beam is (7.53)

$$\frac{d^2 R}{dz^2} = \frac{1}{2}\beta_p^2 \frac{1}{R} + \beta_L^2 K \frac{1}{R^3} - \beta_L^2 f(z)^2 R, \tag{9.45}$$

where $\beta_p = \omega_p/u$, $\beta_L = \omega_L/u$, $f(z)$ is the term in braces in (9.42), and u_0 has been replaced by u because the electron velocity is not constant. Now let $R = R_e + \Delta R$, where $R_e = r/b$ is the normalised electrostatic beam edge, and ΔR the small perturbation caused by the magnetic field. By definition

$$\frac{d^2 R_e}{dz^2} = \frac{1}{2}\beta_p^2 \frac{1}{R_e}. \tag{9.46}$$

Therefore, making use of (9.43)

$$\frac{d^2(\Delta R)}{dz^2} \approx \beta_L^2 R_e \left[\frac{R_0^4}{R_e^4} - f(z)^2 \right], \tag{9.47}$$

where it has been assumed that $\Delta R \ll R_e$. At the edge of the cathode the term in square brackets is zero because from (9.40) $R_0^2/R_e^2 = r_0^2/r_c^2 = f(z_c)$. In order to keep the perturbation as small as possible we wish to keep this term small by matching the flux line to the electrostatic beam edge as closely as possible [31]. It is particularly important to do this close to the cathode where u is small and, therefore, β_L is large. Thus we expect that the best choice of the aperture dimensions will be close to those which make the slope of the flux line at the edge of the cathode equal to the angle of convergence of the beam [32]. Thus a good estimate of a can be obtained from

$$\left(\frac{dr}{d\zeta} \bigg/ \frac{dz}{d\zeta} \right)_c = -\tan\theta_0. \tag{9.48}$$

To find how the choice of aperture dimensions affects the beam, it is necessary to integrate the equations of motion of the electrons under the combined influence of the potentials on the electrodes, the space-charge of the electrons, and the applied magnetic field. To do this properly requires the use of a computer program such as EGN2 [20] to find a self-consistent solution to the problem. However, the effects of the aperture dimensions on the beam can be illustrated by a simple model which assumes that the electron flow is conical between the cathode and the theoretical anode. Then (9.45) is integrated with the initial conditions $r = r_a$ and $dr/dz = -\tan\theta_1$ from that point onwards (see Worksheet 9.1). Since it has been assumed that the electrons follow paraxial paths, it is necessary to determine the effective position of the pole-piece by setting $z = a\zeta$, rather than by using (9.41), to ensure that description of the magnetic field is self-consistent. Using the worksheet it can be shown that
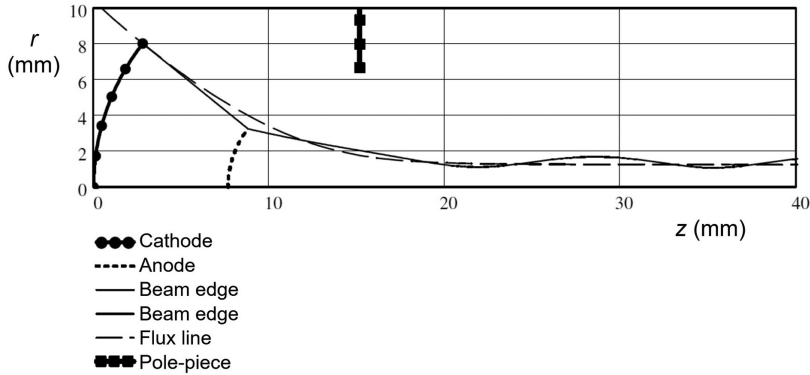
Figure 9.15: Magnetic focusing of the beam in the gun illustrated in Figure 9.12.

there is a value of the pole-piece dimensions for which the electrons form a beam with minimum scalloping. It is found that the slope of the flux line passing through the edge of the cathode is then close to $-\tan\theta_0$. It has also been found that the flux linked to the beam at the plane of the electrostatic beam waist is close to the cathode flux, a criterion which was proposed for PPM focusing by Legarra et al. [33].

Figure 9.15 shows the computed position of the pole-piece aperture, the flux line through the cathode edge, and the magnetically confined beam profile for the gun illustrated in Figure 9.12 when $m = 2$ and a is given by (9.48). The beam shows a small amount of scalloping which can be reduced, but not quite eliminated, by adjusting the diameter and axial position of the aperture. The gun has a large angle of convergence so the use of the paraxial ray approximation is not really valid. Nevertheless the model does show the way in which the beam is affected by the choice of the aperture parameters.

Once an initial design has been chosen it can be modelled accurately, and optimised, using a gun simulation program [20–22, 25, 26]. If necessary, further control over the distribution of the magnetic field can be provided by placing a soft iron cylinder around the gun, though this has little effect on the field at the cathode unless it is placed close to it [31]. It is sometimes also useful, at least for developmental purposes, to place a coil, known as a bucking coil, outside the tube in the plane of the cathode so that the cathode flux can be adjusted to give optimum beam transmission. When a very small beam is required, as in some millimetre-wave tubes, a region of increasing magnetic field may be used to compress the beam [34–36].

The method described above cannot be applied to a magnetically shielded gun because no flux passes through the cathode. An alternative approach is based upon an ideal field which is zero up to the beam waist and equal to the Brillouin field thereafter [37–39]. Since this ideal field cannot be realised in practice we assume that there is a short transitional region where the field is given by

$$B = \frac{B_0}{2}(1 - \cos\alpha), \tag{9.49}$$

where $0 \leq \alpha \leq \pi$. This expression is a good approximation to that given by (9.42). A more accurate approximation is given in [40]. The hole in the pole-piece is assumed to be small so that the magnetic field is effectively zero at the plane of the cathode. The effects of the real, and ideal, magnetic fields can therefore be computed from the equation for the focal length of a magnetic lens (7.46). Since the beam radius is already close to its final value in the region where the magnetic field is changing we can assume that r is constant and remove it from the integral. The condition for the equivalence of the real and ideal magnetic fields is then [39]

$$\int_1^2 B_{real}^2 \, dz = \int_1^2 B_{ideal}^2 \, dz,$$ (9.50)

where the field is changing between planes 1 and 2. This can be written in terms of α

$$\frac{1}{4}\int_0^\pi (1 - \cos\alpha)^2 \, d\alpha = \int_{\alpha_w}^\pi \, d\alpha,$$ (9.51)

where α_w is the position of the beam waist. It is straightforward to show that this condition is satisfied when $\alpha_w = 5\pi/8$ and that, therefore, that the magnetic field should be $0.69 B_0$ at the beam waist to achieve smooth Brillouin flow. It is found that this conclusion is insensitive to the exact form of the variation of magnetic field [38]. Worksheet 9.1 gives a similar result when m is close to unity.

9.3.2   PPM Focusing

The principles for designing the magnetic field for a beam entering a periodic permanent magnet (PPM) stack are the similar to those for solenoid focusing. It was shown in Chapter 7 that the equilibrium conditions for an electron beam, having a given mean beam radius and cathode flux, are the same for both solenoid and PPM focusing when the RMS PPM field is equal to the solenoid field. The maximum diameter of the PPM focused beam is then greater than that of the equivalent solenoid focused beam by the amplitude of the ripple. The diameter of the PPM focused beam is greatest when the magnitude of the axial magnetic flux density is greatest. Thus, in theory, the beam should be launched with the desired maximum radius, and with $dr/dz = 0$, at a plane where the amplitude of magnetic field is greatest.

A method for designing the magnetic field in the entry region was described in [41] (see Worksheet 9.2). An ideal variation of the magnetic field was defined in which the electron trajectories followed the magnetic flux lines between the cathode and the electrostatic beam waist. At that plane the beam entered a cosinusoidally varying PPM field, as shown by the solid line in Figure 9.16. Thus, in the entry region, the flux linked to the beam is constant and equal to the cathode flux. The field profile in Figure 9.16 cannot be realised in practice. However, following an approach similar to that used in (9.50), it was shown that correct launching of the
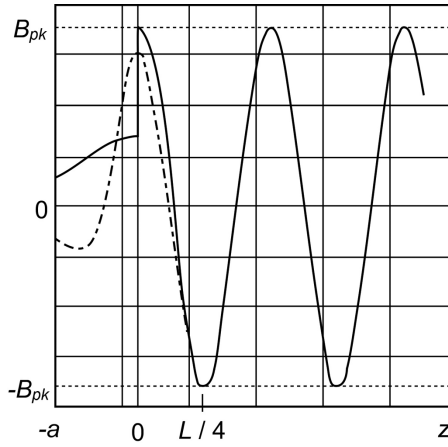
Figure 9.16: Magnetic field for the entry of the beam from a Pierce gun into a PPM stack showing the ideal profile (solid curve) and a real profile (dashed curve) (copyright 1967, IEEE, reproduced, with permission, from [41]).

beam can be achieved by the use of an equivalent magnetic flux density B_e that satisfies the equation

$$\int_{-a}^{L/4} B_e^2 \, dz = \int_{-a}^{0} B_1(z)^2 \, dz + \int_{0}^{L/4} B_{pk}^2 \cos^2\left(\frac{2\pi z}{L}\right) dz, \qquad (9.52)$$

where B_1 is the flux density between the theoretical anode $(z = -a)$ and the entrance plane of the PPM field, B_{pk} is the peak PPM flux density, and the first zero of the flux density is at $z = L/4$. It should be noted that this equation assumes that the electron velocity is constant, and the beam radius is approximately constant, throughout the transition region. This equation imposes the criterion that the focal lengths of the magnetic lenses, represented by the real and ideal magnetic field profiles, should be the same. A field profile satisfying this equation is shown by the dashed curve in Figure 9.16. Rawls et al. showed that nearly identical beam profiles were produced using the real and ideal field profiles. Beam transmission of 98% was achieved in a tube designed in this way. The application of this method to the case where the cathode is shielded is described in [39].

An alternative method for designing the field in the entry region is shown in Figure 9.17 [33]. In region 1 the field is designed in the same way as for solenoid focusing with a target magnetic field B_0. In region 3 the peak magnetic field is $\sqrt{2}B_0$, and the beam has mean radius b with ripple δ. The transition region 2 can be regarded as a quarter section of a PPM stack in which the peak magnetic field is B_0. At the end of the section the beam radius must be $(1-\delta)b$ and the electron trajectories parallel to the axis. If the beam radius is b at the start of region 2 there is no net inward force and the beam does not converge as required. Thus the beam
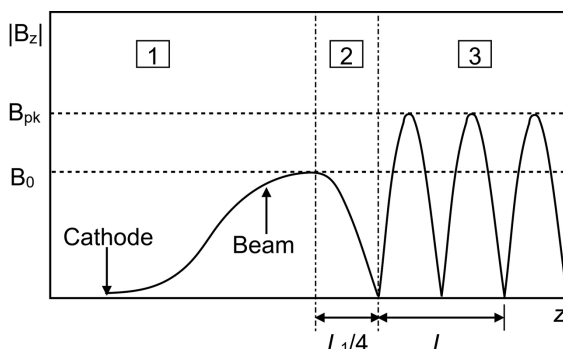
Figure 9.17: Magnetic field profile at the entry of a PPM stack (copyright 1983, IEEE, reproduced, with permission, from [33]).

radius must be greater than b at the end of region 1 and the radius of the beam waist must be less than b so that there is a net outward force there leading to the necessary overshoot.

9.4   Other Pierce Guns

The methods used to design Pierce electron guns for cylindrical beams can also be used to design guns to produce beams having other cross-sections. Pierce guns for producing sheet electron beams have become of greater interest in recent years because of their potential for use in linear-beam tubes at frequencies of 100 GHz and above. Hollow beams have occasionally been used in klystrons and TWTs, but their principal use is in gyrotrons. Most high power tubes with hollow beams have magnetron injection guns as described in Section 9.5.2.

9.4.1   Guns for Sheet Beams

In principle the design of a Pierce gun to produce a sheet beam proceeds in exactly the same way as that described for a cylindrical beam, in Section 9.2 [42]. The current is assumed to be drawn from a space-charge limited cathode which is a section of a circular cylinder given by (5.70). The focal length of the anode slit can be determined in a manner similar to that for a circular hole by considering the transverse impulse given to an electron passing through it. The universal beam spreading curve for a sheet beam can be derived in the same way as that for a cylindrical beam. The design, construction and testing of sheet-beam guns has been described in [43–45].

9.4.2   Guns for Hollow Beams

Pierce guns for the production of hollow cylindrical beams are similar to those for solid beams, except that the electrons are emitted from a ring on the surface of the cathode.

The anode aperture may be a round hole [42, 46–48] or a circular slot [49, 50]. In the former case the perveance is determined from the solution for a spherical space-charge-limited diode in a manner similar to that for a solid beam. In the latter, the solution for a toroidal space-charge-limited diode is used [51]. In each case focusing electrodes are required both inside, and outside, the beam. The anode lens effect for a circular hole is given by (9.5) and the equivalent expression for an annular hole is given in [51]. The spreading of a hollow beam under the influence of space-charge forces is identical to that for a solid beam having the same diameter, current, and voltage. However, because there is no space-charge in the centre of the beam, the trajectories of electrons on the inner edge of the beam do not spread. Thus the design of the gun must ensure that the outer electrons are converging towards the axis whilst the inner electrons are moving parallel to the axis as they leave the anode lens. Experimental results for toroidal electron guns having perveances of up to 15 μperv are described in [51].

9.5   Beam Control Electrodes

It is often desirable to be able to switch the electron beam on and off electronically, or to control the current in the beam independently of the accelerating voltage. This requires the use of a third electrode so that the gun is a triode rather than a diode. It has the advantage that the voltage and power required to control the beam may be much less than if the anode voltage is switched or controlled directly. Three techniques described in the literature are discussed below: Modulating anode; Control focus electrode; and Control grid [6, 52, 53]. Their typical characteristics are summarised in Table 9.2 [52, 53] where μ is the amplification factor as defined in (6.1) and μ_c is the ratio of the anode voltage to the negative grid voltage at cut-off.

9.5.1   Modulating Anode and Control Focus Electrode

When the anode of a diode electron gun is isolated from the body of the tube its potential can be varied independently. It is then known as a modulating anode.

Table 9.2: Comparison of methods of beam modulation using control electrodes
(© 1994, Springer Nature. Reproduced, with permission, from [53]).

Parameter	Modulating anode	Control focus electrode	Intercepting grid	Non-intercepting grid
μ	1 to 3	2 to 10	30 to 75	
μ_c	1 to 3	2 to 10	30 to 150	30 to 250
Capacitance	30 pF	50 to 100 pF	10 to 20 pF	30 to 50 pF
Grid current	0%	0%	10 to 20%	0.1 to 1%
Maximum grid power	n/a	n/a	5 W	5 W

The perveance of the diode gun is fixed by its geometry so that the current depends upon the potential difference between the cathode and the anode in the usual way. However, the velocity of the electrons in the interaction region is determined by the potential difference between the cathode and the body of the tube. Thus the beam current and the electron velocity can be varied independently of one another. It is necessary to take care that the beam is focused correctly over the desired range of voltages and currents. In a second method the focus electrode is isolated from the cathode so that its voltage can be varied independently. Because the focusing of the beam is seriously disrupted by changes in this voltage the technique is only suitable for switching the beam on and off. The negative voltage (with respect to the cathode) that is required to reduce the current density to zero at the centre of the cathode is comparable with the positive voltage applied to the anode. The negative voltage can be reduced somewhat by including a central pin at the same potential [32, 52]. In both of these methods the control voltage is large but the control power is small because the electrode can be designed to intercept very little current. Further information about the design of these two types of beam control electrode is given in [32].

9.5.2   Intercepting Control Grid

Control of the electron beam can be achieved with voltages much less than the anode voltage if a control grid is placed close to the cathode and at a uniform distance from it. The grid is manufactured from a refractory metal, such as molybdenum, or from pyrolitic graphite, which can withstand the high temperatures close to the cathode surface whilst maintaining very precise dimensions. The beam current then depends upon the voltage applied to the control grid in exactly the same way as in a triode (see Chapter 6). The current is cut off if the grid voltage is sufficiently negative with respect to the cathode. The negative grid voltage required to make the gun cut-off is typically 1–3% of the anode voltage while a similar positive voltage is needed to turn the gun fully on [53]. To minimise interception the grid must have transparency (the ratio of the holes in the grid to the total grid area) in the range 80–90% [32, 52]. If the grid voltage is positive then it intercepts 10–20% of the beam current [53]. Intercepted electrons may liberate secondary electrons from the grid and there can also be thermionic emission if the temperature of the grid is high enough [54]. For these reasons guns with intercepting grids are limited to low beam powers and to tubes, such as the IOT, in which the grid voltage is positive for only a small proportion of the time.

The arrangement of the apertures in the grid can take the form of an array of circular or hexagonal holes, or of the spaces between radial and concentric 'wires'. The section of the cathode opposite each hole in the grid behaves like an individual electron gun, and it is desirable that these should all have similar characteristics so that their currents vary together. Where the grid takes the form of a 'honeycomb' array of circular holes it can be much more robust, though the transparency may be as low as 60%, and the cathode surface may be dimpled to improve the electron optics [32, 52].

The properties of a gridded gun can be calculated as described for a triode in Chapter 6. Because the grid is very close to the cathode surface it is possible to consider an equivalent planar geometry in which the spacing between the cathode and the anode is chosen to give the same electric field at the surface of the cathode in the absence of the grid (c.f. (6.11)) [54]. The dimensions of the actual grid can be converted to those of an equivalent two-dimensional grid of round wires [19]. Hence the dependence of the beam current on the grid and anode voltages can be computed, as described in Section 6.4. In general, the current between the grid and the anode is not space-charge limited. It should be noted that the grid dimensions and spacings used in gridded guns are such that island formation usually occurs close to cut-off.

The electron trajectories in the region of the grid exhibit strong lens effects caused by the local fields in a manner similar to that shown in Figure 6.8. Thus, many of the electrons have appreciable transverse components of velocity at the beam waist, and the beam is less well-collimated than the beam from a well-designed diode gun. It may be necessary to remove the beam halo by passing the beam through a tunnel (a 'beam scraper') only slightly larger than the intended beam diameter.

9.5.3   Non-Intercepting Control Grid

The problems of intercepted current make a gun with an intercepting control grid unsuitable for use in high power tubes (with the exception of IOTs), and the collimation of the beam is unsatisfactory for TWTs. These problems can be greatly reduced by the introduction of a second grid (the *shadow grid*) very close to the cathode surface which is maintained at the cathode potential, or very close to it. The arrangement of this grid is such that it is a projection of the control grid so that any electrons which pass through the shadow grid also pass through the control grid without being intercepted by it. The intercepted grid current in such a gun can be 0.1% or less [6, 53, 55]. Figure 9.18 shows the computed electron trajectories in
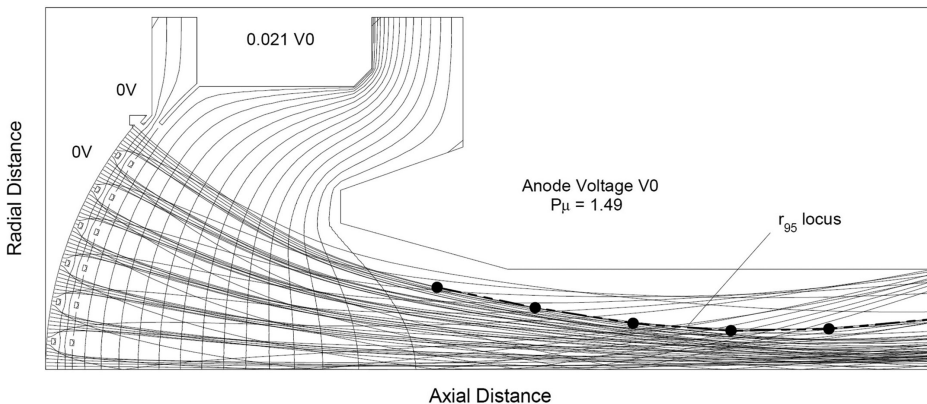
Figure 9.18: Electron trajectories in a shadow-gridded gun
(copyright 1984, IEEE, reproduced, with permission, from [56]).

a shadow-gridded gun [56]. It can be seen that lens effects in the region of the grid produce an appreciable beam halo. Although these trajectories carry only a small fraction of the total beam current they are very difficult to focus. True has shown that (9.28) applies to gridded guns if the beam radius is taken as that enclosing 95% of the beam current. He has also shown that the normalised radial component of the beam velocity can be calculated from the empirical relationship [56]

$$\frac{\bar{u}_r}{u_0} = 0.0015 \times \theta_0^\circ. \tag{9.53}$$

Thus the properties of the beam halo are determined to a large extent by the angle of convergence of the gun. It is desirable to keep this angle as small as possible to make the beam easier to focus. The procedure for coupling the design of a shadow gridded gun with PPM focusing is described in detail in [56]. Because the lens effects in the region of the grid depend upon the grid voltage and hence, the gun perveance, it is found that good collimation of the beam is only achieved over a limited range of operating conditions.

A number of variants of the basic shadow-gridded gun have been found to give improved beam collimation [55, 57]. If the shadow grid is placed in contact with the cathode the emitting area of the cathode is reduced, but it is found that the beam collimation is appreciably improved. The shadow grid may also be embedded in the surface of the cathode, or be formed by a thin patterned layer of non-emissive coating deposited on it. Improved beam collimation can also be achieved by isolating the shadow grid from the cathode and adjusting its potential independently in a tetrode gun [57]. An ultra-laminar gun with three grids is described in [55]. More complex grid structures are used in dual mode guns to enable beams with high and low powers to be produced by the same gun [6].

9.6   Crossed-Field Electron Guns

We saw in Section 8.2.2 that a stable sheet electron beam can exist in a region in which there are uniform static electric and magnetic fields which are perpendicular to each other and to the direction of electron motion. The variation in the electrostatic potential within the beam is given by (8.22). Now, useful interaction between an electron beam, and the RF field of an electromagnetic structure, requires that the spread of electron velocities should be small compared with their mean velocity. Therefore the thickness of a sheet beam in crossed fields must be small. A simple method for launching such a beam where the whole structure is immersed in a uniform magnetic field $B_z = -B_0$ is shown in Figure 9.19. An electron which starts from rest on the surface of the cathode has velocity $u_x = 2E_1/B_0$ when it is moving parallel to the x axis. In order for it to continue to move in a straight line parallel to the x axis $u_x = E_0/B_0$. Thus a necessary condition for the launching of a sheet beam is $E_1 = E_0/2$. It is also necessary for the potential at the interface to be continuous,
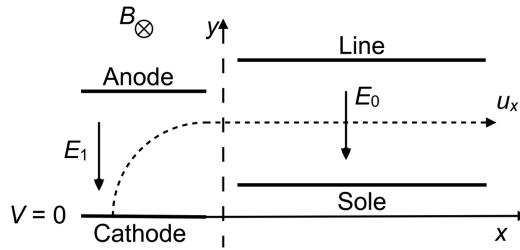
Figure 9.19: Arrangement of a simple crossed-field electron gun.
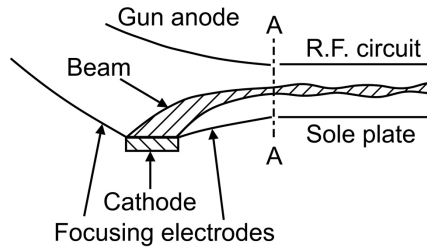
Figure 9.20: Arrangement of a Kino gun
(copyright 1960, IEEE, reproduced, with permission, from [59]).

and equal to that required by the conservation of energy [58]. This arrangement can produce a useful beam, but the width of the emitting strip of the cathode in the x direction must be small and, therefore, the current is limited by the current density available.

9.6.1   Kino Gun

A more practical arrangement is the Kino gun shown in Figure 9.20 [59]. The electron orbits in the gun (to the left of the plane A–A) are those for a space-charge limited magnetron diode given by (8.36) and (8.46) and shown in Figure 8.7. These trajectories are the correct ones to use because, although the diode is cut off, the electrons are extracted as a beam and do not return to the cathode. All the trajectories within the gun have the same form apart from a shift in the direction parallel to the surface of the cathode. Because the top of the trajectory is fairly flat it is possible for the cathode to be have a finite length, though this must be short compared with the cyclotron wavelength $\left(\lambda_c = 2\pi u_x / \omega_c \right)$. Thus, this type of gun is known as a 'short gun'. The whole length of the gun from the centre of the cathode to the plane A–A is approximately $\lambda_c / 2$. The shapes of the focus electrodes and the gun anode are used, as in the Pierce gun, to maintain the correct potential distribution along the edges of the beam. The design and testing of a gun of this type is described in [60].

   The current which can be delivered by a short gun is limited by the maximum current which can be obtained from the cathode. Thus, if greater beam current

is required, it may be necessary to use a cathode which is several cyclotron wavelengths long. However, we have seen that there is then uncertainty as to which of the possible Slater orbits represents reality.

For this type of gun Kino proposed a different solution for the flow of space-charge which assumed that the electrons are emitted from the cathode with finite velocity. The electron trajectories are then found to be parabolic. In certain circumstances the initial velocity required for this solution to be valid is of the same order as the thermal velocities of the electrons. This theory has been criticised by Vaughan [61] who showed that this solution only exists in specific circumstances which do not correspond to many practical cases. He investigated solutions in which the initial velocity of the electrons was an additional parameter and showed that, in a typical case, a well-collimated beam could be formed in which some trajectories showed cycloidal motion. The design and testing of an experimental 'long gun' based on Kino's theory is described in [62]. The gun gave the expected value of the current but the beam thickness was greater than expected. This was attributed to RF instabilities in the gun. Reviews of the literature on that subject are given in [63, 64]. It was found that the noise could be considerably reduced by operating the cathode in the temperature-limited regime; by tilting the cathode to a small angle so that there was a small normal component of the magnetic field [65]; or by placing a grid in front of the cathode [66] which could also be used for modulating the beam.

9.6.2   Magnetron Injection Gun

Figure 9.21 shows the arrangement of a magnetron injection gun which produces a hollow cylindrical beam [8]. This gun bears a superficial resemblance to the Kino gun described in the previous section, rotated about an axis of symmetry. But there is the important difference that the magnetic field is essentially parallel to the electron motion rather than perpendicular to it. The cathode takes the form of a truncated cone, and the whole gun is immersed in a strong axial magnetic field. Thus the region around the cathode is essentially a cut-off cylindrical magnetron diode whose space-charge layer forms a virtual cathode. The tapering of the cathode and the anode produces an axial electric field which draws the electrons into the beam. Focus electrodes are used to maintain the correct potential distribution along the edge of the beam in the gun region. In the transition region the radial electric field is gradually reduced to zero and the beam tends to its final dimensions. The area convergence which can be obtained is greater than that of a hollow-beam Pierce gun.

The design principles for a magnetron injection gun can be illustrated by reference to Figure 9.21. As there is no space-charge inside the hollow beam it follows that the inner electrons must have no angular velocity. Therefore, by the application of Busch's Theorem (7.39), the flux linked to the smaller end of the cathode must be equal to that linked to the inner electrons of the beam. If it is assumed that the whole gun is immersed in a uniform axial magnetic field B_0 then $r_i = r_a$. The
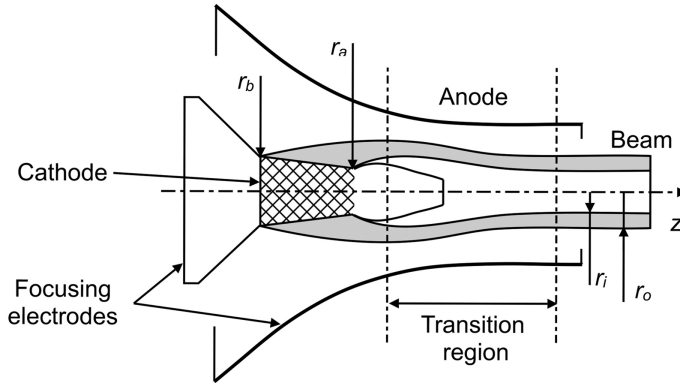
Figure 9.21: Arrangement of a magnetron injection gun
(copyright 1965, Elsevier, reproduced, with permission, from [8]).

condition for space-charge balanced flow of the outer electrons is identical to that
for a solid beam given by (7.54)

$$\beta_p^2 = 2\beta_L^2 \left(1 - K\right),\tag{9.54}$$

where $\beta_p = \omega_p/u_0$ is defined for a solid beam having the same current, velocity and
outer radius, and $\beta_L = eB_0/2m_0u_0$. Hence from (7.52) and (9.54)

$$K = \left(\frac{r_b}{r_o}\right)^4 = 1 - \frac{\beta_p^2}{2\beta_L^2}.\tag{9.55}$$

Thus, given the desired parameters of the beam and the strength of the magnetic
field, the larger radius of the cathode can be determined from (9.55). Now the sec-
ond term on the right-hand side of this equation is equal to unity for Brillouin
focusing of the equivalent solid beam. But, since $r_b > r_i$, it follows that $K > \left(r_i/r_o\right)^4$.
The ratio of the magnetic field to the Brillouin field given by (7.57) requires that

$$m = \frac{B_0}{B_B} = \frac{1}{\sqrt{1 - K}}.\tag{9.56}$$

Hence

$$m > \frac{1}{\sqrt{1 - \left(r_i/r_o\right)^4}}.\tag{9.57}$$

If, for example, $r_i/r_o = 0.9$ then $m > 1.7$. It is evident that the minimum magnetic
field required to focus a hollow beam is appreciably greater than for the equivalent
solid beam. From the geometry of the cathode the beam current is

$$I_0 = \frac{\pi\left(r_b^2 - r_a^2\right)}{\sin\phi} J_c,\tag{9.58}$$
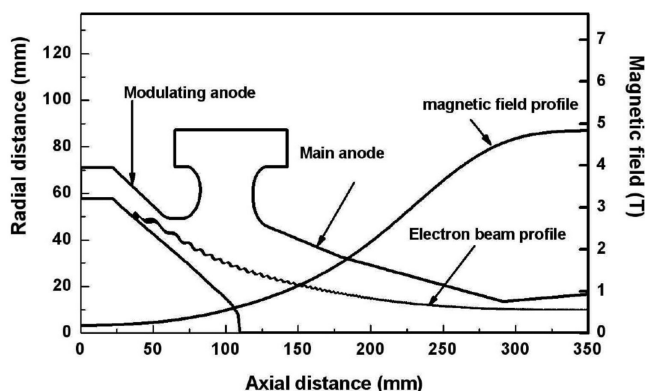
Figure 9.22: Electrode shapes, magnetic field profile and electron trajectories of a magnetron injection gun for a gyrotron
(copyright 2010, The Electromagnetics Academy, reproduced, with permission from [71]).

where ϕ is the angle between the slope of the cathode and the axis and J_c is the cathode current density. This equation determines the angle ϕ and, hence, the slant length of the cathode. Methods for calculating the electron trajectories and the electrode shapes have been described in [67–70].

Magnetron injection guns are occasionally used in high power linear-beam tubes but their main use is in gyrotrons (Chapter 17). Figure 9.22 shows the arrangement of a typical magnetron injection gun for use in a gyrotron [71]. Comparing this with Figure 9.21 it can be seen that the slant length of the cathode is small, that the anode is divided into two sections whose potentials can be adjusted separately, and that the magnetic field increases with distance so that the beam is compressed. A detailed description of the design methods for guns for gyrotrons, including two examples, is given in [72].

References

[1]   J. R. Pierce, *Theory and Design of Electron Beams*. New York: D. van Nostrand, 1949.

[2]   M. B. Müller, 'New points of view in the design of electron guns for cylindrical beams of high space charge', *Journal of the British Institution of Radio Engineers,* vol. 16, pp. 83–94, 1956.

[3]   U. Tiwary and B. N. Basu, 'Noniterative method for the synthesis of convergent Pierce electron guns', *IEEE Transactions on Electron Devices,* vol. 34, pp. 1218–1222, 1987.

[4]   B. N. Basu, *Electromagnetic Theory and Applications in Beam-Wave Electronics*. Singapore: World Scientific, 1996.

[5]   J. R. M. Vaughan, 'Synthesis of the Pierce gun', *IEEE Transactions on Electron Devices,* vol. 28, pp. 37–41, 1981.

[6]   A. S. Gilmour, Jr., *Principles of Traveling Wave Tubes*. Norwood, MA: Artech House, 1994.

[7] A. H. W. Beck, *Thermionic Valves: Their Theory and Design*. Cambridge: Cambridge University Press, 1953.

[8] J. F. Gittins, *Power Travelling-Wave Tubes*. London: English Universities Press, 1965.

[9] H. Z. Sar-El, 'Revised theory of Pierce-type electron guns', *Nuclear Instruments and Methods,* vol. 203, pp. 21–33, 1982.

[10] G. R. Brewer, 'Formation of high-density electron beams', *Journal of Applied Physics,* vol. 28, pp. 7–15, 1957.

[11] R. D. Frost *et al.*, 'Electron guns for forming solid beams of high perveance and high convergence', *Proceedings of the IRE,* vol. 50, pp. 1800–1807, 1962.

[12] W. E. Danielson *et al.*, 'A detailed analysis of beam formation with electron guns of the Pierce type', *Bell System Technical Journal,* vol. 35, pp. 375–420, 1956.

[13] C. Yang *et al.*, 'Improved noniterative method for the synthesis of convergent Pierce electron guns', *IEEE Transactions on Electron Devices,* vol. 53, pp. 2849–2852, 2006.

[14] C. K. Birdsall, 'Aperture lens formula corrected for space charge in the electron stream', *IRE Transactions on Electron Devices,* vol. 4, pp. 132–134, 1957.

[15] K. Amboss, 'The effect of the anode aperture in conical flow Pierce guns', *Journal of Electronics and Control,* vol. 13, pp. 545–572, 1962.

[16] W. Liu and S. Liu, 'PIC simulation study of electron gun with rotational surface cathode', *Frontiers of Electrical and Electronic Engineering in China,* vol. 6, pp. 556–562, 2011.

[17] R. True, 'Gridded Pierce gun design and the PPM focussing of beams from gridded electron guns', in *International Electron Devices Meeting*, pp. 32–35, 1982.

[18] C. C. Cutler and M. E. Hines, 'Thermal velocity effects in electron guns', *Proceedings of the IRE,* vol. 43, pp. 307–315, 1955.

[19] R. True, 'Calculation and design of grids in Pierce guns', in *1989 International Electron Devices Meeting, IEDM '89*, pp. 215–218, 1989.

[20] W. B. Herrmannsfeldt, 'Numerical design of electron guns and space charge limited transport systems', *Nuclear Instruments and Methods in Physics Research,* vol. 187, pp. 245–253, 1981.

[21] J. Petillo *et al.*, 'The MICHELLE three-dimensional electron gun and collector modeling tool: theory and design', *IEEE Transactions on Plasma Science,* vol. 30, pp. 1238–1264, 2002.

[22] L. Kumar and E. Kasper, 'On the numerical design of electron guns', *Optik,* vol. 72, pp. 23–30, 1985.

[23] J. E. Boers, 'Digital computer analysis of axially symmetric electron guns', *IEEE Transactions on Electron Devices,* vol. 12, pp. 425–435, 1965.

[24] V. Hamza, 'Convergence and accuracy criteria of iteration methods for the analysis of axially symmetric and sheet beam electrode shapes with an emitting surface', *IEEE Transactions on Electron Devices,* vol. 13, pp. 485–493, 1966.

[25] B. M. Lewis *et al.*, 'Design of an electron gun using computer optimization', *IEEE Transactions on Plasma Science,* vol. 32, pp. 1242–1250, 2004.

[26] R. True, 'The deformable relaxation mesh technique for solution of electron optics problems', in *International Electron Devices Meeting*, pp. 257–260, 1975.

[27] G. Faillon, 'Technical and industrial overview of RF and microwave tubes for fusion', *Fusion Engineering and Design,* vol. 46, pp. 371–381, 1999.

[28] T. Lee *et al.*, 'A fifty megawatt klystron for the Stanford Linear Collider', in *International Electron Devices Meeting*, pp. 144–147, 1983.

[29] M. N. Pinto *et al.*, 'An electron gun design for a C-band TWT', in *SBMO/IEEE MTT-S International Microwave & Optoelectronics Conference*, pp. 937–940, 2011.

[30] W. R. Smythe, *Static and Dynamic Electricity*, 2nd ed. New York: McGraw-Hill, 1950.

[31] R. B. Nelson, 'Shaping the magnetic field for convergent confined-flow electron guns', in 6th *International Conference on Microwave and Optical Generation and Amplification*, Cambridge, England, pp. 60–64, 1966.

[32] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, UK: Research Studies Press, 1995.

[33] J. R. Legarra *et al.*, 'A convergent confined-flow focusing system for millimeter wave tubes', in *1983 International Electron Devices Meeting*, pp. 137–140, 1983.

[34] K. Amboss, 'Studies of a magnetically compressed electron beam', *IEEE Transactions on Electron Devices,* vol. 16, pp. 897–904, 1969.

[35] J. A. Seeger, 'Magnetic compression of axially symmetric Brillouin-focused electron beams', *IEEE Transactions on Electron Devices,* vol. 16, pp. 1–5, 1969.

[36] E. A. Ash, 'Compression and reflection of Brillouin beams', *Journal of Electronics and Control,* vol. 15, pp. 402–417, 1963.

[37] M. Müller, 'Magnetische Elektronenoptik in Langstrahl-Verstärker Rohren', *Telefunken Zeiting,* vol. 26, pp. 95–101, 1953.

[38] V. Bevc *et al.*, 'On the design of the transition region of axisymmetric, magnetically focused beam valves', *Journal of the British Institution of Radio Engineers,* vol. 18, pp. 696–705, 1958.

[39] J. L. Rawls, 'An analogue study of periodic permanent magnet focusing', *International Journal of Electronics,* vol. 25, pp. 165–175, 1968.

[40] A. L. Igritsky, 'Focusing a beam of electrons in the region of entry into the magnetic field of a traveling-wave tube', *Radio Engineering and Electron Physics,* vol. 8, pp. 117–123, 1963.

[41] J. L. Rawls *et al.*, 'PPM focusing of convergent beams emerging from partially shielded cathodes', *IEEE Transactions on Electron Devices,* vol. 14, pp. 301–305, 1967.

[42] N. B. Kramer and E. G. Todd, 'Design of hollow and strip beam guns', *IEEE Transactions on Electron Devices,* vol. 10, pp. 394–404, 1963.

[43] S. K. Wong *et al.*, 'The design of sheet-beam electron-gun for high-power microwave sources', in *Particle Accelerator Conference*, pp. 1043–1045, 1999.

[44] K. T. Nguyen *et al.*, 'High-perveance W-band sheet-beam electron gun design', in *ICOPS 2008: IEEE 35th International Conference on Plasma Science*, pp. 179–180, 2008.

[45] J. Pasour *et al.*, 'Demonstration of a 100-kW solenoidally focused sheet electron beam for millimeter-wave amplifiers', *IEEE Transactions on Electron Devices,* vol. 58, pp. 1792–1797, 2011.

[46] J. R. M. Vaughan, 'Synthesis of a hollow-beam gun based on the Pierce design', in *International Electron Devices Meeting*, pp. 496–499, 1981.

[47] J. R. M. Vaughan, 'Synthesis of a hollow-beam Pierce gun', *IEEE Transactions on Electron Devices,* vol. 34, pp. 468–472, 1987.

[48] J. R. M. Vaughan, 'Corrections to "Synthesis of a hollow-beam Pierce gun"', *IEEE Transactions on Electron Devices,* vol. 34, p. 1885, 1987.

[49] L. A. Harris, 'Toroidal electron guns for hollow beams', *Journal of Applied Physics,* vol. 30, pp. 826–836, 1959.

[50] H. L. Hartnagel, 'Electrostatically focused electron gun for hollow beams', *Proceedings of the Institution of Electrical Engineers,* vol. 111, pp. 57–66, 1964.

[51] H. Hartnagel, 'Toroidal hollow-beam gun', *Proceedings of the Institution of Electrical Engineers,* vol. 111, pp. 1821–1829, 1964.

[52] A. Staprans *et al.*, 'High-power linear-beam tubes', *Proceedings of the IEEE,* vol. 61, pp. 299–330, 1973.

[53] L. Sivan, *Microwave Tube Transmitters*. London: Chapman and Hall, 1994.

[54] H. E. Gallagher, 'Gridded electron guns for high average power', *IRE Transactions on Electron Devices,* vol. 9, pp. 234–241, 1962.

[55] G. Miram *et al.*, 'Gridded gun design considerations for pulse and cw operation', in *Vakuumelektronik und Displays*, Garmisch Partenkirchen, Germany, pp. 225–233, 1989.

[56] R. True, 'A theory for coupling gridded gun design with PPM focusing', *IEEE Transactions on Electron Devices,* vol. 31, pp. 353–362, 1984.

[57] R. True, 'An ultra-laminar tetrode gun for high duty cycle applications', in *International Electron Devices Meeting*, pp. 286–289, 1979.

[58] O. Doehler, 'Static characteristics of injection crossed-field tubes', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 151–163, 1961.

[59] G. S. Kino, 'A design method for crossed-field electron guns', *IRE Transactions on Electron Devices*, vol. 7, pp. 179–185, 1960.

[60] T. A. Midford and G. S. Kino, 'Some experiments with a new type of crossed-field gun', *IRE Transactions on Electron Devices,* vol. 8, pp. 324–330, 1961.

[61] J. R. M. Vaughan, 'Generalization and synthesis of the Kino gun', *IEEE Transactions on Electron Devices,* vol. ED-28, pp. 41–48, 1981.

[62] T. A. Midford and G. S. Kino, 'Experiments with a new type adiabatic crossed-field gun', *IRE Transactions on Electron Devices,* vol. 9, pp. 431–439, 1962.

[63] T. van Duzer and J. R. Whinnery, 'Noise in crossed-field electron beams', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 327–357, 1961.

[64] R. M. Gilgenbach *et al.*, 'Crossed-field devices', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 289–342, 2005.

[65] M. L. Sisodia and R. P. Wadhwa, 'Noise reduction in crossed-field guns by cathode tilt', *Proceedings of the IEEE,* vol. 56, pp. 94–95, 1968.

[66] R. J. Espinosa and R. R. Moats, 'Broad-band injected-beam crossed-field amplifiers', *IEEE Transactions on Electron Devices,* vol. 24, pp. 13–21, 1977.

[67] C. J. Edgcombe, 'Synthesis of electrodes for axisymmetric systems', *IEE Proceedings – A,* vol. 135, pp. 125–134, 1988.

[68] T. Okoshi, 'An improved design theory of a magnetron injection gun', *IEEE Transactions on Electron Devices,* vol. ED-11, pp. 349–356, 1964.

[69] W. E. Waters, 'A theory of magnetron injection guns', *IEEE Transactions on Electron Devices,* vol. 10, pp. 226–234, 1963.

[70] G. S. Kino and N. J. Taylor, 'The design and performance of a magnetron-injection gun', *IRE Transactions on Electron Devices,* vol. 9, pp. 1–11, 1962.

[71] U. Singh *et al.*, 'Numerical simulation of magnetron injection gun for 1 MW 120 GHz gyrotron', *Progress in Electromagnetics Research Letters,* vol. 16, pp. 21–34, 2010.

[72] J. M. Baird and W. E. S. Lawson, 'Magnetron injection gun (MIG) design for gyrotron applications', *International Journal of Electronics,* vol. 61, pp. 953–967, 1986.

10 Electron Collectors and Cooling

10.1 Introduction

Even the best vacuum tubes do not convert all of the DC input power into radio-frequency power. Therefore the electrons which emerge from the interaction region still carry appreciable power. This must be dissipated safely, in the form of heat, on the electrodes on which the spent electrons are collected. The maximum permissible temperature of the surface is limited by the need to avoid:

- physical damage including re-crystallisation of the copper (at 700 C [1]), local melting, and fatigue fractures caused by cyclical heating in pulsed devices;
- liberation of adsorbed gasses by exceeding the bake-out temperature (350–550 °C) used during evacuation of the tube [2–3]. It has been suggested that the surface temperature of copper should be limited to 300–350 °C [4].

The maximum power density which can be tolerated on the surface depends on the heat transfer from the collector electrodes to some external heat sink, or heat exchanger, and on the method by which the heat is finally dissipated (see Section 10.5).

The impact of electrons on metallic surfaces (usually OHFC copper) results in the liberation of secondary electrons, and the generation of X-rays (see Chapter 18). When the primary electron energies are small (a few keV) the X-rays generated cannot penetrate through the metal wall of the tube, and there is no external hazard. At higher voltages X-rays may escape and suitable shielding must be provided. It is also necessary to ensure that there is no path by which X-rays can pass from the inner surface of the tube through insulators. This is because the attenuation of the X-rays by the insulators (usually ceramics) is much less than that by metals. It is usually necessary to provide X-ray shielding for tubes with an operating voltage greater than 15 kV [5–6] (see Section 20.8).

In gridded tubes and magnetrons the spent electrons are collected directly on the anode of the tube. The dimensions of the anode are limited by the frequency of the tube and it is important to consider the power density on the anode surface during the design process. Because the anode is positive with respect to the rest of the tube any secondary electrons which may be liberated are re-collected on the anode and do not cause any problems. The exception to this is when the potential of the anode of a tetrode falls below that of the screen grid so that secondary electrons from the

anode are then collected on the grid. This is undesirable in power tetrodes and that region of operation must be avoided (see Chapters 6 and 12).

In klystrons, travelling-wave tubes (TWTs), inductive output tubes (IOTs), and gyrotrons the electrons are collected on an electrode, designed for the purpose known as the *collector*. Because this electrode does not form part of the RF interaction structure of the tube it is possible to allow the beam to expand under the influence of its space charge so that the local power density is reduced. It is usual for the collector to be electrically isolated from the body of the tube so that the body current can be monitored separately to protect the tube. The power density on the surface of the collector can be reduced still further by operating the collector at a potential which is negative with respect to the tube body. This technique, known as collector depression, is discussed in Section 10.3. The advantages of collector depression are [7]:

• The DC input to the tube is reduced, and the overall efficiency is increased.
• The power which must be dissipated is reduced so that the cooling pumps and fans can be smaller, and less power is required to run them.
• The intensity and penetrating power of the X-rays generated by high voltage tubes is reduced, so that less shielding is required.

However, it is more difficult to cool electrodes which are not at earth potential. In addition, great care must be taken to ensure that no primary electrons are reflected by the collector, and that any secondary electrons liberated are trapped within it. Electrons which flow back into the interaction region form an unwanted feedback path which may cause the tube to become unstable. The advantages of collector depression can be increased by the use of multi-element depressed collectors. However, the use of many electrodes increases the complexity and cost of a tube and its power supply, and it is necessary to balance the advantages and disadvantages in determining the design of the collector. The development of multi-element depressed collectors was stimulated especially by the need to achieve the highest possible overall efficiencies for tubes for use in satellites and spacecraft.

10.2   Linear Beam Tube Collectors

At the exit of a linear beam tube the spent beam is freed from the magnetic focusing field and allowed to spread out, as shown in Figure 7.8. The kinetic energy of the electrons can then be dissipated on the surface of a collector which typically has the form shown in Figure 10.1. The emission of secondary electrons in single stage collectors is not a problem, provided that they cannot pass out of the collector and flow back through the interaction region of the tube. For this reason it is important that the magnetic focusing field does not penetrate into the collector where it can serve to guide electrons back towards the interaction region. When the impact energies of the electrons on the surface of the collector are high (above about 1 keV) then the dominant effect is the elastic reflection of the primary electrons (see Section 18.3).
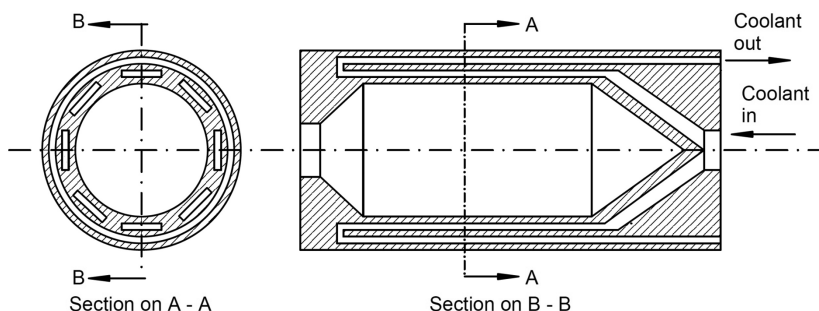
Figure 10.1: Arrangement of a liquid cooled collector.
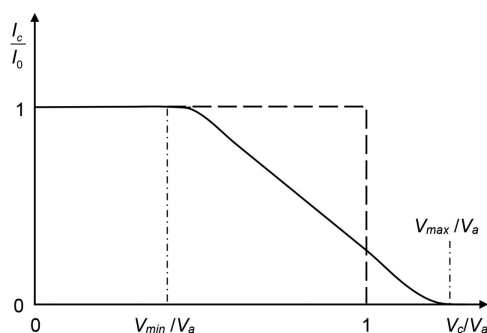
Figure 10.2: Schematic spent-beam distribution curve.

For this reason the collector usually has the form of a long tube with a conical end so that electrons which strike the walls cannot be reflected directly back towards the tube body. It is necessary to ensure that the collector does not have any resonant RF modes which coincide with the operating frequency of the tube, or its harmonics. If any resonances are excited then the RF fields may perturb the electron paths in unexpected ways.

The interaction which takes place in any kind of linear beam tube results in a net transfer of energy from the electron beam to the RF output. Therefore the average kinetic energy of the electrons in the spent beam is less than that at the input to the tube. The electrons in the spent beam do not all have the same velocity on leaving the tube, and some of them usually have velocities greater than the initial beam velocity. Figures 10.2 shows, schematically, the fraction of the beam current which would be collected on an electrode at a potential $-V_c$ with respect to the tube body, normalised to the potential of the cathode $(-V_a)$[1]. This curve assumes that no primary electrons are reflected and that any secondary electrons generated are re-collected by the electrode.

---

[1] Note that in this Chapter the potentials of the electrodes are referred to the potential of the tube body rather than to the potential of the cathode.

The detailed shape of the spent-beam curve depends upon the type of tube, and the specific operating conditions. However, the spent-beam curves for all tubes show the same features:

- As the potential of the collecting electrode is made more negative the whole of the beam current is collected until a potential $-V_{min}$ is reached. Thus all electrons have energies greater than V_{min} electron volts.
- When $|V_c| > |V_{min}|$ the fraction of current collected decreases monotonically as the collection potential is made more negative.
- For potentials beyond $-V_{max}$ the collected current falls to zero indicating that no electrons have energies greater than V_{max} electron volts. It is common to find that $V_{max}/V_a > 1$, which shows that some of the electrons have been accelerated by the RF interaction.

It should be noted that a graph such as that shown in Figure 10.2 represents the time-average of the velocity distribution. In practice the electron velocity distribution varies during the RF cycle.

The power carried by electrons which enter the collector in an energy range (in electron volts) dV at energy V is $V\,dI$. It follows that the total power carried by the spent beam is represented by the area under the curve in Figure 10.2. The dashed rectangle shows the energy distribution in the unmodulated beam, and the area of the rectangle represents the DC beam power $I_0 V_a$. If losses in the tube are neglected the difference between the area of the rectangle and the area under the spent-beam curve represents the RF power output of the tube. Since the power in the spent beam is less than that in the initial beam it follows that the thermal dissipation in the collector is less than the initial beam power. In most cases, however, the collector is designed to be able to dissipate the full DC beam power. If this is not the case, it is essential that the DC input power to the tube is cut, whenever the RF input is removed, to avoid damaging the collector by overheating it.

In order to design a collector we need to know how the power is distributed on its surface. For simplicity let us consider an unmodulated beam of radius b, having uniform current density and electron velocity, emerging from a magnetic field mB_B into a concentric cylindrical collector whose wall radius is r_w, as shown in Figure 10.3. Each electron within the beam will follow the same normalised path given by

$$Z = S(R), \tag{10.1}$$

where $Z = \beta_p z$, R is the radius of the trajectory normalised to its initial value, and the function S is defined by the right-hand side of (7.82). The axial position at which a trajectory whose initial radius is r strikes the wall of the collector is given by

$$Z = S\left(\frac{r_w}{r}\right). \tag{10.2}$$
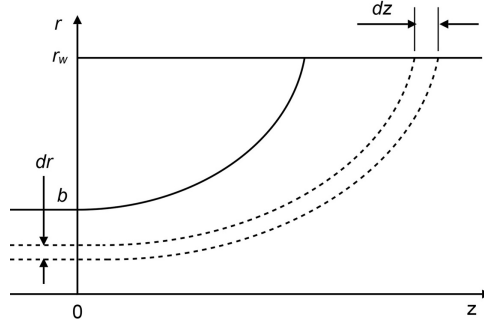
Figure 10.3: Expansion of the spent beam into the collector.
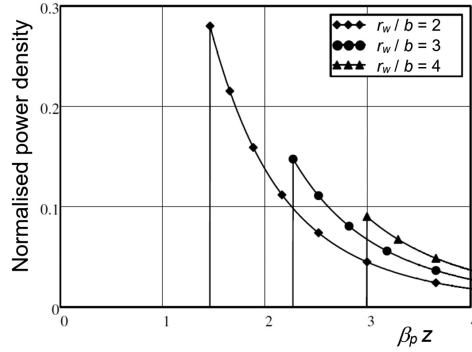
Figure 10.4: Power density on the surface of a collector.

The electrons in an annulus of radial thickness dr strike the collector along a section of normalised length dZ such that

$$\frac{dZ}{dr} = \frac{d}{dr} S\left(\frac{r_w}{r}\right). \tag{10.3}$$

The ratio of the power density at the wall of the collector to that in the initial beam is given by the ratio of the area of the element of the wall to that of the annulus. Thus, the normalised collector power density is

$$P_c = \frac{r}{r_w}\frac{dZ}{dr} = \frac{r}{r_w}\left\{\frac{d}{dr} S\left(\frac{r_w}{r}\right)\right\}^{-1}. \tag{10.4}$$

Since Z and P_c can be calculated as functions of r from (10.2) and (10.4) it is possible to plot the variation in the power density, normalised to the power density in the initial beam, against axial position, as shown by the example in Figure 10.4 (see Worksheet 10.1). It is evident that the power density is greatest at the point where

the beam first strikes the wall of the collector and it diminishes rapidly thereafter with increasing Z.

The power carried by the spent beam is less than that in the unmodulated beam by the RF power extracted from it. There are also two other factors which affect the power density on the collector wall when the beam is modulated:

- The beam is bunched so that the peak current is greater than the DC beam current I_0. The extent of the increase depends upon the type of tube and the drive level. At saturation the peak current in klystrons and TWTs may be as high as $4I_0$.
- The average velocity of the electrons in the bunch is less than the DC velocity.

A fuller discussion of these issues is to be found in Chapters 12, 13, and 14. For an IOT, only the second point applies because the peak current in the bunch is equal to the peak cathode current (see Chapter 12). As a consequence of both these factors the beam spreads more rapidly than it would if it were not modulated. In addition, the electrostatic potential within the beam is depressed by the increased space-charge in the bunch, and by the increase in the radius of the conducting boundary surrounding the beam. The potential depression can be estimated from (7.8) with $r = 0$. If we denote the multiplication factors for the current and the velocity by F_I and F_u then (7.8) can be written in the form

$$\frac{V}{V_a} = \frac{F_I}{F_u} \beta_p^2 b^2 \left[1 + 2 \ln \left(r_w / b \right) \right], \tag{10.5}$$

where r_w is the inner radius of the collector and relativistic effects have been ignored. So, for example, in a tube where $F_I = 2$, $F_u = 0.6$, $\beta_p b = 0.12$ and $r_w / b = 10$ the ratio $V/V_a = 0.27$. The reduction in the electron velocity caused by the potential depression can, therefore, be big enough to be important. It increases the charge density and causes the potential to be still further depressed so that, in extreme cases the ratio can reach unity. A virtual cathode is then formed at the entrance to the collector and some of the electrons are reflected by it. Because it is highly undesirable for any electrons to find their way back into the interaction region, the formation of a virtual cathode must be avoided. The easiest way to achieve this is to arrange for the inner wall of the collector to be conical in the region between its entrance and the point at which the electrons first strike it. It should be possible to keep the ratio r_w / b to less than 4 in the critical region. In the example given above the voltage ratio could be reduced to around 0.18 in this way.

Particle in cell (PIC) code simulations of collectors show that, if that the space potential depression is small compared with the beam voltage, the bunches spread out into the collector without diffusing very much in the axial direction. It is therefore possible to estimate the power density distribution on the inside of the collector wall at different points in the RF cycle by computations using the instantaneous values of the parameters.

10.3   Collector Depression

Examination of the schematic spent beam distribution curve in Figure 10.2 shows that the collector can be biased negatively to potential $V_{\min}$ with respect to the tube body without causing any of the electrons to be repelled. The electrons then strike the collector with reduced energies so that energy is recovered. This is achieved using the power supply arrangement shown in Figure 10.5. The collector is then said to be *depressed* and the net DC input power to the tube is

$$P_{DC\,in} = I_c\left(V_a - V_c\right) + I_b V_a = I_0 V_a - I_c V_c. \tag{10.6}$$

where I_c is the collector current, I_b is the body current and the potential of the collector with respect to the tube body is $-V_c$. Substituting from (10.6) into (1.18) the RF efficiency is

$$\eta_{rf} = \frac{P_2}{I_0 V_a - I_c V_c}. \tag{10.7}$$

where P_2 is the useful RF output power at the fundamental frequency, and it has been assumed that the gain of the tube is great enough for the RF input power to be negligible. The RF efficiency without collector depression is

$$\eta_{rf0} = \frac{P_2}{I_0 V_a}, \tag{10.8}$$

so the RF efficiency has been increased by depressing the collector. The depressed collector efficiency increases, as V_c is increased, up to the point at which electrons start to be reflected. An improvement in efficiency by a factor in the range 1.3 to 1.4 is typical of what can be achieved by simple techniques. The improvement in
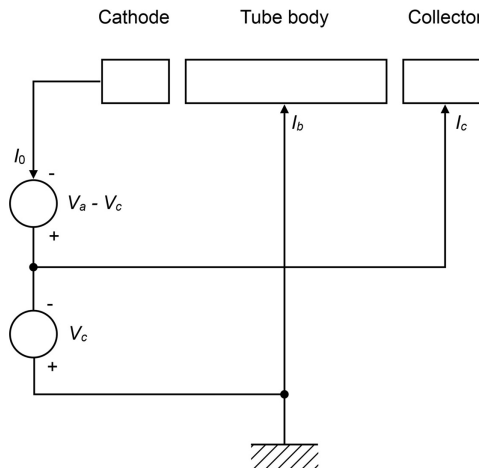
Figure 10.5: Circuit diagram for a tube with a single-stage depressed collector.
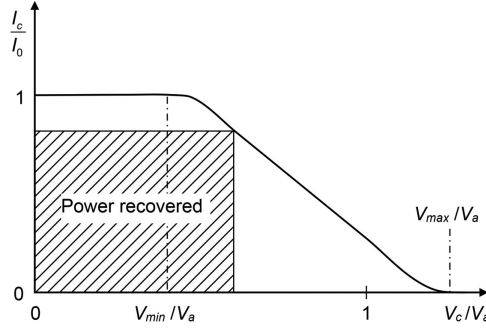
Figure 10.6: Power recovered by a single-stage depressed collector.

efficiency is a result of the reduction in the heat dissipated, but the price paid is increased complexity of the power supply and the problem of cooling a high voltage electrode. If the potential depression is increased further then some of the slower electrons are reflected. They can be prevented from entering the tube body by including an earthed electrode between the magnetic pole piece and the collector on which any reflected electrons can be collected [8]. The collector then has two electrodes, but it is known as a single-stage depressed collector because only one electrode is depressed. The extent to which the efficiency of a tube can be increased is limited by the reflection of electrons.

Since the power dissipated in the collector is the integral under the spent beam curve it follows that the effect of depressing the collector is to reduce the integral by the shaded area shown in Figure 10.6. The power recovered is given by

$$P_{rec} = I_c V_c. \tag{10.9}$$

This increases until $V_c = V_{min}$ and decreases for greater potential depression because some electrons are reflected and collected at zero potential. The power carried by the spent beam entering the collector is

$$P_{ent} = (1 - \eta_e) I_0 V_a, \tag{10.10}$$

where η_e is the electronic efficiency defined in (1.20) and RF losses within the tube are assumed to be negligible. Since the body current is typically of the order of 1% of the beam current we will assume that $I_c \approx I_0$. The efficiency of the collector is defined as the ratio of the power recovered to the power entering the collector

$$\eta_{col} = \frac{I_0 V_c}{(1 - \eta_e) I_0 V_a} = \frac{V_c}{(1 - \eta_e) V_a}. \tag{10.11}$$

Then from (10.7) the RF efficiency with a depressed collector can be written

$$\eta_{rf} = \frac{P_2}{I_0 V_a \left(1 - \eta_{col} (1 - \eta_e)\right)} \tag{10.12}$$
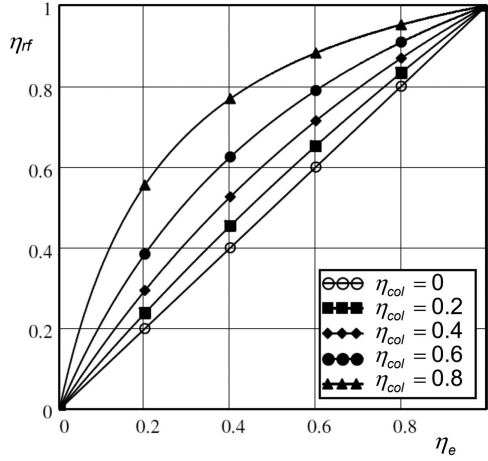
Figure 10.7: Dependence of the RF efficiency of a tube on electronic efficiency and collector efficiency
(copyright 1970, NASA, reproduced, with permission, from [9]).

so that, making use of (1.22)

$$\eta_{rf} = \frac{\eta_e \eta_c}{\left(1 - \eta_{col}\left(1 - \eta_e\right)\right)},\tag{10.13}$$

where η_c is the circuit efficiency, and interception of current on the tube body has been neglected in this simplified derivation. This equation does not depend upon the details of the collector so it is also valid for the multi-element collectors discussed in the next section.

If the effects of RF losses are ignored $\eta_c = 1$. Then the effectiveness of collector depression can be shown by plotting η_{rf} against η_e with η_{col} as a parameter. Figure 10.7 shows that, for a given electronic efficiency, the RF efficiency increases as the collector efficiency increases. The increase in the RF efficiency is greatest when the electronic efficiency is low.

10.3.1   Multi-Element Depressed Collectors

Consideration of Figure 10.6 shows that still greater improvements in efficiency can be achieved by using a multi-element depressed collector, as shown in Figure 10.8. The electrons are sorted into energy ranges and collected on electrodes at potentials corresponding to the lowest energy in each range. The equation for the collector efficiency (10.11) now takes the form

$$\eta_{col} = \frac{\sum_{n=1}^{N} I_{cn} V_{cn}}{\left(1 - \eta_e\right) I_c V_a}.\tag{10.14}$$
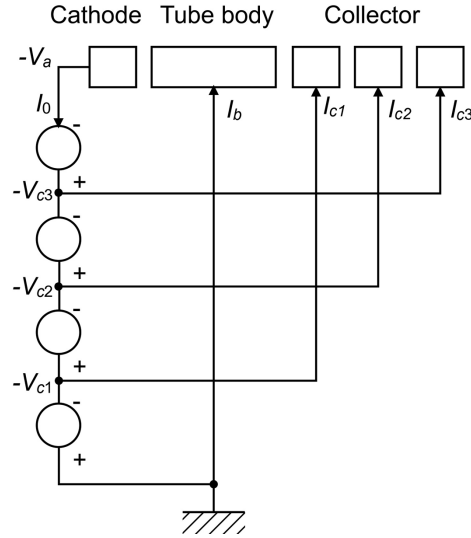
Figure 10.8: Circuit diagram for a tube with a three stage depressed collector.
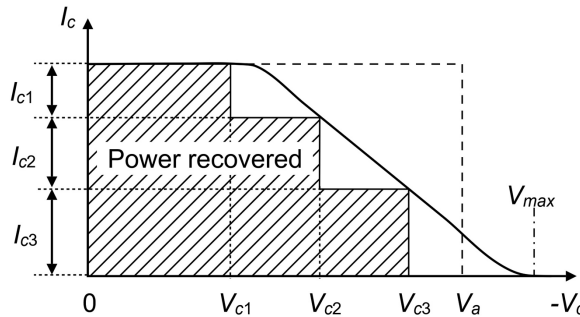
Figure 10.9: Schematic spent-beam diagram for a tube with a three-stage depressed collector.

where V_{cn} is the potential of the nth collector electrode, I_{cn} is the current collected by it and

$$I_c = \sum_n I_{cn}.\tag{10.15}$$

The corresponding spent-beam curve is shown in Figure 10.9 and it is clear that the shaded area, which represents the power recovered, is much greater than in Figure 10.6. Thus it is possible to increase the efficiency of a collector by increasing the number of electrodes. Note that Figure 10.9 assumes that no current is collected at earth potential. In practice it is usual for the potential of the first electrode to be equal to that of the tube body and a small current is collected by it even with a well-designed collector. The potentials of the electrodes do not need to be equally spaced as shown in Figure 10.9 but can be chosen to maximise the power recovered.

Useful insight into the performance of a tube with a multi-element depressed collector can be gained by considering a simplified model adapted from that described by Kosmahl [10]. Let us assume that the spent-beam curve in Figure 10.9 can be approximated by a trapezium whose sloping side is defined by voltages $V_{\min}$ and $V_{\max}$. The collector has N depressed electrodes at equally spaced potentials which collect equal fractions of the beam current. The power entering the collector is

$$P_{ent} = \frac{1}{2} I_0 \left(V_{\max} + V_{\min} \right), \tag{10.16}$$

where it is assumed that the body current is zero. From the geometry of the diagram the power recovered is

$$P_{rec} = P_{ent} - \frac{1}{2N} I_0 \left(V_{\max} - V_{\min} \right). \tag{10.17}$$

Thus the collector efficiency is

$$\eta_{col} = \frac{P_{rec}}{P_{ent}} = 1 - \frac{\left(V_{\max} - V_{\min} \right)}{N \left(V_{\max} + V_{\min} \right)}. \tag{10.18}$$

Now the electronic efficiency of the tube is

$$\eta_e = \frac{P_{DC\,in} - P_{ent}}{P_{DC\,in}} = 1 - \frac{V_{\max} + V_{\min}}{2V_a} \tag{10.19}$$

so that

$$V_{\max} + V_{\min} = 2V_a \left(1 - \eta_e \right) \tag{10.20}$$

and, substituting for $\left(V_{\max} + V_{\min} \right)$ in (10.18)

$$\eta_{col} = 1 - \frac{1}{N} \left(\frac{V_{\max}}{V_a \left(1 - \eta_e \right)} - 1 \right). \tag{10.21}$$

If it is assumed that there are no accelerated electrons, so that $V_{\max} = V_a$, then (10.21) reduces to

$$\eta_{col} = 1 - \frac{\eta_e}{N \left(1 - \eta_e \right)}. \tag{10.22}$$

In the form of this equation given by Kosmahl the factor N is replaced by $(N - 1)$ because of the presence of an additional electrode at potential V_a which collects no current [10]. Equation (10.22) shows that the collector efficiency increases as N increases, as would be expected, but that it decreases as the electronic efficiency increases because the fraction of the power entering the collector which can be recovered decreases. Substituting for the collector efficiency from (10.22) in (10.13),

and ignoring circuit losses, we find that the efficiency of the tube with a depressed collector is

$$\eta_{rf} = \frac{N}{N+1}.$$ (10.23)

Thus, this simple model predicts that the efficiency of the tube depends only on the number of stages in the collector and is independent of the electronic efficiency. The overall efficiency increases as the number of stages increases but the improvement obtained by adding extra stages is small if N is greater than 5. Thus, for most tubes, it is not worth using more than four or five stages because of the additional cost and complexity of the collector and of the power supply for the tube.

The theoretical efficiency indicated by (10.23) is not achieved in practice because the spent beam curve is not normally a trapezium, and because no account has been taken of the effect of electrons which enter the collector with energies greater than V_a. Equation (10.23) also assumes that each electron is collected as efficiently as possible. This is not always the case, as is shown in the next section. In particular, if the electronic efficiency changes as a result of changes in the drive level, then the minimum electron energy V_{min} changes. With fixed electrode potentials the number of stages which are effective changes, and the potentials are not generally optimal. In practice the electrode potentials are chosen to maximise the shaded area below the spent-beam curve. It is not necessarily the case that the best overall efficiency is obtained by designing a tube to achieve the best electronic efficiency and then adding a multi-element depressed collector. This is because the shape of the spent beam curve can vary considerably as the design of the tube is changed. It is therefore necessary to ensure that the shape of the spent-beam curve is such that as much of the energy as possible can be recovered by the collector [11]. Changes in the design of the tube may also result in changes in the circuit efficiency.

10.3.2 Non-Ideal Multi-Element Depressed Collectors

In an ideal multi-element collector each electron would be collected at the most negative potential possible for its energy. Thus all electrons for which $V_n \geq V_e > V_{n+1}$ would be collected at potential V_n. It has been shown experimentally that this situation can be approached quite closely, provided that the space-charge density is very low [12]. However, as the space-charge density is increased some of the electrons in this energy range continue to be collected at potential V_{n-1} so that the collection efficiency is reduced [12–13]. To show how this affects the collector efficiency we will assume that the probability that an electron with energy V_e is collected at potential V_n is

$$p_n(V_e) = \frac{V_e - V_n}{V_{n+1} - V_n}$$ (10.24)

and the probability that it is collected at potential V_{n-1} is

$$p_{n-1}(V_e) = 1 - p_n(V_e). \tag{10.25}$$

If $V_e < V_1$ then the electrons are collected at potential $V_0 = 0$ so that $p_0 = 1$. It is also assumed that $p_N = 1$ when $V_e > 2V_N - V_{N-1}$. Figure 10.10 shows the probability that electrons will be collected on each electrode of a five-stage collector as a function of the normalised electron energy (see Worksheet 10.2). The dashed rectangles show the current collected by an ideal collector and the triangles show the current collected according to the model defined by (10.24) and (10.25).
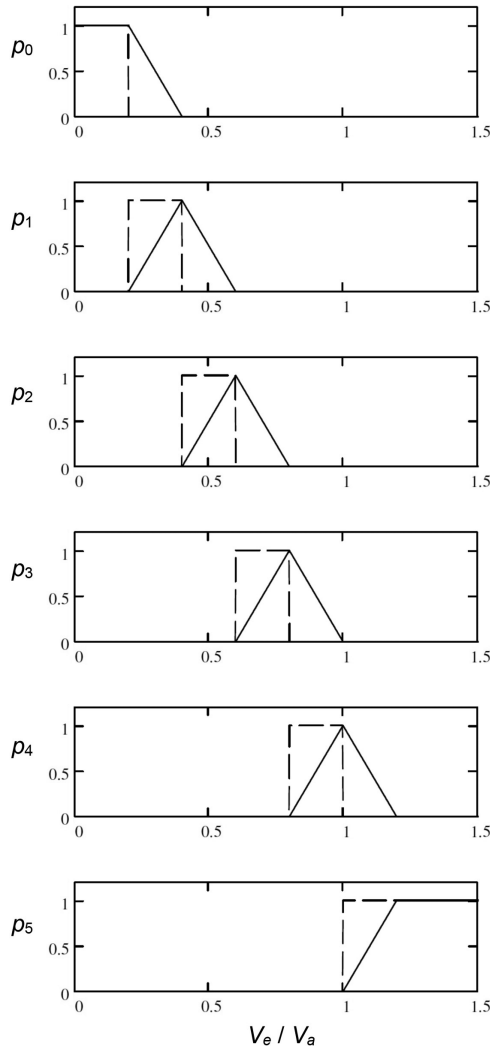
Figure 10.10: Theoretical probabilities that an electron having energy V_e will be collected by the electrodes of a five-stage depressed collector.
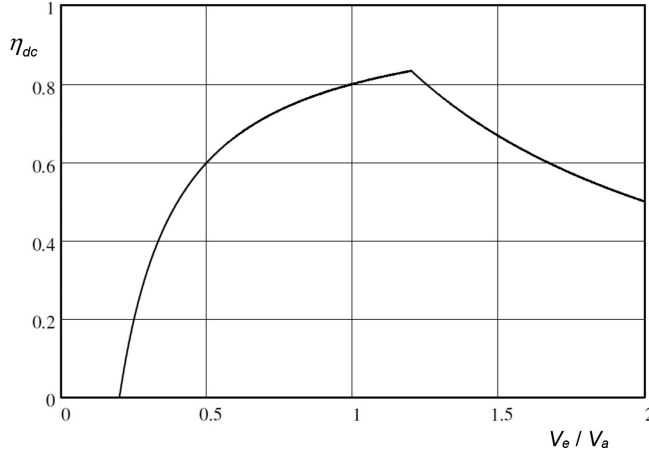
Figure 10.11: Theoretical DC collector efficiency as a function of electron energy.

The DC collector efficiency, that is the collector efficiency for mono-energetic electrons, can be computed as a function of the normalised electron energy of the electrons from

$$\eta_{dc}(V_e) = \frac{1}{V_e} \sum_n p_n(V_e) \cdot V_e. \tag{10.26}$$

Figure 10.11 shows the DC collector efficiency for a five-stage collector with equally spaced voltages. It is found that the shape of this curve varies very little if the linear variation of probability is changed to a sine or sine squared curve (see Worksheet 10.2). The curve is similar to that computed for an experimental five-stage collector [13]. The differences can be explained by the collection of some of the current on electrodes at less favourable potentials. This is caused partly by imperfections in the design, and partly by secondary electrons.

It has been shown that the RF collector efficiency can be computed to good accuracy from the DC collector efficiency and the spent-beam curve [13]. If the current collected on a single depressed electrode at a potential V_c is I_c then the fraction of the total current in the small energy range dV_e at energy V_e is

$$dI_c = \left(\frac{dI_c}{dV_c} \right)_{V_e} dV_e \tag{10.27}$$

and the power entering the collector associated with this current is

$$dP_{ent} = dI_c V_e = \left(\frac{dI_c}{dV_c} \right)_{V_e} V_e \, dV_e. \tag{10.28}$$

Thus the total power entering the collector is given by

$$P_{ent} = \int_{V_{min}}^{V_{max}} \left(\frac{dI_c}{dV_c} \right)_{V_e} V_e \, dV_e, \qquad (10.29)$$

where the integral is taken over the full range of electron energies in the spent beam. Now the fraction of the power entering the collector in the energy range dV_e which is recovered is $\eta_{dc}(V_e)$. Thus, the total power recovered is

$$P_{rec} = \int_{V_{min}}^{V_{max}} \left(\frac{dI_c}{dV_c} \right)_{V_e} V_e \, \eta_{dc}(V_e) \, dV_e. \qquad (10.30)$$

The collector efficiency with an RF modulated beam is found by dividing (10.30) by (10.29). Modern multi-element depressed collectors have collector efficiencies exceeding 80% and therefore the majority of the electrons are being collected at the most favourable potential [14–16] (see Worksheet 10.2).

10.4   Design of Multi-Element Depressed Collectors

The design of a multi-element depressed collector begins with knowledge of the spent-beam curve for the tube. The shape of the curve depends upon the type of tube, and upon the level of the RF input power (see Figures 12.19, 13.15, and 14.16). It also depends on the frequency within the working bandwidth of the tube. Although the energy distribution in the spent beam can be determined experimentally, it is now usual to obtain the information from a large-signal computer model. For this purpose a 2.5-dimensional model is needed since the radial and angular velocities of the electrons are important. The number of collector stages is chosen as a compromise between collector efficiency and complexity, and the potentials of the stages are chosen to maximise the energy recovered.

The task of designing a multi-element depressed collector is one of arranging the electrodes in such a way that the electrons are sorted according to their initial energies, and collected with minimum impact energies. Since these impact energies may be of the order of a few hundred electron volts, the generation of secondary electrons is more important than reflection of primary electrons, though both processes occur (see Section 18.3). It is important that any secondary electrons, or reflected primary electrons, are collected on the surface from which they were emitted to avoid degradation of the collector efficiency. It is also important that there should be no back-streaming electrons flowing from the collector into the interaction region. In a multi-element depressed collector the presence of a number of electrodes at different potentials means that, in general, the electrons are flowing through a succession of electrostatic lenses which may be either converging or diverging. Since the electrons have a range of initial vector velocities, and the effects
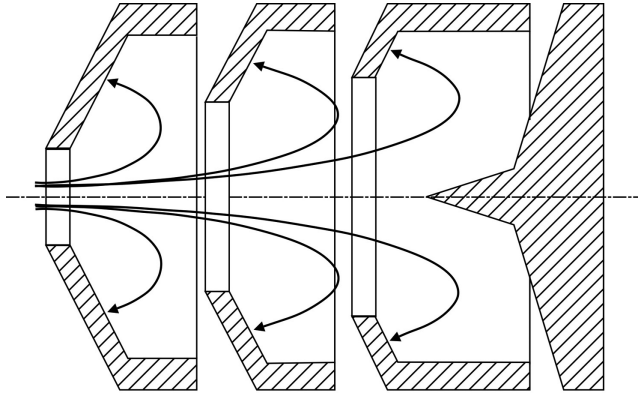
Figure 10.12: Arrangement of a symmetrical multi-element depressed collector.

of the space-charge field cannot be ignored, the problem of synthesising the design of the collector is a difficult one.

Multi-element depressed collectors can be classified as either symmetric (that is axi-symmetric) or asymmetric [17]. Efficient symmetric collectors have the general arrangement shown in Figure 10.12. The electrons are collected on the backs of the electrodes so that any secondary electrons are emitted into a retarding field and tend to be returned to the electrode from which they were emitted. Two essential features of collectors of this type are the conical shape of the last electrode, and the spike which protrudes from it along the axis of symmetry. The purpose of the spike is to provide a radial electric field to aid the dispersion of the beam in a region where the space-charge forces are weak. It also opposes the motion of secondary electrons generated on the cone towards the axis. Ideally the potential of the cone and the spike should be $-1.5V_a$ so that few, if any, electrons would have sufficient energy to reach them. It is not always convenient to provide this potential from the power supply to the tube, and it is more usual for the final electrode to be at the potential of the cathode.

Kosmahl proposed a 'Dispersive-Lens collector' in which the electrodes were placed along equipotential surfaces determined from a given variation of the potential along a boundary surrounding the collector [17–18]. This choice of the positions of the electrodes means that the electric fields within the collector are independent of the sizes of the apertures in the electrodes. The best aperture sizes are those which are just large enough to allow all the electron trajectories to pass through them. Since the field variation is smooth within the collector the lens effects are minimised. In principle it should be possible to design a collector of this type by finding mutually consistent solutions of Poisson's equation, and the equations of motion of the electrons, subject to a prescribed potential distribution on a cylindrical or spherical boundary surrounding the beam. A Dispersive-Lens collector built for a 200W 12 GHz coupled-cavity TWT had a collector efficiency of 81% and it was thought that 90% should be achievable [17].

An alternative approach is the 'Individual-Lens collector' [17] which has a similar form, except that the potentials of the electrodes are not dependent on a prescribed boundary potential distribution, but may be chosen freely. The potential distribution around each electrode can then form either a converging or a diverging lens and it is found that this can be beneficial to the operation of the collector. An Individual-Lens collector is normally larger that a Dispersive-Lens collector for the same tube but the collector efficiency which can be achieved is greater. The design of an Individual-Lens collector for a dual-mode TWT was described by Dayton et al. [19].

Most of the early work on efficient multi-element depressed collectors concentrated on symmetric designs because, at that time, no computer programs existed which could be used to study and optimise asymmetric collectors. Nevertheless, it was recognised that it was not possible to design a symmetric collector which was completely free from back-streaming [18]. Hechtel pointed out that, if space-charge effects are ignored, the electric field between a pair of electrodes can be used as an electron velocity spectrometer [20]. He observed that some field distributions are inherently focusing whereas others are defocusing. Since electrons, having the same velocity, enter the collector with a range of angles relative to the axis it is desirable to employ a focusing field which causes electrons which start from a point with the same velocity, but at different angles, to arrive at a single point when they strike an electrode. In particular he showed that the fields in symmetric collectors of the types described above are inherently defocusing. However, collectors where the beam is not injected on the axis of symmetry the field can be made to be focusing. Hechtel showed that, of the types of collector he studied, the hyperbolic asymmetric collector (Figure 10.13) could be expected to have higher efficiency than
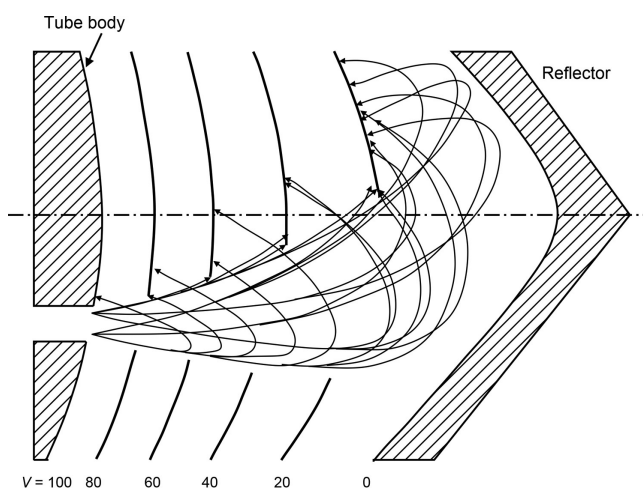
Figure 10.13: Arrangement of Hechtel's hyperbolic asymmetric collector (copyright 1977, IEEE, reproduced, with permission, from [20]).
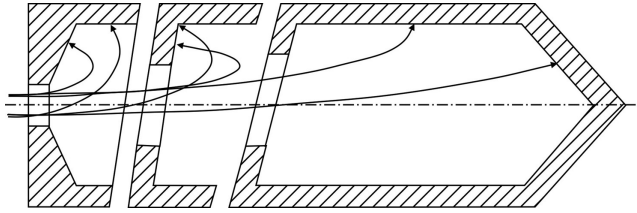
Figure 10.14: Arrangement of an asymmetric multi-element depressed collector.

a parabolic symmetric collector but this conclusion was challenged by Kosmahl [17]. The use of an asymmetric design also greatly reduces the possibility of back-streaming. Hechtel's hyperbolic collector is potentially rather bulky and the theoretical electrode shapes must be approximated by planes and cones for simplicity of manufacture.

The wide variety of possible configurations for symmetric and asymmetric collectors which can exist was classified by Mihran and Neugebauer [9]. The collectors which have been described so far have zero magnetic field and non-uniform electric field. An alternative type of collector which is commonly used because it is compact is illustrated in Figure 10.14. The collector elements are cylindrically symmetric but meet at planes which are at an angle to the axis. This produces a transverse component of the electric field which sweeps secondary electrons sideways to avoid back-streaming. An alternative way of introducing asymmetry into a collector is by use of a transverse magnetic field. This can conveniently be achieved empirically by placing a small permanent magnet alongside an axisymmetric collector. Some high-efficiency collectors employ combinations of tilted electric fields and magnetic fields, including continuation of the PPM field into the collector [16].

The design and optimisation of multi-element depressed collectors requires the use of three-dimensional computer models which can take account of space-charge, reflected primary electrons, and secondary electrons [21–24]. More recently, genetic algorithms have been used for efficient optimisation of the shapes and potentials of the electrodes of both symmetric and asymmetric collectors [22, 25–27].

10.4.1   Suppression of Secondary Electrons

Problems associated with secondary electrons can be reduced by ensuring that the secondary electron emission coefficient of the electrode surfaces is low (see Section 18.3). This is achieved by using materials which have low secondary electron emission coefficients, such as carbon and graphite, and/or by creating a textured surface which tends to trap the secondary electrons. This was recognised very early in the development of depressed collectors, and tests were carried out with electrodes which were carbonised or coated with lamp-black (carbon black) [8]. These coatings have very low secondary electron yield but are unsuitable for use in tubes

because of their poor adhesion [28]. Subsequently, several techniques have been developed [29–30]:

- Oxygen free high conductivity copper is commonly selected for collector electrodes because of its high thermal and electrical conductivity and ease of fabrication. However, the secondary electron yield of copper is greater than unity over a wide range of primary impact energies commonly found in depressed collectors. For this reason surface coatings of materials such as titanium carbide, and surface texturing by ion bombardment, have been used to reduce the secondary electron yield. A low-cost technique for depositing carbon on copper electrodes is described in [31].
- Pyrolytic graphite is a strong, lightweight, material with low secondary electron emission yield. It was found in comparative tests that a collector with electrodes made of pyrolytic graphite had an efficiency which was several percentage points better than a collector having copper electrodes with the same dimensions. Pyrolytic graphite with a textured surface gave results which were better then those with a smooth surface and virtually identical to those obtained using carbon black [28].
- High-purity isotropic graphite has adequate strength for use in collectors and is readily machined. The secondary electron yield when there is a surface texture is close to that for carbon black and slightly lower than that for textured pyrolytic graphite [32].

10.4.2   Reconditioning the Spent Electron Beam

When a linear-beam tube is operated at, or close to, saturation the spent electron beam is strongly bunched and the electrons can have appreciable radial velocities. It is more difficult to collect such a beam efficiently than an unmodulated beam. That is because electrons, which start from the same point with different radial velocities, follow different paths through the collector. It has been found that, if the spent beam does not emerge abruptly from the magnetic focusing field, but passes through a region where the beam is allowed to expand adiabatically to a larger diameter, then the space-charge and the radial velocities are reduced and the beam is easier to collect [17, 33–34]. The magnetic field in this region typically includes a field reversal. Beam conditioning, in this way, is most useful for tubes with high electronic efficiencies and/or high perveances, such as klystrons and coupled-cavity TWTs, because the disorder in the spent-beam is greatest. However, it has been shown that the use of a refocusing region is not necessary for most helix TWTs. In that case it is sufficient to allow the beam to drift through a short length of the PPM field after it has left the helix and before it enters the collector [35]. As the beam has not been allowed to expand, it is possible for the collector to be smaller than is the case when a refocusing section is employed.

10.5   Cooling

The greatest part of the heat generated in a tube is in the collector of a linear-beam tube or gyrotron, or in the anode of a magnetron or gridded tube. The body of the tube is heated by interception of electrons, and by RF losses in the interaction structure. Thus attention must be paid to adequate cooling of all parts of a tube. The different parts of the tube may be cooled by conduction, by natural or forced convection of air or liquid, or by boiling liquid. These methods are reviewed briefly below. All are ways of transferring heat from the very high power density at the tube surface to some surface where the surface area is larger, the power density is lower, and the heat can be dissipated more readily. It is necessary to take care that the efficiency of the cooling system is not reduced by the accumulation of dust, by contamination of coolants, or by corrosion of the surfaces. It is also necessary to consider the ambient temperature conditions in which the equipment will be stored and used. If the ambient temperatures fall below freezing then it is necessary to ensure that liquid coolants do not freeze, and have low enough viscosities for the cooling system to work properly when it is started up. If the ambient temperatures are high then the efficiency of air cooling systems is reduced and there is a risk that liquid coolants may boil, or that parts of the collector may overheat. General information about cooling systems is to be found in [5–6, 36].

10.5.1   Conduction Cooling

Low power tubes, such as helix TWTs, which generate not more than a few hundred watts of heat can be cooled by conduction from the tube to a suitable heat sink. In some mobile applications the heat sink can be the external surface of the vehicle (e.g. an aircraft or a ship). In spacecraft the heat must be radiated into space and heat pipes may be used to convey the heat from the tube to the radiating surface [37]. Alternatively, radiating fins can be mounted directly on the outside of the collector. The tube is then mounted so that these fins are on the outside of the spacecraft and can radiate directly into space.

10.5.2   Air Cooling

Natural, or forced, convection of air can be used to cool regions of a tube where the power density is not too high. The usual method is to use an array of cooling fins through which air can flow, or can be forced by a fan. This has the advantage that it is not necessary to provide liquid coolants. It is important that the air intake is fitted with an effective filter to prevent the accumulation of dust on the cooling surfaces.

10.5.3   Liquid Cooling

At higher power levels it is common to pump a coolant through suitably placed narrow channels, as shown in Figure 10.1. The coolant may be de-ionised water, but other liquids (e.g. fluorocarbons) have been used. The channels must be small enough to ensure turbulent flow. There is a trade-off between the power required to pump the liquid around the system and the thermal power which can be removed. It is necessary to avoid cavitation and local boiling since those can cause hot spots. If, for example, the heat is conducted though a copper wall 1 cm thick, cooled by water at a maximum temperature of 100 °C, then the maximum power density is 700 W cm^{-2}. This figure can be increased to 2 kW cm^{-2} by using very small pipes close to the surface to be cooled [4]. In multi-phase cooling water is pumped through narrow tubes and permitted to boil locally. Power densities up to 2.5 kW cm^{-2} can be handled in this way [36]. The output water temperature is below 100 °C and the steam generated condenses back into the coolant.

10.5.4   Vapour Phase Cooling

In vapour phase cooling the part of the tube to be cooled is immersed in a liquid coolant (usually de-ionised water) which is allowed to boil. The vapour is circulated by natural convection to a heat exchanger where it is condensed. Figure 10.15 shows one possible arrangement. The effective surface area for heat transfer can be increased by arranging a number of large bore cooling channels around the collector. The method has the advantage that the temperature of the tube is kept constant. The power density is up to 150 W cm^{-2} [36, 38].

In a development of this technique, known as Hypervapotron® cooling, the surface of the part to be cooled is provided with a large number of narrow bore blind channels into which the liquid flows. At high power densities the heat flux causes explosive local boiling which expels the heated liquid from the channels. Cooler liquid is sucked back in and the process is repeated [4, 38]. Power densities of up to 2 kW cm^{-2} are possible.
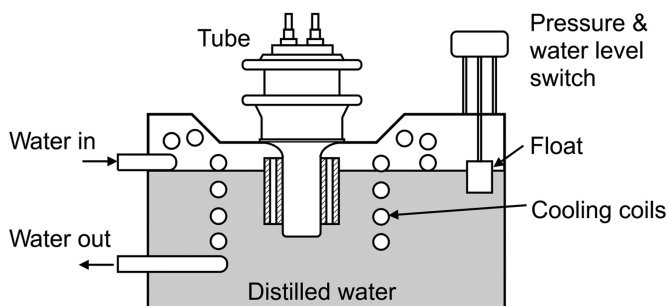
Figure 10.15: Vapour phase cooling of a power gridded tube (copyright 1976, EEV, reproduced, with permission of Teledyne e2v, from [39]).

References

[1] C. Shibata *et al.*, 'High-power (500 kW) c.w. magnetron for industrial heating', *Electrical Engineering in Japan,* vol. 111, pp. 94–100, 1991.

[2] W. P. Bennett, 'Large-power-tube design, processing and testing', in R. S. Burnap, ed., *Electron Tube Design*. Harrison, NJ: RCA, pp. 763–791, 1962.

[3] A. S. Gilmour, Jr., *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-Field Amplifiers and Gyrotrons*. Norwood, MA: Artech House, 2011.

[4] G. Faillon, 'Technical and industrial overview of RF and microwave tubes for fusion', *Fusion Engineering and Design,* vol. 46, pp. 371–381, 1999.

[5] L. Sivan, *Microwave Tube Transmitters*. London: Chapman and Hall, 1994.

[6] Varian, *Technical Manual: Installation, Operation, Maintenance, Care and Handling Instructions, General: Microwave Tubes, Magnetron Tubes, Electron Tubes*, Publication no. T.0.00-25-251, 1 October 1979.

[7] F. Sterzer, 'Improvement of traveling-wave tube efficiency through collector potential depression', *IRE Transactions on Electron Devices,* vol. 5, pp. 300–305, 1958.

[8] H. J. Wolkstein, 'Effect of collector potential on the efficiency of traveling-wave tubes', *RCA Review,* vol. 19, pp. 259–282, 1958.

[9] T. G. Mihran and W. Neugebauer, 'Analytical study of a depressed collector for linear beam microwave amplifiers', NASA Lewis Research Center, Cleveland, OH, 1970.

[10] H. G. Kosmahl, 'How to quickly predict the overall TWT and the multistage depressed collector efficiency', *IEEE Transactions on Electron Devices,* vol. 27, pp. 526–529, 1980.

[11] D. S. Komm *et al.*, 'Advances in space TWT efficiencies', *IEEE Transactions on Electron Devices,* vol. 48, pp. 174–176, 2001.

[12] T. Okoshi *et al.*, 'The tilted electric field soft-landing collector and its application to a traveling-wave tube', *IEEE Transactions on Electron Devices,* vol. 19, pp. 104–110, 1972.

[13] W. Neugebauer and T. G. Mihran, 'A ten-stage electrostatic depressed collector for improving klystron efficiency', *IEEE Transactions on Electron Devices,* vol. 19, pp. 111–121, 1972.

[14] H. G. Kosmahl and P. Ramins, 'Small-size 81- to 83.5-percent efficient 2- and 4-stage depressed collectors for octave-bandwidth high-performance TWT's', *IEEE Transactions on Electron Devices,* vol. 24, pp. 36–44, 1977.

[15] P. Ramins and T. A. Fox, '90- to 93-percent efficient collector for operation of a dual-mode traveling-wave tube in the linear region', *IEEE Transactions on Electron Devices,* vol. 26, pp. 1662–1664, 1979.

[16] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices.* Berlin: Springer-Verlag, pp. 1–84, 2008.

[17] H. G. Kosmahl, 'Modern multistage depressed collectors – a review', *Proceedings of the IEEE,* vol. 70, pp. 1325–1334, 1982.

[18] H. G. Kosmahl, 'A novel, axisymmetric, electrostatic collector for linear beam microwave tubes', NASA Lewis Research Center, Cleveland, OH, 1971.

[19] J. A. Dayton, Jr. *et al.*, 'Experimental verification of a computational procedure for the design of TWT-refocuser-MDC systems', *IEEE Transactions on Electron Devices,* vol. 28, pp. 1480–1489, 1981.

[20] J. R. Hechtel, 'A novel electrostatic-focusing depressed collector for linear beam tubes', *IEEE Transactions on Electron Devices,* vol. 24, pp. 45–52, 1977.

[21] L. Kumar *et al.*, '3-dimensional simulation of multistage depressed collectors on microcomputers', *IEEE Transactions on Electron Devices,* vol. 42, pp. 1663–1673, September 1995.

[22] S. Coco *et al.*, 'Shape optimization of multistage depressed collectors by parallel evolutionary algorithm', *IEEE Transactions on Magnetics,* vol. 48, pp. 435–438, 2012.

[23] J. Petillo *et al.*, 'The MICHELLE three-dimensional electron gun and collector modeling tool: theory and design', *IEEE Transactions on Plasma Science,* vol. 30, pp. 1238–1264, 2002.

[24] J. J. Petillo *et al.*, 'Recent developments to the MICHELLE 2-D/3-D electron gun and collector modeling code', *IEEE Transactions on Electron Devices,* vol. 52, pp. 742–748, 2005.

[25] T. K. Ghosh and R. G. Carter, 'Optimization of multistage depressed collectors', *IEEE Transactions on Electron Devices,* vol. 54, pp. 2031–2039, 2007.

[26] Z. Liang *et al.*, 'Numerical optimization of a multistage depressed collector with secondary electron emission for an X-band gyro-BWO', *IEEE Transactions on Plasma Science,* vol. 37, pp. 2328–2334, 2009.

[27] T. K. Ghosh, 'Three-dimensional modelling and optimization of multistage depressed collectors', PhD, Engineering Department, Lancaster University, Lancaster, UK, 2002.

[28] A. N. Curren and T. A. Fox, 'Traveling-wave tube efficiency improvement with textured pyrolytic graphite multistage depressed collector electrodes', *IEEE Electron Device Letters,* vol. 2, pp. 252–254, 1981.

[29] P. Ramins *et al.*, 'Verification of an improved computational design procedure for TWT-dynamic refocuser-MDC systems with secondary electron emission losses', *IEEE Transactions on Electron Devices,* vol. 33, pp. 85–90, 1986.

[30] A. N. Curren, 'Carbon and carbon-coated electrodes for multistage depressed collectors for electron-beam devices: a technology review', *IEEE Transactions on Electron Devices,* vol. 33, pp. 1902–1914, 1986.

[31] B. T. Ebihara *et al.*, 'TWT efficiency improvement by a low-cost technique for deposition of carbon on MDC electrodes', *IEEE Transactions on Electron Devices,* vol. 34, pp. 490–493, 1987.

[32] P. Ramins and B. T. Ebihara, 'Isotropic graphite multistage depressed collectors – a progress report', *IEEE Transactions on Electron Devices,* vol. ED-36, pp. 817–824, 1989.

[33] N. Stankiewicz, 'Analysis of spent beam refocusing to achieve optimum collector efficiency', *IEEE Transactions on Electron Devices,* vol. 24, pp. 32–36, 1977.

[34] J. A. Dayton, Jr. *et al.*, 'Analytical prediction and experimental verification of TWT and depressed collector performance using multidimensional computer programs', *IEEE Transactions on Electron Devices,* vol. 26, pp. 1589–1598, 1979.

[35] P. Ramins *et al.*, 'A reexamination of spent beam refocusing for high-efficiency helix TWTs and small MDCs', *IEEE Transactions on Electron Devices,* vol. 35, pp. 539–548, 1988.

[36] EIMAC, *Care and Feeding of Power Grid Tubes*, 5th ed. San Carlos, CA: CPI Inc. Eimac Division, 2003.

[37] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, UK: Research Studies Press, 1995.

[38] P. Gerlach and C. Kalfon, 'New technologies in power-grid tubes and their impact in high-power UHF-TV operation', *Proceedings of the IEEE,* vol. 70, pp. 1335–1345, 1982.

[39] EEV, '*Preamble – Tetrodes*', ed. Chelmsford, UK: Teledyne e2v, 1976.

11   Beam-Wave Interaction

11.1   Introduction

Microwave tubes depend for their operation on the interaction between an electron beam and an electromagnetic structure as shown in Figure 11.1. In the majority of tubes the geometry is rotationally symmetric. The current in the beam is modulated by the RF electric field of the structure. The RF current in the beam then induces currents in the structure which result in a transfer of energy from the beam to the structure. The electromagnetic structure of a klystron is made up of cavity resonators (see Chapter 3), whilst in a travelling-wave tube (TWT) it is a slow-wave structure (see Chapter 4). This chapter develops the theory of these two types of interaction as a foundation for the discussion of klystrons (in Chapter 13), and TWTs (in Chapter 14). The theory of crossed-field tubes is developed in Chapters 15 and 16, and the theory of gyrotrons is covered in Chapter 17.

The equations describing beam-wave interactions are non-linear, and it is necessary to use numerical methods to find solutions to them. However, useful insight can be gained by assuming that the modulation of the beam is small, compared with its DC properties. In small-signal theory the equations are replaced by linearised approximations which can be solved more easily. In a further approximation ballistic theory ignores the space-charge forces that the electrons exert on each other. We shall begin by considering the ballistic theory of the interaction between an electron beam and the RF fields of a cavity resonator.

The effects of space-charge can be introduced by treating the electron beam as a compressible fluid. The space-charge waves which propagate on a linear electron beam were first studied by Hahn [1] and Ramo [2]. A considerable body of literature exists on this subject and the reader is referred to [3] and [4] for convenient summaries of much of it. This chapter deals only with space-charge wave theory as far as it is needed to understand klystrons and TWTs. Many other possible types of space-charge wave devices have been studied in the past including resistive and inductive wall amplifiers, velocity jump amplifiers, double beam amplifiers, and beam-plasma amplifiers, but these are not of practical importance (see [3, 4] and [5] for further information). It is possible to develop a space-charge wave theory of crossed-field devices [3] but this is of limited value because, as will be seen in Chapters 15 and 16, the operation of these tubes is always non-linear.
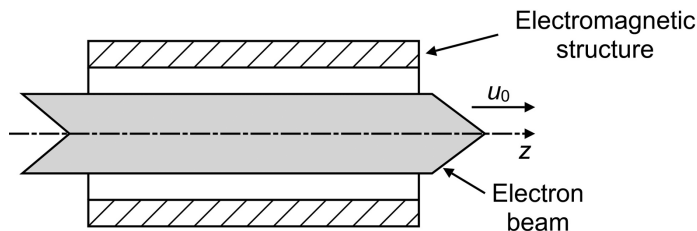
Figure 11.1: Arrangement of a linear-beam tube.

Under large-signal conditions electron overtaking can occur so that the electron velocities are no longer single-valued functions of position. Under these conditions space-charge forces cannot be neglected and the interaction must be modelled numerically by tracking the motion of sample electrons under the influence of both circuit and space-charge fields. The large-signal theory of beam-wave interactions is discussed in Section 11.8.

The analysis of problems in beam-wave interaction presents a number of pitfalls which are a result of the negative charge on the electron. Referring to Figure 11.1 we see that, for an electron beam, the convection current in the beam is in the negative z direction. That is, the conventional current in a linear-beam tube flows from the collector to the gun. The reversal of the direction of the current presents conceptual problems. The author has found that it is much easier to develop the analysis assuming that the electrons are positively charged, and this approach is followed in this book. The results for a beam of negatively charged electrons are easily recovered, if necessary, by reversing the signs of all the currents and voltages. In many cases this makes no difference, as we shall see. For convenience, the charge carriers will be referred to as 'electrons' even though their charge has been assumed to be positive.

11.2   Ballistic Theory of Interaction with a Gap

In many microwave tubes the electron beam interacts with the electric field of one or more gaps in the surrounding drift tube. These are connected to external cavity resonators or other electromagnetic structures. In the region of the gap the RF electric field is usually much greater than the space-charge field. Useful information about the interaction between the electrons and the gap can therefore be obtained by ballistic analysis which neglects the space-charge forces.

11.2.1   Beam Modulation by a Gridded Gap

Figure 11.2 shows an unmodulated electron beam passing through a gridded inter-action gap having an RF voltage across it. Gridded gaps are not used in tubes because the grids would rapidly be destroyed by the power of the electrons intercepted by
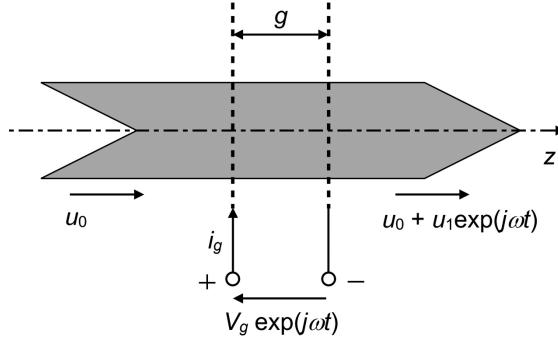
Figure 11.2: Modulation of an electron beam by the RF electric field in the gap between a pair of grids (electrons positive).

them. Nevertheless it is useful to begin by studying this case because the analysis is straightforward and provides insight into the interaction. The sign convention used is that the gap voltage is positive when charges are accelerated. The direction of the gap current i_g is such that power is flowing from the gap to the beam.

The electric field of the gap is given by

$$E_z = -\frac{V_g}{g}\exp(j\omega t), \tag{11.1}$$

where V_g is the amplitude of the RF voltage across the gap. The non-relativistic equation of motion of an electron is

$$\ddot{z} = \frac{eV_g}{m_0 g}\exp(j\omega t). \tag{11.2}$$

Multiplying (11.2) by $\dot{z}$, and integrating with respect to time, gives the change in kinetic energy of an electron (in electron volts) produced by the gap

$$V_1 = \frac{m_0}{2e}\dot{z}^2 = \frac{V_g}{g}\int_{-g/2}^{g/2}\exp(j\omega t)\,dz. \tag{11.3}$$

In this equation the limits of integration are the times at which an electron enters and leaves the gap. Now, if the RF component of the beam velocity is small compared with the DC component (u_0), and the electron crosses the centre of the gap when $t = t_0$, then the time at which it reaches position z is given by

$$t = \frac{z}{u_0} + t_0. \tag{11.4}$$

Substituting for t from (11.4) into (11.3) gives

$$V_1 = \frac{V_g\exp(j\omega t_0)}{g}\int_{-g/2}^{g/2}\exp(j\beta_e z)\,dz, \tag{11.5}$$
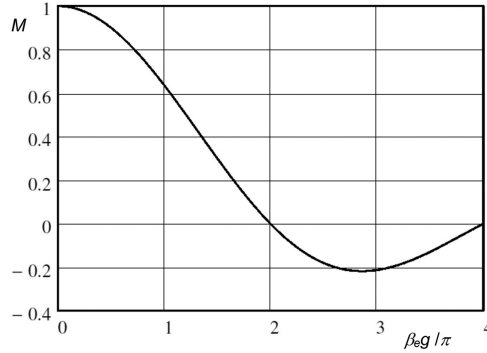
Figure 11.3: Dependence of the coupling factor for a gridded gap on the electron transit angle.

where the electronic propagation constant is defined by

$$\beta_e = \frac{\omega}{u_0}. \tag{11.6}$$

Evaluating the integral we obtain

$$V_1 = \frac{\sin(\theta/2)}{\theta/2} V_g \exp(j\omega t_0) = M V_g \exp(j\omega t_0), \tag{11.7}$$

where $\theta = \beta_e g$ is the *gap transit angle* and M is the *gap coupling factor*. Figure 11.3 shows how the gap coupling factor varies with the transit angle (see Worksheet 11.1). We note that M is always less than or equal to unity because the field changes during the finite transit time of the electrons across the gap. Since M tends to unity as the transit angle tends to zero it follows that the action of a finite gridded gap is equivalent to that of a very narrow gridded gap with gap voltage MV_g. The beam is said to be *velocity modulated*.

11.2.2   Ballistic Electron Bunching

The kinetic energy of electrons leaving the gap can be written

$$\frac{1}{2} m_0 u^2 = e\left(V_0 + M V_g \exp(j\omega t_0)\right), \tag{11.8}$$

where V_0 is the potential of the cathode relative to the drift region, including the effects of space-charge and beam rotation. V_0 and V_g are both positive. As the electrons drift downstream from the gap the variation in velocity causes them to become bunched together so that the beam current is modulated. If the modulation is small we can neglect second-order quantities and the velocity of an electron leaving the gap is

$$u = u_0 \left(1 + \frac{M V_g}{2 V_0} \exp(j\omega t_0)\right). \tag{11.9}$$
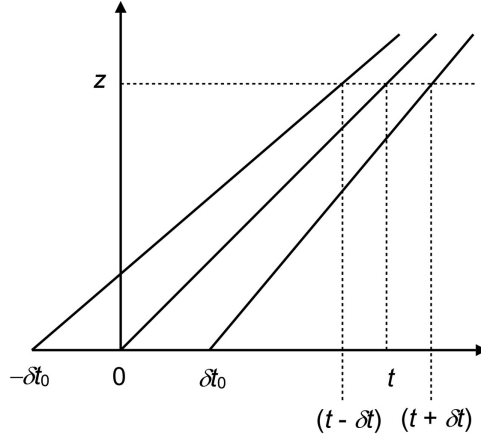
Figure 11.4: Ballistic bunching of a velocity modulated electron beam.

The electron reaches a plane z from the gap at time

$$t = t_0 + \frac{z}{u_0}\left(1 - \frac{MV_g}{2V_0}\exp(j\omega t_0)\right), \tag{11.10}$$

as shown in Figure 11.4. Under small-signal conditions the electron trajectories do not cross and therefore the current at z is related to the current at $z = 0$ by

$$(I_0 + I_1)dt = I_0\,dt_0. \tag{11.11}$$

Differentiating (11.10) with respect to t_0 we obtain

$$\frac{dt}{dt_0} = 1 - j\beta_e z\left(\frac{MV_g}{2V_0}\right)\exp(j\omega t_0). \tag{11.12}$$

where β_e is defined in (11.6). Substituting for dt_0/dt in (11.11) gives the RF convection current at the plane z

$$I_1 = jI_0\left(\frac{MV_g\beta_e z}{2V_0}\right)\exp j(\omega t - \beta_e z), \tag{11.13}$$

since $\omega t_0 = \omega t - \beta_e z$. Thus, in a frame of reference moving with uniform velocity u_0, the phase of the charge bunches leads the phase of the gap voltage by 90°. This conclusion still holds when the charges are negative. However, because I_0 is negative the conventional RF current I_1 is maximum in the negative direction at the centre of the charge bunches.

It can be shown that, for longer drift lengths, the amplitudes of the harmonic currents in a bunched electron beam are given by [6–8]

$$|I_n| = 2I_0 J_n\left(\frac{nMV_g\beta_e z}{2V_0}\right) \quad n = 1, 2, \cdots \tag{11.14}$$

where J_n is the Bessel function of the first kind and order n. This equation, which ignores space-charge forces, is valid even when electron overtaking occurs. However, except in the special case of frequency-multiplier tubes, it is desirable to avoid electron overtaking as far as possible. Also, in practical tubes the drift length between interaction gaps is normally much less than the distance to the first maximum of (11.14) and space-charge forces are not negligible. The bunching of a velocity-modulated beam, including the effects of space-charge forces, is discussed in Section 11.4.

11.2.3   Beam Loading of a Gridded Gap

When (11.8) is integrated over one RF cycle it is found that the mean energy of the modulated beam is equal to the energy of the unmodulated beam. Thus first-order analysis suggests that no power is required to modulate the beam. This is a consequence of the assumption that the electrons cross the gap without any change in velocity. In practice those electrons which are accelerated cross the gap in a shorter time, and the effective value of M is increased. Thus these electrons gain more energy than is shown in the first-order analysis. Similarly, those electrons which are retarded lose less energy than shown in the first-order analysis. Hence the power carried by the beam is increased by velocity modulation and this can be verified by numerical integration of the equations of motion (see Worksheet 11.3).

To calculate the power delivered to the beam by the gap we consider the gap to be made up of a series of elementary gaps whose position is z and whose length is dz. The transit angle of each elementary gap is small so that $M = 1$ and the gap voltage is $V_g dz/g$. Then from (11.13) the current at a remote plane z_1 due to the modulation by an elementary gap is

$$dI_1(z_1) = \frac{1}{2} G_0 \frac{V_g}{g} j\beta_e(z_1 - z)\exp\left[-j\beta_e(z_1 - z)\right] dz, \tag{11.15}$$

where $G_0 = I_0/V_0$ and the phase is referred to the phase of the gap voltage at z. Note that, for negative charges the signs of both I_0 and V_0 are reversed so that G_0 is unchanged. Thus the RF current at z_1 is

$$I_1(z_1) = \frac{1}{2} G_0 \left(\frac{V_g}{g}\right)\int_{-g/2}^{z_1} j\beta_e(z_1 - z)\exp\left[-j\beta_e(z_1 - z)\right] dz. \tag{11.16}$$

Making the substitution $\alpha = \beta_e(z_1 - z)$ (11.16) becomes

$$I_1(z_1) = \frac{1}{2} G_0 \left(\frac{V_g}{\beta_e g}\right)\int_0^{\beta_e(z_1 + g/2)} j\alpha\exp(-j\alpha)\, d\alpha. \tag{11.17}$$

Evaluating the integral gives

$$I_1(z_1) = \frac{1}{2} G_0 \left(\frac{V_g}{\beta_e g}\right)\left[-\beta_e(z_1 + g/2)e^{-j\beta_e(z_1 + g/2)} + j e^{-j\beta_e(z_1 + g/2)} - j\right]. \tag{11.18}$$

The power flow from the gap to the beam is given by [9]

$$P = \frac{1}{2} \int_{-g/2}^{g/2} E_z\left(z_1\right) I_1\left(z_1\right)^* \, dz_1 = \frac{V_g}{2g} \int_{-g/2}^{g/2} I_1\left(z_1\right)^* \, dz_1,$$

(11.19)

where the asterisk denotes the complex conjugate. Substituting for I_1 from (11.18) and performing the integration we find that

$$P = \frac{1}{2} G_0 \left(\frac{V_g}{\theta}\right)^2 \left[\left(1 - \cos\theta - \frac{\theta}{2}\sin\theta\right) + j\left(\frac{\theta}{2}\cos\theta + \frac{\theta}{2} - \sin\theta\right)\right],$$

(11.20)

where $\theta = \beta_e g$. The real part of this expression is the active power required to modulate the beam

$$P_a = \frac{1}{2} G_0 \frac{\sin\theta/2}{\theta/2}\left(\frac{\sin\theta/2}{\theta/2} - \cos\theta/2\right)\frac{V_g^2}{2} = G_b \frac{V_g^2}{2},$$

(11.21)

where G_b is the *beam loading conductance*. It is straightforward to show that

$$\frac{G_b}{G_0} = -\frac{1}{4}\theta\frac{d}{d\theta}\left(M\left(\theta\right)\right)^2 = -\frac{1}{2}M\theta\frac{dM}{d\theta}.$$

(11.22)

When relativistic effects are important (11.22) becomes [10]

$$\frac{G_b}{G_0} = -\frac{1}{\gamma_R\left(1 + \gamma_R\right)}M\theta\frac{dM}{d\theta},$$

(11.23)

where

$$\gamma_R = \frac{1}{\sqrt{1 - u_0^2/c^2}}.$$

(11.24)

Note that the symbol γ_R is used in this chapter for the relativistic factor to avoid confusion with the use of γ for radial propagation constants. The reactive power transferred to the beam is given by the imaginary part of (11.20)

$$P_r = -\frac{1}{2} G_0 \frac{\cos\left(\theta/2\right)}{\theta/2}\left(\frac{\sin\left(\theta/2\right)}{\theta/2} - \cos\left(\theta/2\right)\right)\frac{V_g^2}{2} = -B_b\frac{V_g^2}{2},$$

(11.25)

where B_b is the *beam loading susceptance*. The negative sign in (11.25) arises because the reactive power flow into a susceptance B is given by

$$P_r = \frac{1}{2}VI^* = \frac{1}{2}V\left(jBV\right)^* = -jB\frac{V^2}{2}.$$

(11.26)

Thus the presence of the beam can be represented in an equivalent circuit by the admittance $Y_b = G_b + jB_b$. Figure 11.5 shows how the beam loading conductance, and susceptance, of a gridded gap vary with the transit angle. Because G_b/G_0 is
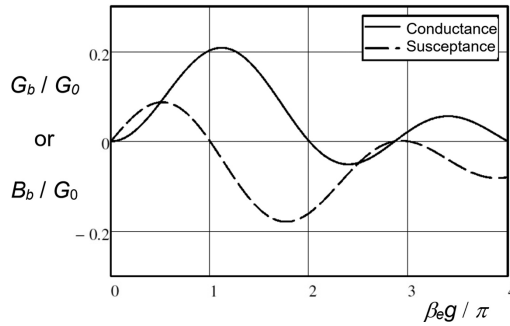
Figure 11.5: Beam loading conductance, and susceptance, of a gridded gap.
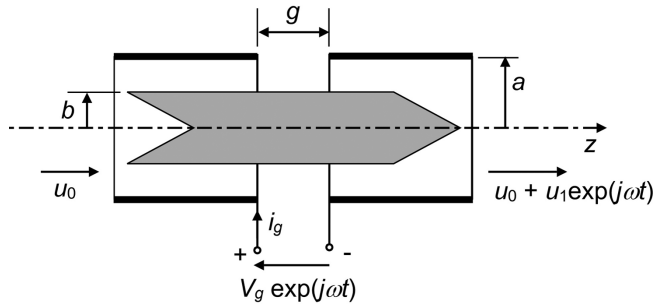
Figure 11.6: Arrangement of a gridless interaction gap (electrons positive).

normally less than 0.2 the power required to modulate the beam is only a small fraction of the DC beam power. The beam loading conductance may be negative when the transit angle is greater than 2π. If the sum of the beam loading conductance and the conductance of the external circuit connected to the gap is negative then the system is unstable and can break into self-oscillation (see Section 13.4.1).

11.2.4   Beam Modulation by a Gridless Gap

Figure 11.6 shows the arrangement of a gridless gap of the type generally used in microwave tubes. The axial component of the electric field within the gap can be found using the method described in Section 3.5. It is assumed that E_z is known as a function of z when $r = a$ as discussed in Section 3.5.3. The Fourier transform of this function is

$$\Gamma(\beta) = \int_{-\infty}^{\infty} E_z(a, z) \exp(j\beta z)\, dz. \tag{11.27}$$

Since the fields must satisfy the wave equation in free space we know that E_z varies with radius as $I_0(\gamma r)$ where the radial propagation constant $\gamma = \sqrt{\beta^2 - \omega^2/c^2}$. Thus the electric field within the gap is given by the inverse Fourier transform

$$E_z(r,z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{I_0(\gamma r)}{I_0(\gamma a)} \Gamma(\beta) \exp(-j\beta z)\, d\beta. \qquad (11.28)$$

We saw, in Section 3.5.3, that the electric field distribution in the region occupied by the electron beam depends only weakly on the profile of the field at $r = a$. We shall therefore assume, for the moment, that $E_z = -V_g/g$ when $r = a$ and $|z| \le g/2$, and zero elsewhere, so that

$$\Gamma(\beta) = -V_g \frac{\sin(\beta g/2)}{\beta g/2}. \qquad (11.29)$$

It is convenient to remove the dependence on r by taking the average of the first term in the integral in (11.28) across the beam. We shall also assume that the electron beam is in confined flow and that the RF current is nearly constant across the beam. The average electric field acting on the electrons is then

$$E_z(z) = -\frac{V_g}{2\pi} \int_{-\infty}^{\infty} \frac{2}{\beta b} \cdot \frac{I_1(\gamma b)}{I_0(\gamma a)} \cdot \frac{\sin(\beta g/2)}{\beta g/2} \exp(-j\beta z)\, d\beta \qquad (11.30)$$

and the magnitude of the effective gap voltage is

$$V_1 = -\int_{-\infty}^{\infty} E_z(z) \exp(j\beta_e z)\, dz. \qquad (11.31)$$

Substituting for the field from (11.30) and exchanging the order of the integration gives

$$V_1 = \frac{V_g}{2\pi} \int_{-\infty}^{\infty} \frac{2}{\beta b} \cdot \frac{I_1(\gamma b)}{I_0(\gamma a)} \cdot \frac{\sin(\beta g/2)}{\beta g/2} \left\{ \int_{-\infty}^{\infty} \exp\left(-j(\beta_e - \beta)z\right) dz \right\} d\beta. \qquad (11.32)$$

Following Branch [9] we note that

$$\int_{-\infty}^{\infty} \exp\left(-j(\beta_e - \beta)z\right) dz = 2\pi\, \delta(\beta_e - \beta), \qquad (11.33)$$

where $\delta(z)$ is the Dirac delta function which is defined as a pulse having zero width, infinite height and unit area. Thus the integral with respect to z is zero except when $\beta = \beta_e$ and then

$$V_1 = \frac{2}{\beta_e b} \cdot \frac{I_1(\gamma_e b)}{I_0(\gamma_e a)} \cdot \frac{\sin(\beta_e g/2)}{\beta_e g/2} \cdot V_g, \qquad (11.34)$$

where β_e has been used in place of β in the calculation of γ. Then (11.34) can be written as

$$V_1 = MV_g = \mu_c\mu_d V_g,\tag{11.35}$$

where the *axial coupling factor* given by

$$\mu_d = \frac{\sin(\beta_e g/2)}{\beta_e g/2}\tag{11.36}$$

has exactly the same form as the coupling factor for a gridded gap given by (11.7). If the electric field at $r = a$ is not uniform this expression must be replaced by the Fourier transform of the gap field for unit gap voltage. When the ends of the drift tubes have knife edges so that the field in the gap is given by (3.89) then

$$\mu_d = J_0(\beta_e g/2),\tag{11.37}$$

where J_0 is the Bessel function of the first kind [11].

The average of the radial variation of the field (known as the *radial coupling factor*) is, for confined flow,

$$\mu_c = \frac{2}{\gamma_e b} \cdot \frac{I_1(\gamma_e b)}{I_0(\gamma_e a)}.\tag{11.38}$$

An alternative average is based upon an analysis of the exchange of power between the field and the beam [9]

$$\mu_b = \frac{\sqrt{I_0^2(\gamma_e b) - I_1^2(\gamma_e b)}}{I_0(\gamma_e a)}.\tag{11.39}$$

In practical cases, the difference between the numerical values given by (11.38) and (11.39) is negligible and either expression may be used. For a beam in Brillouin flow the RF current is a surface current at $r = b$ (see Figure 7.6) so that the radial coupling factor is

$$\mu_B = \frac{I_0(\gamma_e b)}{I_0(\gamma_e a)}.\tag{11.40}$$

Figure 11.7 shows curves of the radial coupling factor for Brillouin flow, and confined flow, plotted against $\gamma_e b$ with b/a as a parameter. It should be noted that (11.40) does not include the effect of the radial component of the electric field on the electron motion. However Brillouin flow is of limited practical importance because it is normally necessary to use a magnetic field greater than the Brillouin field to ensure sufficient beam stiffness. We shall therefore assume that (11.38) is a satisfactory approximation if $m > 1.5$. The effects of the radial component of the electric field may be important if the beam stiffness is small.
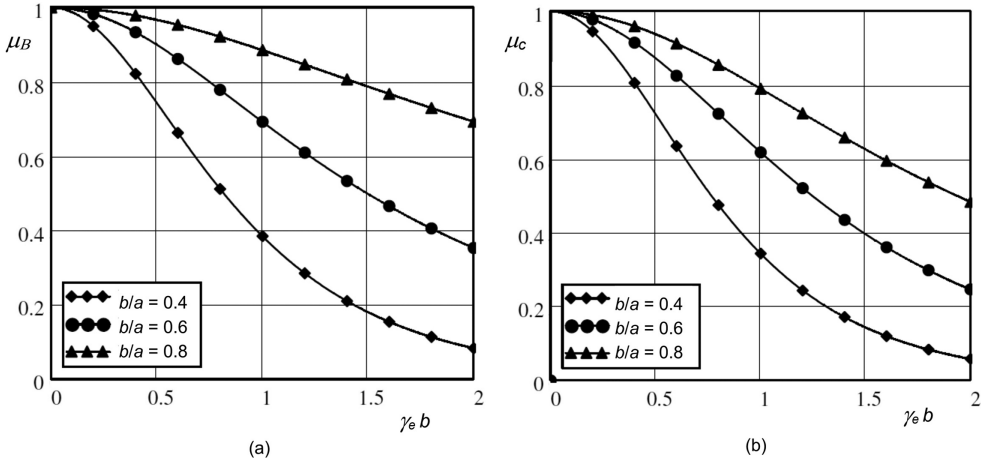
Figure 11.7: The radial gap coupling factor of an ungridded gap for: (a) Brillouin flow, and (b) confined flow.

11.3   Space Charge Waves on Linear Electron Beams

The RF properties of electron beams under small-signal conditions, including the effects of space charge, can be represented by space-charge waves. The non-relativistic equations which govern the motion of the electrons in a linear electron beam are the continuity equation

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}. \tag{11.41}$$

Newton's second law of motion

$$\frac{d\mathbf{u}}{dt} = \frac{e}{m_0}\left(\mathbf{E}_b + \mathbf{E}_c\right), \tag{11.42}$$

where $\mathbf{E}_b$ is the electric field arising from the space charge of the electrons and $\mathbf{E}_c$ is the field applied to the beam by an external microwave structure. The convection current in the electron beam is related to the charge density and the velocity by

$$\mathbf{J} = \rho\mathbf{u} \tag{11.43}$$

and, finally, the relationship between the charge density and the space-charge field is given by Poisson's equation

$$\nabla \cdot \mathbf{E}_b = \frac{\rho}{\varepsilon_0}. \tag{11.44}$$

These four equations involve four dependent variables so, in principle, they can be solved. The general solutions are rather complicated but the properties of space-charge waves can be understood from a simplified solution.

As a first stage of simplification let us assume that the electrons are moving parallel to the z axis and that the beam extends to infinity in the transverse direction. We will also assume that all quantities vary only with z and with time. The four equations (11.41) to (11.44) then become:

$$\frac{\partial J_z}{\partial z} = -\frac{\partial \rho}{\partial t}, \tag{11.45}$$

$$\frac{du_z}{dt} = \frac{\partial u_z}{\partial t} + u_z \frac{\partial u_z}{\partial z} = \frac{e}{m_0}\left(E_{zb} + E_{zc}\right), \tag{11.46}$$

$$J_z = \rho u_z, \tag{11.47}$$

and

$$\frac{\partial E_{zb}}{\partial z} = \frac{\rho}{\varepsilon_0}. \tag{11.48}$$

The left-hand side of (11.46) is expanded as the sum of: the rate of change of the velocity with time at constant position, and the rate of change with position as the motion of the electron is followed. In these equations the vector quantities J, u and E are assumed to possess only z components. The set of equations is non-linear because (11.46) and (11.47) contain terms which are the product of two variables.

The next stage is to linearise the equations by assuming the propagation of waves so that all the dependent variables can be written in the form

$$a = a_0 + a_1 \exp j\left(\omega t - \beta z\right), \tag{11.49}$$

where $a_1 \ll a_0$ and ω and β are the same for all variables. It is then possible to separate each equation into constant and time-varying parts. The constant part of the solution does affect the velocities of the electrons and it has been common to assume that it is cancelled out by the static field of a background of positive ions. Although positive ions do sometimes accumulate in the region of an electron beam, their effects are undesirable, and such accumulations are to be avoided. In the theory of space-charge waves it is better to take account of DC space-charge potential depression by including its effects as a correction to the DC electron velocity, as shown in Section 7.2. The small-signal equations for the time-varying quantities are

$$\beta J_1 = \omega \rho_1 \tag{11.50}$$

$$j\left(\omega - \beta u_0\right) u_1 = \frac{e}{m_0}\left(E_{b1} + E_{c1}\right) \tag{11.51}$$

$$J_1 = \rho_0 u_1 + \rho_1 u_0 \tag{11.52}$$

$$-j\beta E_{b1} = \frac{\rho_1}{\varepsilon_0}. \tag{11.53}$$

For the moment we will assume that the RF field of the external circuit (E_{c1}) is zero. Then E_{b1} can be eliminated between (11.51) and (11.53) to give

$$(\omega - \beta u_0)u_1 = \frac{e}{\varepsilon_0 m_0} \cdot \frac{\rho_1}{\beta}. \tag{11.54}$$

Similarly J_1 can be eliminated between (11.50) and (11.52) to give

$$(\omega - \beta u_0)\rho_1 = \beta \rho_0 u_1. \tag{11.55}$$

Equations (11.54) and (11.55) both relate u_1 to ρ_1 and they can be satisfied simultaneously only if

$$(\omega - \beta u_0)^2 = \frac{e\rho_0}{\varepsilon_0 m}. \tag{11.56}$$

The terms in this equation have the dimensions of angular frequency squared and the expression on the right-hand side is recognised as the electron plasma frequency which was introduced in (7.47) and defined by

$$\omega_p^2 = \frac{e\rho_0}{\varepsilon_0 m_0}. \tag{11.57}$$

It should be noted that this remains a positive quantity if the charge is negative because ρ_0 is also negative. When the DC electron velocity is large enough for relativistic effects to be important the rest mass of the electron is replaced by the longitudinal mass $\gamma^3 m_0$ (see (1.7) and [3]). The physical significance of the plasma frequency can be understood by considering the case when the electrons have no DC component of velocity so that the solution of (11.56) is

$$\omega = \omega_p. \tag{11.58}$$

This is the natural frequency of oscillation of a stationary cloud of electrons which has been disturbed from equilibrium. The possible solutions to (11.56) can therefore be written

$$\beta_\pm = \frac{\omega \mp \omega_p}{u_0} = \beta_e \mp \beta_p. \tag{11.59}$$

It is convenient to display these solutions in the form of a dispersion diagram, as shown in Figure 11.8. The slope of each line is equal to the DC beam velocity u_0. The lines defined by (11.59) represent two possible wave solutions (β_+ and β_-) whose phase velocities (ω/β) are, respectively, greater than and less than u_0. The waves are compression waves of charge density known as fast (β_+), and slow (β_-), *space-charge waves*. The properties of these waves will be examined in more detail in Section 11.3.3.

Since it has been assumed that the beam has infinite diameter and the electron motion is confined to the axial direction it follows that the RF electric field only has a longitudinal component, as shown in Figure 11.9(a). This field exerts the
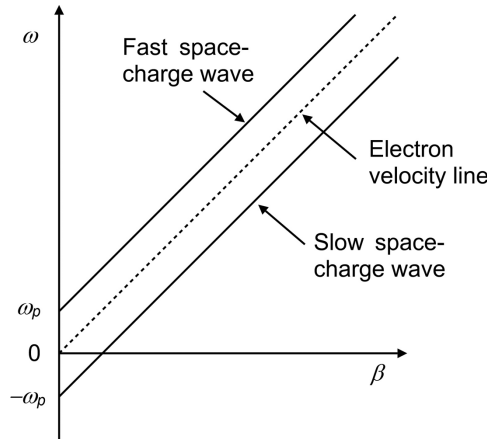
Figure 11.8: Dispersion diagram for space-charge waves neglecting the effects of radial boundaries.
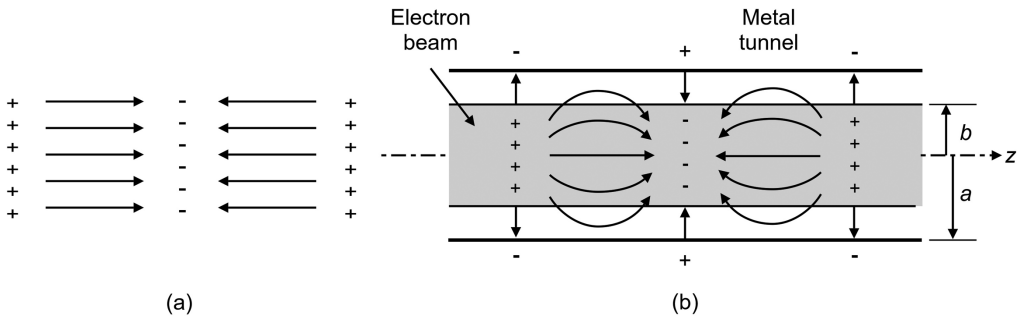
Figure 11.9: Electric field due to space charge: (a) without radial boundaries and (b) with radial boundaries.

forces on the electrons which cause the plasma oscillations. The displacement current density within the beam is found, using (11.53), to be

$$\frac{\partial D_z}{\partial t} = j\omega\varepsilon_0 E_{b1} = -\frac{\omega}{\beta}\rho_1. \tag{11.60}$$

Comparing (11.60) with (11.50) we see that the displacement current density is equal and opposite to the convection current density J_1 so that the total RF current density is zero. It follows that there is no RF magnetic field in an infinite beam. We shall see in the next section that this conclusion is modified when the beam is finite in the transverse direction.

11.3.1 Effect of Radial Boundaries

Any practical electron beam must have finite dimensions normal to the direction of electron motion. The RF electric field is then able to spread out into the charge-free

region outside the beam, as shown in Figure 11.9(b). The space-charge forces acting on the electrons are therefore less than those for the infinite beam illustrated in Figure 11.9(a). Thus the plasma frequency is lower, in finite beam, than it is in an infinite beam. This effect can be represented by multiplying the right-hand side of (11.53) by a factor p^2 where $0 \leq p^2 \leq 1$. The right-hand sides of equations (11.54) and (11.56) are, likewise, multiplied by p^2. Thus the conclusions of the analysis in the previous section are unchanged, provided that the plasma frequency is replaced by the effective plasma frequency (ω_q) defined by

$$\omega_q = p\omega_p, \tag{11.61}$$

where p is known as the *plasma frequency reduction factor*.

We saw in Section 7.4.3 that the diameter of a magnetically focused electron beam increases as the beam current is increased, except in the limiting case when the magnetic field tends to infinity. Thus, in order to determine the value of p in any particular case it is necessary to solve Maxwell's equations for a cylindrical electron beam within a concentric metal cylinder in the presence of a finite axial magnetic field. Under small-signal conditions the variations in the beam radius can be represented by charges on the surface of a beam whose radius is constant. This problem was studied by Brewer and Mackenzie [12, 13] who showed that the plasma frequency reduction factor is the solution of the equation

$$\left(1 - \frac{1}{p^2}\right) \frac{1}{\tau b} \frac{I_1(\tau b)}{I_0(\tau b)} = \frac{1}{\beta b} \left\{ \frac{I_1(\beta b) K_0(\beta a) + I_0(\beta a) K_1(\beta b)}{I_0(\beta b) K_0(\beta a) - I_0(\beta a) K_0(\beta b)} \right\} \tag{11.62}$$

where a and b are the tunnel radius and the beam radius (see Figure 11.9(b)), $I_0(z)$, $I_1(z)$, $K_0(z)$, and $K_1(z)$ are modified Bessel functions of the first and second kinds, and β is the axial propagation constant of the waves. It is assumed that the phase velocity of the waves is small compared with the velocity of light so that the axial and radial propagation constants are identical in the region outside the beam. Inside the beam the radial propagation constant τ is given by

$$\tau^2 = \beta^2 \left\{ \frac{\omega_p^2 / \omega_q^2 - 1}{\omega_p^2 / \omega_0^2 - 1} \right\}, \tag{11.63}$$

where ω_0 is defined by

$$\omega_0^2 = \omega_q^2 - 4\left(\frac{e\Phi_k}{2\pi m_0 b^2}\right)^2 \tag{11.64}$$

and Φ_k is the magnetic flux linked to the cathode. Making use of (7.52) and (7.58) we can rewrite (11.64) as

$$\omega_0^2 = \left[p^2 - 2(m^2 - 1)\right]\omega_p^2, \tag{11.65}$$

where m is the ratio of the focusing field to the Brillouin field which was defined in (7.56). Equation (11.65) shows that ω_0^2 is positive and equal to ω_q^2 when $m = 1$, that it

passes through zero when $m \sim 1.02$, and is then negative for all higher values of m. The denominator in (11.63) is positive when $m = 1$ and has an infinity and changes sign as ω_0 passes through zero. Thus τ is real for values of m close to unity, and imaginary for higher values of m. The Bessel functions on the left-hand side of (11.62) are computed with real or imaginary arguments as appropriate. Equation (11.62) can be solved numerically for any desired values of βb, a/b and m (see Worksheet 11.2). Since β is the propagation constant of the waves given by

$$\beta_{\pm} = \beta_e \mp \beta_q, \tag{11.66}$$

where $\beta_q = \omega_q / u_0$, the value of p should be calculated separately for the fast and slow space-charge waves. An initial estimate can be found by setting $\beta = \beta_e$ and refined by iteration if necessary. Curves of p versus βb for typical values of b/a and m are shown in Figure 11.10. When the effects of radial boundaries are taken into account the dispersion diagram for the fast and slow space-charge waves takes the form shown in Figure 11.11. Comparison with Figure 11.8 shows that negative values of ω do not normally occur and that the dispersion curves are asymptotic to those in Figure 11.8 for large values of β. For beam voltages greater than about 20 kV the assumption that the phase velocity of the waves is small compared with the velocity of light is not valid. It is reasonable to suppose that β should then be replaced by $\gamma = \beta\sqrt{1 - u_0^2/c^2}$ in (11.62) and (11.63), and in Figure 11.10. It was found that this correction was necessary to obtain good results for the klystron discussed in Section 13.4.3.

Brewer [12] showed how the division of the RF beam current between beam radius modulation (surface current) and charge density modulation (body current) depends on the magnetic field. Figure 2 of that paper shows these currents as being in anti-phase with each other. There is apparently a value of the magnetic field for which there is zero net convection current. This must be wrong, on physical grounds, and seems to be a consequence of using different sign conventions for the charge in e/m_0 and ρ. When this error is corrected it is found, using Brewer's normalisation, that the body current should be negative. The way in which these two currents vary with magnetic field is shown in Figure 11.12 when $\beta b = 0.5$ and $b/a = 0.5$. The curves do not vary very much if different parameter values are chosen. When this figure is compared with Figure 7.6 it can be seen that the two analyses give identical results in the limit of small RF currents.

11.3.2   Induced Current

When a space-charge wave is excited on an electron beam, which is surrounded by a continuous metal drift tube, the radial RF electric field must be terminated by induced charges on the wall of the drift tube (see Figure 11.9(b)). The movements of these charges constitute a longitudinal RF induced current in the wall. It is easy to see that this must be so by considering Maxwell's equation

$$\int \mathbf{H} \times \mathbf{ds} = \iint \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot \mathbf{dS}. \tag{11.67}$$
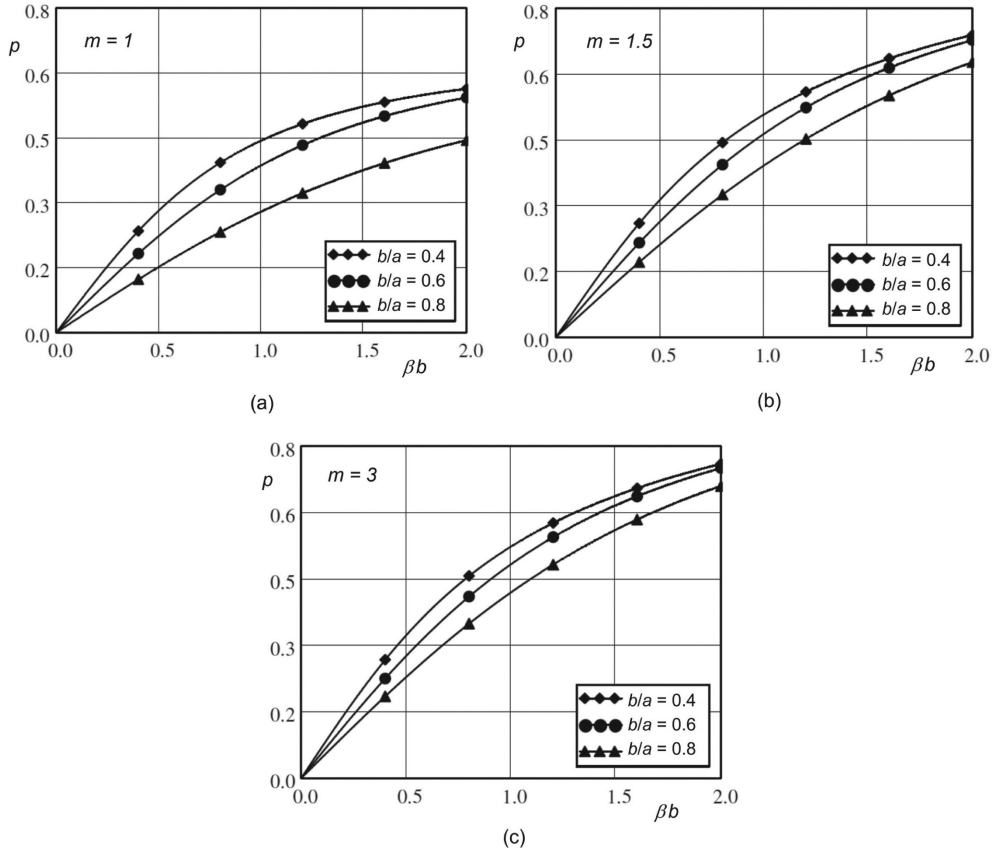
Figure 11.10: Plasma frequency reduction factor curves: (a) $m = 1$, (b) $m = 1.5$ and (c) $m = 3$ (Figure 11.4(a) copyright 1956, IEEE, reproduced, with permission, from [12]).
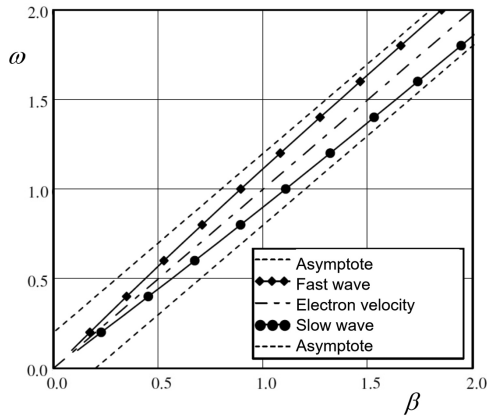
Figure 11.11: Dispersion diagram for space-charge waves including the effects of radial boundaries (nominal scales).
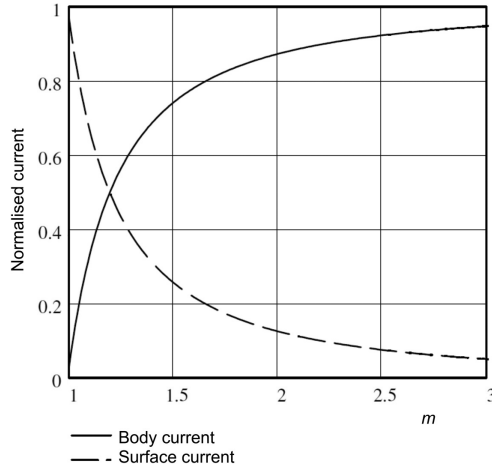
Figure 11.12: Dependence of the RF surface and body currents in an electron beam on the magnetic field parameter m.

If the system is cylindrically symmetrical and the integral of **H** is taken in the θ direction along a line just inside the drift tube wall, then the boundary conditions at the wall require the current in the wall to be equal and opposite to the integral of **H**. The two terms on the right-hand side of (11.67) are, respectively, the convection current and the axial component of the displacement current within the drift tube. Thus (11.67) requires that the induced current in the wall should be

$$i_W = -\left(I_C + I_D\right), \tag{11.68}$$

where I_C and I_D are the longitudinal convection and displacement currents within the drift tube. The convection current can be computed using the results of Brewer (with the sign of the body current reversed). The displacement current (using Brewer's normalisation) given by

$$\left(I_D\right)_{norm} = \frac{I_1\left(\tau b\right)}{\tau b} + \left[\frac{1 - \beta b \cdot I_1\left(\beta b\right) K_0\left(\beta a\right) + I_0\left(\beta a\right) K_1\left(\beta b\right)}{I_0\left(\beta b\right) K_0\left(\beta a\right) - I_0\left(\beta a\right) K_0\left(\beta b\right)}\right] \cdot \frac{I_0\left(\tau b\right)}{\left(\beta b\right)^2}. \tag{11.69}$$

is in antiphase with the convection current. The wall current can be computed using (11.69), or by integrating H_θ over the surface of the drift tube taking account of the sign of the convection current. The results obtained by these two methods agree with each other, confirming the correction to the sign of the body current. Figure 11.13 shows the ratio (q) of the magnitude of the wall current to the convection current as a function of βb, and of the ratio b/a, for $m = 1$, 1.5 and 3. The reduction in the normalised wall current is explained by the increase in the axial displacement current as the separation between the beam edge and the wall increases, and as the wavelength on the electron beam decreases. The changes resulting from changes in the magnetic focusing field, and, therefore, in the proportions of surface

and body current in the beam, are not so marked. The differences between the curves for $m = 1.5$ and $m = 3.0$ are almost indistinguishable and further increases in m make no difference. In a microwave linear-beam tube it is normally desirable that the coupling between the beam and the surrounding circuits should be as strong as possible. For this reason βb is normally not greater than 0.5, and b/a is not less than 0.5. It can be seen from Figure 11.13 that this choice is necessary to ensure that the induced wall current is high. It is also evident that, because the curves fall steeply with increasing βb, the coupling between the beam and the wall at harmonics of the fundamental frequency falls off sharply with increasing order of harmonic. The curves in Figure 11.13(c) are identical to those in Figure 11.7(b) except that the horizontal axis is βb rather than γb because it has been assumed that the electron velocity is small compared with the velocity of light. If this restriction is removed it is seen that the magnitude of the induced current in the wall is the product of the radial coupling factor and the convection current in the beam. Thus the use of γb in place of βb here, and in Figure 11.10, ensures the symmetry of the theory.
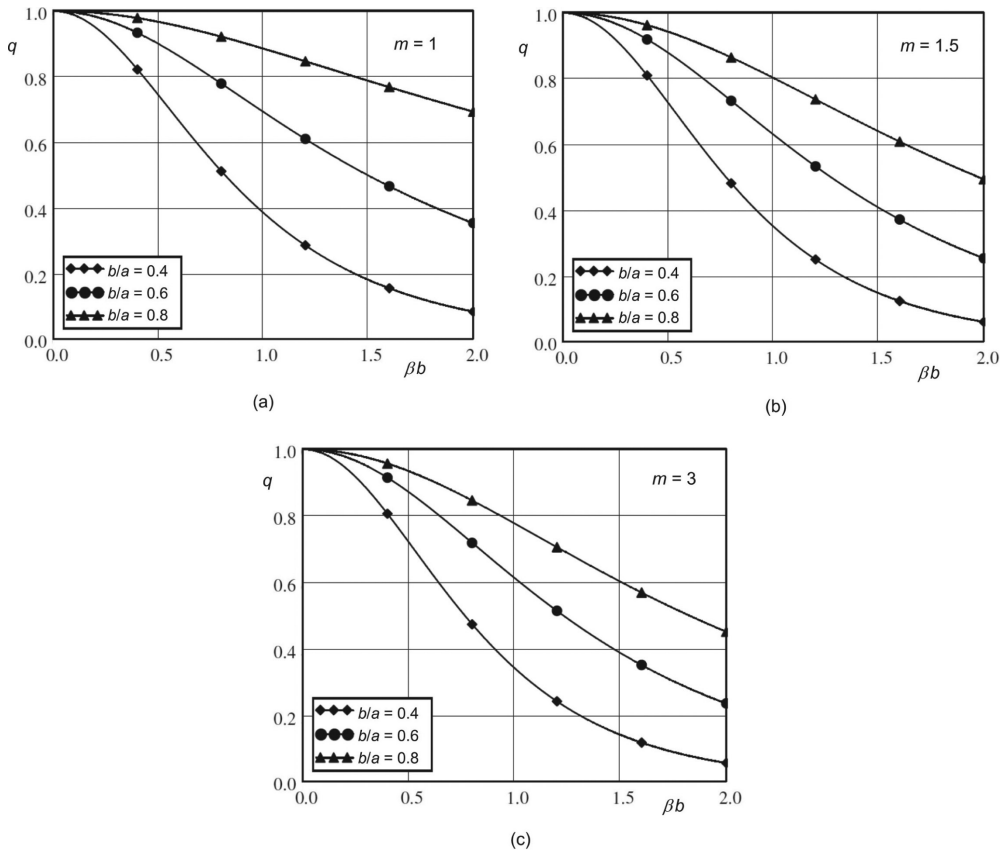
Figure 11.13: Ratio of the RF current in the wall of a drift tube to the RF convection current in a concentric electron beam: (a) $m = 1$, (b) $m = 1.5$ and (c) $m = 3$.

In linear-beam tubes of all kinds the drift tube enclosing the electron beam is not continuous, but is broken at intervals by interaction gaps. In the case of the klystron these gaps only occupy a small portion of the total length of the tube and it is satisfactory to ignore them when computing the reduced plasma frequency. In TWTs the gaps occupy a much greater proportion of the length of the tube so that the assumption of a continuous tunnel is no longer appropriate. A possible way of compensating for this is to use an effective tunnel radius [14]. It should be understood that this correction to ω_q is distinct from the interaction between the space-charge waves and the gapped tunnel which is dealt with in Section 11.6.

11.3.3   Transmission Line Representation of Space-Charge Waves

It is often useful to represent space-charge waves by equivalent transmission line modes. To do this we define an RF voltage, known as the *beam kinetic voltage* (V_1), by invoking the non-relativistic form of the principle of conservation of energy

$$\frac{1}{2} m_0 \left(u_0 + u_1 \exp\left(j\omega t\right)\right)^2 = e\left(V_0 + V_1 \exp\left(j\omega t\right)\right). \tag{11.70}$$

Note that V_0 is the potential of the cathode relative to the drift space. It corresponds to the DC beam velocity u_0 allowing for the effects of space-charge potential depression and rotation in the magnetic field. For small signals the left-hand side of (11.70) can be expanded to give

$$u_0^2 + 2u_0 u_1 \exp\left(j\omega t\right) = \frac{2e}{m_0}\left(V_0 + V_1 \exp\left(j\omega t\right)\right), \tag{11.71}$$

where the second-order term u_1^2 has been neglected. This equation is true at all times and so the RF and DC terms must balance separately to give

$$V_1 = \frac{u_0 u_1}{e/m_0} \tag{11.72}$$

and

$$V_0 = \frac{u_0^2}{2\left(e/m_0\right)}. \tag{11.73}$$

Dividing (11.72) by (11.73) we obtain

$$\frac{V_1}{V_0} = \frac{2u_1}{u_0}. \tag{11.74}$$

Note that this equation remains the same if the sign of the charge is reversed. The RF current density is obtained from (11.50) and (11.52) as

$$J_1 = \frac{\omega \rho_0}{(\omega - \beta u_0)} u_1 \tag{11.75}$$

so that, making use of (11.66) and (11.74), for the two space-charge waves

$$J_\pm = \pm \frac{\omega}{2\omega_q} \cdot \frac{J_0}{V_0} V_\pm. \tag{11.76}$$

If it is assumed that the DC and RF current densities are uniform then (11.76) can be integrated across the cross-section of the beam to give the RF convection currents of the space-charge waves

$$I_\pm = \pm \frac{\omega}{2\omega_q} \cdot \frac{I_0}{V_0} V_\pm. \tag{11.77}$$

These waves can be represented in terms of an equivalent transmission line. The characteristic admittance, known as the *electronic admittance*, is

$$Y_e = \frac{\omega}{2\omega_q} \cdot \frac{I_0}{V_0}, \tag{11.78}$$

which is positive regardless of the sign of the electronic charge. The kinetic power flows in the waves are then

$$P_\pm = \pm \frac{1}{2} |V_\pm|^2 Y_e, \tag{11.79}$$

where the plus and minus signs refer to the fast and slow space-charge waves respectively. Equation (11.79) reveals the surprising fact that the slow space-charge wave carries negative power. The direction of the power flow is certainly positive because, as Figure 11.11 shows, the slow-wave has a positive group velocity. The explanation of this paradox is to be found in equation (11.76) which shows that, in the slow space-charge wave, the electron density and the velocity are in anti-phase with one another. Thus when $u > u_0$, $\rho < \rho_0$, and when $u < u_0$, $\rho > \rho_0$. It follows that the average kinetic energy of a beam carrying a slow space-charge wave is less than the kinetic energy of the unmodulated beam. In order to set up a slow space-charge wave, power must be removed from the beam. This unexpected feature of the slow space-charge wave provides the key to understanding all kinds of microwave linear-beam tubes. When the beam velocity is relativistic the results are qualitatively the same but the electronic admittance is given by [3]

$$Y_e = \frac{\omega I_0}{\gamma_R (\gamma_R + 1) \omega_q V_0}, \tag{11.80}$$

where the relativistic factor γ_R is defined in (11.24) and ω_q is calculated using the relativistic value of u_0 but without other relativistic corrections.

The analysis above ignores the electromagnetic power associated with space-charge waves. It can be shown that the ratio of the electromagnetic power to the kinetic power in a space-charge wave is of the order of ω_q / ω [15]. For tubes having low space-charge this ratio is small and the effects of electromagnetic power can be ignored. In many high power tubes the ratio is around 10% and the electronic admittance should probably be corrected for this effect.

11.3.4   Space-Charge Waves on Non-Ideal Electron Beams

In general, the aim of the tube designer is to create an electron beam in which the current density and the axial velocity are uniform throughout and there is no scalloping. In that way the properties of the beam are close to those of the ideal beam assumed by theory. In practice all beams are non-ideal to some extent but, provided the variations from the ideal are not too great, it is possible to use the theory developed above.

We saw in Chapter 7 that the effects of space-charge potential depression and beam rotation should be taken into account when calculating the mean axial velocity of the electrons. These effects also mean that there is a small variation in the electron velocities across the beam. This variation is normally ignored, but it could contribute to the noise carried by the beam through the two-stream instability (see [4]) and to the Landau damping of space-charge waves [16–18]. These effects are not normally important in high power microwave tubes.

If the current density is not uniform across the beam but does not vary along its length then that affects the calculation of the plasma frequency reduction factor. Some idea of the magnitude of the correction needed can be estimated from the curves for the plasma frequency reduction factors of hollow beams [19]. If the beam is scalloping then both the plasma frequency and the plasma frequency reduction factor vary along the length of the beam. Where the beam radius is a maximum the local plasma frequency is a minimum and the two effects tend to cancel one another out. It has been shown that, when the variations are averaged along the beam, the mean value of ω_q does not depart very much from the value for the uniform beam [20]. When the beam is focused by a periodic permanent magnet system there is always a periodic variation of the beam radius. The author has found that good results are generally obtained if the equilibrium beam radius, calculated as described in Section 7.6.1, is taken to be the effective radius of the beam.

11.3.5   Higher-Order Modes

The analysis in Section 11.2.1 assumed that the waves which are excited on the electron beam are in the cylindrical TM_{01} mode. However, like circular waveguides, electron beams can also support higher-order space-charge modes such as those illustrated in Figure 11.14. These waves exist as pairs of fast and slow waves like the fundamental modes discussed above. The higher-order modes are usually ignored in the analysis of microwave tubes and only those with azimuthal symmetry could normally be excited by the fields of axially symmetrical structures. Further information can be found in [3, 4].
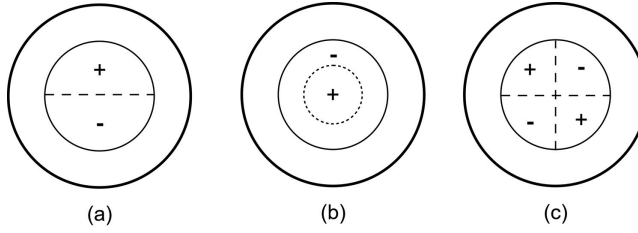
Figure 11.14: Examples of higher-order space-charge wave modes: (a) TM_{11}, (b) TM_{02} and (c) TM_{21}.

11.3.6   Cyclotron Waves

We have seen that when an electron beam is confined by an axial magnetic field B_z the electrons follow helical paths. The space-charge waves already discussed are associated with axial displacements of the electrons. Another set of modes exists for which the displacements are azimuthal. These modes, known as cyclotron waves, occur in fast and slow wave pairs in the same manner as the space-charge waves with propagation constants

$$\beta_\pm = \frac{\omega \mp \omega_c}{u_0},$$ (11.81)

where the cyclotron frequency $\omega_c = eB_z/m_0$. These waves are normally not excited in linear beam tubes because the azimuthal component of the electric field of the electromagnetic structure is usually zero. A fuller discussion of cyclotron modes can be found in [4, 21].

11.4   Space-Charge Wave Theory of the Interaction between a Beam and a Gap

The interaction between an electron beam and a narrow, gridded, gap $(M = 1)$ can be expressed in terms of the amplitudes of the transmission line analogue of space-charge waves. If the beam is initially unmodulated the boundary conditions are

$$V_+ + V_- = V_g$$ (11.82)

and, because the convection current must be continuous at the gap,

$$I_+ + I_- = Y_e \left(V_+ - V_- \right) = 0,$$ (11.83)

where time variation as $\exp(j\omega t)$ has been assumed and the symbol I indicates a conventional current. The solution to (11.82) and (11.83) is

$$V_+ = V_- = \frac{1}{2} V_g.$$ (11.84)

At a plane which is z from the gap the beam kinetic voltage is

$$V_1 = V_+ \exp(-j\beta_+ z) + V_- \exp(-j\beta_- z)$$
$$= V_g \cos(\beta_q z)\exp(-j\beta_e z), \tag{11.85}$$

where time variation $\exp(j\omega t)$ has been assumed. Similarly the RF current is

$$I_1 = I_+ \exp(-j\beta_+ z) + I_- \exp(-j\beta_- z)$$
$$= jY_e V_g \sin(\beta_q z)\exp(-j\beta_e z). \tag{11.86}$$

The corresponding equation derived using ballistic theory is (11.13). The two equations give the same result for short drift lengths. It can be seen that bunching of the beam can be thought of as beating between the fast and slow space-charge waves. When $\beta_q z = \pi/2$ the velocity modulation is zero and the current modulation is maximum. Similarly, when $\beta_q z = \pi$ the velocity modulation is maximum and the current modulation is zero. This pattern is repeated periodically.

The space-charge wave theory of the interaction between a beam and a gridless gap can be developed following the same approach as in Section 11.2.4. A small element of length dz at position z within the field of the gap functions like a narrow gridded gap so that the change in the amplitude of the kinetic voltage produced is

$$dV_1 = -E_z(z)dz \tag{11.87}$$

and the RF current leaving the elementary gap is zero. This can be expressed in terms of the amplitudes of the space-charge waves using (11.84)

$$dV_\pm = \frac{1}{2}dV_1 = -\frac{1}{2}E_z(z)dz. \tag{11.88}$$

The amplitudes of the space-charge waves at a plane z_1 are found by superimposing the amplitudes produced by each element dz. Thus

$$V_\pm(z_1) = -\frac{1}{2}\int_{-\infty}^{z_1} E_z(z)\exp(-j\beta_\pm(z_1 - z))\, dz. \tag{11.89}$$

Now the amplitudes of the waves do not depend on z_1, provided that it lies beyond the point where the field of the gap has fallen to zero. We may therefore extend the upper limit of integration to infinity and write the amplitudes of the space-charge waves as

$$V_\pm = -\frac{1}{2}\int_{-\infty}^{\infty} E_z(z)\exp(-j\beta_\pm z)\, dz. \tag{11.90}$$

Equation (11.90) has the same form as (11.31) so we conclude that

$$V_\pm = \frac{1}{2}M(\beta_\pm)V_g = \frac{1}{2}M_\pm V_g, \tag{11.91}$$

where $\beta_\pm$ has replaced β_e in the calculation of the gap coupling factor. Hence the beam kinetic voltage at a plane outside the field of the gap which is distance z from the centre of the gap is

$$V_1 = \frac{1}{2}V_g\left(M_+\exp\left(j\beta_q z\right)+M_-\exp\left(-j\beta_q z\right)\right)\exp\left(-j\beta_e z\right). \qquad (11.92)$$

Similarly the RF convection current is

$$I_1 = \frac{1}{2}Y_e V_g\left(M_+\exp\left(j\beta_q z\right)-M_-\exp\left(-j\beta_q z\right)\right)\exp\left(-j\beta_e z\right). \qquad (11.93)$$

11.4.1   Current Induced in a Gap by Space-Charge Waves

We saw, in Section 11.3.2, that the RF current of a space-charge wave induces a current in a surrounding conducting drift tube. This current is determined by the need to satisfy the boundary condition for the azimuthal component of the RF magnetic field $\left(H_\theta\right)$ on the surface of the drift tube. Figure 11.15 shows the coupling between a modulated beam and an external circuit through a gap in the drift tube. The sign convention used for the gap voltage and current is that they are positive when the gap field is accelerating, and power is flowing from the gap to the beam.

To first order, H_θ is unchanged by the presence of the gap and the boundary conditions require that the tangential magnetic field is continuous at the gap. Thus the RF magnetic field of the beam is a source which drives the current in the circuit external to the gap. In circuit terms this is represented by a current generator connected across the gap. The current of this generator must be the average of the induced wall current across the gap. Thus, for the two space-charge waves

$$i_{g\pm} = -\frac{1}{g}\int_{-g/2}^{g/2} i_{W\pm}\exp\left(-j\beta_\pm z\right)dz, \qquad (11.94)$$
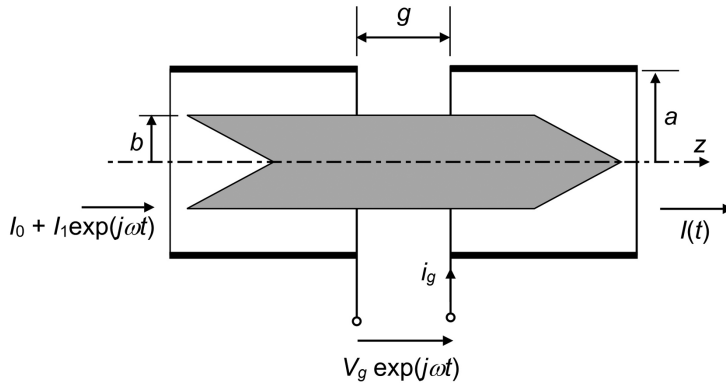
Figure 11.15: Current induced in a gap by a modulated beam (electrons positive).

where time variation at frequency ω has been assumed and i_W is in phase with RF convection current I_1. Carrying out the integration we find that

$$i_{g\pm} = -\frac{\sin(\beta_\pm g/2)}{\beta_\pm g/2} i_{W\pm} = -\mu_d(\beta_\pm) i_{W\pm}. \tag{11.95}$$

Thus the gap current is related to the induced current in the wall by the axial gap coupling factor μ_d. Now we saw, in Section 11.3.2, that the wall current is less than the convection current in the beam by a factor which is equal to the radial coupling factor defined by (11.38) for a beam in confined flow, and (11.40) for a Brillouin beam. Thus the current induced in the gap is given by

$$i_g = (M_+ I_+ + M_- I_-) = Y_e \begin{bmatrix} M_+ & -M_- \end{bmatrix} \begin{bmatrix} V_+ \\ V_- \end{bmatrix}. \tag{11.96}$$

11.4.2   Beam Loading of a Gridless Gap

The power carried by a modulated beam is the sum of the powers of the fast and slow space-charge wave (we recall that the slow wave carries negative power). Equation (11.91) shows that the two space-charge waves excited by the gap do not have exactly equal amplitudes. Thus active power supplied to the beam by the gap is

$$P_g = \frac{Y_e}{2}\left(|V_+|^2 - |V_-|^2\right). \tag{11.97}$$

Substituting for the voltages from (11.91) we obtain

$$P_g = \frac{Y_e}{8}\left(M_+^2 - M_-^2\right)|V_g|^2 = \frac{1}{2}G_b|V_g|^2, \tag{11.98}$$

where G_b is the *beam loading conductance.* This can be written

$$\frac{G_b}{G_0} = -\frac{1}{4}\theta\frac{\Delta M^2}{\Delta\theta}, \tag{11.99}$$

where $G_0 = I_0/V_0$, $\theta = \beta_e g$, $\Delta M^2 = M_+^2 - M_-^2$ and $\Delta\theta = \theta_+ - \theta_-$. Equation (11.99), which is unchanged by changing the sign of the charge on the electron, may be compared with (11.22) which was derived using ballistic analysis. It is found that the results given by these two equations are virtually identical if $\omega_q/\omega < 0.1$, which is the situation in many linear-beam tubes. Thus it is usually satisfactory to ignore space-charge effects when calculating the gap coupling factor and the beam loading [22, 23].

The beam loading susceptance of a gridless gap can be calculated using the equations in [24]. In general

$$Y_b = \frac{1}{2}Y_e\left[I_{av}'(j\beta_- g) - I_{av}'(j\beta_+ g)\right], \tag{11.100}$$

where Y_e is given by (11.40) and the definition of I_{av}' is given in the paper. For the special case where the electric field is uniform in the gap the real part of I_{av}' is given by

$$\mathrm{Re}\left[I_{av}'\left(j\beta g\right)\right] = \frac{1}{2}\left\{\frac{\sin\left(\beta g/2\right)}{\beta g/2}\cdot\frac{\sqrt{I_0^2\left(\gamma b\right)-I_1^2\left(\gamma b\right)}}{I_0\left(\gamma a\right)}\right\}^2 = \frac{1}{2}M\left(\beta\right)^2, \qquad (11.101)$$

where the definition of the radial coupling factor is given by (11.73) and the notation has been changed to be consistent with that used in this book. The real part of the beam loading admittance given by (11.100) is identical to (11.99). The imaginary part of I_{av}' for a uniform field is given by

$$\mathrm{Im}\left[I_{av}'\left(j\beta g\right)\right] = -\frac{\beta g - \sin\left(\beta g\right)}{\beta^2 g^2}\cdot\frac{\left[I_0^2\left(\gamma b\right)-I_1^2\left(\gamma b\right)\right]}{I_0^2\left(\gamma a\right)}$$

$$\cdots - 2\beta g\sum_{n=1}^{\infty}\left\{\frac{A_n}{\left(\beta^2 g^2 + r_n^2\right)^2} - \frac{B_n}{\beta^2 g^2 + r_n^2}\right\}. \qquad (11.102)$$

In this equation

$$r_n = \lambda_n\frac{g}{a}\sqrt{1-\left(\frac{\beta a}{\lambda_n}\right)^2\left(\frac{u_0}{c}\right)^2}, \qquad (11.103)$$

where λ_n are the zeroes of the Bessel function $J_0\left(\lambda\right)$,

$$A_n = 2\left(\frac{x_n + y_n}{1-\left(\beta a/\lambda_n\right)^2\left(u_0/c\right)^2}\right)\cdot\left(\frac{r_n - 1 + \exp\left(-r_n\right)}{r_n}\right), \qquad (11.104)$$

$$B_n = \left(\frac{x_n + y_n}{1-\left(\beta a/\lambda_n\right)^2\left(u_0/c\right)^2}\right)\cdot\left(\frac{3 - 2r_n - \left(3 + r_n\right)\exp\left(-r_n\right)}{r_n^3}\right)\cdots$$

$$+ \frac{2x_n\left(r_n - 1 + \exp\left(-r_n\right)\right)}{r_n^3}, \qquad (11.105)$$

$$x_n = \frac{g^2}{a^2}\frac{J_0^2\left(\lambda_n b/a\right)}{J_1^2\left(\lambda_n\right)}, \qquad (11.106)$$

$$\text{and} \quad y_n = \frac{g^2}{a^2}\cdot\frac{J_1^2\left(\lambda_n b/a\right)}{J_1^2\left(\lambda_n\right)}. \qquad (11.107)$$

Then the beam loading susceptance can be calculated. The summation in (11.102) converges rapidly so that three terms are usually sufficient to give good accuracy (see Worksheet 11.1). Figure 11.16 shows examples of curves of beam loading conductance and susceptance calculated in this way for a number of gridless gaps [10].
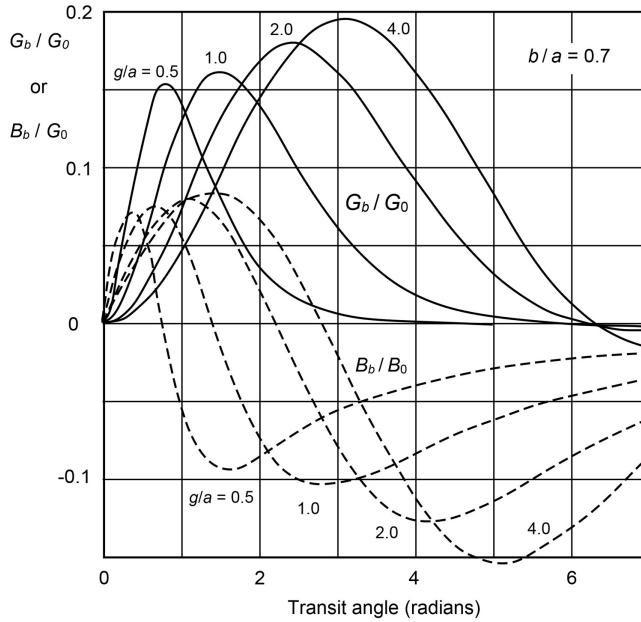
Figure 11.16: Beam-loading conductance and susceptance of gridless gaps of length g when $b/a = 0.7$
(copyright 1967, IEEE, reproduced, with permission, from [10]).

11.4.3   Beam Interaction with a Passive Gridless Gap

Bers [22] showed that the effect of a gridless gap can conveniently be represented by two drift regions separated by an equivalent narrow gridded gap. The interaction between a beam and a passive gap can be represented by the equivalent circuit shown in Figure 11.17. The admittance Y_L represents the properties of the circuit connected to the gap. The voltage induced in the gap by a modulated beam is

$$V_g = -\frac{i_g}{Y_T} = -\frac{Y_e}{Y_T}\begin{bmatrix} M_+ & -M_- \end{bmatrix}\begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix}, \tag{11.108}$$

where the initial modulation of the beam at the plane of the gap is represented by the amplitudes of the space-charge waves $(V_{1\pm})$ and $Y_T = Y_b + Y_L$. The negative sign in (11.108) is present because power must be transferred from the beam to a passive circuit. The modulation added to the beam by the gap voltage is given by (11.91). Thus the modulation of the beam leaving the gap is

$$\begin{bmatrix} V_{2+} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix} - \frac{Y_e}{2Y_T}\begin{bmatrix} M_+ \\ M_- \end{bmatrix}\begin{bmatrix} M_+ & M_- \end{bmatrix}\begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix}$$

$$= \begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix} - \frac{Y_e}{2Y_T}\begin{bmatrix} M_+^2 & M_+M_- \\ M_+M_- & M_-^2 \end{bmatrix}\begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix}. \tag{11.109}$$
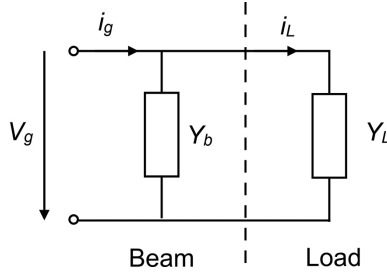
Figure 11.17: Equivalent circuit for the interaction between an electron beam and a passive gap (electrons positive).

The matrices for the drift regions take the form

$$\begin{bmatrix} V_{2+} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} \exp(-j\beta_+ z) & 0 \\ 0 & \exp(-j\beta_- z) \end{bmatrix} \begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix} \qquad (11.110)$$

and it is assumed that the drift regions extend outside the space occupied by the field of the gap. Thus the interaction between an electron beam and a series of passive gaps can be represented by cascaded matrices (see Section 13.2). The interaction with a series of gaps that are connected together is discussed in Section 11.6.

The relationship between the transmission line and space-charge wave parameters describing the modulation of the beam can be written

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ Y_e & -Y_e \end{bmatrix} \begin{bmatrix} V_+ \\ V_- \end{bmatrix}. \qquad (11.111)$$

Inverting this equation

$$\begin{bmatrix} V_+ \\ V_- \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & Z_e \\ 1 & -Z_e \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}. \qquad (11.112)$$

From (11.95) the modulation of a beam by a gap is

$$\begin{bmatrix} V_+ \\ V_- \end{bmatrix} = \frac{1}{2} \begin{bmatrix} M_+ \\ M_- \end{bmatrix} V_g. \qquad (11.113)$$

After the gap, from (11.111) and (11.113),

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ Y_e & -Y_e \end{bmatrix} \begin{bmatrix} M_+ \\ M_- \end{bmatrix} V_g = \frac{1}{2} \begin{bmatrix} M_+ + M_- \\ Y_e(M_+ - M_-) \end{bmatrix} V_g = \begin{bmatrix} M \\ G_b/M \end{bmatrix} V_g, \qquad (11.114)$$

where

$$M = \frac{1}{2}(M_+ + M_-) \qquad (11.115)$$

and, from (11.98),

$$G_b = \frac{Y_e}{4}\left(M_+^2 - M_-^2\right). \tag{11.116}$$

Similarly from (11.96) and (11.112) the induced current is

$$i_g = \frac{Y_e}{2}\begin{bmatrix} M_+ & -M_- \end{bmatrix}\begin{bmatrix} 1 & Z_e \\ 1 & -Z_e \end{bmatrix}\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} \dfrac{G_b}{M} & M \end{bmatrix}\begin{bmatrix} V_1 \\ I_1 \end{bmatrix}, \tag{11.117}$$

where V_1 and I_1 describe the modulation of the beam entering the gap. The effect of the gap on a modulated beam is then given by

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} - \frac{1}{Y_T}\begin{bmatrix} M \\ G_b/M \end{bmatrix} i_g. \tag{11.118}$$

Substituting for i_g from (11.117) gives, after some rearrangement [25]

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 - \dfrac{G_b}{Y_T} & -\dfrac{M^2}{Y_T} \\ -\dfrac{G_b^2}{M^2 Y_T} & 1 - \dfrac{G_b}{Y_T} \end{bmatrix}\begin{bmatrix} V_1 \\ I_1 \end{bmatrix}. \tag{11.119}$$

This equation is unchanged if the sign of the charge on the electron is reversed.

11.5   Continuous Interaction with a Slow-Wave Structure

Figure 11.18 shows, schematically, the continuous interaction between an electron beam and a uniform slow-wave structure (circuit). At any given frequency the circuit can be regarded as a transmission line having characteristic impedance Z_c on which the propagation constant is β_0. The potential difference between the surface of the structure and a surrounding earthed conductor is

$$V_c = V_1 \exp j\left(\omega t - \beta_0 z\right). \tag{11.120}$$
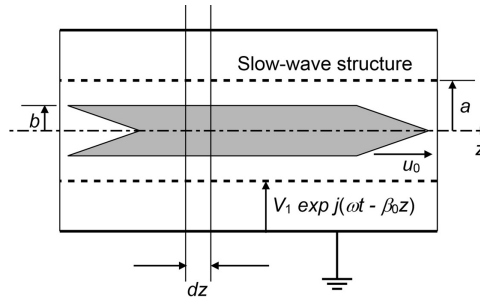
Figure 11.18: Interaction between an electron beam and a uniform slow-wave structure (electrons negative).

The amplitude of the axial component of electric field at the surface of the circuit is given by (4.11). When the phase velocity is much less than the velocity of light this becomes

$$E_z(a) = -\frac{\partial V_c}{\partial z} = j\beta_0 V_c. \tag{11.121}$$

If it is assumed that the electron beam is in confined flow (which has been seen to be a good approximation if m > 1.5) then the effective electric field acting on the electrons is the average over the cross-section of the beam so that

$$E_z = j\beta_0 \mu_c(\gamma_0) V_c, \tag{11.122}$$

where the radial coupling factor μ_c is given by (11.38) and $\gamma_0 = \sqrt{\beta_0^2 - \omega^2/c^2}$. The relationship between the electric field acting on the beam and the power flow along the circuit can be written as

$$\frac{|E_z|^2}{2\beta_0^2 P} = \mu_c^2 Z_c = Z_C, \tag{11.123}$$

where Z_C is the *Coupling Impedance* of the circuit averaged over the cross-section of the beam. In the limiting case when the radius of the beam is vanishingly small then (11.123) becomes

$$\frac{|E_z(0)|^2}{2\beta_0^2 P} = \frac{Z_c}{I_0(\gamma_0 a)^2} = Z_P, \tag{11.124}$$

where Z_P is known as the *Pierce Impedance* of the circuit [26].

Now consider the interaction taking place in a short length dz, as shown in Figure 11.18. The changes in the amplitudes of the space-charge waves caused by the field of the circuit are given by (11.88). These must be added to those caused by the propagation of the waves already existing on the beam so that

$$\frac{\partial V_\pm}{\partial z} = -j\beta_\pm V_\pm + \frac{1}{2}\mu_c \frac{\partial V_c}{dz}. \tag{11.125}$$

The signs in this equation assume that the charge on the electron is negative. This is essential for consistency with the conventional voltages and currents on the circuit. The amplitude of the wave on the circuit (V_c) can be expressed as the sum of the amplitudes of the forward and backward waves so that (11.125) becomes

$$\frac{\partial V_\pm}{\partial z} = -j\beta_\pm V_\pm + \frac{1}{2}\mu_c \frac{\partial}{dz}(V_f + V_b). \tag{11.126}$$

Note that we have assumed here that the beam does not load the circuit appreciably so that the modulation of the beam can be expressed in terms of the circuit voltage. This is justified because the beam impedance is much greater than that of the circuit.

The RF displacement current flowing from the beam into the element of the circuit can be found by applying the equation

$$\iint_S \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot d\mathbf{S} = 0 \tag{11.127}$$

to the closed surface bounded by the planes z and $z + dz$ and the surface of the circuit at $r = a$. Then the radial displacement current flowing into the element of the circuit is

$$di_R = -\frac{\partial}{\partial z}(I_C + I_D)dz = -\mu_c \frac{\partial I_C}{\partial z}dz, \tag{11.128}$$

where I_C and I_D are the convection current and the axial displacement current, as defined in (11.68). The radial displacement current is divided equally between the forward and backward directions so that the increments in the forward and backward waves on the circuit are

$$dV = -\frac{1}{2}\mu_c Z_c \frac{\partial I_C}{\partial z}dz. \tag{11.129}$$

This increment is added to amplitudes of the waves which already exist on the circuit. The convection current can be expressed as the sum of the currents in the fast and slow space-charge waves so that the forward and backward waves on the circuit satisfy

$$\frac{\partial V_f}{\partial z} = -j\beta_0 V_f + \frac{1}{2}\mu_c Z_c Y_e \frac{\partial}{\partial z}(V_+ - V_-) \tag{11.130}$$

and

$$\frac{\partial V_b}{\partial z} = j\beta_0 V_b - \frac{1}{2}\mu_c Z_c Y_e \frac{\partial}{\partial z}(V_+ - V_-). \tag{11.131}$$

Note that we have assumed that the coupling is weak so that the characteristic impedances of the modes are not significantly altered by it. We have also assumed that the coupling takes the form of induction of current on the structure. This is justified because the impedance of the beam is much higher than that of the structure.

Equations (11.126), (11.130) and (11.131) can be rearranged as coupled-mode equations (see Section 1.3.8). The result is

$$j\frac{\partial}{\partial z}\begin{bmatrix} V_f \\ V_b \\ V_+ \\ V_- \end{bmatrix} = \begin{bmatrix} \beta_0 & 0 & -A\beta_+ & -A\beta_- \\ 0 & -\beta_0 & A\beta_+ & A\beta_- \\ B\beta_0 & -B\beta_0 & \beta_+ & 0 \\ B\beta_0 & -B\beta_0 & 0 & \beta_- \end{bmatrix}\begin{bmatrix} V_f \\ V_b \\ V_+ \\ V_- \end{bmatrix}, \tag{11.132}$$

where

$$A = \frac{1}{2}\mu_c Z_c Y_e \qquad (11.133)$$

and

$$B = \frac{1}{2}\mu_c. \qquad (11.134)$$

If the voltages are normalised, following Louisell [21], by dividing them by the square root of the appropriate characteristic impedance, then the magnitudes of the coupling terms are typically of the order of 1% of the diagonal terms. This confirms the validity of the assumption that the coupling as weak, and that the propagation constants and characteristic impedances of the modes are not changed by it. The propagation constants of the four coupled modes are the eigenvalues of the matrix so that

$$\big| [C] - \beta[I] \big| = 0, \qquad (11.135)$$

where $[C]$ is the matrix in (11.132) and $[I]$ is the unit matrix. Expanding the determinant gives the dispersion equation of the coupled system

$$\big[\beta_0^2 - \beta^2 \big]\big[\beta_q^2 - (\beta_e - \beta)^2 \big] = -\mu_c^2 Z_c Y_e \beta_q \beta_0 \beta^2. \qquad (11.136)$$

Substitution for Y_e using (11.78) gives

$$\big[\beta_0^2 - \beta^2 \big]\big[\beta_q^2 - (\beta_e - \beta)^2 \big] = -\frac{1}{2}\big(\mu_c^2 Z_c G_0\big)\beta_e \beta_0 \beta^2. \qquad (11.137)$$

where $G_0 = |I_0/V_0|$. This is a quartic equation in β which therefore has four roots at any given frequency. The bracketed term on the right-hand side of this equation is the ratio of the coupling impedance defined by equation (11.123) to the DC impedance of the beam. It may also be noted that the right-hand side of the equation is independent of ω_q.

To examine the solutions of (11.137) let us consider first the case when the coupling between the beam and the circuit is removed, by setting $\mu_c = 0$. The solutions are then

$$\beta = \pm\beta_0 \qquad (11.138)$$

for the forward and backward waves modes on the circuit in the absence of the beam, and

$$\beta = \beta_e \mp \beta_q \qquad (11.139)$$

for the fast and slow space-charge waves on the beam. Figure 11.19(a) shows the uncoupled modes defined by these equations. The axes are normalised to the *synchronous point* (ω_s, β_s) which is defined here as the point of intersection of

(a)

(b)

—————— Circuit waves (real parts)
— — · Space-charge waves (real parts)
- - - - · Coupled modes (imaginary parts x 50)
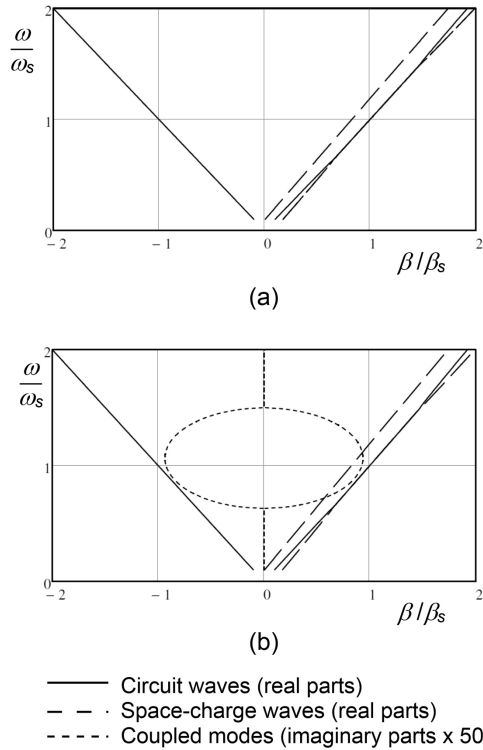
Figure 11.19: Continuous interaction between an electron beam and a uniform slow-wave structure: (a) uncoupled modes and (b) coupled modes. The real parts are shown by solid and dashed lines. The imaginary parts multiplied by 50 by the dotted lines. At the synchronous point $A = 0.0014$, $B = 0.25$, and $\omega_q/\omega = 0.1$.

the forward circuit wave, and the slow space-charge wave, curves in the absence of coupling.[1] The parameter values are typical of those found in helix TWTs but for simplicity it has been assumed that the reduced plasma frequency is constant. Coupled-mode theory (see Section 1.3.8) leads us to expect type C coupling between the forward wave on the circuit, and the slow space-charge wave, in the region where they are close to synchronism. Figure 11.19(b) shows the coupled-mode diagram computed from (11.132). The coupling produces a region where a pair of complex conjugate solutions exists. These are known as the *growing wave* and the *decaying wave*. It is evident that the backward wave on the circuit is hardly affected by the coupling. Similarly, the fast space-charge wave is not much affected by the coupling, especially if the space-charge density is high.

The solutions of (11.137) were first derived by Pierce, who assumed that the backward wave was unchanged, and solved the resulting the cubic equation [26, 27]. He employed normalised parameters which have remained part of the common

---

[1] This definition differs from that commonly used $(\beta_0 = \beta_e)$ but, in the author's opinion the definition used here is to be preferred.

language of TWT theory although they are not an ideal choice. The Pierce parameters describing a loss-less tube are:

- The *gain parameter*

$$C = \left| \frac{I_0 Z_C}{4 V_0} \right|^{\frac{1}{3}} = \left(\frac{G_0 Z_C}{4} \right)^{\frac{1}{3}},$$

(11.140)

which resembles the constant term on the right hand side of (11.137).
- The *velocity parameter*

$$b = \frac{1}{C} \left(\frac{u_0}{v_p} - 1 \right),$$

(11.141)

where v_p is the phase velocity of the waves on the circuit.
- The *space-charge parameter*

$$QC = \frac{1}{4C^2} \left(\frac{\omega_q / \omega}{1 + \omega_q / \omega} \right)^2.$$

(11.142)

Curves giving the solutions to the quartic equation (11.136) in terms of the Pierce parameters were published by [28, 29]. Today it is better to program a computer to find the eigenvalues of (11.132) directly (see Worksheet 11.4).

Useful insight can be obtained from the two-wave approximation which assumes that only coupling between the slow space-charge wave and the forward wave on the circuit is important. The approximate equation is obtained from (11.132) by dropping the terms in V_b and V_+ to give

$$j \frac{\partial}{\partial z} \begin{bmatrix} V_f \\ V_- \end{bmatrix} = \begin{bmatrix} \beta_0 & -A\beta_- \\ B\beta_0 & \beta_- \end{bmatrix} \begin{bmatrix} V_f \\ V_- \end{bmatrix}.$$

(11.143)

The eigenvalues of this matrix are the solutions of the quadratic dispersion equation

$$\beta^2 - (\beta_0 + \beta_-)\beta + \left(1 + \frac{1}{4} \mu_c^2 Z_c Y_e \right) \beta_0 \beta_- = 0,$$

(11.144)

which can be solved using the usual formula to give

$$\beta = \frac{1}{2}(\beta_0 + \beta_-) \pm \frac{1}{2} \sqrt{(\beta_0 - \beta_-)^2 - \beta_0 \beta_- \mu_c^2 Z_c Y_e}.$$

(11.145)

This solution reveals a number of important things about the travelling-wave interaction. If the term under the square root is negative then β has a pair of complex conjugate roots. These are the growing and decaying waves whose amplitudes vary with z as $\exp(\pm \alpha z)$ where

$$\alpha = \frac{1}{2} \sqrt{\mu_c^2 Z_c Y_e \beta_0 \beta_- - (\beta_0 - \beta_-)^2}.$$

(11.146)

The real part of each of these roots is given by the first term of (11.145) which is the mean of the two uncoupled roots. The validity of the two-wave approximation increases as the plasma frequency increases because the coupling between the slow space-charge wave and the forward wave on the circuit is reduced. Table 11.1 shows a comparison between the values of the four propagation constants computed from (11.132), and the approximate values obtained from (11.145) (see Worksheet 11.4). The same parameter values are used as in Figure 11.18. The figures in brackets are those for the uncoupled modes. It can be seen that the difference between the two sets of values is small.

At the synchronous point (11.145) becomes

$$\beta = \beta_0 \left(1 \pm j \frac{\mu_c}{2} \sqrt{Z_c Y_e} \right),$$

(11.147)

where the positive sign corresponds to the growing wave. This is very close to the condition that the imaginary part of β is greatest. Thus the maximum gain per circuit wavelength $\left(\lambda_0 = 2\pi/\beta_0 \right)$ is given by

$$\text{Gain} = 20 \log_0 \left[\exp\left(\pi \mu_c \sqrt{Z_c Y_e} \right) \right] = 27.3 \mu_c \sqrt{Z_c Y_e} \quad \text{dB } \lambda_0^{-1}.$$

(11.148)

The parameter $\mu_c^2 Z_c Y_e$ is typically in the range 0.01 to 0.02 corresponding to gain of around 3 to 4 dB per wavelength. It is not surprising that the gain increases with coupling impedance (corresponding to increasing interaction field for a given power flow), and with the electronic admittance of the beam (corresponding to increasing current and decreasing voltage).

If we define $x = \beta_-/\beta_0$, $y = 2\alpha/\beta_0$ and $C' = \mu_c^2 Z_c Y_e$ then (11.146) can be written as

$$y = \sqrt{C'x - (x-1)^2}.$$

(11.149)

Figure 11.20 shows curves of y against x for a range of typical values of C'. Careful examination of these curves shows that they are not quite symmetrical. The peak of the curve lies very close to $x = 1$ but moves gradually to the right as C'

Table 11.1: Comparison between 2-wave and 4-wave models of the continuous interaction between an electron beam and a slow-wave structure

ω/ω_s	Model	β/β_0	β/β_0	β/β_0	β/β_0
0.5	2 waves	1.010	1.076	(0.744)	(−1.000)
	4 waves	1.012	1.075	0.742	−1.000
1.0	2 waves	1.001+0.019i	1.001−0.019i	(0.829)	(−1.000)
	4 waves	1.001+0.019i	1.001−0.019i	0.828	−1.000
1.5	2 waves	0.986+0.005i	0.986−0.005i	(0.858)	(−1.000)
	4 waves	0.988	0.985	0.856	−1.000
2.0	2 waves	0.996	0.962	(0.872)	(−1.000)
	4 waves	0.997	0.961	0.871	−1.000
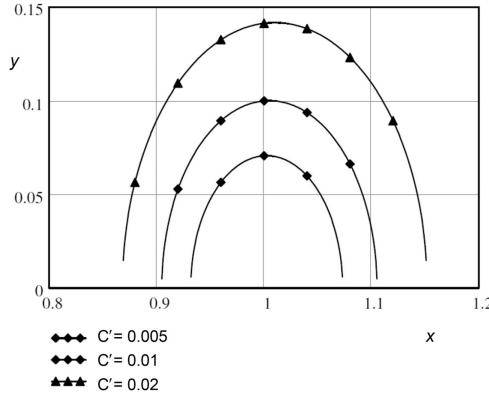
Figure 11.20: Illustration of the dependence of the bandwidth of a TWT calculated from (11.149) on $C' = \mu^2 Z_c Y_e$.

increases. The curves show that the gain per unit length is very nearly proportional to $\sqrt{C'}$. Figure 11.20 also shows that the bandwidth of the interaction increases as C' increases. Since the bandwidth has been defined in terms of the ratio β_-/β_0 the bandwidth in terms of frequency depends also upon the angle of intersection between the dispersion curves for these two modes. The angle can be reduced by reducing the effective plasma frequency of the beam, and also by adjusting the shape of the dispersion curve of the slow-wave circuit. Although the curves in Figure 11.20 show the general properties of the travelling-wave interaction, it is important to remember that the gain curves for real TWTs differ from them in detail because of dependence of ω_q, μ_c, Z_c, and Y_e on frequency.

The analysis presented so far has ignored the effect of losses in the slow-wave structure. In practice all real slow-wave structures have some transmission loss in the absence of the electron beam. Also it is common to introduce additional loss to ensure stability against oscillations. The effects of the cold loss of the circuit can be included in the theory by replacing β_0 for the forward wave by $(\beta_0 - j\alpha_0)$. The two-wave approximation to the solution of the dispersion equation (11.145) shows that the growing wave then grows as $\exp(\alpha - \alpha_0/2)$. The effect of this is to reduce the overall gain of a section of slow-wave structure by half the cold loss in decibels. A more exact figure can be arrived at in particular cases by allowing for the cold loss when solving (11.132).

11.6   Discrete Interaction with a Slow-Wave Structure

In high-power TWTs the slow-wave structure is not continuous and the field acting on the beam is the superposition of the fields of a series of gaps [30]. It is self-evident that this is true for coupled-cavity, ring-bar, and other similar periodic slow-wave structures (see Chapter 4). In fact the field of a helix should really also be considered in the same way, particularly for the space harmonic fields. Whilst it
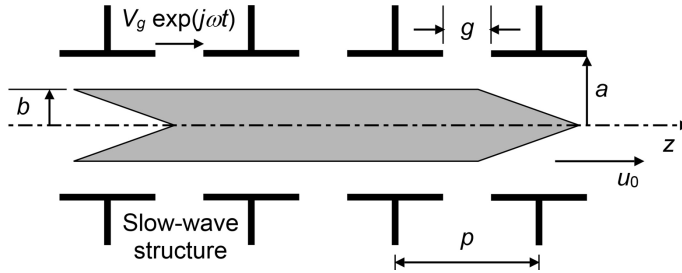
Figure 11.21: Interaction between an electron beam and a periodic slow-wave structure (electrons negative).

is possible to model the forward-wave interaction in a helix TWT as continuous, it does not provide an adequate description for backward wave interactions.

In order to model the interaction between an electron beam and a periodic slow-wave structure let us consider a single cell of the circuit whose pitch (length) is p, as shown in Figure 11.21. For consistency the polarity of the gap voltage has been reversed so that the electron charge is negative. Within this cell we will use the drift-gap-drift model described in Section 11.4.3. This model is satisfactory for the circuits used in space-harmonic TWTs because the gaps are small compared with the pitch of the circuit, and the fields of adjacent gaps do not overlap to any appreciable extent. There is a problem, however, if the model is applied to forward-fundamental coupled-cavity TWTs because the length of the gap is then comparable with the pitch of the circuit and the fields of adjacent gaps do overlap. Any attempt to model the simultaneous interaction between an electron and the fields of two, or more, gaps results in great difficulties. The best solution seems to be to regard the total electric field within one cell of the circuit as 'belonging' to one gap. This approach results in discontinuities in the field at the junctions between the cells, and all the induced current is regarded as 'belonging' to that cell rather than being divided between adjacent cells. Nevertheless, the results obtained show that the approximation is acceptable. These issues will not be pursued any further here and we will assume that overlapping between the fields of adjacent gaps can be neglected.

The matrix describing the interaction between the beam and a localised gap is found, by reference to Figure 11.17, and using (11.91) and (11.96) to be

$$
\begin{bmatrix} V_{2+} \\ V_{2-} \\ i_L \end{bmatrix} = \begin{bmatrix} 1 & 0 & M_+/2 \\ 0 & 1 & M_-/2 \\ M_+ Y_e & -M_- Y_e & Y_b \end{bmatrix} \begin{bmatrix} V_{1+} \\ V_{1-} \\ V_g \end{bmatrix}. \tag{11.150}
$$

This matrix is unchanged when the signs of the currents and voltages are reversed. The term Y_b is typically of the order of 4×10^{-6} which is very small compared with the other terms and can, therefore, be neglected. The physical explanation for this is that the beam impedance of several hundred kilohms is much greater than the total impedance of the circuit (a few hundred ohms) so that the beam does not load the circuit to an appreciable extent. This must be distinguished from the situation

which exists when the beam interacts with a resonant circuit where the beam and circuit impedances are comparable with each other. Equation (11.150) can be written in partitioned form as

$$\begin{bmatrix} v_2 \\ i_L \end{bmatrix} = \begin{bmatrix} I & B_2 \\ B_3 & Y_b \end{bmatrix} \begin{bmatrix} v_1 \\ V_g \end{bmatrix}$$

(11.151)

where $[v] = \begin{bmatrix} V_+ \\ V_- \end{bmatrix}$, $[I]$ is the 2×2 unit matrix, $[B_2] = \dfrac{1}{2} \begin{bmatrix} M_+ \\ M_- \end{bmatrix}$ and $[B_3] = Y_e [M_+ \quad -M_-]$.

The generalised matrix for localised coupling to a transmission line, when there is no localised energy storage is, from (4.22),

$$\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_g \end{bmatrix} = \begin{bmatrix} 1 & 0 & -NZ_c/2 \\ 0 & 1 & NZ_c/2 \\ N & N & 0 \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \\ i_L \end{bmatrix},$$

(11.152)

where Z_c is the characteristic impedance of the line and $N = \sqrt{Z_T/Z_c}$ from (4.28). Equation (11.152) can be written in partitioned form as

$$\begin{bmatrix} V_2 \\ V_g \end{bmatrix} = \begin{bmatrix} I & S_2 \\ S_3 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ i_L \end{bmatrix}$$

(11.153)

where $[V] = \begin{bmatrix} V_f \\ V_b \end{bmatrix}$, $[S_2] = \dfrac{NZ_c}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $[S_3] = [N \quad N]$. Eliminating V_g and i_L between (11.151) and (11.153) yields

$$\begin{bmatrix} V_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} I - Y_b [S_2][S_3] & [S_2][B_3] \\ [B_2][S_3] & I \end{bmatrix} \begin{bmatrix} V_1 \\ v_1 \end{bmatrix} = [G] \begin{bmatrix} V_1 \\ v_1 \end{bmatrix}.$$

(11.154)

We will assume that the gap is at the centre of the cell so that the drift matrix $[D]$ for half the pitch is

$$\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_{2+} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} \exp(-j\beta_0 p/2) & 0 & 0 & 0 \\ 0 & \exp(j\beta_0 p/2) & 0 & 0 \\ 0 & 0 & \exp(-j\beta_+ p/2) & 0 \\ 0 & 0 & 0 & \exp(-j\beta_- p/2) \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{1b} \\ V_{1+} \\ V_{1-} \end{bmatrix}.$$

(11.155)

Thus the complete representation of the cell is

$$[V_2] = [D][G][D][V_1] = [P][V_1].$$

(11.156)

The normal modes of the coupled system are the eigenvalues of the matrix $[P]$. These are found in the usual way by requiring that $[V_2] = \lambda [V_1]$ so that the values of λ are the solutions of the dispersion equation

$$\left| [P] - \lambda [I] \right| = 0.$$

(11.157)
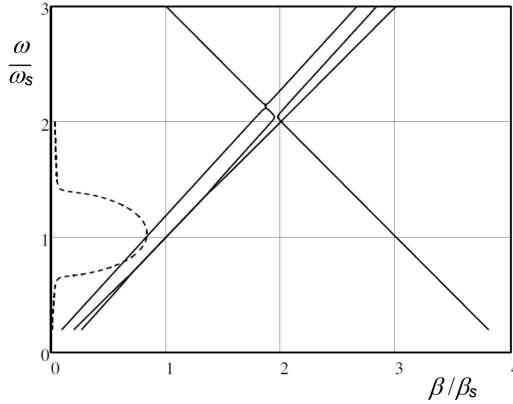
Figure 11.22: Dispersion diagram for coupling between an electron beam and a periodic slow-wave structure. The imaginary part $(\times 50)$ is shown by the dotted line.

Since $[P]$ is a 4×4 matrix, equation (11.157) has four roots which may be expressed in terms of the propagation constants of the coupled system using

$$\lambda_n = \exp(-j\beta_n p). \tag{11.158}$$

Thus (11.157) represents the properties of a discretely coupled system in the same way that (11.135) represents the properties of a continuously coupled system. For any particular combination of beam and circuit it is possible to find a numerical solution to (11.157) (see Worksheet 11.5). Figure 11.22 shows the coupled modes obtained for a typical case where ω and β are normalised to the synchronous values. The predicted region of gain is very similar to that seen in Figure 11.18(b). There is, however, one important difference because the periodic nature of the interaction also permits interactions with space-harmonic waves.

The interaction with the $m = -1$ space harmonic of the backward wave on the circuit is shown in more detail in Figure 11.23. This diagram shows Type B coupling between the backward circuit wave and the fast space-charge wave, and Type D coupling between the backward circuit wave and the slow-space-charge wave (see Section 1.3.8). These interactions are predicted by coupled-mode theory. To understand their implications we must apply the correct boundary conditions. For coupling between the backward wave and the fast space-charge wave the only permitted solutions are those in which the power in the forward wave is progressively reflected into the backward wave. Thus, for a sufficiently long region of coupling, the power in the forward wave is completely reflected leading, to a stop band in the dispersion diagram. For coupling between the backward wave and the slow space-charge wave, however, the transfer of power from the slow space-charge wave to the backward wave causes the amplitude of the space-charge wave to grow. It is therefore possible for a backward wave with finite amplitude at the input to be created by the injection of a very low noise power in the space-charge wave. The result is the possibility of backward-wave oscillation, which is discussed in the next section.
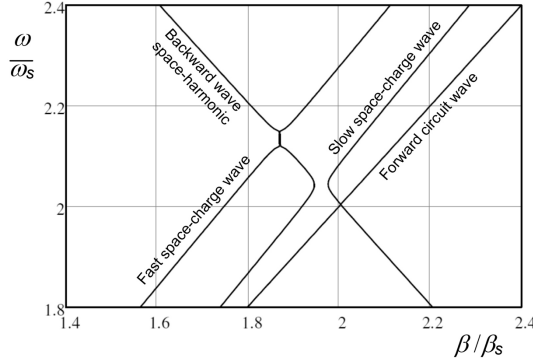
Figure 11.23: Detail of the dispersion diagram for coupling between an electron beam and a periodic slow-wave structure showing the interaction with the backward wave.

11.7   Backward-Wave Interactions

It was shown in the previous section that the space-charge waves can interact with the $m = -1$ space harmonic of the backward wave on the structure. This leads to the possibility of amplification of the backward wave. Since energy must be conserved for the system comprising the beam and the slow-wave structure, it is not possible for any amplification to result from the interaction with the fast space-charge wave. The possibility of an interaction between the backward circuit wave and the slow space-charge wave has generally been studied using the theory of a continuous interaction [3, 31–33]. The equations are developed in exactly the same manner as in Section 11.5 except that the sign of the characteristic impedance of the slow-wave structure is reversed because the power flow is in the negative direction. In order to understand the implications of this it is sufficient to use the two-wave approximation so that (11.144) becomes

$$\beta^2 - (\beta_{-1} + \beta_-)\beta + \left(1 - \frac{1}{4}\mu_c^2 Z_c Y_e\right)\beta_{-1}\beta_- = 0, \tag{11.159}$$

where $\beta_{-1} = -\beta_0 + 2\pi/p$ and circuit impedance Z_c is that for the appropriate space harmonic. The solutions to this equation are

$$\beta = \frac{1}{2}(\beta_{-1} + \beta_-) \pm \frac{1}{2}\sqrt{(\beta_{-1} - \beta_-)^2 + \beta_{-1}\beta_- \mu_c^2 Z_c Y_e} \tag{11.160}$$

and we see at once that they are always real. At the synchronous point $\beta_{-1} = \beta_-$ and

$$\beta = \beta_{-1}\left(1 \pm \frac{\mu_c}{2}\sqrt{Z_c Y_e}\right). \tag{11.161}$$

Thus the amplitude of the backward wave must vary as

$$V_b = \left[A\exp\left(j\beta_{-1}z\mu_c\sqrt{Z_c Y_e}/2\right) + B\exp\left(-j\beta_{-1}z\mu_c\sqrt{Z_c Y_e}/2\right)\right]\exp(-j\beta_{-1}z), \tag{11.162}$$

where A and B are constants. If a backward wave interaction takes place in a system of length L where the output end of the circuit is matched, then V_b must be finite when $z = 0$ and zero when $z = L$. The conditions for this are that $A = B$ so that the amplitude of the backward wave varies as $\cos\left(\beta_{-1} z \mu_c \sqrt{Z_c Y_e}/2\right)$ and $\beta_{-1} L \mu_c \sqrt{Z_c Y_e}/2 = \pi/2$. Thus the oscillations are just possible when

$$Y_e = \left(\frac{\pi}{\beta_{-1} L \mu_c}\right)^2 Y_c. \tag{11.163}$$

Substituting for Y_e from (11.78) gives the least current for which oscillation can occur

$$I_0 = \frac{2\omega_q V_0}{\omega}\left(\frac{\pi}{\beta_{-1} L \mu_c}\right)^2 Y_c. \tag{11.164}$$

For helix slow-wave structures μ_c is calculated for the TM_{11} mode. More exact results for the start-oscillation current can be obtained by solving the full dispersion equation and applying the appropriate boundary conditions. Good agreement has been shown between the theoretical and experimental values of start-oscillation current [34]. If the current is insufficient to permit oscillations to occur, and a backward wave is launched at the output end of the structure, then the boundary conditions are modified and an amplified wave appears at the input end. However, because the interaction produces a standing wave, the growth is sinusoidal rather than exponential, as in the case of the forward-wave interaction.

Equation (11.159) was derived assuming that all quantities varied sinusoidally with time. To examine the rate of growth of the oscillations with time it is necessary to allow the frequency to be complex [31]. We therefore replace ω by $(\omega - j\delta)$ so that

$$\beta_{-1} \rightarrow \beta_{-1} - \frac{j\delta}{v_p}, \tag{11.165}$$

where v_p is the phase velocity of the wave on the circuit. The sign of the imaginary part is positive because the slope of the dispersion curve of the backward wave is negative. Similarly

$$\beta_- \rightarrow \beta_- + \frac{j\delta}{u_0}. \tag{11.166}$$

It can be shown that, when there are no reflections at the output of the section, oscillations can only occur at the synchronous point [31]. Then (11.160) becomes

$$\beta = \beta_{-1} + \frac{j\delta}{2}\left(\frac{1}{v_p} - \frac{1}{u_0}\right) \pm \frac{1}{2}\sqrt{\mu_c^2 Z_c Y_e \beta_{-1}^2 - \delta^2 \left(\frac{1}{v_p} + \frac{1}{u_0}\right)^2}, \tag{11.167}$$

where it is assumed that $\delta \ll \omega$ so that second-order terms may be neglected. It can be seen that β is now complex. When the boundary conditions are applied it is found that δ increases as the beam current is increased beyond the start-oscillation

current. Thus, in general, both β and ω are complex when a backward wave interaction occurs, confirming the existence of an absolute instability. Their values are found by solving the dispersion equation subject to the appropriate boundary conditions [33]. Once oscillation has commenced the amplitude increases with time until it is limited by non-linear effects.

11.8 Large-Signal Modelling of Beam-Wave Interactions

The fundamental equations for the interaction between an electron beam and an RF electric field in one dimension are (11.45) to (11.48). The theory of space-charge waves was developed by linearising these equations using the small-signal assumption in (11.49). The result is an Eulerian analysis which treats the electron beam as a compressible fluid. It is possible to develop this analysis so that it applies under large-signal conditions [35–37]. The method breaks down whenever electrons overtake one another so that the properties of the beam at a given plane are no longer a single-valued function of position.

In general, large-signal theory must use Lagrangian analysis in which the electron beam is described by a set of sample electrons. The motions of these electrons are found by self-consistent numerical integration of the equations of motion, taking account of the space-charge field and the circuit field. It is possible to use commercially available particle in cell (PIC) codes for this purpose [38]. These codes model the whole electromagnetic system of the tube, including the electron dynamics, using either a finite difference or finite element method. The model can include the motion of large numbers of sample electrons in three dimensions so that it is possible to model a complete tube. This approach is very time consuming even with the fastest modern computers.

For many purposes it is better to use specially constructed models in which the properties of the electromagnetic structure outside the interaction region are modelled by equivalent circuits. This has the advantage that their properties are known and do not have to be re-computed at each time step. It is also possible to achieve good results using quite small numbers of sample electrons each of which is associated with a disc, or ring, of charge. The disc models divide the unmodulated beam into identical transverse slices and track the electron motion only in the axial direction using radially averaged values of the electric field. The motion of a disc is defined by two coordinates $(z, \dot{z})$ which are functions of time and the model is therefore one-dimensional. Ring models subdivide the discs into concentric rings of charge which are able to move independently of one another. In these models azimuthal symmetry is assumed, the motion of a disc is defined by five coordinates $(r, z, \dot{r}, \dot{\theta}, \dot{z})$, and they are therefore known as 2.5-dimensional models. It is usual to track the motion of the set of sample electrons which enter the interaction space in one RF cycle. In disc models 24 identical discs are usually sufficient.

The fundamental equations describing the problem ((11.45) to (11.48)) show that the variables depend on both distance and time. In order to carry out the numerical

integration of the equations it is necessary to choose one of these to be the independent variable. When the independent variable is z (a distance stepping model) the coordinates of the electrons in a 1D model are $(t, \dot{z})$. Conversely in a time stepping model their coordinates are $(z, \dot{z})$. As the solution proceeds it is necessary to compute the current induced in the electromagnetic structure at intervals, and the electric field of the space-charge. Both of these are continuous functions which have to be reconstructed from the information available from the sample electrons. The convection current at a plane, which is strictly periodic in time, is determined from the times at which the sample electrons cross that plane. The electric field of the space-charge is found from the positions of the sample electrons at a given time. Thus the convection current can be computed directly in a distance stepping model, but the space-charge field must be calculated by interpolation. Conversely, in a time stepping model the space-charge field can be computed directly, but interpolation is necessary to find the induced current. Examples of computer programs using both methods are to be found in the literature (see, for example: time step [39, 40], and distance step [41, 42]).

The variation of the beam current with time at any plane is computed by associating a periodically repeated pulse of current with each sample electron. This current can be expressed as a set of harmonic currents by Fourier analysis, and the results summed over all the electrons, to give the amplitude and phase of each harmonic of the beam current. It might be thought that this method would be subject to numerical noise when the modulation of the beam is small, but that has not been found to be a problem.

The calculation of the space-charge field is more difficult because the distribution of charge density in space is not periodic. Thus, if the time-step method is used, it is necessary either to assume an approximate periodic distribution, or to track more than a single wavelength of electrons so that the forces on the electrons at the ends of the group are calculated correctly. It is important to be able to model electrons which overtake one another, and also those which overtake those in an adjacent wavelength. With the distance-step method the positions of the electrons are found by interpolation on data from previous steps, or by extrapolation to future steps. This does not raise as many problems as might appear at first because the space-charge forces have short range, and those arising from discs at a distance greater than half a wavelength are usually negligible. Also, as we shall see, it is necessary to make multiple passes through the interaction space to find a self-consistent solution so it is possible to store the arrival times of the electrons at each plane. Thus data from a previous iteration can be used when calculating the space-charge field of electrons whose positions are beyond the current plane. The actual calculation of the space-charge field from the positions of the electrons can be achieved by assigning charges to the nearest nodes of a spatial grid and then solving Poisson's equation numerically [40, 43]. However, the space-charge calculation forms the innermost of a number of nested loops in the computational algorithm and it therefore has a very important effect on the time taken by the computation. For this reason it is better to compute the field of a single disc, or ring, and store the results in a look-up

table [41, 44] (see Worksheet 11.6). The infinitely thin discs used in early models led to problems with singularities in the field as electrons passed each other. That problem is removed by using discs having finite thickness. In 1D models it is possible to use compressible discs defined by the charge between each pair of sample electrons [45] but the additional complexity is not justified by the results.

In order to find the motion of the electrons it is also necessary to know the self-consistent field of the electromagnetic structure surrounding the beam. If the structure is excited from an external source then the initial field is known. However, the interaction between the beam and the structure changes that field, and it is necessary to make multiple passes through the structure in order to achieve a self-consistent solution. The self-consistency can be checked by computing the net flow of energy through a closed boundary surrounding the interaction space because that should be zero in the steady state. When the electromagnetic structure has no external source, but is driven by a modulated beam, iteration is again necessary to achieve a self-consistent solution. Many computer codes have been written for the simulation of klystrons and TWTs but, almost without exception, they are proprietary and not generally available outside the organisations which produced them. An exception to this is AJ-DISK which is a 1D large-signal klystron model [11]. Simple Mathcad models for klystrons and helix TWTs can be found in Worksheets 13.3, 13.5, and 14.3.

11.8.1   Large-Signal Model of the Beam–Gap Interaction

The large-signal interaction between an electron beam and the RF field of an interaction gap can be investigated by numerical integration of the equations of motion of the electrons. To obtain accurate results it is necessary to use a three-dimensional model [46, 47]. However, the main features can be studied qualitatively using a one-dimensional disc model of the electron beam [48–51]. In Worksheet 11.6 the motion of a single wavelength of electrons is tracked in time, and the space-charge forces are calculated by assuming that the positions of the electrons are periodic in space. This assumption is not strictly valid, but it is adequate for purposes of illustration. The results of such calculations are commonly displayed in the form of an Applegate diagram as shown in Figure 11.24 (c.f. Figure 11.4). The slopes of the electron trajectories are proportional to their velocities. It is convenient to normalise the time and position by plotting $\theta = \beta_e z$ against $\phi = \omega t$ because then electrons which have velocity u_0 have unit slope. The figure shows the trajectories of 12 electrons distributed through one RF period when gap centre is at $z = 0$, the normalised effective RF voltage $M V_g / V_0 = 0.2$, and space-charge forces are neglected. The maximum accelerating field is at $\omega t = \pi/2$ and the formation of the bunch is clearly visible.

For many purposes it is convenient to use a modified Applegate diagram in which the axes are interchanged. The phase is plotted relative to that of an electron travelling at constant velocity (i.e. $\phi' = \phi - \theta$). In this diagram the trajectory of an electron whose velocity is u_0 is horizontal, while increased (reduced) velocity is indicated by
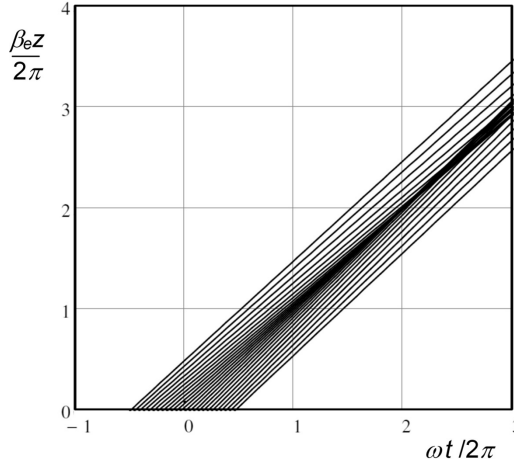
Figure 11.24: Applegate diagram for ballistic bunching of an electron beam $\left(MV_g/V_0 = 0.2\right)$.

a negative (positive) slope. Figure 11.25(a) shows a modified Applegate diagram, including space-charge effects, when $\omega_q/\omega = 0.1$ and the beam is modulated by a gap at $z = 0$ with $MV_g/V_0 = 0.1$. The relative phase of the gap voltage is $\pi/2$, so that it is accelerating when $0 < \phi' < \pi$ and retarding when $-\pi < \phi' < 0$. Under these conditions the bunching is just in the small-signal regime and the diagram shows repeated plasma oscillations whose wavelength is approximately $10\lambda_e$. These oscillations show that there is a periodic exchange of energy between the kinetic energy of the electrons and the potential energy stored in the bunch. At the plane of where the bunching is greatest the trajectories are almost parallel to the horizontal axis, indicating that the electron velocities are approximately u_0, as predicted by small-signal theory. The relative phase of the bunch is close to zero and it therefore leads the phase of the gap voltage by approximately 90°, as expected from (11.13). However the diagram is not quite symmetrical about $\phi' = 0$ because a finite input gap has been modelled and second-order effects are not negligible. In Figure 11.25(b) MV_g/V_0 has been increased to 0.2, the bunching is more pronounced, and the plane of the bunch has moved closer to the input gap. There is now a visible spread in the electron velocities at the plane of the bunch, and the subsequent trajectories are only approximately periodic. When the gap voltage is increased further (Figure 11.25(c), $MV_g/V_0 = 0.4$) the electron trajectories cross one another so that the bunch is not well-defined and there is a considerable spread of velocities.

The harmonic components of the beam current, as a function of position, can be calculated from the arrival times of the sample electrons at successive planes in z, as described above. Figure 11.26 shows how the harmonic currents vary with axial position for three different values of the effective gap voltage. In Figure 11.26(a) $\left(MV_g/V_0 = 0.02\right)$ the behaviour is very close to that predicted by small-signal theory, and the amplitude of the first current harmonic is proportional to $\sin\left(\beta_q z\right)$. It should be noted, however, that even in this case the amplitudes of the higher harmonic components are not zero. The remaining figures show that, as the effective
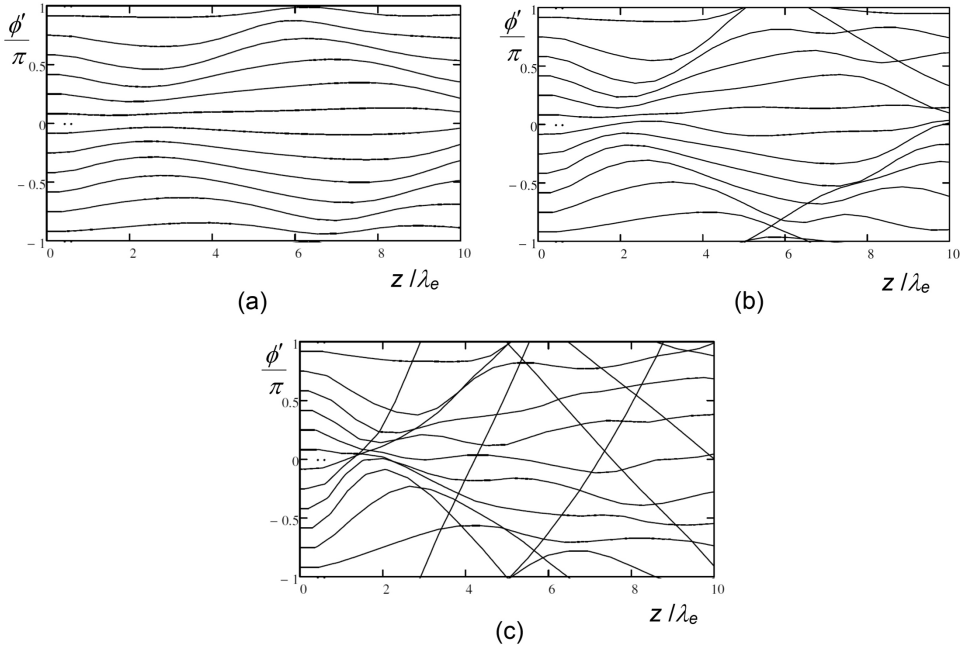
Figure 11.25: Modified Applegate diagrams for bunching of an electron beam with space-charge $\left(\omega_q / \omega = 0.1\right)$: (a) $MV_g / V_0 = 0.1$, (b) $MV_g / V_0 = 0.2$ and, (c) $MV_g / V_0 = 0.4$.
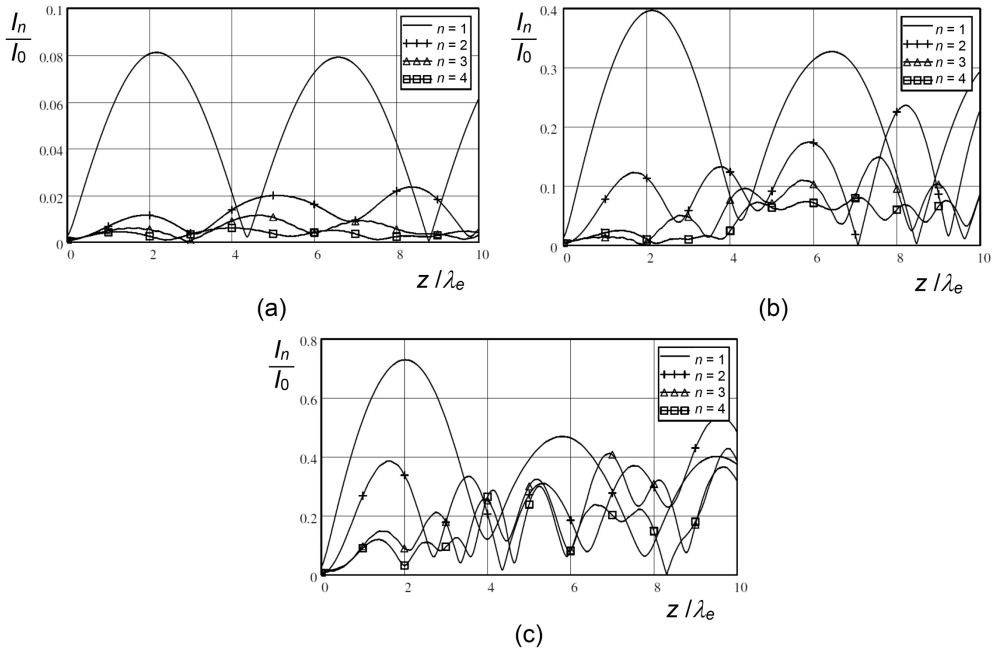
Figure 11.26: Normalised harmonic currents as a function of axial position for $\omega_q / \omega = 0.1$ and different values of the effective gap voltage: (a) $MV_g / V_0 = 0.02$, (b) $MV_g / V_0 = 0.1$ and (c) $MV_g / V_0 = 0.2$.

gap voltage is increased, the maximum of the first harmonic increases, and moves closer to the bunching gap, whilst the amplitudes of the higher harmonics increase more rapidly, so that the behaviour is increasingly non-linear.

11.8.2   Modulation of an Electron Beam by a Gap

When an electron beam passes through an interaction gap the field of the gap is generally much greater than the space-charge field. It is therefore possible to neglect the space-charge field to a first approximation when investigating the interaction. Figure 11.27(a) shows the profile of the radially averaged axial electric field acting on the electrons for a typical case $(b = 0.6a, g = a)$. The dashed vertical lines show the positions of the edges of the gap in the beam tunnel. The gap used for this illustration has a small-signal gap coupling factor $M = 0.79$ and the ratio of the gap voltage to the beam voltage is 1.52 corresponding to $MV_g/V_0 = 1.2$. The modified Applegate diagram in Figure 11.27(b) shows that at this gap voltage the electrons which lose energy (moving upwards in the figure) remain in the gap much longer than those which gain energy. They therefore experience a much greater change in
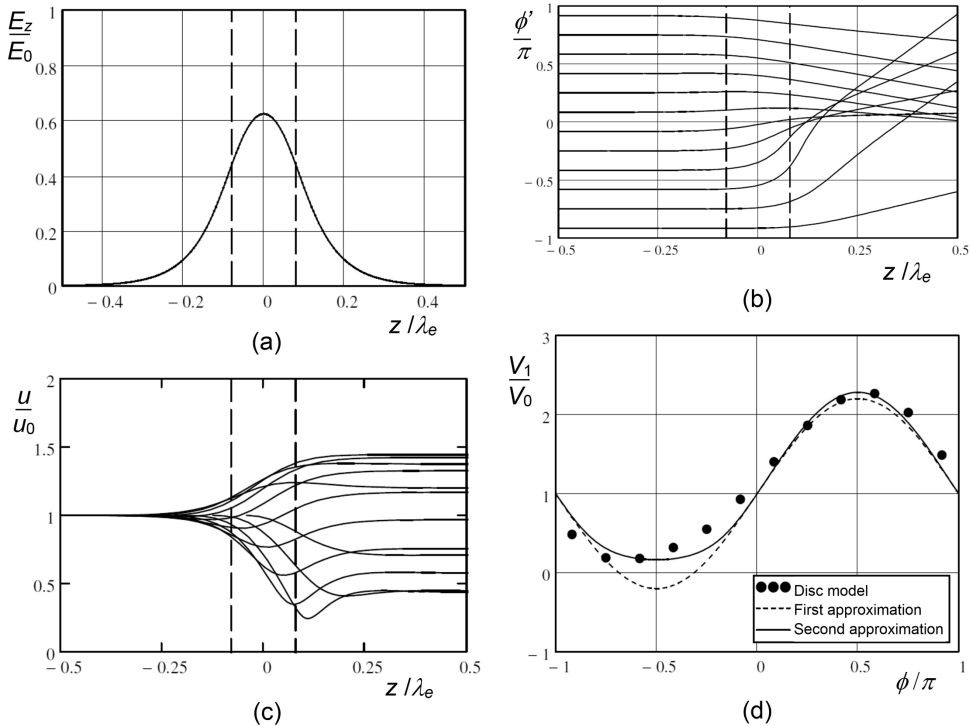
Figure 11.27: Modulation of an electron beam by a gap under large-signal conditions when space-charge is neglected: (a) variation of the axial electric field with position, (b) the modified Applegate diagram, (c) the variation of the electron velocities with position and (d) the dependence of the final electron energy on phase $\left(\beta_e a = \beta_e g = 1; \beta_e b = 0.6; MV_g/V_0 = 1.2\right)$.

the phase of the field. In extreme cases the phase of the field is reversed during the transit time of an electron. This behaviour is also illustrated in Figure 11.27(c), which shows the normalised velocities of 12 electrons as they pass through the gap. Only those electrons with the greatest final velocities are accelerated continuously, all the others stay in the gap long enough to experience a reversal of the field. The normalised velocities of the slowest and fastest electrons are approximately 0.45 and 1.45 as they leave the gap. Figure 11.27(d) shows the ratio of the final electron energy to the initial electron energy as a function of the phase at which each electron would reach the centre of the gap if the field of the gap were zero. It is evident that under these conditions there is asymmetry between those electrons which are accelerated and those which are retarded. The dotted line in the figure shows the electron energies computed using the small-signal approximation (11.8). That model gives poor results especially for the electrons which are retarded because it assumes that they pass through the gap with constant velocity. Examination of Figure 11.26(c) suggests that better results could be obtained assuming that an electron has its initial velocity (u_i) up to the centre of the gap, and its final velocity $(u_f(\phi))$ thereafter. Then the final energy is given by

$$V_f(\phi) = V_0 + \frac{1}{2}\left(M(u_i) + M(u_f(\phi))\right)V_g \cos(\phi), \qquad (11.168)$$

which can be solved by iteration to find V_f. The results of this second approximation are shown by the solid line in Figure 11.27(d). This equation predicts the velocities of the fastest and slowest electrons quite well but fails to reproduce the phase shift calculated by the disc model. The accuracy of (11.168) increases as the gap voltage is reduced.

The beam loading conductance can be calculated by computing the change in the kinetic energy of an unmodulated beam produced by the field of the gap [52] (see Worksheet 11.3). Using this method it is found that the beam loading conductance is given by (11.23) up to $MV_g/V_0 = 0.6$. The results are only slightly affected by the inclusion of space-charge forces, as has been shown by the similarity between the results of ballistic analysis (11.22) and space-charge wave theory (11.99).

11.8.3   Current Induced in a Passive Gap by a Modulated Beam

We have seen that, when a modulated beam passes through an interaction gap, a current is induced in the gap. Under small-signal conditions the induced current can be calculated by assuming that the magnitude of the RF current in the beam is constant within the gap. The gap voltage is the product of the induced current and the impedance presented to the gap by beam loading and by the external circuit. Under large-signal conditions this assumption is no longer valid and the current must be calculated by finding self-consistent values of the induced current and the gap voltage [53]. It is normally the case that the external circuit is passive and therefore the time average of the power in the beam can only be reduced by the interaction.
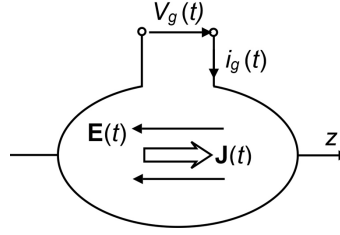
Figure 11.28: Field and circuit representations of the interaction between a modulated electron beam and a gap in the wall of the surrounding drift tube.

Figure 11.28 shows the interaction between the convection current $(\mathbf{J}(t))$ and the electric field of a gap in the wall of the surrounding drift tube $(\mathbf{E}(t))$. The sign convention used for the gap voltage (V_g) and the induced current (i_g) is the same as in Figure 11.15. The induced current is in antiphase with the conventional current of the electrons and, therefore, in phase with the convection current. The instantaneous power transferred to the beam from the gap is given by [54]

$$i_g(t)V_g(t) = -\oiiint \mathbf{J}(t) \cdot \mathbf{E}(t)\, dv. \tag{11.169}$$

In a one-dimensional model this becomes

$$i_g(t)V_g(t) = -\int_{z_1}^{z_2} I(z,t) E_z(z,t)\, dz, \tag{11.170}$$

where I is the convection current in the beam, E_z is the axial component of the electric field averaged over the cross-section of the beam, and the integral is taken over the region in which $E_z \neq 0$. Now the electric field is proportional to the gap voltage so we may write

$$E_z(z,t) = -\frac{V_g(t)}{g} f(z). \tag{11.171}$$

Substituting in (11.170) gives the instantaneous induced current

$$i_g(t) = \frac{1}{g}\int_{z_1}^{z_2} I(z,t) f(z)\, dz. \tag{11.172}$$

But the currents are harmonic in time so that they can be expressed as a Fourier series. Then, equating the harmonics on the two sides of the equation gives the complex amplitude of the nth harmonic of the induced current in terms of the complex amplitude of the nth harmonic of the convection current

$$i_{gn} = \frac{1}{g}\int_{z_1}^{z_2} f(z) I_n(z)\, dz. \tag{11.173}$$
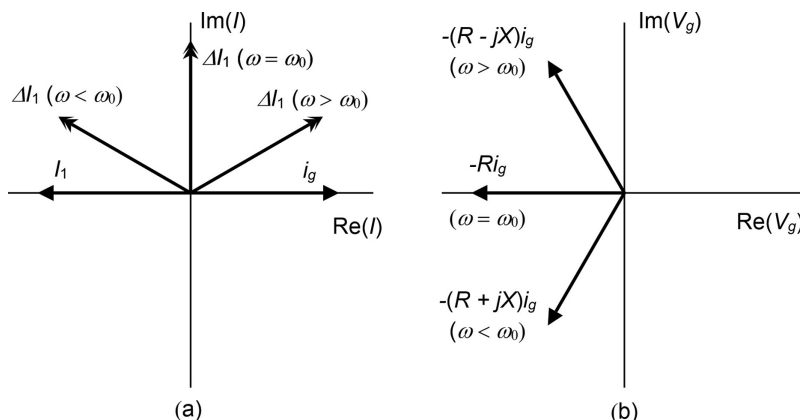
Figure 11.29: Phasor diagrams for the interaction between a modulated electron beam and a passive cavity resonator: (a) current phasors, and (b) gap voltage phasors.

In small-signal theory it is assumed that the induced current depends only on the modulation of the beam entering the gap. The gap current associated with the re-modulation of the beam by the field of the gap is represented by the beam loading admittance. In large-signal theory the gap current computed by numerical evaluation of (11.173) includes both effects. The cavity resonator outside the gap presents impedance Z_g to the current. That impedance is negligible except close to the signal harmonic to which it has been tuned. Thus the gap voltage is

$$V_g(t) = -Z_g(n\omega)i_{gn}\exp(jn\omega t). \qquad (11.174)$$

Mutually consistent values of the complex induced current and the complex gap voltage can be found by iteration.

The impedance of the gap is real when the signal frequency harmonic is equal to the resonant frequency of the cavity (ω_0), it has positive (inductive) reactance when $\omega < \omega_0$, and negative (capacitive) reactance when $\omega > \omega_0$ (see Section 3.2.1). The phase angle of the impedance varies from $\pi/2$ to $-\pi/2$ as the signal frequency increases (see Figure 3.4). The interaction between the modulated beam and the cavity can be understood by considering the phasor diagrams in Figure 11.29. The phases are referred to the instant at which the centre of the bunch lies at the centre of the gap. The phase of the RF convection current (I_1) is $-\pi$, and the phase of the induced current in the gap is zero, as shown in Figure 11.29(a). Figure 11.29(b) shows the positions of the phasor representing the gap voltage at three values of the modulation frequency (ω). The phase of the current modulation added to the beam by the gap (ΔI_1) lags the phase of the gap voltage by approximately $\pi/2$. The phases of the additional modulation for the three frequencies are shown in Figure 11.29(a). In many cases the additional modulation is the dominant modulation downstream from the gap. The special case of the output gap of an IOT or a klystron is considered in more detail in the next section. The application to electron

bunching in klystrons, by cavities whose resonant frequency differs from the modulation frequency, is considered in section 13.3.3.

11.8.4   Power Transfer in an Output Gap

The resonant frequency of the output cavity of an IOT or a klystron is normally very close to the frequency of modulation. The impedance presented to the induced current in the gap is therefore largely resistive. It is dominated by the external load in order to minimise the loss of RF power in the cavity.

To investigate how the power transferred to the output gap depends upon the current waveform, and upon the properties of the gap, we will consider the set of idealised current waveforms. These have been constructed to give the greatest possible value of I_1/I_0 for a given number of harmonics, subject to the condition that the total current can never be negative. They are defined by

$$I(\phi,n) = 2\pi(1+\cos\phi)^n ,$$

(11.175)

where n is an integer. These functions are maximally flat at the zeroes of current. The DC current is given by

$$I_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} I(\phi,n)\,d\phi$$

(11.176)

and the harmonic currents are

$$I_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} I(\phi,n)\cos(m\phi)\,d\phi.$$

(11.177)

Table 11.2 shows the peak currents, and the amplitudes of the harmonic currents, of these idealised waveforms normalised to the DC current. Figure 11.30 shows the first four waveforms of the set. As the number of harmonics is increased the amplitude of the first harmonic increases and the width of the bunch decreases. For $n \geq 4$ very nearly all the electrons are grouped within a phase range of 180°.

Table 11.2: Peak currents and harmonic current components of idealised bunching waveforms

Number of harmonics	1	2	3	4	5	6
I_{pk}/I_0	2.00	2.67	3.20	3.66	4.06	4.43
I_1/I_0	1.000	1.333	1.500	1.600	1.667	1.714
I_2/I_0		0.333	0.600	0.800	0.952	1.071
I_3/I_0			0.100	0.229	0.357	0.476
I_4/I_0				0.029	0.079	0.143
I_5/I_0					0.008	0.026
I_6/I_0						0.002
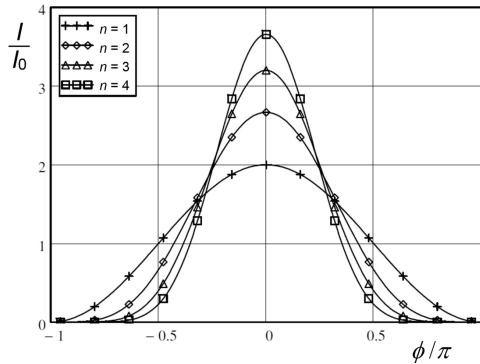
Figure 11.30: Idealised current waveforms having the maximum possible amplitude of the first harmonic for different numbers of harmonic components.

The range of usefulness of these equations can be investigated by comparing the results with those obtained directly using the disc model, without space-charge, so that the effects of potential energy are not considered. It is assumed that the electrons all enter the gap with velocity u_0. This problem was studied by Hechtel [51] who showed that the optimum conversion efficiency decreased with increases of the normalised tunnel radius $(\gamma_e a)$ and the normalised gap length $(\beta_e g)$. Detailed results were only obtained for $I_1/I_0 = 1.44$ and the waveform assumed was not given. The same problem can be studied using Worksheet 11.6 with the waveforms given by (11.175) and very similar results are obtained. Under small-signal conditions the maximum power delivered to the gap is given by

$$P_{gap} = \frac{1}{2} M I_1 V_g. \tag{11.178}$$

If the RF beam current and the gap voltage are fixed then the power is proportional to the gap coupling factor (M) so that it decreases as M decreases. That is the case when the normalised tunnel radius $(\gamma_e a)$ and the normalised gap length $(\beta_e g)$ increase, exactly as found by Hechtel.

The physical origin of the beam loading conductance of an unmodulated beam is the imbalance between the powers transferred to electrons which are accelerated or retarded. When the input beam is modulated, the proportions of electrons in these two classes change. Then the time-average transfer of power between the beam and the gap depends upon the magnitude of the modulation, and upon the phase difference between the bunch centre and the field of the gap. The power delivered to the gap is the time-average of the product of the current and the effective gap voltage. Making use of (11.168) this can be written approximately as

$$P_{gap} = -\frac{1}{4\pi} \int_{-\pi}^{\pi} I(\phi)\big(M(u_0) + M(u_f(\phi))\big) V_g \cos(\phi + \Phi)\, d\phi, \tag{11.179}$$
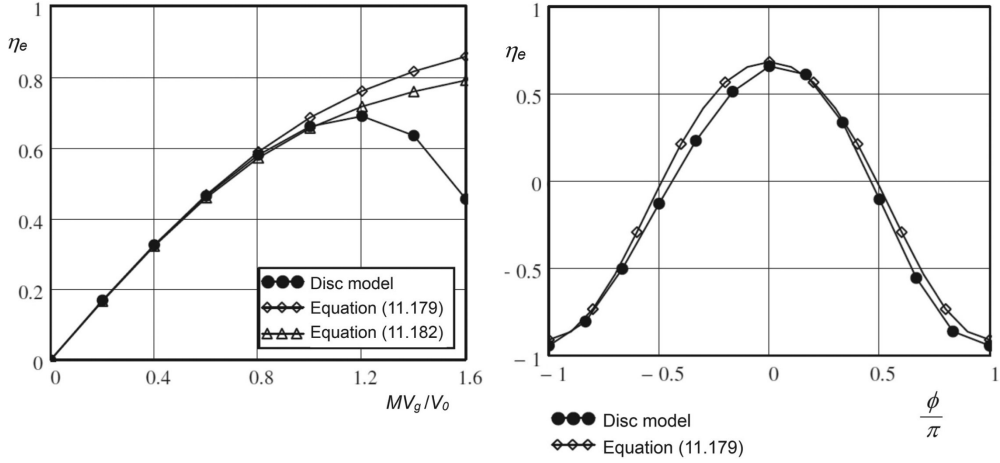
Figure 11.31: Efficiency of the extraction of energy from an ideally bunched electron beam comparing results from the disc model and equations (11.179) and (11.182): (a) with $\Phi = 180°$, and (b) with $MV_g/V_0 = 1.0$ ($\beta_e a = \beta_e g = 1; \beta_e b = 0.6$).

where $\phi = \omega t$, $I(\phi)$ is the beam current at the centre of the gap, and Φ is the phase of the gap voltage when the bunch is at the centre of the gap. This equation is found to give good agreement with the power computed using the disc model when the bunch length is $180°$ or less. Figure 11.31(a) shows a comparison between the electronic efficiency calculated using the disk model, and that computed using (11.179), for extraction of energy from an idealised bunch with $n = 6$. The disk model shows that the transfer of energy is greatest when $MV_g/V_0 = 1.2$. The electronic efficiency decreases for greater gap voltages because some electrons remain in the gap for a long time, so that they experience a large change of phase and are re-accelerated, taking energy from the gap. For the ideally bunched beam, and the one-dimensional model without space-charge, considered here, this effect that sets a limit on the electronic efficiency of the interaction. The energy of the slowest electron leaving the gap is around $V_{min} = 0.15\,V_0$ when $g = a$. This figure decreases slowly as the length of the gap is reduced.

Figure 11.31(b) compares the results of the disc model with those from (11.179) for energy transfer from the bunch to the field when $MV_g/V_0 = 1.0$ and the phase of the bunch relative to the field of the gap is varied. It is found that these results are almost independent of the normalised length of the gap.

By analogy with small-signal theory we can define an effective large-signal gap coupling factor by

$$M_{eff} = \frac{2P_{gap}}{I_1 V_g},\tag{11.180}$$

where I_1 is the first harmonic of the beam current under large-signal conditions. Then the efficiency of the transfer of power from the beam to the gap is

$$\eta_e = \frac{M_{eff} I_1 V_g}{2 I_0 V_0} = \frac{1}{2} \cdot \frac{M_{eff}}{M} \cdot \frac{I_1}{I_0} \cdot \frac{M V_g}{V_0}. \tag{11.181}$$

The value of M_{eff} can be calculated using the disk model or, approximately, using (11.179). When the beam is tightly bunched the majority of the electrons cross the gap centre close to the phase of the maximum retarding field. Then the exit velocities of the electrons are approximately $u_{min} = u_f(\pi)$ and a useful approximation to M_{eff} is

$$M_{eff} \approx \frac{1}{2}\left(M(u_0) + M(u_{min})\right). \tag{11.182}$$

The electronic efficiency calculated using (11.182) is close to that given by the disc model up to $M V_g / V_0 = 1.0$, as shown in Figure 11.30(a) when $V_0 = 25$ kV so that relativistic effects are not important.

The effective coupling factor that gives maximum efficiency was calculated using the disc model, without space-charge, for I_1 / I_0 in the range 1.33 to 1.67; $\gamma_e a$ in the range 0.2 to 3.0; and $\beta_e g$ in the range 0.1 to 5.0. When the results were plotted against the small-signal coupling factor (M) the result was the universal curve, shown in Figure 11.32. The symbols represent different combinations of the parameters. The curve is closely fitted by the empirical approximation

$$M_{eff} = \frac{1}{2}\left(M + M^6\right). \tag{11.183}$$

The corresponding values of V_g / V_0 were about 1.5 when $M = 0.8$. They increased at varying rates, depending upon the parameters chosen, as M decreased. It is important to remember that this result was obtained using a one-dimensional disc model in which space-charge was ignored and a uniform electric field in the gap was assumed. The value of the effective gap coupling factor for understanding the large-signal operation of IOTs and klystrons, and the effects when the bunches are
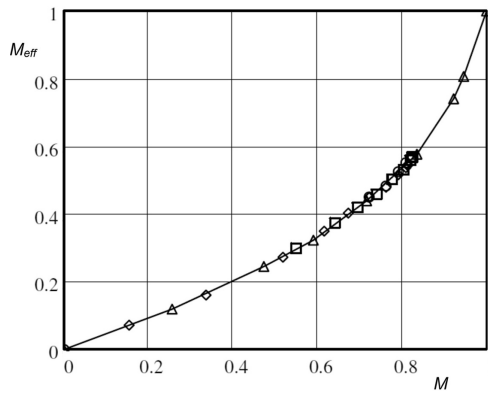
Figure 11.32: Dependence of the effective gap coupling factor giving maximum efficiency upon the small-signal coupling factor.

not ideal, are explored in Chapters 12 and 13. The effects of space-charge on the maximum efficiency of energy transfer to the gap are discussed in Section 13.3.6.

References

[1] W. Hahn, 'Small signal theory of velocity modulated electron beams', *Gen. Elec. Rev,* vol. 42, pp. 258–270, 1939.

[2] S. Ramo, 'The electronic-wave theory of velocity-modulation tubes', *Proceedings of the IRE,* vol. 27, pp. 757–763, 1939.

[3] A. H. W. Beck, *Space-Charge Waves and Slow Electromagnetic Waves.* London: Pergamon Press, 1958.

[4] R. E. Trotman, *Longitudinal Space-Charge Waves.* London: Chapman & Hall, 1966.

[5] M. T. Vlaardingerbroek and K. R. U. Weimer, 'Beam-plasma amplifier tubes', *Phillips Research Review,* vol. 27, pp. 275–284, 1966.

[6] K. Spangenburg, *Vacuum Tubes.* New York: McGraw-Hill, 1948.

[7] D. L. Webster, 'Cathode-ray bunching', *Journal of Applied Physics,* vol. 10, pp. 501–508, 1939.

[8] B. B. van Iperen and H. J. C. A. Nunnink, 'Harmonics in velocity-modulated cylindrical electron beams', *Philips Research Reports,* vol. 20, pp. 432–461, 1965.

[9] G. M. Branch, Jr., 'Electron beam coupling in interaction gaps of cylindrical symmetry', *IRE Transactions on Electron Devices,* vol. 8, pp. 193–207, 1961.

[10] E. J. Craig, 'The beam-loading admittance of gridless klystron gaps', *IEEE Transactions on Electron Devices,* vol. 14, pp. 273–278, 1967.

[11] G. Caryotakis, 'Klystrons', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimeter-Wave Power Electronics.* Piscataway, NJ: IEEE Press, pp. 107–170, 2005.

[12] G. R. Brewer, 'Some effects of magnetic field strength on space-charge-wave propagation', *Proceedings of the IRE,* vol. 44, pp. 896–903, 1956.

[13] L. A. MacKenzie, 'Space charge waves in finite magnetic fields', in *The Fourth International Conference on Microwave Tubes*, The Hague, The Netherlands, pp. 663–668, 1962.

[14] V. Srivastava and R. G. Carter, 'Effect of boundaries on the space charge potential in coupled cavity travelling wave tubes', *IEE Proceedings I: Solid-State and Electron Devices,* vol. 133, pp. 185–188, 1986.

[15] W. H. Louisell and J. R. Pierce, 'Power flow in electron beam devices', *Proceedings of the IRE,* vol. 43, pp. 425–427, 1955.

[16] R. L. Gunshor, 'Decay of space-charge waves on electron beams', *Journal of Applied Physics,* vol. 37, pp. 1904–1911, 1966.

[17] M. Caulton, 'Damping of waves in electron beams', *Journal of Applied Physics,* vol. 38, pp. 1839–1855, 1967.

[18] M. Caulton *et al.*, 'Experimental evidence of Landau damping in electron beams', *Journal of Applied Physics,* vol. 33, pp. 800–803, 1962.

[19] G. M. Branch and T. G. Mihran, 'Plasma frequency reduction factors in electron beams', *IRE Transactions on Electron Devices,* vol. 2, pp. 3–11, 1955.

[20] G. M. Branch, Jr. *et al.*, 'Space-charge wavelengths in electron beams', *IEEE Transactions on Electron Devices,* vol. 14, pp. 350–357, 1967.

[21] W. H. Louisell, *Coupled Mode and Parametric Electronics.* New York: Wiley, 1960.

[22] A. Bers, 'Linear space-charge theory of gap interaction between an electron beam and electromagnetic fields', in *Microwave and Optical Generation and Amplification*, Munich, pp. 53–60, 1960.

[23] C.-L. Huang *et al.*, 'AC-space-charge effects on gap coupling coefficient of a klystron cavity', *IEEE Transactions on Plasma Science,* vol. 40, pp. 828–834, 2012.

[24] E. J. Craig, 'Relativistic beam-loading admittance', *IEEE Transactions on Electron Devices*, vol. 16, pp. 139–139, 1969.

[25] G. Faillon, 'Klystrons de puissance à large bande', *Revue Technique Thomson-CSF,* vol. 8, p. 139, June 1976.

[26] J. R. Pierce, *Traveling-Wave Tubes*. Princeton, NJ: D. van Nostrand, 1950.

[27] J. R. Pierce, 'Theory of the beam-type traveling-wave tube', *Proceedings of the IRE,* vol. 35, pp. 111–123, 1947.

[28] C. K. Birdsall and G. R. Brewer, 'Traveling wave tube characteristics for finite values of C', *Transactions of the IRE Professional Group on Electron Devices,* vol. 1, pp. 1–11, 1954.

[29] G. R. Brewer and C. K. Birdsall, 'Traveling-wave tube propagation constants', *IRE Transactions on Electron Devices,* vol. 4, pp. 140–144, 1957.

[30] G. S. Kino *et al.*, 'Small-signal and large-signal theories for the coupled-cavity TWT', in *6th International Conference on Microwave and Optical Generation and Amplification*, Cambridge, UK, pp. 49–53, 1966.

[31] D. L. Bobroff, 'The buildup of oscillations in an electron beam backward-wave oscillator', *IEEE Transactions on Electron Devices,* vol. 12, pp. 307–312, 1965.

[32] B. O. Nilsson and C. E. Hagstrom, 'A two wave theory of traveling-wave tubes and backward-wave oscillations', *IEEE Transactions on Electron Devices,* vol. 22, pp. 869–880, 1975.

[33] K. Minami *et al.*, 'Analysis of starting currents in a backward wave oscillator with finite structure length', *Journal of the Physics Society Japan,* vol. 61, pp. 3566–3575, 1992.

[34] H. R. Johnson, 'Backward-wave oscillators', *Proceedings of the IRE,* vol. 43, pp. 684–697, 1955.

[35] S. K. Datta *et al.*, 'Nonlinear Eulerian hydrodynamical analysis of helix traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 45, pp. 2055–2062, 1998.

[36] J. G. Wöhlbier *et al.*, 'The multifrequency spectral Eulerian (MUSE) model of a traveling wave tube', *IEEE Transactions on Plasma Science,* vol. 30, pp. 1063–1075, 2002.

[37] C. C. Motta, 'A large-signal analysis of a ring-bar TWT', in *IEEE 34th International Conference on Plasma Science*, pp. 875–875, 2007.

[38] T. M. Antonsen, Jr. *et al.*, 'Advances in modeling and simulation of vacuum electronic devices', *Proceedings of the IEEE,* vol. 87, pp. 804–839, 1999.

[39] B. E. Carlsten and P. J. Tallerico, 'Self-consistent klystron simulations', *Nuclear Science, IEEE Transactions on Nuclear Science,* vol. 32, pp. 2837–2839, 1985.

[40] J. R. M. Vaughan, 'Calculation of coupled-cavity TWT performance', *IEEE Transactions on Electron Devices,* vol. 22, pp. 880–890, 1975.

[41] V. Srivastava and R. G. Carter, 'A fast large-signal model for coupled-cavity TWTs', *IEEE Transactions on Electron Devices,* vol. 35, pp. 2068–2076, November 1988.

[42] D. Chernin *et al.*, 'A three-dimensional multifrequency large signal model for helix traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 48, pp. 3–11, 2001.

[43] T. Shintake, 'FCI – field charge interaction program for high power klystron simulations', in *Proceedings of the 1989 IEEE Particle Accelerator Conference*, vol. 1, pp. 94–96, 1989.

[44] J. R. Hechtel, 'The effect of potential beam energy on the performance of linear beam devices', *IEEE Transactions on Electron Devices,* vol. 17, pp. 999–1009, 1970.

[45] C. J. Edgcombe, 'Increased efficiency for klystron amplifiers', *IEE Journal on Solid-State and Electron Devices,* vol. 1, pp. 62–68, 1977.

[46] T. G. Mihran, 'The effect of space charge on bunching in a two-cavity klystron', *IRE Transactions on Electron Devices,* vol. 6, pp. 54–64, 1959.

[47] H. G. Kosmahl and L. U. Albers, 'Three-dimensional evaluation of energy extraction in output cavities of klystron amplifiers', *IEEE Transactions on Electron Devices,* vol. 20, pp. 883–890, 1973.

[48] S. E. Webber, 'Ballistic analysis of a two-cavity finite beam klystron', *IRE Transactions on Electron Devices,* vol. 5, pp. 98–108, 1958.

[49] S. E. Webber, 'Large signal bunching of electron beams by standing-wave and traveling-wave systems', *IRE Transactions on Electron Devices,* vol. 6, pp. 365–372, 1959.

[50] S. E. Webber, 'Some calculations on the large signal energy exchange mechanisms in linear beam tubes', *IRE Transactions on Electron Devices,* vol. 7, pp. 154–162, 1960.

[51] J. R. Hechtel, 'DC-to-RF energy conversion in ungridded klystron gaps', *IEEE Transactions on Electron Devices,* vol. 16, pp. 212–217, 1969.

[52] J. R. M. Vaughan, 'A model for the klystron cavity gap', *IEEE Transactions on Electron Devices,* vol. 32, pp. 2482–2484, 1985.

[53] T. Kageyama, 'A large-signal analysis of broad-band klystrons with design applications', *IEEE Transactions on Electron Devices,* vol. ED-24, pp. 3–12, 1977.

[54] B. E. Carlsten *et al.*, 'Accuracy of the equivalent circuit model using a fixed beam impedance for klystron gain cavities', *IEEE Transactions on Plasma Science,* vol. 26, pp. 1745–1749, 1998.

12   Gridded Tubes

12.1   Introduction

Gridded tubes include triodes, tetrodes, and inductive output tubes (IOTs). In these devices the current emitted by a thermionic cathode is modulated by the voltage applied to a control grid as discussed in Chapter 6. The time varying current delivers RF power to an output circuit. The technology of triodes and tetrodes is mature and most of the literature dates from 1960 or earlier. Up to that time small valves were still the dominant type of active electronic device, and much of the literature concentrates on them. The subject was fully discussed in a number of books which contain much greater detail than is possible here [1–5]. The state of the art of high power triodes and tetrodes was reviewed in [6]. The principle of the IOT was known for many years, but this type of tube did not become commercially important until the development of pyrolytic graphite grids [7, 8]. Triodes and tetrodes are now mainly used in high power oscillators and amplifiers up to around 200 MHz. Inductive output tubes are high power amplifiers in the range 200 MHz to 1.3 GHz [9, 10].

12.1.1   Gridded Tube Amplifiers

Before we proceed to examine the properties of gridded tubes in detail it is necessary to understand how they are used in power amplifiers. Figure 12.1 shows the circuit of a common cathode tetrode amplifier with a tuned anode (tank) circuit. In UHF amplifiers this circuit takes the form of a cavity resonator, but at lower frequencies lumped components are used. At very low frequencies a resistive anode load may be used, but this is unsatisfactory in the VHF band and above because of the effects of parasitic capacitance. At the resonant frequency the load in the anode circuit comprises the shunt resistance of the resonator in parallel with the load resistance. If the load impedance has a reactive component then it detunes the resonator and can be compensated for. The DC electrode potentials are maintained by the power supplies shown and the capacitors provide RF coupling and decoupling.

An alternative circuit which is frequently used is the common grid circuit shown in Figure 12.2. This circuit is easier to construct, using tubes with coaxial electrodes, and there is better isolation between the input and the output circuits because the RF voltage on both grids is zero. The other important difference is that in this
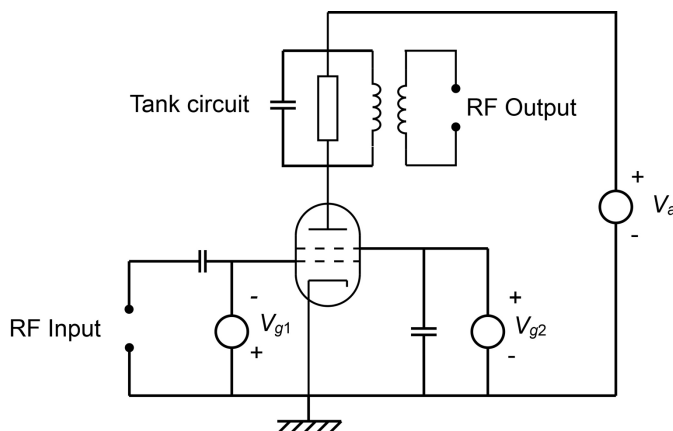
Figure 12.1: Common cathode tetrode amplifier.
(copyright 1992, CERN, reproduced, with permission, from [11]).
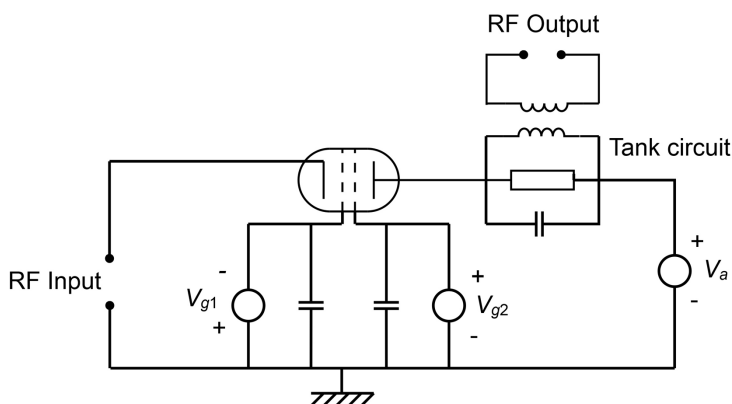
Figure 12.2: Common grid tetrode amplifier
(copyright 1992, CERN, reproduced, with permission, from [11]).

circuit the full anode current flows in the input circuit with the result that the input impedance and the gain are both lower than for the grounded cathode circuit.

12.1.2   Classes of Amplification

Figure 12.3 shows typical characteristic curves of a tetrode. In the absence of an RF input the currents flowing to the electrodes are determined solely by the DC voltage applied to them. The tube is then said to be at its quiescent (Q) point. Tuned amplifiers are normally operated in one of three modes known as classes A, B, and C as described below. The difference between these lies in the choice of the Q point. The Q point marked in Figure 12.3 lies on the curve of zero anode current at an anode voltage of 10 kV. This is achieved by setting the control grid voltage to –240 V. The screen grid is at RF ground and a DC potential of 900V.

$U_{G1} = f(U_A)$ Parameter $= I_A$ ——————
$U_{G2} = 900$ V   Parameter $= I_{G2}$ —·—·—·—
                    Parameter $= I_{G1}$ — — — — —
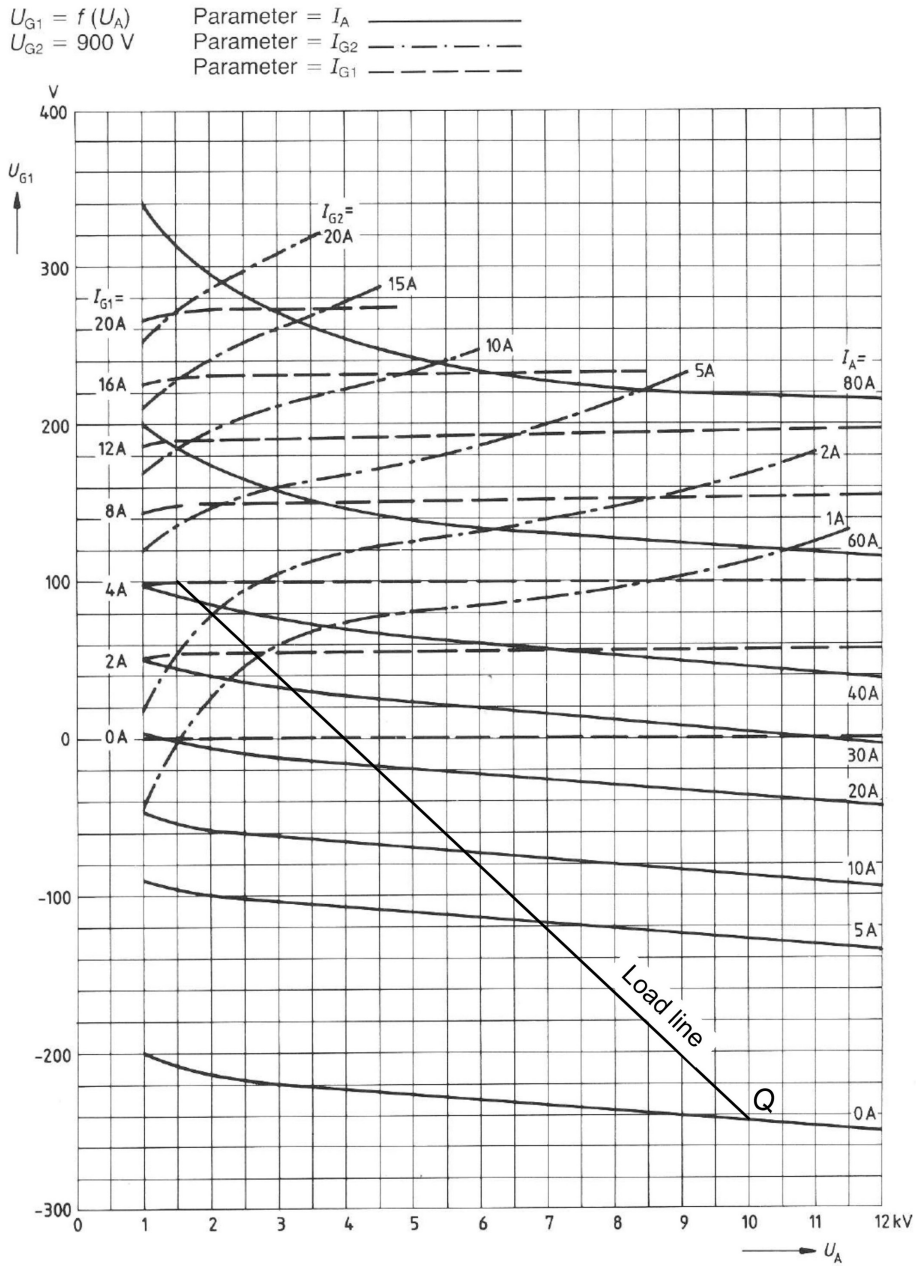
Figure 12.3: RS 2058 tetrode characteristic curves with load line (copyright 1986, Siemens, reproduced, with permission of Thales Electron Devices, from [12]).

When an RF input voltage is applied between the cathode and the control grid, the RF cathode current is in phase with the RF grid voltage. If the anode load is tuned to resonance then it presents a pure resistance to the fundamental component of the RF current. The tank circuit is not resonant at the higher harmonics of the anode current. Thus, the RF voltage on the anode is sinusoidal and in antiphase with the control grid voltage. The tube operates on a straight load line passing through the Q point whose slope is determined by the anode load resistance. A typical load line is shown in Figure 12.3. When the instantaneous control grid voltage is 100 V the anode current is a little over 40 A and the anode voltage is 1.5 kV.

The operation of a gridded-tube power amplifier can be explored using a simple theory in which it is assumed that the tube obeys the 3/2 power law, and the grid currents are negligible. The analysis is presented here in terms of a triode, but tetrode amplifiers show the same behaviour. Let the control grid voltage be

$$V_g(t) = V_{go} + V_{g1}\cos(\omega t),\tag{12.1}$$

where ω is the frequency of operation. The anode current is given by (6.1),

$$I_a(\theta) = K\left(V_{go} + V_{g1}\cos(\theta) + \frac{V_a}{\mu}\right)^{\frac{3}{2}},\tag{12.2}$$

when the term in brackets is positive. The conduction angle α is defined as the range of θ for which the anode current is non-zero. Thus from (12.2)

$$V_{go} + \frac{V_a}{\mu} = -V_{g1}\cos\left(\frac{\alpha}{2}\right).\tag{12.3}$$

Substituting into (12.2) gives

$$I_a(\theta) = KV_{g1}^{\frac{3}{2}}\left(\cos(\theta) - \cos(\alpha/2)\right)^{\frac{3}{2}}.\tag{12.4}$$

The peak anode current is

$$I_{pk} = KV_{g1}^{\frac{3}{2}}\left(1 - \cos(\alpha/2)\right)^{\frac{3}{2}},\tag{12.5}$$

so that the anode current may be written

$$I_a(\theta) = I_{pk}\left(\frac{\cos(\theta) - \cos(\alpha/2)}{1 - \cos(\alpha/2)}\right)^{\frac{3}{2}}.\tag{12.6}$$

Figure 12.4 shows normalised anode current waveforms for some representative conduction angles. The classes of amplification are defined in terms of the conduction angle as shown in Table 12.1.

Using Fourier analysis, the anode current can be expressed as

$$I_a(t) = I_{a0} + I_{a1}\cos(\omega t) + \cdots\tag{12.7}$$

Table 12.1: Power amplifier classes

Class	Conduction angle (degrees)
A	*360*
AB	*180 to 360*
B	*180*
C	*<180*
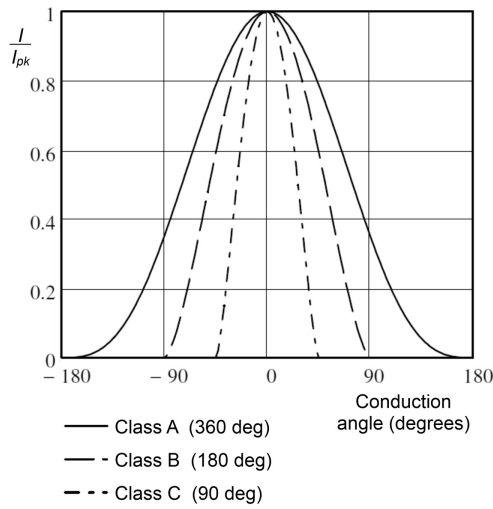
Figure 12.4: Normalised anode current waveforms for representative conduction angles.

If it is assumed that the output circuit is resonant at the signal frequency, the anode voltage can be written

$$V_a = V_{a0} - V_{a1} \cos(\omega t), \tag{12.8}$$

because V_{a1} is in antiphase with I_{a1}. The DC input power is

$$P_0 = I_{a0} V_{a0} \tag{12.9}$$

and the RF output power is

$$P_2 = \frac{1}{2} I_{a1} V_{a1}. \tag{12.10}$$

Then the electronic efficiency is

$$\eta_e = \frac{P_2}{P_0} = \frac{I_{a1} V_{a1}}{2 I_{a0} V_{a0}}. \tag{12.11}$$
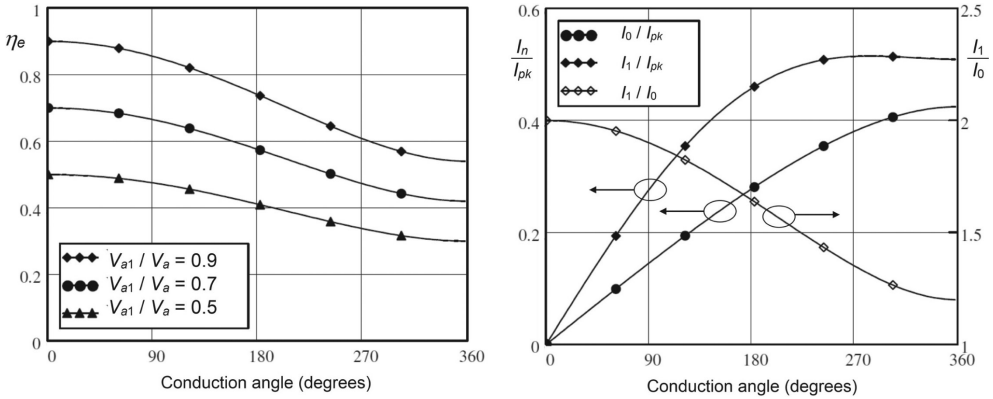
Figure 12.5: (a) Variation of electronic efficiency with conduction angle, and (b) variation of DC and RF current with conduction angle.

Because the amplifier is a power amplifier the amplitudes of the RF voltage and current should be as great as possible. They are limited only by the need to ensure that the anode voltage is always greater than the grid voltage in a triode, or the screen grid voltage in a tetrode, and that the tube ratings are not exceeded. Figure 12.5(a) shows the efficiency as a function of the conduction angle and of V_{a1}/V_{a0}. It can be seen that the efficiency increases as V_{a1}/V_{a0} increases and as the conduction angle decreases. The reason for this can be seen in Figure 12.5(b), which shows the variation of the DC and RF currents with conduction angle. The peak current is limited by the ratings of the tube. Therefore the DC current, the RF current, and the power output of the amplifier, decrease as the conduction angle decreases. However, the DC current decreases more rapidly than the RF current so that the efficiency increases for a given ratio V_{a1}/V_{a0} (see Worksheet 12.1).

The input power of a common cathode amplifier is

$$P_1 = \frac{1}{2}V_{g1}I_{g1},\tag{12.12}$$

where I_{g1} is the RF current flowing to the control grid. For a common grid amplifier it is

$$P_1 = \frac{1}{2}V_{g1}\left(I_{g1}+I_{a1}\right) \approx \frac{1}{2}V_{g1}I_{a1}.\tag{12.13}$$

The grid current is zero except when the grid is positive so $I_{g1} \ll I_{a1}$ and the grid current can be ignored in the first approximation. The gain of the amplifier in decibels

$$G_{dB} = 10\log\left(\frac{P_2}{P_1}\right)\text{dB}.\tag{12.14}$$

Thus the gain of a common cathode amplifier is always greater than that of a common grid amplifier having the same output power. As the conduction angle

decreases, the amplitude of the RF grid voltage must be increased to maintain the same peak anode current. Therefore the input power increases and the gain decreases.

The power dissipated on the anode is the difference between the DC input power and the RF output power. Therefore it decreases, as the conduction angle is reduced, at constant RF output power. In many cases the performance of a tube is limited by grid temperature raised by radiation from the cathode, electron bombardment, and RF power dissipation [13]. The RF power dissipation is proportional to the square of the RF anode current. It is therefore possible to increase the peak anode current, while maintaining constant RF anode current, as the conduction angle decreases. However, grid dissipation caused by electron bombardment increases as the peak anode current increases, and this sets a limit to the peak anode current. For many purposes class B operation offers a good compromise between high efficiency and high gain.

Two tubes driven in anti-phase can be connected to the same output circuit so that each conducts in turn to form a push-pull amplifier. The tubes are operated approximately in class B but the conduction angle is increased slightly to make the output waveform more nearly sinusoidal (reduction of 'cross-over distortion'). Further information on this and other types of power amplifier can be found in [13–16].

12.2   Triodes

Figure 12.6 shows a schematic view of a high power triode. The cathode, grid, and anode are usually in the form of concentric cylinders with the anode on the outside for ease of cooling. Triodes for use at microwave frequencies have planar geometry. The cathode may be an indirectly heated oxide cathode, or be made of thoriated
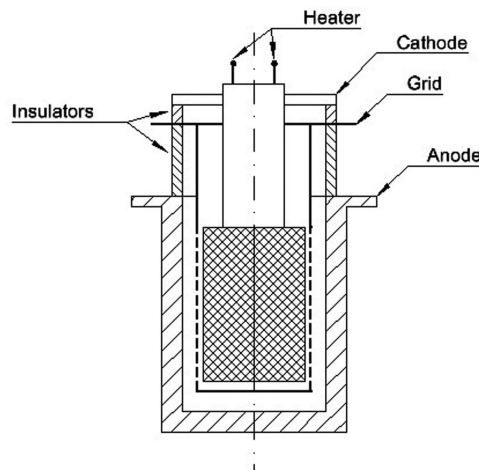
Figure 12.6: Arrangement of a high power triode.
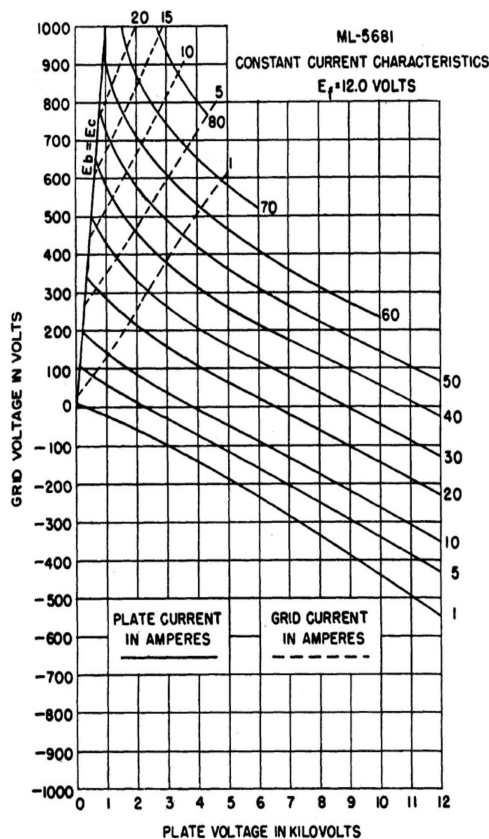
Figure 12.7: Characteristic curves of the ML-5681 high power triode (copyright 1952, IEEE, reproduced, with permission, from [19]).

tungsten wire which is heated directly by the passage of current through it. The grid is formed from metallic wire or sheet, or from pyrolytic graphite. The anode is made of oxygen-free high conductivity copper and cooled by forced convection of air or water or by vapour phase cooling (see Chapter 10). The insulators are usually alumina ceramics, though glass was used in older tubes. Further details of the construction of these tubes can be found in [6, 13, 17, 18].

Figure 12.7 shows the constant current characteristic curves of a typical high-power triode [19]. As long as the grid potential is negative with respect to the cathode no electrons are collected by the grid. However, when the grid potential is positive some of the electrons strike the grid and may liberate secondary electrons. The grid current shown is therefore a balance between the primary and secondary electron currents. The anode voltage must always exceed the grid voltage to avoid collection on the grid of secondary electrons emitted from the anode. The chief disadvantage of the triode is that the capacitance between the grid and the anode provides unwanted feedback between the input and output circuits. This limits the maximum operating frequency to around 200 MHz for coaxial triodes and 4 GHz for planar triodes.

12.2.1   Case Study: The ML-5681 Triode

The theory of the triode developed in Chapter 6 can be illustrated by computing the characteristic curves of a triode for which data is available [19, 20]. This tube, designed to give 50 kW output power at up to 110 MHz, had coaxial geometry with a helical grid supported on 16 equally spaced longitudinal rods. The electron source was a thoriated tungsten filament comprising 16 parallel wires equally spaced in a circle and centred mid-way between the grid support rods. The characteristic curves of this tube are shown in Figure 12.7. Examination of the data in this figure reveals the following information:

- The anode current curves are approximately parallel throughout the region in which the grid current is zero. The slopes of these curves correspond to values of μ in the range 22–25 with a stated design value of 23. The variation in μ is presumably the effect of space-charge between the grid and the anode, which has been ignored in the theory.
- The curves of constant cathode current, computed by adding the grid current to the anode current, are extensions of the straight sections of the anode current curves. From this it may be concluded that, if there is emission from the grid, few electrons are collected by the cathode.
- A plot of total current to the power 2/3 against the grid voltage is close to being a straight line. Thus the tube is free from island formation (see Section 6.4.2). The departures from the straight line at higher currents are a consequence of the variations in μ.
- Figure 12.8 shows experimental data for the ratio of the grid current to the anode current for anode currents in the range 10A to 80A and grid currents in the range 1A to 20A. The points are a good fit to the empirical line given by

$$\frac{I_g}{I_a} = 0.67\sqrt{\frac{V_g}{V_a}} - 0.2, \tag{12.15}$$

showing that space-charge effects are not important. Figure 12.8 also shows lines derived from (6.61) and (6.66), neither of which is a good fit to the experimental data (see Worksheet 12.2). The greater slope of the line given by (12.15) may be the result of the collection of electrons on the grid support wires, and the offset may be caused by focusing effects of the fields in three dimensions.

The approximate internal dimensions of the tube are given in the paper or may be found by measurements on the photographs in it. The results are summarised in Table 12.2. The amplification factor computed from the data in Table 12.2 using (6.36) is 23.9, which lies in the range found from the characteristic curves. The use of equations derived for planar triodes is justified because the distance between the filament and the anode is small compared with the inner radius of the anode. It has been found from experience that a filament of the kind used in this tube is equivalent to a continuous

Table 12.2: Key dimensions of the ML-5681 triode

Parameter	Dimensions	Normalised dimensions
Grid diameter	88 mm	
Grid spacing (a)	3.5 mm	1.0
Grid wire radius (r)	0.3 mm	0.086
Cathode-grid spacing (d_1)	3.18 mm	0.91
Grid-anode spacing (d_2)	9.5 mm	2.72
Effective cathode length (L_c)	165 mm	
Number of grid sections (N_s)	16	
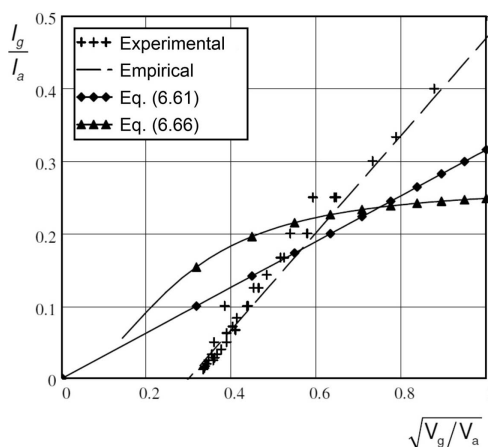
Figure 12.8: ML-5681 triode: comparison between measured and calculated grid currents.

cathode whose active sections have a width equal to twice the distance between the filament and the grid ([5] p.191). Thus the area of the equivalent cathode is

$$A_{eff} = 2N_s d_1 L_c. \tag{12.16}$$

The effective area corresponds to only about one-third of the perimeter of the grid so it is safe to assume that the screening factor of the grid is not affected by the grid support rods. The cathode current can be computed for any values of the grid and anode voltages using (6.56) and (12.16). The normalised cathode-grid spacing is close to unity so there is virtually no island formation and this is confirmed by direct calculation. The characteristic curves computed using (12.15) for the grid current shown in Figure 12.9 are in good agreement with those in Figure 12.7.

The validity of these characteristic curves as a description of the tube at its maximum operating frequency can be established by examining the time taken for electrons to move from the cathode to the grid since it is in that region where they are moving most slowly. For typical class B operation with a minimum anode voltage of 1.5 kV, and a maximum grid voltage of 515 V, the transit time given by (5.20)
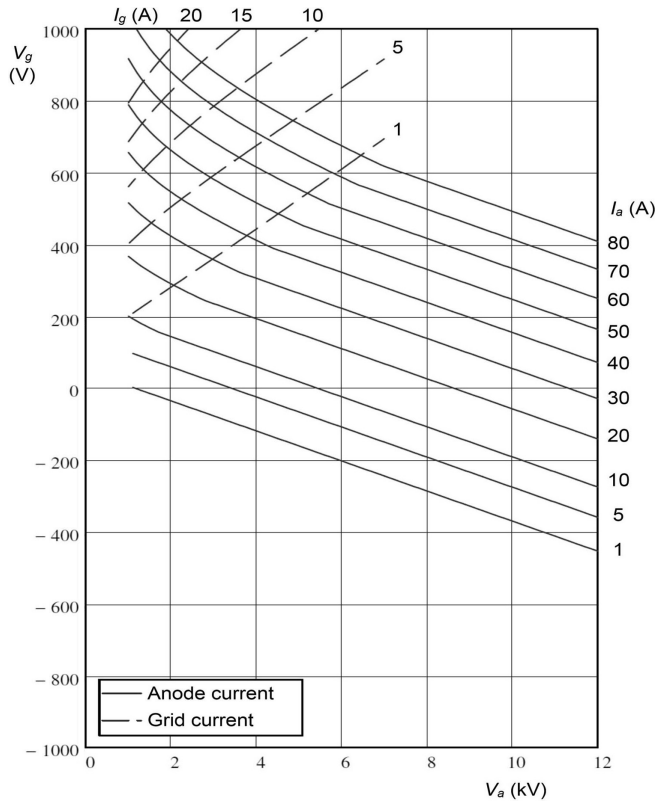
Figure 12.9: ML-5681 triode: theoretical characteristic curves.

using the effective diode voltage is 8% of an RF period. The increase in the magnitude of the RF current computed using (5.95) is 5%. The conclusion that the use of the static characteristics is valid under these conditions is confirmed by the figures given for a class B amplifier in the tube data sheet. At the same frequency the effective length of the filament is 1/16 of the free space wavelength. The overall conclusion is that the methods described above provide a good initial estimate of the properties of a tube. They are therefore useful for conceptual design calculations, and for gaining insight into the relationship between the dimensions of a tube and its electrical properties.

12.3   Tetrodes

Figure 12.10 shows the arrangement of a high-power tetrode in which a second grid (the *screen grid*) has been placed between the control grid and the anode. The characteristic curves of a typical high power tetrode are shown in Figure 12.3. They resemble the curves of a triode shown in Figure 12.7, but the dependence of the anode current on the anode voltage is much less because the screen grid keeps the
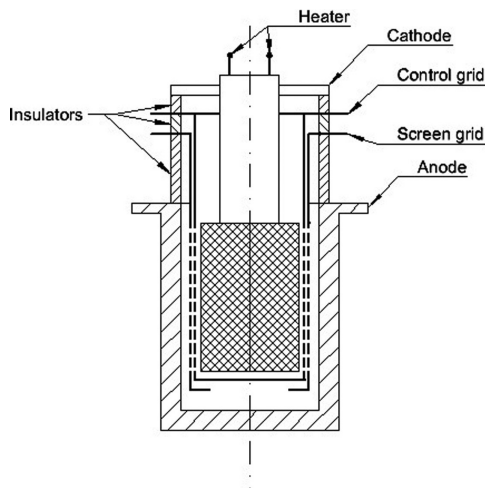
Figure 12.10: Arrangement of a high-power tetrode.

effective anode voltage almost constant. It is important that the anode potential should not fall below that of the screen grid because then secondary electrons liberated from the anode are collected on the screen grid and the thermal dissipation on that grid is increased. The collection of secondary electrons also distorts the characteristic curves by reducing the anode current. The current collected by the control grid, when it is positive with respect to the cathode, is seen to be largely independent of the anode voltage. It is usual in tetrodes to arrange the position of the wires of the screen grid so that they are shielded, to some extent, by the wires of the control grid so that fewer electrons are intercepted by the screen grid.

12.3.1   Case Study: The RS 2058 Tetrode

The theory of the tetrode, described in Chapter 6, can be validated by comparison with experimental results in the same way as for a triode. If the detailed dimensions are not available it is still possible to fit the model to experimental data and infer the dimensions of the tube. As an example we will consider the RS 2058 tetrode whose characteristic curves for one value of screen grid voltage are shown in Figure 12.3 [12]. This tube is a coaxial tetrode with a directly heated thoriated tungsten cathode. It is designed for use as an RF amplifier at frequencies up to 220 MHz with an output power of 64 kW CW and up to 150 kW pulsed. The ratings of this tube for operation as a class B, grounded grid, RF power amplifier are shown in Table 12.3 [12].

The first step is to analyse the characteristic curves to determine whether there is island formation. This can be done by plotting the 2/3 power of the total current against the control grid voltage when the anode and screen grid voltages are constant. Figure 12.11 shows the plot for an anode voltage of 6 kV and various screen grid voltages. The lines are effectively straight for control grid voltages greater than

Table 12.3: RS 2058 tetrode: Ratings for operation as a class B, grounded grid, power amplifier

Frequency	220 MHz
Anode voltage	12 kV
Screen grid voltage	1000 V
Control grid voltage	–350 V
Peak RF control grid voltage	320 V
DC cathode current	35 A
Peak cathode current	100 A
Anode dissipation	90 kW
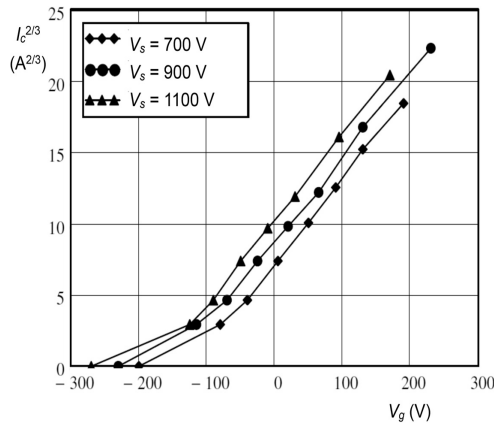
Figure 12.11: RS 2058 tetrode: graph of cathode current to the power 2/3 against grid voltage determined from the characteristic curves for 6 kV anode voltage and three screen grid voltages.

–100 V and have appreciable curvature for lower voltages, showing that island formation is occurring. We saw in Section 6.4.2 that the straight portions of these curves correspond to an analysis in which island formation is ignored. Thus, setting $E_x = 0$ in (6.87), the intercept of the projection of the straight line on the horizontal axis is

$$V_g = - D'_g \left(V_s + D_s V_a \right). \tag{12.17}$$

If values of D'_g and D_s are assumed then V_g can be calculated for pairs of values of V_s and V_a and compared with values calculated directly from the characteristic curves. Using iteration, values of D'_g and D_s can be found which minimise the RMS difference between the data and the theoretical model.

Next we note, from (6.46) and (6.87) that, at cut-off

$$V_g = - D'_g \left(a/2 \right) \cdot \left(V_s + D_s V_a \right) \tag{12.18}$$

Table 12.4: Constants of the RS 5058 tetrode determined from the characteristic curves

D_s	D_g'	$D_g'(a/2)$
0.019	0.15	0.23

where, from (6.88),

$$D_g'(y) = \frac{D_g(y)}{1 + D_s(1 + d_3/d_2)}, \tag{12.19}$$

where $D_g(y)$ is given by (6.45). Then $D_g'(a/2)$ can be calculated from the voltages at cut-off. These calculations were carried out for $V_s = 700\,\text{V}$, $900\,\text{V}$ and $1100\,\text{V}$ and $V_a = 6\,\text{kV}$ and $12\,\text{kV}$ with the results shown in Table 12.4.

The same constants can be calculated from the dimensions of the tetrode using the theory described above. In the absence of information to the contrary we will assume that the spacings and wire diameters of the two grids are the same. We can take $a = 1\,\text{mm}$ without loss of generality since all the equations involve ratios of the dimensions. The tube is then defined in terms of four constants: r, d_1, d_2 and d_3 which must be chosen so that the calculated values of the penetration factors are equal to the values given in Table 12.4. Since there are four unknowns, but only three equations, the solution obtained is not unique. However, it is found that, by using the results of the calculation as trial values, the process converges onto a stable solution after a few iterations. This solution has the property that the values are almost constant if one of the trial values is altered slightly. The results were used to calculate the set of curves shown in Figure 12.12 for comparison with those in Figure 12.11. The area of the cathode was chosen so that the current calculated agreed with that in the characteristic curves. It can be seen that the two sets of curves are in good agreement with one another.

The characteristic curves can also be computed for comparison with Figure 12.3 by first finding the current intercepted by the control grid, as described in Section 6.4.1. The primary current division factor is calculated from (6.62) with $D = D_g$, and the ratio of the grid current to the anode current is found using (6.61) with V_a replaced by V_e. This turns out to be quite a good approximation to the grid current in this case. Figure 12.13 shows the curves of constant anode current for comparison with Figure 12.3. The curves of constant control grid current (not shown) are very similar to those in Figure 12.3 (see Worksheet 12.3).

The agreement between the calculated and measured characteristic curves is good, apart from the curvature at low anode voltages caused by the collection of current on the screen grid. This current is a small fraction of the total current except at large positive control grid voltages and low anode voltages, when the space-charge between the screen grid and the anode becomes important in determining the trajectories of the electrons, as discussed in Section 6.7.2. Under other operating conditions the trajectories are focused between the wires of the screen
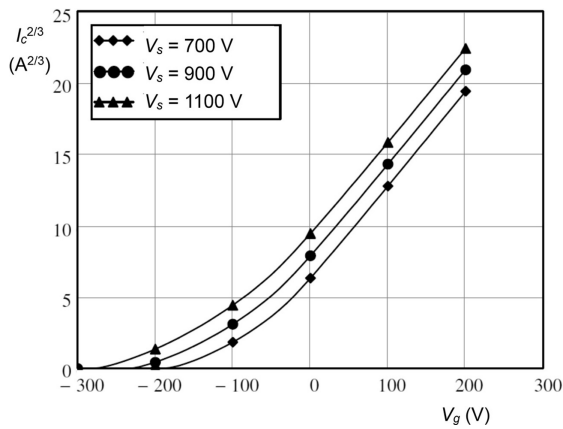
Figure 12.12: RS 2058 tetrode: graph of cathode current to the power 2/3 against grid voltage computed for 6 kV anode voltage and three screen grid voltages.
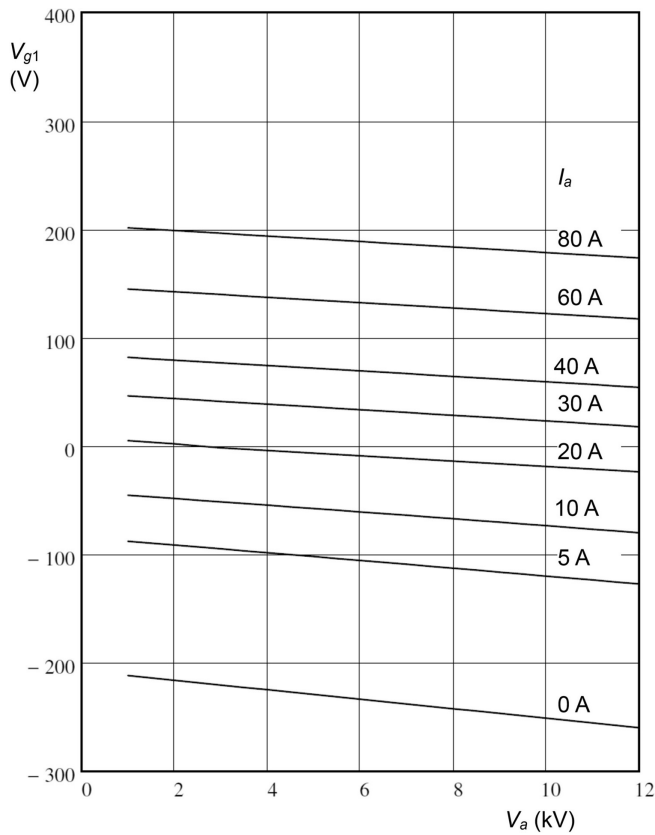
Figure 12.13: Calculated characteristic curves of the RS 2058 tetrode $(V_s = 900\,\text{V})$.

Table 12.5: Computed dimensions for the RS 2058 tetrode

Parameter	Possible dimensions	Normalised dimensions
Grid spacing (a)	2.4 mm	1.0
Grid wire radius (r)	0.145 mm	0.06
Cathode-control grid spacing (d_1)	0.8 mm	0.33
Control grid – screen grid spacing (d_2)	2.2 mm	0.92
Screen grid-anode spacing (d_3)	20.7 mm	8.58
Cathode diameter	37 mm	
Cathode length (L_c)	85 mm	
Anode inner diamter	84 mm	

grid by the electrostatic fields, and the intercepted current is small [21]. There is no simple way of estimating the screen grid current.

The final values of the normalised parameters for the tube are shown in Table 12.5. It must be emphasised that these have been obtained by fitting the model to the measured data and are not the actual parameters of the tube. The scaling of the parameters can be estimated from the transit angle for typical class B operation. If we assume that $\beta_e d_1 = 0.1$, and that the amplitude of the RF control grid voltage is 350 V, then $d_1 = 0.8$ mm and all the other dimensions can be found by scaling from the normalised values.

If the length of the active region is taken to be $\lambda/16 = 85$ mm then the diameter of the cathode is 37 mm, which is plausible for a thoriated tungsten mesh cathode that emits over its whole surface [13]. The internal diameter of the anode is then 84 mm which can be compared with the external diameter of the anode water cooling jacket (125 mm). The maximum cathode current density is 1 A cm^{-2} which is well within practical limits for this kind of cathode [13]. It may be concluded that the methods adopted have yielded a plausible set of dimensions for this tube.

12.4   Design of Triodes and Tetrodes

Information about design procedures for high power triodes and tetrodes is available in a number of sources [6, 13, 19, 22]. Triodes and tetrodes are used in power oscillators for industrial heating, FM and AM broadcast transmitters, SSB communications transmitters, pulsed radar transmitters and RF power sources for particle accelerators. They have also been used as high power switches. Because of limitations of space we will only consider the design of tetrodes for use as CW power amplifiers operating in conditions close to class B. The reader is referred to the sources cited for further information. To show how a tetrode might be designed we will consider the design of a tube similar to the RS 2074 tetrode which delivers up to 600 kW of CW output power at frequencies up to 110 MHz [12]. The steps in the design calculations are as follows:

i) From Figure 12.5(a) it is expected that the conversion efficiency of the tube operating in class B will be around 70%. For an output power of 600 kW the DC input power required is 860 kW. Tubes of this kind generally have thoriated tungsten cathodes and operate at anode voltages up to around 20 kV [13]. But high anode voltages lead to greater risk of voltage breakdown both inside and outside the tube and so we will choose a figure of 15 kV. (Note that, for purposes of illustration, the use of the exact figures in the data sheet has been avoided.) Then the DC current is 60 A and, from Figure 12.5(b), the peak current is 204 A. It is desirable for the control grid voltage to be positive for only a small part of the cycle to minimise the control grid current, as shown by the load line in Figure 12.3. From that figure we see that the sum of the two grid currents is typically about 20% of the maximum anode current. Thus the cathode must be able to deliver a peak current of 255 A.

ii) Let us assume that the screen grid voltage is 10% of the DC anode voltage and that the penetration factor of the control grid (D_g) is 0.2. Then the control grid voltage at cut-off is approximately

$$V_{g0} = D_g V_s = -300 \text{ V}. \tag{12.20}$$

Now the maximum positive control grid voltage in a class B amplifier must not be too great to avoid excessive interception of current on the control grid. We will assume that the maximum control grid voltage is 100 V so that the amplitude of the RF control grid voltage is $V_{g1} = 400$ V. The spacing between the cathode and the control grid must be small enough to avoid transit time effects at the maximum operating frequency. Let us assume that $\beta_e d_1 = 0.05$ and that $\beta_e = \omega / \left(\sqrt{2e V_{g1}/m_0} \right)$ so that $d_1 = 0.86$ mm.

iii) To avoid island formation let us choose $a = d_1$ and assume that $d_2 = d_1$. Then the radius of the control grid wires can be found by making the value of D_g calculated from (6.36) equal to the figure assumed above. If the grid is to be a tungsten or pyrolitic graphite mesh then the penetration factor can be calculated by using the geometrical screening factor in the calculations [23, 24].

iv) Let us assume that the penetration factor of the screen grid (D_s) is 0.02. The screen grid wire radius, and the distance between the screen grid and the anode, are adjusted to achieve this figure. It may be noted that the wires of the screen grid are often thicker than those of the control grid to make it more robust and that, typically, $d_3 \sim 10 d_2$.

v) The current density can now be calculated as a function of the electrode voltages using (6.93). The area of the cathode is computed from the peak current calculated above and, hence, the current density at the maximum control grid voltage with the chosen screen grid voltage and minimum anode voltage.

vi) If it is assumed that the active length of the tube is $\lambda/16$ then the diameter of the cathode and of the other electrodes can be calculated.

vii) The control grid current can be calculated as described in Section 6.4.1 and then the characteristic curves can be plotted.

Table 12.6: Leading dimensions of a 600 kW, 110 MHz tetrode

Parameter	Dimensions
Grid spacing (a)	0.86 mm
Control grid wire radius (r_g)	0.04 mm
Screen grid wire radius (r_s)	0.052 mm
Cathode-control grid spacing (d_1)	0.86 mm
Control grid–screen grid spacing (d_2)	0.86 mm
Screen grid-anode spacing (d_3)	6.9 mm
Cathode diameter	38 mm
Cathode length (L_c)	170 mm
Anode inner diameter	55 mm

The complete set of dimensions calculated in this way is shown in Table 12.6. The figures are comparable with those of the RCA 4648 tetrode [6] apart from the diameters of the grid wires which may be too thin. The peak cathode current density is 1.3 A cm^{-2}, which is within the practical range [13]. The perimeter of the anode is well below the value which would permit the excitation of the TE$_{11}$ mode. The set of calculations described above can readily be incorporated in a worksheet which enables the effects of different design choices to be examined quickly (see Worksheet 12.4). In practice the results of this calculation would be informed by experience and would then be confirmed by more detailed calculations [22].

The anode power density in class B operation is just under 1 kW cm^{-2} which is comfortably within the limit for power dissipation with turbulent forced convection cooling [25] (see Section 10.5.3). The electrode spacings must be great enough to avoid voltage breakdown. Whilst it is difficult to calculate figures for the geometry of such a tube, it may be noted that the voltages and spacings are similar to those of the RCA 4648 tetrode [6]. The heater power required for the cathode can be estimated using the method described in [18].

This section has discussed the principles of design of coaxial tetrodes and it could equally be applied to triodes and to planar tubes. Information about the technology used can be found in the literature [6, 13, 18]. A number of variants of the standard coaxial design exist including [6, 13, 26]:

- Tubes comprising multiple unit triodes or tetrodes connected in parallel within a single vacuum envelope. This system has the advantage that it can readily be scaled to higher power levels, provided that the diameter does not become so big that the TE$_{11}$ mode can be excited.
- Tubes in which the cathode is outside the anode.
- Beam power tubes in which additional electrodes are used to focus the electrons, so that the interception of electrons by the grids is reduced.
- Tubes in which there is a second anode insulator, so that the coaxial space within the anode can be located in the centre of the external resonator, rather than at
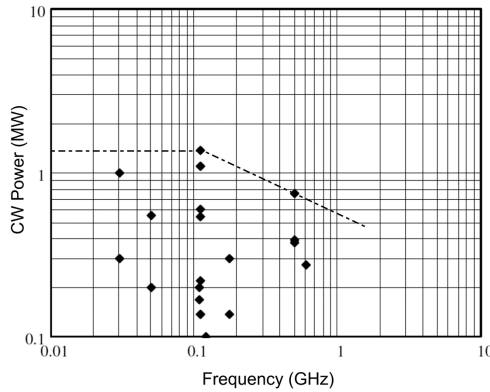
Figure 12.14: State of the art of commercially available high power tetrodes (CW Class B operation).

one end. It is then possible to arrange that the standing wave in the resonator has an anti-node of RF voltage at the centre of the anode, and a node of RF current. Thus the heat dissipation caused by RF currents in the screen grid is much less than in a conventional tetrode. This technique enables the RF power output to be doubled at a given frequency [26].

Figure 12.14 shows the state of the art of commercially available high power tetrodes.

12.5 Design of Triode and Tetrode Amplifiers

The process by which a tetrode amplifier can be designed is best explained by means of an example [27]. This is based upon a 62 kW, 200 MHz, amplifier used in the CERN SPS [28]. The example was chosen because sufficient information is available about the amplifier to verify the results of the calculations. The amplifier uses a single RS2058CJ tetrode [12] operating with a DC anode voltage of 10 kV, and 900 V screen grid bias. The design procedure described below is based upon that given in [29, 30]. The actual amplifier is operated in class AB but quite close to class B. For simplicity class B operation is assumed in the calculations which follow.

i) The first step is to estimate the probable efficiency of the amplifier. We will assume that the minimum anode voltage is 1.5 kV. Using the method described in Section 12.1.2 we expect that the efficiency of the amplifier will be around 72% (see Figure 12.5(a)). This estimate can be adjusted later, if necessary, when the actual efficiency has been calculated. Then the DC power input required to obtain the desired output power is

$$P_2 = 62 / 0.72 = 86 \text{ kW}. \tag{12.21}$$

The DC anode voltage was chosen to be $V_{a0} = 10\,\text{kV}$ so the mean anode current is

$$I_{a0} = 86/10 = 8.6 \; A. \tag{12.22}$$

ii) The theoretical value of I_{pk} from the Fourier analysis is $3.6I_{a0}$ but when other factors including island formation, space-charge effects and grid current are taken in to account it is found that this figure lies in the range 3.6 to 4.4. If we take the factor to be 4.0 then

$$I_{pk} = 4 \times 8.6 = 34 \; A. \tag{12.23}$$

iii) Next we construct the load line on the characteristic curves for the tube, as shown in Figure 12.3, by joining the point (1.5 kV, 34 A) to the quiescent point (10 kV, 0 A). We note that this requires the control grid voltage to swing slightly positive with a maximum of +70 V.

iv) The DC and RF currents are found by numerical Fourier analysis of the anode current waveform, using values read from Figure 12.3 at intervals of 15°. The anode voltages are given by

$$V_a = 10 - 8.5 \cos \theta, \tag{12.24}$$

where θ is the phase angle. The results are shown in Table 12.7. Then using the formulae given in [30]

$$I_{a0} = \frac{1}{12}(0.5I_0 + I_{15} + I_{30} + I_{45} + I_{60} + I_{75}) \tag{12.25}$$

and

$$I_{a1} = \frac{1}{12}(I_0 + 1.93I_{15} + 1.73I_{30} + 1.41I_{45} + I_{60} + 0.52I_{75}), \tag{12.26}$$

where the subscripts refer to the phase angles in degrees. Using the data from Table 12.7 it is found that:

$$I_{a0} = 8.9 \; \text{A}, \tag{12.27}$$

$$I_{a1} = 15.0 \; \text{A}. \tag{12.28}$$

v) The DC input power is

$$P_0 = I_{a0} V_{a0} = 89 \; \text{kW}. \tag{12.29}$$

The amplitude of the RF voltage is

$$V_{a1} = 10.0 - 1.5 = 8.5 \; \text{kV} \tag{12.30}$$

and the RF output power is

$$P_2 = \frac{1}{2} V_{a1} I_{a1} = 64 \; \text{kW}, \tag{12.31}$$

Table 12.7: Anode currents at 15° phase intervals taken from Figure 12.3

Phase (degrees)	0	15	30	45	60	75	90
V_a (kV)		1.5	1.8	2.6	4.0	5.7	7.8	10.0
I_a (A)		34	32.5	28	18	8	3	0

which is very close to the desired value and gives an efficiency of 72%, as originally assumed.

vi) The effective load resistance is

$$R_L = V_{a1} / I_{a1} = 570 \ \Omega. \tag{12.32}$$

The source impedance of the output of the amplifier can be found by noting that, if the RF load resistance is zero and the anode voltage is constant, the peak anode current is 46 A for the same RF voltage on the control grid. Thus the short circuit RF current is 20 A and the anode source resistance (R_a) is 1.7 kΩ.

vii) To find the input impedance of the amplifier we note that the amplitude of the RF control grid voltage is

$$V_{g1} = 245 + 70 = 315 \ \text{V} \tag{12.33}$$

and that, for grounded grid operation, the amplitude of the RF input current is the sum of the RF anode and control grid currents

$$I_{c1} = I_{a1} + I_{g1} \approx I_{a1}. \tag{12.34}$$

viii) The amplitude of the RF control grid current (I_{g1}) may be obtained by reading the control grid currents off Figure 12.3 at 15° intervals and employing Eq. (12.26). The result is 0.67 A, which is small compared with the RF anode current, and can be neglected in the first approximation. The RF input resistance is then

$$R_1 = V_{g1} / I_{a1} = 20 \ \Omega. \tag{12.35}$$

ix) Finally, we note that the input power is

$$P_1 = \frac{1}{2} V_{g1} I_{a1} = 2.5 \ \text{kW} \tag{12.36}$$

and that the power gain of the amplifier is

$$G_{dB} = 10 \log (64 / 2.5) = 14 \ \text{dB}. \tag{12.37}$$

Table 12.8 shows a comparison between the figures calculated above and those reported in [31]. The differences between the two columns of Table 12.8 are attributable to the difference between the actual class AB operation and the class B operation assumed in the calculations.

Table 12.8: Comparison between actual, and calculated, parameters of the amplifier described in [31]

Parameter	Actual	Calculated
Anode voltage (kV)	10	10
Anode current (A)	9.4	8.9
Screen grid voltage (V)	900	900
Control grid voltage (V)	−200	−245
RF output power (kW)	62	64
RF input power (kW)	1.8	2.5
Gain (dB)	15.4	14
Efficiency (%)	64	72
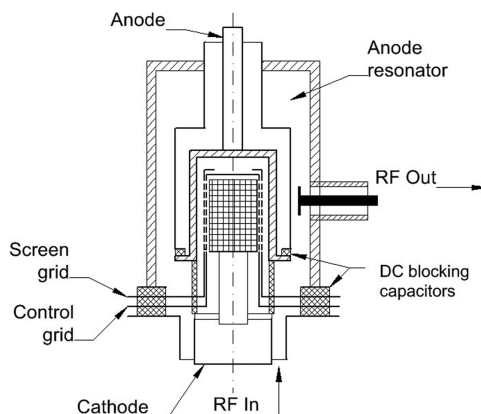
Figure 12.15: Arrangement of a tetrode amplifier (copyright 2011, CERN, reproduced, with permission, from [10]).

12.5.1   Practical Details

Figure 12.15 shows a simplified diagram of a tetrode amplifier. The tube is operated in the grounded grid configuration with coaxial input and output circuits. The outer conductors of the coaxial lines are at ground potential and they are separated from the grids by DC blocking capacitors. The anode resonator is a re-entrant coaxial cavity which is separated from the anode by a DC blocking capacitor. The output power is coupled through an impedance matching device to a coaxial line. The anode HT connection and cooling water pipes are brought in through the centre of the resonator.

The electrodes of the tube form coaxial lines with characteristic impedances of a few ohms. We have seen above that the input impedance of the amplifier is typically a few tens of ohms, and the output impedance a few hundred ohms. Thus the terminations of both the input and output lines are close to open-circuits. The

anode resonator therefore has one end open circuited and the other short circuited, and it must be an odd number of quarter wavelengths long at resonance. Typically the resonator is 3/4 of a wavelength long. In that case the point at which the output coaxial line is coupled into the resonator can be used to transform the impedance to provide a match. There is also a voltage node towards the lower end of the outer part of the resonator and this can be used to bring connections through to the anode. The higher-order modes of the cavity can be troublesome, and it is usually necessary to damp them by the selective placing of lossy material within the cavity, or by using coupling loops connected to external loads. The cathode heater connections must incorporate some means of decoupling from the RF circuit. Further information about the design of gridded tube amplifiers can be found in [17, 28, 32].

12.6   Inductive Output Tubes (IOTs)

Tetrode amplifiers suffer from the disadvantage that the same electrode, the anode, is part of both the DC and the RF circuits. The output power is limited by screen grid and anode dissipation. In addition, the electron velocity is least when the current is greatest because of the voltage drop across the output resonator. This means that the gap between the screen grid and the anode must be small to avoid transit time effects. To get high power at high frequencies it is necessary to employ high-velocity electrons, and to have a large collection area for them. This can be achieved by separating the electron collector from the RF output circuit. The possibility that this might be done was originally recognised by Haeff in 1939, but it was not until 1982 that the concept became practical, following the development of laser-machined pyrolytic graphite grids [7, 8]. Haeff called his invention the 'Inductive Output Tube' (IOT) but it is also known by the proprietary name Klystrode ®.

Figure 12.16 shows a schematic diagram of an IOT. The electron beam is formed by a gridded Pierce electron gun and confined by an axial magnetic field (not shown). The RF input circuit is formed by the cathode and the grid in a manner similar to that in a tetrode amplifier. The gun is biased so that the conduction angle is
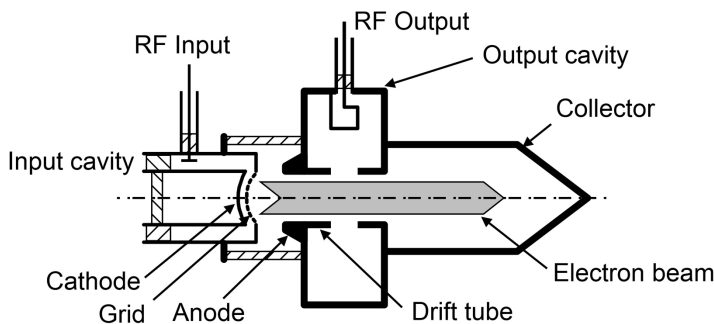
Figure 12.16: Arrangement of an inductive output tube (IOT)
(copyright 2010, IEEE, reproduced, with permission, from [33]).
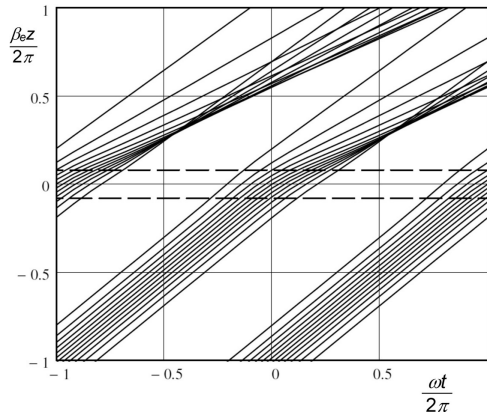
Figure 12.17: Electron trajectories in an IOT.

approximately 180° and the electron bunches formed by the cathode and the grid are accelerated through the constant potential difference between the cathode and the anode. A detailed diagram of the input cavity resonator for an IOT is given in [34]. The bunched electron beam passes through a cavity resonator where the azimuthal magnetic field induces a current in the cavity, hence the name of the tube (see Section 11.8.4). Because the cavity is tuned to the repetition frequency of the bunches the RF electric field in the interaction gap is maximum in the retarding sense when the centre of a bunch is at the centre of the gap. This resembles the operation of a class B amplifier with the important difference that the input electron velocity is constant throughout the RF cycle. Figure 12.17 shows a typical Applegate diagram for an IOT (see Worksheet 12.5). Because the electron velocity is high it is possible to use a much longer output gap than in a tetrode. Other advantages of the IOT over the tetrode are that it does not need a DC blocking capacitor in the RF output circuit, because the cavity is at ground potential, and that it has higher isolation between input and output and a longer life than an equivalent tetrode amplifier. These advantages are offset to some extent by the need for a magnetic focusing field. The typical gain is 20–23 dB which is appreciably higher than that of a tetrode amplifier and high enough for the input to a 60 kW tube to be supplied by a solid state driver stage.

IOTs have been designed for use in UHF TV transmitters, and for powering particle accelerators [35–38]. For both these applications the linearity of the power transfer curve is a valuable asset (see Figure 12.22(a)). Where increased bandwidth is required (e.g. for digital TV transmission) a double-tuned output cavity is used [34, 37] (see Section 3.2.4). A general review of the principles and technology of IOTs and related devices is given in [39].

12.6.1 Bunch Formation

The beam current is given as a function of time by (12.6). To avoid transit time effects, and to minimise the RF voltage required to modulate the beam, the grid

must be placed very close to the cathode so that $\beta_e d_1 \ll 0.7$ at the maximum effective diode voltage (see Section 5.7). Calculations based on the 1.3 GHz tube discussed in Section 12.6.5, for example, suggest that the separation between the grid and the cathode should be less than 0.3 mm when the amplitude of the RF grid voltage is 350 V. At lower forward voltages the transit angle is increased and transit time effects become important. The electron velocity at the grid is smaller at the start and the finish of each bunch than it is at the bunch centre. Hence the bunch waveform is distorted from that shown in Figure 12.4 because the currents at the start and the finish of the bunch are delayed with respect to the bunch centre [40]. The very small spacing between the cathode and the grid must be maintained accurately at the working temperature of the gun. It sets a limit on the maximum frequency of operation and means that island formation is likely to occur. The cathode and the grid must be manufactured with high precision and have high thermal stability to ensure that the intended separation is maintained, and to avoid the risk of a short circuit between them.

The penetration factor of the grid may vary over its surface for two reasons. In the first place the electric field of the anode is not normally constant over the surface of the grid [41]. It may be necessary to compensate for that by varying the spacing of the grid wires and/or the distance between the grid and the cathode. In the second place the diameter of the grid may be an appreciable fraction of the free-space wavelength. If we use the same guidelines as in the design of triodes and tetrodes then the cathode diameter should be less than $\lambda/8$. For the example used above where the frequency is 1.3 GHz the cathode diameter should be less than 30 mm. If the penetration factor is not constant for any reason, or if island formation exists, then the effective power law of the gun is greater than 3/2 (see Section 6.4.2). The input circuit of an IOT is effectively that of a common grid amplifier so that the RF input power is given by (12.13).

Other ways of modulating the beam, which could permit operation at higher frequencies, have been investigated theoretically including: photo-cathodes [42–44], gated field emission arrays [45], and deflection modulation [46].

12.6.2   Space-Charge Debunching

After the electrons have left the grid they are accelerated by the field between the grid and the anode, and space-charge forces are normally negligible. However, when the electron bunches have passed into the drift tube, then space-charge forces dominate. The radial space-charge forces are counteracted by the force of the magnetic focusing field, but the axial space-charge forces tend to cause debunching so that the conduction angle increases. The effect of space-charge forces can be calculated by numerical integration of the equations of motion of a set of electrons distributed within the bunch, as described in Section 11.8. Since any increase in the bunch length (conduction angle) causes a reduction in conversion efficiency (see Figure12.5(a)) it is desirable to keep the drift length as short as possible.

12.6.3   Power Transfer in the Output Gap

The extraction of power from the electron bunches in the output gap can be computed using a disk model or by using the simplified theory in Section 11.8.4 (see Worksheet 12.5). If the effective output load resistance, including the effects of losses in the output cavity, is R_L then the minimum energy of electrons leaving the gap is

$$V_{\min} = V_a - M_{eff}^2 I_1 R_L, \tag{12.38}$$

where V_a is the voltage of the anode relative to the cathode, and the effective gap coupling factor is given by (11.182). Self-consistent values of $V_{\min}$ and $u_{\min}$ can be found by iteration using (11.182) and (12.38). Then the electronic efficiency is

$$\eta_e = \frac{M_{eff} I_1 V_g}{2 I_0 V_a}. \tag{12.39}$$

The usefulness of this approach for modelling the performance of an IOT is demonstrated in the case study in Section 12.6.5. For typical dimensions, and class B operation, $I_1/I_0 = 1.65$, $M = 0.8$, $M_{eff} = 0.53$ and $MV_g/V_a = 1.2$ when the gap voltage is adjusted to give the optimum efficiency. Inserting these figures in (12.39) gives an electronic efficiency of 66%. This is shown by the dashed line in Figure 12.18(a) for comparison with the efficiency of CW IOTs given in manufacturers' data sheets. Figure 12.18(b) shows data from the same sources for CW output power.

The maximum output power of an IOT is limited by the DC current which can be drawn from the cathode, while avoiding excessive interception by the grid, and by the maximum voltages which can be applied without voltage breakdown. Information about the design of a 1.3 GHz, 100 kW CW, IOT is given in [47]. A 500 kW pulsed IOT at 425 MHz is described in [48]. The use of PIC codes to model IOTs has been described in [40, 49, 50]. The possibility of developing higher power IOTs with multiple beams, annular beams, or radial beams has been studied in [51–56].
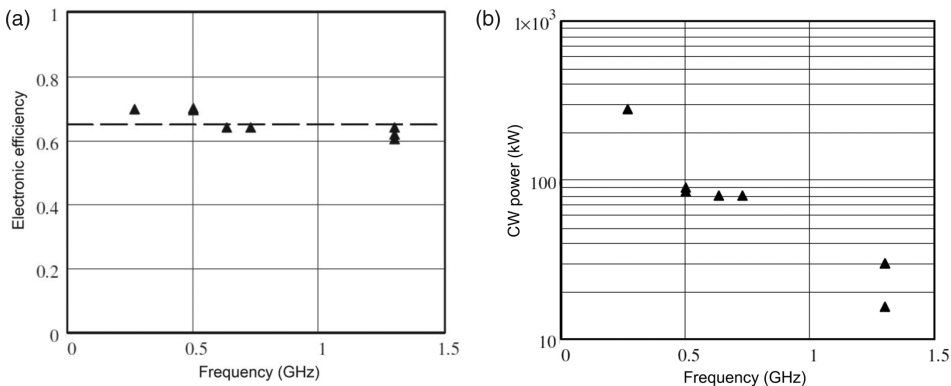
Figure 12.18: State of the art of commercial CW IOTs: (a) efficiency, and (b) output power.

12.6.4   IOT Collectors

The conversion efficiency of a IOT at low drive levels can be improved considerably by the use of a multi-element depressed collector. The energy of the electrons in the spent beam can be expressed as a function of the phase angle as

$$V_s(\theta) = V_a - M_{eff} V_{gap} \cos \theta. \tag{12.40}$$

The slowest electrons are those at the centre of the bunch where $\theta = 0$. In the simple model used here the bunch profile is symmetrical, and the part of the beam current which has energy greater than V_s is given by

$$I_s(\theta) = \frac{1}{\pi} \int_{\theta}^{\pi} I(\theta) \, d\theta, \tag{12.41}$$

where θ lies in the range $(0 : \pi)$. The spent beam curves can be plotted, using these equations, for comparison with the results from the disc model, shown in Figure 12.19(a). These curves were obtained by adjusting the RF grid voltage to give an output power of 12 kW in each case. Figure 12.19(b) shows how the spent-beam curves computed using the simple model vary with output power. Although these are superficially similar to those for a klystron (see Section 13.3.7) there are some important differences. The IOT does not exhibit saturation because the maximum power is limited by the electron gun. The tube is typically operated in class AB so that the few electrons crossing the output gap in the accelerating phase of the gap field are only mildly accelerated. Thus there are fewer fast electrons than is normally the case for a klystron. In addition, the mean beam current drops as the output power is reduced so that the electronic efficiency remains high at low output power levels.

Figure 12.20 shows typical dependence of the RF efficiency of an IOT with a five-stage depressed collector on the normalised output power (see Worksheet 12.6). Because the electronic efficiency is higher at low output power, than is the
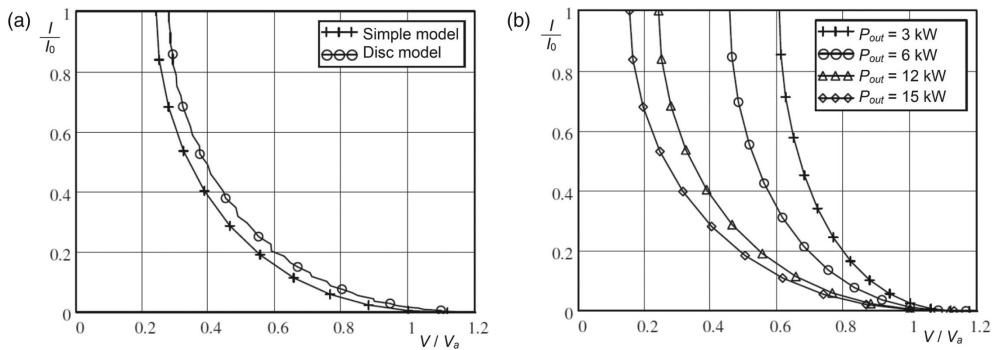
Figure 12.19: Spent-beam curves for an IOT: (a) comparison between different models, and (b) spent-beam curves for different output powers.
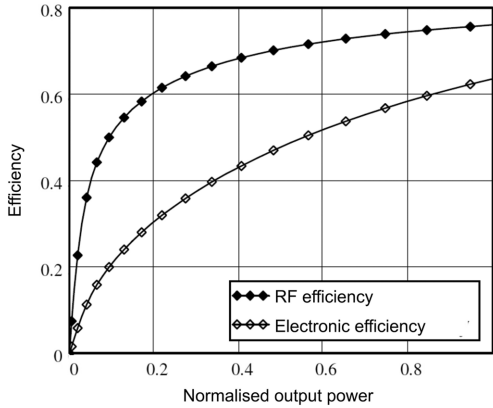
Figure 12.20: Dependence of the efficiency of an IOT with a five-stage depressed collector on normalised output power.

case with a klystron, the overall efficiency is also high at low output powers. It has been shown that, with careful design of the collector, almost constant efficiency of around 65% can be achieved over the range from 25% to 100% of full output power [34, 57, 58]. This performance is very valuable for applications such as digital television broadcasting where the signal has a high peak to average power ratio.

12.6.5   Case Study: The 116LS IOT

The 116LS IOT was designed to generate up to 16 kW CW at 1.3 GHz for powering particle accelerators [59, 60]. This tube has been modelled, using the methods described above, for comparison with the published experimental data [33]. Where information about the tube is not available appropriate assumptions can be made as described below. The description that follows differs slightly from that previously published.

It is assumed that the properties of the electron gun can be described by (12.2) in which K, μ and the power law are treated as adjustable constants. The anode voltage is 25 kV and the DC grid bias voltage $\left(V_{g0}\right)$ is –105 V, as stated in [59]. Then the DC and fundamental RF currents $\left(I_0, I_1\right)$ can be calculated as functions of the amplitude of the RF grid voltage $\left(V_{g1}\right)$. Since IOTs are operated as common grid amplifiers the RF input impedance is given by

$$R_{in} = \frac{V_{g1}}{I_1\left(V_{g1}\right)}. \tag{12.42}$$

If the RF source impedance is R_s then the RF input power is given by

$$P_{in} = \frac{1}{8}\left(\frac{R_{in} + R_s}{R_{in}}\right)^2 \frac{V_{g1}^2}{R_s}. \tag{12.43}$$
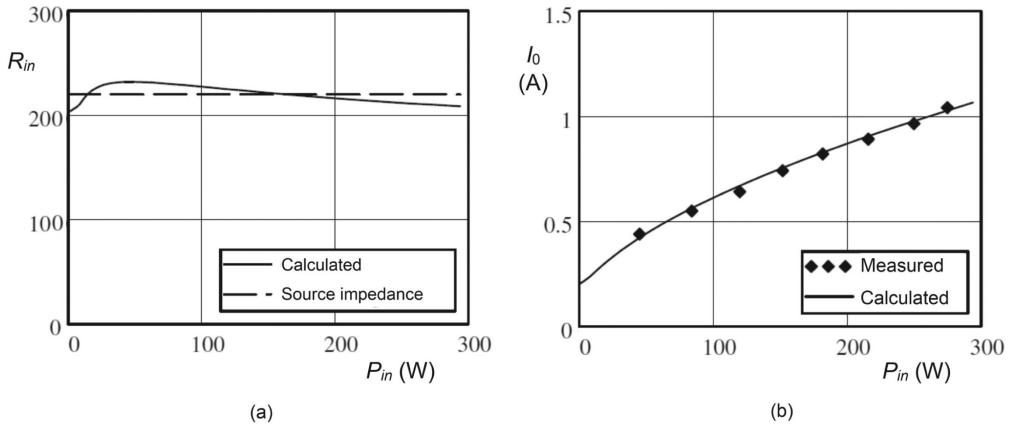
Figure 12.21: (a) Input impedance, and (b) DC beam current, of an IOT as functions of the RF input power
(copyright 2010 IEEE, reproduced, with permission, from [33]).

It should be noted that I_0 R_{in} and P_{in} are functions of V_{g1} with R_s as a parameter. The input impedance and the DC beam current can therefore be plotted against the RF input power, as shown in Figure 12.21(a). If the source impedance is assumed to be 220 Ω then a good fit to the experimental data can be obtained, as shown in Figure 12.21(b), by setting $K = 0.42 \text{ mA.V}^{-1.5}$, $\mu = 150$ and the power law to 1.5. The principles governing the design of the type of gun used in this tube were discussed in Section 12.6.1.

Since dimensions of the output gap are not given in [59] it is assumed that: $\beta_e a = 1.0$, $\beta_e b = 0.6$, and $\beta_e g = 1.0$ where $\beta_e = \omega/u_0$, a is the drift tube radius, b is the beam radius and g is the length of the output gap. These parameters are typical of those used in IOTs and klystrons. The gap coupling factor can then be calculated as a function of the electron velocity. Starting from an initial guessed value of the velocity (u_{min}) of the slowest electrons leaving the gap self-consistent values of u_{min} and V_{gap} are found for each drive power level using (12.38), as described above. The load resistance R_L is used as an adjustable parameter. Note that this procedure differs slightly from that described in [33]. The results using the modified procedure show better agreement with experiment. Graphs of the output power and efficiency against the input power were plotted (see Figure 12.22) and it was found that a good fit to experimental data could be achieved by setting $R_L = 28 \text{ k}\Omega$. The 1 dB bandwidth of this tube was 3 MHz. If it is assumed that this is determined by the properties of the output cavity then $Q_L = 220$ and $R/Q = 120$.

The output power and efficiency were also calculated using the disc model without space-charge. The values of V_{gap} were determined by iteration to give the correct load resistance at each input power level. The results of these calculations are shown in Figure 12.22 for comparison with the results of the simple model. It should be noted that the curve of output power against input power is very nearly a straight
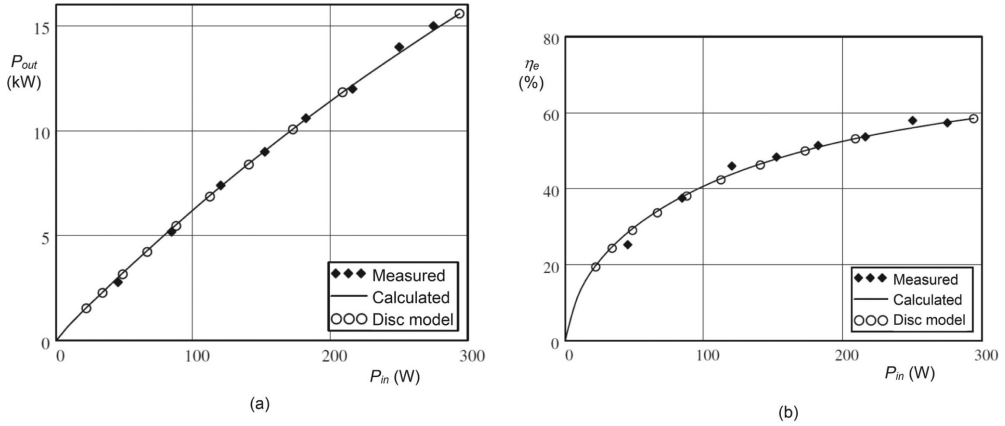
Figure 12.22: (a) RF output power, and (b) electronic efficiency of an IOT as functions of RF input power.
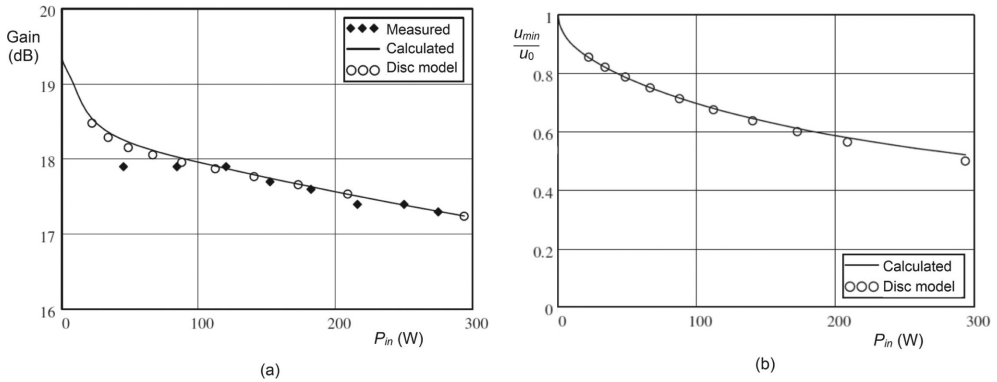
Figure 12.23: (a) Gain, and (b) normalised minimum electron velocity of an IOT as functions of RF input power.

line. The shape of this curve is found to depend upon the quiescent point selected, and it is most nearly linear when there is a small quiescent current. Thus an IOT has performance which is close to that of an ideal linear amplifier. Figure 12.23 shows the gain, and the normalised minimum electron velocities, plotted against RF input power. The gain of this tube was a little lower than that typically achieved in IOTs (around 20 dB). It was increased to around 21 dB in an improved version of the tube that showed evidence of island formation [60]. The minimum electron velocity in the spent beam, normalised to the DC beam velocity, is seen to be around 0.5 so that no electrons are reflected by the output gap. The comparison between the results from the simple model and the disc model shows that the performance of an IOT can be modelled with good accuracy using the simple model, as described above.

The effects of space-charge can be included in the disc model if the drift length is known. For example, the assumption of a drift length five times the length of the gap, gave a reduction in efficiency of five percentage points, and a reduction in gain of 1 dB at an input power of 300 W. Agreement with the measured data could be achieved by changing the load resistance assumed.

References

[1] W. G. Dow, *Fundamentals of Engineering Electronics*. New York: John Wiley, 1937.
[2] A. H. W. Beck, *Thermionic Valves*. Cambridge: Cambridge University Press, 1953.
[3] R. S. Burnap, *Electron Tube Design*. Harrison, NJ: Radio Corporation of America, 1962.
[4] J. W. Gewartowski and H. A. Watson, *Principles of Electron Tubes*. Princeton, NJ: D. van Nostrand Co. Inc., 1965.
[5] K. R. Spangenberg, *Vacuum Tubes*. New York: McGraw-Hill, 1948.
[6] T. E. Yingst *et al.*, 'High-power gridded tubes – 1972', *Proceedings of the IEEE,* vol. 61, pp. 357–381, 1973.
[7] A. V. Haeff and L. S. Nergaard, 'A wide-band inductive-output amplifier', *Proceedings of the IRE,* vol. 28, pp. 126–130, 1940.
[8] D. Preist and M. Shrader, 'The klystrode – an unusual transmitting tube with potential for UHF-TV', *Proceedings of the IEEE,* vol. 70, pp. 1318–1325, 1982.
[9] R. G. Carter, 'Acceleration technologies for charged particles: an introduction', *Contemporary Physics,* vol. 52, pp. 15–41, 2011.
[10] R. G. Carter, 'R.F. power generation', in R. Bailey, ed., *Proceedings of the CERN Accelerator School 'RF for Accelerators', 8–17 June 2010, Ebeltoft, Denmark*, Geneva: CERN, pp. 173–207, 2011.
[11] R. G. Carter, 'Review of RF power sources for particle accelerators', in S. Turner, ed., *CERN Accelerator School: RF Engineering for Particle Accelerators,* vol. 1. Geneva: CERN, pp. 269–300, 1992.
[12] Siemens, *Transmitting Tubes Data Book 1986/87*. Siemens AG, 1986.
[13] W. P. Bennett, 'Large-power-tube design, processing and testing', in R. S. Burnap, ed., *Electron Tube Design*. Harrison, NJ: RCA, pp. 763–791, 1962.
[14] N. O. Sokal, 'RF power amplifiers, classes A through S-how they operate, and when to use each', in *Electronics Industries Forum of New England, 1997. Professional Program Proceedings*, pp. 179–252, 1997.
[15] F. H. Raab *et al.*, 'Power amplifiers and transmitters for RF and microwave', *IEEE Transactions on Microwave Theory and Techniques,* vol. 50, pp. 814–826, 2002.
[16] S. Kostic *et al.*, 'Increasing efficiency and output power of HFHPTA by injection of two harmonics', *IEEE Transactions on Broadcasting,* vol. 47, pp. 32–37, 2001.
[17] EIMAC, *Care and Feeding of Power Grid Tubes*, 5th ed. San Carlos, CA: CPI Inc. Eimac Division, 2003.
[18] P. Gerlach and C. Kalfon, 'New technologies in power-grid tubes and their impact in high-power UHF-TV operation', *Proceedings of the IEEE,* vol. 70, pp. 1335–1345, 1982.
[19] R. H. Rheaume, 'A coaxial power triode for 50-kW output up to 110 Mc', *Proceedings of the IRE,* vol. 40, pp. 1033–1037, 1952.
[20] Machlett Laboratories Inc., 'ML-5681 Data Sheet', 1962.

[21] C. S. Bull, 'The alignment of grids in thermionic valves', *Journal of the Institution of Electrical Engineers – Part III: Radio and Communication Engineering,* vol. 92, pp. 86–92, 1945.

[22] S. G. McNees, 'Two megawatts RF power tetrode', *IEEE Transactions on Nuclear Science,* vol. NS-20, pp. 422–423, 1973.

[23] Y. Kusunose, 'Calculation of characteristics and the design of triodes', *Proceedings of the Institute of Radio Engineers,* vol. 17, pp. 1706–1749, 1929.

[24] L. C. Scholz, 'Calculation of fields and currents', in R.S. Burnap, ed., *Electron Tube Design*. Harrison, NJ: Radio Corporation of America, pp. 159–201, 1962.

[25] G. Faillon, 'Technical and industrial overview of RF and microwave tubes for fusion', *Fusion Engineering and Design,* vol. 46, pp. 371–381, 1999.

[26] G. Clerc *et al.*, 'A new generation of gridded tubes for higher power and higher frequencies', in *Particle Accelerator Conference*, Vancouver, BC, pp. 2899–2901, 1997.

[27] R. G. Carter, 'R.F. Generation', in R. Bailey, ed., *High Power Hadron Machines: CERN-2013-001*. Geneva: CERN, pp. 45–69, 2013.

[28] W. Herdrich and H. P. Kindermann, 'RF power amplifier for the CERN SPS operating as LEP injector', *IEEE Transactions on Nuclear Science,* vol. NS-32, pp. 2794–2796, October 1985.

[29] EIMAC, 'Tube performance computer for RF amplifiers', CPI Inc., 1969.

[30] EEV, 'Preamble – Tetrodes'. Chelmsford, UK: Teledyne e2v, 1976.

[31] H. P. Kindermann *et al.*, 'The RF power plant of the SPS', *IEEE Transactions on Nuclear Science,* vol. NS-30, pp. 3414–3416, August 1983.

[32] T. Fujisawa *et al.*, 'The radiofrequency system of the RIKEN ring cyclotron', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 292, pp. 1–11, 1990.

[33] R. G. Carter, 'Simple model of an inductive output tube', *IEEE Transactions on Electron Devices,* vol. ED-57, pp. 720–725, 2010.

[34] R. S. Symons, 'The constant efficiency amplifier', in *Annual Broadcasting Engineering Conference*, pp. 523–530, 1997.

[35] H. Bohlen *et al.*, 'Inductive output tubes for particle accelerators', in *European Particle Accelerator Conference*, Lucerne, Switzerland, pp. 1111–1113, 2004.

[36] H. Bohlen *et al.*, 'IOT RF power sources for pulsed and cw linacs', in *LINAC 2004*, Luebeck, Germany, pp. 574–577, 2004.

[37] R. Heppinstall and G. Clayworth, 'The inductive output tube', *EBU Technical Review,* pp. 43–52, 1997.

[38] R. Heppinstall and G. Clayworth, 'The inductive output tube – a modern UHF amplifier for the terrestrial television transmitter', *GEC Review,* vol. 13, pp. 76–85, 1998.

[39] E. G. Zaidman and M. A. Kodis, 'Emission gated device issues', *IEEE Transactions on Electron Devices,* vol. 38, pp. 2221–2228, 1991.

[40] K. Nguyen *et al.*, 'Analysis of the 425-MHz klystrode', *IEEE Transactions on Electron Devices,* vol. 38, pp. 2212–2220, 1991.

[41] R. True, 'Calculation and design of grids in Pierce guns', in *1989 International Electron Devices Meeting, IEDM '89*, pp. 215–218, 1989.

[42] J. Welch, 'Results from the SLAC Lasertron', *NASA STI/Recon Technical Report N,* vol. 88, p. 26647, 1988.

[43] M. Shrader *et al.*, 'Pre-bunched beam devices: efficient sources of UHF and microwave power', in *International Electron Devices Meeting*, pp. 342–345, 1985.

[44] E. Garwin *et al.*, 'An experimental program to build a multimegawatt lasertron for super linear colliders', *IEEE Transactions on Nuclear Science,* vol. 32, pp. 2906–2908, 1985.

[45] M. A. Kodis *et al.*, 'Operation and optimization of gated field emission arrays in inductive output amplifiers', *IEEE Transactions on Plasma Science,* vol. 24, pp. 970–981, 1996.

[46] C. Wilsen *et al.*, 'Gridless IOT for accelerator applications', in *IEEE Particle Accelerator Conference*, pp. 2556–2558, 2007.

[47] C. Yates *et al.*, 'Performance results of a 1.3 GHz 100 kW CW IOT', in *IEEE International Vacuum Electronics Conference*, Monterey, CA, pp. 107–108, 2014.

[48] D. H. Preist and M. B. Shrader, 'A high-power Klystrode with potential for space application', *IEEE Transactions on Electron Devices,* vol. 38, pp. 2205–2211, 1991.

[49] P. Schütt *et al.*, 'Computer simulations of inductive output tubes', in *European Particle Accelerator Conference*, pp. 1233–1235, 1998.

[50] H. P. Freund *et al.*, 'Time-domain simulation of inductive output tubes', *IEEE Transactions on Plasma Science,* vol. 35, pp. 1081–1088, 2007.

[51] D. Preist, 'Prospects for very high power high efficiency RF generators', in *Microwave and Optical Generation and Amplification*, Cambridge, UK, pp. 146–150, 1966.

[52] H. Bohlen *et al.*, 'Ways to decrease high-voltage requirements in high-power microwave tubes', in *Displays and Vacuum Electronics*, Garmisch-Partenkirchen, Germany, pp. 269–273, 1998.

[53] H. Bohlen, 'Inductive output tubes: history and present status', in *Displays and Vacuum Electronics*, Garmisch-Partenkirchen, Germany, pp. 93–98, 2001.

[54] H. S. Kim and H. S. Uhm, 'Analytical calculations and comparison with numerical data for annular klystrode', *IEEE Transactions on Plasma Science,* vol. 29, pp. 875–880, 2001.

[55] L. Ives *et al.*, '12.2: A 350 MHz, 200 kW CW, multiple beam IOT', in *IEEE International Vacuum Electronics Conference (IVEC)*, Monterey, CA, pp. 215–216, 2010.

[56] E. Wright *et al.*, 'High-power multiple-beam IOT design', in *IEEE Thirteenth International Vacuum Electronics Conference (IVEC)*, Monterey, CA, pp. 143–144, 2012.

[57] C. Yates *et al.*, 'Performance characteristics of an MSDC IOT amplifier', *IEEE Transactions on Electron Devices,* vol. 48, pp. 116–121, 2001.

[58] R. Symons *et al.*, 'Prototype constant-efficiency amplifiers', *IEEE Transactions on Broadcasting,* vol. 47, pp. 147–152, 2001.

[59] J. Orrett *et al.*, 'IOT testing at the ERLP', in *European Particle Accelerator Conference,* Edinburgh, Scotland, pp. 1382–1384, 2006.

[60] A. Wheelhouse, 'Improved 1.3 GHz inductive output tube for particle accelerators', in *European Particle Accelerator Conference*, Edinburgh, Scotland, pp. 1373–1375, 2006.

13 Klystrons

13.1 Introduction

We saw in the previous chapter that the performance of an IOT is limited in both power and frequency by the use of a control grid to modulate the electron beam. At higher frequencies and powers it is necessary to modulate the beam in some other way. In the klystron this is achieved by velocity modulation of an electron beam by a cavity resonator excited by an external RF source, as described in Section 11.2.1. When an output cavity, tuned to the signal frequency, is placed in the region where the beam is bunched the result is the simple two-cavity klystron illustrated in Figure 13.1. Radio frequency current is induced in the second cavity in exactly the same way as in an IOT. At resonance the cavity presents a resistive impedance to the induced current, and the phase of the field across the gap is in anti-phase with the RF beam current. Electrons which cross the gap within ±90° of the bunch centre are retarded and give up energy to the field of the cavity. Since more electrons cross the second gap during the retarding phase, than the accelerating phase, there is a net transfer of energy to the RF field of the cavity. Thus the klystron operates as an amplifier by converting some of the DC power input into RF power in the output cavity.

In practice the gain and efficiency of a two-cavity klystron are too low to be of practical value. It is therefore usual to add further cavity resonators in order to increase the gain, efficiency and bandwidth of the tube. Figure 13.2 shows the arrangement of a multi-cavity klystron. The electron beam is formed by a diode electron gun whose perveance is typically in the range $0.5 - 2.0\ \mu\mathrm{A\,V}^{-1.5}$. The function of all the cavities, except the last, is to form electron bunches from which RF power can be extracted by the output cavity. The first and last cavities are tuned close to the centre frequency and have low Q factors determined largely by the coupling to the external waveguides. The intermediate (idler) cavities normally have high Q and are tuned to optimise the performance of the tube. The long electron beam is confined by an axial magnetic field to avoid interception of electrons on the walls of the drift tube. Radio frequency power passes into and out of the vacuum envelope through ceramic windows, and the spent electrons are collected by a collector in exactly the same way as in an IOT.
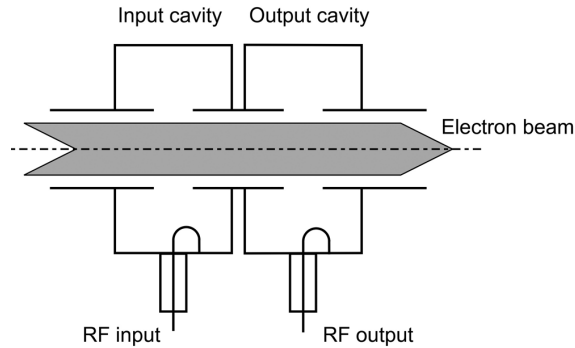
Figure 13.1: Arrangement of a two-cavity klystron.
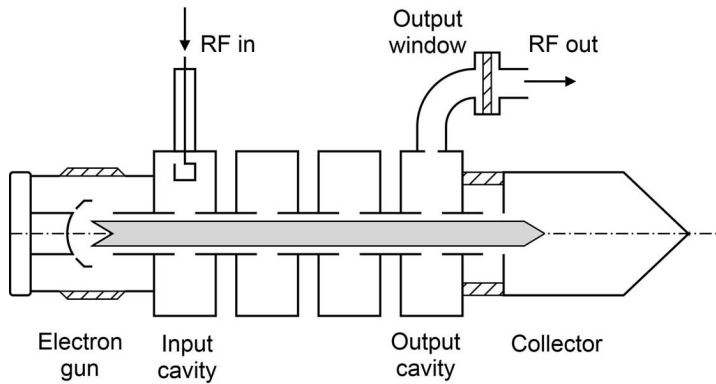
Figure 13.2: Arrangement of a multi-cavity klystron.

The tube illustrated in Figure 13.2, whose cavities are wholly within the vacuum envelope, is known as an integral cavity klystron. An alternative arrangement is to insert a cylindrical ceramic window in the cavity as shown in Figure 13.3 so that the outer portion is at atmospheric pressure. Tubes with this kind of cavity are known as external cavity klystrons [1–3]. The external part of the cavity is made in two halves which are bolted around the tube body. A good electrical connection at the joint is ensured by the use of metallic spring fingers. The cavity can be tuned by sliding walls as shown. The tuning and coupling arrangements are made simpler by the use of external cavities because they are outside the vacuum envelope. A tuning range of 50–60% is possible with this arrangement. The advantages of external cavity klystrons are: easier tuning; easier coupling; and cheaper replacement tubes. However, they are limited to powers under 100 kW and frequencies less than 1 GHz by voltage breakdown in the part of the cavity at atmospheric pressure, and by the need to ensure good electrical contact at the junction between the external cavity and the tube body. The window is in line of sight with the electron beam and may be damaged by direct electron bombardment or a multipactor discharge (see Section
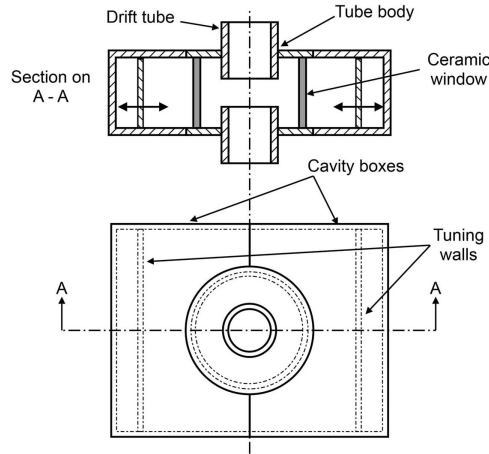
Figure 13.3: Arrangement of an external klystron cavity.

18.8). The task of replacing a failed tube is made more difficult because of the need to transfer the external cavities to the new tube, but they should then need little or no retuning.

This chapter considers the theory and design of klystrons. Small signal models are based on the space-charge wave theory in Section 11.4, while large signal models represent the electron beam by discs, or rings, of charge as described in Section 11.8. Klystrons fall into two main classes:

- Broadband tubes have bandwidths up to 10% at frequencies from UHF to around 18 GHz. Their applications include radar, UHF television transmitters, satellite ground stations and tropospheric scatter systems. [1, 4].
- Narrow band, high efficiency, tubes are used chiefly in particle accelerators and plasma heating and in nuclear fusion experiments at frequencies below 4 GHz [5, 6].

The design of tubes in these classes is discussed in Section 13.4. There is a brief review of the principles of multiple beam and sheet beam klystrons. Further information can be found in [3, 7–10]

13.2   Small-Signal Klystron Theory

Under small-signal conditions a klystron can be modelled using the space-charge wave theory introduced in Chapter 11. The cavities are represented by equivalent circuits, as shown in Figure 13.4. The coupling factors, and beam loading admittances, of the interaction gaps can be calculated using (11.36), (11.38), and (11.100). In practical cases they can be determined by cold and hot test measurements, or by numerical modelling [11, 12].
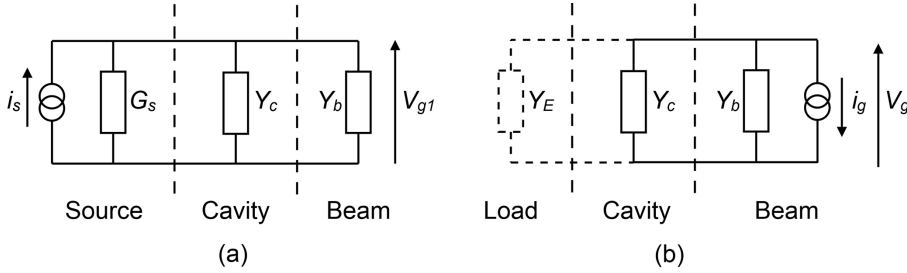
Figure 13.4: Klystron equivalent circuits: (a) input cavity, and (b) idler and output cavities.

The beam loading conductance (G_b) is typically around 0.15 times the DC beam conductance (G_0) when $g = a$, and the gap transit angle is 1 radian (see Figure 11.16). Thus, for a non-relativistic beam, we can write

$$G_b \sim 0.15 K \sqrt{V_a},\tag{13.1}$$

where K is the perveance and V_a is the beam voltage. If $K = 0.5\,\mu\text{A V}^{-1.5}$ and $V_a = 10\,\text{kV}$, both of which are at the lower end of the range of typical values, then $G_b \sim 7.5\,\mu\text{S}$. The R/Q of a typical re-entrant cavity is of the order of $100\,\Omega$. The practical unloaded Q is around 6000 at 1 GHz, and decreases by a factor of 4 for every decade increase in frequency, if the surfaces of the cavity are smooth (see Chapter 3). Thus the shunt conductances of the cavities range from around 1 μS at 1 GHz, to 10 μS, or more, at 30 GHz. This is comparable with the beam loading conductance. The reactive component of the beam loading is normally negligible [13] but, if necessary, it can be compensated for by slight adjustments to the tuning of the cavities.

13.2.1   Input Cavity

The equivalent circuit of the input cavity is shown in Figure 13.4(a). The source current and conductance are transformed from those for the external source by the input coupler (see Section 3.6). Then the input gap voltage is

$$V_g = \frac{i_s}{G_s + Y_c + Y_b},\tag{13.2}$$

where the cavity admittance is obtained from (3.5) as

$$Y_c = G_c\left(1 + jQ_U\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)\right),\tag{13.3}$$

where G_c is the shunt conductance. The forward power in the input waveguide is

$$P_{in} = \frac{i_s^2}{8G_s}.\tag{13.4}$$

Then, from (13.2) and (13.4) the input gap voltage is

$$\left|V_{g1}\right|^2 = \frac{8G_s}{\left|G_s + Y_c + Y_b\right|^2} P_{in}. \tag{13.5}$$

If the input is matched at resonance then $G_s \simeq G_b + G_c$, where G_b is the real part of Y_b, and the power transferred to the beam is

$$P_1 = \frac{1}{2}\left|V_g\right|^2 G_b = \frac{4G_s G_b}{\left|2G_s + jQ_U G_c \dfrac{2\Delta\omega}{\omega_0}\right|^2} P_{in}, \tag{13.6}$$

where $\Delta\omega = \omega - \omega_0$. But $Q_E G_s = Q_U G_c$ so that (13.6) becomes

$$P_1 = \frac{1}{\left(1 + Q_E\left(\Delta\omega/\omega\right)\right)^2} \cdot \frac{G_b}{G_s} P_{in} \tag{13.7}$$

This equation shows that the gain of the input cavity is maximum at resonance and that it is reduced by losses in the cavity. Away from resonance the input is no longer matched and the gain is reduced by reflected power. The modulation of the beam by the input gap can be expressed in terms of the space-charge wave amplitudes using (11.91)

$$\begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} M_+ \\ M_- \end{bmatrix} V_{g1}. \tag{13.8}$$

The transformation of the RF beam voltage and current by a drift length d is given by (11.110)

$$\begin{bmatrix} V_{2+} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} \exp(-j\beta_+ d) & 0 \\ 0 & \exp(-j\beta_- d) \end{bmatrix} \begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix} = [D]\begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix}. \tag{13.9}$$

13.2.2   Idler Cavities

Figure 13.4(b) shows the equivalent circuit for an idler cavity including the possibility of an external load (though this is unusual). The beam is represented by a current source having magnitude i_g and admittance equal to the beam admittance (Y_b). In small-signal analysis it is usual to use the beam-loading admittance of an unmodulated beam in all cavities, though this is not strictly correct [14, 15]. If the imaginary part of Y_b is small compared with the real part (G_b) then Y_b can be replaced by G_b without serious error. The cavity extracts power from the modulated beam and uses it to add fresh modulation to the beam. Because the cavity impedance is high, a large gap voltage is produced by a small induced current. If the cavity losses are neglected, the kinetic power in the beam leaving the cavity is equal to that entering. Analysis of the interaction between the beam and the cavity is represented by the transfer matrix in (11.109)

$$\begin{bmatrix} V_{2+} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} 1 - \dfrac{Y_e}{2Y_T} M_+^2 & -\dfrac{Y_e}{2Y_T} M_+ M_- \\ -\dfrac{Y_e}{2Y_T} M_+ M_- & 1 - \dfrac{Y_e}{2Y_T} M_-^2 \end{bmatrix} \begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix} = [C] \begin{bmatrix} V_{1+} \\ V_{1-} \end{bmatrix}, \tag{13.10}$$

where

$$Y_T = Y_c + Y_b + Y_e \tag{13.11}$$

is the admittance of the cavity, the beam, and the external circuit, in parallel with one another. If the klystron has N cavities separated by $(N-1)$ drift spaces the voltage and current modulation at the output gap is given by

$$\begin{bmatrix} V_{N+} \\ V_{N-} \end{bmatrix} = \frac{1}{2} [D_{N-1}][C_{N-1}] \dots [C_2][D_1] \begin{bmatrix} M_+ \\ M_- \end{bmatrix} V_{g1}, \tag{13.12}$$

where the properties of the gaps and of the drift spaces may differ from one another.

13.2.3   Output Cavity

The voltage induced in the output gap is given by (11.108) as

$$V_{gN} = -\frac{Y_e}{Y_T} [M_+ \quad M_-] \begin{bmatrix} V_{N+} \\ V_{N-} \end{bmatrix}. \tag{13.13}$$

The output power is

$$P_{out} = \frac{1}{2} \mathrm{Re}(Y_E) |V_{gN}|^2 \tag{13.14}$$

so that, finally, the power gain in decibels is

$$G_{dB} = 10 \log\left(\frac{P_{out}}{P_{in}}\right). \tag{13.15}$$

This method of calculation is incorporated in Worksheet 13.1.

An alternative formulation using the transmission line representation of the beam employs the cavity transfer matrix in (11.119) [13]

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 - \dfrac{G_b}{Y_T} & -\dfrac{M^2}{Y_T} \\ -\dfrac{G_b^2}{M^2 Y_T} & 1 - \dfrac{G_b}{Y_T} \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}. \tag{13.16}$$

The drift matrix is

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \exp(-j\beta_e z) \begin{bmatrix} \cos\beta_q z & jZ_e \sin\beta_q z \\ jY_e \sin\beta_q z & \cos\beta_q z \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}, \tag{13.17}$$

where time variation as $\exp(j\omega t)$ has been assumed. At the input, from (11.114)

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} M \\ G_b/M \end{bmatrix} V_{g1} \qquad (13.18)$$

and, at the output, from (11.117)

$$V_{gN} = \frac{1}{Y_{TN}} \begin{bmatrix} \dfrac{G_b}{M} & M \end{bmatrix} \begin{bmatrix} V_N \\ I_N \end{bmatrix}. \qquad (13.19)$$

The small-signal can be computed by cascading the transfer matrices as before.

The voltage at the output combines contributions from the individual cavities so that there is a peak of voltage at the resonant frequency of each cavity. In addition the phasor addition of the contributions of the cavities produces maxima and minima as the voltages add and subtract in phase. It can be shown that the complex gain can be written

$$G(s) = A \frac{\displaystyle\prod_{n=2}^{N-1}(s - z_n)}{\displaystyle\prod_{n=1}^{N}(s - p_n)}, \qquad (13.20)$$

where s is the complex frequency and A is a constant [3, 7]. This function has poles in the complex frequency plane when $s = p_n$ and zeroes when $s = z_n$.

The way in which the poles and zeroes arise can be illustrated by a simple approximate analysis based on cascaded matrices, as described above. The tube is considered to be made up of series of stages, each of which comprises one cavity with the following drift length. As the modulation grows along the length of the tube the RF current in the beam increases while the electron velocity remains approximately constant. The RF current in the beam entering a stage drives the gap voltage of the cavity. The velocity modulation imposed by the gap is transformed into additional current modulation by the drift region. Thus, to a first approximation, the current gain in a stage can be computed by ignoring the velocity modulation at its start. This analysis assumes that the contribution to the gain from cavities which are not adjacent is negligible. The overall gain of the tube is found by summing the gains of the stages in decibels.

The interaction with a gap derived from (13.16) is

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} -\dfrac{M^2}{Y_T} \\ 1 - \dfrac{G_b}{Y_T} \end{bmatrix} I_1, \qquad (13.21)$$

where I_1 is the RF beam current before the gap. The modulation of the beam after a drift length d is found by pre-multiplying (13.21) by the drift matrix in (13.17) so that the RF current at the end of a drift region of length d is given by

$$I_3 = \exp\left(-j\beta_e d\right)\left\{-j\frac{Y_e}{Y_T}M^2\sin\left(\beta_q d\right)+\left(1-\frac{G_b}{Y_T}\right)\cos\left(\beta_q d\right)\right\}I_1. \qquad (13.22)$$

Then the current gain in decibels is

$$G_{dB} = 10\log\left\{\left|1-\frac{G_b}{Y_T}\right|^2\cos^2\left(\beta_q d\right)+\left|\frac{Y_e}{Y_T}M^2\right|^2\sin^2\left(\beta_q d\right)\right\}. \qquad (13.23)$$

If there is no added loss, or external loading, then $Y_T \approx G_b$ at the resonant frequency of the cavity. Then the second term is dominant and the gain contributed by the stage is approximately

$$G_{dB} \simeq 20\log\left(\frac{M^2 Y_e}{Y_T}\sin\left(\beta_q d\right)\right). \qquad (13.24)$$

For typical values of the parameters this has a value in the range 20 to 40 dB. When the signal frequency is far from resonance Y_T is dominated by the cavity susceptance so that

$$Y_T \simeq j\frac{Q_U}{R_c}\frac{2\Delta\omega}{\omega_0}. \qquad (13.25)$$

If we now assume that $|Y_T| \gg G_b$ then (13.22) becomes

$$\left|\frac{I_3}{I_1}\right| = \cos\left(\beta_q d\right)-\frac{Y_e}{2}\frac{R_c}{Q_U}\frac{\omega_0}{\Delta\omega}M^2\sin\left(\beta_q d\right). \qquad (13.26)$$

This expression is zero when

$$\frac{\Delta\omega}{\omega_0} = \frac{Y_e}{2}\cdot\frac{R_c}{Q_U}M^2\tan\left(\beta_q d\right), \qquad (13.27)$$

which shows that, when $\beta_q d < 90°$ the frequency of the zero lies above the resonant frequency of the cavity, and that it increases as $\beta_q d$ increases. The frequency of the zero lies below the resonant frequency of the cavity when $\beta_q d > 90°$. These general conclusions are confirmed by more accurate calculations as illustrated, for typical parameters, in Figure 13.5. The maxima and minima of the curves correspond to the poles and zeroes of the transfer function of the stage in the complex frequency plane. The transfer functions of the input and output cavities, which do not have zeroes, only contribute poles to the overall transfer function. Hence, for a klystron with N cavities there are N poles and $N-2$ zeroes. The zeroes produce dips in the gain, and an important part of the design of a klystron is the selection of the cavity and drift-tube parameters in such a way that the variation in gain within the band is minimised. The design of broad-band klystrons is discussed in Section 13.4.1.
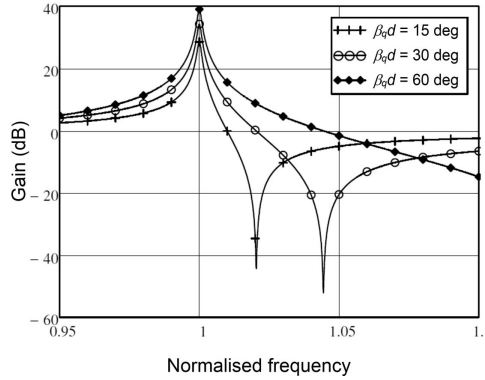
Figure 13.5: Gain of a klystron stage (cavity plus drift length) as a function of frequency and drift length $\left(R/Q = 100, Q = 4000, M = 0.79, Gb = 10\,\mu\text{S}, Y_e = 2\,\text{mS} \right)$.

13.2.4   Simplified Small-Signal Model

A simplified small-signal model of a klystron can be constructed by assuming that the contribution to the RF beam current made by a gap is negligible so that (13.18) becomes

$$V_1 = MV_{g1}. \tag{13.28}$$

The RF beam current is given as a function of position by (11.86)

$$I_2(z) = jY_e M_1 V_{g1} \sin(\beta_q z) \exp(-j\beta_e z). \tag{13.29}$$

and the voltage induced in a cavity by this current is

$$V_{g2} = -\frac{M_2 I_2}{Y_{T2}} = -j\frac{Y_e}{Y_{T2}} M_1 M_2 V_{g1} \sin(\beta_q d)\exp(-j\beta_e d), \tag{13.30}$$

where d is the drift length between the cavities. The voltage induced in each cavity excites waves which, in turn, induce voltages in each of the subsequent cavities, and so on. Because the analysis is linear, the final voltage induced in the output cavity is the phasor sum of the voltages induced by all the waves which have been excited. This can be thought of in terms of a number of signal paths in parallel, each of which has its own gain and phase shift as a function of frequency. These paths represent all the possible connections between the input and the output. Thus for a five-cavity tube, for example, we find: one path which connects the input and the output directly; three which take in one intermediate cavity; three which take in different pairs of cavities; and one which takes in all three intermediate cavities – a total of eight paths [7, 16, 17]. This model is implemented in Worksheet 13.2.

13.2.5   Overall Performance

Figure 13.6 shows a comparison between the calculated and measured small-signal gain for a 3kW klystron for use in satellite ground stations. The cavities of this tube are tuneable and their exact frequencies in the experimental tube are not known. The frequencies of the experimental results were scaled by 1.25% to match the calculated results. The figure shows results from Worksheets 13.1 and 13.2. The gain was found to be quite sensitive to the beam radius. In the example shown, the best results were obtained when the radius was reduced by about 10% relative to the theoretical value. It is often difficult to obtain a reliable value for the beam radius, especially if the current density is not uniform, though good results are usually given by using the equilibrium beam radius in the calculations. This tube has four stages (five cavities) so that the average small signal gain is about 11 dB per stage. The cavities are tuned for broad-band operation (see Section 13.4.1). If the cavities in the model are all tuned to the same frequency the gain per stage at the band centre is increased to 19 dB. However some tubes suffer from feedback oscillations if they are synchronously tuned so it does not follow that this tube could be operated with synchronous tuning.

13.3   Large-Signal Behaviour of Klystrons

Unlike an IOT, the DC power input to a klystron is constant, regardless of the level of the RF output power. Therefore it is usually desirable to operate a tube with high RF power output to obtain the best conversion efficiency. Under large-signal conditions the behaviour of a tube becomes non-linear and the linearised theory in the previous section is not valid. It is therefore necessary to use a large-signal model as described in Section 11.8. Details of time-step and distance-step models, and of the application of PIC codes, can be found in the literature [18–20]. The main features of the large-signal behaviour of a klystron can be illustrated using a simple
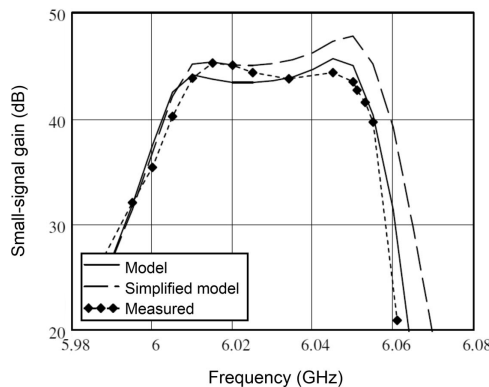
Figure 13.6: Comparison between calculated and measured small-signal gain of a klystron.

time-step disc model (see Worksheet 13.3). The power transferred to the output gap tends to be an over-estimate when computed by a one-dimensional model, because the radial variations of the space-charge field and the field of the gap, and the radial motion of electrons, have been neglected [21, 22].

We saw in Section 11.8.4 that the RF power extracted from a bunched beam by the output cavity increases with the first harmonic of the beam current (I_1). In the small-signal regime, the amplitude of I_1 at a fixed plane varies linearly with the normalised input gap voltage (see (13.29)). When the RF power in the beam is small, the output gap voltage is small compared with V_a so that the electrons cross the gap with average velocities close to u_0. The gap coupling factor is then independent of the level of modulation, and the output power is proportional to the input power. As the RF power in the beam increases the behaviour of the tube becomes non-linear because:

- The amplitude of the fundamental RF current is not linearly dependent on the input gap voltage.
- The position at which the amplitude of the current is greatest varies with the drive level (see Figure 11.26).
- The effective coupling factor of the output gap decreases because the electrons at the centre of the bunch slow down as they pass through the gap.

The plot of P_{out} against P_{in} is then not linear at high drive levels but *saturates* as shown in Figure 13.7. The difference between the gain at saturation and the small-signal gain, known as the *gain compression* of the tube, is typically in the region of 5 to 7 dB. The output power of a klystron therefore depends on both the effectiveness with which the beam is bunched, and on the power conversion efficiency in the output gap. When the frequency of the input signal is varied within the working band of the tube it is found that the saturated output power is constant to first order. However, the saturated gain and hence the input power required to saturate the tube, vary with frequency [13].
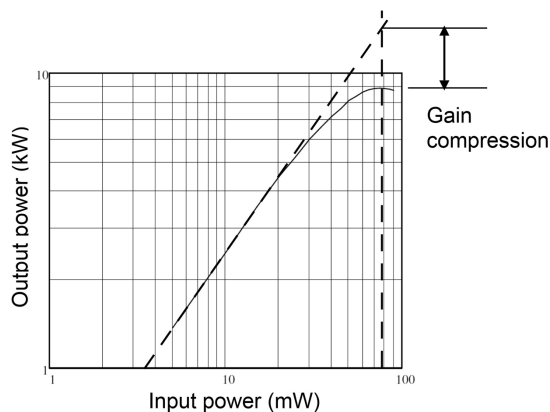
Figure 13.7: Typical non-linear power transfer characteristics of a klystron.

13.3.1 Klystron Sections

It is convenient to consider the design of a multi-cavity klystron, such as that illustrated in Figure 13.2, in terms of three sections, as shown in Figure 13.8 [23]. Each section is made up of one, or more, stages each comprising a cavity with the following drift length. The cavities can be represented by the equivalent circuits shown in Figure 13.4 if the gap coupling factors are adjusted for large-signal effects where necessary. The purpose of the first section is to provide sufficient gain to produce a strongly bunched beam over the working bandwidth of the tube. The second section increases the fundamental RF current by compressing the bunches to enable the maximum power to be extracted from them by the output section. The factors influencing the design of each section are discussed below.

The objective of the bunching process is that the final value of $|I_1/I_0|$ should be as great as possible, with minimum velocity spread and few crossing trajectories. It has been suggested [24] that a useful figure of merit for bunching is

$$ F_1 = \frac{1}{2}\left|\frac{I_1}{I_0}\right|\frac{u_{\min}}{u_0}, \qquad (13.31) $$

where $u_{\min}$ is the velocity of the slowest electron within the bunch. The maximum possible value of this figure of merit is $F_1 = 1$, and it decreases sharply when electron crossovers occur. It has been found to be a good predictor of efficiency at saturation and to be useful for the optimisation of the bunching [13, 25]. A definition based on the minimum electron energy in the input bunch could also be justified. However, the velocity of the slowest electron is not a good measure of the velocity spread in the bunch as a whole [26]. An improved figure of merit is

$$ F_2 = \frac{1}{2}\left|\frac{I_1}{I_0}\right|\left(1 - \frac{\Delta u}{u_0}\right), \qquad (13.32) $$

where Δu is the root mean square of the velocity spread in the bunch. The application of this figure of merit to the design of klystrons is discussed in Section 13.3.3.
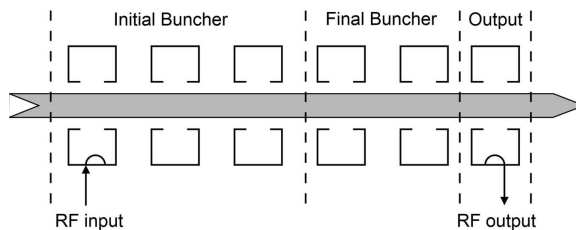
Figure 13.8: The sections of a klystron.

13.3.2 Initial Bunching Section

The input cavity is normally working in the small-signal regime so that the analysis in Section 13.2.1 is valid. The resonant frequency of the cavity is often close to the centre frequency of the tube but it may be elsewhere in the working frequency band in a broadband klystron (see Section 13.4.1). The normalised voltage in the input gap can be estimated by recalling that the RF power transferred to the beam is given by (11.98)

$$P_g = \frac{1}{2} G_b \left| V_g \right|^2, \tag{13.33}$$

and that the output power of the tube is

$$P_{out} = \eta_e I_0 V_0. \tag{13.34}$$

Thus the gain is

$$G = \frac{2 \eta_e I_0 V_0}{G_b \left| V_g \right|^2} = \frac{2 \eta_e G_0}{G_b} \left| \frac{V_0}{V_g} \right|^2. \tag{13.35}$$

Now, typically for a broadband klystron, $\eta_e \sim 0.5$, $G_b / G_0 \sim 0.15$ and the power gain is 40 dB so that $G = 10^4$. Substituting these figures into (13.35) we obtain the normalised input gap voltage

$$\frac{M V_g}{V_0} \sim 0.02, \tag{13.36}$$

where it has been assumed that, typically, $M \sim 0.8$.

Klystron input cavities may be affected by multipactor discharge (see Section 18.8), especially between the tips of the drift tubes. These discharges are usually of low power but they add loading to the cavity and introduce noise [27–29]. The problem may be reduced or eliminated by coating the surface with a material having a low secondary electron emission coefficient, by castellation of the drift-tube tips, or by changing the surface topology in some other way [3].

The remaining cavities in the initial bunching section make a major contribution to the gain of the tube. Their parameters are chosen to equalise the gain over the bandwidth of the tube, and to create well-formed bunches for compression by the following section. Although small-signal theory is sometimes used to model bunch formation we shall see that the role of harmonic modulation of the beam can be important. The resonant frequencies of the cavities generally lie within the bandwidth of the tube. Their Q factors may be adjusted by the addition of lossy coatings or by coupling to external loads [30]. In this section the RF power in the beam is small compared with the DC beam power. Thus, the power dissipated in the cavity losses is small and has a negligible effect on the overall efficiency of the tube. The gap voltage induced in a cavity is determined by the RF current in the input beam

so that, using small-signal theory, the amplitude of the velocity modulation added to the beam is

$$|V_1| = \frac{M^2 I_1}{\left| G_b + G_c + j\dfrac{Q_U}{R_c} \cdot \dfrac{2\Delta\omega}{\omega_0} \right|}, \tag{13.37}$$

where ω_0 is the resonant frequency of the cavity. Taking $R_c/Q_0 = 100\ \Omega$ as a typical figure, it can be seen that the imaginary term in the denominator becomes comparable with the beam conductance estimated from (13.1) when the frequency deviation is of the order of 1%. For greater frequency deviations the imaginary term is dominant and (13.37) can be written approximately as

$$\left|\frac{V_1}{V_0}\right| \simeq G_0 \left(M^2 \frac{R_c}{Q_U} \right) \left(\frac{\omega_0}{2\Delta\omega} \right) \frac{I_1}{I_0}. \tag{13.38}$$

Thus, if large bandwidth is required, it is desirable to maximise the beam conductance by choosing a high beam perveance, and to maximise $M^2 \left(R_c/Q_0 \right)$ [13, 31]. This parameter depends essentially on the choice of the normalised radius of the drift tube and on the gap transit angle. It has been found that the optimum transit angle is 1.4 radians [13].

13.3.3   Final Bunching Section

The cavities in this section are tuned to frequencies above the working band of the tube so that the phase of the gap voltage leads the phase of the RF beam current by just under 90°. The centre of the bunch crosses the gap when the RF field is zero and the bunch is compressed. The extent to which the bunch can be compressed in this way depends upon both the form of the input bunch and the space-charge density in the beam. The influence of these factors on bunch compression by a single cavity can be explored using the disk model in Worksheet 13.4. Figure 13.9 shows, as an example, the compression of bunches in a 25 kV, 1 μPerv, beam when the current waveform in the input beam is defined by (11.175) with $n = 1.5$ so that $I_1/I_0 = 1.2$. The depression of the potential by the space charge of the bunches causes some of the kinetic energy in the beam to be converted to potential energy. The initial electron velocity has been reduced to account for this. The figure of merit F_2 is greatest, at a distance $z/\lambda_e = 0.845$ from the centre of the gap, when the normalised gap voltage $MV_g/V_a = 0.33$. The normalised RF current at this plane is $I_1/I_0 = 1.63$. Figure 13.9(b) shows the computed current waveform with the ideal current waveform for $n = 4$ for comparison. The bunching can also be represented by plotting the relative phase at which sample electrons reach a plane against their initial phase, as shown in Figure 13.9(c) [32]. These results were obtained using the parameters $\beta_e a = 0.5$, $\beta_e b = 0.3$, $\beta_e g = 0.5$. When the case in Figure 13.9 was recomputed with all the parameter values doubled
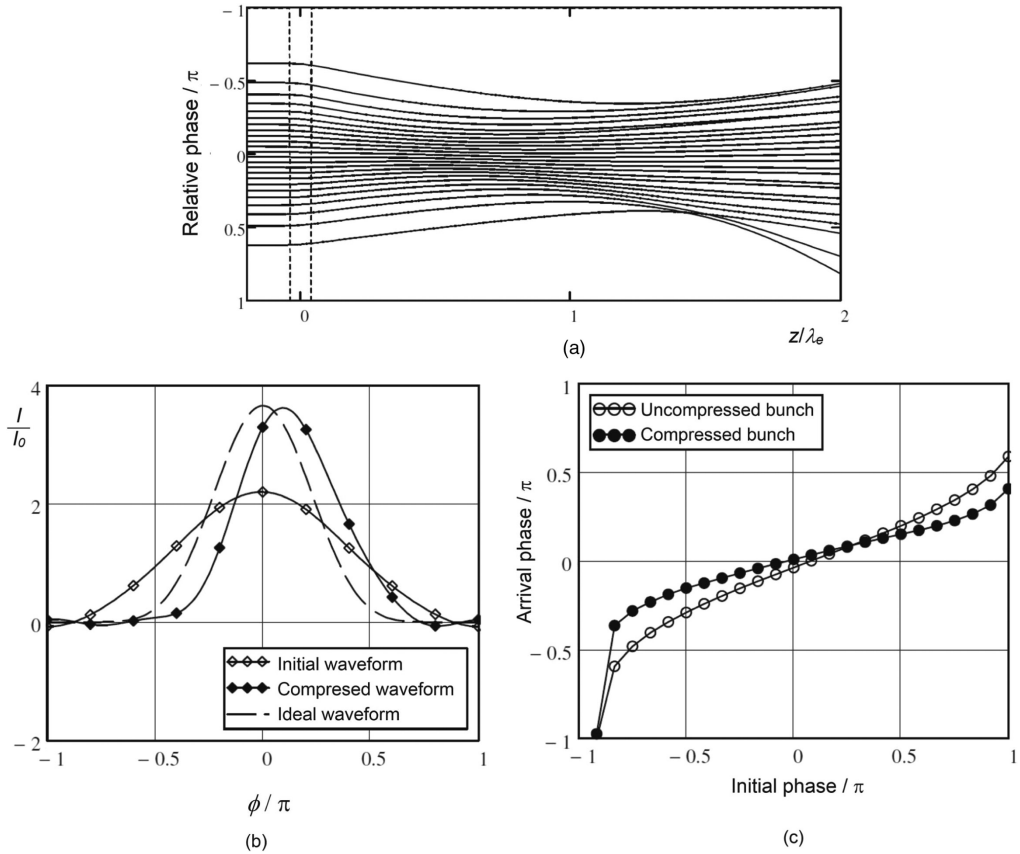
Figure 13.9: Compression of a bunch by a cavity: (a) electron trajectories (b) current waveforms, and (c) phases of sample electrons.

the maximum normalised RF current was 1.78 at $z/\lambda_e = 1.005$. Thus the detailed results depend upon the choice of parameters but it is to be expected that the general trends will be the same.

The effects of perveance and of the initial bunching of a beam having the same parameters are shown in Figure 13.10. Figure 13.10(a) shows that the normalised RF current in the final bunches increases, tending to a limit which depends upon the perveance, as the initial bunches become tighter. However, for perveance greater than 1.0, it is found that a point is reached where the initial bunches are already optimal and cannot be compressed further if the maximisation of F_2 is used as the criterion. The figure also shows that, beyond a certain point, further tightening of the input bunches yields only a marginal improvement in the final bunches. Figure 13.10(b) shows that the distance from the cavity to the plane of optimal bunching decreases as the output bunches get tighter, and also as the perveance increases.
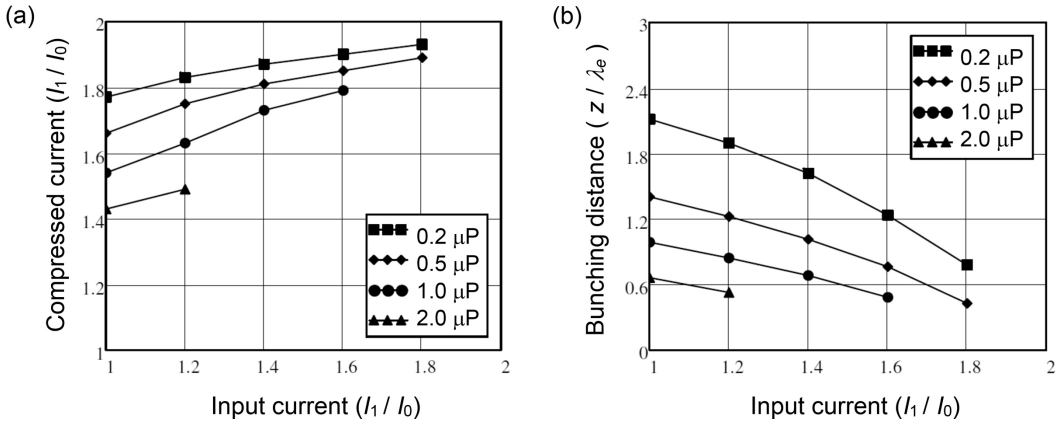
Figure 13.10: (a) Variation of normalised RF current in compressed bunches, and (b) variation of the optimal drift length with the normalised RF current in the input beam and with beam perveance.

13.3.4 Output Section

The output cavity is normally tuned close to resonance at the band centre so that the gap field is in anti-phase with the bunches to remove as much energy from them as possible. From the discussion in Section 11.8.4 we know that the electronic efficiency is given by

$$\eta_e = \frac{1}{2}\left|\frac{I_1}{I_0}\right| M_{eff} \frac{V_g}{V_0},$$ (13.39)

where V_g is the output gap voltage and M_{eff} is the effective gap coupling factor defined by (11.180). The ratio $\left|I_1/I_0\right|$ is determined by the bunching in the previous sections. It has a theoretical maximum value of 2 for a delta-function bunch. The electronic efficiency which can be achieved is determined by a number of factors:

i) The tendency for some electrons to be re-accelerated leads to an optimum value for energy transfer to the gap, as shown in Figure 11.31. In extreme cases some electrons may be reflected by the gap. If these pass back down the drift tube towards the electron gun they may provide positive feedback causing the tube to break into oscillation. We saw in Section 11.8.4 that if $M = 0.8$ then the optimum gap voltage is $1.5V_0$ and $M_{eff} \sim 0.55$ so that $M_{eff}V_g/V_0 \sim 0.82$. This analysis neglected the effects of space-charge which are discussed in Section 13.3.6.

ii) The electrons leaving the gap must have sufficient residual velocity to drift out of the field of the gap, without the formation of a virtual cathode at the entrance to the collector. For example, if $M_{eff}V_g/V_0 = 0.9$ then the electrons leave the gap with about 10% of their initial energy and velocities around 30% of their initial velocity. This is probably close to the limit set by this requirement. We note that this limit may be greater than that set in (i) above but it is also influenced by

potential depression caused by the high space-charge density just after the gap (see Section 13.3.6).

iii) The velocity spread in the bunches as they enter the output gap tends to reduce the exit velocities of the slowest electrons and to increase the risk that they may be reflected. For this reason it is desirable for the velocity spread to be small.

iv) Storage of potential energy in the bunches reduces the kinetic energy available for conversion to RF power. The increase in the space-charge density in the bunch entering the output gap leads to increased potential depression which can be estimated from (7.8) by replacing the mean beam current by the peak current. Thus the velocity of the bunch is less than u_0 and the effective gap voltage must be reduced accordingly. When the velocity of the bunch is reduced by the gap the space-charge density and the stored potential energy are increased. The power balance equation is

$$P_{DC} = P_{RF} + P_{KE} + P_{PE},$$ (13.40)

where P_{RF} is the power transferred to the gap, P_{KE} and P_{PE} are the kinetic and potential power in the spent beam, and the RF power used to modulate the beam is neglected. Thus the electronic efficiency is

$$\eta_e = 1 - \frac{P_{KE}}{P_{DC}} - \frac{P_{PE}}{P_{DC}}.$$ (13.41)

The importance of the potential energy term increases as the average charge density in the beam increases. It can lead to a reduction in efficiency computed from the kinetic power by several percentage points in beams with high perveance [33]. This suggests a possible alternative figure of merit for bunching

$$F_3 = \frac{V_b}{V_a} \cdot F_2,$$ (13.42)

where V_b is the energy of the bunched beam in eV calculated from the kinetic energies of the electrons.

v) It is essential to avoid voltage breakdown in the output gap (see Section 18.7.1). The calculations which have been presented so far assume that the gap voltage can be as large as necessary to obtain the optimum extraction of energy from the bunches. The breakdown condition for a radio-frequency gap is often taken to be given by Kilpatrick's formula [34]

$$f = 1.6E_{\max}^2 \exp\left(-8.5/E_{\max}\right),$$ (13.43)

where f is the frequency in MHz and $E_{\max}$ is the maximum electric field on the surfaces of the gap in MV m^{-1}. The maximum electric field is much greater than the average field in the gap, defined by $E_0 = V_g/g$, because of the concentration of the field on the tips of the drift tubes. These are commonly rounded to reduce the maximum field as far as possible. An example of the output gap design for a high

power klystron is given in [35]. In this 50 MW tube operating at 2856 MHz the output gap voltage at saturation was 418 kV, corresponding to $V_g/V_a = 1.32$. The maximum electric field was 36 MV m^{-1} and the average gap field was 26 MV m^{-1} giving a field enhancement factor of 1.4. Substitution of these figures into (13.43) gives the maximum safe frequency of operation of the gap as 1.6 GHz. However, it is known that fields up to twice this limit can be sustained if the surfaces are clean and well-prepared [34]. Thus the criterion in (13.43) may be regarded as conservative.

Performance beyond that of a single output cavity can be achieved by using two or more coupled output cavities in which the gap field is less than it would be in an equivalent single cavity [36, 37]. This arrangement, known as an extended inter-action output cavity, is usually operated in the π-mode with the fields in adjacent cavities in antiphase. The cavity resembles a short section of coupled-cavity slow-wave structure with short-circuited ends. A description of a five-section, π-mode, extended interaction cavity is given in [38]. At very high powers it has been found better to operate the cavities in the 2π-mode so that adjacent cavities are in phase with one another [39].

The efficiency of klystrons decreases with increasing frequency, as shown by the data for CW klystrons in Figure 13.11(a). The decrease is the result of design compromises and the effects of increasing losses in the output cavity. The figure shows the trend line

$$\eta_e = 0.65 - 0.0225f, \tag{13.44}$$

where f is the frequency in GHz. Figure 13.11(b) shows the data for pulsed klystrons with the same trend line. The data follows the same trend apart from two outlying points.

Large-signal modelling of klystrons shows that, at saturation, the harmonic currents in the beam have large amplitudes and, in particular for the second harmonic, $I_2/I_0 \sim 1$. However, because the output cavity is resonant, a low impedance is presented to the harmonic current, and the output power at harmonic frequencies is small. Measurements on a SLAC high power klystron at a fundamental power of 31 MW showed second and third harmonic outputs of –24 dBc and –39 dBc respectively [40]. Clearly this would not be the case if any of the higher-order modes of the output cavity were to present appreciable impedance to the beam at harmonic frequencies. It is therefore important to check the frequencies of the higher-order modes as part of the design of the cavity.

13.3.5 Output Coupling

The discussion to this point has assumed that the coupling of the output cavity to the external load has been arranged so that it is matched. In practice it is not always possible to ensure that the load is perfectly matched to the output of the klystron even when a circulator is placed between them. It is therefore necessary to consider the effect on the performance of the tube if the load is mismatched. For simplicity

(a)

(b)
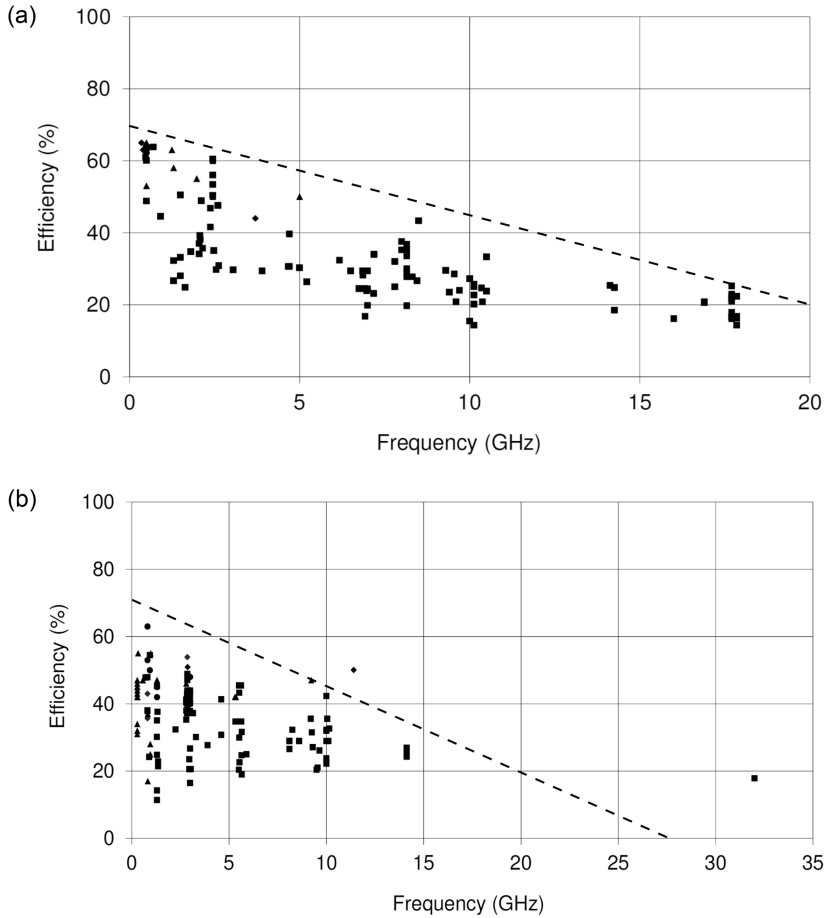
Figure 13.11: Efficiencies of klystrons from published data sheets plotted against frequency: (a) CW tubes, and (b) pulsed tubes.

we will assume that the current induced in the cavity is not affected by the change in the match and that the cavity is resonant. From Figure 13.4(b) we see that the load is driven by a current source i_g whose source admittance is $G_c + G_b$. From (11.108) the gap voltage is

$$V_g = \frac{i_g}{G_b + G_c + Y_L},\tag{13.45}$$

where $Y_L = G_L + jB_L$ is the load admittance referred to the gap. When the load is correctly matched to the tube $B_L = 0$, $G_L = G_{L0} = G_c + G_b$ and

$$V_{g0} = \frac{i_g}{2G_{L0}}\tag{13.46}$$

as expected. Dividing (13.45) by (13.46) gives

$$\frac{V_g}{V_{g0}} = \frac{2G_{L0}}{G_{L0} + G_L + jB_L} = \frac{2}{1 + g_L + jb_L}, \tag{13.47}$$

where the lower case symbols represent admittances which have been normalised to the source conductance. This can be written

$$\frac{V_g}{V_{g0}} = 1 + |\rho| \exp(j\phi), \tag{13.48}$$

where $|\rho|$ is the magnitude of the voltage reflection coefficient of the load and ϕ is its phase. If the tube is designed to give maximum efficiency with a matched load then the output power is reduced when the amplitude and phase of the load are changed (see Figure 11.31). Contours of constant load power can be plotted on a Smith chart of the normalised load admittance referred to the output flange of the tube. The result is known as a Rieke diagram (see Figure 13.12). The contours of constant power are not circles because the effective gap coupling factor depends on the gap voltage. This effect increases as the mismatch increases. The shaded area of the diagram represents a region in which the tube must not be operated because the gap voltage would be too great. The use of the output flange, rather than the gap, as the reference plane has caused the contours to be rotated about the centre of the chart relative to their theoretical positions.

13.3.6   Theoretical Limits to Efficiency

It has been observed by a number of authors that the maximum efficiency of a klystron decreases with increasing perveance [41–44]. Taking the average of the empirical formulae suggested gives
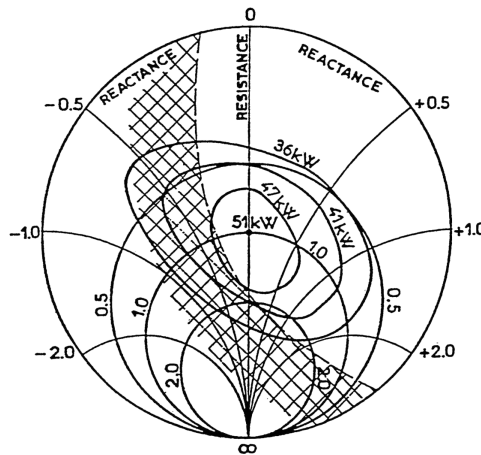
Figure 13.12: Example of a Rieke diagram for a klystron
(courtesy of Thales Electron Devices and Georges Faillon).
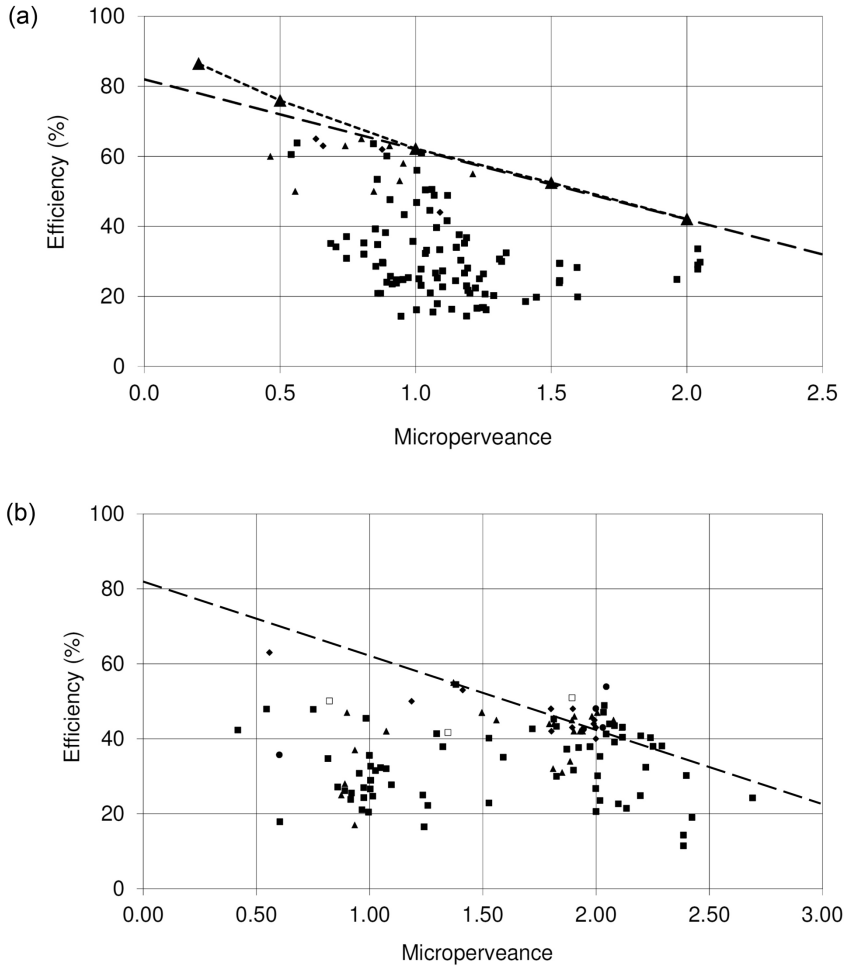
Figure 13.13: Efficiencies of klystrons from published data sheets plotted against perveance showing the trend line given by (13.43): (a) CW tubes (the triangles show points computed using Worksheet 13.4), and (b) pulsed tubes.

$$\eta_e = 0.82 - \left(0.2\times10^{-6}\right)\frac{I_0}{V_a^{1.5}}. \tag{13.49}$$

Figure 13.13(a) shows points obtained from published data sheets for CW klystrons with the line given by (13.49) for comparison. It should be noted that this line does not take account of the decrease in efficiency with increasing frequency shown in Figure 13.11. The two effects can be combined to produce a general expression for the maximum possible efficiency

$$\eta_e = 0.85 - 0.2\times\mu\text{Perv.} - 0.0225\times f\left(GHz\right). \tag{13.50}$$

The data for pulsed klystrons is shown with the same trend line in Figure 13.13(b). It is clear that the efficiencies of pulsed tubes can be greater than predicted by (13.49).

The reason for this is that many pulsed tubes operate at higher beam voltages than CW tubes (see Tables 13.2 and 13.3). At these voltages relativistic effects become important and it is possible to remove more kinetic energy from the electrons.

The variation of efficiency with perveance can be explained by the effects of space-charge density. Thus we expect that the maximum attainable efficiency will be related to ω_q^2 / ω^2. If relativistic effects are ignored then

$$\left(\frac{\omega_q}{\omega} \right)^2 = \frac{p^2}{2\pi\varepsilon_0 \sqrt{2e/m_0}} \cdot \frac{1}{\left(\beta_e b\right)^2} \cdot \frac{I_0}{V_a^{1.5}}. \tag{13.51}$$

Now, typically, $\beta_e b < 1$ and $b/a \sim 0.6$. In this parameter range p is approximately proportional to $\beta_e b$ for a magnetically focused electron beam, as can be seen from Figures 11.10(b) and (c). Thus ω_q^2 / ω^2 is approximately proportional to the perveance of the beam for all tubes.

The disc model in Worksheet 13.4 was used to explore the reasons for the dependence of efficiency on perveance for non-relativistic electrons. The normalised beam and gap parameters were taken to be $\beta_e a = 0.5$; $\beta_e b = 0.3$; $\beta_e g = 0.5$ giving a small signal gap coupling factor $M = 0.945$. Thus the energy given up by the electrons was generally not limited by re-acceleration in the output gap. These parameters are similar to those used in tubes designed for high efficiency. The effective coupling factor taken from Figure 11.32 was 0.81. The beam entering the penultimate cavity was assumed to be ideally bunched, as described in Section 13.4.2. For each value of the perveance the case giving the highest value of $|I_1/I_0|$ in Figure 13.10(a) was chosen. The voltage across the penultimate gap was adjusted to maximise F_2 and the output gap was placed at the plane where the figure of merit was maximum. The amplitude of the output gap voltage was then adjusted to minimise the final kinetic energy of the electrons. The relative phase of the output gap voltage was fixed at $-180°$ which is close to the phase at which the bunches experience the maximum retarding field. This procedure was found to give the best conversion efficiency. The results, shown as triangles on Figure 13.13(a), are similar to those given by (13.49). They are an indication of the maximum efficiency which can be achieved when transverse motion of the electrons is not important, other design constraints do not apply, and the power loss in the output cavity is negligible. In all the cases examined, apart from 0.2 µPerv, the efficiency was found to be limited by the reflection of slow electrons by a potential minimum after the output gap (see Figure 13.14). From these results it appears that the maximum efficiency of non-relativistic tubes tends to 85% in the limit of small perveance and low frequencies, as suggested in (13.50).

The space-charge density in the beam (which increases with increasing perveance) limits the maximum RF current in the compressed bunches entering the output cavity. It also limits the energy which can be extracted from the bunches because of the need to avoid reflection of electrons by potential depression in the spent beam. In order to achieve the highest possible efficiency at a given perveance:
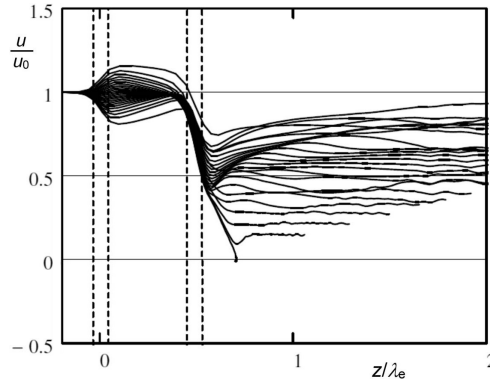
Figure 13.14: Normalised electron velocities for maximum efficiency with a 1 μPerv beam.

- The beam entering the output cavity must be tightly bunched, with minimum velocity spread, and the fraction of electrons in the inter-bunch region must be as small as possible. The techniques used to reduce this fraction in high efficiency tubes are discussed in Section 13.4.2.
- The small-signal gap coupling factor of the output cavity must be increased to the point where reflection of electrons by potential depression in the spent beam is the limiting factor.

The theoretical limit may not be achieved in practice because of:

- The need to avoid voltage breakdown in the output gap.
- The effects of radial motion of electrons.
- Design compromises, especially in broad-band tubes.
- Losses in the output cavity, which increase as the frequency increases.

13.3.7   Electron Collection

In some cases the overall efficiency of a klystron can be increased by collector depression. The technique is not suited to high-power tubes because of the difficulty of cooling multiple high-voltage electrodes (see Chapter 10). The distribution of the energies of the electrons in the spent beam can be computed using a disc model. However, it has been found that useful results can be obtained for a broadband tube using an approximate method (see Worksheet 13.5). This can be illustrated by applying it to the tube described in [45] which had a saturated efficiency (η_{sat}) of 51%. For such a tube the beam is not strongly bunched at the output cavity so it is reasonable to model the current waveform at saturation using (11.175) with $n = 2$. The waveform at other drive levels is assumed to be

$$\frac{I(\theta, X)}{I_0} = 1 + 1.33X \cos\theta + 0.33X^2 \cos 2\theta \qquad (13.52)$$

where $\theta = \omega t$ and X is the ratio of the voltage at the input gap to the value at saturation. Then the amplitude of the first harmonic current is

$$\frac{I_1}{I_0} = 1.33X. \tag{13.53}$$

The normalised dimensions of the output gap are taken to be $\beta_e a = 1$, $\beta_e b = 0.6$, $\beta_e g = 1$ and the effective voltage of the output gap at saturation is given by

$$\frac{M_{eff} V_g}{V_a} = 2\eta_{sat} \frac{I_0}{I_1} = 0.77. \tag{13.54}$$

The normalised energy of the slowest electrons is $V_{min}/V_a = 0.23$ and the effective gap coupling factor at saturation is found, from (11.182), to be $M_{eff} = 0.6$. Then the normalised resistance of the output load is

$$R_L \frac{I_0}{V_a} = \frac{V_g}{V_a} \cdot \frac{I_0}{M_{eff} I_1}, \tag{13.55}$$

where losses in the output cavity have been neglected. For output levels below saturation the effective gap voltage $V_{eff}(X)$ can be computed as a function of X by iteration. Then the energy of electrons in the spent beam as a function of the phase of the output gap voltage is given by

$$V_s(\theta, X) = V_a - V_{eff}(X)\cos(\theta). \tag{13.56}$$

Since the current waveform is symmetrical the current collected by an electrode whose potential is $V_s(\theta, X)$ is

$$I_C(\theta, X) = \frac{1}{\pi} \int_\theta^\pi I(\theta, X)\, d\theta. \tag{13.57}$$

The output backoff, in decibels, is given by

$$OBO = 20 \log\left(\frac{M_{eff}(X)}{M_{eff}(1)} X\right). \tag{13.58}$$

Figure 13.15 shows the spent-beam curves computed at saturation and three different levels of output backoff. It should be noted that an appreciable fraction of the electrons have been accelerated so that $V_s > V_a$. This is typical of the properties of the spent beam in a klystron and it limits the improvement in efficiency which can be achieved by collector depression. Despite the simplifications and approximations used to calculate them, these curves are very similar to those calculated using a disc model [45]. Note that their shapes are very different from the idealised curve in Figure 10.9, and from the typical curve for a TWT in Figure 14.16.
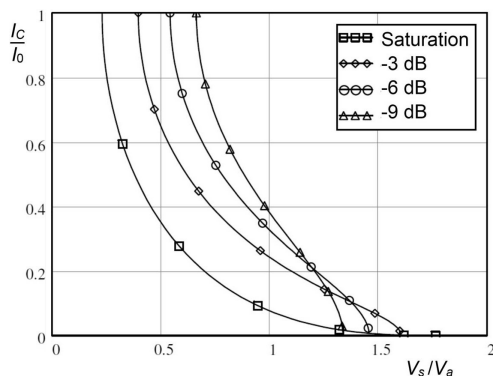
Figure 13.15: Typical spent-beam curves for a broad-band klystron at saturation, and three levels of output backoff.

The tube used in this example was designed for UHF TV transmission [45]. In analogue TV service it was operated close to saturation for brief periods during the synchronising pulses. For the greater part of the time the average power was less than half of that at saturation [46, 47]. Figure 13.15 shows that the range of energies in the spent beam is then relatively small so that an efficient multi-stage collector can be designed. In the two cases cited five-stage depressed collectors were used. The power required was then found to be less than half that without col-lector depression leading to a substantial reduction in the operating costs of the transmitter.

The other main use of klystrons with depressed collectors is in space where high overall efficiency is essential. An experimental pulsed 5.3 GHz klystron for a satellite borne synthetic aperture radar, having an RF efficiency of 40%, achieved an overall efficiency of nearly 50% with a four-stage collector (includ-ing a spike at cathode potential) [48]. The spent-beam curve of this tube is shown in Figure 7.4 in [3]. Experiments with a klystron having an RF efficiency of 40–50% gave an overall efficiency at saturation of around 70% with a ten-stage collector [49].

13.3.8   Terminal Characteristics

The performance of a klystron is appreciably affected by variations in the beam voltage, signal frequency, and output match. Figure 13.7 shows that, when a klys-tron is operated at or close to saturation, the output power is insensitive to varia-tions of input power and, by extension, to variations of beam voltage. The effects on the phase of the output signal are more serious because of the distance from the input to the output. If the distance from the centre of the input gap to the centre of the output gap is L then the phase difference between the input and the output is approximately

$$\phi \approx \frac{\omega L}{u_0}.$$

(13.59)

Thus, if the normal beam voltage is 90 kV, the tube length is 1.17 m, and the frequency is 500 MHz, the sensitivity of phase to changes in the beam voltage is –5.8 degrees per kV.

The small-signal transfer function of a klystron contains both poles and zeroes as shown by (13.20). If the zeroes are outside the working band of the tube then the transfer function is essentially that of a multi-resonator filter. The performance of the klystron as a system element is governed by this. It is found that the filling times of the cavities are much greater than the electron transit time along the tube. Thus the transient behaviour is determined by the Q factors of the cavities. It can be determined to good accuracy from the characteristics of the tube in the frequency domain [50, 51]. Under large-signal conditions the non-linearity of the tube leads to signal distortion and the generation of inter-modulation products [52, 53] (see Section 1.6.4).

13.4   Klystron Design

The design of any klystron is likely to start from a statement of requirements which includes the frequency, output power, gain, and bandwidth. Although many tubes are designed by scaling from existing devices it is instructive to consider the conceptual design process. Many of the design options are common to all tubes and the choices of parameters are usually constrained by physical factors. A typical sequence of design decisions is:

i)   The DC beam power required is obtained from the RF output power using an estimate of the tube efficiency based upon previous experience (see Figures 13.11 and 13.13). This estimate depends upon whether the primary design goal is high efficiency or broad bandwidth and it can be revised in the light of subsequent calculations. Thus

$$P_{DC} = P_{RF} / \eta_{rf}.$$

(13.60)

ii) The beam current and voltage are selected to give acceptable perveance, convergence and cathode loading in the gun. The choice is influenced by whether high gain or high efficiency is most important. If the perveance is K then

$$V_a = \left(\frac{P_{DC}}{K}\right)^{0.4}$$

(13.61)

and

$$I_0 = P_{DC} / V_a.$$

(13.62)

The beam voltage must be low enough to avoid voltage breakdown in the gun [41, 54].

iii) The normalised beam and tunnel radii $(\gamma_e b$ and $\gamma_e a)$ are chosen to achieve a compromise between strong coupling to the gaps and the strength of magnetic focusing field required for adequate beam stiffness. The beam filling factor (b/a) is typically in the range 0.5 to 0.6 to achieve low beam interception and $\gamma_e a \leq 1.0$. The tunnel radius must be small enough to ensure that the beam tunnel is strongly cut-off to avoid unwanted coupling between the cavities.

iv) The cathode radius is chosen to achieve a compromise between low cathode loading (to achieve a long lifetime) and small area convergence (to get good beam optics).

v) The choice of the number of cavities, their dimensions, and the drift lengths depends upon the desired gain, bandwidth and efficiency. This is discussed further below. It may be necessary to make a compromise between the drift lengths and the cavity heights (which increase with increasing R/Q).

The conceptual design process is best carried out using a spreadsheet, or other computer aid, which enables the results of design options to be investigated rapidly. The computer aids include simple small-signal and large-signal models of the kind illustrated in Worksheets 13.1, 13.2, and 13.3. Figure 13.16 shows the technological limits to the power of klystrons [55]. Further information about factors limiting the design of klystrons is given in [41].

Once a promising conceptual design has been found it is possible to proceed to detailed design in which every aspect is checked using computer models. Particular aspects include the design of the cavities and couplers; the electron optics of the gun, the focusing system and the collector; and the cooling system. The large-signal performance must be predicted more accurately by using a three-dimensional large-signal model.
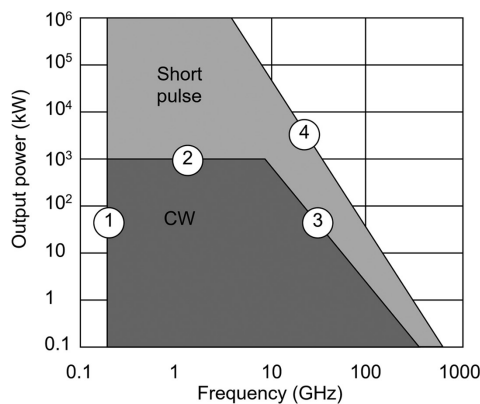
Figure 13.16: Approximate technology limitations on klystron performance: (1) size (CW and pulsed); (2) Output window (CW); (3) output cavity power density and cyclical fatigue (CW); and (4) radio-frequency breakdown (pulsed) (courtesy of H.P. Bohlen).

13.4.1   Broad-band Klystrons

The design of broad-band klystrons for telecommunications, television broadcast-ing, and radar has been discussed by a number of authors [2, 13, 30, 31, 56]. The general principles are to use stagger tuning of the cavities and a high perveance beam to reduce the Q of the cavities by high beam loading. The cavities may also be loaded internally, or externally, to reduce their Q factors. A small-signal model may be used to find the best tuning of the cavities for broad bandwidth, but a large-signal model must be used to find the effects of cavity tuning on efficiency. The zeroes in the transfer function of the tube are arranged to give minimum rip-ple in the gain of the tube, within its operating band, by suitable choice of the drift lengths (see Section 13.2.3). This can be achieved by placing cavity poles close to zeroes which fall within the band [3, 7, 17]. Good results have been obtained when the frequencies of the cavities increase progressively down the tube. The drift length preceding each cavity is adjusted so that the pole of one cavity is close to the zero of the previous one [30, 31]. It is usual to tune one or more final bunching cavities to frequencies lying above the operating band to increase the efficiency of the tube. Because the zeroes attributable to the individual sections of the tube generally lie above the cavity resonant frequencies, the gain usually falls more rapidly above the band than below it (see Figure 13.6). Some examples of tuning patterns are shown in Table 13.1.

The bandwidth of the output cavity can be increased by coupling it to the out-put waveguide by a filter, such as a second resonant cavity (see Section 3.2.4), or by using an extended interaction cavity [13, 47]. Extended interaction cavities may also be used elsewhere in place of conventional cavities to produce an extended interaction klystron (EIK) [36]. Under some circumstances the current modulation

Table 13.1: Examples of cavity tuning patterns for broad-band klystrons (relative frequencies)

Cavity number	[13]	[31]	[30]
1	0.975	0.956	0.96
2	0.955	0.960	0.95
3	1.015	0.985	1.01
4	1.045	0.985	1.04
5	1.050	1.02	1.06
6	1.092	1.02	1.08
7	1.000	1.047	1.10
8	–	1.05	1.00
9	–	1.05	–
10	–	1.067	–
11	–	1.00	–
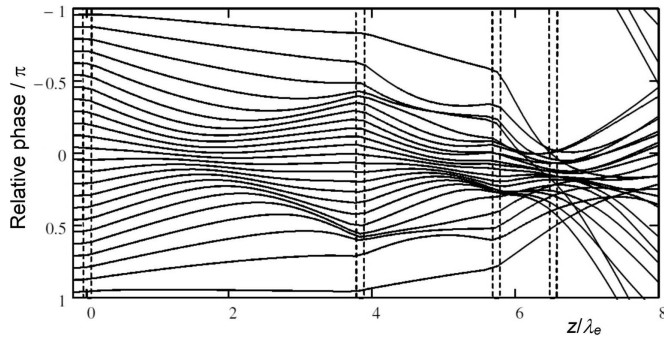
Figure 13.17: Trajectories in a high-efficiency klystron with long drift lengths.

excited on the beam by one cavity can deliver power to the same cavity so that the beam loading conductance becomes negative and the cavity self-oscillates [57]. These are known as monotron oscillations and are the basis of monotron and extended interaction oscillators [58–61]. In the ultra-wide bandwidth Clustered-Cavity™ klystron the individual cavities are replaced by closely-spaced clusters of cavities [17, 62].

13.4.2   High-Efficiency Klystrons

High-efficiency klystrons have been developed for powering particle accelerators and for plasma heating in nuclear fusion experiments [7, 35, 38, 39, 43, 63–65]. To achieve high efficiency the fundamental RF beam current entering the output cavity and the effective gap coupling factor should both be as high as possible (see Section 13.3.6). The initial bunching process should gather as many electrons as possible into a phase range of 180° while avoiding crossing trajectories. The bunches can then be compressed using several stages, if necessary (see Figure 13.9).

Two techniques for achieving high efficiency were discussed in [23]. The paper does not give full details of the parameters of the tubes considered but the calculations can be made using Worksheet 13.4 if some assumptions are made about the missing information. Figure 13.17 shows three stage bunching of a 1.0 μPerv beam. The first stage creates bunches with $I_1/I_0 = 0.86$ at $z/\lambda_e = 2.28$. It can be seen that the inner trajectories start to diverge at that plane whilst the outer ones continue to converge. The phase focusing effect of the second cavity is greatest at ±90° with respect to the bunch centre. The best bunch compression is achieved by placing the second cavity after the plane for optimum bunching so that more of the electrons are captured into the bunch. The second stage of compression gives $I_1/I_0 = 1.3$ at $z/\lambda_e = 5.4$. The third cavity, which is also placed a little beyond the plane of optimal bunching, gives $I_1/I_0 = 1.8$ at $z/\lambda_e = 6.6$. The growth of the first and second harmonics of the RF beam current along the tube is shown in Figure 13.18. Figures 13.17 and 13.18 are similar to Figure 4 in [23]. The chief disadvantage of this method of achieving high efficiency is that the drift lengths are long and increase as the
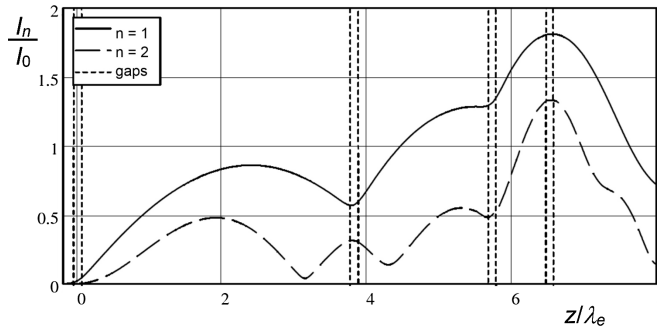
Figure 13.18: Growth in the first and second harmonic currents on the beam in a high-efficiency klystron with long drift lengths.

perveance decreases. The bandwidth of high-efficiency tubes designed in this way is comparable with that of conventional klystrons. An example of a tube of this type is discussed in Section 13.4.3. Further information can be found in [25, 32, 65, 66].

At an early date in the development of klystrons it was observed that much better initial bunching could be achieved if the gap field had a saw tooth waveform [67]. Although this cannot be achieved in practice, a useful approximation can be achieved by using one, or more, cavities tuned to the second harmonic of the signal frequency [23]. Such a cavity can be thought of as compensating for the astigmatism in the preceding fundamental mode cavity. Figure 13.19 shows the effect of a second harmonic cavity in a 0.5 µPerv tube, based on Lien's data. The first cavity acting alone produces a bunch with $I_1/I_0 = 0.9$ at $z/\lambda_e = 2.9$. A cavity tuned close to the second harmonic frequency is placed after the first cavity at $z/\lambda_e = 1.03$. The tuning of the cavity is adjusted so that the outlying electrons are gathered into the bunch. Then $I_1/I_0 = 0.95$ at $z/\lambda_e = 3.2$ while the amplitude of I_2/I_0 decreases. At this plane about 90% of the electrons are contained in a phase range of ±90° with respect to the bunch centre, as shown in Figure 13.20. The final bunching stage then comprises two cavities. The first, placed at $z/\lambda_e = 1.8$, increases the bunching to give $I_1/I_0 = 1.4$ at $z/\lambda_e = 3.6$. The second at $z/\lambda_e = 3.2$ produces a final bunch with $I_1/I_0 = 1.8$ at $z/\lambda_e = 3.9$. The growth in the first and second harmonics of the beam current is shown in Figure 13.21. Figures 13.19 and 13.21 show the same general features as figures 2 and 3 in [23]. Comparison with Figure 13.17 shows that the use of a harmonic cavity greatly reduces the total drift length required to produce tight electron bunches. This technique is usually used in high efficiency tubes. Its chief disadvantage is that it is inherently narrow band.

Klystrons which have been developed specifically for use in accelerators are commonly known as super-power klystrons. Tables 13.2 and 13.3 summarise the state of the art for these tubes. The beam voltage is limited by the need to avoid voltage breakdown in the electron gun. It can be seen from the tables that the typical beam voltages are higher for pulsed tubes than for continuous wave tubes because the breakdown voltage is higher for short pulses than for steady voltages. The beam
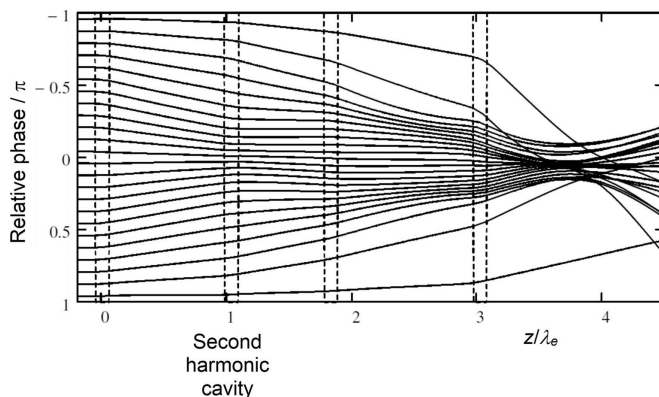
Figure 13.19: Trajectories in a high efficiency klystron with a second harmonic cavity.
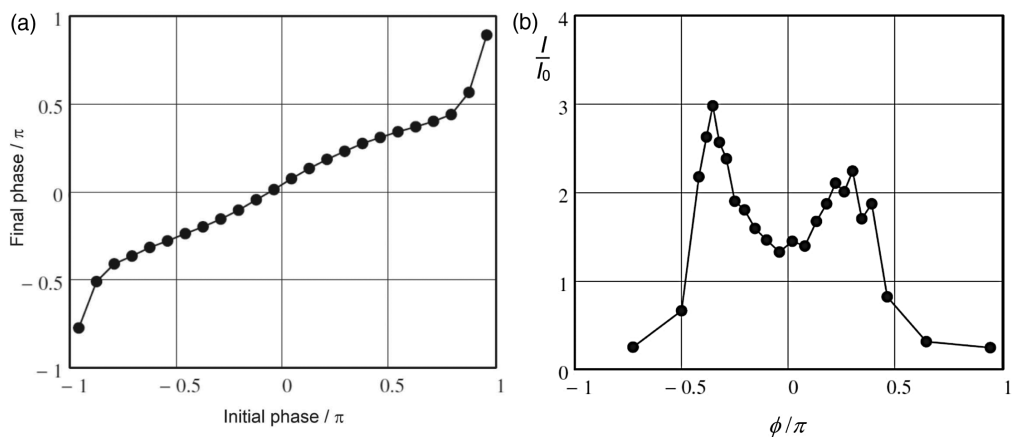
Figure 13.20: Bunching at $z/\lambda_e = 3$ in a high efficiency klystron with a second harmonic cavity: (a) arrival phase plotted against initial phase, and (b) current waveform.
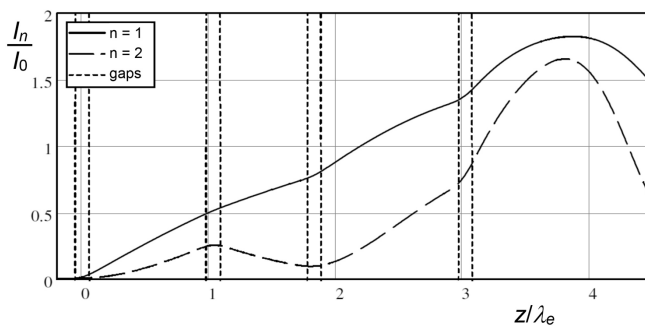
Figure 13.21: Growth in the first and second harmonic currents on the beam in a high efficiency klystron with a second harmonic cavity.

Table 13.2: Characteristics of typical continuous wave super-power klystrons

Tube	TH 2089	VKP-7952	TH 2103C*	
Manufacturer	Thales	CPI	Thales	
Frequency	352	700	3700	MHz
Beam voltage	100	95	73	kV
Beam current	20	21	22	A
RF output power	1.1	1.0	0.7	MW
Gain	40	40	50	dB
Efficiency	65	65	44	%

* This tube was developed for heating plasmas in nuclear fusion experiments

Table 13.3: Characteristics of typical pulsed super-power klystrons

Tube	[39]	[68]	[69]	
Frequency	2.87	3.0	11.4	GHz
Pulse length	1.0	1.0	1.6	μsec
Beam voltage	475	610	506	kV
Beam current	620	780	296	A
RF output power	150	213	75	MW
Gain	59	58	60	dB
Efficiency	51	44	50	%

current is limited by the current density which is available at the cathode, and by the area of the cathode which decreases with frequency. The saturation current density of thermionic cathodes is greater for short (microsecond) pulses than for DC operation.

13.4.3   Case Study: The SLAC 5045 Klystron

The design of the SLAC 5045 50 MW klystron has been described in detail and provides a convenient case study [35, 70]. The requirement was for a pulsed tube at 2856 MHz with 6 μs pulses at a pulse repetition frequency of 180 Hz. From Figure 13.13(b) it can be seen that an efficiency of at least 45% should be attainable at a perveance of $2.0 \, \mu\text{A V}^{-1.5}$. Although a higher efficiency could be obtained by reducing the perveance, a higher beam voltage would be required, leading to increased problems with voltage breakdown. The beam voltage could be reduced by choosing a higher perveance but the efficiency would be reduced and it would be more difficult to produce a well-collimated beam. The DC beam power required is 100 MW giving a voltage of 315 kV and current of 350 A for the perveance chosen (see (13.60) to (13.62)). If the beam filling factor $b/a = 2/3$

then the effective beam voltage is reduced to 290 kV by space-charge potential depression (see (7.8)).

The normalised beam and tunnel radii chosen were $\gamma_e b = 0.5$ and $\gamma_e a = 0.75$ so that $b = 10$ mm and $a = 15$ mm. Using these beam parameters with Worksheet 13.4, as described in Section 13.3.6, the theoretical maximum electronic efficiency was calculated to be 52%. This confirms that the limiting efficiency increases as the beam voltage increases as a result of relativistic effects. The plasma frequency including the effects of relativity is 790 MHz and the Brillouin field is 0.63 T. From Figure 13.10(a) we see that the maximum RF current at the output cavity is $1.5I_0$. From Table 11.2 and Figure 11.30 it is estimated that the peak current in the bunch is $3.2I_0$. Figure 7.5 shows that the stiffness of the beam should be adequate if the solenoid field is greater than 1.75 times the Brillouin field. The power consumed by the solenoid produces a significant reduction in the overall efficiency of a pulsed tube. We therefore choose the minimum field ratio giving a solenoid field of 0.11 T. The power required by the solenoid has been eliminated in some more recent tubes by using PPM focusing [43]. The tube has a Pierce electron gun of conventional design with a dispenser cathode and an average loading of $5.6\,\mathrm{A\,cm^{-2}}$. The product of the potential difference between the electrodes of the gun and the maximum electric field is $7.1 \times 10^3\,\mathrm{kV^2 mm^{-1}}$ which is close to the maximum for this pulse length given in [54]. Careful selection of the electrode materials together with extensive polishing and careful conditioning resulted in satisfactory rates of voltage breakdown. Because the beam is strongly relativistic it is important to use $\gamma_e = \sqrt{\beta_e^2 - \omega^2/c^2}$ rather than β_e when calculating the plasma frequency reduction factor. The reduced plasma frequency is 210 MHz and the reduced plasma wavelength is 1.08 m.

The details of the interaction structure of the tube are shown in Table 13.4. Most of the data is taken from [35] but the cavity frequencies are those given in [70]. The tube had been redesigned after the first paper to reduce its gain and increase its stability, but it is not clear whether the changes were restricted to retuning of the cavities. The coupling factors of the gaps were computed as described in Section 11.2.4 using the field profile given in (3.90) with $kg = 4$.

This tube was modelled using Worksheet 13.3 with beam voltage 315 kV, beam current 354 A, for comparison with the data given in [70]. The cavity voltages in the first four cavities were obtained using the small-signal model in Worksheet 13.1. The results obtained are shown in Table 13.5.

Figures 13.22 and 13.23 show the electron trajectories computed and the growth in the RF current along the tube. The first three drift angles are approximately $\beta_q z \approx 18°$ so that the first four cavities and the drift lengths which follow them form the initial bunching section of the tube as described in Section 13.3.2. The long drift length after the fourth cavity serves to increase the bunching as described in Section 13.4.2. The effective gap coupling factor calculated using (11.180) is 0.4. This is smaller than the optimum figure given in Figure 11.32 and suggests that the external Q of the output cavity may not be optimal. The differences between the results of measurement and calculation may be because radial motion of the electrons has

Table 13.4: Data for the SLAC 5045 klystron

Cavity No.	Cavity frequency (MHz)	R / Q (Ohms)	Q_0	Q_{ext}	Gap $\beta_e g$	M	Drift z/λ_q
1	2860	80	2000	200	0.53	0.879	0.052
2	2870	75	2000		0.56	0.878	0.051
3	2890	87	2000		0.64	0.873	0.051
4	2910	96	2000		0.86	0.856	0.257
5	2970	96	2000		0.90	0.852	0.102
6	2853	85	2000	21	1.26	0.814	

Table 13.5: SLAC 5045 klystron: comparison between measured and calculated and results at saturation

	Measured [70]	Calculation
Power input (W)	500	50
Power output (MW)	50	47
Efficiency (%)	45	42
Gain (dB)	50	60

been ignored or because the current density in the beam is not uniform. Overall the results confirm the validity of Worksheet 13.3 as a means of studying the design of klystrons.

A tube of this kind can be designed by working forwards through the tube adjusting the position and properties of each cavity in turn to optimise the bunching figure of merit for the next stage. It is difficult to achieve an optimum design by hand because of the large number of parameters which may be adjusted. If a sufficiently reliable large signal computer model is available the overall performance can be optimised using a search algorithm [71].

13.5   Other Klystrons

A number of variants of the klystron are of importance for high-power, high-efficiency, high-frequency and broad-band operation. The multiple-beam klystron and the sheet beam klystron are discussed briefly below and further information can be found in [7]. The reflex klystron oscillator in which the electron beam passes twice through a single cavity, having been reflected by a negative electrode, is of renewed interest as a possible source of power at sub-millimetre wavelengths [72, 73]. The electro-statically focused klystron has the attraction that it does not require a heavy focusing magnet, but the need to combine electrostatic focusing with the
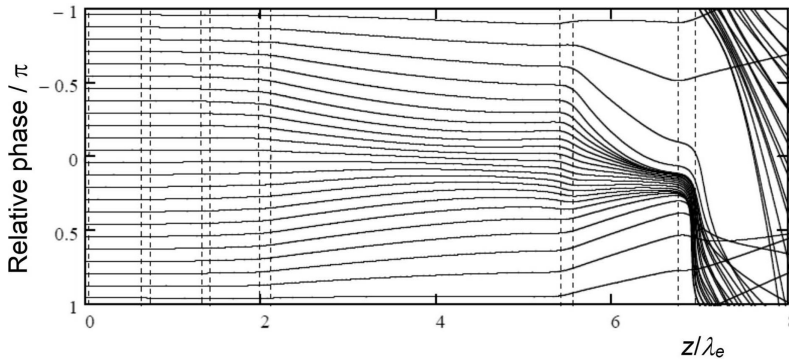
Figure 13.22: Computed electron trajectories in the SLAC 5045 klystron.
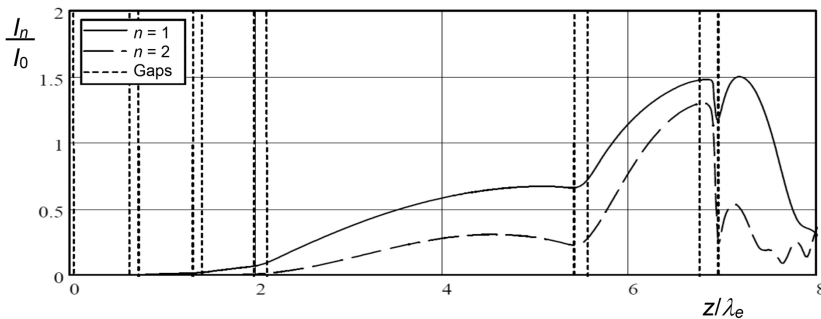
Figure 13.23: Computed harmonic current growth in the SLAC 5045 klystron.

interaction structure limits the value of this concept [3, 74]. The relativistic klystron, in which powers in the gigawatt region are generated by a very intense, high energy, electron beam, lies outside the scope of this book [75].

13.5.1   Multiple-Beam Klystrons

We have seen that the efficiency of a klystron is determined by the perveance of the electron beam so that, to get high efficiency, it is necessary to use a high-voltage, low-current beam. The use of high voltages produces problems with voltage breakdown and it is therefore difficult to obtain very high power with high efficiency. One solution to this problem is to use several electron beams within the same vacuum envelope. A klystron designed in this way is known as a multiple-beam klystron (MBK). The individual beams have low perveance to give high efficiency whilst the output power is determined by the total power in all the beams.

The principle of the MBK has been known for many years [76] but, until recently, the only such tubes constructed were for military applications in the former Soviet Union [77–79]. The first MBK designed specifically for use in particle accelerators was the Thales type TH1801 having seven beams whose performance is shown in

Table 13.6: Characteristics of a multiple beam klystron

Type	TH 1801	
Frequency	1300	MHz
Beam voltage	115	kV
Beam current	133	A
Number of beams	7	
Power	9.8	MW
Pulse length	1.5	msec
Efficiency	64	%
Gain	47	dB

Section on A - A
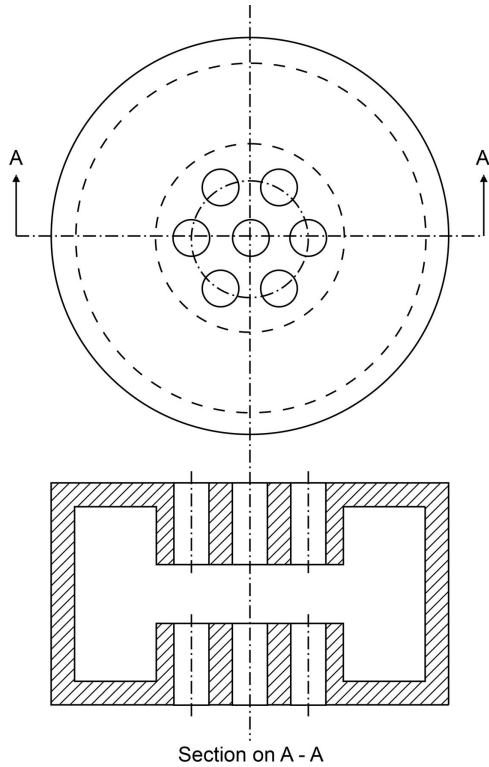
Figure 13.24: Arrangement of a cavity for a multiple beam klystron.

Table 13.4 [80]. Figure 13.24 shows an example of a cavity similar to those used in that tube. Note that the need to accommodate the beams constrains the shape of the cavity. The gap coupling factor of the central gap differs from those of the other gaps because the fields are not identical. The drift tubes may be provided with noses to increase the fields in the interaction gaps though this can make the cavities more difficult to construct. Coaxial cavities may be used in place of simple

re-entrant cavities [81, 82]. One of the main design challenges for an MBK is in the electron optics, since only a slight non-uniformity in the magnetic field can cause the beams to be deflected sideways. Further information about multiple-beam klystrons is given in [7].

13.5.2   Sheet Beam Klystrons

Multiple-beam klystrons are not well-suited to high frequencies because of the difficult electron optics. An alternative way of combining high power with low perveance is to use a sheet electron beam [7, 83, 84]. This approach is particularly of interest for devices operating at millimetre and sub-millimetre wavelengths.

References

[1] R. S. Symons, 'Klystrons for UHF television', *Proceedings of the IEEE,* vol. 70, pp. 1304–1312, 1982.

[2] G. O. Chalk and C. N. O'Loughlin, 'Klystron amplifiers for television', *Electronic Equipment News,* February 1968.

[3] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, England: Research Studies Press Ltd., 1995.

[4] L. Sivan, *Microwave Tube Transmitters*. London: Chapman and Hall, 1994.

[5] R. G. Carter, 'R.F. power generation', in R. Bailey, ed., *Proceedings of the CERN Accelerator School 'RF for Accelerators', 8–17 June 2010, Ebeltoft, Denmark*. Geneva: CERN, pp. 173–207, 2011.

[6] R. G. Carter, 'Acceleration technologies for charged particles: an introduction', *Contemporary Physics,* vol. 52, pp. 15–41, 2011.

[7] G. Caryotakis, 'Klystrons', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimeter-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 107–170, 2005.

[8] A. Staprans *et al.*, 'High-power linear-beam tubes', *Proceedings of the IEEE,* vol. 61, pp. 299–330, 1973.

[9] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer-Verlag, pp. 1–84, 2008.

[10] A. S. Gilmour, Jr., *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-Field Amplifiers and Gyrotrons*. Norwood, MA: Artech House, 2011.

[11] J. R. M. Vaughan, 'The input gap voltage of a klystron', *IEEE Transactions on Electron Devices,* vol. 32, pp. 2510–2511, 1985.

[12] B. E. Carlsten and P. Ferguson, 'Numerical determination of the matching conditions and drive characteristics for a klystron input cavity with beam', *IEEE Transactions on Electron Devices,* vol. 44, pp. 894–900, 1997.

[13] G. Faillon, 'Klystrons de puissance à large bande', *Revue Technique Thomson-CSF,* vol. 8, pp. 289–331, June 1976.

[14] B. E. Carlsten *et al.*, 'Accuracy of the equivalent circuit model using a fixed beam impedance for klystron gain cavities', *IEEE Transactions on Plasma Science,* vol. 26, pp. 1745–1749, 1998.

[15] R. S. Symons and J. R. M. Vaughan, 'Modification of klystron beam loading by initial velocity modulation of the beam', in *International Electron Devices Meeting*, pp. 885–888, 1990.

[16] G. R. White, 'Small-signal theory of multicavity klystrons', *IRE Transactions on Electron Devices,* vol. 6, pp. 449–457, 1959.

[17] R. S. Symons and R. M. Vaughan, 'The linear theory of the Clustered-Cavity™ Klystron', *IEEE Transactions on Plasma Science,* vol. 22, pp. 713–718, 1994.

[18] T. M. Antonsen, Jr. *et al.*, 'Advances in modeling and simulation of vacuum electronic devices', *Proceedings of the IEEE,* vol. 87, pp. 804–839, 1999.

[19] R. G. Carter, 'Computer modelling of microwave tubes – a review', in *2nd IEEE International Vacuum Electronics Conference 2001*, Noordwijk, Netherlands, pp. 393–396, 2001.

[20] L. D. Ludeking *et al.*, 'Computational modeling', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway NJ: IEEE Press, pp. 507–585, 2005.

[21] T. G. Mihran, 'The effect of space charge on bunching in a two-cavity klystron', *IRE Transactions on Electron Devices,* vol. 6, pp. 54–64, 1959.

[22] H. G. Kosmahl and L. U. Albers, 'Three-dimensional evaluation of energy extraction in output cavities of klystron amplifiers', *IEEE Transactions on Electron Devices,* vol. 20, pp. 883–890, 1973.

[23] E. L. Lien, 'High-efficiency klystron amplifiers', in *Eighth International Conference on Microwaves and Optical Generation and Amplification*, Amsterdam, The Netherlands, pp. 11–21 to 11–27, 1970.

[24] T. G. Mihran, 'The effect of drift length, beam radius, and perveance on klystron power conversion efficiency', *IEEE Transactions on Electron Devices,* vol. 14, pp. 201–206, 1967.

[25] P. J. Tallerico, 'Design considerations for the high-power multicavity klystron', *IEEE Transactions on Electron Devices,* vol. 18, pp. 374–382, 1971.

[26] T. G. Mihran *et al.*, 'Electron bunching and output gap interaction in broad-band klystrons', *IEEE Transactions on Electron Devices,* vol. 19, pp. 1011–1017, 1972.

[27] R. F. Koontz *et al.*, 'Anomalous electron loading in SLAC 5045 klystron and relativistic klystron input cavities', in *IEEE Particle Accelerator Conference*, pp. 159–161, 1989.

[28] J. R. M. Vaughan, 'Multipactor', *IEEE Transactions on Electron Devices,* vol. 35, pp. 1172–1180, 1988.

[29] C. Hill and R. G. Carter, 'Investigation of possible multipactor discharge in a klystron input cavity', in *2006 IEEE International Vacuum Electronics Conference held jointly with 2006 IEEE International Vacuum Electron Sources*, pp. 81–82, 2006.

[30] R. L. Metivier, 'Broadband klystrons for multimegawatt radars', *Microwave Journal,* vol. 14, pp. 29–32, April 1971.

[31] G. Faillon, 'A 200 kilowatts S band klystron with TWT bandwidth capability', in *International Electron Devices Meeting*, pp. 283–286, 1973.

[32] A. Y. Baikov *et al.*, 'Toward high-power klystrons with RF power conversion efficiency on the order of 90%', *IEEE Transactions on Electron Devices,* vol. 62, pp. 3406–3412, 2015.

[33] J. R. Hechtel, 'The effect of potential beam energy on the performance of linear beam devices', *IEEE Transactions on Electron Devices,* vol. 17, pp. 999–1009, 1970.

[34] W. Peter *et al.*, 'Criteria for vacuum breakdown in RF cavities', *IEEE Transactions on Nuclear Science,* vol. 30, pp. 3454–3456, 1983.

[35] T. Lee *et al.*, 'A fifty megawatt klystron for the Stanford Linear Collider', in *International Electron Devices Meeting*, pp. 144–147, 1983.

[36] M. Chodorow and T. Wessel-Berg, 'A high-efficiency klystron with distributed interaction', *IRE Transactions on Electron Devices,* vol. 8, pp. 44–55, 1961.

[37] T. G. Lee, 'Multiple extraction cavities for high-power klystrons', *IEEE Transactions on Electron Devices,* vol. 40, pp. 1329–1334, 1993.

[38] W. Luebke and G. Caryotakis, 'Development of a one megawatt CW Klystron', *Microwave Journal,* vol. 9, pp. 43–47, 1966.

[39] T. Lee *et al.*, 'The design and performance of a 150-MW klystron at S band', *IEEE Transactions on Plasma Science,* vol. 13, pp. 545–552, 1985.

[40] W. Fowkes and E. Wu, 'Multimode harmonic power output measurement of SLAC high power klystrons', SLAC-PUB-3009, November 1982.

[41] R. S. Symons, 'Scaling laws and power limits for klystrons', in *International Electron Devices Meeting*, pp. 156–159, 1986.

[42] A. Beunas *et al.*, 'A high power long pulse high efficiency multi beam klystron', in *5th Modulator-Klystron Workshop for Future Linear Colliders*, 2001.

[43] R. M. Phillips and D. W. Sprehn, 'High-power klystrons for the Next Linear Collider', *Proceedings of the IEEE,* vol. 87, pp. 738–751, 1999.

[44] E. Jensen and I. Syratchev, 'CLIC 50 MW L-band multi-beam klystron', in *AIP Conference Proceedings*, p. 90, 2006.

[45] E. W. McCune, 'A UHF-TV klystron using multistage depressed collector technology', in *IEEE International Electron Devices Meeting*, pp. 160–163, 1986.

[46] E. W. McCune, 'Klystron performance using a multistage depressed collector', in *1987 International Electron Devices Meeting*, pp. 157–159, 1987.

[47] W. Schmidt, 'Multi-stage-depressed-collector klystron for high-efficiency UHF transmitter in high and medium power range', in *International Broadcasting Convention*, pp. 44–45, 1990.

[48] D. Perring, 'A design for a pulsed, PPM focused depressed collector klystron for space (SAR) applications', in *Vakuumelektronik und Displays*, Garmisch-Partenkirchen, Germany, 1989.

[49] W. Neugebauer and T. G. Mihran, 'A ten-stage electrostatic depressed collector for improving klystron efficiency', *IEEE Transactions on Electron Devices,* vol. 19, pp. 111–121, 1972.

[50] R. G. Carter and R. O. Jenkins, 'Studies of the transient response of a klystron', in *IEEE International Vacuum Electronics Conference*, Monterey, USA, pp. 312–313, 2008.

[51] T. L. Lavine *et al.*, 'Transient analysis of multicavity klystrons', in *IEEE Particle Accelerator Conference*, pp. 126–128, 1989.

[52] K. Yamamoto, 'Nonlinearities of multicavity klystron amplifiers', *IEEE Transactions on Electron Devices,* vol. 24, pp. 648–654, 1977.

[53] J. G. Wohlbier and J. H. Booske, 'Nonlinear space charge wave theory of distortion in a klystron', *IEEE Transactions on Electron Devices,* vol. 52, pp. 734–741, May 2005.

[54] G. Faillon, 'Technical and industrial overview of RF and microwave tubes for fusion', *Fusion Engineering and Design,* vol. 46, pp. 371–381, 1999.

[55] H. Bohlen, 'Vacuum electronic device limitations for high-power RF sources', in *CWRF 2008*, CERN, Geneva, 2008.

[56] T. Kageyama, 'A large-signal analysis of broad-band klystrons with design applications', *IEEE Transactions on Electron Devices,* vol. ED-24, pp. 3–12, 1977.

[57] G. M. Branch, Jr., 'Electron beam coupling in interaction gaps of cylindrical symmetry', *IRE Transactions on Electron Devices*, vol. 8, pp. 193–207, 1961.

[58] J. J. Barroso, 'Design facts in the axial monotron', *IEEE Transactions on Plasma Science,* vol. 28, pp. 652–656, 2000.

[59] J. J. Barroso, 'Electron bunching in split-cavity monotrons', *IEEE Transactions on Electron Devices,* vol. 56, pp. 2150–2154, 2009.

[60] B. E. Carlsten and W. B. Haynes, 'Discrete monotron oscillator', *IEEE Transactions on Plasma Science,* vol. 24, pp. 1249–1258, 1996.

[61] W. R. Day and J. A. Noland, 'The millimeter-wave extended interaction oscillator', *Proceedings of the IEEE,* vol. 54, pp. 539–543, 1966.

[62] R. S. Symons *et al.*, 'An experimental Clustered-Cavity™ klystron', in *International Electron Devices Meeting*, pp. 153–156, 1987.

[63] E. G. Schweppe *et al.*, 'Design and results of a 1.3 MW CW klystron for LEP', in *Particle Accelerator Conference*, vol. 2, pp. 1178–1180, 1993.

[64] C. O'Loughlin *et al.*, 'Progress in high power klystron manufacturing at EEV', in *European Particle Accelerator Conference*, pp. 1906–1908, 1994.

[65] T. G. Mihran *et al.*, 'Design and demonstration of a klystron with 62 percent efficiency', *IEEE Transactions on Electron Devices,* vol. 18, pp. 124–133, 1971.

[66] V. A. Kochetova *et al.*, 'Criteria for optimum bunching and the shape of the optimum electron bunch in a drift klystron', *Radio Engineering and Electron Physics,* vol. 26, pp. 85–91, 1981.

[67] A. H. W. Beck, *Space-Charge Waves and Slow Electromagnetic Waves*. London: Pergamon Press, 1958.

[68] S. Choroba *et al.*, 'Performance of an S-band klystron at an output power of 200MW', in *XIX International Linac Conference*, Chicago, IL, pp. 917–919, 1998.

[69] D. Sprehn *et al.*, 'Current status of the next linear collider X-band klystron development program', in *EPAC 2004*, Lucerne, Switzerland, pp. 1090–1092, 2004.

[70] G. T. Konrad, 'High power RF klystrons for linear accelerators', SLAC-PUB-3324, 1984.

[71] C. J. Lingwood *et al.*, 'Automatic optimization of a klystron interaction structure', *IEEE Transactions on Electron Devices,* vol. 60, pp. 2671–2676, 2013.

[72] J. W. Gewartowski and H. A. Watson, *Principles of Electron Tubes*. Princeton, NJ: D. van Nostrand, 1965.

[73] J. R. Fletcher *et al.*, 'Design considerations for submillimeter-wave reflex klystrons', *IEEE Transactions on Microwave Theory and Techniques,* vol. 52, pp. 2344–2351, 2004.

[74] B. Vancil *et al.*, 'A medium power electrostatically focused multiple-beam klystron', *IEEE Transactions on Electron Devices,* vol. 54, pp. 2582–2588, 2007.

[75] M. V. Fazio *et al.*, 'A 500 MW, 1 μs pulse length, high current relativistic klystron', *IEEE Transactions on Plasma Science,* vol. 22, pp. 740–749, 1994.

[76] M. R. Boyd *et al.*, 'The multiple-beam klystron', *IRE Transactions on Electron Devices,* vol. 9, pp. 247–252, 1962.

[77] A. N. Korolyov *et al.*, 'Multiple-beam klystron amplifiers: performance parameters and development trends', *IEEE Transactions on Plasma Science,* vol. 32, pp. 1109–1118, 2004.

[78] E. A. Gelvich *et al.*, 'The new generation of high-power multiple-beam klystrons', *IEEE Transactions on Microwave Theory and Techniques,* vol. 41, pp. 15–19, 1993.

[79] A. S. Pobedonostev *et al.*, 'Multiple-beam microwave tubes', in *IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 1131–1134, 1993.

[80] A. Beunas and G. Faillon, '10 MW/1.5ms, L-band multi-beam klystron', in *Displays and Vacuum Electronics*, Garmisch-Partenkirchen, Germany, pp. 257–262, 1998.

[81] T. Habermann *et al.*, 'High-power high-efficiency L-band multiple-beam klystron development at CPI', *IEEE Transactions on Plasma Science,* vol. 38, pp. 1264–1269, 2010.

[82] R. Zhang and Y. Wang, 'Design of the RF circuit for a coaxial cavity high-power multiple-beam klystron', *IEEE Transactions on Electron Devices,* vol. 61, pp. 909–914, 2014.

[83] G. Caryotakis *et al.*, 'Design of a 11.4 GHz, 150-MW, sheet beam, PPM-focused klystron', *High Energy Density and High Power RF,* vol. 691, pp. 22–33, 2003.

[84] J. Pasour *et al.*, 'Sheet beam extended interaction klystron (EIK) in W band', in *IEEE 14th International Vacuum Electronics Conference*, Paris, France, pp. 1–2, 2013.

14 Travelling-Wave Tubes

14.1 Introduction

The output of a klystron or an IOT incorporates a resonant circuit so that it is effectively operating as a class B or C amplifier. The instantaneous bandwidth is therefore 10% or less. In order to get greater bandwidth the electron beam must interact with a travelling wave supported by a non-resonant structure. We saw in Section 11.5 that gain occurs when the slow space-charge wave on a beam interacts with the forward wave of a slow-wave structure. A device which operates in this way is known as a travelling-wave tube (TWT). TWTs can be divided into two main classes: those in which the interaction between the beam and the slow-wave structure can be regarded as continuous (e.g. helix TWTs, see Section 14.2.1); and those in which it is discontinuous (e.g. coupled-cavity TWTs, see Section 14.2.2). However, TWTs of all kinds have much in common so that it is convenient to discuss them together, while drawing attention to the differences where those are important. The next section discusses the small-signal theory of both types of TWT. This is followed by a review of the properties of TWTs under large signal conditions and of various aspects of their design. The final section considers the design of TWTs for different applications requiring, respectively, wide bandwidth, high efficiency, high frequency, or high power.

The slow space-charge wave carries negative kinetic power and its amplitude increases as power is transferred to the slow-wave structure along the length of the tube. If the match at the output of the tube is not perfect then the reflected signal passes back down the slow-wave structure, attenuated by the cold loss of the structure. A further reflection occurs if the match at the input of the structure is not perfect, leading to the possibility of feedback oscillations. The maximum permissible gain in a section of a TWT can be found by considering the equivalent network for a single section shown in Figure 14.1. The RF terminations are matched 'cold' (in the absence of the beam) to achieve around 20 dB return loss. The matches in a 'hot' tube are affected by the beam loading [1]. If the forward gain of the section is A, the voltage reflection coefficients of the terminations are ρ, and the cold loss of the structure is L then output voltage is

$$V_{out} = A\left(V_{in} \pm V_{out}\rho^2 L\right) \tag{14.1}$$
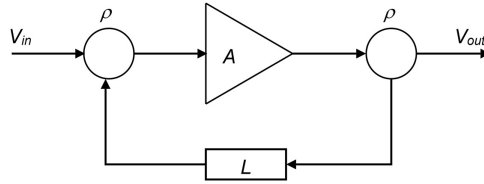
Figure 14.1: Equivalent network for a single section of a TWT.

in which the two signs represent the extreme cases where the reflected signal is in phase, and in anti-phase, with the input signal. Equation (14.1) can be re-written as

$$V_{out} = \frac{A}{1 \pm A\rho^2 L} V_{in}.$$ (14.2)

If the reflection is in anti-phase then (14.2) shows that it is possible for the loop gain to be greater than unity so that the tube will self-oscillate. For smaller loop gain the variation in the electrical length of the section with frequency means that a ripple is superimposed upon the gain curve of the tube as the phase of the reflected signal varies [1]. This discussion can be illustrated by some typical figures. Suppose that $\rho = 0.1\,(-20\text{ dB})$, $L = 0.5\,(-6\text{ dB})$ and the gain ripple should not exceed 1 dB peak to peak. Then $A\rho^2 L \leq 0.058$ so that the maximum voltage gain of the section is $11.5\,(21\text{ dB})$. Thus the gain of a single section of a TWT must normally be restricted to 20–23 dB to achieve acceptable gain ripple. To achieve greater overall gain the tube must comprise two or more sections which are separated by breaks (known as *severs*) in the slow-wave structure, or by attenuators which block the passage of the reflected wave.

14.1.1   Helix and Helix-Derived TWTs

The properties of slow-wave structures were discussed in Chapter 4 where we saw that the helix slow-wave structure provides the best approximation to a wide-band dispersion-free structure. TWTs incorporating helix slow-wave structures can achieve bandwidths up to three octaves or more. The output power is limited to a few hundred watts (depending on the frequency). Helix TWTs find applications in communications systems (especially satellite systems), electronic countermeasures and instrumentation [2–5].

Figure 14.2 shows the general arrangement of a helix travelling-wave tube body. The electron beam is generated by a Pierce electron gun, it is PPM focused, and the spent beam is commonly collected by a multi-element depressed collector. The helix is usually formed from tungsten or molybdenum tape, having a rectangular cross-section to ensure good thermal contact with the support rods. The rods themselves are made of a ceramic such as anisotropic boron nitride (ABN) or synthetic diamond [6]. Beryllia is now usually avoided because of its toxicity. The helix may be brazed to the support rods to reduce the thermal resistance [7]. The input and
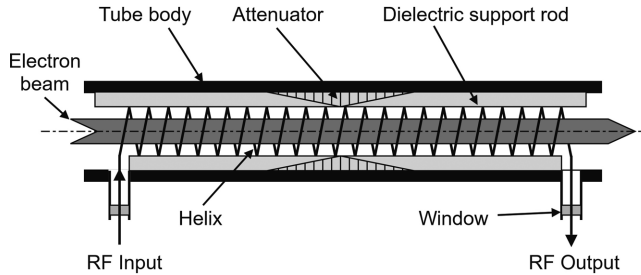
Figure 14.2: Arrangement of the body of a helix TWT. The Pierce electron gun, PPM focusing system, and the collector are not shown.

output connections may be made using coaxial lines. The tube is normally divided into two sections by an attenuator created by spraying lossy material onto the support rods [5, 8, 9]. Meander-line structures (see Section 4.2.2) have been considered for use in low-cost planar TWTs and for micro-fabrication for millimetre wave TWTs [10].

Because the helix is supported by dielectric rods, through which all heat dissipated on it must be conducted, the mean power of a helix TWT is limited by heat dissipation in the helix. In addition, the reduction in the coupling impedance as the pitch of the helix is increased makes them unsuitable for use at voltages above 25 kV. Their performance can vary with time because of changes in the properties of the attenuators and support rods [11, 12]. They are also susceptible to 'Power Holes' which cause unexpected dips in output [13]. Higher powers can be obtained at the expense of bandwidth by using the ring-bar and ring-loop structures (see Section 4.4). Compared with the helix, ring-bar and ring-loop structures are dispersive, and therefore narrow band, but they are capable of higher voltage operation. The hot bandwidth is typically around 10%. Because these structures are supported by dielectric rods in the same way as a helix their mean power capability is similar and they are therefore used chiefly for medium power pulsed applications. Examples of such tubes are: a 200 kW pulsed ring-bar TWT with bandwidth 1.2 to 1.4 GHz and a 15kW pulsed ring-loop TWT with bandwidth 3.1 to 3.5 GHz.

14.1.2 Coupled-Cavity TWTs

TWTs able to generate high peak and mean powers employ all-metal structures such as the folded waveguide, and coupled-cavity, slow-wave structures described in Sections 4.5 and 4.6. Figure 14.3 shows the arrangement of the body of a coupled-cavity TWT. The interaction is discontinuous, like that in a klystron, rather than continuous as in a helix TWT. At each interaction gap the gap voltage adds to the beam modulation, and the beam modulation induces a current in the cavity. The cavity voltages are also linked via the forward and backward waves propagating on the structure. The structure is usually divided into two or more sections by severs. Each section may be terminated internally by sever loads made from lossy
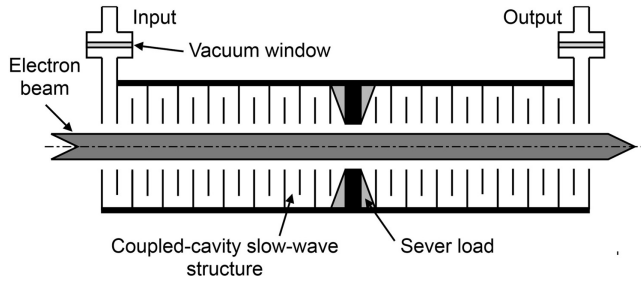
Figure 14.3: Arrangement of the body of a coupled-cavity TWT.
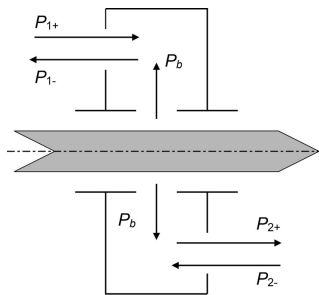
Figure 14.4: Power flows in one cavity of a coupled-cavity TWT.

ceramic material as shown. Alternatively the sever power may be coupled out into waveguides and dissipated in microwave loads either inside, or outside, the vacuum envelope [14]. Coupled-cavity TWTs are used for high power communication links and frequency agile and pulse compression radar [2].

Figure 14.4 shows the power flows in a cavity which combine to determine the gap voltage, namely the forward and backward wave powers P_+ and P_-, and the power transferred from the beam P_b. The power transferred from the beam divides equally in the forward and backward directions, and the associated signal voltages are added as phasors to those of the forward and backward waves. The backward wave assumes a more important role than in helix tubes because the backward wave components do not cancel each other to the same extent. This effect is particularly important at small values of the phase shift per cavity.

14.1.3   Energy Conversion in a TWT

The process by which energy is transferred from the electron beam to the slow-wave structure of a TWT is essentially the same for TWTs of all kinds. For simplicity we will consider a helix TWT where the dispersion diagram for uncoupled forward waves is shown schematically in Figure 14.5. We recall that the gain per unit length is greatest at the synchronous point where the slow-space-charge wave
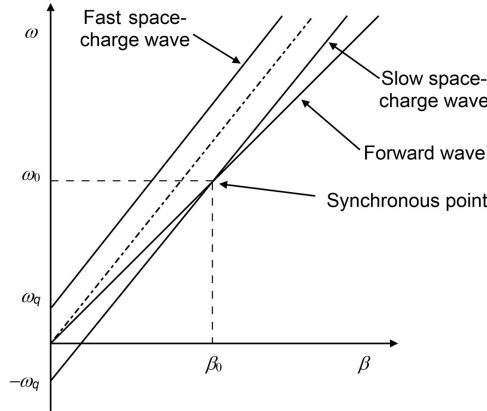
Figure 14.5: Schematic dispersion diagram of a helix TWT.

is synchronous with the forward wave on the helix (see Figure 11.20). This means that the DC beam velocity (u_0) represented by the slope of the dashed line must be greater than the phase velocity of the wave on the helix $(v_p = \omega_0/\beta_0)$.

The behaviour of the electrons as they move through the wave on the helix can be understood by considering a simple model in which space-charge forces and the spatial growth of the wave are ignored [15]. In the laboratory frame of reference the electrostatic potential at the synchronous point is

$$V = \mu_c V_0 \cos(\omega_0 t - \beta_0 z),\qquad(14.3)$$

where V_0 is the amplitude of the wave on the helix. Now let us define a coordinate system which is moving with the wave so that $z' = z - v_p t$. In this system the potential is

$$V' = \mu_c V_0 \cos(\beta_0 z').\qquad(14.4)$$

If an electron has velocity u in the laboratory frame then its velocity in the moving frame is

$$u' = u - v_p.\qquad(14.5)$$

Because the beam voltage in helix TWTs is less than 10 kV we can ignore relativistic effects so that the kinetic energy of the electron in the moving frame is

$$T' = \frac{1}{2}m_0 u'^2 = \frac{1}{2}m_0\left(u - v_p\right)^2 = \frac{1}{2}m_0 u^2 + \frac{1}{2}m_0 v_p^2 - m_0 u v_p.\qquad(14.6)$$

The total energy of the electron in the moving frame is the sum of the kinetic and potential energies

$$W' = \frac{1}{2}m_0 u^2 + \frac{1}{2}m_0 v_p^2 - m_0 u v_p - e\mu_c V_0 \cos(\beta_0 z').\qquad(14.7)$$

This function does not vary with time and it must therefore be a constant for the electron. The constant can be determined by observing that $u = u_0$ when $V_0 = 0$ so that

$$\frac{u^2}{v_p^2} - 2\frac{u}{v_p} - A\cos(\theta) - \frac{u_0}{v_p}\left(\frac{u_0}{v_p} - 2\right) = 0, \tag{14.8}$$

where $A = 2\mu_c e V_0 / m_0 v_p^2$ is the normalised amplitude of the wave and $\theta = \beta_0 z'$. The parameter A can be expressed in terms of the beam voltage as

$$A = \frac{\mu_c V_0}{V_a} \cdot \frac{u_0^2}{v_p^2}. \tag{14.9}$$

The solution of (14.8) is

$$\frac{u}{v_p} = 1 \pm \sqrt{A\cos(\theta) + \left(\frac{u_0}{v_p} - 1\right)^2}. \tag{14.10}$$

The right-hand side of (14.10) has real values for all θ when $A \leq A_0 = \left(u_0/v_p - 1\right)^2$. For larger values of A, real values of u/v_p only exist for a limited range of values of θ, indicating that the electron has been trapped by the wave. Figure 14.6 shows how the normalised velocity of the electron varies with θ for a range of values of A when $u_0/v_p = 1.05$. This figure, in which the trajectory of an electron is defined by the coordinates (z', u), is known as a *phase space diagram*. It is a useful tool for understanding the operation of a TWT. From (14.9) we find that the normalised effective helix voltage at the point where trapping commences is given by

$$\frac{\mu_c V_0}{V_a} = \left(1 - \frac{v_p}{u_0}\right)^2. \tag{14.11}$$

This has the numerical value 0.0023 for the case shown in Figure 14.6. Thus, up to the point at which the electrons begin to be trapped, the change in the mean electron velocity is small, and the transfer of energy to the wave is a small correction. Further information, including the relativistic form of (14.7) which applies to coupled-cavity TWTs, is given in [15].

In a travelling-wave tube the electrons enter the field of the wave with uniform velocity when the amplitude of the wave is small. They then follow a curve such as Orbit 2 in Figure 14.6 and move through the wave from left to right, being accelerated on the left-hand side of the diagram and retarded on the right-hand side. The electrons tend to bunch together in the retarding phase of the wave. The phase of the bunch relative to the wave remains constant while individual electrons enter and leave the bunch, and the average kinetic energy of the electrons decreases. A helpful analogy is given by the behaviour of a steady stream of cars moving along a hilly road [5]. The energy lost by the electrons increases the power in the wave which therefore increases in amplitude. This process continues until the amplitude of the
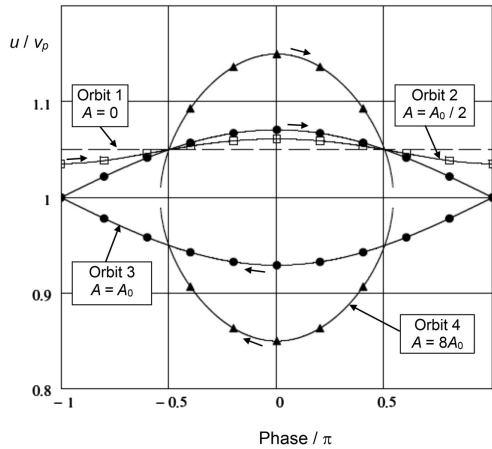
Figure 14.6: Phase space diagram for an electron for which $u_0/v_p = 1.05$ for various values of the normalised potential A showing trapping when $A \geq A_0$.

wave is great enough to reduce the relative velocity of some electrons to zero so that they are trapped by the wave, as shown by Orbit 3. Any further increase in the amplitude of the wave (A) causes the trapped electrons to move in a closed orbit, such as that shown by Orbit 4, around which they move in a clockwise direction.

Figure 14.7 shows examples of phase space diagrams, for the same parameters, which have been computed by integrating the trajectories of sample electrons through the field of a helix when space-charge effects are ignored (see Worksheet 14.2). Figures 14.7(a) and (b) show the bunching of the electrons up to the point at which trapping commences. Although many of the electrons are in the retarding phase of the field $(0 \leq \phi/\pi \leq 1)$ a substantial proportion remain in the accelerating phase $(-1 \leq \phi/\pi \leq 0)$. The mean velocity of the electrons remains close to the initial value. Figure 14.7(c) shows the majority of electrons following an orbit such as Orbit 4 in Figure 14.6, but a few lie outside it. The electrons which are trapped in the retarding field have lost appreciable energy, and the amplitude of the wave increases substantially. The slowest electrons are just about to move back into the accelerating phase of the field. Figure 14.7(d) corresponds to saturation of the output power. The electrons in the accelerating phase have begun to gain energy from the field but this is compensated by those in the retarding phase which have lost energy. The total kinetic energy is a minimum but there are still a few electrons whose velocity is greater than u_0.

14.2   Small-Signal Theory

The basic theory of interaction between an electron beam and a slow-wave structure was considered in Chapter 11 for both continuous and discrete interactions. It was shown that the coupling between the normal modes of the slow-wave structure
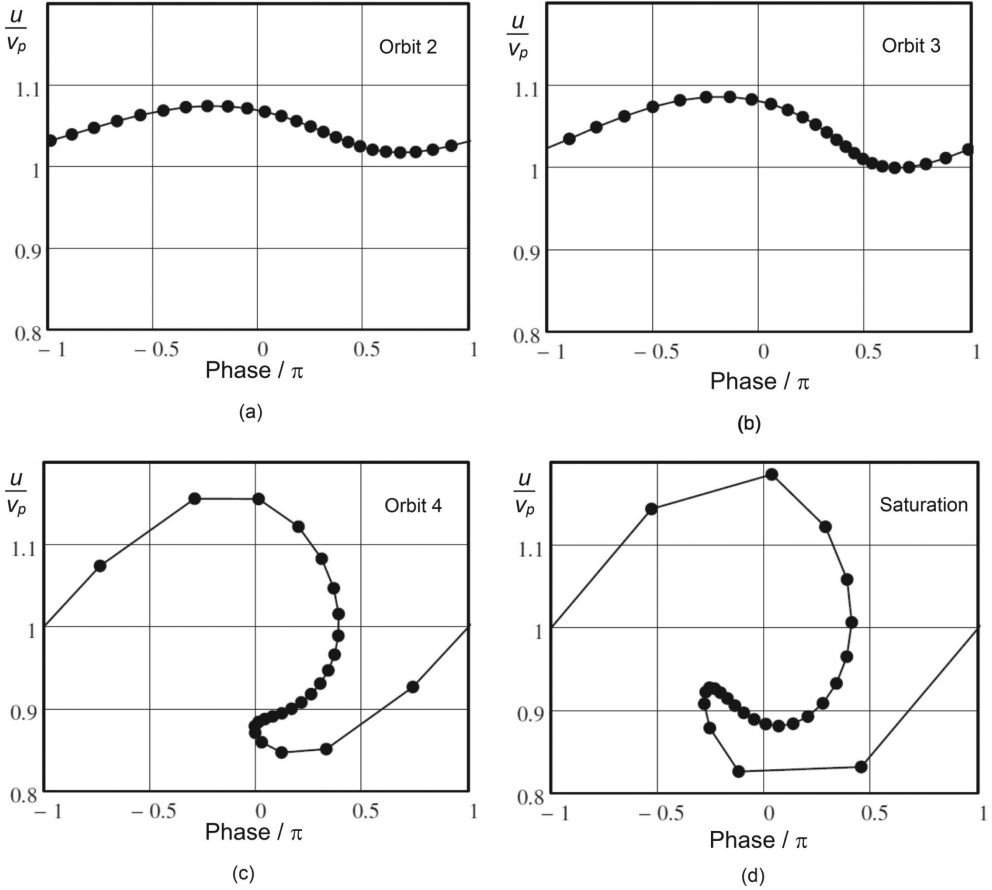
Figure 14.7: Phase space diagrams computed from the electron trajectories for the same parameters as Figure 14.6 $\left(u_0/v_p = 1.05\right)$.

and of the electron beam results in four waves, three in the forward direction and one in the opposite direction. In order to model a complete TWT the boundary conditions at the ends of each section of the tube are used to find the amplitudes of the four waves.

14.2.1   Small-Signal Theory of Helix TWTs

Figure 14.8 shows a section of a slow-wave structure of length L which is continuously coupled to an electron beam. The two ends of the circuit are connected to external transmission lines having characteristic impedance Z_c which is assumed to be the same as the characteristic impedance of the helix in the absence of the beam. The signal in the uncoupled regions can be expressed in terms of the voltages of the forward and backward waves on the transmission line $\left(V_f, V_b\right)$, and the fast and
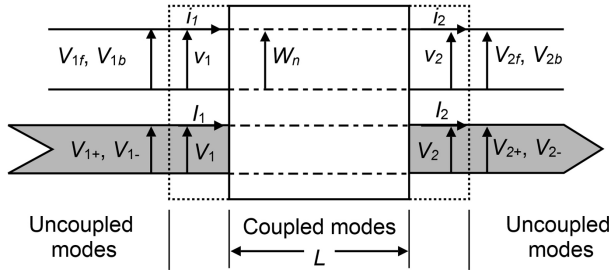
Figure 14.8: External connections to a section of a TWT.

slow space-charge waves on the electron beam (V_+, V_-). In the coupled region the signal is represented by the voltages on the helix of the four coupled modes (W_n). At the interfaces between the coupled and uncoupled regions the helix voltage v, the helix current i, the beam kinetic voltage V, and the RF beam current I. must be continuous. To construct a small-signal model of a section of a helix TWT we must find the relationships between the mode voltages and the voltages and currents at the interfaces.

Consider, first, the relationships between the voltage and current on the transmission line and the amplitudes of the uncoupled forward and backward waves. These can conveniently be written as a matrix equation

$$\begin{bmatrix} v \\ i \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ Y_c & -Y_c \end{bmatrix} \begin{bmatrix} V_f \\ V_b \end{bmatrix}. \tag{14.12}$$

Similarly for the electron beam

$$\begin{bmatrix} V \\ I \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ Y_e & -Y_e \end{bmatrix} \begin{bmatrix} V_+ \\ V_- \end{bmatrix}. \tag{14.13}$$

Where Y_e is the electronic admittance of the beam. These two equations can be combined in the form

$$\begin{bmatrix} v \\ i \\ V \\ I \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ Y_c & -Y_c & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & Y_e & -Y_e \end{bmatrix} \begin{bmatrix} V_f \\ V_b \\ V_+ \\ V_- \end{bmatrix} = [C_1][U], \tag{14.14}$$

where the column vector $[U]$ contains the amplitudes of the uncoupled normal waves.

To find the relationships for the coupled modes we express each variable in terms of the amplitudes of the normal mode voltages on the helix. Setting $\partial/\partial z = -j\beta$ in (11.130) gives

$$\beta V_f = \beta_0 V_f + \frac{1}{2}\beta\mu_c Z_c I \tag{14.15}$$

since $I = (V_+ - V_-)Y_e$. Similarly, from (11.131) we obtain

$$\beta V_b = -\beta_0 V_b - \frac{1}{2}\beta\mu_c Z_c I. \tag{14.16}$$

Adding (14.15) to (14.16)

$$\beta v = \beta_0 Z_c i \tag{14.17}$$

or

$$i = \frac{\beta}{\beta_0 Z_c} v. \tag{14.18}$$

Subtracting (14.16) from (14.15) gives

$$\beta Z_c i = \beta_0 v + \beta\mu_c Z_c I. \tag{14.19}$$

Then, substituting for i from (14.18) and re-arranging we obtain

$$I = \frac{\beta^2 - \beta_0^2}{\beta_0 \beta\mu_c Z_c} v. \tag{14.20}$$

In the same way (11.126) yields

$$\beta V_+ = \beta_+ V_+ + \frac{1}{2}\beta\mu_c v \tag{14.21}$$

and

$$\beta V_- = \beta_- V_- + \frac{1}{2}\beta\mu_c v. \tag{14.22}$$

Subtracting (14.22) from (14.21) gives

$$\begin{aligned} \beta Z_e I &= \beta_+ V_+ - \beta_- V_- \\ &= (\beta_e - \beta_q)V_+ - (\beta_e + \beta_q)V_- \\ &= \beta_e Z_e I - \beta_q V \end{aligned} \tag{14.23}$$

or

$$V = \left(\frac{\beta_e - \beta}{\beta_q}\right)Z_e I. \tag{14.24}$$

Substituting for I from (14.20) gives

$$V = \left(\frac{\beta_e - \beta}{\beta_q}\right)\left(\frac{\beta^2 - \beta_0^2}{\beta_0 \beta}\right)\frac{Z_e}{\mu_c Z_c} v. \tag{14.25}$$

The magnitude of each physical variable can then be expressed as the sum of the voltages of the coupled normal modes:

$$v = \sum_{n=1}^{4} W_n, \tag{14.26}$$

$$i = \sum_{n=1}^{4} \frac{\beta_n}{\beta_0 Z_0} W_n, \tag{14.27}$$

$$I = \sum_{n=1}^{4} \frac{\beta_n^2 - \beta_0^2}{\beta_n \beta_0} \cdot \frac{W_n}{\mu_c Z_c}, \tag{14.28}$$

and

$$V = \sum_{n=1}^{4} \left(\frac{\beta_e - \beta_n}{\beta_q} \right) \left(\frac{\beta_n^2 - \beta_0^2}{\beta_n \beta_0} \right) \frac{Z_e}{\mu_c Z_c} W_n, \tag{14.29}$$

where W_n are the amplitudes of the helix voltages of the normal modes in a uniform coupled section, and β_n are the solutions of the dispersion equation (11.136). Equations (14.26) to (14.29) can be written in matrix form as

$$\begin{bmatrix} v \\ i \\ V \\ I \end{bmatrix} = [C_2] \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \end{bmatrix}. \tag{14.30}$$

Finally, equations (14.14) and (14.30) can be combined to give the connection matrix which relates the amplitudes of the coupled modes to those of the uncoupled waves

$$[W] = [C_2]^{-1} [C_1][U] = [C][U]. \tag{14.31}$$

where $[W]$ is the column vector on the right hand side of (14.30) and $[C]$ is the coupling matrix. Since the four coupled waves are normal modes, their propagation through the circuit can be expressed as a diagonal transfer matrix $[S]$ where

$$S(z)_{nn} = \exp(-j\beta_n z). \tag{14.32}$$

Then the vector at the output of the section is expressed in terms of the vector at the input by

$$\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_{2+} \\ V_{2-} \end{bmatrix} = [C]^{-1} [S(L)][C] \begin{bmatrix} V_{1f} \\ V_{1b} \\ V_{1+} \\ V_{1-} \end{bmatrix} = [T(L)] \begin{bmatrix} V_{1f} \\ V_{1b} \\ V_{1+} \\ V_{1-} \end{bmatrix}, \tag{14.33}$$

where L is the length of the section. At the input of the first section of a TWT the forward wave is generally known, and the beam is unmodulated. The backward

wave in the input transmission line is not known but, if the output end of the circuit is well-matched, it can be assumed that there is no backward wave there. Thus the boundary conditions are mixed, with three specified at the input, and one at the output. The amplitude of the input wave can be set to 1 without loss of generality since the analysis is linear and all other parameters will then be scaled to unity. Inserting the boundary conditions into (14.33) yields

$$\begin{bmatrix} V_{2f} \\ 0 \\ V_{2+} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} T(L_1) \end{bmatrix} \begin{bmatrix} 1 \\ V_{1b} \\ 0 \\ 0 \end{bmatrix}, \tag{14.34}$$

where L_1 is the length of the section and V_{1b} is the backward wave in the input waveguide. The presence of the beam changes the match at the input of the section. Thus the backward wave is the sum of the wave reflected by that mismatch and the backward wave emerging from the tube. From the second row of the matrix we know that

$$T(L_1)_{2,1} + T(L_1)_{2,2} V_{1b} = 0 \tag{14.35}$$

so that V_{1b} can be calculated and substituted into the right-hand side of (14.33). The output vector can then be calculated directly with the check that $V_{2b} = 0$. For a single-section tube the gain is

$$G_{dB1,2} = 20 \log |V_{2f}|. \tag{14.36}$$

At any point within the section the wave amplitudes are given by

$$W(z) = S(z)[C] \begin{bmatrix} 1 \\ V_{1b} \\ 0 \\ 0 \end{bmatrix} \tag{14.37}$$

and the amplitude of the forward wave, in decibels, as a function of position is

$$A(z) = 20 \log \left| \sum_n W_n(z) \right|, \tag{14.38}$$

where the summation excludes the backward wave. Figure 14.9 shows a typical plot of the forward wave amplitude in decibels against distance along the helix. After some distance the amplitude of the growing wave is much greater than that of the other two forward waves. The graph is then almost a straight line corresponding to exponential growth of the growing wave. The boundary conditions at $z = 0$ mean that some of the input wave is reflected and that all three forward waves have non-zero amplitudes. Thus, the projection of the graph back to $z = 0$ intercepts the vertical axis at a negative value known as the *launching loss* whose value is typically in the range 3 to 6 dB.
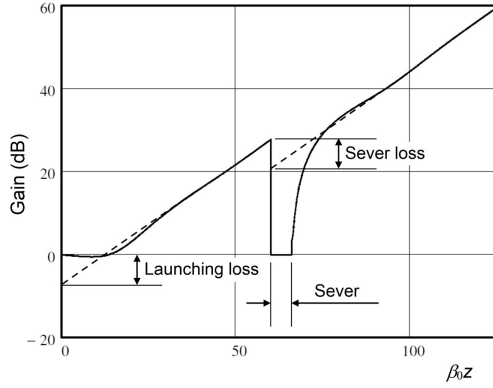
Figure 14.9: Signal growth along a two-section TWT.

When a TWT is divided into two or more sections by severs, the signal is carried between the sections by the space-charge waves on the beam. Since these are normal modes of the beam they propagate without change of amplitude, but their relative phase at the start of the next section depends upon the length of the sever. The uncoupled space-charge wave amplitudes at the end of a sever of length L_s are

$$\begin{bmatrix} V_{3+} \\ V_{3-} \end{bmatrix} = \begin{bmatrix} \exp - j\beta_+ L_s & 0 \\ 0 & \exp - j\beta_- L_s \end{bmatrix} \begin{bmatrix} V_{2+} \\ V_{2-} \end{bmatrix}. \tag{14.39}$$

If the terminations of the second section are matched then the amplitude of the forward wave at the input of the section, and the backward wave at its output, are both zero. Inserting these boundary conditions into (14.33), and denoting the input and output of the section by subscripts 3 and 4 respectively we obtain

$$\begin{bmatrix} V_{4f} \\ 0 \\ V_{4+} \\ V_{4-} \end{bmatrix} = \begin{bmatrix} T(L_2) \end{bmatrix} \begin{bmatrix} 0 \\ V_{3b} \\ V_{3+} \\ V_{3-} \end{bmatrix}, \tag{14.40}$$

where L_2 is the length of the section. The value of the unknown parameter V_{3b} in the vector on the right-hand side of the equation can be determined from the second row of the matrix using

$$T(L_2)_{2,2} V_{3b} + T(L_2)_{2,3} V_{3+} + T(L_2)_{2,4} V_{3-} = 0 \tag{14.41}$$

and inserted into the input vector to enable the output vector to be determined. The small-signal gain of the whole tube is then found by assuming that $V_{1f} = 1$ so that

$$G_{dB\,1,4} = 20 \log |V_{4f}|. \tag{14.42}$$

The wave amplitudes at the start of the second section can be determined using (14.31). Because the boundary conditions must be satisfied at the beginning of the section, the amplitude of the growing wave at the start of the section is less than that at the end of the previous section. This difference is the *sever loss*, shown in Figure 14.9 which is typically about 6 dB [16]. Like the launching loss this figure depends upon the parameters of the tube and the frequency of operation. This model is implemented in Worksheet 14.1.

The method can readily be extended to tubes with more than two sections, or those in which the parameters of the sections differ from one another. It can also be applied to tubes where the parameters of the helix are not constant throughout a section, if each uniform sub-section is modelled by a separate transfer matrix and the results cascaded to represent the whole section. The same approach can be used if a distributed attenuator is used in place of a physical sever, provided that the propagation constants of the sub-sections are adjusted to represent the cold loss profile of the attenuator.

14.2.2   Small-Signal Theory of Coupled-Cavity TWTs

In Chapter 11 it was shown that the transfer matrix for the interaction between the electron beam and a single cavity within a coupled-cavity slow-wave structure is given by (11.156)

$$
\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_{2+} \\ V_{2-} \end{bmatrix} = [P] \begin{bmatrix} V_{1f} \\ V_{1b} \\ V_{1+} \\ V_{1-} \end{bmatrix}.
\tag{14.43}
$$

Thus the transfer matrix for a section of a TWT comprising N cavities is

$$
\begin{bmatrix} V_{2f} \\ V_{2b} \\ V_{2+} \\ V_{2-} \end{bmatrix} = [P_N][P_{N-1}]\ldots[P_2][P_1] \begin{bmatrix} V_{1f} \\ V_{1b} \\ V_{1+} \\ V_{1-} \end{bmatrix},
\tag{14.44}
$$

which allows for the possibility that the cavities may not be identical to one another. Equation (14.44) takes the place of (14.33) for continuous coupling, and the boundary conditions can be applied in the same way as before. The chief difference between continuous and discrete coupling is in the effect of the backward wave generated by the interaction. With continuous coupling the backward wave components induced in each element of the slow-wave structure tend to cancel one another out and the net backward wave is small. With discrete coupling, however, this does not happen if the phase shift per cavity is small, that is, at frequencies below the synchronous point. If the continuous interaction theory is used with the parameters of a coupled-cavity tube it is found that, in this region, the forward wave gain is overestimated and the backward wave amplitude is under-estimated

[17]. Thus, in a coupled-cavity TWT, appreciable power may appear at the input port of a section as a result of the interaction within it.

14.3   Large Signal Effects

At large signal levels a TWT becomes non-linear and exhibits power saturation like a klystron. The large-signal behaviour of a TWT can be modelled with a Lagrangian model using discs or rings of charge, as described in Section 11.8. Early work in the field used normalised variables based on the Pierce parameters [18–20] but, on the whole, this makes the results difficult to interpret. More recent authors have used physical parameters [21–25]. Eulerian models, which are extensions of small-signal methods, require less computational effort than Lagrangian models but have decreasing accuracy as saturation is approached [26, 27]. Some modelling has been carried out using three-dimensional PIC codes but very large computer resources are needed for this purpose [28–30]. For most purposes one-dimensional disk models provide the best compromise between accuracy and computational efficiency. However, one-dimensional models become less accurate when $\gamma_e b > 1$ because radial variations in the electric field then mean that the motion of an electron varies appreciably with its radial position. In addition, these models cannot predict radial motion of the electrons and the current intercepted by the slow-wave structure.

The results presented below were generated using a one-dimensional model of a helix TWT similar to that in [31]. In a model of this kind it is necessary to track the electrons through a section of the tube and then use iteration to achieve a self-consistent solution. The process is then repeated for subsequent sections. However, the author has found that results which closely resemble those given by such a model can be obtained using a single pass if the voltage on the helix is first calculated using a small-signal model. As saturation is approached this method over-estimates the forward power on the helix which must therefore be calculated, instead, from the loss of kinetic energy of the sample electrons. The difference between the forward wave power assumed using small-signal theory and that calculated from the change in the kinetic energies of the electrons means that the model does not represent a self-consistent solution. However, it is found that the results are generally very close to those obtained from a more accurate model. The only exception is that the harmonic currents in the beam are too great. This is because the excessive helix voltage close to the output confines the trapped bunches more tightly than if the correct voltage profile were used. Better results are obtained if the amplitude of the wave on the helix is adjusted to show saturation, as described in Section 14.3.6. This model has been implemented in Worksheet 14.2. It has not been validated for use as a design tool, but it provides a simple way of exploring the large-signal behaviour of a TWT.

Figure 14.10 shows the growth in the forward wave on a helix TWT which is operated at the synchronous point. The tube has a single uniform section and

Table 14.1: Parameters of the TWT used in the illustrations

Beam voltage	6.0 kV
Beam current	135 mA
Beam radius	0.34 mm
Helix radius	0.68 mm
Frequency	11.7 GHz
Propagation constant (β_0)	1700 m^{-1}
Pierce impedance	37 Ω
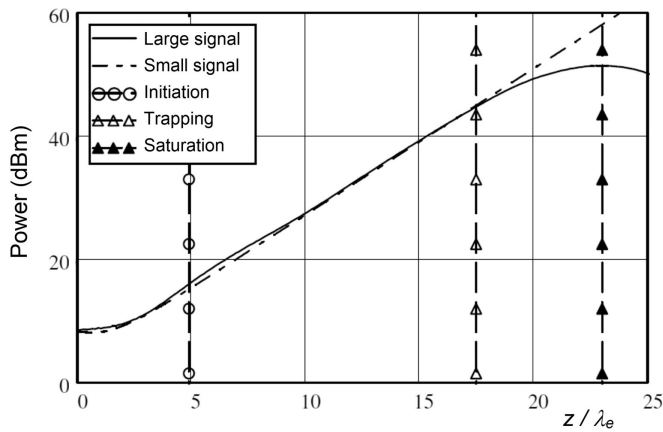
Figure 14.10: Non-linear signal growth in a single-section TWT.

parameters, similar to those used in X band space TWTs, as given in Table 14.1. They are identical to those in Figure 14.7, with the difference that space-charge effects have been included. The inclusion of space-charge forces reduces the saturated efficiency computed from 21% to 17%.

It is convenient to think of the operation of a TWT in terms of four regions as shown by the vertical dividing lines in Figure 14.10:

  i) An initial region in which the interaction between the wave on the helix and the electron beam is established.
 ii) A region in which the tube is linear and obeys small-signal theory. This region may be taken to end at the point where the velocity of the slowest electrons is equal to the phase velocity of the growing wave on the helix (v_w) so that they are on the point of being trapped by the wave.
iii) The non-linear region in which the electrons are increasingly trapped and the output power saturates.
 iv) The region beyond saturation in which the tube is overdriven.

The details of the interaction can be understood by considering further results from the model which were obtained under the same conditions.
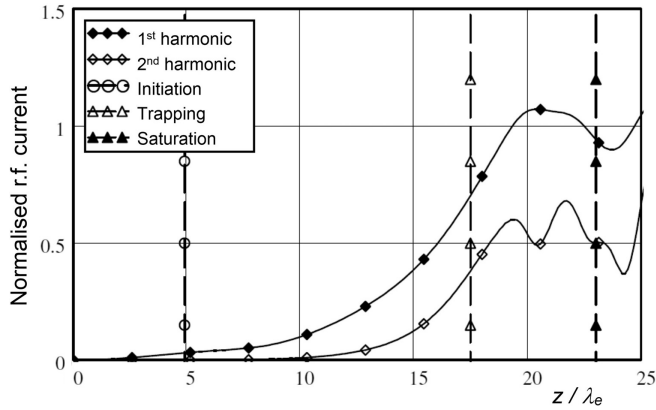
Figure 14.11: Amplitudes of the first and second harmonics of the beam current in a TWT as a function of axial position.
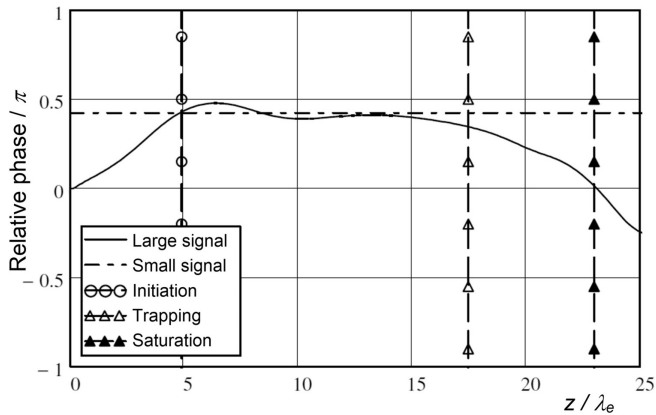
Figure 14.12: Phase of the first harmonic of the beam current in a TWT relative the wave on the helix as a function of axial position.

Figures 14.11 and 14.12 show the amplitudes of the first and second harmonics of the beam current, and the phase of the first harmonic relative to the phase of the wave on the helix, as functions of axial position. In the second region the amplitude of the current grows exponentially. The phase is almost constant, and close to the maximum retarding phase $(\pi/2)$. It should be noted that the amplitude of the second harmonic current is appreciable even in this region where the behaviour of the tube is apparently linear. The small-signal relative phase, calculated from (14.20) by assuming that the wave on the helix is dominated by the growing wave, is shown by the broken line in Figure 14.12. In the third region energy is taken from the bunches as more and more of the electrons are trapped, the phase of the first harmonic current moves away from the maximum retarding phase, and the magnitude of the current grows more slowly and eventually passes through a maximum.
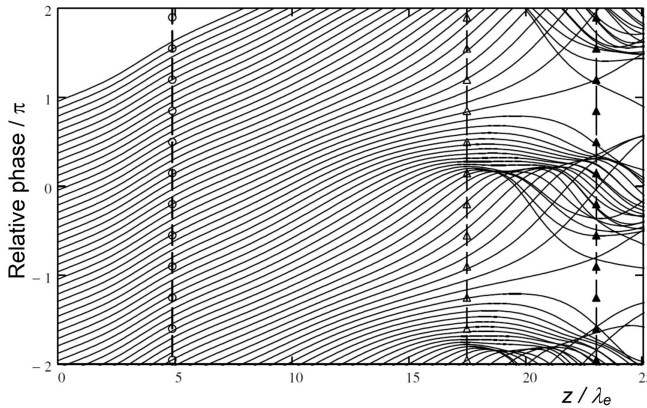
Figure 14.13: Typical modified Applegate diagram for a single-section TWT.

As the output power saturates, the phase of the first harmonic current passes from the retarding to the accelerating phase of the wave on the helix.

Figure 14.13 shows the modified Applegate diagram for the section. In this diagram the phase velocity of the growing wave on the helix (v_w) is used as the reference, and the trajectories of electrons whose velocities are greater than v_w have positive slopes. In the first region, and most of the second region, the perturbation of the trajectories is imperceptible and the electrons move through the wave with almost constant velocity. By the end of the second region the bunching is visible and the electrons are just beginning to be trapped. More and more of the electrons are trapped by the wave as they move into the third region. It can be seen that the trapped electrons oscillate within the potential well, as described in Section 14.1.3. A detailed experimental study of electron trapping was reported in [32].

Figure 14.14 shows phase space diagrams at different points along the tube for comparison with Figure 14.7. At small-signal levels (Figures 14.7(a) and 14.14(a)) the figures are very similar, except that the effect of the space-charge forces is to move the bunch centre closer to the maximum retarding phase. It can be seen that the phase of the maximum charge density is close to that for the minimum velocity, as is the case for a slow space-charge wave. At the point at which trapping commences (Figures 14.7(b) and 14.14(b)) the normalised velocity of the slowest electron $u/v_w = 1$. The phase of the bunch has started to move towards the neutral phase, and the space-charge forces produce a distortion of the diagram. The plane at which trapping commences is very close to that for the zero space-charge case. When the slowest electron is just beginning to be re-accelerated (Figure 14.14(c)), the space-charge forces prevent the electrons at the front of the bunch from entering the trapped orbit, but they remain in the retarding phase of the helix wave. At saturation (Figure 14.14(d)), the effects of space-charge produce a much more disordered diagram in which the velocities of the slowest electrons are greater than in the space-charge free case. This explains the reduction in saturated efficiency caused by the space-charge forces.
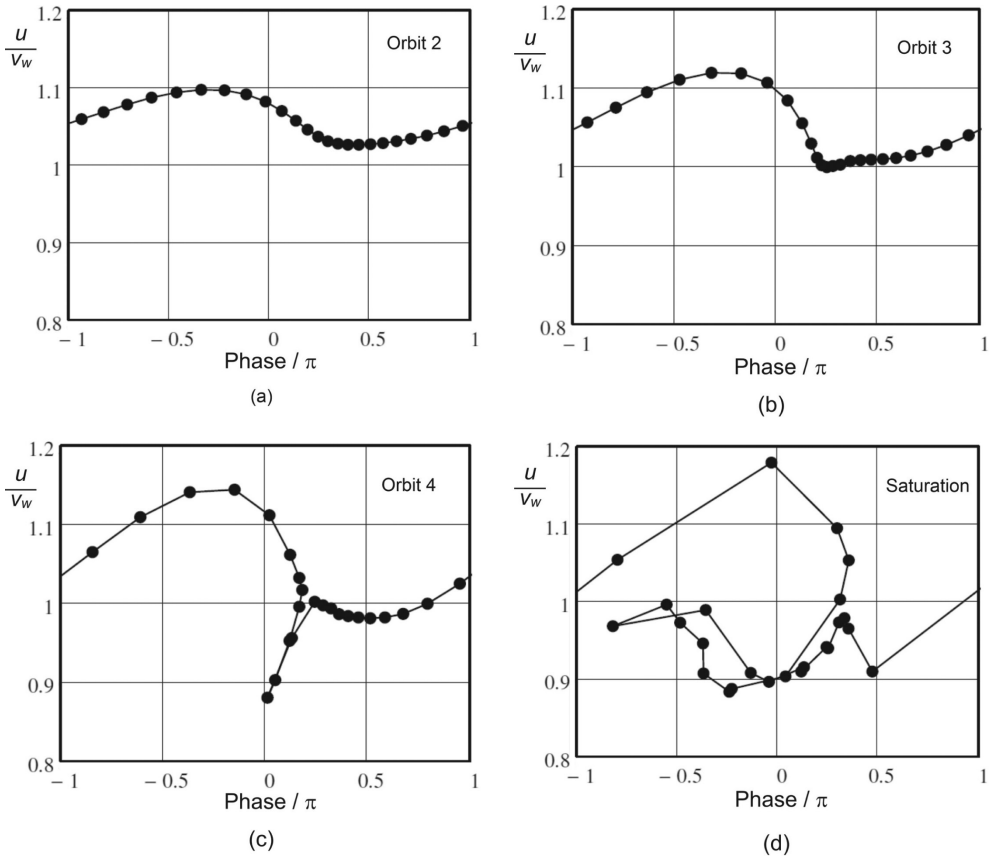
Figure 14.14: Phase space diagrams computed from the electron trajectories for the same parameters as in Figure 14.7 with the inclusion of space-charge forces.

Finally, Figure 14.15 shows the variation in the mean velocity of the electrons, normalised to the phase velocity of the growing wave on the helix. In the first region, and for most of the second region, the mean electron velocity is constant and greater than the phase velocity of the wave. In the third region the electrons lose energy and, at saturation, the mean electron velocity is a minimum. Beyond saturation the electrons start to regain energy from the wave and the mean velocity increases. Figure 14.16 shows the spent-beam distribution curve at saturation. Comparing this to the idealised curve in Figure 10.9, we can see that the overall efficiency of this tube could be increased substantially by using a multi-element depressed collector.

The results which have been presented in Figures 14.10 to 14.16 all relate to operation at the synchronous point. The first three regions which have been described correspond approximately to the initial bunching, final bunching, and output regions in a klystron. The results show the typical behaviour of a helix TWT under
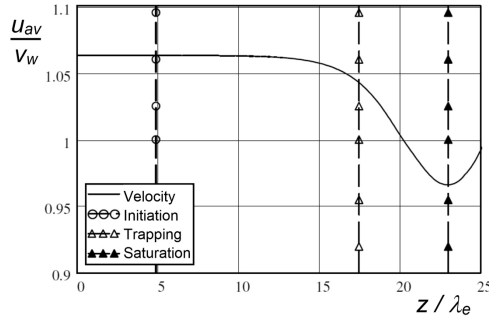
Figure 14.15: Mean electron velocity, normalised to the helix phase velocity, as a function of axial position in a TWT.
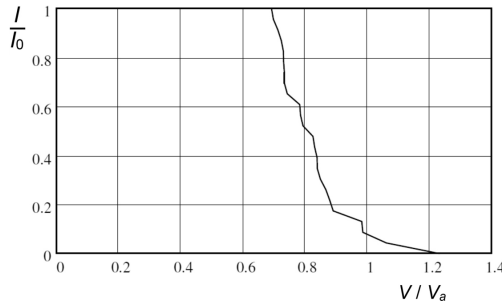
Figure 14.16: A typical spent-beam distribution curve at saturation.

large-signal conditions but variations in detail are found when the operating point is changed (see Section 14.3.3).

14.3.1   Dimensionless Parameters

Although TWTs show complicated non-linear behaviour, the principles of dimensional analysis can be applied to them (see Section 1.7.1) [33]. Thus efficiency, which is a dependent dimensionless parameter, is a function of the set of independent dimensionless parameters describing the tube. If we restrict our attention to uniform tubes having a solid beam and cylindrical symmetry then the requirement for geometrical similarity is satisfied by fixing the value of b/a. Since this is typically in the range 0.5 to 0.6 for many tubes we shall assume that it is constant. It is also necessary to assume that the length of the output section of the tube is large enough to ensure that the efficiency is not reduced and that the effects of harmonics are negligible (see Sections 14.3.3 and 14.3.4). Then the scale of the tube can be defined by the single dimensionless parameter $\gamma_e a$ because the phase velocity, coupling impedance, and plasma frequency reduction factor are all functions of it. We will assume that the magnetic focusing field is strong enough to suppress radial motion of the electrons and that relativistic corrections can be ignored. The

remaining independent dimensionless parameters can be found by rewriting the small-signal determinantal equation (11.137) as

$$\left[\Gamma_0^2 - \Gamma^2\right]\left[\Gamma_q^2 - \left(1 - \Gamma^2\right)\right] = -\frac{1}{2}\Gamma_1\Gamma_0\Gamma^2, \tag{14.45}$$

where

$$\Gamma = \frac{\beta}{\beta_e} \tag{14.46}$$

is a complex dependent parameter. The independent parameters (which can also be expressed in terms of the Pierce parameters) are:

$$\Gamma_0 = \frac{\beta_0 - j\alpha_0}{\beta_e} \quad \left(= 1 + Cb - jCd\right), \tag{14.47}$$

where α_0 defines the loss per unit length of the circuit. For a loss-less circuit this becomes the ratio of the initial electron velocity to the phase velocity of the slow-wave structure

$$\Gamma_0 = \frac{\beta_0}{\beta_e} = \frac{u_0}{v_p} \quad \left(= 1 + Cb\right); \tag{14.48}$$

the ratio of the coupling impedance to the impedance of the beam

$$\Gamma_1 = \frac{I_0 \mu_c^2 Z_0}{V_0} \quad \left(= 4C^3\right); \tag{14.49}$$

and the ratio of the reduced plasma frequency to the signal frequency

$$\Gamma_q = \frac{\beta_q}{\beta_e} = \frac{\omega_q}{\omega} \quad \left(= \frac{2C\sqrt{QC}}{1 - 2C\sqrt{QC}}\right). \tag{14.50}$$

We saw in Section 13.3.6 that this parameter can also be expressed in terms of the dimensionless beam radius and the perveance of the electron beam. Thus an alternative dimensionless parameter to Γ_q is

$$\Gamma_K^2 = \frac{1}{2\pi\varepsilon_0\sqrt{2e/m_0}} \cdot \frac{I_0}{V_a^{1.5}} = \left(\frac{\beta_e b\,\Gamma_q}{\omega_q/\omega_p}\right)^2, \tag{14.51}$$

which is the perveance divided by a constant $\left(33.0\ \mu\text{A V}^{-1.5}\right)$. The presence of the parameter C in (14.47) to (14.50) means that the Pierce parameters are only independent of one another when C is constant. The parameters defined here have the advantage that their physical significance is clearer and that it is easier to vary them independently. Two tubes which are described by the same set of independent dimensionless parameters are expected to have the same saturated efficiency and gain (apart from the effects of cold loss).

Where full data is available for an existing tube it is possible to develop a new tube by scaling, if the two tubes are described by the same set of independent dimensionless parameters. Schindler [34] showed that if $\gamma_e a$ and C are kept constant then QC is constant when both b/a and the beam perveance are constant, and the velocity parameter is chosen to give either maximum gain or maximum efficiency. Thus it is possible to scale a tube in such a way that all the dimensionless parameters are simultaneously constant. There is limited scope for scaling a tube to higher power levels, at the same frequency, because of the need to keep the temperature of the slow-wave structure below an acceptable level.

14.3.2   Dependence of Efficiency on Normalised Parameters

An approximate theory of the dependence of the efficiency of a TWT, with no cold loss, on its operating point was given by Slater [15]. It was assumed that, at saturation, all the electrons were trapped having a velocity (u_{sat}) as much below the phase velocity of the growing wave and the initial velocity was above it. Thus

$$v_w - u_{sat} \approx u_0 - v_w, \tag{14.52}$$

where v_w is the phase velocity of the growing wave. Strictly speaking this should be the phase velocity under large signal conditions but the error involved in using the small-signal phase velocity is small [32]. Then the efficiency is approximately

$$\eta_e \approx 1 - \frac{u_{sat}^2}{u_0^2} = 1 - \left(\frac{2v_w}{u_0} - 1 \right)^2. \tag{14.53}$$

When space-charge effects are included the small-signal propagation constant of the growing wave is given, in the two-wave approximation, by (11.145) so that

$$\frac{u_0}{v_w} = \frac{\beta_0 + \beta_-}{2\beta_e} = \frac{1}{2} \left(\frac{u_0}{v_p} + 1 + \frac{\omega_q}{\omega} \right). \tag{14.54}$$

At the synchronous point $(\omega = \omega_s)$

$$\frac{u_0}{v_w} = 1 + \frac{\omega_q}{\omega_s} = \frac{u_0}{v_p}. \tag{14.55}$$

Substituting from (14.55) into (14.53) gives

$$\eta_e \approx 1 - \left(\frac{1 - \omega_q/\omega_s}{1 + \omega_q/\omega_s} \right)^2. \tag{14.56}$$

Thus to a first approximation the efficiency should depend only on the normalised reduced plasma frequency. If we assume that $\omega_q/\omega_s \ll 1$ then the term in brackets can be expanded to give

$$\eta_e \approx 4\frac{\omega_q}{\omega_s} \tag{14.57}$$

so that the efficiency increases with increasing space-charge density. We note that this is the opposite to the case for klystrons.

Figure 14.14(d) shows the distribution of the velocities of sample electrons at saturation when the tube defined in Table 14.1 is operated at the synchronous point and $u_0/v_w = 1.063$. It can be seen that, in this case at least, the assumption made in (14.52) is fairly good. The electronic efficiency calculated from (14.56) is 22% which may be compared with the efficiency of 17.2% calculated using the disc model. The chief error arises from the fact that only 20 out of 24 sample electrons have been trapped, so that it would be reasonable to reduce the efficiency estimated from that equation by a capture ratio of $20/24 = 0.83$. The result is 18.5%. The remaining difference between the results can be explained by noting that the mean velocity of the electrons which are not trapped is a little greater than the initial velocity. This reduces the effective capture ratio to 0.78.

In terms of the Pierce parameters (14.56) becomes

$$\frac{\eta_e}{C} \approx 8\sqrt{QC}. \tag{14.58}$$

An experimental study of the factors affecting the efficiency of a TWT [35] assumed that

$$\frac{\eta_e}{C} \approx 4\theta H, \tag{14.59}$$

where θ is a function of the Pierce loss parameter d (defined in (14.47)) such that $\theta = 1$ when the slow-wave structure is loss-less. The function H was obtained from an unpublished memorandum by Pierce and plotted as a graph of H against QC. Comparing (14.59) with (14.58) we see that $H = 2\sqrt{QC}$ and this function is a found to be a very good fit to the curve in [35] when $QC \geq 0.25$. The discrepancy for small values of QC is not surprising because the two-wave approximation assumed in (14.54) is not valid when the space-charge is small. When experimental results were plotted against $L/C = 54.5 d$ (where L is the cold loss in dB per circuit wavelength, C is the Pierce gain parameter, and d is the Pierce loss parameter) it was found that θ was given to a useful approximation by [35, 36]

$$\log_{10}\theta = -\frac{L}{25C} = -2.18 d. \tag{14.60}$$

Thus the efficiency of a TWT operating at the synchronous point can be estimated using (14.59), or (14.56) if cold loss is neglected. If the effects of space-charge potential depression are ignored (13.51) can be written at the synchronous point as

$$\frac{\omega_q}{\omega_s} = \frac{p}{\beta_e b}\sqrt{\frac{\mu\text{Perv}}{33}}$$ (14.61)

so that the efficiency can be estimated from the perveance. Typically, the beam filling factor $b/a \sim 0.67$ and $\beta_e b \sim 0.6$, so that $p/\beta_e b \sim 0.55$ for many tubes. When the efficiencies of over fifty TWTs of different types were plotted against $0.55\sqrt{\mu\text{Perv}/33}$ it was found that nearly all the points lay between theoretical curves calculated using (14.56) with multiplying factors 0.3 and 0.7. The multiplying factor takes account of the capture ratio, and the effects of space-charge potential depression and the cold loss of the helix. Other factors affecting the efficiency include design choices such as the choice of synchronous point, and tapering of the slow-wave structure (see Section 14.3.7). There was good correlation between points derived from the data in [37] and a theoretical curve with a multiplying factor of 0.7.

The space defined by the Pierce parameters was explored in a classic series of experiments by Cutler [38]. The results were summarised in two graphs showing contours of η_e/C plotted on the $(QC, \gamma_e b)$ plane for maximum gain per unit length (synchronism) and maximum output power. These graphs showed that:

 i) at synchronism $\eta_e/C > 3$ when $\gamma_e b \sim 0.6 \pm 0.05$ and $QC \sim 0.25 \pm 0.05$;
ii) at maximum efficiency $\eta_e/C > 3.5$ when $\gamma_e b \sim 0.5 \pm 0.1$ and $QC \sim 0.2 \pm 0.05$.

It was assumed that η_e/C was constant when all the other parameters were constant. The Pierce parameters of the tube, defined in Table 14.1, are $\gamma_e b = 0.57, C = 0.062$ and $QC = 0.22$. Calculations using the model in Worksheet 14.2 gave the results $\eta_e/C = 2.7$ at synchronism and $\eta_e/C = 3.4$ at maximum efficiency, which are close to Cutler's figures.

Cutler's experimental results included phase space diagrams which closely resemble those in Figures 14.7 and 14.14. The graphs also included comparisons with the results of large-signal calculations which showed poor agreement because the radial variation of the field of the slow-wave structure had been neglected in them. Cutler concluded that there is a broad optimum in the region of $\gamma_e b = 0.5, C \leq 0.14$ and $QC = 0.2$ where it should be possible to achieve efficiencies greater than 30%. The results in this paper have been widely used to guide the design of TWTs. However, the use of QC as a space-charge parameter, and the fact that $H = 1$ when $QC = 0.25$, have led to the widespread assumption that the efficiency depends on C rather than on ω_q/ω [8, 14, 39, 40]. This misunderstanding arose because Slater's original work, which ignored the effects of space-charge, suggested that $\eta_e \approx 2C$. Pierce and Cutler were aware that the correct expression should be $\eta_e \approx -4Cy_1$ where y_1, which is the growing wave parameter, ranges between $-1/2$ and -2 depending upon the operating point and the value of the space-charge parameter. With this correction the assumptions are exactly the same as have been made in the derivation of (14.56). Thus the assumption that the efficiency is proportional to C assumes, wrongly, that y_1 is constant.

Other experiments described in a less well-known paper by Caldwell and Hoch [41] covered similar ground though with a more restricted range of parameters. It

was concluded that there was no simple relationship between the efficiency and the Pierce parameters. The saturated efficiency at the synchronous point was found to be independent of the beam current but related to the normalised reduced plasma frequency. These results, therefore, tend to support the validity of (14.57) as an approximation.

14.3.3   Dependence of Efficiency on the Operating Point

Equation (14.53) shows that the efficiency of a TWT with no circuit loss should be determined by the ratio u_0/v_w at the operating point. It is useful to note that u_0/v_w is always greater than unity when the tube has gain, whereas u_0/v_p can be less than unity. Thus, the efficiency given by (14.53) is always positive. The propagation constant of the growing wave in the two-wave approximation is given by (11.145) as

$$\beta = \frac{1}{2}(\beta_0 + \beta_-)$$
(14.62)

so that

$$\frac{u_0}{v_w} = \frac{\beta_0 + \beta_-}{2\beta_e} = \frac{1}{2}\left[\frac{\beta_0 + \beta_q}{\beta_e} + 1\right].$$
(14.63)

Thus, the ratio u_0/v_w can be increased by one, or more of:

  i) reducing β_e by increasing the beam voltage, keeping the frequency, the reduced plasma frequency and the pitch of the helix constant [42];
 ii) increasing β_0 by increasing the frequency, keeping the beam voltage and helix pitch constant;
iii) increasing β_0 by reducing the pitch of the helix (and thus reducing v_p), keeping the beam voltage and the frequency constant;
 iv) increasing β_q by increasing the beam current, keeping the beam voltage and the frequency constant.

It should be noted that the effect of changing u_0/v_w by different means does not yield exactly the same results, because other parameters may be changed also. Thus, the effects of increasing the beam voltage at constant DC power, constant perveance, and constant current differ from one another. Similarly, when the frequency is changed, the impedance of the helix also changes. A simple helix has positive dispersion (i.e. v_p decreases with frequency) and the slow space-charge wave has negative dispersion so that u_0/v_w increases with increasing frequency. When the helix is heavily loaded it may have negative dispersion and then the ratio u_0/v_w may be kept almost constant if the dispersion of the helix is chosen correctly [43].

 If the beam voltage of a TWT is varied at constant perveance then all the normalised parameters, except u_0/v_p, are almost constant. Curve (a) in Figure 14.17 shows how the small-signal growth rate $\mathrm{Im}(\beta)/\beta_0$ varies with u_0/v_w as the beam
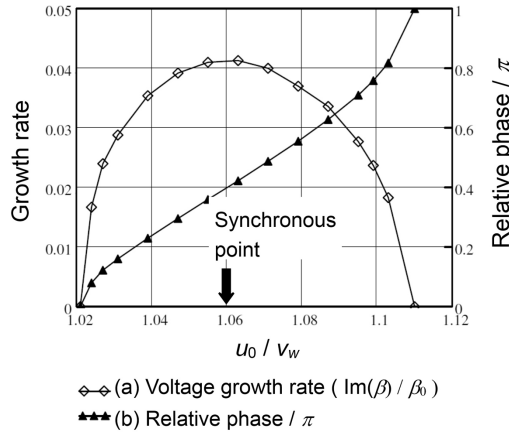
Figure 14.17: Effect of the velocity parameter u_0/v_w under small-signal conditions on (a) the normalised growth rate $\left(\text{Im}\left(\beta\right)/\beta_0\right)$ and, (b) the phase of the RF beam current relative to the helix voltage.

voltage is varied at constant perveance, using the same parameters as before. The normalised reduced plasma frequency is almost constant and equal to 0.06. Curve (b) in the figure shows the phase of the RF beam current relative to the wave on the slow-wave structure under small-signal conditions when the growing wave is fully established. The bunch is in the retarding phase of the field for all values of u_0/v_w for which the gain per unit length is non-zero. When the gain is maximum the phase of the bunch is close to the maximum retarding field. As the ratio u_0/v_w decreases, the bunch moves towards the neutral phase where the forces acting on the bunch tend to compress it.

The dependence of the saturated efficiency, and of the maximum normalised RF beam current, on the velocity parameter u_0/v_w are shown in Figure 14.18. The graphs were computed for a uniform section of a TWT by changing the beam voltage at constant perveance, using Worksheet 14.2 and the parameters in Table 14.1. The Pierce parameters $\left(C = 0.06 \text{ and } QC = 0.22\right)$ were approximately constant. Curve (a) shows that the saturated efficiency increases with u_0/v_w, as expected, and that it reaches a maximum close to the top of the range for which the tube has gain. Curve (b), obtained by multiplying the result from (14.53) by a capture ratio of 0.78, is a good fit to curve (a). Figure 14.11 shows that the RF beam current typically reaches a maximum somewhere between the plane at which the electrons begin to be trapped, and the plane of saturation. Curve (c) in Figure 14.18 shows the magnitude of the maximum value of the normalised RF beam current $\left(I_1/I_0\right)$ as a function of u_0/v_w. The maximum RF current increases as u_0/v_w decreases and the phase of the bunch moves closer to the neutral phase. The curve has a maximum towards the lower end of the range for which the tube has gain. We shall see that the conditions for maximum efficiency, and maximum RF current, have an important role in TWT design.
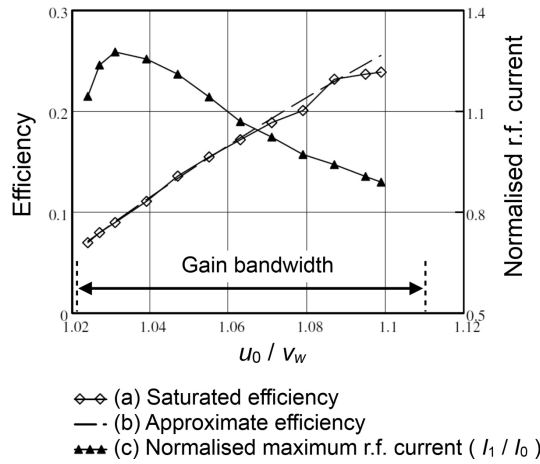
Figure 14.18: Variation of the large-signal properties of a TWT as the beam voltage is changed at constant perveance ($u_0/v_w = 1.06$ at the synchronous point): (a) saturated efficiency; (b) saturated efficiency estimated from (14.53) multiplied by 0.78 and; (c) maximum magnitude of the normalised RF beam current.

The most comprehensive theoretical investigation of the effect of changes in the operating point of a TWT was carried out by Rowe using the Pierce parameters [19]. The difficulty with using these parameters is that changes in the operating point generally result in changes in two or more of the parameters. Thus it is not straightforward to use Rowe's results to predict the behaviour of a tube when the operating point is changed. However it is found that the Pierce parameters C and QC are approximately constant when the beam voltage is changed at constant perveance. Rowe's results show that the efficiency then increases linearly with the Pierce velocity parameter b and reaches a maximum close to $b = 2$ which lies just below the point where the gain falls to zero. Figure 14.19 shows curves of efficiency against u_0/v_w, obtained by re-plotting Rowe's data [19]. The lowest point on each curve is the synchronous point. The figure also shows theoretical curves calculated using (14.53) for capture ratios of 0.8 and 0.6. The computed data shows the same trend as the theoretical curves. An increased value of C is needed to maintain the capture ratio as the space-charge is increased. Other results given by Rowe showed that the output power is reduced by 10 dB when the cold loss per circuit wavelength divided by the Pierce parameter C is increased by about 25 dB, in agreement with (14.60). These results, which were obtained using a ring model of the beam, also showed that the best efficiency was obtained with a beam filling factor (b/a) in the range 0.5 to 0.75 corresponding to an optimum value of $\gamma_e b$ in the range 0.75 to 1.0.

Further insight into the effects of variations in the operating point is given by Figures 14.20 and 14.21, which relate respectively to the conditions for maximum RF current and maximum efficiency (see also [44]). In the former, the RF current is greatest at saturation, only two of the sample electrons have not been captured by the wave and the velocities of the trapped electrons do not differ very much from v_w.
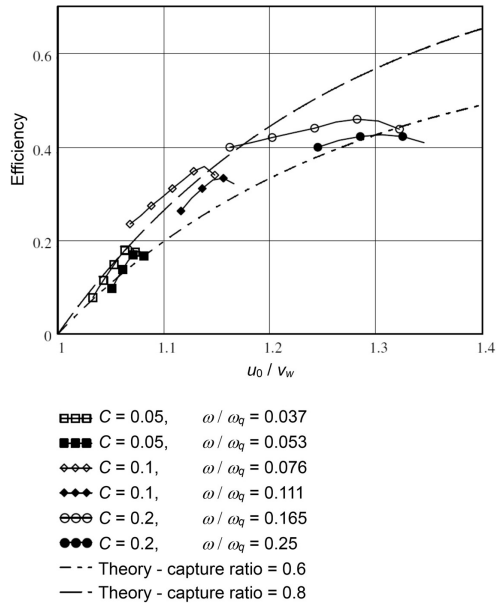
Figure 14.19: Computed efficiency of a TWT as a function of u_0/v_w derived from data given in [19] with C and ω_q/ω as parameters.

In the latter the maximum RF current occurs just after the plane at which trapping commences, and fewer of the electrons have been trapped. The trapped electrons have a much greater variation of velocity and at least two of the slowest electrons have drifted into the following bunch. Thus the bunc hes are much better formed in the first case but the small change in energy when they are trapped by the wave means that the efficiency is low.

An experimental study by Gerchberg and Niclas [45] showed that, under conditions for maximum gain, the phase of the circuit field is almost constant. Thus, the change in the relative phase seen in Figure 14.12 is almost entirely due to the changing phase of the RF beam current as the electrons lose energy. However, under conditions for maximum power output, the phases of both the current and the circuit wave changed with power level until the input power was a few dB below that required to saturate the tube. For greater drive levels the phase of the current remained almost constant while the phase of the circuit voltage continued to change until the two were in quadrature at saturation. This behaviour is reflected in kink in the phase curve in Figure 14.21(d).

14.3.4   Effect of a Sever

The tube modelled in the preceding sections is impracticable because the gain is too high for a single-section tube to be stable. Thus, in practice, it would be divided into two, or more, sections by severs. The efficiency of a tube is not affected by a sever provided that the small-signal gain of the output section exceeds a critical
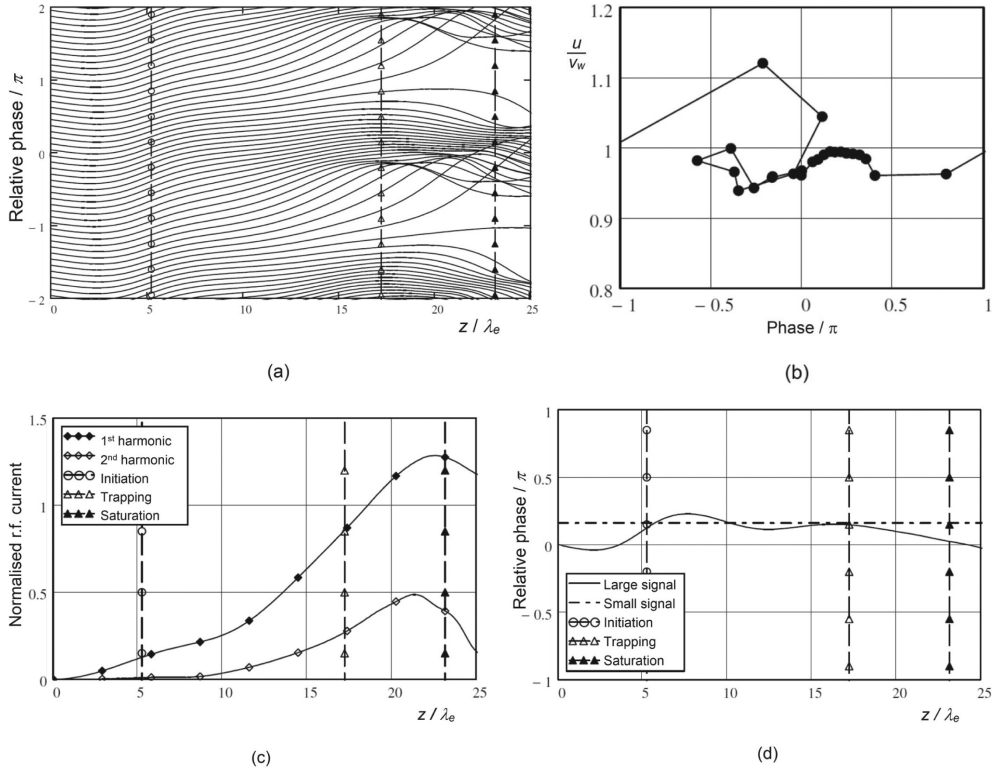
Figure 14.20: TWT behaviour at the working point at which the RF current is greatest $(u_0/v_w = 1.031)$: (a) modified Applegate diagram, (b) phase space plot at saturation, (c) normalised first and second harmonic currents, and (d) phase of the RF beam current relative to the wave.

value [46–48]. When the gain falls below this value the efficiency is rapidly reduced. Thus Scott found that the efficiency of a particular tube was halved when the small-signal gain of the output section was 5 dB less than the critical value. The most important physical reason for the reduction in efficiency is the dispersion of the electron bunches in the sever, and in the initial part of the output section where the electric field of the circuit is weak. To avoid this it is necessary for the sever to be placed at a position where small-signal conditions apply [42]. If the space-charge density in the beam is increased, then the length of the output section required to avoid a reduction in efficiency also increases [47]. From Figure 14.9 it can be seen that, for the tube modelled, the length of the output section required to re-establish the interaction after the sever corresponds to a small-signal gain of about 10 dB. Similarly, from Figure 14.10, we see that the small-signal gain from the plane at which trapping commences, to the plane of saturation, is about 15 dB. It seems reasonable to suppose that the efficiency will not be affected by the sever if the interaction is fully re-established by the plane at which trapping commences. Thus, in this case, the small-signal gain should be at least 25 dB in the output section. This
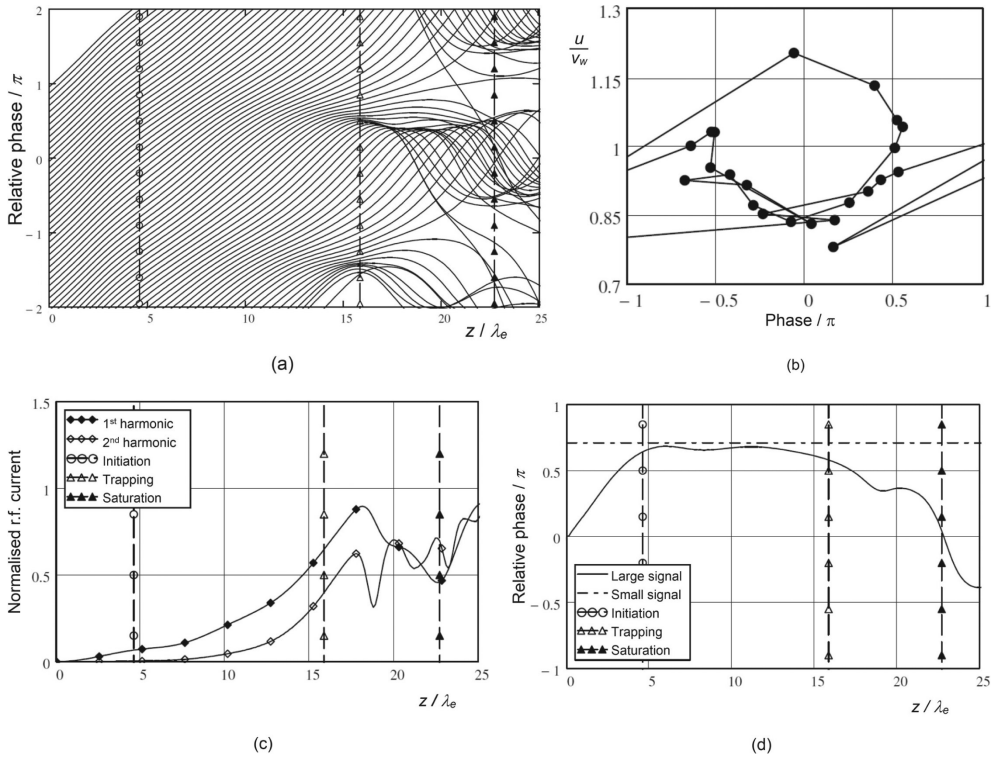
Figure 14.21: TWT behaviour at the working point at which the efficiency is greatest $(u_0/v_w = 1.095)$: (a) modified Applegate diagram, (b) phase space plot at saturation, (c) normalised first and second harmonic currents, and (d) phase of the RF beam current relative to the wave.

is comparable with the results in [47, 48]. If the interaction is linear at the plane of the sever, then the reduction in efficiency caused by dissipation of power in the sever load, or in the attenuator, is negligible.

When the tube is operated away from the synchronous point, the small-signal gain of the output section decreases, and the efficiency may be reduced compared with that of a single-section tube. This effect is greater when u_0/v_w is greater than the value at the synchronous point. Thus, in broad-band tubes the position of the sever should be chosen to achieve satisfactory performance at the highest frequency [48, 49]. It has also been shown that the efficiency of a severed TWT can be greater than that of an unsevered tube if the length of the section after the sever is slightly greater than the critical length [47, 49]. This is caused by a change in the phase of the bunches relative to the phase of the wave on the slow-wave structure after the sever. Overall, the length of the final section of a TWT is a compromise between the requirements to maximise the efficiency, and to minimise the gain ripples. This problem becomes more acute when the space-charge density is high.

14.3.5   Harmonics

At high drive levels the output of a TWT contains appreciable harmonic power. Figure 14.22 shows typical beam current waveforms at saturation for three different values of the velocity parameter u_0/v_w. The current waveforms were produced by superimposing the current pulses defined by the phases of successive pairs of electrons given by

$$I(\phi) = \begin{cases} \dfrac{\omega Q}{\phi_{j+1} - \phi_j} & \text{if } \phi_j \le \phi < \phi_{j+1} \\ 0 & \text{otherwise} \end{cases}.$$

(14.64)

where $\phi = \omega t$ and Q is the charge carried by one disc. It can be seen that all the bunches have a high harmonic content. As u_0/v_w increases the width of the bunch

(a)

(b)

(c)
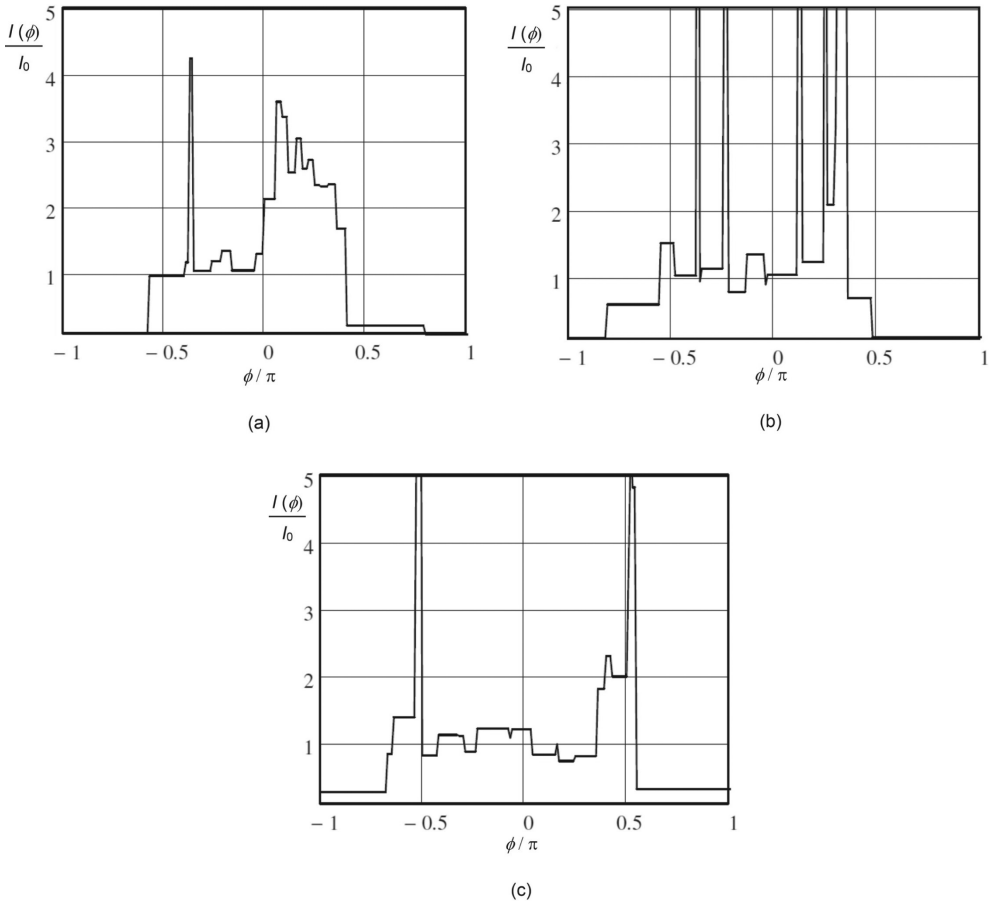
Figure 14.22: Typical beam current waveforms at saturation for (a) $u_0/v_w = 1.031$, (b) $u_0/v_w = 1.063$, and (c) $u_0/v_w = 1.095$.

increases and it has a tendency to split into two parts, one in the accelerating, and one in the decelerating, phase of the field. The helix can propagate waves at harmonics of the signal frequency and, although the coupling impedance decreases rapidly with increasing order of harmonic, some power is coupled to the helix. Thus there may be appreciable power output at the second, and higher, harmonic frequencies depending upon the dispersion and impedance characteristics of the helix.

There are special problems in broad-band helix tubes when the harmonics of signals at the low frequency end of the band may lie within the operating range of the tube. It is not unknown for the harmonic power to exceed the fundamental power [20, 50]. The total RF power at saturation is then approximately equal to the saturated output power of an equivalent tube in which harmonic generation is negligible. The level of harmonic generation is found to depend only slightly on the gain per unit length of the tube. However, it is strongly affected by the ratio of the phase velocities at the first and second harmonic frequencies $\left(v_{p2}/v_{p1}\right)$. There is strong second harmonic generation, with a corresponding reduction in the efficiency at the first harmonic frequency, when this parameter is slightly less than unity. However, when the ratio is reduced still further, the relative amplitude of the second harmonic is much smaller and the efficiency at the first harmonic is only a little less than that calculated when the effects of the second harmonic are ignored. Typically the increased level of the second harmonic occurs in the range $0.94 < v_{p2}/v_{p1} < 1.0$ [20]. The second harmonic power increases as the coupling impedance at the second harmonic frequency increases. This behaviour can be explained in qualitative terms by the effect on the second harmonic beam current of the interaction at both the first, and second, harmonic frequencies. If these two processes are in anti-phase with one another, some degree of cancellation will occur, and the amplitude of the second harmonic wave on the helix is reduced. The output of a tube is also affected by internal, and external, reflection of signals at the second harmonic frequency. It has been observed that the performance of a TWT can be altered by the injection of a second harmonic signal [51]. Thus, in general, the effects of the second harmonic (at least) must be included in large-signal simulations of helix TWTs. However, the dispersion of the helix in the tube modelled in the preceding sections is great enough for harmonic effects to be negligible.

14.3.6   Transfer Characteristics

In the previous sections the large-signal behaviour of a TWT has been considered as a function of distance along the tube at constant drive level. This is related to the terminal properties of the tube, represented by the amplitude and phase of the signal at the output relative to that at the input. Figure 14.23(a) shows a typical plot of output power against input power for a TWT (both in decibels relative to saturation) which is very similar to that for a klystron in Figure 13.7. At low drive levels the graph is a straight line having unit slope. The non-linearity in amplitude
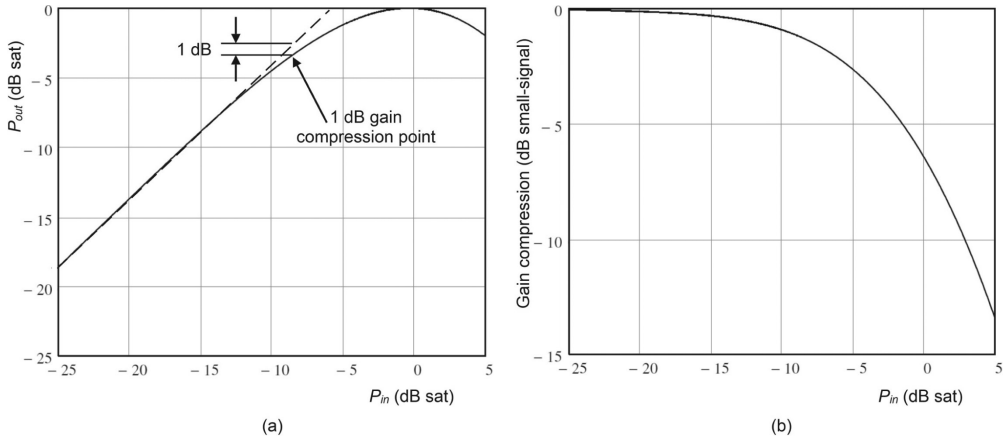
Figure 14.23: Typical characteristics of a TWT: (a) amplitude transfer, and (b) gain compression.
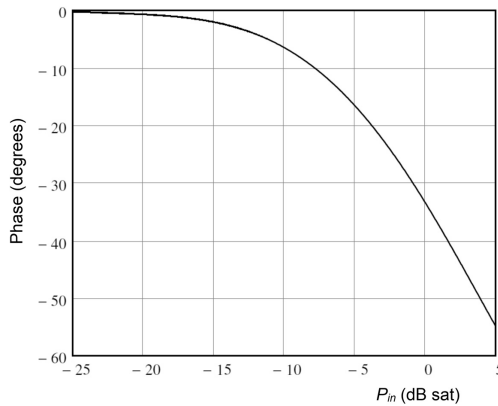
Figure 14.24: Typical phase transfer characteristic of a TWT.

is shown by the gain compression curve in Figure 14.23(b) which shows the gain relative to small-signal gain as a function of input power. The linear region of tube is often taken to extend up to the 1 dB compression point shown in Figure 14.23(a).

Because the mean electron velocity decreases as the power output is increased, the phase of the signal at the output of a TWT is a function of the input power, as illustrated in Figure 14.24. Under small-signal conditions the output phase must be independent of the input power. It can be seen from Figures 14.23 and 14.24 that the phase is a more sensitive indicator of non-linearity than the amplitude. This is confirmed by Figure 14.12 which shows that the phase of the current departs appreciably from the small-signal value within the region where the amplitude variation is linear. It is common to specify the linearity of a TWT in terms of the derivatives of the curves in Figures 14.23(a) and 14.24, as shown in Figure 14.25(a) and (b). These are the amplitude modulation to amplitude modulation (AM/AM) conversion
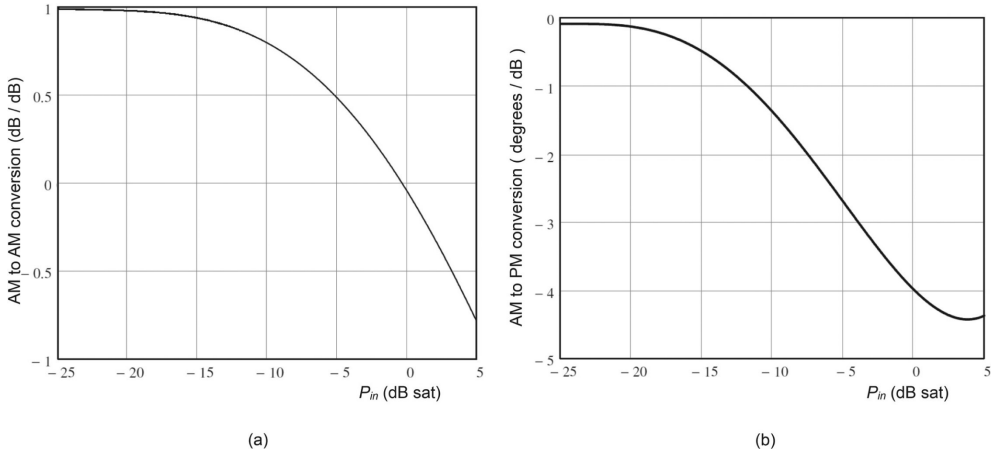
Figure 14.25: Typical derivative transfer characteristics of a TWT: (a) AM/AM conversion, and (b) AM/PM conversion.

expressed in dB/dB, and the amplitude modulation to phase modulation (AM/PM) conversion expressed in degrees/dB. For telecommunications applications it is desirable that both these parameters should be as small as possible (see Section 1.6.4).

A good approximation to the graph of output power against position can be found by assuming that the tube is linear up to the plane at which electron trapping commences, and that the gain then follows a cubic curve until the slope is zero at the plane of saturation. In this region the gain in decibels is given by

$$G_{dB} = 20\,\alpha\log(e)\left\{(z - z_t) - \frac{(z - z_t)^3}{3(z_s - z_t)^2}\right\}, \tag{14.65}$$

where α is the imaginary part of the small-signal propagation constant of the growing wave, z_t is the plane at which trapping commences, and z_s is the plane of saturation. The positions of these two planes can be established, using a large-signal disc model, with the assumption that the wave amplitude on the helix is given by small-signal theory. Then (14.65) can be used to correct for the gain compression in the large-signal region. The planes z_t and z_s are hardly changed by the revision of the profile of the wave on the helix and it is found that the model then has greater self-consistency (see Worksheet 14.2). Thus the gain compression at saturation is

$$\Delta G_{dB} = -\frac{20}{3}\,\alpha\log(e). \tag{14.66}$$

When the input power to a tube is increased by 1 dB the effect is equivalent to increasing the linear region of the tube by a section whose length would provide 1 dB of small-signal gain. This remains true provided that small-signal conditions exist over some part of the length of the tube. The small-signal gain per unit length is

$$\frac{dG_{0,dB}}{dz} = 20\,\alpha\log(e). \tag{14.67}$$

Then, differentiating (14.65) and dividing by (14.67) we find an expression for the AM/AM conversion in dB/dB

$$\frac{dG}{dz} \cdot \frac{dz}{dG_{0,dB}} = 1 - \frac{\left(z - z_t\right)^2}{\left(z_s - z_t\right)^2} \tag{14.68}$$

This gives a parabolic variation from unity in the linear region to zero at saturation, which may be compared with Figure 14.25(a). The AM/AM conversion is least when the distance between z_t and z_s is least, which generally occurs close to the synchronous point.

A simplified model of AM/PM conversion can be derived from small-signal theory by neglecting the backward wave [52]. Then, from (11.130), the forward wave is given by

$$\frac{\partial V_f}{\partial z} = -j\beta_0 V_f + \frac{1}{2}\mu_c Z_c \frac{\partial I_1}{\partial z}. \tag{14.69}$$

Now let

$$V_f = V_{f0}(z)\exp j\left(\omega t - \beta_0 z\right) \tag{14.70}$$

so that the phase of the voltage is referred to the phase of a wave travelling with the cold phase velocity of the helix. Substituting in (14.69) the wave voltage at $z + dz$ is given by

$$V_{f0}(z+dz) = V_{f0}(z) + \frac{1}{2}\mu_c Z_c \frac{\partial I_1}{\partial z} dz. \tag{14.71}$$

But we may write

$$\frac{\partial I_1}{\partial z} dz = \frac{\partial I_1}{\partial t} dt = j\omega I_1 \frac{dz}{u_0} \tag{14.72}$$

so that

$$V_{f0}(z+dz) = V_{f0}(z) + \frac{1}{2} j\omega\mu_c Z_c I_1(z)\frac{dz}{u_0}\exp\left(j\phi(z)\right), \tag{14.73}$$

where $\phi(z)$ is the phase of the RF beam current with respect to the wave. Then, if change of the amplitude of the wave in the element dz is small, the rate of change of the phase of V_{f0} with z is

$$\frac{d\Phi}{dz} = \frac{1}{2}\beta_e \mu_c Z_c \frac{I_1(z)}{V_{f0}(z)}\cos\left(\phi(z)\right), \tag{14.74}$$

which depends upon the ratio of the amplitude of the RF beam current to the amplitude of the wave voltage, and also upon the phase difference between them. These are determined by the non-linear interaction between the electron beam and the wave on the helix that also leads to the generation of harmonic currents. If the ratio of the phase velocities of the helix at the first and second harmonic frequencies lies

in the range for which appreciable harmonic interaction takes place then the AM/PM conversion is affected [53]. The AM/PM conversion tends to be small when the variation in $\phi(z)$ with z is small, and this occurs at low values of the ratio u_0/v_w [43].

14.3.7   Tapers

We have seen that, at saturation, the phase of the first harmonic of the beam current relative to the voltage wave is zero, and no further energy can be extracted from the electrons. However, if the slow-wave structure is modified, so that the phase velocity of the wave is reduced, then it is possible to extract more energy from the electrons, provided that they remain trapped by the wave. Such a change in the structure is referred to as a taper [16, 42, 54]. The change in the phase velocity in a taper can be continuous, or in a series of steps. It is found that the saturated power output of a tube can be increased in this way by about 2–3 dB. In general tapering the slow-wave structure reduces its characteristic impedance so that the amplitude of the wave corresponding to a given power is reduced. Thus the wave does not trap the electrons as effectively. A point is reached, eventually, where further tapering yields no improvement because the space-charge forces reduce the trapping of the electrons by the wave. Also, since some of the electrons are not trapped by the wave at saturation, the effect of the taper may be to move them into the accelerating phase of the field. Hence it is easiest to increase the efficiency of a TWT using an output taper if the beam perveance is low, and the bunches are well-formed [44]. Conversely, little improvement in efficiency can be obtained by reducing the phase velocity when an untapered tube is operated at the condition for maximum efficiency. However, it is sometimes found that the relative phase tends to π as saturation is approached in an untapered tube, operated at maximum efficiency, and that there is a big variation in the velocities of the electrons within the bunch. It has been shown that the efficiency can then be increased by the use of a positive taper, in which the phase velocity increases. This taper recaptures the bunch so that the relative phase decreases towards saturation in the usual manner [45, 55].

Early papers on tapering concluded that the best efficiency is obtained when the taper starts at the saturation plane of an untapered tube, and that the detailed profile of the taper does not have much effect on the efficiency achieved [16, 42]. However, this strategy tends to produce an AM/AM transfer curve which is concave just before saturation so that the tube is non-linear over a greater range of values of the input power. More recent studies have shown that the taper may start before saturation; typically at the plane of maximum RF current [44, 56]. Since it is usually desirable to optimise the performance of a tube over a band of frequencies, the taper should be designed to give the greatest improvement in efficiency at the end of the band where the untapered efficiency is low. In helix TWTs the variation in phase velocity is usually achieved by changing the pitch of the helix which means that the gaps between the turns and, hence, the coupling impedance, are reduced. Since only a single parameter is varied there is necessarily a compromise between the properties of the helix at different frequencies within the working band. However,

in coupled-cavity TWTs the dispersion curve can be altered in a number of ways so that it is possible to design tapers whose properties are closer to optimum throughout the band [44, 54, 57]. It is also possible to increase the efficiency of a tube by dividing the slow-wave structure into two or more sections which have uniform pitch. The sections are insulated from one another and have DC voltages applied to them so that the electron bunches are accelerated, moving them into the retarding phase of the wave [42]. This adds considerably to the complexity of the tube and does not appear to have been used except in experimental devices.

Tapering the slow-wave structure can also be used to enhance the performance of TWTs in other ways including [42, 58, 59]:

• reducing harmonic output power;
• improving linearity and reducing the AM/PM distortion;
• equalising the performance across the frequency band;
• improving tolerance to changes in the beam voltage.

It has already been observed that the AM/AM conversion of a tube depends upon the distance between the plane at which trapping commences, and the plane of saturation. In an untapered tube this distance is least close to the synchronous point, but it can be further reduced by the use of a tapered structure. We have also seen that the AM/PM conversion is strongly influenced by the change in the phase of the RF beam current with respect to the wave voltage. Thus, improved linearity can be achieved by adjusting the phase velocity of the slow-wave structure to keep the ratio u/v_w constant as the electrons lose energy to the wave [60]. Such a taper, known as a dynamic velocity taper, can achieve improved linearity together with an increase of 1 to 2 dB in the efficiency.

14.3.8   Stability

A particular problem with helix TWTs, especially those designed to work at high power levels, or broad bandwidths, is a tendency to backward wave oscillations (BWO) when the beam interacts with the $n = -1$ space-harmonic of the backward-wave on the helix. The starting conditions for oscillation can be found using small-signal theory (see Section 11.7). Backward-wave oscillations are characterised by a threshold beam current above which the oscillation occurs, and a sharply defined frequency which is dependent on the beam voltage. Equation (11.164) shows that the start-oscillation current is inversely proportional to the square of the length of the section. Thus, the risk of backward-wave oscillations increases sharply as the length of a section is increased. The stability of a tube with given small-signal gain can be increased by increasing the number of severs, with the penalty of additional tube length to compensate for the increase in sever loss. It has been shown, however, that the stability of a section can also be increased by introducing a step change in the helix pitch at some point in its length. If the average pitch of the section remains constant, then the effect on the forward-wave gain is small, whilst the stability against backward-wave oscillations is determined by the stability of each

uniform section [61–63]. Other techniques for increasing stability include the use of resonant loss using patterns printed on the support rods [64], and the use of external filter structures to damp the unwanted mode [65, 66]. This analysis assumes that only a single mode can propagate on the helix. To avoid the possibility of feedback oscillations involving higher-order modes it is important that they are strongly cut-off. The lowest mode of concern is the TE_{11} waveguide mode in the space between the helix and the shield and it is important that the cut-off frequency of this mode is above the working band of the tube.

In coupled-cavity TWTs oscillations can occur at the band edges if the beam voltage is reduced. At these frequencies the group velocity tends to zero so that the slow-wave structure supports a standing wave. The coupling impedance of a normal coupled-cavity structure tends to infinity at the upper cut-off frequency (see Figure 4.28(b)) [67]. Band-edge oscillations are caused by the interaction of the slow space-charge wave with a backward-wave space-harmonic close to the band edge. The frequency of oscillation varies only slightly with beam voltage. These oscillations present particular problems in cathode pulsed TWTs because the beam voltage passes through the critical range for oscillation on the rise and fall of the pulse. The result is a characteristic RF pulse shape with the oscillations appearing as 'Rabbits' Ears' on the rise and fall of the pulse. If this pulse were viewed using a spectrum analyser it would be seen that the 'Rabbits' Ears' were at a different frequency from the main part of the pulse. They can be suppressed by careful design, and by ensuring that the pulse rises and falls too quickly for the oscillation to build up. The start-oscillation conditions can be calculated using the methods described in [68] or by using a small-signal gain calculation program. Frequency selective loss is sometimes introduced to prevent these oscillations [69]. It has also been shown that they can be suppressed by tapering the slow-wave structure [70]. A third approach is to use a slow-wave structure designed to work in the inverted mode (see Section 4.5) so that the coupling impedance is zero at the upper cut-off frequency [71]. There is some reduction in performance because the coupling impedance at the centre of the pass-band is reduced. It is necessary to ensure that the beam cannot interact with the first higher-order mode, which then has infinite coupling impedance at its lower cut-off frequency. Backward-wave oscillations may occur in the higher order modes of a coupled-cavity structure, and it may also be possible for feedback oscillations to occur involving a higher-order mode [72]. Not all higher-order modes have appreciable interaction fields but it is necessary to check for possible problems of this kind when a tube is being designed.

14.4   TWT Design

The design procedure for TWTs is similar to that for klystrons described in Section 13.4. However, the great variety of types of TWT, and of the slow-wave structures incorporated in them, makes it much more difficult to make generalisations about

the process. The sections which follow discuss some of the chief issues involved in the design of TWTs of different types.

14.4.1 Case Study: An Octave Bandwidth Helix TWT

Helix TWTs with a bandwidth of about one octave have been used in many communications and radar systems. The conceptual design procedure can be illustrated by considering a tube with the specification in Table 14.2 [39].

The design calculations can conveniently be carried out using a spreadsheet so that effects of different design choices can be investigated quickly and easily (see Worksheet 14.3). The principal steps in the calculations for the tube in Table 14.2 are:

i)   Since the power output at the band edges is to be at least 1 kW, the RF power at the band centre must be at least 1.26 kW. The bandwidth required means that only a helix slow-wave structure is suitable. However, the mean output power of 10 W is well within the thermal capabilities of such a structure in this frequency band. The DC beam power required is found by dividing the RF power by an estimate of the efficiency of the tube at the band centre based on previous experience.

ii) The perveance of the electron gun is chosen in the range 0.1 to $2.0\ \mu A\ V^{-1.5}$. In this case a high perveance is desirable to reduce the risk of DC voltage breakdown, to ensure that the coupling impedance of the helix slow-wave structure is as high as possible, and that the electronic efficiency is high. The anode voltage (V_a) and the beam current (I_0) can then be calculated.

iii) The beam filling factor (b/a) is chosen in the range 0.5 to 0.7. Then the space-charge potential depression as the beam passes through the helix can be calculated using (7.8). The potential on the axis (V_h) is used to calculate the mean velocity of the electrons (u_0) (see Section 7.2).

iv) The normalised beam radius at the band centre $(\gamma_e b)$ is chosen in the range 0.5 to 1.0. A large value of this parameter leads to an appreciable radial variation in the RF field of the helix acting on the electrons, and reduced coupling impedance through the radial coupling factor. Since $\gamma_e b$ increases with frequency these effects are particularly important at the top of the frequency band of the tube. The beam radius (b) and the helix radius (a) can be calculated.

v)   It is now possible to calculate the plasma frequency of the beam and the Brillouin field. The magnetic field factor m is chosen to give the best compromise between

Table 14.2: Outline statement of requirements for a pulsed TWT

Frequency	2.0 to 4.0 GHz
Power output (1 dB band edges)	1.0 kW
Duty cycle	1%
Gain (at 1 kW)	33 dB

beam stiffness and the size and weight of the focusing magnet. This tube is suitable for PPM focusing and it is necessary to check that the peak field strength can be achieved without saturating the pole-pieces. Typically m is in the range 1.5 to 3.0. The magnetic field parameter α is chosen in the range 0.1 to 0.2 to give an acceptable ripple in the beam radius (see Section 7.6.1). The peak magnetic field and the period of the PPM stack can then be calculated.

vi) The plasma frequency reduction factor can now be calculated as a function of frequency. The saturated efficiency of the tube at the synchronous point is estimated using (14.56) multiplied by an estimate of the capture ratio (see Figure 14.19). This figure which is based on the helix voltage must be multiplied the ratio V_h/V_a to give the estimated ratio of the RF power to the DC power input to the tube. The preceding calculations can be repeated using this revised estimate now, or deferred until a more accurate figure can be calculated using a large-signal disc model.

vii) The synchronous frequency is chosen to lie within the operating band of the tube. It will usually lie close to the centre of the intended band, though, that is not always the best choice. The synchronous point can be adjusted, if necessary, to ensure that the maximum gain occurs at the centre frequency.

viii) The slow-wave structure is designed to achieve the correct phase velocity at the synchronous point. The sheath helix model can be used for initial calculations because the phase velocity is not seriously in error, and the chief effect of errors in the coupling impedance is in the gain per unit length. Assumptions must be made about the shield radius and the dielectric loading. Vane loading is not necessary for a tube having an octave bandwidth. From this information the pitch of the helix can be calculated together with the phase velocity, and coupling impedance, at frequencies within the working band.

ix) When the preceding steps have been completed sufficient information is available for the small-signal and large-signal performance to be calculated. One sever should be sufficient to ensure that the tube is stable. Initial estimates of the lengths of the two sections can be made by using the two-wave approximation to find the gain per unit length at the lowest frequency where the electrical length of the tube is least. The minimum requirement of 33 dB of gain at the 1 dB band edges implies a saturated gain of at least 34 dB at the band centre. If it is assumed that the launching loss, sever loss and gain compression are each 6 dB then the total small-signal gain required is 52 dB at the band centre. This can conveniently be achieved using two sections of equal length so that there is sufficient gain in the output section to avoid loss of efficiency. The length of the attenuator can be estimated by noting that tapered loss over three wavelengths is normally sufficient to ensure a good match. Thus the total length of the central attenuator should be at least six wavelengths.

Table 14.3 shows the comparison between the tube parameters calculated using the method described above (see Worksheet 14.3) and those given by [39]. The methods are essentially the same but there are some differences between the

Table 14.3: Comparison between the tube parameters given in [39] and those calculated using Worksheet 14.3

Parameter	Bliss [39]	WS14.3	Revised	Units
Estimated efficiency	15%	15%	26%	
Gun perveance	1.96	1.9	1.9	$\mu A\,V^{-1.5}$
Beam voltage (V_a)	7400	7200	5800	V
Beam current (I_0)	1.25	1.16	0.84	A
Filling factor (b/a)	0.5	0.5	0.5	
Synchronous helix voltage (V_h)	6100	6100	4900	V
Normalised beam radius $\gamma_e b$	0.75	0.7	0.7	
Beam radius	1.65	1.73	1.55	mm
Helix radius	3.3	3.46	3.10	mm
Brillouin field		0.059	0.059	T
Peak magnetic field	0.09	0.09	0.09	T
Magnet period	17.8	19.7	18.0	mm
Active length of helix	178	192	172	mm
v_p/c at synchronism	0.14	0.134	0.120	

design choices made. The estimated efficiency was 26% and the calculations were repeated using this estimate which was found to be self-consistent. The results of the revised calculations are shown in the fourth column of the table. Figure 14.26(a) shows the computed dispersion curves of the helix and the slow space-charge wave. Figure 14.26(b) shows the small-signal gain of a single-section tube 176 mm in length calculated using the two-wave model and assuming a fixed total of 12 dB for the launching and sever losses.

The performance of a single-section tube having the parameters in the fourth column of Table 14.3 was calculated at the centre frequency using Worksheet 14.2. The saturated output power was 1.1 kW corresponding to an electronic efficiency of 23%. The same result is obtained from (14.56) if the capture ratio is reduced from 0.7 to 0.6. These figures ignore the effects of the cold loss of the helix, of mismatches at the ends of the helix, and the radial motion of the electrons. Calculations using Worksheet 14.2 with the data in the second column of Table 14.3 gave a saturated efficiency of 22%. The design in [39] was for a tube with 33 dB gain at saturation. If the input and output sections of the circuit are assumed to be of equal length then the gain after the sever may be insufficient to allow the full efficiency to be reached. The design assumption of 15% efficiency may reflect that expectation. The high electronic efficiency of the tube means that little improvement could be achieved by tapering the helix.

The overall efficiency could be increased by the collector depression using the spent-beam curve computed by the large-signal disc model. A drift section can be added at the end of the tube to allow the bunches to disperse under the influence
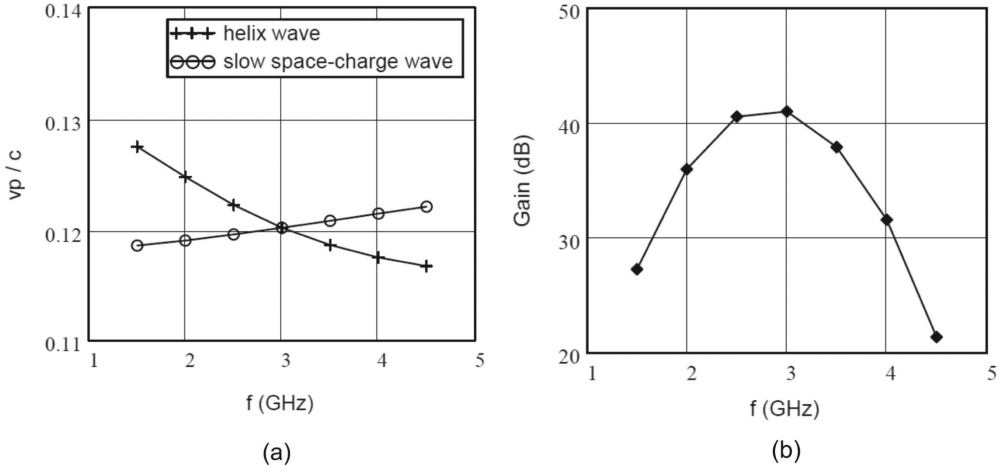
Figure 14.26: S-band helix TWT: (a) dispersion curves for the helix and for the slow space-charge wave, and (b) small-signal gain of a single section versus frequency.

of space-charge forces. However, debunching does not affect the overall shape of the spent-beam curve very strongly [73]. In practice, the shapes of the spent-beam curves for TWTs vary with the design of the tube, and with frequency within the operating bandwidth. Figure 14.27 shows the spent beam curve at saturation for the tube whose parameters are in the fourth column of Table 14.3. At present there is no simple model of the large-signal behaviour of a TWT corresponding to those used in this book for klystrons and IOTs. However, the maximum normalised electron energy is typically 1.1 over a range of drive levels, and the spent-beam curve for helix TWTs is often approximately linear, as shown by the dashed line in Figure 14.27 [73]. With these assumptions the power in the spent beam is

$$P_{spent} = \frac{1}{2} P_{dc} \left(1.1 + \frac{V_s}{V_a} \right), \tag{14.75}$$

where V_s is the energy of the slowest electrons in eV. The electronic efficiency is then

$$\eta_e = 1 - \frac{1}{2} \left(1.1 + \frac{V_s}{V_a} \right) \tag{14.76}$$

or

$$\frac{V_s}{V_a} = 0.9 - 2\eta_e. \tag{14.77}$$

For the tube modelled above $\eta_e = 0.23$ at saturation so that from (14.77) the normalised energy of the slowest electron is $V_s/V_a = 0.44$, which is close to the value shown in Figure 14.27. The energy of the slowest electron can be estimated in the same way when the tube is operated with reduced output power. This information
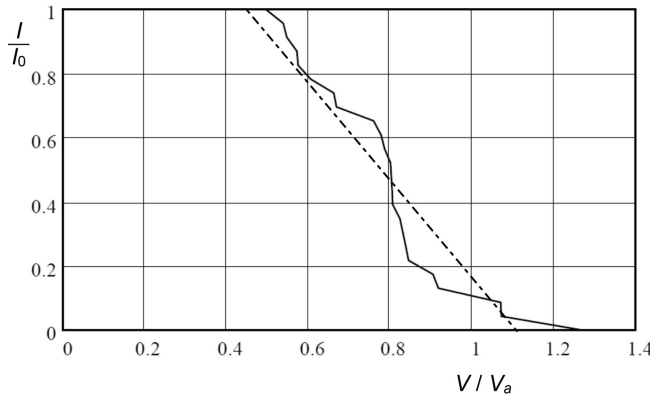
Figure 14.27: Spent beam curve at saturation for the S-band helix TWT.

can be used to construct an approximate model which reveals some of the features of depressed collector operation of a TWT (see Worksheet 10.4).

An alternative way of estimating the overall efficiency of a TWT with a depressed collector was suggested by Kosmahl [37]. The method is based on a trapezoidal approximation to the spent-beam curve such that the normalised energy of the slowest electron is given by

$$\frac{V_s}{V_a} \approx 1 - f(K)\sqrt[3]{\eta_e K}, \tag{14.78}$$

where K is the microperveance and $f(K)$ is an empirical function plotted in the paper. Fitting a second-order function to the curve plotted gives

$$f(K) \approx 0.073K^2 - 0.37K + 1.25. \tag{14.79}$$

Assuming that $K = 1.9$ and $\eta_e = 0.23$ in (14.78) gives $V_s/V_a = 0.39$ which is close to the figure calculated above. Further information about the design of TWT collectors can be found in [73–77].

14.4.2   Millimetre-Wave Helix TWTs

The design of millimetre wave TWTs with helix slow-wave structures is discussed in [36, 78–80]. The discussion which follows examines the design of the 90 W, 34 GHz pulsed TWT described in [80]. The conversion efficiency at saturation of this tube was 6.1% so that a DC beam power of 1.47 kW was required. The perveance chosen was $0.08\,\mu\text{A V}^{-1.5}$, corresponding to a beam current of 115 mA at a voltage of 12.75 kV. The beam voltage is unusually high for a helix TWT and it has a number of disadvantages: reduced coupling impedance; increased risk of voltage breakdown; and increased problems with backward wave oscillations. It does, however, allow the radii of the beam and the helix to be greater than if a lower voltage had been chosen. The normalised beam and helix radii were 0.7 and 1.4. These are

a little larger than those suggested by Cutler for maximum efficiency [38] but are close to the choice recommended in [39]. The Brillouin field for this beam is 0.124 T, while the RMS PPM field was 2.26 T leading to a much greater beam stiffness than would normally be used at lower frequencies. The reason for this is that in such a small beam the effects of transverse thermal velocities are much greater than those of space-charge forces [36, 78]. When thermal velocities are important, the magnetic field required to confine a beam with equilibrium radius b is given by [81]

$$B^2 = B_B^2 + \frac{8m_0 kT}{e^2 b^2} \left(\frac{r_c}{b}\right)^2 + \left(\frac{r_c}{b}\right)^4 B_k^2, \tag{14.80}$$

where B_B is the Brillouin field in the absence of thermal velocities. This equation is equivalent to (7.54) with the addition of the second term to account for the effects of thermal velocities. This term becomes increasingly important when an electron beam is scaled to higher frequencies, keeping the plasma frequency and the area convergence of the gun constant, because of the presence of the beam radius in the denominator. In the case of the tube described above the second term corresponds to a magnetic field of 0.054 T at a cathode temperature of 1100 °K, which is comparable with the Brillouin field. To minimise this term it is desirable to make the area convergence of the gun as small as possible, consistent with acceptable cathode loading. In this case the area convergence was 35:1 with a cathode loading of 2.4 A cm^2. Thus the maximum perveance of the electron gun in a millimetre wave tube is determined by the current which can be collimated by the maximum practical magnetic field.

To obtain stable PPM focusing the ratio of the plasma wavelength to the PPM period was 3.4. The 80 W, 30–35 GHz pulsed helix TWT described by Ohtomo et al. had very similar parameters [79]. At millimetre wavelengths the cold loss of the helix has a major effect on the efficiency. For example, the efficiency of a 0.5 W, 54 GHz, TWT was increased by a factor of 3 by plating the molybdenum helix with copper [78].

More recently the design of a 94 GHz pulsed TWT with saturated output power greater than 150 W was described in [82]. This tube had an inter-digital line slow-wave structure and a beam perveance of $0.058\,\mu\text{A V}^{-1.5}$, giving a conversion efficiency of 5.5%. The data given in this paper suggests that its parameters were approximately $\gamma_e b = 1.9$ and $\omega_q/\omega = 0.022$ giving a theoretical efficiency of 8.4% corresponding to an effective capture ratio (including the effects of losses) of 0.65. The overall efficiency was increased to over 15% by the use of a single stage depressed collector. Other work aimed at the development of TWTs at still higher frequencies is reviewed in [83].

14.4.3   High Efficiency Helix TWTs

A very important application of helix TWTs is in the field of satellite communications both as uplink and downlink power amplifiers. The requirements for space tubes are especially demanding since they must combine [5]:

- Operating lifetime greater than 15 years in order to keep the number of spare tubes carried to a minimum.
- High total electrical efficiency because of the limited power available on the satellite and the difficulty of dissipating waste heat in space.
- High linearity to avoid co-channel interference arising from intermodulation products.
- Low mass to minimise the cost of launching the satellite.

The advances in the performance of these tubes since they were first used in Telstar 1 in 1962 have been remarkable [5, 84–86]. The state of the art is represented by tubes giving continuous powers of 200 W at Ku-band (10.7–12.75 GHz) and 130 W at Ka-band (17–22 GHz) with overall efficiencies of 68% and 66% respectively [5]. It should be noted that the operating bandwidths of these tubes are small compared with the bandwidth of which helix TWTs are capable. A detailed discussion of the design of space TWTs is given in [87]. Table 14.4 shows some examples of tubes whose design has been described in the literature.

The lifetime of a TWT depends chiefly on the lifetime of the cathode, which is determined by the point at which the production of free barium is no longer sufficient to support the desired current density (see Section 18.5). M-type dispenser cathodes have achieved lifetimes greater than 100,000 hours (11.4 years) with continuous operation at up to 2 A cm^{-2} [86]. The lifetime of a cathode depends not only upon its construction but also on the current density. This is because a high current density requires a high operating temperature that causes the evaporation rate of the barium to increase. For this reason the electron guns used in space TWTs generally have low perveance because this allows the cathode loading to be small without high area convergence. A further limitation is set by the PPM focusing which means that the perveance must be less than about $0.5\,\mu\text{A V}^{-1.5}$, depending upon the type of permanent magnets used.

Table 14.4: Examples of TWTs for space communications

Frequency (GHz)	Power(W)	Overall efficiency	Application	Reference
4.0–4.2	3.5	34%	Downlink	[88]
17.7–21.2	75	45%	Downlink	[89]
27.5–29.5	40	25%	Uplink	[90]
5.85–6.425	3000	24%	Uplink	[91]
14.0–14.5	600	26%	Uplink	[91]
17.7–21.2	75	54%	Downlink	[92]
10.7–12.7	65	60%	Downlink	[93]
59–64	20	40%	Downlink	[94]
3.4–4.2	70	70%	Downlink	[95]
10.9–12.75	130	72%	Downlink	[95]

The overall efficiency of a space TWT depends upon the design of the slow-wave structure and also on the performance of the multi-element depressed collector. We have seen that the efficiency of a tube with constant helix pitch depends principally upon the beam perveance. For example, a tube with a beam perveance of $0.2\,\mu A\,V^{-1.5}$ and a uniform helix might have an electronic efficiency of about 11%. This could be increased by making the synchronous point below the centre frequency but with the penalty of increased non-linearity. The use of a simple taper would increase the efficiency to around 15% but, again, with a reduction in linearity. The design goals of high efficiency, and high linearity, are linked because it is necessary to operate the tube some way below its saturated output power to achieve acceptable linearity. Thus, to take an example, if output backoff of 3 dB is necessary to achieve acceptable linearity, the operating efficiency is half the saturated efficiency. A method for determining the operating parameters of a tube to give the best efficiency for a given linearity is described in [87].

It has been found that it is possible to increase the efficiency and the linearity of a TWT simultaneously by using the double taper profile as shown in Figure 14.28(a). We have seen that when the electron velocity is reduced relative to the phase velocity of the helix the maximum RF beam current increases, and the difference in the phase between small-signal conditions and saturation is reduced (see Figure 14.20). The same thing happens when the phase velocity of the helix is increased, while keeping the electron velocity constant. Thus, the effect of the section of helix with increased phase velocity is to draw the electrons into the bunch more effectively, in a manner similar to that of the detuned penultimate cavity in a klystron. It is found that the section can be designed so that the phase lag introduced decreases with increasing drive level. Hence the section can simultaneously increase the electronic efficiency, by improving the bunching, and increase the linearity by reducing the AM/PM conversion, compared with a similar tube with a conventional negative taper. The optimisation of a tube with a double taper is a complex multi-variable problem but a satisfactory initial design can be achieved systematically using a large-signal disk model [56, 96, 97]:

i)   A uniform TWT is modelled to find the helix phase velocities giving maximum gain (v_{p0}), maximum bunching (v_{p1}), and maximum efficiency (v_{p2}) at the centre frequency of the band over which the tube is to work. This is normally achieved by changing the pitch of the helix.

ii) A TWT is modelled at the centre frequency with the pitch profile shown in Figure 14.28(b). The input signal level is chosen to be the drive at which the tube is intended to saturate. The length z_0 is chosen to give at least 20 dB of gain so that the growing wave is fully established by the end of the section. The position z_1 is found within a second section of arbitrary length which maximises the amplitude of the first harmonic of the RF beam current.

iii) A TWT is modelled at the centre frequency with the pitch profile shown in Figure 14.28(a), and the same drive level as before. The length of the third section (z_{12}) is adjusted so that the phase of the bunch relative to the wave on
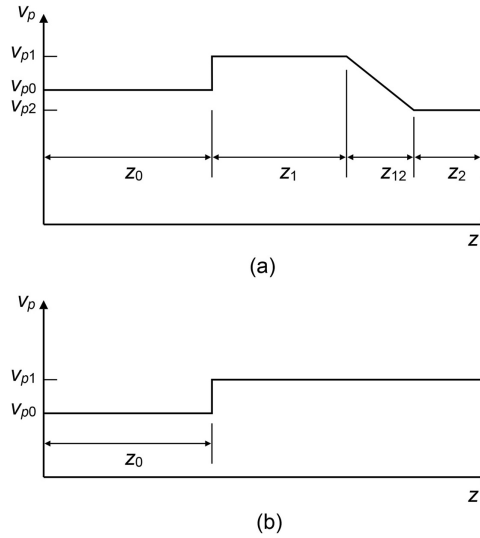
Figure 14.28: (a) Double-taper profile used in high-efficiency space TWTs, and (b) single-step taper used to design the taper in (a).

the helix is constant as the bunch loses energy. This is essentially a dynamic velocity taper (see Section 14.3.7). It has been found that a good initial approximation is to make this section at least three wavelengths long.

iv) The length of the fourth section is chosen so that the output power just reaches saturation at the centre frequency.

v) Finally, the phase velocity of the first section may be adjusted so that the gain is flattened across the frequency band, which may be up to 20% of the centre frequency.

All modern space TWTs employ double tapers of this kind, together with high efficiency multi-element depressed collectors [5, 84, 85], to achieve high operating efficiency at a working point which is backed off from saturation by a few decibels. The efficiency can be increased still further by reducing the cold loss of the helix by adding a high-conductivity coating, and by using support rods with low dielectric losses [92–94].

14.4.4   Ultra-Broad-band TWTs

We have seen that a TWT having an octave bandwidth can be designed using a helix slow-wave structure. The bandwidth is then limited by the dispersion of the structure. However, when the dispersion of the structure is modified by vane or dielectric loading (see Section 4.3.2) then synchronism between the beam and structure waves can be maintained over a much greater bandwidth. Tubes have been built in this way whose bandwidth exceeds three octaves [98]. The design of tubes of this kind presents its own challenges [99]:
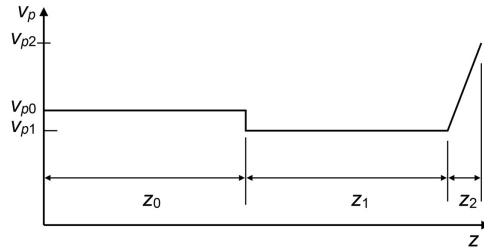
Figure 14.29: Double taper used in very wide bandwidth helix TWTs (copyright 2009, IEEE. Reproduced, with permission, from [103]).

- The slow-wave structure must have negative dispersion so that synchronism with the slow space-charge wave is maintained over a broad band.
- The gain in the output section must be equalised to avoid forward wave oscillations at frequencies where the gain is greatest, and the section must be long enough to give good efficiency at all frequencies.
- Means must be provided to suppress backward wave oscillations.
- The generation of harmonics must be suppressed.

The first method proposed for achieving a very broad bandwidth used a slow-wave structure in which a first section with positive dispersion was followed by a section with negative dispersion [98]. The section with negative dispersion (vane loading) was optimised for efficiency in the upper half of the band, while the unloaded first section provided gain at the lower frequencies. It was found that the harmonic performance was improved considerably if the relative phase velocities of the two sections were chosen correctly. This concept has recently been re-examined [100].

A uniform slow-wave structure can be used if it is designed so that the increased gain at high frequencies, caused by the greater electrical length, is compensated by reduction in synchronism and/or coupling impedance [101, 102]. Methods for suppressing harmonics and backward-wave oscillations have been reviewed in Sections 14.3.5 and 14.3.8 respectively. Tubes designed in this way tend to suffer from low efficiency at high frequencies because the electron bunch is divided into slow electrons which are trapped by the wave, and fast electrons which are not. The use of a positive phase velocity taper has been found to improve the efficiency at the top of the operating band and, therefore, the useful bandwidth of the TWT [55, 103]. This technique has also been found to reduce the output power at the second harmonic. Further improvements in the uniformity of the efficiency over a wide bandwidth have been obtained using the helix pitch profile, shown in Figure 14.29, in which a section of helix with reduced pitch to improve the bunching is followed by a short positive taper [103]. As an alternative to variation in the pitch of the helix the vane loading may be varied to achieve the same velocity profile [104].

14.4.5   Coupled-Cavity TWTs

General reviews of the performance and design of coupled-cavity TWTs are given in [2, 14, 105]. The design procedure for coupled-cavity TWTs follows the same general pattern as that for helix TWTs. It is necessary to decide, first, whether a forward fundamental or a space harmonic structure is required. The normalised beam radius $(\gamma_e b)$ should be around 0.6 to ensure good efficiency. The point at which the slow-space charge wave is synchronous with the circuit wave is chosen so that the hot band of the tube is centred within the cold pass-band of the structure. Typically, $\beta_0 p \sim 0.6\pi$ for a forward fundamental tube, and 1.5π for a space-harmonic tube. The width of the cold pass-band of the structure is usually around twice that of the hot band, but it is possible for a tube to have useful output power outside the cold pass-band of the structure. The slow-wave structure can then be designed, and the final design checked, by small-signal and large-signal gain calculations using discrete interaction models. If the tube is to be PPM focused, as is normal except at the highest power levels, then the PPM system design must proceed in parallel with the design of the slow-wave structure to ensure that the two are mutually compatible. Particular care is necessary to ensure that the tube is stable against oscillations at the cut-off frequencies of the structure, and in higher-order modes (see Section 14.3.8).

The design of tapers for coupled-cavity TWTs is discussed in [44]. In an untapered tube the frequencies for maximum gain and maximum efficiency are different. The gain at maximum efficiency may be too small for the tube to be saturated with the available drive power. An important function of the taper in a coupled-cavity TWT is, therefore, to align the gain and efficiency bands. Figure 14.30 shows, for a particular tube, the improvement in efficiency given by different step tapers at the frequencies for maximum bunching (f_b), and maximum gain (f_g). The improvement is greatest at maximum bunching, but also varies markedly with the change in the phase velocity in the taper (v_{pt}). At maximum gain the improvement is smaller and fairly insensitive to the change in phase velocity. At maximum efficiency (f_e)
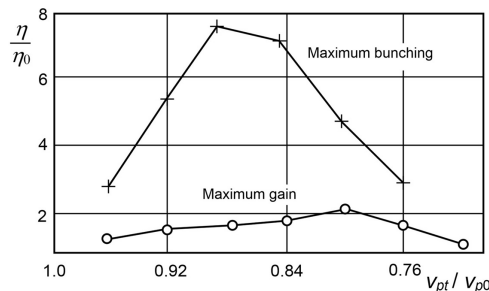
Figure 14.30: Increase in efficiency produced by a taper at the frequencies of maximum bunching and maximum gain
(copyright 1991, IEE, reproduced, with permission of The IET, from [44]).
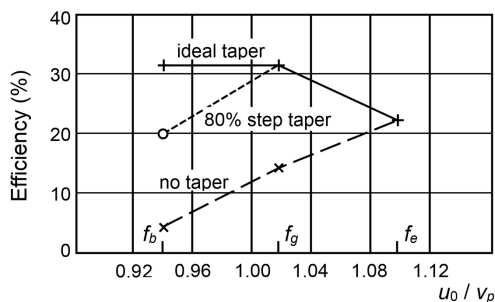
Figure 14.31: Saturated efficiency versus frequency curves for a tube without a taper, with an 80% one-step taper and with an ideal differential taper showing the conditions for maximum bunching, gain and efficiency
(copyright 1991, IEE, reproduced, with permission of The IET, from [44]).

a taper produced no further improvement. Clearly an ideal taper would give maximum efficiency at every frequency within the band of the tube. Figure 14.31 shows the effect on the efficiency of a single step taper, chosen to give the greatest efficiency at the frequency of maximum bunching, and of an ideal differential taper.

In a coupled-cavity TWT the dispersion diagram of the slow-wave structure can be changed by:

• changing the pitch of the structure;
• changing the resonant frequency of the cavities;
• changing the strength of the coupling between the cavities.

Figure 14.32(a) shows the dispersion diagram of an untapered space-harmonic structure, together with the dispersion line for an ideal differential taper in which the change in phase velocity is optimum at each frequency. Figure 14.32(b) shows the dispersion diagrams corresponding to each method of tapering compared with the standard structure. Evidently a dispersion diagram which is close to the ideal curve can be produced by using a combination of the different possible changes to the dimensions of the structure. This approach suggests that in designing a coupled-cavity TWT:

• The basic design should be such that f_b and f_e lie at the bottom and top of the frequency band of the tube. Then f_g will be in the centre of the band.
• The taper should be designed to give maximum efficiency at f_b.

It was found that good results were obtained by this method [44]. Table 14.5 shows some examples of TWTs with coupled-cavity, and other discrete interaction slow-wave structures, whose design has been described in the literature.

14.4.6   Hybrid Tubes

The bunched electron beams produced in both klystrons and TWTs are not ideal because not all the electrons are grouped into the bunches, and because of the spread
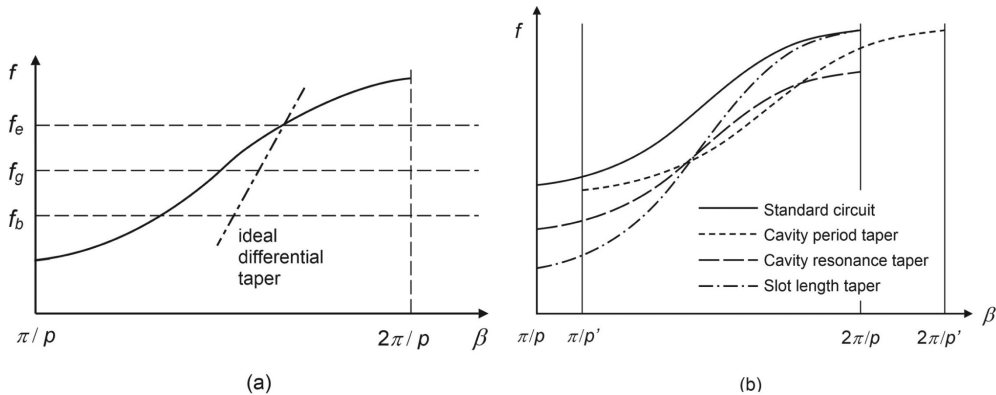
Figure 14.32: Dispersion diagrams for a typical coupled-cavity slow-wave structure showing: (a) the standard circuit and the ideal differential taper, and (b) the standard circuit (solid line), cavity period taper (dotted line), cavity resonance taper (dashed line) and slot length taper (chain dotted line)
(copyright 1991, IEE, reproduced, with permission of The IET, from [44]).

Table 14.5: Examples of TWTs with coupled-cavity and other discrete interaction slow-wave structures

Frequency GHz	Bandwidth %	Power	Efficiency %	Slow-wave structure	Reference
2.75	10	1.0 MW pk	13	Space harmonic	[106]
2.6–2.95	12	4.2 MW pk	37	Cloverleaf	[107]
5.4–5.9	9	4.1 MW pk	34	Cloverleaf	[107]
9.0	12	1.2 MW pk	25	Centipede	[108]
11.9–12.1	14	240 W cw	30	Space harmonic	[109]
3	14	160 kW pk	35	Stub-supported ring	[110]
5.25–5.75	9	26 kW pk	16	Inverted space-harmonic	[71]
90	20	100 W cw	2	Double ladder	[111]
22.3–24.8	7	1 kW	23	Space harmonic	[112]
28.3–30.0	2	500 W cw	31	Double staggered slot	[113]
90	1	400 W pk	15	Folded waveguide	[114]
33.5	7	700 W pk	7	Folded waveguide	[115]

of electron energies. A possible way of overcoming this is to create bunches which have near ideal properties before extracting the energy from them in a standing-wave, or a travelling-wave, output structure. The IOT is a simple example of a device of this kind which should have higher efficiency than a conventional klystron or a TWT. This concept was investigated theoretically by Meeker and Rowe [116] who showed

that high efficiencies should be attainable, but that they may be considerably reduced by spreading of the bunches caused by space-charge forces. In an experimental investigation Lichtenberg created bunches at 1 GHz using grid modulation, together with an inductively tuned cavity [117]. The bunches were injected into a helix slow-wave structure which was also fed with an input signal. It was found that when the system was optimised the peak efficiency was around 40%, which was double that of the TWT on its own. The key finding was that it was necessary for the initial bunches to have both current and velocity modulation to counteract the effects of space-charge. The propagation characteristics of a beam of this kind were investigated in [118].

It is difficult to form a prebunched beam using grid modulation at frequencies much above 1 GHz because of transit-time effects. A possible alternative is to combine the properties of a klystron bunching section with a TWT output section. We have seen that a very high degree of bunching with little velocity spread can be achieved in a well-design klystron. A hybrid tube of this kind is known as a Twystron®. This tube, operating at megawatt power levels, can achieve bandwidth, efficiency and gain flatness superior to those which can be realised in either a klystron or a TWT alone [105, 119–121].

The development of cold cathode technologies has offered the possibility of premodulating an electron beam at higher frequencies [122]. An experimental TWT of this kind developed an output power of up to 280 mW in the range from 3.9 to 6.75 GHz and this has been increased to around 50 W by subsequent development [123, 124].

References

[1] S. O. Wallander, 'Reflexions and gain ripple in TWT's', *IEEE Transactions on Electron Devices,* vol. 19, pp. 655–660, 1972.

[2] J. T. Mendel, 'Helix and coupled-cavity traveling-wave tubes', *Proceedings of the IEEE,* vol. 61, pp. 280–298, 1973.

[3] L. Sivan, *Microwave Tube Transmitters*. London: Chapman and Hall, 1994.

[4] J. H. Booske *et al.*, 'Traveling-wave tubes', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE, pp. 171–245, 2005.

[5] G. Faillon *et al.*, 'Microwave Tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer-Verlag, pp. 1–84, 2008.

[6] D. K. Abe and J. P. Calame, 'Advanced materials technologies', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimeter-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 649–689, 2005.

[7] E. D. Maloney, 'Developments in earth-station and satellite tubes reflect information demand', *Microwave Systems News,* pp. 48, 50, 52–55, March 1984.

[8] A. S. Gilmour, Jr., *Principles of Traveling Wave Tubes*. Norwood, MA: Artech House, 1994.

[9] D. K. Abe and J. P. Calame, 'Advanced material technologies', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 649–689, 2005.

[10] B. R. Potter *et al.*, 'High-power printed circuit traveling wave tubes', in *International Electron Devices Meeting*, pp. 521–524, 1973.

[11] D. M. Goebel *et al.*, 'Gain stability of traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 46, pp. 2235–2244, 1999.

[12] D. M. Goebel, 'Theory of long term gain growth in traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 47, pp. 1286–1292, 2000.

[13] D. Chernin *et al.*, '"Power holes" and nonlinear forward and backward wave gain competition in helix traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 50, pp. 2540–2547, 2003.

[14] J. F. Gittins, *Power Travelling-Wave Tubes.* London: English Universities Press, 1965.

[15] J. C. Slater, *Microwave Electronics.* New York: D. van Nostrand, 1950.

[16] N. H. Pond and R. J. Twiggs, 'Improvement of traveling-wave tube efficiency through period tapering', *IEEE Transactions on Electron Devices,* vol. 13, pp. 956–961, 1966.

[17] G. S. Kino *et al.*, 'Small-signal and large-signal theories for the coupled-cavity TWT', in *6th International Conference on Microwave and Optical Generation and Amplification*, Cambridge, UK, pp. 49–53, 1966.

[18] P. K. Tien *et al.*, 'A large signal theory of traveling-wave amplifiers', *Proceedings of the IRE,* vol. 43, pp. 260–277, 1955.

[19] J. E. Rowe, *Nonlinear Electron-Wave Interaction Phenomena.* New York: Academic Press Inc., 1965.

[20] N. J. Dionne, 'Harmonic generation in octave bandwidth traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 17, pp. 365–372, 1970.

[21] D. Chernin *et al.*, 'A three-dimensional multifrequency large signal model for helix traveling wave tubes', *IEEE Transactions on Electron Devices,* vol. 48, pp. 3–11, 2001.

[22] V. Srivastava and S. N. Joshi, 'Improved nonlinear model for multisignal analysis of helix TWTs', *IEE Proceedings H, Microwaves, Antennas and Propagation,* vol. 139, pp. 129–134, 1992.

[23] V. Srivastava and R. G. Carter, 'A fast large-signal model for coupled-cavity TWTs', *IEEE Transactions on Electron Devices,* vol. 35, pp. 2068–2076, November 1988.

[24] D. J. Connolly and T. A. O'Malley, 'Computer program for analysis of coupled-cavity traveling-wave tubes', NASA Lewis Research Center, Cleveland, OH, 1977.

[25] J. R. M. Vaughan, 'Calculation of coupled-cavity TWT performance', *IEEE Transactions on Electron Devices,* vol. 22, pp. 880–890, 1975.

[26] S. K. Datta *et al.*, 'Nonlinear Eulerian hydrodynamical analysis of helix traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 45, pp. 2055–2062, 1998.

[27] J. G. Wöhlbier *et al.*, 'The multifrequency spectral Eulerian (MUSE) model of a traveling wave tube', *IEEE Transactions on Plasma Science,* vol. 30, pp. 1063–1075, 2002.

[28] N. Brealey, 'Computer simulation of helix travelling wave tubes', in *High Power Microwave Generation and Applications*, SIF, Bologna, Italy, pp. 549–556, 1992.

[29] C. L. Kory *et al.*, 'Traveling-wave tube amplifier model to predict high-order modulation intersymbol interference', in *Third IEEE International Vacuum Electronics Conference*, Monterey, CA, pp. 308–309, 2002.

[30] Z. Zhu *et al.*, 'Particle-in-cell simulation of helix traveling-wave tube', in *IEEE International Vacuum Electronics Conference*, Rome, Italy, pp. 131–132, 2009.

[31] V. Srivastava and S. N. Joshi, 'One-dimensional nonlinear model for helix TWTs', *IETE Technical Review,* vol. 6, pp. 500–507, 1989.

[32] G. Dimonte and J. Malmberg, 'Destruction of trapping oscillations', *Physics of Fluids,* vol. 21, pp. 1188–1206, 1978.

[33] E. Buckingham, 'On physically similar systems; illustrations of the use of dimensional equations', *Physical Review,* vol. 4, pp. 345–376, 1914.

[34] M. J. Schindler, 'Can traveling-wave tubes be scaled? ', *Microwave Journal,* vol. 9, pp. 43–47, 1966.

[35] C. C. Cutler and D. J. Brangaccio, 'Factors affecting traveling wave tube power capacity', *Transactions of the IRE Professional Group on Electron Devices,* vol. PGED-3, pp. 9–24, 1953.

[36] W. E. Danielson *et al.*, 'Design of a 100 mW helix travelling-wave amplifier at 50 Gc/s', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 105, pp. 405–408, 1958.

[37] H. G. Kosmahl, 'How to quickly predict the overall TWT and the multistage depressed collector efficiency', *IEEE Transactions on Electron Devices,* vol. 27, pp. 526–529, 1980.

[38] C. C. Cutler, 'The nature of power saturation in traveling wave tubes', *Bell System Technical Journal,* vol. 35, pp. 841–876, 1956.

[39] E. E. Bliss, 'Traveling-wave tube design', in R. S. Burnap, ed., *Electron Tube Design*. Harrison, NJ: Radio Corporation of America, pp. 898–928, 1962.

[40] A. H. W. Beck, *Space-Charge Waves and Slow Electromagnetic Waves*. London: Pergamon Press, 1958.

[41] J. J. Caldwell, Jr. and O. L. Hoch, 'Large signal behavior of high power traveling-wave amplifiers', *IRE Transactions on Electron Devices,* vol. 3, pp. 6–17, 1956.

[42] O. Sauseng, 'Efficiency enhancement of travelling wave tubes by velocity resynchronisation', in *7th International Conference on Microwave and Optical Generation and Amplification*, Hamburg, Germany, pp. 16–29, 1968.

[43] L. Winslow, 'Phase velocity dispersion shaping as a design parameter in traveling wave tubes', in *International Electron Devices Meeting,* pp. 350A–350C, 1977.

[44] V. Srivastava and R. G. Carter, 'Design of phase velocity tapers in coupled-cavity TWTs', *IEE Proceedings H Microwaves, Antennas and Propagation,* vol. 138, pp. 469–474, 1991.

[45] R. W. Gerchberg and K. B. Niclas, 'The positively tapered traveling-wave tube', *IEEE Transactions on Electron Devices,* vol. 16, pp. 827–828, 1969.

[46] A. I. Denisov, 'Effect of sectionalisation on the saturation power of a traveling wave tube', *Radio Engineering and Electron Physics,* vol. 6, pp. 1416–1417, 1961.

[47] A. I. Denisov, 'The effects of sectionalization on saturation power of TWT for large values of the space charge parameter', *Radio Engineering and Electron Physics,* vol. 11, pp. 617–625, 1966.

[48] A. W. Scott, 'Why a circuit sever affects traveling-wave tube efficiency', *IRE Transactions on Electron Devices,* vol. 9, pp. 35–40, 1962.

[49] V. Srivastava and R. G. Carter, 'Determination of sever positions in a coupled-cavity TWTs', *IEE Proceedings H Microwaves, Antennas and Propagation,* vol. 138, pp. 55–60, 1991.

[50] A. Sangster, 'Traveling-wave interaction in structures with non-zero impedance at harmonics of the drive frequency', in *Proc. 6th Int. Conf. Microwave and Optical Generation and Amplfication*, pp. 125–130, 1966.

[51] O. Sauseng *et al.*, 'Reduction of intermodulation distortion with harmonic injection for wideband traveling-wave tubes', in *International Electron Devices Meeting*, pp. 411–414, 1975.

[52] E. Ezura and T. Kano, 'Measured and theoretical nonlinear phase distortion in traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 22, pp. 890–897, 1975.

[53] J. G. Wöhlbier and J. H. Booske, 'Mechanisms for phase distortion in a traveling wave tube', *Physical Review E,* vol. 69, p. 066502, 2004.

[54] D. J. Bates and A. W. Scott, 'The effect of circuit tapering on the efficiency bandwidth characteristics of dispersive traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 10, pp. 89–94, 1963.

[55] S.-S. Jung *et al.*, 'Positive phase-velocity tapering of broadband helix traveling-wave tubes for efficiency enhancement', *Applied Physics Letters,* vol. 80, pp. 3000–3002, 2002.

[56] V. Srivastava *et al.*, 'Design of helix slow-wave structures for high efficiency TWTs', *IEEE Transactions on Electron Devices,* vol. 47, pp. 2438–2443, December 2000.

[57] D. J. Connolly, 'Efficiency enhancement of coupled-cavity TWT's through cavity resonance tapering', *IEEE Transactions on Electron Devices,* vol. 26, pp. 1576–1580, 1979.

[58] L. Winslow, 'RF loss as a design parameter in traveling wave tubes', in *International Electron Devices Meeting*, pp. 374–376, 1974.

[59] H. Hirata, 'Analysis of phase and intermodulation distortion of a travelling-wave tube', *International Journal of Electronics,* vol. 83, pp. 249–270, 1997.

[60] H. G. Kosmahl and J. C. Peterson, 'A TWT amplifier with a linear power transfer characteristic and improved efficiency', NASA Lewis Research Center, 1984.

[61] B. Epsztein and G. Kantorowicz, 'Suppression of backward-wave oscillations in multi-kilowatt helix TWT's', in *3rd European Microwave Conference*, p. C.11.3, 1973.

[62] C. Hagström and O. Nilsson, 'Start oscillation current in tapered BWO'S', in *3rd European Microwave Conference,* p. C.11.5, 1973.

[63] J. L. Zhang *et al.*, 'Backward-wave suppression for broadband helix traveling-wave tubes using the phase velocity variation of output circuit', *IEEE Transactions on Electron Devices,* vol. 59, pp. 2263–2267, 2012.

[64] C. E. Hobrecht, 'Resonant loss for helix traveling wave tubes', in *International Electron Devices Meeting*, pp. 348–350, 1977.

[65] A. J. Sangster *et al.*, 'Backward-wave suppression in a very wide-band helix travelling-wave tube using a slow waveguide filter', *IEE Proceedings H, Microwaves, Antennas and Propagation,* vol. 138, pp. 79–85, 1991.

[66] C. E. Hobrecht and J. L. Putz, 'Traveling wave tube oscillation suppression', Varian Associates Inc., Palo Alto, CA, 1978.

[67] J. A. Ruetz, 'Resonant circuit oscillations in travelling-wave tubes', in *4th International Congress on Microwave Tubes*, Eindhoven, Netherlands, pp. 94–98, 1962.

[68] A. J. Bahr, 'A coupled-monotron analysis of band-edge oscillations in high-power traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 12, pp. 547–556, 1965.

[69] E. Glass, 'Suppression of spurious modes in high-power traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 30, pp. 1798–1806, 1983.

[70] C. C. Lo, *Studies of the Effect of Circuit Tapering on TWT Performance.* Stanford, CA: Stanford University Microwave Lab, 1964.

[71] J. R. Frey and I. Tammaru, 'A coupled-cavity TWT operating in the inverted slot mode', in *International Electron Devices Meeting*, pp. 504–506, 1981.

[72] A. Karp and G. T. Hunter, 'Higher order modes and instabilities in coupled-cavity TWT's', *IEEE Transactions on Electron Devices,* vol. 33, pp. 1890–1895, 1986.

[73] J. A. Dayton, Jr. *et al.*, 'Analytical prediction and experimental verification of TWT and depressed collector performance using multidimensional computer programs', *IEEE Transactions on Electron Devices,* vol. 26, pp. 1589–1598, 1979.

[74] A. C. Schram, 'TWT efficiency improvement using multi-stage collectors', *Microwave Journal,* vol. 18, pp. 31–33, 81, 1975.

[75] H. G. Kosmahl and P. Ramins, 'Small-size 81- to 83.5-percent efficient 2- and 4-stage depressed collectors for octave-bandwidth high-performance TWT's', *IEEE Transactions on Electron Devices,* vol. 24, pp. 36–44, 1977.

[76] H. G. Kosmahl, 'Modern multistage depressed collectors – a review', *Proceedings of the IEEE,* vol. 70, pp. 1325–1334, 1982.

[77] P. Ramins and T. A. Fox, '90- to 93-percent efficient collector for operation of a dual-mode traveling-wave tube in the linear region', *IEEE Transactions on Electron Devices,* vol. 26, pp. 1662–1664, 1979.

[78] H. L. McDowell *et al.*, 'A half-watt CW traveling-wave amplifier for the 5–6 millimeter band', *Proceedings of the IRE,* vol. 48, pp. 321–328, 1960.

[79] R. H. Ohtomo *et al.*, 'Recent advances in an EHF helix TWT', in *1984 International Electron Devices Meeting*, pp. 502–505, 1984.

[80] H. Sloley *et al.*, 'High power, high frequency helix TWT's', in *Military Microwaves' 86*, pp. 360–365, 1986.

[81] G. Herrmann, 'Optical theory of thermal velocity effects in cylindrical electron beams', *Journal of Applied Physics,* vol. 29, pp. 127–136, 1958.

[82] W. Gerum *et al.*, '94-GHz TWT for military radar applications', *IEEE Transactions on Electron Devices,* vol. 48, pp. 72–73, 2001.

[83] J. H. Booske *et al.*, 'Vacuum electronic high power terahertz sources', *IEEE Transactions on Terahertz Science and Technology,* vol. 1, pp. 54–75, 2011.

[84] D. S. Komm *et al.*, 'Advances in space TWT efficiencies', *IEEE Transactions on Electron Devices,* vol. 48, pp. 174–176, 2001.

[85] G. Kornfeld and E. Bosch, 'From history to future of satellite TWT amplifiers', *Frequenz,* vol. 55, pp. 258–262, 2001.

[86] J. D. Wilson *et al.*, 'Advances in space traveling-wave tubes for NASA missions', *Proceedings of the IEEE,* vol. 95, pp. 1958–1967, 2007.

[87] R. Strauss *et al.*, 'Traveling wave tubes for communication satellites', *Proceedings of the IEEE,* vol. 65, pp. 387–400, 1977.

[88] M. Bodmer *et al.*, 'The satellite traveling-wave tube', *Bell System Technical Journal,* vol. 42, pp. 1703–1748, 1963.

[89] J. Heney and R. Tamashiro, 'A 20 GHz, 75 watt, helix TWT for space communications', in *Proceedings of the Sixteenth Annual Electronics and Aerospace Conference and Exposition*, pp. 69–74, 1983.

[90] H. Hashimoto *et al.*, 'A 30 GHz 40 watt helix traveling-wave tube', in *International Electron Devices Meeting*, pp. 133–136, 1983.

[91] M. Takahashi *et al.*, 'Non-brazed helix TWT attained 3kW output at C-band and 600W at Ku-band', in *International Electron Devices Meeting*, pp. 167–170, 1986.

[92] R. Tamashiro and S. Aldana, '60 percent efficient K-band TWT using a new diamond rod technology', in *International Electron Devices Meeting*, pp. 187–190, 1989.

[93] H. Safa and A. Pelletier, 'Efficiency improvements in a 12 GHz-50 W space TWT', in *International Electron Devices Meeting*, pp. 191–194, 1989.

[94] G. K. Kornfeld *et al.*, '60-GHz space TWT to address future market', *IEEE Transactions on Electron Devices,* vol. 48, pp. 68–71, 2001.

[95] W. L. Menninger *et al.*, '70% efficient Ku-band and C-band TWTs for satellite downlinks', *IEEE Transactions on Electron Devices,* vol. 52, pp. 673–678, 2005.

[96] P. Thouvenin *et al.*, 'New helix tapers boost space TWT efficiency to 55%, broadband', in *International Electron Devices Meeting*, pp. 477–480, 1987.

[97] M. K. Alaria *et al.*, 'Design and development of helix slow-wave structure for Ku-band TWT', *IEEE Transactions on Plasma Science,* vol. 39, pp. 550–554, 2011.

[98] J. P. Rymer and M. J. Cascone, 'Three octaves with one TWT', in *1982 International Electron Devices Meeting*, pp. 30–31, 1982.

[99] P. L. Walchli, 'Multiple octave traveling wave tubes', in *1981 International Electron Devices Meeting*, pp. 707–707, 1981.

[100] S. K. Datta *et al.*, 'Analytical exploration of ultrawideband helix slow-wave structures using multidispersion phase velocity taper', *IEEE Transactions on Plasma Science,* vol. 37, pp. 311–316, 2009.

[101] H. J. Kim *et al.*, 'Experimental investigation of broadband vaned helix traveling-wave tube', *Japanese Journal of Applied Physics,* vol. 45, pp. 292–299, 2006.

[102] S.-S. Jung *et al.*, 'Wide-band semivane and heavily dielectric loaded helix traveling-wave tubes', *IEEE Transactions on Plasma Science,* vol. 30, pp. 1009–1016, 2002.

[103] T. K. Ghosh *et al.*, 'Design of helix pitch profile for broadband traveling-wave tubes', *IEEE Transactions on Electron Devices,* vol. 56, pp. 1135–1140, 2009.

[104] A. B. Danilov *et al.*, 'A method for reducing the second harmonic level in wideband traveling wave tubes', *Journal of Communications Technology and Electronics,* vol. 58, pp. 353–356, 2013.

[105] A. Staprans *et al.*, 'High-power linear-beam tubes', *Proceedings of the IEEE,* vol. 61, pp. 299–330, 1973.

[106] M. Chodorow *et al.*, 'The design and characteristics of a megawatt space-harmonic traveling-wave tube', *IRE Transactions on Electron Devices,* vol. 6, pp. 48–53, 1959.

[107] J. A. Ruetz and W. H. Yocom, 'High-power traveling-wave tubes for radar systems', *IRE Transactions on Military Electronics,* vol. MIL-5, pp. 39–45, 1961.

[108] T. Roumbanis *et al.*, 'A megawatt X-band TWT amplifier with 18% bandwidth', in *High Power Microwave Tubes Symposium*, Fort Monmouth, NJ, pp. 114–129, 1962.

[109] H. G. Kosmahl *et al.*, 'High-efficiency, 200-watt, 12-gigahertz traveling wave tube', NASA Lewis Research Center, Cleveland, OH, 1974.

[110] H. Desmur *et al.*, '160-kW pulsed S-band TWT', *Microwave Journal,* vol. 17, pp. 53–56, 1974.

[111] B. G. James, 'Coupled-cavity TWT designed for future mm-wave systems', *Microwave Systems News,* vol. 16, September 1986.

[112] J. D. Wilson *et al.*, 'A high-efficiency ferruleless coupled-cavity traveling-wave tube with phase-adjusted taper', *IEEE Transactions on Electron Devices,* vol. 37, pp. 2638–2643, 1990.

[113] J. R. Legarra *et al.*, 'A 500-W coupled-cavity TWT for Ka-band communication', *IEEE Transactions on Electron Devices,* vol. 52, pp. 665–668, 2005.

[114] A. J. Theiss *et al.*, 'High-average-power W-band TWT development', *IEEE Transactions on Plasma Science,* vol. 38, pp. 1239–1243, 2010.

[115] H. Gong *et al.*, 'Experimental investigation of a high-power Ka-band folded wave-guide traveling-wave tube', *IEEE Transactions on Electron Devices,* vol. 58, pp. 2159–2163, 2011.

[116] J. E. Rowe and J. G. Meeker, 'Interaction of premodulated electron streams with propagating circuits', *International Journal of Electronics,* vol. 9, pp. 439–466, 1960.

[117] A. J. Lichtenberg, 'Prebunched beam traveling-wave tube studies', *IRE Transactions on Electron Devices,* vol. 9, pp. 345–351, 1962.

[118] J.-G. Jin *et al.*, 'Propagation characteristics of a premodulated electron beam', *IEEE Transactions on Plasma Science,* vol. 26, pp. 794–798, 1998.

[119] A. LaRue, 'Development of a broadband S-band amplifier', Varian Associates Inc., Palo Alto, CA, 1972.

[120] P. S. Matthews *et al.*, 'Comparison of the performance of a travelling-wave tube and a hybrid tube designed for 12% bandwidth, multi-megawatts output at L-band', in *IEEE International Electron Devices Meeting*, pp. 134–137, 1975.

[121] P. S. Matthews *et al.*, 'The centipede structure in practical megawatt tubes', in *8th European Microwave Conference*, Paris, France, pp. 753–760, 1978.

[122] D. R. Whaley *et al.*, 'Application of field emitter arrays to microwave power amplifiers', in *27th IEEE International Conference on Plasma Science*, p. 122, 2000.

[123] D. R. Whaley *et al.*, 'Experimental demonstration of an emission-gated traveling-wave tube amplifier', *IEEE Transactions on Plasma Science,* vol. 30, pp. 998–1008, 2002.

[124] D. R. Whaley *et al.*, 'Operation of a low-voltage high-transconductance field emitter array TWT', in *IEEE International Vacuum Electronics Conference*, Monterey, CA, pp. 78–79, 2008.

15 Magnetrons

15.1 Introduction

This chapter, and the next, are concerned with tubes which employ electron beams in crossed electric and magnetic fields, as described in Chapter 8. Magnetron oscillators are considered in this chapter, beginning with a review of the basic principles of their operation. We shall refer to these devices as 'magnetrons' for convenience. Although both linear and circular geometries are possible (see Figure 8.1), only the second of these is normally used because it minimises the volume in which a uniform magnetic field must be maintained. In addition, all high power devices are based on electron emission from a cylindrical cathode as illustrated in Figure 8.1(d) because it is not possible to make stable, high current, injected beams. The rotating cloud of electrons interacts with the RF field of a slow-wave structure which also forms the anode as shown in Figure 15.1. A number of different forms of anode are in use, and these are discussed in Section 15.3. There follows a review of the properties of magnetrons, and a discussion of ways of modelling them. The final section of the chapter is concerned with the design of magnetrons. Crossed field amplifiers, which are closely related to magnetrons, are dealt with in Chapter 16.

The magnetron was the first type of microwave tube to be used in systems, following its successful development as the power source for radar during World War II. Modern tubes have the same basic features as the early tubes but are the result of over seventy years of engineering development [1–3]. An overall efficiency of 50% is typical for a pulsed tube but figures as high as 90% are achieved in CW tubes designed specifically for high efficiency. In comparison with a linear beam tube, at the same frequency and power level, the magnetron is smaller, simpler, and cheaper, and works at a lower voltage. The tuning range is small, usually 2–10%. Magnetrons are, therefore, preferred to linear beam tubes when an oscillator is suitable and only a narrow tuning range is needed. The main applications of magnetrons are in radar, in domestic and industrial microwave ovens, and in medical linear accelerators [3–5]. Relativistic magnetrons, which deliver powers up to several GW with pulse durations of some tens of nanoseconds, lie outside the scope of this book [6, 7].
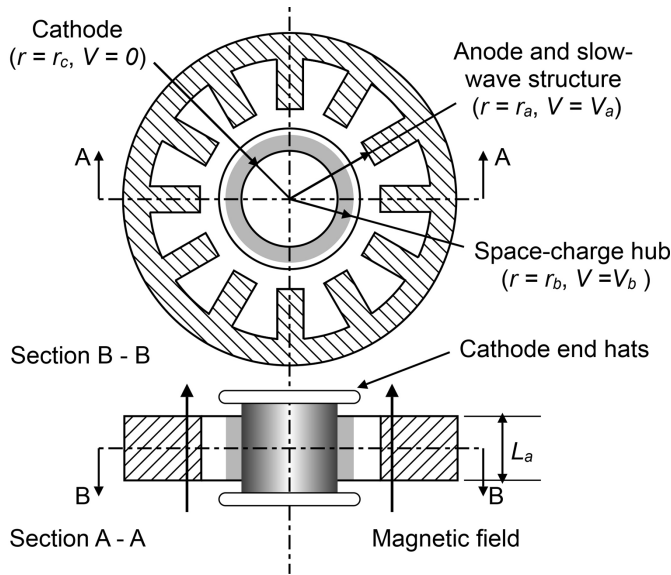
Figure 15.1: General arrangement of a magnetron.

15.2   Basic Principles

Figure 15.1 shows the general arrangement of a magnetron. The electrons are emitted from a cylindrical cathode. In most types of magnetron an indirectly heated matrix cathode is used (see Section 18.5.3). The exceptions are tubes for domestic or industrial microwave ovens in which the cathode is a directly heated thoriated tungsten helix (see Section 18.5.1) [4]. The axial length of the interaction region (L_a) is normally small compared with the free-space wavelength to avoid variation of the interaction in the axial direction. The cathode end hats limit the flow of electrons in the axial direction. The cathode is usually supported by a coaxial extension (see Figure 15.25). The output power may be removed from the anode by probe, or iris, coupling to a waveguide and window at right angles to the axis of the tube. Alternatively, one or more of the anode vanes may be connected to an antenna on the axis of the tube which radiates into the external waveguide though a domed window (see Figure 15.26). The other possible arrangement, in which the cathode is outside the anode, suffers from the disadvantage that the power density on the anode is greater [8].

The magnetron is based upon the cut-off magnetron diode, discussed in Chapter 8. Before oscillations commence much of the space between the cathode and the anode is occupied by a rotating cloud of electrons. It is believed the cloud is in a state of oscillation comprising azimuthal space-charge modes with a background of broadband noise (see Section 8.6). However, the properties of magnetrons can be understood by assuming that the cloud takes the form of a uniform space-charge hub whose outer radius is r_b, as shown in Figure 15.1. A magnetron

is always operated at an anode voltage which is less than the Hull cut-off voltage (V_H) given by (8.59). Thus, no current can flow through the tube in the absence of an interaction with the RF field of the anode. The properties of the space-charge hub are reviewed in Section 15.2.1.

The anode supports a number of resonant modes which can be decomposed into pairs of counter-rotating slow electromagnetic waves. These have radial and tangential components of the electric field, as discussed in Section 15.2.2. If the angular velocities of the electrons are approximately synchronous with one of the travelling waves on the anode a collective interaction can take place. The modes are initially excited by the fluctuations in the electron cloud. There is competition between different possible modes of oscillation which grow at different rates. It is therefore important to ensure that the desired mode eventually dominates. Because the slow-wave structure is a closed circle the device is an oscillator. The strength of the interaction grows until it is limited by non-linear effects.

The nature of the interaction between the electrons in the hub and the RF field of the anode is illustrated in Figure 15.2. This shows the motion of the electrons in a frame of reference which is moving with their steady velocity u_0 which is greater than the phase velocity of the RF electric field of the anode. Planar geometry has been assumed for simplicity. In this frame of reference the steady electric and magnetic forces are in balance. Electrons, which are on the surface of the electron cloud, move under the combined influence of the RF electric field, and the static magnetic field, as shown by the arrows in Figure 15.2(a). Their new positions and the new shape of the electron cloud are shown in Figure 15.2(b). The charge density in the electron cloud is approximately constant so that the electrons are concentrated regions of retarding field. The wave on the perturbed cloud resembles a slow space-charge wave on a Brillouin focused linear electron beam. Electrons which are in a retarding phase of the tangential RF electric field of the anode transfer energy to it. Because their tangential velocity is reduced there is an imbalance between the outward electric force and the inward magnetic force on them and they drift outwards towards the anode. In so doing they move into a region where their potential energy is lower so that they gain kinetic energy and are able to maintain synchronism with
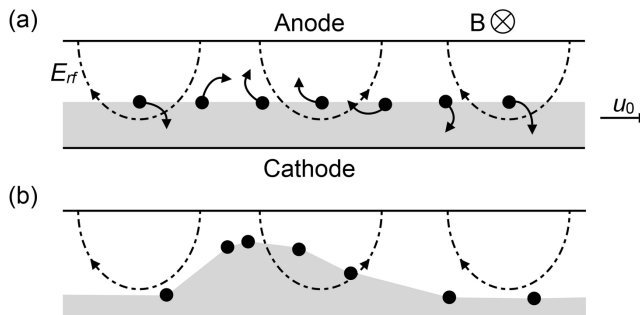
Figure 15.2: Motion of electrons in a magnetron.

the rotating RF field. Conversely, those electrons which are accelerated by the RF field move inwards towards the cathode. The perturbation of the charge cloud, and the growth in the magnitude of the RF electric field, are mutually reinforcing. It is found that the electrons can only reach the anode as a result of the interaction if the anode voltage is greater than the *threshold voltage* (V_T) which depends upon the magnetic field. The RF power delivered to the anode is very small until this voltage is exceeded. Thus a working magnetron operates at an anode voltage such that $V_T < V_a < V_H$. The equation for the threshold voltage is derived in Section 15.2.3.

When stable oscillation has been established the electron bunches take the form of 'spokes' of charge based on a space-charge hub, as shown in Figure 15.3. A more detailed view is shown in Figure 15.36. The spokes move with an angular velocity which is synchronous with the wave, and the potential energy lost by the electrons is converted directly into RF energy in the field of the anode. Those electrons which return to the cathode may strike it with high energies and liberate secondary electrons. The back-bombardment of the cathode caused in this way means that the emission from the cathode is a combination of thermionic and secondary electron emission. It is difficult to determine the fraction of electrons falling into each category, but it is known that this can have important effects on the properties of the magnetron [9]. Electron bombardment also causes heating of the cathode, which may be compensated by reducing the current in the cathode heater. In some cases it is possible to operate a magnetron without any external heating of the cathode. The way in which the electronic efficiency depends upon the magnetic field is explored in Section 15.2.4.

The description of the operation of a magnetron in terms of an inner space-charge hub and spokes of charge is satisfactory as a basis for understanding the operation of a magnetron and the principal features of its performance. But it is
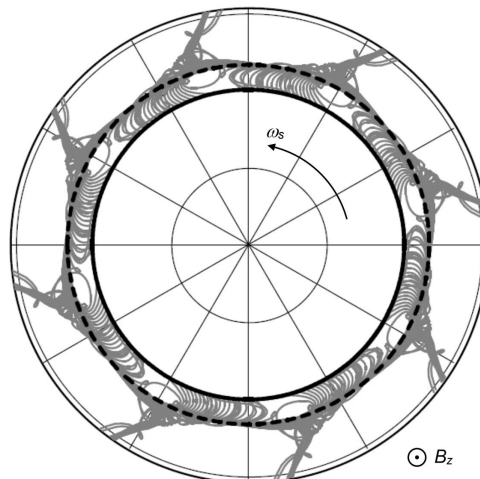
Figure 15.3: The space-charge hub and spokes in an oscillating magnetron.

important to realise that the actual behaviour of the electrons is much more complex than is assumed in this simple model. We saw in Chapter 8 that, in practice, the electrons execute complex cycloidal motions about guiding centres which are in uniform motion. However, this motion is usually rapid compared with the drift of the guiding centres so that it is effectively averaged out and does not play a major role in the interaction. This is the justification for using a uniform hub as a model. It is assumed that the RF field of the anode does not act on the electrons within the hub. Its properties are therefore unchanged when electrons are drawn from it to form the spokes. A more detailed discussion of the interaction in a magnetron can be found in Section 15.6.

15.2.1   Hub Model

The DC potential within the space-charge hub is given by (8.68)

$$V(r) = \frac{e}{8m_0} B_z^2 r^2 \left(1 - \frac{r_c^2}{r^2}\right)^2 . \tag{15.1}$$

We recall that this equation does not depend on any assumptions made about the space-charge distribution. The potential at the surface of the hub (V_b) is found by setting $r = r_b$ in this equation. Outside the hub the potential is given by (8.72)

$$V(r) = \frac{e}{4m_0} B_z^2 r_b^2 \left(1 - \frac{r_c^4}{r_b^4}\right) \cdot \ln\left(\frac{r}{r_b}\right) + \frac{e}{8m_0} B_z^2 r_b^2 \left(1 - \frac{r_c^2}{r_b^2}\right)^2 . \tag{15.2}$$

Then the radial component of the DC electric field outside the hub is

$$E_0 = -\frac{e}{4m_0} B_z^2 r_b^2 \left(1 - \frac{r_c^4}{r_b^4}\right) \frac{1}{r} . \tag{15.3}$$

The radius of the hub can be determined in any particular case by numerical solution of (15.2) when $V = V_a$ at $r = r_a$. The potential on the surface of the hub is then found by substitution into (15.1). The angular velocity of an electron on the surface of the hub is found to be

$$\omega_b = \frac{\sqrt{2(e/m_0)V_b}}{r_b} \tag{15.4}$$

by using the principle of conservation of energy. It should be noted that this angular velocity is not normally equal to the angular velocity of the travelling electromagnetic wave on the anode.

15.2.2   Interaction Field

Although practical magnetrons have cylindrical geometry, it is helpful to examine first the planar geometry shown in Figure 15.4 because that is sometimes used as an
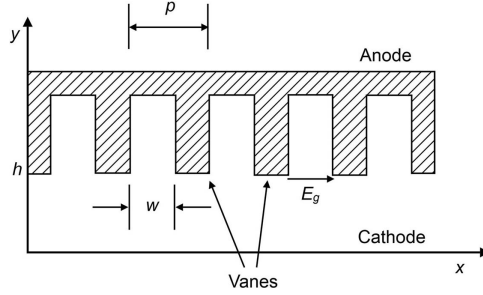
Figure 15.4: Arrangement of a planar magnetron.

approximation to the cylindrical case. The design of magnetron anodes is discussed in Section 15.3. For the present it is sufficient to note that the slow electromagnetic wave is generated by the RF voltages between the tips of a regular series of metallic vanes.

Let us assume that a wave travelling in the x direction with frequency ω and propagation constant β_0 has uniform electric field amplitude E_g between the vane tips. Then, from (4.37), the tangential electric field of the mth space-harmonic wave on the surface of the anode can be written

$$E_{xm}(x,h) = E_g \frac{w}{p}\left(\frac{\sin(\beta_m w/2)}{\beta_m w/2}\right)\exp j(\omega t - \beta_m x), \tag{15.5}$$

where $\beta_m = \beta_0 + 2m\pi/p$ and the other symbols are defined in Figure 15.4. Since the phase velocity of the wave is much less than c, and the transverse dimensions are small compared with the free-space wavelength, we can use the quasi-static approximation. Then the electrostatic potential of the mth space-harmonic wave on the surface of the anode is given by

$$V_m(x,h) = -\int E_{xm}(x,h)\,dx$$
$$= -j\frac{wE_g}{\beta_m p}\left(\frac{\sin(\beta_m w/2)}{\beta_m w/2}\right)\exp j(\omega t - \beta_m x). \tag{15.6}$$

Thus the amplitudes of the space-harmonics of the potential wave on the surface of the anode are

$$V_{sm} = \frac{w}{\beta_m p}\left(\frac{\sin(\beta_m w/2)}{\beta_m w/2}\right)E_g. \tag{15.7}$$

If the wave is a standing wave the amplitudes of the forward and backward waves are $E_g = V_g/2w$ where V_g is the magnitude of the RF voltage between the tips of the vanes. The total potential distribution is then the sum of the space-harmonic waves. For the special case of the π mode the $m = 0$ space harmonic of the forward wave and the $m = -1$ space harmonic of the backward wave both have the same propagation constant and the same amplitude. The amplitudes of all the other

space-harmonics with positive phase velocities are smaller so that the RF potential on the surface of the anode can be approximated by the sum of the equal forward and backward waves given by

$$V_\pm(x,h) = -j\frac{V_g}{\pi}\left(\frac{\sin(\pi w/2p)}{\pi w/2p}\right)\exp j(\omega t \mp \pi x/p),\tag{15.8}$$

since $\beta_0 p = \pi$. Thus the amplitude of the standing potential wave is

$$V_s = \frac{2}{\pi}\left(\frac{\sin(\pi w/2p)}{\pi w/2p}\right)V_g.\tag{15.9}$$

Figure 15.5 shows a comparison between the RF potential on the surface of the anode, modelled by the piecewise linear function which we have assumed, and the potential of the standing sinusoidal wave, when $w = p/2$. It can be seen that in this case, which is close to the situation in most magnetrons, the sinusoidal wave is a good approximation. Thus the peak-to-peak amplitude of the wave is approximately equal to the gap voltage, as one would expect.

The amplitude of the quasi-static potential of the standing wave in the π-mode must satisfy Laplace's equation in two dimensions, hence

$$V(x,y) = V_s\frac{\sinh(\pi y/p)}{\sinh(\pi h/p)}\cos(\pi x/p).\tag{15.10}$$

Then, taking the gradient of the potential

$$E_x(x,y) = V_s\frac{\pi}{p}\cdot\frac{\sinh(\pi y/p)}{\sinh(\pi h/p)}\sin(\pi x/p),\tag{15.11}$$

and

$$E_y(x,y) = -V_s\frac{\pi}{p}\cdot\frac{\cosh(\pi y/p)}{\sinh(\pi h/p)}\cos(\pi x/p).\tag{15.12}$$
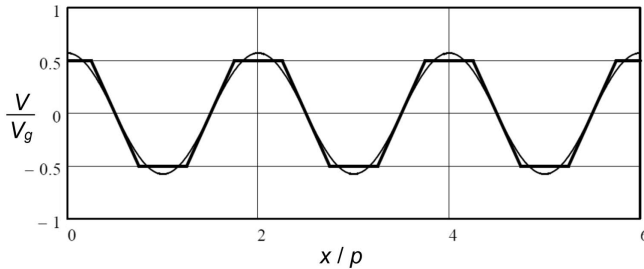
Figure 15.5: Comparison between the piecewise linear variation of electric potential on the surface of a magnetron anode, and the sum of the first forward and backward space-harmonic waves.
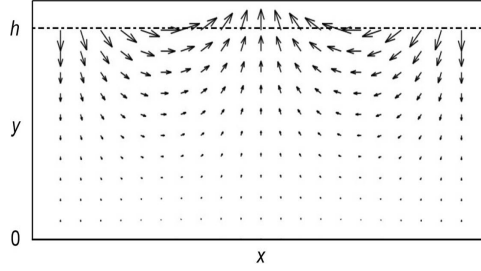
Figure 15.6: R.F. electric field of the π-mode of a planar magnetron.

Figure 15.6 shows the electric field of two adjacent cavities when the cavity pitch is equal to the distance between the cathode and the anode. It should be noted that the x component of the field is strongest close to the surface of the anode and that it decays rapidly with increasing distance towards the cathode. This justifies the use of a single term in the expansion of the field.

The analysis for circular geometry follows the same approach. Laplace's equation takes the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2}=0. \tag{15.13}$$

If there are n wavelengths around the anode then, separating the variables, the potential takes the form

$$V(r,\theta)=V_r(r)\cos(n\theta). \tag{15.14}$$

Substituting for V in (15.13) gives

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V_r}{\partial r}\right)-\frac{n^2}{r^2}V_r=0. \tag{15.15}$$

It can be shown by substitution that the general solution of this equation is

$$V_r = A r^n + B r^{-n} \tag{15.16}$$

where A and B are constants. The relationship between the constants is found by applying the boundary conditions $V = 0$ when $r = r_c$, and $V = V_s \cos(n\theta)$ when $r = r_a$. Substitution into (15.14) gives

$$V(r,\theta)=V_s\left(\frac{R^n-R^{-n}}{R_a^n-R_a^{-n}}\right)\cos(n\theta) \tag{15.17}$$

where $R = r/r_c$ and $R_a = r_a/r_c$. Then

$$E_r = -\frac{\partial V}{\partial r}=-\frac{nV_s}{Rr_c}\left(\frac{R^n+R^{-n}}{R_a^n-R_a^{-n}}\right)\cos(n\theta) \tag{15.18}$$
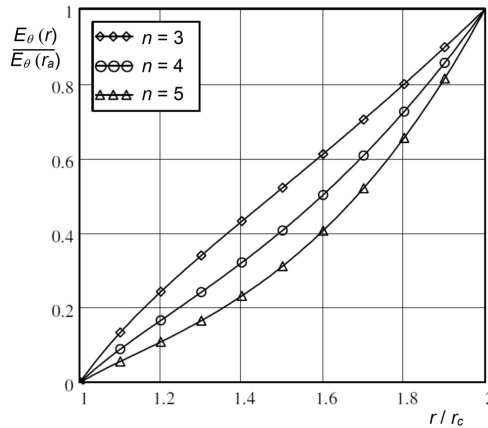
Figure 15.7: Variation of the tangential electric field between the cathode and the anode when $R_a = 2$ for three values of n.

and

$$E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta} = \frac{nV_s}{Rr_c}\left(\frac{R^n - R^{-n}}{R_a^n - R_a^{-n}}\right)\sin(n\theta) \tag{15.19}$$

Figure 15.7 shows the variation of the tangential electric field between the cathode and the anode for $n = 3, 4, 5$ when $R_a = 2$. It can be seen that the tangential field falls off rapidly with increasing distance from the surface of the anode and that the rate of change increases with n. Figure 15.8 shows the arrow plot of the π-mode field, for one quadrant of an anode with 12 resonators, obtained from (15.18) and (15.19).

When the phase shift per cavity is π the standing-wave field is the sum of equal forward and backward space-harmonics as in the planar case. The amplitude of the standing potential wave on the surface of the anode is related to the RF gap voltage by (15.9) where $p = 2\pi r_a/N_v$, N_v is the number of vanes and the gap between the vane tips $w = p - t$. The amplitudes of the travelling waves are then

$$V_1 = \frac{1}{\pi}\left(\frac{\sin\left(wN_v/4r_a\right)}{wN_v/4r_a}\right)V_g. \tag{15.20}$$

15.2.3 Threshold Condition for Oscillations

The threshold voltage can be calculated by considering the conditions under which electrons can just reach the anode. The RF field of a cylindrical anode is decomposed into a pair of equal waves propagating in opposite directions whose tangential electric fields vary as $\exp j(\omega t \pm n\theta)$ where n (an integer) is the number of wavelengths around the anode. The condition for constant phase is

$$\frac{d\theta}{dt} = \pm\frac{\omega}{n} = \pm\omega_s \tag{15.21}$$
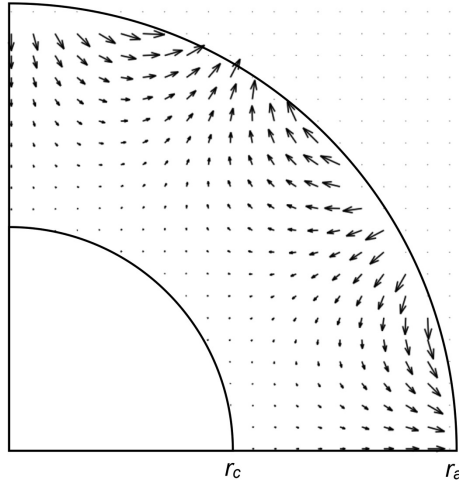
Figure 15.8: RF electric field of a 12-cavity magnetron anode operating in the π-mode when $r_a = 2r_c$.

so that a plane of constant phase rotates about the axis with angular velocity ω_s. When oscillations are just able to start, the RF field of the anode is negligible, the velocities of the electrons are tangential to the surface of the anode and synchronous with the wave. Then, using the principle of conservation of energy, the anode voltage is

$$V_0 = \frac{m_0}{2e} r_a^2 \omega_s^2. \tag{15.22}$$

This is the *Characteristic Voltage* which is the least anode voltage for which oscillations are possible. Since the diode is just cut off the corresponding magnetic field is obtained by substituting V_0 in (15.1) and setting $r = r_a$ to give

$$B_0 = \frac{2m_0 \omega_s}{e} \cdot \frac{1}{\left(1 - r_c^2/r_a^2\right)} \tag{15.23}$$

This is known as the *Characteristic Field*. The point (B_0, V_0) lies on the Hull cut-off curve (8.59) so that

$$V_0 = \frac{e}{8m_0} B_0^2 r_a^2 \left(1 - \frac{r_c^2}{r_a^2}\right)^2. \tag{15.24}$$

Dividing (15.1) by (15.24) the Hull cut-off curve can be expressed as

$$\frac{V_b}{V_0} = \frac{B_z^2}{B_0^2}, \tag{15.25}$$

where V_b is the potential at the surface of the space-charge hub. If the magnetic field is increased, the diode is more strongly cut off, and the radius of the hub decreases.

To show that oscillations can still start when $B_z > B_0$ we consider the dynamics of the electrons in the presence of the RF field of the wave on the anode. The equations of motion of an electron in cylindrical polar co-ordinates are [9]

$$m_0 \left[\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right] = F_r \qquad (15.26)$$

$$\frac{d}{dt} \left(m_0 r^2 \frac{d\theta}{dt} \right) = r F_\theta, \qquad (15.27)$$

where F_r and F_θ are the components of the force acting on the electron. Multiplying (15.26) by dr/dt and (15.27) by $d\theta/dt$ and adding gives

$$\frac{d}{dt} \left\{ \frac{m_0}{2} \left[\left(\frac{dr}{dt} \right)^2 + r^2 \left(\frac{d\theta}{dt} \right)^2 \right] \right\} = F_r \frac{dr}{dt} + F_\theta r \frac{d\theta}{dt}. \qquad (15.28)$$

This equation is a statement of the work-energy theorem. The left-hand side is the rate of change of the kinetic energy of the electron with time and the right-hand side is the rate of working of the forces acting on it. If the forces are derived from a potential field the equation can be integrated to obtain the law of conservation of energy. Unfortunately this is not possible in the present case because the forces vary with time in a fixed coordinate system. This difficulty can be overcome by changing to a coordinate system which is rotating with the synchronous velocity, so that the RF electric field of the wave can be derived from a static potential. To do this we define

$$\theta' = \theta - \omega_s t \qquad (15.29)$$

then (15.26) and (15.27) become

$$m_0 \left[\frac{d^2 r}{dt^2} - r \left(\frac{d\theta'}{dt} \right)^2 - 2r\omega_s \frac{d\theta'}{dt} - r\omega_s^2 \right] = F_r \qquad (15.30)$$

and

$$\frac{d}{dt} \left(m_0 r^2 \frac{d\theta'}{dt} + m_0 r^2 \omega_s \right) = r F_\theta. \qquad (15.31)$$

Then, by the same procedure as before, we obtain the equation in the rotating frame of reference, which is equivalent to (15.28).

$$\frac{d}{dt} \left\{ \frac{m_0}{2} \left[\left(\frac{dr}{dt} \right)^2 + r^2 \left(\frac{d\theta'}{dt} \right)^2 \right] \right\} = F_r \frac{dr}{dt} + F_\theta r \frac{d\theta'}{dt} + \frac{d}{dt} \left(\frac{1}{2} m_0 \omega_s^2 r^2 \right). \qquad (15.32)$$

Now

$$F_r = -e \left(E_r + r \frac{d\theta}{dt} B_z \right) = -e \left(E_r + r \frac{d\theta'}{dt} B_z + r\omega_s B_z \right) \qquad (15.33)$$

and

$$F_\theta = -e\left(E_\theta - \frac{dr}{dt}B_z\right), \tag{15.34}$$

where the electric field, which is the superposition of the DC and RF electric fields in the rotating system of coordinates, does not vary with time.

Substituting for F_r and F_θ in (15.32) we obtain, after simplification,

$$-\frac{d}{dt}\left\{\frac{m_0}{2e}\left[\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta'}{dt}\right)^2\right]\right\} = E_r\frac{dr}{dt} + r\omega_s B_z\frac{dr}{dt} + E_\theta r\frac{d\theta'}{dt} - \frac{d}{dt}\left(\frac{m_0}{2e}\omega_s^2 r^2\right). \tag{15.35}$$

Integration during the time taken for an electron to move from the cathode to the anode gives

$$-\left\{\frac{m_0}{2e}\left[\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta'}{dt}\right)^2\right]\right\}_1^2 = \int_{r_c}^{r_a} E_r\,dr + \omega_s B_z\int_{r_c}^{r_a} r\,dr + \int_{\theta_c'}^{\theta_a'} rE_\theta\,d\theta' - \left(\frac{m_0}{2e}\omega_s^2 r^2\right)_1^2. \tag{15.36}$$

At the cathode $dr/dt = 0$ and $d\theta'/dt = -\omega_s$. To find the threshold condition for oscillation we assume that an electron which just reaches the anode does so with near-zero radial and tangential velocities in the rotating frame. Substituting these boundary conditions into the left-hand side of (15.36) gives

$$V_T = \frac{1}{2}\omega_s B_z\left(r_a^2 - r_c^2\right) - \frac{m_0}{2e}\omega_s^2 r_a^2 + \int_{\theta_c'}^{\theta_a'} rE_\theta\,d\theta', \tag{15.37}$$

where

$$V_T = -\int_{r_c}^{r_a} E_r\,dr. \tag{15.38}$$

In the limit of a vanishingly small RF electric field the last term in (15.37) tends to zero and

$$V_T = \frac{1}{2}\omega_s B_z\left(r_a^2 - r_c^2\right) - \frac{m_0}{2e}\omega_s^2 r_a^2, \tag{15.39}$$

where V_T is the threshold (or Hartree) voltage. This is the minimum DC voltage at which electrons can reach the anode under the influence of the RF field, for a given magnetic field. The threshold voltage can be given a simple physical explanation [10]. The first term on the right hand side of (15.39) is the voltage induced in the rotating spoke, considered as a conductor. The second term is the kinetic energy in electron volts of a synchronous electron which reaches the anode with zero radial velocity.

Combining (15.39) with (15.22) and (15.23) we obtain

$$\frac{V_T}{V_0} = \frac{2B_z}{B_0} - 1. \tag{15.40}$$
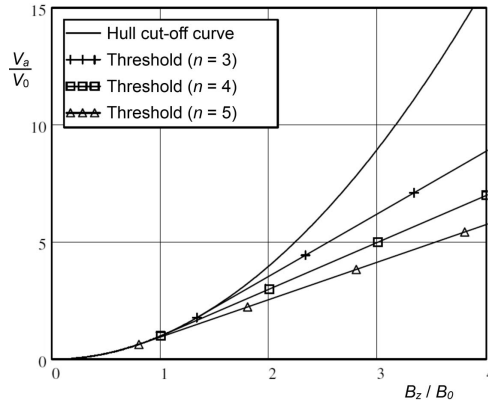
Figure 15.9: Threshold voltages for the $n = 3$, 4 and 5 modes of an eight-cavity magnetron. The axes are normalised to the characteristic voltage, and the characteristic field, for the $n = 4$ mode.

It is straightforward to show that (15.40) is a straight line which is tangential to the cut-off curve (15.25) at the point (B_0, V_0). The characteristic field is proportional to ω_s so that each possible mode has its own characteristic point and threshold line. Figure 15.9 shows the normalised cut-off curve for an eight-vane magnetron, together with the threshold lines for $n = 4$ (the π mode) and the adjacent modes. If the magnetic field is greater than the characteristic field for the π mode no current can flow until the anode voltage exceeds the threshold voltage for that mode. Once the threshold voltage has been exceeded oscillations build up from the noise already present in the electron hub. The operation of the magnetron as an oscillator requires the anode voltage to remain below the cut-off voltage for the given magnetic field. Thus the possible region of operation lies between the threshold voltage line and the cut-off voltage curve. The way in which the choice of operating point affects the properties of the magnetron are discussed in Section 15.4.

The assumptions involved in the derivation of (15.39) can be questioned. This is illustrated by Figure 15.10 which shows electron trajectories, in the synchronous frame of reference, in a magnetron which is just conducting. The trajectories were computed for a cylindrical magnetron but they are displayed in Cartesian coordinates for convenience. A finite RF electric field was assumed. The tangential velocities of the electrons on the surface of the hub exceed the synchronous velocity so that they are initially in the direction of positive θ'. The trajectory which finally reaches the surface of the anode has already travelled through several RF periods. It eventually reaches a position where the tangential retarding RF electric field is strong enough to cause the electron to drift towards the anode. The continuing drift of the electron to the right must be arrested before it leaves the retarding phase of the field $(0° - 180°)$. The drift motion in the synchronous frame of reference is reversed and the electron then moves outward until it reaches the anode. The tangential velocity on impact is equal to the synchronous velocity, but the calculations show that the impact energy is approximately $0.2\,eV_a$. Thus the assumption that the RF electric field, and the energies of electrons reaching the anode, are
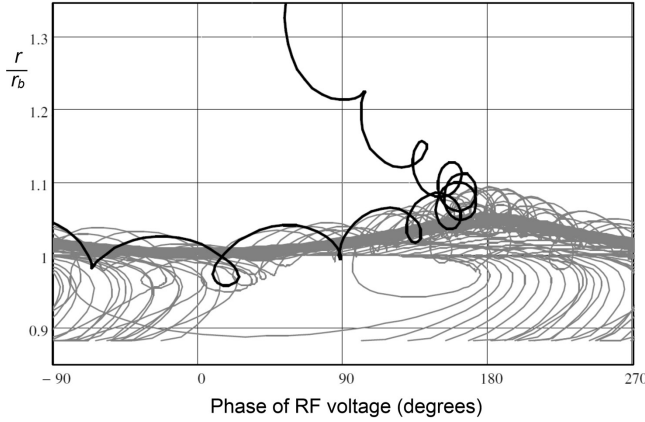
Figure 15.10: Electron trajectories in a magnetron which is just conducting.

both vanishingly small at the threshold voltage is seen to be false. Nevertheless, the threshold voltages computed from (15.39) are found to be close to the results of experimental measurements. It must be concluded that the effect of the finite RF electric field is balanced by the impact energy of the electron on the anode.

The requirement that the RF electric field should be finite means that the initiation of oscillations cannot be explained by the theory described above. An alternative possibility is to consider the threshold condition in terms of the synchronism between the slow space-charge wave on the hub and the wave on the anode [11]. We have seen that, in a circular magnetron, the angular velocity of the electrons on the surface of the hub (ω_b) exceeds the synchronous velocity. The angular velocity is synchronous at the characteristic voltage and magnetic field, and the difference increases as the magnetic field, and the anode voltage, are increased. In the example shown in Figure 15.10 where $V_a \approx V_T$ and $B_z = 3.34 B_0$ it was calculated that $\omega_b = 1.46 \, \omega_s$.

Space-charge waves on electron beams in crossed-fields take the form of perturbations of the surface of the electron beam with constant charge density. Thus they resemble the space-charge waves on linear beams with Brillouin focusing. It has been shown that the propagation constants in a planar magnetron take the form

$$u_0 \beta_\pm = \omega - \frac{\omega_c}{2} \mp \left[\left(\frac{\omega_c}{2} \right)^2 + \frac{\omega_p^2}{1 + \coth \beta d} \right]^{\frac{1}{2}}, \qquad (15.41)$$

where d is the distance between the anode and the surface of the hub [12]. We will assume that this equation is approximately correct for a cylindrical magnetron. The plasma frequency on the surface of the hub is given by (8.67) so that (15.41) becomes

$$u_0 \beta_\pm = \omega - \frac{\omega_c}{2} \mp \frac{\omega_c}{2} \left[1 + \frac{2 \left(1 + r_c^4 / r_b^4 \right)}{1 + \coth \left(\beta (r_a - r_b) \right)} \right]^{\frac{1}{2}}. \qquad (15.42)$$
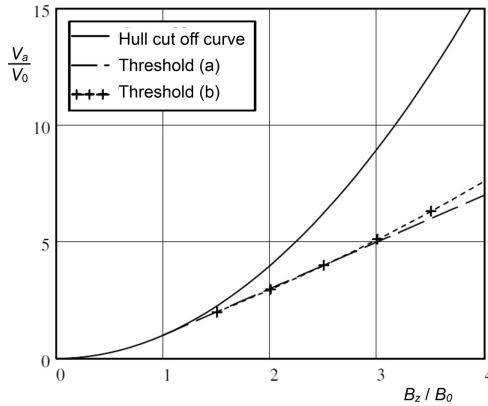
Figure 15.11: Comparison between threshold voltages computed for the 4J50 magnetron using: (a) (15.39) and (b) (15.44).

Making use of the synchronous condition

$$\beta_- = \frac{n}{r_b} \tag{15.43}$$

we find that the threshold voltage is the root of

$$\omega - n\omega_b - \frac{\omega_c}{2} + \frac{\omega_c}{2}\left[1 + \frac{2\left(1 + r_c^4/r_b^4\right)}{1 + \coth\left(n\left(r_a/r_b - 1\right)\right)}\right]^{\frac{1}{2}} = 0 \tag{15.44}$$

for a given value of B_z, since ω_b and r_b are then functions of V_a only. Figure 15.11 shows an example of the threshold voltage computed from this equation, compared with that given by (15.39). There is excellent agreement between the two curves for small values of the magnetic field. The difference at larger magnetic fields can be attributed to the difference between the results for planar and cylindrical magnetrons as the distance between the anode and the surface of the hub increases. We saw in Section 8.6 that, in the absence of external RF fields, the space charge hub is perturbed by oscillations that are strongest when there is an integral number of wavelengths around the hub. If the anode voltage is set to the value given by (15.44) then the slow space-charge wave with mode number n is synchronous with the wave on the anode. The currents induced in the anode by the space-charge wave transfer energy from the wave to the anode. There it is dissipated in resistive losses and in the external load. Because the slow space-charge wave carries negative energy its amplitude increases through the transfer of energy until the point is reached where the fields are strong enough for spoke formation. Thus space-charge wave theory explains why the threshold voltage is given correctly by (15.39) even when the RF fields are vanishingly small.

15.2.4   Electronic Efficiency

The electronic efficiency of a magnetron can be estimated by noting that the RF output power is given by

$$P_{rf} = P_{dc} - P_a - P_c,$$ (15.45)

where the terms on the right-hand side are, respectively, the DC input power and the thermal power dissipated on the anode and on the cathode. Thus the electronic efficiency is given by

$$\eta_e = 1 - \frac{P_a + P_c}{P_{dc}}.$$ (15.46)

The power dissipated on the cathode is typically less than 5% of the input power so that it may usually be neglected in the first approximation (though it can be as great as 50% at low output powers) [13]. Since the spokes are stationary in the rotating frame the mean angular velocity of the electrons at the surface of the anode is ω_s so that that the tangential energy of the electrons is eV_0 from (15.22). If it is assumed that the radial energy of an electron striking the anode is approximately equal to the tangential energy we can write (15.46) as

$$\eta_e \approx 1 - \frac{2V_0}{V_a}.$$ (15.47)

Now a magnetron is normally operated at a voltage close to the threshold voltage so setting $V_a \approx V_T$

$$\eta_e \approx 1 - \frac{2V_0}{V_T}.$$ (15.48)

Substituting for V_0/V_T from (15.40) we obtain

$$\eta_e \approx \frac{2B_z - 3B_0}{2B_z - B_0}.$$ (15.49)

The graph of the efficiency given by (15.49) in Figure 15.12 shows that the efficiency increases rapidly with B_z/B_0 and that high efficiency is possible if a high magnetic field is used. The validity of the approximations made in deriving this result can be examined using the theoretical models in Section 15.6.

15.3   Magnetron Anodes

Figure 15.13(a) shows a plan view of a magnetron anode which incorporates an even number of regularly spaced cavity resonators. The detailed shape of the resonators may vary but the vane-type anode illustrated is the most common [14]. This circuit is a modified two-wire line in which the forward and backward waves are
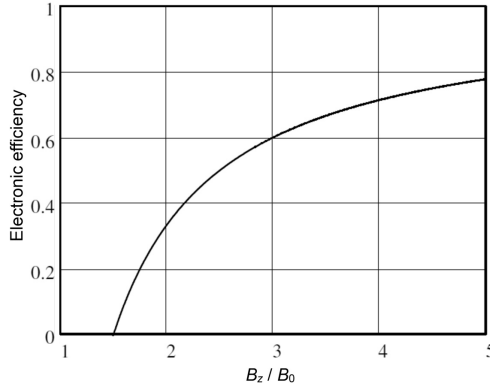
Figure 15.12: Approximate dependence of the efficiency of a magnetron on the normalised magnetic field.
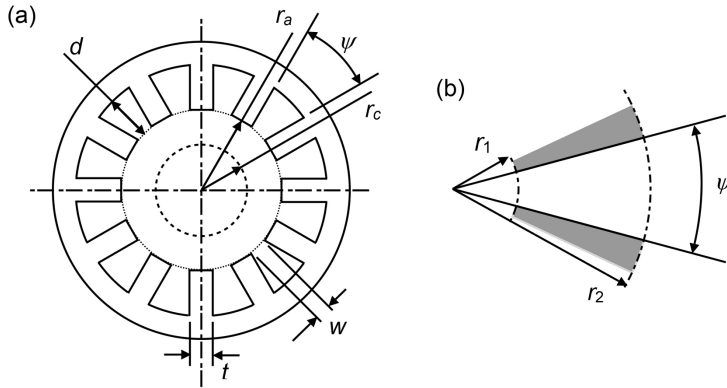
Figure 15.13: (a) Arrangement of a vane-type magnetron anode and (b) detail of one cavity (reproduced, with permission, from [15]).

strongly coupled together at frequencies close to the resonant frequencies of the cavities.

The properties of this structure can be determined from the equivalent circuit shown in Figure 15.14. The admittance at the tips of the vanes of the cavity, represented by L_1 and C_1, is

$$Y_a(\omega) = j\left(\omega C_1 - \frac{1}{\omega L_1}\right) = j\omega C_1\left(1 - \frac{\omega_1^2}{\omega^2}\right), \tag{15.50}$$

where $\omega_1 = 1/\sqrt{L_1 C_1}$ is the resonant frequency of the cavity. The admittance presented at the same terminals by the space between the cathode and the anode is

$$Y_k(\omega, \phi) = \frac{j\omega C_0}{2(1 - \cos\phi)}, \tag{15.51}$$
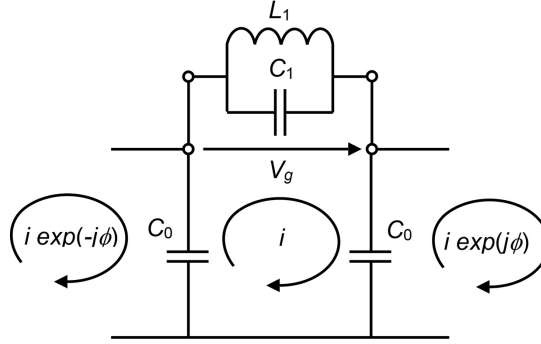
Figure 15.14: Equivalent circuit of one cavity of a magnetron anode.

where C_0 is the capacitance between the tip of a vane and the cathode and ϕ is the phase difference between adjacent cavities. The resonant frequency of the cavity, modified by the admittance of the inner space is given by

$$Y_a(\omega) + Y_k(\omega, \phi) = 0. \tag{15.52}$$

Substituting into this equation from (15.50) and (15.51) we obtain, after a little rearrangement,

$$\cos\phi = 1 - \frac{C_0}{2C_1} \cdot \frac{1}{\left(\omega_1^2/\omega^2 - 1\right)}. \tag{15.53}$$

This is a low-pass filter characteristic such that $\omega = 0$ when $\phi = 0$ and

$$\omega = \frac{\omega_1}{\sqrt{1 + C_0/4C_1}} = \omega_\pi, \tag{15.54}$$

when $\phi = \pi$. Normally the capacitance between an anode vane and the cathode (C_0) is small compared with the capacitance between the vanes (C_1) so that ω_π is slightly less than ω_1. Because the slow-wave structure forms a closed circuit the only possible values of ϕ are those for which there are an integral number of wavelengths around the circuit. Thus

$$\phi_n = \frac{2n\pi}{N_v}; \quad n = 0, 1 \ldots N_v/2, \tag{15.55}$$

where N_v is the number of vanes in the anode and n is the *mode number*. Equation (15.53) can be rearranged as

$$\frac{\omega}{\omega_1} = \left(1 + \frac{C_0}{2C_1} \cdot \frac{1}{1 - \cos\phi}\right)^{-\frac{1}{2}}, \tag{15.56}$$

from which the dispersion curve can be calculated. Figure 15.15 shows examples of dispersion curves calculated for a 12-vane anode for two values of the ratio C_0/C_1.
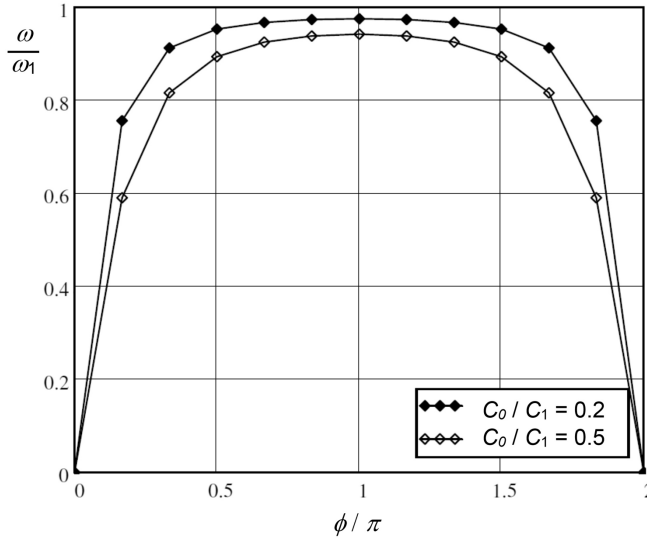
Figure 15.15: Dispersion curves of 12-vane magnetron anodes.

Because the structure is periodic, the field can be expanded as a series of space harmonics. Figure 15.15 shows the possible forward wave modes in the first two Brillouin zones (see Figure 4.6). Thus, the section of each curve in the range $\pi \le \phi \le 2\pi$ represents the $m = -1$ space harmonic of the backward wave. The resonances of a transmission line in the form of a circular loop can be represented as linear combinations of modes having azimuthal variations $\cos n\theta$ and $\sin n\theta$. When the transmission line is periodic, as in the case of a magnetron anode, pairs of modes exist in which the relationship between the phase of the mode and the discontinuities in the line are different. In these modes, known as *degenerate modes*, the nodes and antinodes of the field pattern are exchanged. Figure 15.16 shows, as an example, the two possible orientations of the radial field maxima of the degenerate $n = 4$ modes of a 12-resonator anode. The differing relationships between the field patterns, and the positions of the vanes, mean that the two modes have slightly different frequencies. The asymmetry introduced by the output coupler means that one mode of the pair is strongly coupled to the output waveguide and has a low Q factor. The other is weakly coupled and has a high Q factor so that unwanted oscillations may be a problem. When $\phi = \pi$ the maxima of the tangential electric field coincide with the centres of the gaps between the vanes. This mode cannot be degenerate because the tangential electric field in the alternative orientation is short-circuited by the vanes. For this reason magnetrons are always designed to operate using the π-mode field of the anode. Figure 15.15 shows that the frequencies of other possible modes are very close to the frequency of the π-mode. It is therefore important in the design and operation of magnetrons to ensure that the wrong mode is not excited. The most troublesome modes are those for which $n = N_v/2 - 1$ and $n = N_v/2 + 1$. These are known as the $(\pi - 1)$ and $(\pi + 1)$ modes irrespective of the number of vanes. Because the output coupler has
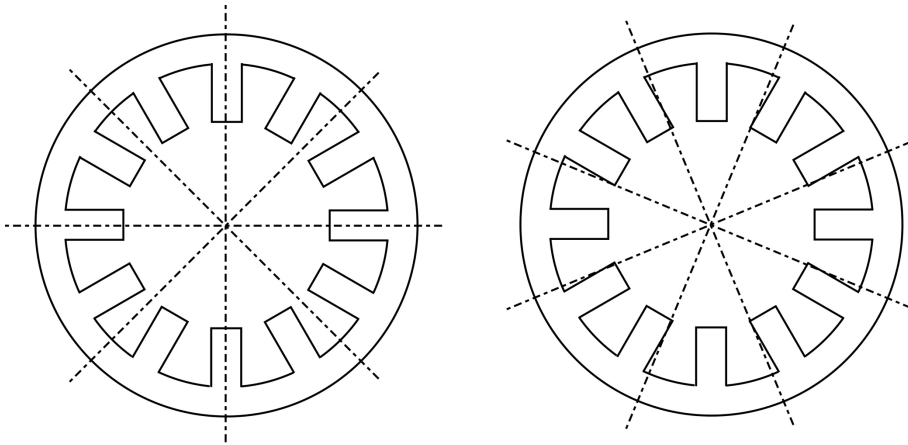
Figure 15.16: Alternative orientations of the maxima of the radial electric field for the $n = 4$ mode of a 12-resonator anode.

been designed with respect to the π-mode the excitation of other modes usually shows itself as a reduction in the output power in addition to a shift in frequency. The change in the interaction may also cause damage to the cathode and arcing within the tube because the impedance is mismatched to that of the power supply. Various techniques for ensuring operation in the correct mode are discussed in Section 15.4.5.

The resonant frequency of the cavities in the anode in Figure 15.13(b) can be estimated by assuming that the depth of each cavity is a quarter of a wavelength so that $\lambda = 4(r_2 - r_1)$. It is found that this formula overestimates the frequency by about 20%. A more exact solution is obtained from the input admittance of a wedge-shaped cavity with open ends given by

$$Y = j\sqrt{\frac{\varepsilon_0}{\mu_0}} \cdot \frac{L_a}{\psi r_a} \cdot \left\{ \frac{J_0(kr_1)Y_1(kr_2) - J_1(kr_2)Y_0(kr_1)}{J_1(kr_1)Y_1(kr_2) - J_1(kr_2)Y_1(kr_1)} \right\}, \tag{15.57}$$

where $k = \omega/c$ is the free-space propagation constant at frequency ω, L_a is the axial length of the anode and $\psi = 2\pi/N_v$ [15]. The length of the anode is normally less than half of the free-space wavelength to avoid the excitation of modes having axial variation of the RF field strength. The radii r_1 and r_2 are measured from the point of intersection of the projections of the vanes (see Figure 15.13(b)) so that

$$r_1 = r_a - \frac{t}{\psi}, \tag{15.58}$$

where $\psi = 2\pi/N_v$. At resonance, $k = k_0 = \omega_0/c$, $Y = 0$ and (15.57) reduces to

$$J_0(k_0 r_1)Y_1(k_0 r_2) - J_1(k_0 r_2)Y_0(k_0 r_1) = 0. \tag{15.59}$$

This equation can be solved for k_0 graphically or by numerical methods. The capacitance of the resonator (C_r) can be calculated from (15.57) since the input admittance of the equivalent circuit is

$$Y = j\omega C_r \left[1 - \left(\frac{\omega_0}{\omega} \right)^2 \right].$$

(15.60)

Differentiating (15.60) with respect to ω yields

$$\frac{dY}{d\omega} = jC_r \left[1 + \left(\frac{\omega_0}{\omega} \right)^2 \right]$$

(15.61)

so that

$$C_r = -\frac{j}{2c} \left(\frac{dY}{dk} \right)_{k=k_0},$$

(15.62)

where c is the velocity of light and the derivative of the Y with respect to k at resonance is calculated from (15.57). The value of the inductance L_1 is then obtained immediately from C_r and ω_0.

This calculation of the resonant frequency assumes that the RF magnetic field of the cavity is zero when $r = r_1$. It therefore ignores the effects of the fringing fields in the space between the anode and the cathode and at the ends of the anode. The effects of the fringing field can be modelled by adding a capacitance C_f in parallel with the equivalent circuit of the cavity. Because the dimensions of the region of fringing field are small compared with the free-space wavelength it is possible to determine the capacitance C_f from an electrostatic solution. The fringing capacitance at the tips of the vanes is given approximately by [16]

$$C_f = \frac{\varepsilon_0 L_a}{\pi} \left\{ 1 + \ln \left(\frac{4}{\pi\mu} \right) \right\},$$

(15.63)

where

$$\mu = \frac{N_v w}{2\pi r_a}.$$

(15.64)

The capacitance of the cavity including the correction for the fringing field is then

$$C_1 = C_r + C_f.$$

(15.65)

The capacitance per vane between the anode and the cathode (C_0) can be estimated by finding upper and lower bounds [17]. An upper bound for C_0 is obtained by assuming that the equipotential surfaces are concentric cylinders to give

$$C_U = \frac{\varepsilon_0 L_a}{\ln(r_a/r_c)} \cdot \frac{2\pi}{N_v}.$$

(15.66)

To find a lower bound we assume that the flux lines between the anode and the cathode are confined to wedge-shaped volumes which terminate on the tips of the vanes so that

$$C_L = \frac{\varepsilon_0 L_a}{\ln(r_a/r_c)} \cdot \frac{t}{r_a}. \tag{15.67}$$

Finally C_0 is taken to be the geometric mean of these two bounds so that

$$C_0 = \sqrt{C_L C_U}. \tag{15.68}$$

Note that this calculation assumes that the cathode is a cylinder. In continuous wave magnetrons the cathode commonly takes the form of a helical wire and the capacitance calculated from (15.68) would need to be reduced to allow for this. We have seen that the dispersion diagram is not strongly dependent on C_0 for typical values so this estimate is sufficiently accurate for our purposes. Further information about the calculation of equivalent circuit parameters for resonators of other shapes and about the effects of the fringing fields at the ends of the anode can be found in [14].

The simple anode structure which has been discussed in this section suffers from the disadvantage that the mode frequencies are too close together to ensure that the π-mode alone is excited, without contamination from the adjacent modes. It is evident that the problem gets more difficult as the number of resonators in the anode is increased. A variety of solutions to this problem have been found in which the properties of the anode are modified by changes in its design. These are discussed in the sections which follow. In addition it has been found that the introduction of $N_v/2$-fold azimuthal periodicity into the magnetic field, the properties of the cathode, or the shape of the anode, can help to ensure that the oscillations start in the π-mode [18–21].

The output from a conventional cavity magnetron is generally made through a slot in the outer wall of one of the resonators or by connecting the inner conductor of a coaxial line to one, or more, of the vanes [3, 14]. In either case the coupling is made into a rectangular waveguide through a vacuum window. Although the magnetron is a fixed-frequency oscillator, it is possible to vary the frequency over a limited range by including metallic tuning elements that can be moved relative to the anode block. Metal bellows are used to take the motion through the vacuum envelope. The tuning elements can perturb the capacitance, the inductance, or both, of the resonant cavities. Pins inserted near the tips of the vanes alter the cavity capacitance, while those inserted near the outer edge of the cavity alter the inductance [3, 22–24]. The frequency can be changed by up to 10% by a mechanical tuner and various methods have been used to vary the frequency from pulse to pulse [3]. Electronic tuning which can vary the frequency by a few percent can be achieved using ferrites, diode switches or multipactor discharges [3, 25]. A greater tuning range can be achieved by coupling the anode to a high Q external cavity as described in Section 15.3.3.

15.3.1 Strapped Anodes

In order to reduce the risk that a mode other than the π-mode is excited, the anodes in practical magnetrons are designed in such a way that the separation of these two modes is increased. Historically, the first method used was the introduction of wires, known as *straps*, which connect the vanes together. In modern magnetrons the straps are generally connected to alternate vanes as shown in Figure 15.17. Two or three straps are added to one, or both ends of the anode. Theoretically the straps do not carry current in the π-mode, since they are connected to vanes which are at the same phase. They do carry current in the other modes whose resonant frequencies are therefore perturbed. The actual effect of strapping is a little different, as will be seen, but it achieves the purpose of increasing the separation between the π and the $\pi-1$ modes.

The geometrical complexity of the strapping makes it difficult to determine their effects accurately except by cold test measurements, or from computational electromagnetics software. However, useful insight can be gained by considering the straps as TEM transmission lines which are connected to alternate vanes, as shown in Figure 15.18, where the admittance of each cavity is $Y = Y_a + Y_k$ from (15.50) and (15.51) [26].

Let us assume that the voltages on the tips of the vanes are given by

$$V = V_0 \exp(jq\phi), \tag{15.69}$$

where ϕ is the phase shift per cavity and $q = 0, \pm 1, \pm 2 \ldots$ is the index number of the vanes such that $q = 0$ when $x = 0$. Now let the voltage on strap 1 relative to the cathode in the range $0 \le x \le 2p$ be

$$V_1 = V_+ \exp(-jkx) + V_- \exp(jkx) \tag{15.70}$$

where V_+ and V_- are the amplitudes of the forward and backward waves, and $k = \omega/c$.

View of vane *B*

View of vane *A*
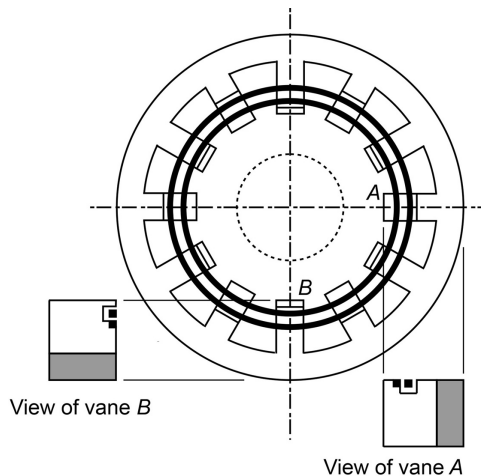
Figure 15.17: Strapped magnetron anode.
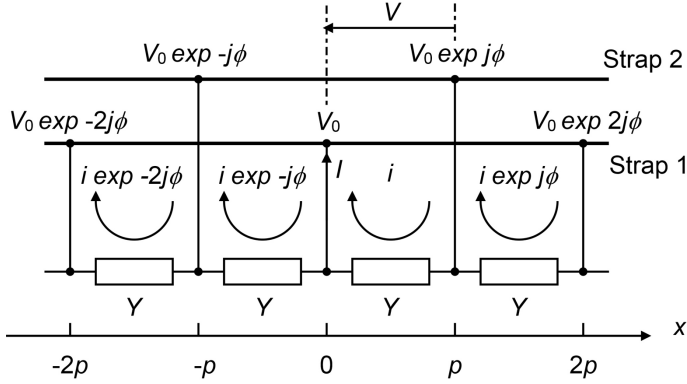
Figure 15.18: Equivalent circuit for a strapped anode.

When $x = 0$

$$V_1 = V_0 = V_+ + V_- \tag{15.71}$$

and, when $x = 2p$

$$V_1 = V_0 \exp(2j\phi) = V_+ \exp(-2jkp) + V_- \exp(2jkp). \tag{15.72}$$

Substituting for V_- in (15.72) from (15.71) gives

$$V_0 \left[\exp(2j\phi) - \exp(2jkp) \right] = -2jV_+ \sin(2kp). \tag{15.73}$$

Similarly, substituting for V_+ gives

$$V_0 \left[\exp(2j\phi) - \exp(-2jkp) \right] = 2jV_- \sin(2kp). \tag{15.74}$$

The current flowing in the positive x direction when x is just greater than zero is given by

$$I_+ = Y_c \left(V_+ - V_- \right), \tag{15.75}$$

where Y_c is the characteristic admittance of the transmission line formed by strap 1. Substituting for V_+ and V_- in (15.75) from (15.73) and (15.74) gives

$$I_+ = \frac{jY_cV_0}{\sin(2kp)} \left(\exp(2j\phi) - \cos(2kp) \right). \tag{15.76}$$

Similarly, the current flowing in the negative x direction when x is just less than zero is given by

$$I_- = \frac{jY_cV_0}{\sin(2kp)} \left(\exp(-2j\phi) - \cos(2kp) \right). \tag{15.77}$$

The current flowing to the strap from the vane is the sum of these two currents

$$I = jY_cV_0\left(\frac{2\cos(2\phi)-2\cos(2kp)}{\sin(2kp)}\right). \quad (15.78)$$

This current can also be written in terms of the loop current circulating in the cavity as

$$I = i\left(1-\exp(-j\phi)\right). \quad (15.79)$$

The voltage driving the loop current into the transmission line is

$$V = V_0\left(1-\exp(j\phi)\right). \quad (15.80)$$

Hence the admittance presented to the cavity by the straps is

$$Y_s\left(\omega,\phi\right) = \frac{i}{V} = jY_c\left(\frac{\cos(2\phi)-\cos(2kp)}{\sin(2kp)(1-\cos(\phi))}\right). \quad (15.81)$$

This result is not the same as that given by Yu and Hess [26] but it is believed to be correct.

The dispersion curve of a strapped anode is found from the solutions of

$$Y_a\left(\omega\right)+Y_k\left(\omega,\phi\right)+Y_s\left(\omega,\phi\right) = 0. \quad (15.82)$$

The characteristic admittance of the straps $\left(Y_c\right)$ is unknown, but it can be used as an adjustable parameter to fit the dispersion curve to experimental data. Figure 15.19 shows, as an example, the dispersion curves, with and without straps, for the anode of the 4J50 magnetron whose dimensions are given in [27] (see Worksheet 15.1). This
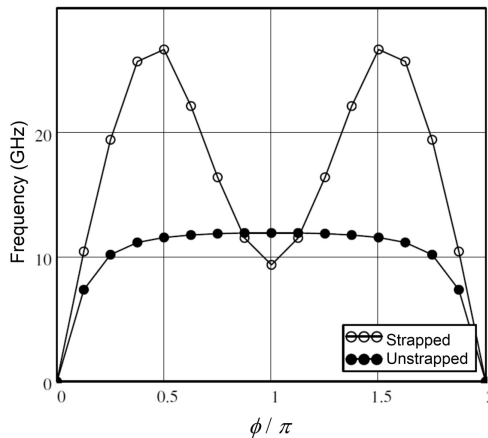
Figure 15.19: Dispersion curves for the anode of the Litton 4J50 magnetron with, and without, straps.

anode has 16 vanes and two strapping rings at each end. The characteristic admittance of the straps was adjusted to give the correct π-mode frequency [13]. The frequency of the $\pi - 1$ mode then agreed with the figure given in [13] to three significant figures. It can be seen that the effect of the straps is to reduce the frequency of the π-mode and to increase the separation between the π and $\pi - 1$ modes. The admittance of the straps is capacitive at the π-mode so that the frequency of this mode is lower than in an unstrapped anode. It should be noted that the strapped anode is a bi-periodic structure so that the π-mode corresponds to a phase shift of 2π in one period (two cavities). The dispersion diagram has a forward-wave characteristic when the phase shift per cavity is small. This must be the case because the structure is a heavily modified form of two-wire line whose dispersion curve therefore passes through the origin. At larger values of phase shift the characteristic becomes backward wave. The technique is limited to frequencies below 16 GHz by manufacturing difficulties. It is not suitable at the highest lower power levels because of problems with voltage breakdown between the straps and the vanes.

We have seen that the modes other than the π-mode exist as pairs whose orientations are determined by the position of the output coupler. If a break is made in the straps at an angle of 45° to the output waveguide, then the orientations of the two modes are determined by it. Both modes are then coupled to the output waveguide to some extent, the Q factors of the modes are reduced, and oscillations are less likely.

15.3.2   Rising Sun Anodes

An alternative anode structure which also makes use of bi-periodicity to increase the separation between the π and $\pi - 1$ modes is the *Rising Sun* structure, shown in Figure 15.20. This structure is useful at high frequencies and high powers where construction of straps is difficult, or they suffer from voltage breakdown. The ratio of the depths of the two resonators d_1/d_2 (where $d = (r_2 - r_1)$ is known as the circuit ratio.

The equivalent circuit of the rising sun structure is shown in Figure 15.21. Analysis of this circuit shows that the dispersion curves are given by

$$\begin{vmatrix} Z_1 + 2Z_0 & -Z_0\left(1 + \exp(2j\phi)\right) \\ -Z_0\left(1 + \exp(-2j\phi)\right) & Z_2 + 2Z_0 \end{vmatrix} = 0, \qquad (15.83)$$

where ϕ is the phase shift per cavity,

$$Z_1 = \left(j\omega C_1 + 1/j\omega L_1\right)^{-1}, \qquad (15.84)$$

$$Z_2 = \left(j\omega C_2 + 1/j\omega L_2\right)^{-1}, \qquad (15.85)$$

and

$$Z_0 = 1/j\omega C_0. \qquad (15.86)$$
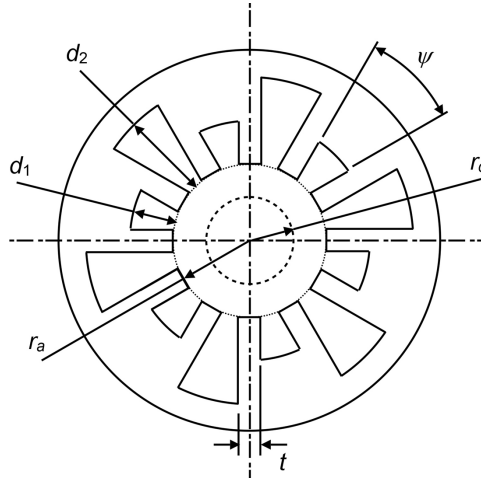
Figure 15.20: Arrangement of a rising sun magnetron anode (reproduced, with permission, from [15]).
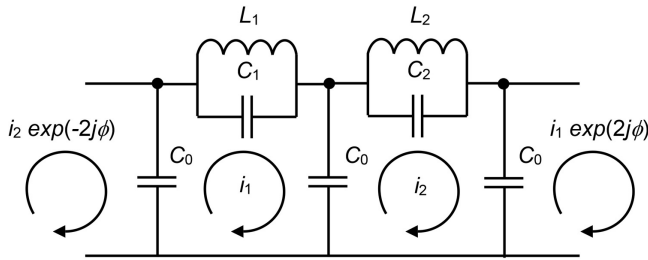
Figure 15.21: Equivalent circuit of a rising sun anode.

Equation (15.83) can be expanded to give

$$\cos^2 \phi = \left(1 + \frac{Z_1}{2Z_0}\right)\left(1 + \frac{Z_2}{2Z_0}\right),$$

$$= \cos \phi_1 \cos \phi_2$$

(15.87)

where ϕ_1 and ϕ_2 are the phase shifts per cavity of uniform structures in which the one set of cavities is replaced by a short circuit, as shown in Figure 15.22. This is not quite accurate because short-circuiting a set of cavities also makes a small change in C_0. Each structure has half as many elements as the full anode and the π-modes of these structures therefore correspond to the $\pi/2$ modes of the full anode. Figure 15.23 shows an example of the coupled, and uncoupled, dispersion curves (see Worksheet 15.2). The coupling between the modes produces a pair of pass-bands with a stop band between them. The upper pass-band is cut-off at $\phi = 0$ and $\phi = \pi$ with a frequency somewhat above the resonant frequencies of the large
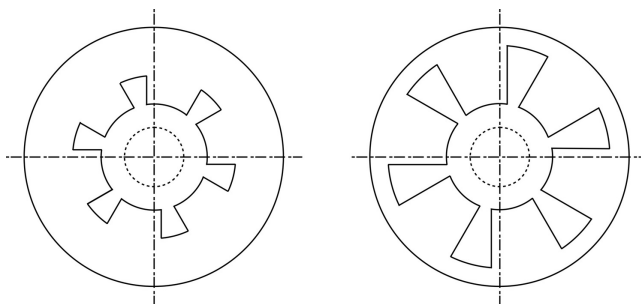
Figure 15.22: Uniform structures which are combined in a rising sun anode.
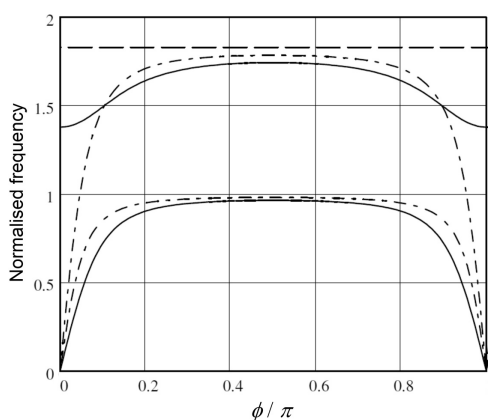
Figure 15.23: Dispersion curves of a typical rising sun anode, normalised to the frequency of the larger cavities (solid lines), together with the dispersion curves of the uniform uncoupled structures (chain dotted lines), and the frequency of the smaller cavities (dashed line) $(N_v = 18, d_2/d_1 = 1.78)$.

cavities. The anode is designed to operate in the π-mode of the upper pass-band and the steep dispersion curve close to that point ensures good mode separation.

Figure 15.24 shows how the mode separation increases with circuit ratio d_2/d_1 for an anode where $N_v = 18$. The use of a large circuit ratio is superficially attractive, but it is found to lead to problems with the excitation of the $\phi = 0$ mode. At the π-mode resonance the fields are different in the large and the small cavities, and this gives rise to a component of the tangential field which is in phase for all cavities. This component increases as the difference between the depths of the cavities increases. When interaction with this mode occurs it causes heavy back-bombardment of the cathode and serious loss of efficiency. An electron which is synchronous with the π-mode sees the $\phi = 0$ mode as an RF field at the signal frequency. If the cyclotron frequency coincides with the signal frequency then a resonance occurs. It is therefore vital to ensure that the cyclotron frequency is well away from the signal frequency in a magnetron with a rising sun anode. It is usual also to avoid values of N_v which are divisible by four because it is found that they can lead to problems associated
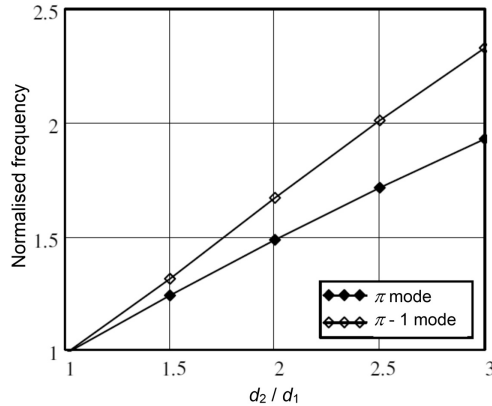
Figure 15.24: Dependence of the π and $\pi-1$ mode frequencies of a rising sun anode with 18 vanes on the circuit ratio d_2/d_1.

with the mode for which $\phi = \pi/2$. Further information about the properties of rising sun anodes can be found in [16].

15.3.3   Coaxial Anodes

Conventional magnetrons have mode and frequency stabilities which are inadequate for some purposes. Their tuning ranges are also rather restricted. These problems can be overcome by coupling the anode to another resonator having a very high Q factor [28–30]. The additional cavity is placed around the anode resulting in the coaxial magnetron, shown in Figure 15.25. The coaxial cavity is coupled to alternate cavities by slots, so that the π-mode of the anode is coupled to the TE_{011} (circular electric) mode of the coaxial cavity. The energy stored in the cavity is much greater than that stored around the vanes. This means that the operating frequency is determined by the resonant frequency of the cavity and can be tuned over a greater range than that of a conventional magnetron. Because the TE_{011} mode does not have any current flow across the end walls of the cavity, it is possible for the tuning plunger to be made non-contacting. The space behind the tuner can hold lossy material to damp out unwanted modes. A comparison between a conventional and a coaxial magnetron is shown in Table 15.1. It is seen that the latter is superior in its tuning range and stability (frequency pushing and pulling are discussed in Sections 15.4.2 and 15.4.3). The disadvantages of the coaxial magnetron are its large size, for a given frequency, and the slow rise of the anode voltage pulse needed to allow time for the energy stored in the coaxial cavity to build up as oscillation commences.

15.3.4   Long Anodes

The power output of a conventional magnetron is limited by heat dissipation in the cathode. To increase the output power it is necessary to increase the diameters

Table 15.1: Comparison between a coaxial and a conventional X-band magnetron

(copyright 1961, Academic Press, reproduced, with permission, from [28])

	Coaxial magnetron	**Conventional magnetron**
Peak power	250 kW	250 kW
Circuit efficiency	85%	70%
Electronic efficiency	65%	60%
Tuning range	20%	10%
Frequency pulling	3 MHz	14 MHz
Frequency pushing	0.08 MHz/A	0.4 MHz/A
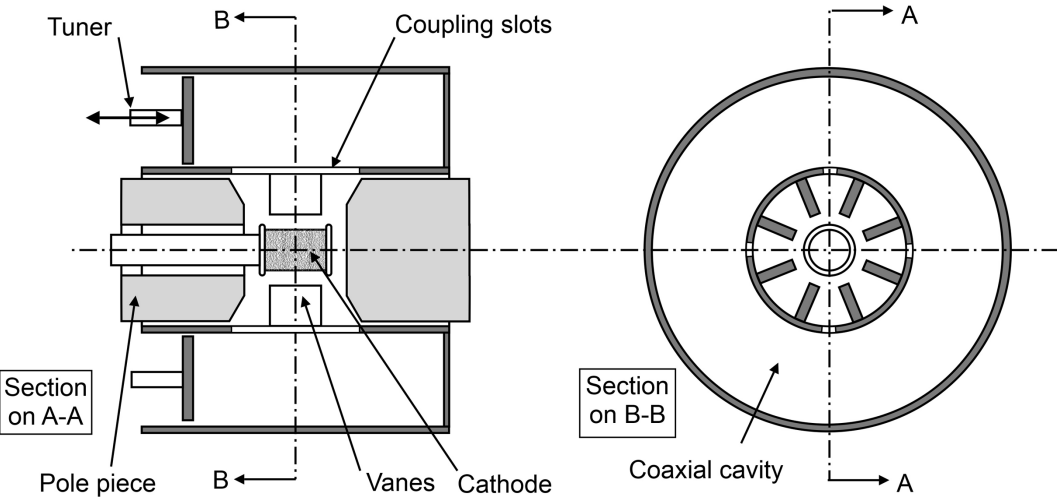
Figure 15.25: Arrangement of a coaxial magnetron.

of the anode and the cathode and, therefore, the number of vanes. This brings with it problems of reduced mode separation. These problems can be overcome by using a magnetron with a much longer anode (known as a *long anode magnetron*) [31–33]. By making the anode length comparable with the free-space wavelength, the cathode area can be increased without increasing its diameter and the number of vanes can be small. These magnetrons are normally made without straps and use the method of output connection to select the correct mode. A long anode can be regarded as a length of waveguide of unusual cross-section which propagates a TE mode in the axial direction. The anode then has resonances at the cut-off frequency of each mode, and also at higher frequencies where there are longitudinal standing waves determined by the boundary conditions at the ends of the anode. The notation for the modes is H_n, H_n', H_n'', etc. where n is azimuthal mode number and the number of primes represents the number of null planes along the length of the anode.
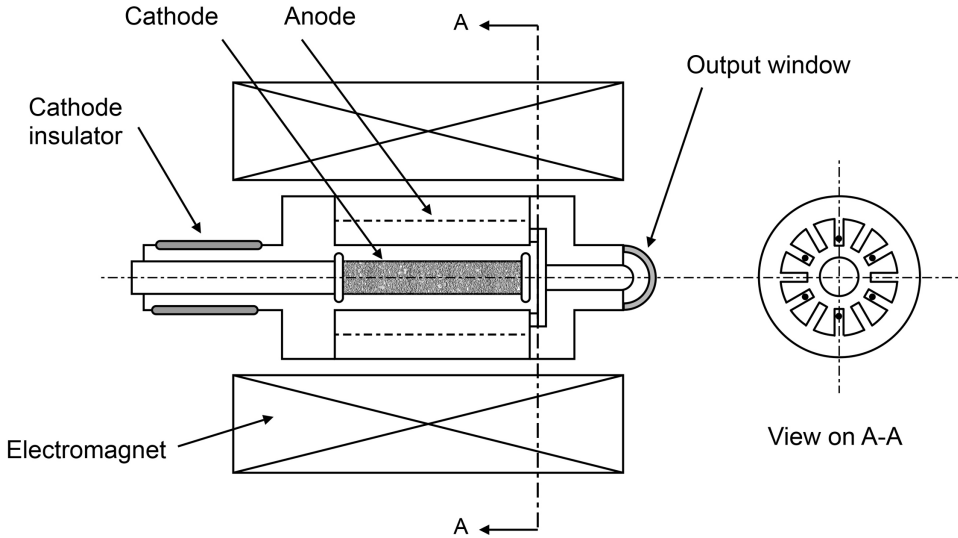
Figure 15.26: Simplified diagram of a long anode magnetron.

The output connection is made by connecting the inner conductor of a coaxial line symmetrically to alternate vanes as shown in Figure 15.26. The strength of the coupling is greatest when the connections are made close to the tips of the vanes, and it decreases as the pitch circle diameter of the connections is increased. The effect of the symmetry of the coupler is to couple the modes together in pairs so that the H_n mode is coupled to the $H_{(N_v/2)-n}$ mode. For example, in a 12 cavity anode the $n = 0$ mode is coupled to the $n = 6$ mode, the $n = 1$ mode to the $n = 5$ mode, and so on. These coupled modes are denoted by $H_{(0,6)}$, $H_{(1,5)}$ etc. since the H_n and $H_{(N_v/2)-n}$ modes do not exist independently. Where one of the component modes is strongly cut off, it represents local storage of energy close to the output coupler, and the axial variation of the electric field strength is determined by the other mode. It is then possible to simplify the notation still further. Thus, if the H_0 mode is strongly cut off, the $H_{(0,6)}$ mode can be referred to as the 6 mode or π mode. This mode is strongly coupled to the TEM mode of the output coaxial line. The line propagates all frequencies so that all the π mode resonances $(\pi, \pi', \pi''$ etc.$)$ are strongly coupled to the output and have loaded Q values determined by it. The non-π modes couple to the higher-order modes of the coaxial line which are cut off, and these modes therefore have high Q factors determined by the losses in the anode. The symmetry of the output coupler allows pure π-modes to be established when the frequency separation from the adjacent modes is as low as 1% so that straps are not necessary. The effect of adding straps was discussed in [34]. The output is coupled directly into a rectangular waveguide via a capacitive probe enclosed by a hemispherical dielectric window. Further information about the design of long-anode magnetrons can be found in [31–33].

15.4   Magnetron Properties

We saw in Section 15.2 that the behaviour of a magnetron can be described in terms of the cut-off and threshold voltages. At the characteristic voltage and field (V_0, B_0) these two voltages coincide. However, it can be seen from Figure 15.12, that a tube must be operated appreciably above the characteristic field in order to achieve satisfactory conversion efficiency. Usually $B/B_0 \geq 3$ and the magnetron is operated at an anode voltage just above the threshold voltage, and well below the cut-off voltage (see Figure 15.9). The threshold voltage for each mode can be considered as the anode voltage at which oscillation in that mode begins. The oscillations in the various possible modes start from the noise in the circulating electron current and grow competitively until one becomes dominant. When the wrong mode becomes dominant the magnetron is said to be *moding*. This is revealed by a change in the frequency, reduced output power and excessive heat dissipation on the cathode. The anode current remains small until the oscillation is fully established. Therefore the impedance presented to the power supply by the magnetron varies during the start-up of the oscillation. This section reviews the practical aspects of the performance of magnetrons.

15.4.1   The Performance Chart

Manufacturers of magnetrons commonly supply information about their characteristics in the form of a *performance chart* like the one shown in Figure 15.27. The diagram shows the anode voltage and current, the magnetic field, and the power output and efficiency. The lines for constant magnetic field are commonly called *Gauss lines*. *Gauss line discontinuities* are associated with discontinuous changes in other properties, as described in Section 15.4.4. The projection of a Gauss line intersects the vertical axis close to the threshold voltage for that magnetic field. For the example shown, the intersections lie a little below the theoretical threshold voltages. Better agreement is obtained by linear projections which are tangents to the Gauss lines at the right-hand side of the diagram. Regions of operation to be avoided because of poor mode stability or excessive heat dissipation may be marked. Note that the axes of this figure are exchanged, compared with the usual presentation of characteristic curves. For most magnetrons with strapped or rising sun anodes the magnetic field is provided by a permanent magnet whose strength is set by the manufacturer. Where the magnetic field is provided by an electromagnet (e.g. long anode magnetrons) it may be varied to some extent by the user. The efficiency increases as the magnetic field is increased. For a fixed magnetic field there is a region of optimum efficiency, as shown in Figure 15.27. It is found that the graph of the RF power against the DC current is approximately a straight line. A small current flows when the power is zero. This offset is caused by current which does not take part in the interaction, mainly emission from the cathode end-caps.
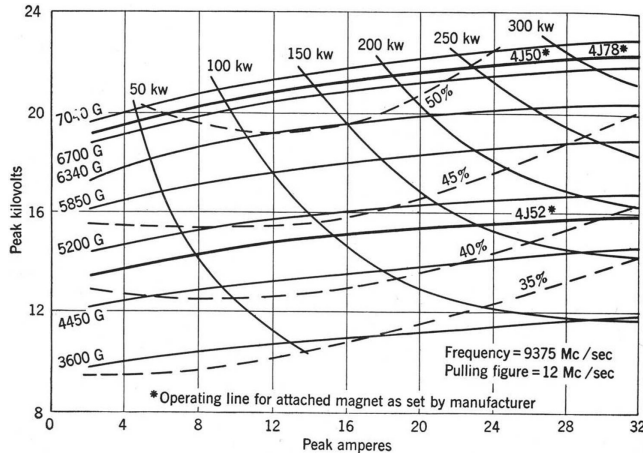
Figure 15.27: Performance chart of the 4J50 magnetron (reproduced, with permission, from [35]).
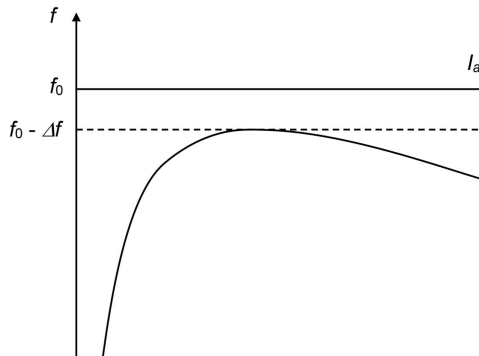
Figure 15.28: Typical frequency pushing curve of a magnetron.

15.4.2 Frequency Pushing

The frequency of oscillation of a magnetron varies with the anode current, as shown in Figure 15.28 where f_0 is the resonant frequency of the anode in the absence of any electrons. The frequency at the peak of the curve is depressed relative to the unloaded resonant frequency of the anode by electron loading. The change in frequency (Δf) is given approximately by

$$\frac{\Delta f}{f_0} \sim \frac{0.6}{Q_L},$$ (15.88)

where the loaded Q of the anode (Q_L) is typically 250 for conventional magnetrons, and 1000 for coaxial magnetrons. The shape of the curve is determined by the

changing phase of the electron spokes with respect to the RF field of the anode (see Section 15.6.7). Tubes are normally designed to be operated on the flattest part of the curve. The residual variation of frequency with current (~100 kHz/A) is called the *frequency pushing* figure of the magnetron [36, 37].

15.4.3   Frequency Pulling

The effect of the output match on the performance of a magnetron is displayed by plotting the frequency and output power against the match on a Rieke diagram, as shown in Figure 15.29. It should be noted that the phase of the mismatch is referred to the output flange of the tube. Thus the line showing zero frequency shift does not normally coincide with the real axis of the impedance chart. The point of maximum power output does not usually coincide with the centre of the chart. This to ensure that the performance of the tube is stable throughout the permitted range of mismatch, that thermal limits are not exceeded, and that internal arcing does not occur. It is usual for magnetrons, other than low power tubes, to be protected from excessive variations in the output match by the inclusion of a circulator in the output waveguide.

To investigate the effect of a mismatched load we note that the admittance of the unloaded resonator is

$$Y_U = G_U \left[1 + jQ_U \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right) \right], \tag{15.89}$$

where G_U is the shunt conductance of the resonator and Q_U its unloaded Q. The external load may be represented by a transmission line whose characteristic
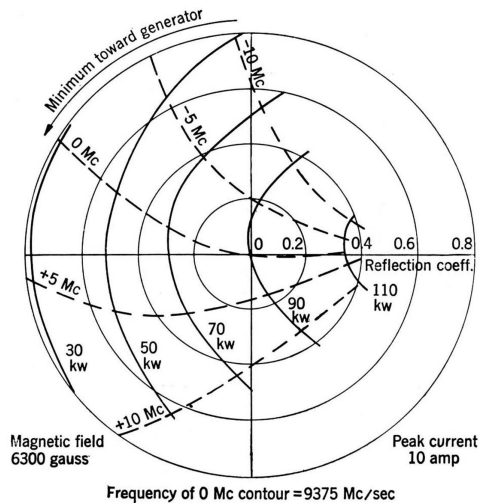
Figure 15.29: Rieke diagram of the 4J50 magnetron (reproduced, with permission, from [35]).

admittance is G_E so that the external loading conductance is G_E when the transmission line is matched. When the line is terminated by a mismatch, the external admittance loading the resonator may be written

$$Y_E = \frac{S + j \tan\theta}{1 + jS \tan\theta} G_E, \tag{15.90}$$

where S is the VSWR on the line and θ its electrical length. It is usual to specify the effect of a mismatched load for $S = 1.5$, and it is then found that the imaginary part of (15.90) is maximum when $\theta = \pm 0.588$ and $Y_E = (1.083 \pm 0.417\,j)G_E$. When the external load is connected to the resonator the imaginary part of $Y_U + Y_E$ is zero at resonance so that

$$Q_U G_U \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right) \pm 0.417 G_E = 0. \tag{15.91}$$

Now $Q_U G_U = Q_E G_E$ so that (15.91) may be written

$$\frac{2\Delta\omega}{\omega_0} = \pm \frac{0.417}{Q_E}, \tag{15.92}$$

where $\omega = \omega_0 + \Delta\omega$. The *frequency pulling figure* (f_p) is defined as the change in the frequency when a load for which $S = 1.5$ is varied through all possible phases. Thus

$$f_p = 0.417 \frac{f_0}{Q_E}. \tag{15.93}$$

This derivation assumes that there is only one possible frequency at which the sum of the imaginary parts of the admittances is zero. When the transmission line connecting the load to the magnetron is long, as may be the case when an antenna is connected directly to the tube, multiple solutions are possible. The frequency of oscillation may then be unstable or show hysteresis effects. These *long line effects* can be avoided by careful design, or by placing an isolator between the magnetron and the transmission line [38].

15.4.4   Spectrum

The frequency spectrum of a magnetron normally includes some output at harmonics of the fundamental frequency. Since the electronic charge does not vary sinusoidally with angle, but is concentrated in the spokes, the RF current flowing to the anode has a high harmonic content. The extent to which this is reflected in the output from the tube depends upon the coincidence between any higher-order modes of the anode with the harmonics of the signal frequency, and also on how well such modes are coupled to the output waveguide. Since these frequencies are tuned by changes in the output match it may be expected that those changes will also affect the harmonic output. It is possible to suppress harmonic emissions by

including a filter in the output waveguide either inside, or outside, the tube [39]. However, since the filter reflects the harmonic signals back into the anode, there may be unexpected changes in the output at the fundamental frequency.

The spectrum of a pulsed magnetron is normally determined by the Fourier Transform of the RF pulse. It is found, however, that the spectrum can be much broader if the magnetron is operated at a low anode current where there is strong frequency pushing. This effect can be explained by the variation in the frequency of oscillation during the pulse. The spectrum of continuous wave cooker magnetrons is found to vary with the operating current. A study of magnetrons, operating at 2.45 GHz, for use in domestic microwave ovens showed three distinct regimes [40]. At low currents (~20% of normal operating current) broadband random noise was observed with a bandwidth of up to 600 MHz. The noise showed distinct peaks corresponding to modulation at a frequency in the range 150 to 300 MHz, which increased with anode current. This low frequency noise can be conducted, and radiated, by the high-voltage cathode leads and its harmonics can cause interference with UHF TV reception. At an anode current around 50% of the normal operating current the noise could take the form of discrete sidebands at frequencies about 300 MHz above and below the fundamental, with power about 30 dB below it. This condition, involving spurious oscillation, was associated with Gauss line discontinuities where a discontinuity in the characteristic curve corresponds to an abrupt change in the spectrum. Finally, at high currents, the spectrum showed no noise above the noise floor of the spectrum analyser. All cooker magnetrons show low-current, high-noise, and high-current, low-noise, regions with a discontinuous transition between them. Where spurious oscillation is associated with the transition it is strongly affected by the external VHF impedance between the cathode and the anode. In particular, it is possible to lock the frequency to a VHF resonant circuit in the cathode leads. The spectrum of a magnetron tends to deteriorate with age showing increasing levels of broadband noise. This change is presumably correlated with decreasing thermionic emission from the cathode and consequent changes in the relative proportions of thermionic and secondary emission.

Some pulsed magnetrons exhibit a phenomenon known as *twinning* in which the spectrum shows two distinct peaks, and two different levels of RF output power and anode current, as shown in Figure 15.30 [41, 42]. This phenomenon is clearly distinguishable from moding (see Section 15.4) because the frequency difference between the two modes is much smaller (typically only a few MHz). It is still not completely understood, but appears to be related to changes in the space-charge cloud that alter the electronic loading of the anode. The changes probably arise from changes in the proportions of thermionic and secondary electrons emitted by the cathode. A detailed study of twinning in pulsed magnetrons showed that the two states do not coexist, but that the tube may switch from one to the other during the pulse. On occasion multiple transitions were observed [42]. It was found that the conditions under which a tube twinned could be altered by changes in the output match, the magnetic field, the heater power, and the cathode coating. However, none of these would completely eliminate the problem from tubes which showed
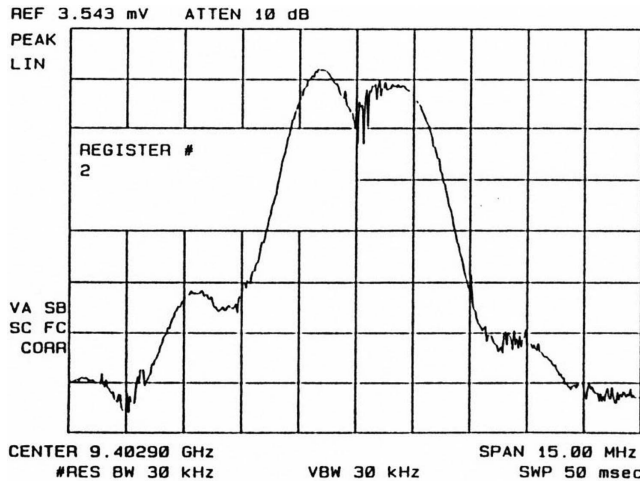
Figure 15.30: Spectrum of a magnetron showing twinning.
(copyright 2000, IEEE, reproduced, with permission, from [42]).

this tendency. Twinning is one of those phenomena, known to, and frustrating to, all tube engineers that may appear or disappear between batches of tubes as a result of minute changes in the production process. The instability appears to be correlated with Gauss line discontinuities. A survey of a wide range of different types of magnetrons showed that twinning is most likely to occur when the ratio of the electron cyclotron frequency to the operating frequency lies in the range 0.75 to 1.25.

15.4.5   Mode Selection, Priming, and Locking

The typical characteristic curves of a magnetron oscillator, in Figure 15.27, show that for a given magnetic field the current remains low until the threshold voltage is reached. It then rises rapidly over a small change of variation of the anode voltage, and settles at an operating point determined by the load line of the power supply. When the magnetron is pulsed the shape of the pulse must be designed to ensure that the correct mode is excited [43, 44]. Figure 15.31 shows typical pulse waveforms for the anode voltage and anode current. For a discussion of pulse modulators see Section 20.3.

The important features of the voltage pulse are:

- The rate of rise of the voltage can be fast as far as V_T but must thereafter be slow enough to ensure that the oscillation is fully established in the π mode before the threshold voltage of the $\pi-1$ mode is reached. Typically this takes of the order of 50–100 RF cycles. During the initial rise of the pulse the current consists mainly of the charging current of the capacitance between the anode and the cathode. If the oscillation is not fully established by the time the voltage reaches V_T then the anode current is low, the magnetron impedance is high, the voltage overshoots and $\pi-1$ mode oscillation is probable.
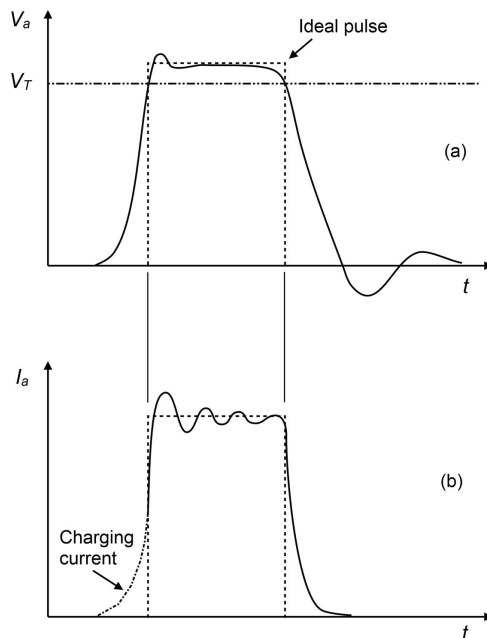
Figure 15.31: Typical voltage and current pulses of a magnetron.

- The pulse top must be flat because, as shown in Figures 15.27 and 15.28, a small voltage swing results in a large current swing, strong frequency pushing, and frequency modulation during the pulse.
- The voltage pulse should fall fast to V_T and must fall fast enough thereafter to ensure that none of the modes beyond the π mode is excited.

If the output port of the tube is connected to a circulator, then a signal can be injected at the π mode frequency. A signal with an amplitude 40 dB below the final output level has the effect of giving that mode a head start over the competing modes. The result is greater certainty that the correct mode will be excited. The technique, known as *Priming*, also determines the phase of the oscillation and provides pulse-to-pulse phase coherence.

The phase and frequency of an oscillator, such as a magnetron, can be locked by an injected signal [45]. The injection power required increases with the square of the frequency difference between the injected signal and the natural frequency of the oscillator. The bandwidth of a free-running magnetron is normally so large that the injected signal must be around 10 dB below the output power to achieve locking [46]. However, if the bandwidth can be reduced by suitable selection of the operating conditions, or by an external control loop, the power required is much less [47–49]. In the same way the frequencies of two or more tubes can be locked to one another [46, 50].

15.5   Particle in Cell Magnetron Models

To model a magnetron fully it is necessary to find self-consistent solutions for the electron motions and the RF fields (including the space-charge field). A three-dimensional model is required to include the effects of longitudinal motion of the electrons, longitudinal variations of the fields, and the effects of fields in the end spaces. It is also necessary to model the combination of thermionic and secondary electron emission from the cathode (which may well be non-uniform), and to model the build-up of oscillations from the initial noise level to saturation. The additional complexity, compared with a model of a klystron or a TWT, requires the use of a fine computational mesh close to the surface of the cathode, and continuation of the calculations over some fifty or more RF cycles. Thus, the computational time is very long, even on a very powerful computer [51–53].

As with klystrons and TWTs, it is possible to achieve some economy in the computational effort: by using a two-dimensional model which ignores axial variations; by pre-calculating the static magnetic field; and by representing the properties of the anode by an equivalent circuit. This avoids the need to re-compute the electromagnetic fields outside the interaction region at each time step. However, because the electrons are collected on the anode it is desirable to include at least part of the space between the vanes in the interaction region. Calculations of this kind can provide valuable insight into the complex processes in a magnetron [54–57]. Three states have been investigated [56, 57]:

- In the first the cathode current is dominated by primary emission, and the DC electric field is zero at the surface of the cathode. The electron trajectories are then similar to first-order Slater trajectories (see Section 8.3.2) and the thickness of the space-charge hub is close to that given by the numerical solution of (15.2). In order to maintain the space-charge hub the current emitted from the cathode must be between three and four times the anode current. The fraction of the DC power dissipated on the surface of the cathode is around 2%. The spokes are narrow as shown in Figure 15.32(a) and the base of the spoke is an unstable region with fluctuating behaviour. High-efficiency magnetrons for microwave ovens and industrial heating, which have thoriated tungsten cathodes with low secondary emission coefficients, are believed to operate in this condition.
- In the second case the cathode current is dominated by secondary emission. Because the distribution of returning electrons is not uniform on the surface of the cathode, the shapes of the spokes vary with time, and regions of low-charge density move through the hub. These fluctuations are considerably greater than those in the primary emission state. The spokes are much thicker in this case as shown in Figure 15.32(b) and the cathode dissipation is of the order of 10% of the DC power. Most high power pulsed magnetrons, which have oxide or dispenser cathodes with high secondary emission coefficients, are believed to operate in this condition.
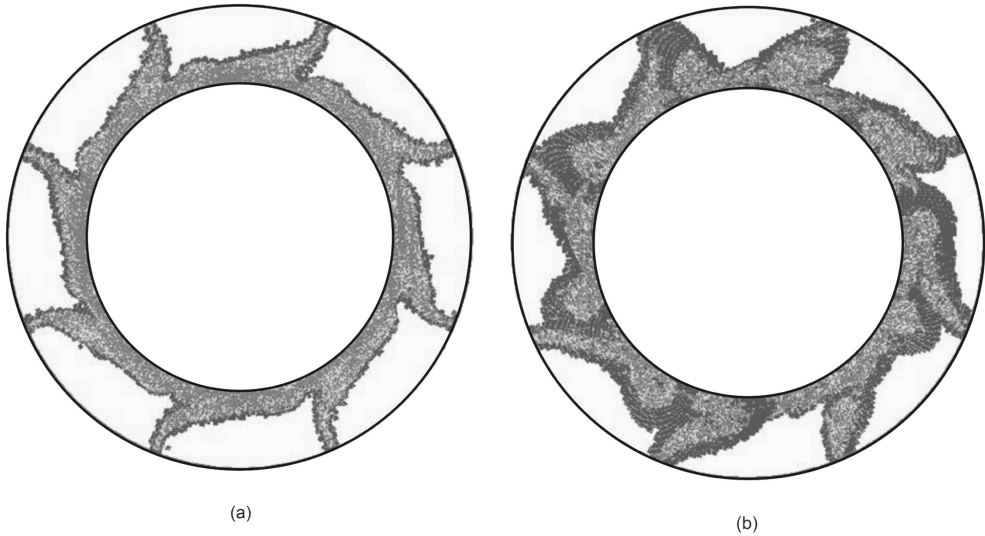
Figure 15.32: Space-charge distribution in a 16-cavity magnetron where the emission is dominated by: (a) thermionic emission, and (b) secondary emission. The spokes rotate in the clockwise direction
(copyright 2004, IEEE, reproduced, with permission, from [57]).

• In the third case there is no secondary emission and the cathode current falls below around three times that required to supply the anode current. It is then limited by the primary emission from the cathode. This state is characterised by a complete absence of fluctuations. It appears to correspond to the low-noise state observed in microwave oven magnetrons when the heater power is reduced.

  Table 15.2 shows a summary of the results obtained for a 4J52 magnetron [56] with experimental data for comparison [35]. The computed results differ considerably from the experimental data so that conclusions drawn from the model must be treated with caution.

15.6   Simple Magnetron Models

Because the modelling of magnetrons using PIC codes is very time-consuming it is useful to have simpler models which can give approximate results quickly. The images in Figure 15.32, and the discussion in the previous section, illustrate the complexity of the problem of modelling a magnetron, and the danger of making assumptions about, for example, the shapes of the spokes. It has been observed that magnetrons may exhibit chaotic behaviour in the sense that very small changes in the operating conditions produce appreciable changes in performance. Twinning is an example of this (see Section 15.4.4). Despite these

Table 15.2: Properties of a 4J52 magnetron $(B_z = 4900 \text{ G})$, computed for different emission regimes (copyright IEEE, 1998, reproduced, with permission, from [56]), compared with experimental data [35]

	Primary emission dominated	Secondary emission dominated	Primary emission limited	Experiment
Anode voltage (kV)	16.5	15.5	17.5	16
V_a/V_T	1.15	1.08	1.22	
Anode current (A)	15.8	20.3	27.8	27
Output power (kW)	140.6	130.6	145.7	180
Efficiency (%)	53.9	41.4	29.9	42
Anode dissipation (%)	20.9	29.8	51.8	
Cathode dissipation (%)	2.1	11.1	5.4	
Circuit loss (%)	23.1	17.7	12.8	
Frequency shift (MHz)	–22.1	–17.9	–25.0	
Thermionic current available (A)	220	2.2	55.0	
Total cathode current (A)	54	77.6	53.5	
Current fraction returned (α)	0.71	0.74	0.48	
Wave voltage / DC voltage	0.28	0.27	0.28	

difficulties it is possible to gain useful insight into the factors affecting the performance of magnetrons using simple methods [13, 37, 58–61]. These models are based upon the motion of the electrons in simplified fields with the following assumptions:

- The electrons are drawn into the spokes from a space-charge hub whose properties are determined by static theory.
- The electric field of the anode is represented by a travelling wave and the RF magnetic field is neglected.
- The effects of all space-harmonics, other than the one synchronous with the spokes, are assumed to average to zero.
- The space-charge field of the spokes is neglected.
- The current flowing to the anode through the spokes is determined from the properties of the hub and the proportion of its perimeter from which electrons are drawn.

The properties of the space-charge hub, including the effects of secondary electron emission from the cathode are discussed in the next section. Following that, three different ways of modelling the spokes are reviewed and compared. It is then shown how the performance chart, the Rieke diagram, and the frequency pushing characteristics of a magnetron can be computed.

15.6.1   The Space-Charge Hub

The conditions affecting the electron emission can be illustrated by a simple discussion [62]. Let us assume that the total current emitted by the cathode is I_c and that a fraction α of this current returns to the cathode so that the anode current is

$$I_a = (1-\alpha)I_c. \tag{15.94}$$

If the average secondary electron emission coefficient of the cathode is δ then the available secondary emission current is

$$I_s = \alpha\delta I_c. \tag{15.95}$$

McDowell found that α was typically in the range 0.65–0.75, except when the current was limited by primary emission [56]. Thus if the product $\alpha\delta$ is greater than unity the total current will be a self-balancing combination of primary and secondary emission. The secondary emission dominated state observed by McDowell will only arise if the heating of the cathode by the heater, and by back bombardment, is insufficient to generate enough primary emission. Thus a magnetron model which assumes that the current is uniform and space-charge limited should correspond to the operating conditions in many magnetrons. However, this kind of model is unable to represent the fluctuating secondary emission from the cathode observed by McDowell when the secondary emission is dominant.

We saw in Section 8.5 that, theoretically, only zero or first-order trajectories can form a stable hub in a cylindrical magnetron. Either of these may be used as a model for the hub. For a given anode voltage and magnetic field the hub radius (r_b) is obtained by numerical solution of (15.2). The electrostatic potential on the surface of the hub is found by setting $r = r_b$ in (15.1), and the angular velocity of electrons on the surface of the hub is given by (15.4). These calculations assume a zero-order hub but, as was shown in Section 8.5, the properties of a first-order hub are the same to within a few percent.

In a zero-order (Brillouin) hub the electrons move in circles concentric with the cathode [13, 37, 60, 61]. The electric field at the surface of the cathode is zero, and the space-charge density at the surface of the hub, obtained from (8.67), is

$$\rho_b = \frac{\varepsilon_0 m_0}{2e}\omega_c^2\left(1+\frac{r_c^4}{r_b^4}\right), \tag{15.96}$$

where $\omega_c = eB_z/m_0$ is the cyclotron frequency. It is assumed that the electrons forming the spokes are drawn from the surface of the space-charge hub with constant charge density. Since the spokes are synchronous with the field of the anode we assume, from (15.34), that

$$\frac{dr}{dt} \approx \frac{E_\theta}{B_z} \tag{15.97}$$

at the base of the spoke. The tangential electric field is obtained by setting $R = R_b$ in (15.19).

Then anode current is given by

$$I_a = n\rho_b r_b L_a \int_{\theta_1'}^{\theta_2'} \frac{E_\theta}{B_z} \, d\theta' \tag{15.98}$$

where the integration is carried out over the angles for which current can flow to a spoke. Three different methods for determining this range of angles are described in the sections which follow. This model suffers from the difficulty that, because there is no radial current within a zero-order hub, the electrons must be assumed to diffuse through the hub from the surface of the cathode. It also assumes that the charge density in the hub is not perturbed by the RF field of the anode. The calculation in (15.98) avoids the assumptions about the shape of the spoke made in [13].

An alternative model of the space-charge hub, suggested by the results obtained by McDowell [56], is to assume that the electrons in the hub follow first-order Slater orbits which start from the cathode and return to it. The radius of the hub and the angular velocities of the electrons on its surface are close to those of a zero-order hub. However, the radial electrostatic field is reduced so that the electrons experience a net inward force. When their tangential velocities are reduced, by interaction with the RF electric field, they can move outwards into the spokes. If it is assumed that the total cathode emission is the maximum possible for a given ratio (r_b/r_c), then the cathode current in a cut-off diode is approximately 0.35 of the space charge limited current in the absence of a magnetic field (see Figure 8.13). When the diode is just conducting, the fraction is 0.7 because no electrons return to the cathode and the cathode current is doubled for given space charge. Thus, if the space-charge hub is essentially the same as that in a diode at the point of cut off, the charge density and the electrostatic potential are cylindrically symmetrical, so we can write

$$(1-\alpha)I_c + 2\alpha I_c \approx 0.7 I_{CL}, \tag{15.99}$$

where I_{CL} is determined using the Child–Langmuir law (5.70). Note that this assumes that the current density has azimuthal variation because the hub is cut off between the spokes but not at their bases. The multiplying constant on the right-hand side can be determined more precisely from r_b/r_c, if necessary (see Figure 8.13 with r_a set to the hub radius). The fraction of the cathode current which flows to the anode $(1-\alpha)$ is obtained from the range of angles over which current can enter the spokes so that

$$\alpha = 1 - \frac{n(\theta_2' - \theta_1')}{2\pi}. \tag{15.100}$$

From (15.99) the current drawn from the cathode is

$$I_c \approx \frac{0.7}{1+\alpha} I_{CL}. \tag{15.101}$$

We note that this tends to the expected values in the limits $\alpha \to 0$ and $\alpha \to 1$. The anode current is found by substitution in (15.94) to be

$$I_a = \left(\frac{1-\alpha}{1+\alpha}\right) 0.7 I_{CL} \qquad (15.102)$$

and the back bombardment current striking the cathode is

$$I_b = \left(\frac{\alpha}{1-\alpha}\right) I_a. \qquad (15.103)$$

Thus the anode current can be obtained directly from α, the hub radius, and the anode voltage. The variation in current with changes of the RF voltage is determined solely by changes in α.

15.6.2   The Rigid Spoke Model

The simplest approach to finding the properties of the spokes was suggested by Welch [37] and subsequently developed by Hull and Vaughan [13, 59]. A necessary, but not sufficient, condition for electrons to enter the spokes is that they should be in the retarding phase of the tangential RF electric field. However, if they are to be able to reach the anode, then the sum of the DC and RF voltages on the anode must also exceed the threshold voltage given by (15.39), as shown in Figure 15.33. Then, in the rotating frame of reference which is synchronous with the wave on the anode,

$$V_a + V_1 \cos n\theta' \geq V_T, \qquad (15.104)$$

where V_1 is the amplitude of the synchronous RF wave, which is a travelling wave in the stationary frame of reference, and a stationary wave in the synchronous frame. Equation (15.104) shows that the tube can conduct when the anode voltage is below
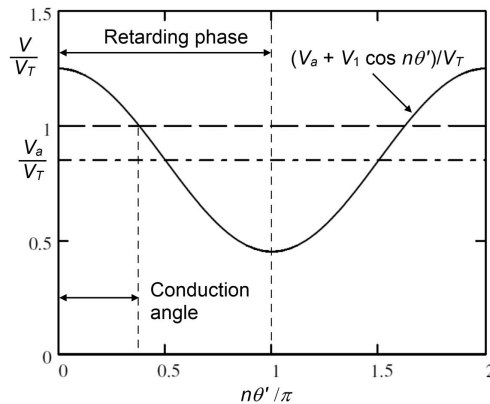
Figure 15.33: Criterion for the conduction angle of a magnetron proposed by Welch (copyright 1953, IEEE, reproduced, with permission, from [37]).

the threshold voltage if the inequality in (15.104) is satisfied. The conduction angle within which an electron can reach the anode ranges from $\theta_1' = 0$ to

$$\theta_2' = \frac{1}{n}\cos^{-1}\left(\frac{V_T - V_a}{V_1}\right), \tag{15.105}$$

and the fraction of the electrons returned to the cathode is

$$\alpha = 1 - \frac{1}{2\pi}\cos^{-1}\left(\frac{V_T - V_a}{V_1}\right). \tag{15.106}$$

Using the data in Table 15.2 the fractions of the current returned to the cathode given by (15.106) are 0.67, 0.71, and 0.64, respectively, for the three cases. The first two figures are close to those in the table so the rigid spoke model is a valid approximation when the current is space-charge limited.

When a zero-order hub is assumed, the anode current, obtained from (15.19) and (15.98), is

$$I_a = \frac{n^2\rho_b L_a V_1}{B_z}\left(\frac{R_b^n - R_b^{-n}}{R_a^n - R_a^{-n}}\right)\int_0^{\theta_c'}\sin\left(n\theta'\right)d\theta' \tag{15.107}$$

Evaluating the integral we find that

$$I_a = \frac{n\rho_b L_a V_1}{B_z}\left(\frac{R_b^n - R_b^{-n}}{R_a^n - R_a^{-n}}\right)(1 - \cos n\theta_2'). \tag{15.108}$$

For a first-order hub, the anode current is given by (15.102), where α is given by (15.106), and the Child–Langmuir current has been calculated from (5.70) for a diode with inner radius r_c, outer radius r_b, length L_a, and applied voltage V_b. The tip of the spoke is assumed to reach the anode at the angle $\theta_s' = \theta_2'/2$.

15.6.3 Guiding Centre Orbits

A more accurate model of the spokes is provided by the motion of the guiding centres of the electron trajectories [60, 61, 63]. In this model the rapid cycloidal motion of the electrons in the spokes, and the space-charge field, are ignored. The papers cited deal only with the case of a planar magnetron but the same approach can also be used for circular geometry. It is assumed that, in the synchronous frame of reference, the radial acceleration and the angular velocity are small so that the equations of motion (15.30) and (15.31) can be rewritten using (15.33) and (15.34) as

$$\left\{\omega_c - 2\omega_s - \frac{d\theta'}{dt}\right\}r\frac{d\theta'}{dt} = -\frac{e}{m_0}E_r - \left\{\omega_c - \omega_s\right\}r\omega_s, \tag{15.109}$$

and

$$\left\{\omega_c - 2\omega_s - 2\frac{d\theta'}{dt}\right\}\frac{dr}{dt} = \frac{e}{m_0}E_\theta, \tag{15.110}$$

where the cyclotron frequency $\omega_c = eB_z/m_0$. Now, typically ω_s and $d\theta'/dt$ are around 10% of ω_c so that the products of these terms can be neglected in the first approximation. Then (15.109) and (15.110) become

$$\frac{d\theta'}{dt} = -\frac{E_r}{B_z r} - \omega_s, \tag{15.111}$$

where E_r is the sum of the DC and RF fields given by (15.3) and (15.18), and

$$\frac{dr}{dt} = \frac{E_\theta}{B_z}. \tag{15.112}$$

We note that (15.112) is identical to (15.34) if the tangential force F_θ is negligible. Now the electric field can be derived from the sum of the DC and RF electrostatic potentials in the rotating frame of reference so that

$$\frac{d\theta'}{dt} = \frac{1}{B_z r} \frac{\partial}{\partial r} \left(V - \frac{m_0}{2e} \omega_s \omega_c r^2 \right) \tag{15.113}$$

and

$$\frac{dr}{dt} = -\frac{1}{B_z r} \cdot \frac{\partial V}{\partial \theta'}. \tag{15.114}$$

These equations can be expressed in terms of a total potential defined by

$$U = V - \frac{m_0}{2e} \omega_c \omega_s \left(r^2 - r_c^2 \right), \tag{15.115}$$

which is zero on the surface of the cathode so that

$$\frac{d\theta'}{dt} = \frac{1}{B_z r} \cdot \frac{\partial U}{\partial r} \tag{15.116}$$

and

$$\frac{dr}{dt} = -\frac{1}{B_z r} \cdot \frac{\partial U}{\partial \theta'}. \tag{15.117}$$

From (15.116) we see that on the surface of the cathode, where $V = 0$, the angular velocity is

$$\frac{d\theta'}{dt} = -\omega_s \tag{15.118}$$

as expected. Substituting for the DC and RF potentials in the rotating frame of reference from (15.1), (15.2) and (15.17) we can write (15.115) as

$$U(r, \theta') = V_0(r) + V_1(r) \cos(n\theta') - \frac{m_0}{2e} \omega_c \omega_s \left(r^2 - r_c^2 \right). \tag{15.119}$$

The kinetic energy of the electrons is

$$T(r,\theta') = \frac{m_0}{2e}\left(\dot{r}^2 + r^2\dot{\theta}'^2\right)$$ (15.120)

and the Hamiltonian for the electron trajectories is

$$H = T + V.$$ (15.121)

This quantity is a constant for the motion of an electron so that the trajectory of an electron which is initially on the surface of the hub at (r_b,θ_0') is given by

$$H(r,\theta') = H(r_b,\theta_0').$$ (15.122)

It is found that, typically, $T \sim 0.1\,U$ so that it can be neglected in the first approximation. Then, from (15.119)

$$\cos(n\theta') = \frac{1}{V_1(r)}\left\{V_0(r_b) - V_0(r) + V_1(r_b)\cos(n\theta_0') + \frac{m_0}{2e}\omega_c\omega_s\left(r^2 - r_b^2\right)\right\}.$$ (15.123)

This equation can be used to calculate the trajectories of the guiding centres of the electrons forming the spokes. Figure 15.34 shows a typical set of trajectories in the retarding phase of the tangential electric field. The radius is normalised to the radius of the hub and the trajectories shown intersect the hub at 30° intervals. The electrons move outwards from the hub because the field is retarding. Their initial tangential velocity is greater than the synchronous velocity so that they move forwards in phase. In this example, electrons with initial phases between 0° and 120° can reach the anode, while those with phases in the range 150° to 180° move into the accelerating phase before reaching the anode. The limiting phase is represented by a trajectory which is moving directly outwards when its phase is 180°. Trajectories lying above, and to the right of, this line are non-physical because they can never
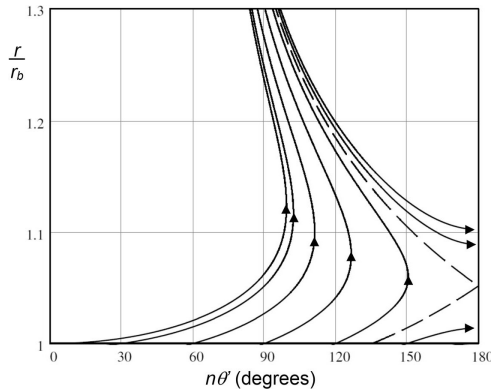
Figure 15.34: Typical guiding centre trajectories which cross the surface of the hub at 30° intervals, and the limiting trajectory (broken line).

be reached by electrons starting from the cathode. The trajectories are plotted in Cartesian coordinates for convenience so that motion to the right corresponds to anticlockwise rotation of the spokes. Setting $d\theta'/dt = 0$ in (15.111), substituting for the radial electric field, and setting $n\theta' = \pi$ gives

$$\frac{e}{4m_0} \cdot \frac{B_z r_b^2}{r^2}\left(1 - \frac{r_c^4}{r_b^4}\right) - \frac{nV_s}{B_z r^2}\left(\frac{r^n + r^{-n}}{r_a^n - r_a^{-n}}\right) - \omega_s = 0, \tag{15.124}$$

which can be solved for the radius r at which $n\theta' = \pi$. Equation (15.123) can then be solved for θ_2' corresponding to the limiting trajectory which is plotted as a dashed line in Figure 15.34. The conduction angle is usually greater than that assumed in the rigid spoke model. The anode current can be calculated for zero- and first-order hubs using (15.108) and (15.102). The empirical correction made by Riyopoulos, that the charge density in a zero-order hub should be approximately half of that given by (15.96) in order to get agreement with experimental data, has not be used here. An empirical model of the space-charge in the spokes is described in [60, 61].

15.6.4   Electron Trajectory Model

The electron trajectories can be calculated directly by numerical integration of the equations of motion (15.30) and (15.31). The trajectories are initially equally spaced on the surface of the space-charge hub with angular velocities given from (15.1) and (15.4) by

$$\frac{d\theta'}{dt} = \frac{1}{2}\omega_c\left(1 - \frac{r_c^2}{r_b^2}\right) - \omega_s. \tag{15.125}$$

The space-charge of the spokes, and modification of the hub by the RF fields, are ignored and the integration is continued until the trajectories end on either the cathode or the anode. The fraction of the current returned to the cathode and the limiting initial angles between which trajectories can reach the anode can be determined directly.

Figure 15.35 shows an example of the trajectories calculated together with the RF voltage and the threshold voltage normalised to the anode voltage (see Worksheet 15.3). Figure 15.36 shows a detail of the trajectories in one RF cycle, plotted using Cartesian coordinates for convenience. The spoke assumed by the rigid spoke model, and the spoke calculated from the guiding centre model, are shown, for comparison. These calculations are for the 4J52 magnetron, at the operating point shown in Figure 15.27. The DC and RF voltages were computed using the rigid spoke model with a zero-order hub. It can be seen that the spoke calculated by the guiding centre model is similar to the computed trajectories. The difference is attributable to the approximations made in the derivation of the guiding centre model. The phase range within which electrons can reach the anode, and the phase of the tip of the spoke in the rigid spoke model, are similar to those from the guiding centre model. The principal difference is that the rigid spoke model ignores the initial angular velocity of the electrons at the base of the spoke.
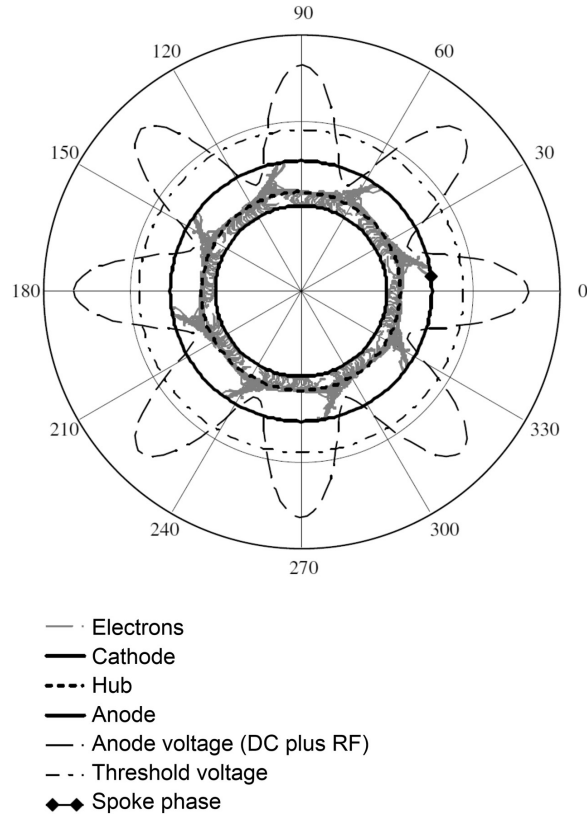
Figure 15.35: Spokes computed for the 4J52 magnetron at its normal operating point with the RF voltage and the threshold voltage normalised with respect to the anode voltage. The rotation of the spokes is counter-clockwise. The position of the spoke tip given by the rigid spoke model is shown for comparison.
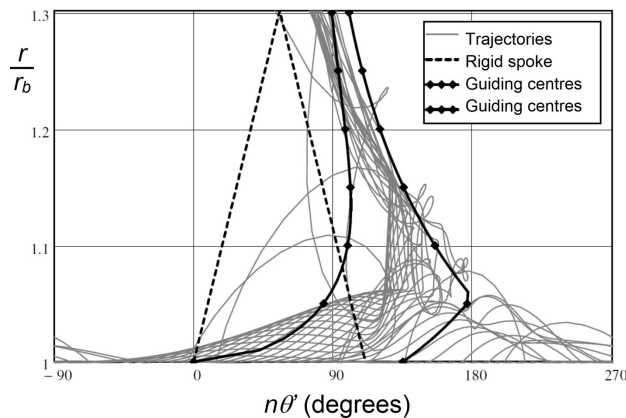
Figure 15.36: Detail of the trajectories of the 4J52 magnetron in one RF cycle plotted using Cartesian co-ordinates for convenience, together with the spokes predicted by the simple and the guiding centre models. Note that counter-clockwise rotation of the spokes is represented by motion to the right in this figure.

15.6.5   Calculation of the Output Power

The energy with which the electrons strike the anode is

$$V_{ia} = \frac{m_0}{2e}\left(\dot{r}^2 + r_a \dot{\theta}^2\right). \tag{15.126}$$

Substituting for the radial and angular velocities at the tip of the spoke, from (15.112) and (15.111), and multiplying by the anode current, gives the power dissipated on the anode. This calculation avoids the arbitrary assumption about the tangential velocity made by Vaughan [13]. It should be noted that the angular velocity of the electrons at the tip of the spoke is normally slightly greater than ω_s. It is not possible to calculate the power dissipated on the cathode in the same manner. Vaughan [13] proposed the empirical expression

$$P_c = \frac{0.04 I_a V_a}{\cos(n\theta)} \tag{15.127}$$

on the basis that this power is typically around 5% of the input power at the normal operating point, but it may be up to 50% at very low inputs when $n\theta_s \to \pi/2$. Because the power dissipated on the cathode is usually small, errors in its estimation do not result in major errors in the calculated performance. In the trajectory calculations the impact energies on the anode and the cathode can be calculated directly. Since the anode current is known, the RF power transferred to the anode, and the electronic efficiency, can be calculated from (15.45) and (15.46) as functions of the anode voltage, the RF wave voltage, and the magnetic field.

A second expression for the RF power delivered to the anode is

$$P_a = \frac{1}{2}\left(G_U + G_E\right)V_g^2 \tag{15.128}$$

where G_U and G_E are the conductances of the anode, and of the external load, and V_g is the magnitude of the RF voltage as measured between adjacent vane tips. Now V_g is related to the wave voltage by (15.20) and therefore (15.128) gives the RF power delivered to the anode as a function of V_1. The operating point can then be determined by solving the simultaneous equations (15.45) and (15.128) for V_a and V_1, given the values of I_a and B_z. Figure 15.37 shows the performance chart calculated for the 4J50/4J52 magnetron for comparison with the experimental results in Figure 15.27. The rigid spoke model with a zero-order hub was used for this calculation. Note that the magnetic fields represented by the Gauss lines are not identical in the two figures, also that this tube differs from the Litton 4J50 modelled by Vaughan [13]. Overall, the rigid spoke model produces results which are close enough to those obtained from experiment for initial exploration of the properties of a proposed design.

Table 15.3 shows a comparison between the results given by these models (see Worksheet 15.3). Overall the results from both models with zero-order hubs agree with experiment to a similar extent. Both sets are closer than the results

Table 15.3: 4J52 magnetron: Comparison between the results given by the models discussed in the text, and experimental data from [35]

Hub order	Zero	First	Zero	First	Experiment
Spoke	Rigid	Rigid	Guiding centre	Guiding centre	
Anode voltage (kV)	16.5	18.2	13.2	17.3	16
V_a/V_T	1.15	1.26	0.92	1.20	
Anode current (A)	27.0	27.0	27.0	27.0	27
Output power (kW)	198	220	174	248	180
Efficiency (%)	44	45	49	53	42
Anode dissipation (%)	32	32	26	20	
Cathode dissipation (%)	5	5	4	4	
Circuit loss (%)	19	19	20	23	
Current fraction returned (α)	0.69	0.64	0.50	0.62	
Wave voltage / DC voltage	0.36	0.35	0.42	0.39	
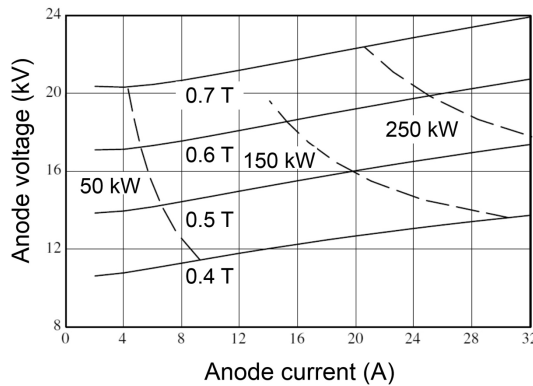
Figure 15.37: Calculated performance chart for the 4J50/4J52 magnetron for comparison with Figure 15.27.

of PIC simulation in Table 15.2. Trajectory calculations, based on the electric fields determined using the rigid spoke model with a zero-order hub, give $\alpha = 0.52$. The reason that this figure is smaller than the one in the table is that some electrons for which $\theta'_1 < 0$ can enter the spokes. Those electrons are accelerated and move forwards just below the surface of the hub until they enter the retarding phase of the field. It can be argued that the guiding centre model gives similar results if trajectories lying inside the hub are permitted. Strictly speaking, the Welch criterion

(see Figure 15.33) determines the range of angles from which electrons leaving the cathode can reach the anode. However, because the cathode current is zero in a zero-order hub, it is necessary to make some assumption about the extent of the bases of the spokes in order to determine the current in them. In addition, the method for determining the current assumes that the radial velocities of the electrons are zero inside the hub, and non-zero at the base of the spoke. Thus there are a number of objections to theories based on the zero-order hub. In principle, the use of a first-order hub could overcome these difficulties. In the rigid spoke model the current drawn from the cathode can be determined directly using the Welch criterion without making any assumptions about the shape of the spoke. The guiding centre model effectively replaces the Welch criterion. The assumption that the RF fields are zero inside the hub means that the conduction angle calculated using the guiding centre model also applies on the surface of the cathode. At present the rigid spoke model based on a zero-order hub appears to give the best agreement with experiment.

15.6.6   The Rieke Diagram

If the load is varied we can write the external admittance as

$$Y_E = G_{E0}\left(\frac{1+\rho\exp j\phi}{1-\rho\exp j\phi}\right) = G_E + jB_E, \tag{15.129}$$

where ρ is the voltage reflection coefficient of the load and ϕ its phase angle. Substituting this expression into (15.128), and applying the same method as before, we can plot the contours of constant power on a Rieke diagram. Also, using (15.129) in combination with (15.89), the frequency shift as a function of reflection coefficient and phase angle is

$$\Delta f = \frac{f_0}{2Q_E} B_E \tag{15.130}$$

from which the contours of constant Δf can be plotted. Figure 15.38 shows the Rieke diagram computed for the 4J50 magnetron, operating with a constant current of 8.8A, for comparison with Figure 15.29 (see Worksheet 15.3). This current was chosen so that the output power into a matched load was the same as in the experimental data. It should be remembered that the zero of the phase is arbitrary because of the convention that the experimental data is referred to the plane of the output flange.

15.6.7   Frequency Pushing

The RF current in the anode circuit is given by

$$I_1 = V_g G_L\left(1 + jQ_L\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)\right). \tag{15.131}$$
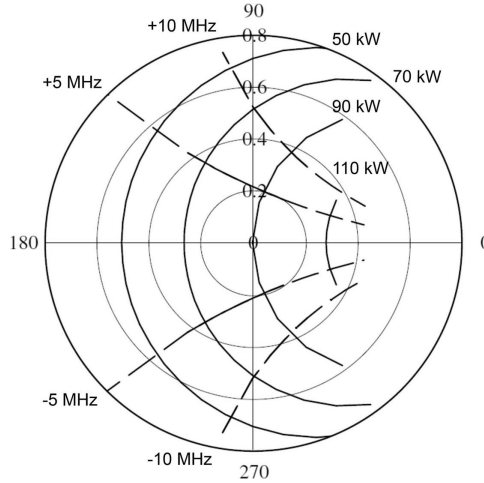
Figure 15.38: Calculated Rieke diagram of the 4J50 magnetron for comparison with Figure 15.29.

Thus the phase angle of the current relative to the voltage is given by

$$\tan \phi = Q_L \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right) \approx Q_L \frac{2 \Delta \omega}{\omega_0}. \tag{15.132}$$

Frequency pushing arises because the frequency must change to match the phase of the RF anode current so that (15.132) is satisfied. The RF current flowing into the anode comprises two components: the conduction current intercepted by the anode, and the induced current from the transverse motion of the charge in the spokes. Reference to Figure 15.36 shows that, at the moment when the RF voltage is maximum on the vane at $\theta = 0$, the position of the tip of the spoke is $\theta = \theta_s$ and moving in the positive direction of θ. To a first approximation we may assume that this is also the effective position of the spoke for calculating the induced current. The power transferred to the anode by the induced current is maximum when the phase of the spoke coincides with the position at which the tangential electric field is retarding $(\theta = \pi/2n)$. Thus, the phase of the spoke relative to the tangential field is

$$\phi_s = n\theta_s - \frac{\pi}{2}. \tag{15.133}$$

The power transferred to the anode by the conduction current is maximum when the phase of the spoke coincides with the vane on which the voltage has its greatest negative value $(\theta = \pi/n)$. Hence the phase of the conduction current is

$$\phi_c = n\theta_s - \pi. \tag{15.134}$$

Comparing (15.134) with (15.133) we see that the conduction current lags the induced current by 90°. Thus the currents can be represented by the phasor
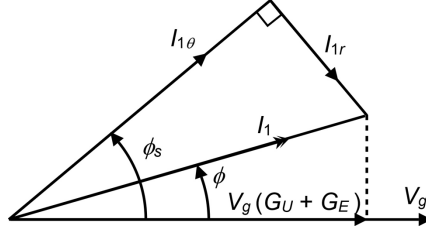
Figure 15.39: Phasor diagram for calculating frequency pushing.

diagram in Figure 15.39 where $I_{1\theta}$ is the induced current, and I_{1r} is the conduction current. The phasor sum of these currents (I_1) is the net current flowing into the anode.

It is difficult to find the induced current directly but, fortunately, this is not necessary. The conduction current can be approximated by a rectangular pulse of phase duration 2Δ and amplitude $(\pi/\Delta)I_a$ where Δ is the width of the tip of the spoke. Then, using Fourier analysis, it can be shown that the amplitude of the first harmonic of the conduction current is

$$I_{1r} = 2I_a \frac{\sin \Delta}{\Delta}. \tag{15.135}$$

Now, from Figure 15.36 we see that Δ is typically of the order of $10°$ so that $I_{1r} = 2I_a$ to a very good approximation. The power transferred to the travelling voltage wave on the anode by the conduction current is

$$P_r = \frac{1}{2} I_{1r} V_1 \cos \phi_c. \tag{15.136}$$

This can also be expressed in terms of the associated component of the current $\left(I_{gr}\right)$ flowing in the anode circuit

$$P_r = \frac{1}{2} I_{gr} V_g \cos \phi_c \tag{15.137}$$

so that

$$I_{gr} = 2I_a \frac{V_1}{V_g}, \tag{15.138}$$

where V_1/V_g is given by (15.20). The component of the RF current which is in phase with the gap voltage is

$$I_g \cos \phi = V_g \left(G_U + G_E\right). \tag{15.139}$$

Then, from the phasor diagram in Figure 15.39,

$$V_g \left(G_U + G_E\right) = I_{g\theta} \cos \phi_s + I_{gr} \sin \phi_s \tag{15.140}$$
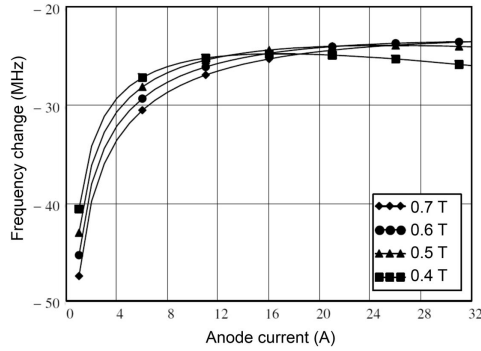
Figure 15.40: Calculated frequency pushing curves for the 4J50/4J52 magnetron.

or

$$I_{g\theta} = \frac{V_g\left(G_U + G_E\right) - I_{gr}\sin\phi_s}{\cos\phi_s},$$ (15.141)

and

$$\tan\phi = \frac{I_{g\theta}\sin\phi_s - I_{gr}\cos\phi_s}{V_g\left(G_U + G_E\right)},$$ (15.142)

so that the frequency pushing curve can be calculated by substitution in (15.132). Figure 15.40 shows the frequency pushing curves calculated for the 4J50/4J52 magnetron for various values of the magnetic field.

15.7   Magnetron Design

The conceptual design of a magnetron is based on the theory described above. For further information about the design of magnetrons see [1, 64]. Discussion of the design of magnetrons of different types can also be found in [29, 33, 65–67].

15.7.1   Dimensionless Parameters

Magnetrons, like other tubes, are commonly designed by scaling from existing designs. We have seen in previous chapters that this means that geometric similarity and dynamic similarity must both be maintained. The first requirement is that all dimensions, and the wavelength at the π mode, are scaled by the same factor. The second requirement means that the tubes must be described by the same normalised parameters. The anode voltage and the magnetic field can be normalised by dividing them by the characteristic voltage (V_0) given by (15.22) and the characteristic field (B_0) given by (15.23). We recall that the point defined by these values lies on

the cut-off curve of the magnetron diode. A scale factor for the current can be defined as the current which flows in a diode when it is just not cut off [9]. We have seen that this current is approximately 0.7 times the Child–Langmuir current, where the constant is a function only of the ratio r_a/r_c provided that the emission from the cathode is great enough to ensure that the current is limited by space-charge (see Section 8.5). The scale factors for power and impedance can then be defined as $P_0 = I_0 V_0$ and $G_0 = I_0/V_0$. The operating point on the performance chart of a tube is uniquely defined by any two of: the magnetic field; the anode voltage; and the anode current. Thus, a universal performance chart can be drawn in terms of dimensionless parameters. This applies to all geometrically similar magnetrons having the same normalised values of G_E/G_0 and G_U/G_0.

An alternative set of dimensionless parameters, which assumes geometrical similarity, is [64]

$$v = \frac{\eta V_a}{r_a^2 \omega_s^2},\tag{15.143}$$

$$b = \frac{\eta B_z}{\omega_s} = \frac{\omega_c}{\omega_s},\tag{15.144}$$

and

$$i = \frac{\eta I_a}{\varepsilon_0 r_a^3 \omega_s^3}.\tag{15.145}$$

15.7.2   Design Parameters

The basic specification of a magnetron includes the frequency, the power output, and the duty cycle. Because the current in a magnetron does not obey the 3/2 power law the concept of perveance is not useful. It is usual, instead, to use estimates of the efficiency, and the DC impedance, to find initial figures for the DC current and voltage. Table 15.4 shows typical values for different types of magnetron based on manufacturers' data sheets. The circuit efficiency is typically 90–95% at frequencies up to 3 GHz, falling to 70–80% at 9 GHz, and 50–60% at 24 GHz.

15.7.3   Design Case Study

This section illustrates the conceptual design of a 500 kW, 915 MHz, CW magnetron for microwave heating described by Shibata in [66] (see Worksheet 15.4). The principal steps of the design calculation below are not identical to those used by Shibata but the parameters have been chosen to give very nearly the same results.

i)   Since this is an exceptionally high-power tube we estimate, conservatively, from Table 15.4 that the efficiency will be 80%. In order to keep the anode voltage as

Table 15.4: Typical impedances and efficiencies for different types of magnetron. The output powers and anode currents for pulsed tubes are peak values

Anode	Duty	Frequency (GHz)	Output power (kW)	Anode current (A)	Impedance (kΩ)	Overall Efficiency (%)
Strapped	CW	0.9	30–100	3–6	3.0–5.0	80–90
Strapped	CW	2.45	0.2–30	2–40	3.0–7.0	65–70
Strapped	pulse	2–4	5–4500	4–130	0.5–1.5	35–50
Strapped	pulse	4–8	200–1000	25–80	0.5–1.0	30–50
Strapped	pulse	8–11	1–2000	2–100	0.4–1.4	30–50
Strapped	pulse	11–18	20–250	8–28	0.8–1.2	20–45
Rising sun	pulse	25–35	40–125	13–28	0.7–1.0	20–25
Coaxial	pulse	8–10	10–350	4–27	0.8–1.1	30–50
Coaxial	pulse	14–17	1–180	2–25	0.8–1.9	20–40
Coaxial	pulse	32–36	25–130	10–20	1.0–1.4	20–30
Long anode	pulse	3	2000–5500	50–200	0.2–0.5	45–55

low as possible the DC impedance is taken to be 2.9 kΩ. The anode voltage and current are then found to be 44.0 kV and 15.2 A.

ii) The normalised magnetic field is related to the electronic efficiency by (15.49). However, this equation does not take account of the power lost by cathode back-bombardment and by RF heating of the anode. If the circuit efficiency is assumed to be 90%, and the cathode back-bombardment power is ignored, then $B_z/B_0 = 6.5$ and, from (15.40), $V_T/V_0 = 12.0$.

iii) Figure 15.27 shows that the working anode voltage of a magnetron is usually close to the threshold voltage. Therefore, we assume that $V_a/V_T = 1$ so that the characteristic voltage $V_0 = 3.7$ kV. The cut-off voltage is found from (15.25) to be 155 kV.

iv) From (15.22) the anode radius is given, as function of the number of vanes, by

$$r_a = \frac{\sqrt{2(e/m_0)V_0}}{2\omega} N_v. \qquad (15.146)$$

The choice of the radius of the space-charge hub is very important for correct magnetron operation. Since the RF fields in a planar magnetron and a cylindrical magnetron are similar it is possible to deduce the design criteria for a cylindrical tube from a consideration of the simpler, planar, geometry. If the normalised depth of the space-charge hub in a planar magetron (d/h) is too small there is a large gap between the surface of the space-charge layer and the anode, and the field of the π mode at the surface of the hub is smaller than that of the $\pi-1$ mode. There is, then, a serious risk of $\pi-1$ mode excitation. If d/h is

too large then the potential difference between the surface of the space charge layer and the anode is small and the tube efficiency is low.

In the π mode the normalised tangential RF field on the surface of a planar hub obtained from (15.11) is

$$\frac{E_x(x,d)}{E_x(x,h)} = \frac{\sinh(\pi d/p)}{\sinh(\pi h/p)} \approx \exp(\pi(d-h)/p).$$
(15.147)

The approximation in (15.147) is valid if $d/p > 2$. Then a figure of merit (the Slater ratio) can be defined by

$$R = -\ln\left|\frac{E_x(x,d)}{E_x(x,h)}\right| = \frac{\pi(h-d)}{p}.$$
(15.148)

We expect that empirical evidence will suggest the best value of R for optimum interaction. Eliminating B_z between (8.12) and (8.25) gives

$$\frac{V_a}{V_H} = \frac{d}{h}\left(2 - \frac{d}{h}\right).$$
(15.149)

Solving for d/h gives

$$\frac{d}{h} = 1 - \sqrt{1 - \frac{V_a}{V_H}}.$$
(15.150)

Then, substituting into (15.148), we have

$$R = \frac{\pi h}{p}\sqrt{1 - \frac{V_a}{V_H}}.$$
(15.151)

In this equation the inaccessible variable d has been replaced by the ratio of the anode voltage to the cut-off voltage. It is not possible to find an equivalent closed-form expression for circular geometry, so (15.151) is adapted by assuming that the planar parameters can be defined in terms of the circular parameters by

$$h = r_a - r_c,$$
(15.152)

and

$$p = \frac{\pi(r_a + r_c)}{N_v}.$$
(15.153)

This definition of p assumes that it is the mean circumference of the interaction region which matters. Note that p is not the pitch of the circular anode. The figure of merit for a cylindrical magnetron, known as the *Modified Slater Factor*, is therefore given by

$$R' = N_v\left(\frac{r_c - r_a}{r_c + r_a}\right)\sqrt{1 - \frac{V_a}{V_H}}.$$
(15.154)

Experience suggests that R' should have a value of about 2 for a wide range of designs. Making use of (15.146) and (15.154) we can tabulate the anode and cathode radii for a range of values of N_v. We recall that the length of the anode should be small, compared with the free-space wavelength, to avoid the excitation of longitudinal modes. Since this is an exceptionally high-powered tube we choose $L_a = 0.16\lambda = 53$ mm. Then the ratio of the anode current to the surface area of the cathode can be calculated. If it is assumed that 75% of the electrons are returned to the cathode then the cathode current is four times the anode current so that the cathode current density can be calculated. The results of this calculation are shown in Table 15.5.

In a very high power CW magnetron the design is limited by the maximum temperature of the vane tips rather than by the cathode current density. If the power dissipated on the cathode is neglected the power dissipated by the is anode is given by

$$P_a = P_{dc} - P_{rf},\qquad(15.155)$$

and the vane tip temperature is

$$T_v = P_a \frac{L_{th}}{\kappa A_v} + T_r,\qquad(15.156)$$

where L_{th} is the thermal path distance to the cooling channel, κ is the thermal conductivity of copper (401 W m^{-1} K^{-1}), A_v is the total cross-sectional area of the vanes normal to the heat flow, and T_r is temperature of the cooling channel. The thickness of the vanes (t) is chosen to that the ratio of the thickness to the pitch is approximately 0.5. The pitch calculated from (15.146) is 20 mm, so let us choose the vane thickness to be 12 mm. To cool the anode effectively water cooling channels must be embedded in the vanes. Let us suppose that the effective thermal path is 10 mm. The vane tip temperatures calculated assuming that $T_r = 293$ K are shown in Table 15.5.

All the numbers of vanes considered result in cathode currents which are comfortably within the capability of thoriated tungsten [66, 68]. This is the preferred choice of cathode material because a back-bombardment power of around 14 kW is anticipated. The experimental tube used a pure tungsten wire cathode which was pre-heated to 1870 °C and operated at 2100 °C. The cathode terminals were water-cooled to reduce the heating by back-bombardment. The maximum anode temperature must be less than 700 °C to avoid re-crystallisation of the oxygen-free high-conductivity copper of which it is made. To allow a safety margin to cover local hot-spots the maximum temperature should be 500 °C (773 °K). It is desirable to keep the number of vanes as small as possible to reduce problems with excitation of the $\pi - 1$ mode. Therefore we choose a 14-vane anode.

v) Once the cathode and anode radii have been determined, the characteristic field can be found from (15.23) and, hence, the magnetic flux density required. The results are: $B_0 = 0.019$ T and $B_z = 0.121$ T. We note that B_z is well below the saturation flux density for soft iron (see Section 19.4.1), and therefore no problems

Table 15.5: Design choices for a 500 kW magnetron

N_v	r_a (mm)	r_c (mm)	J_c (A cm^{-2})	T_v (K)
8	25.2	13.7	1.34	1117
10	31.4	19.4	0.95	952
12	37.7	25.3	0.73	842
14	44.0	31.3	0.59	764
16	50.3	37.3	0.49	705
18	56.6	43.4	0.42	659
20	62.9	49.6	0.37	622
22	69.2	55.7	0.33	593
24	75.5	61.9	0.30	568

with non-uniformity of the field should occur. The magnetic field of the experimental tube described in [66] was provided by a solenoid magnet. The ratio of the cyclotron frequency to the frequency of oscillation is 3.7 so that the difference is much greater than the figure of 20% which is desirable.

vi) The threshold voltage for the $\pi - 1$ mode can be found from (15.40) after calculating V_0 and B_0 with $n = 6$. The result is 50.6 kV which is comfortably above the working voltage.

The experimental magnetron had a coaxial anode to increase the circuit efficiency and the overall efficiency. No details of the properties of the anode are given in [66]. Although they could be calculated we shall infer them from simulation of the tube using the rigid spoke model with a zero-order hub. Table 15.6 shows the leading parameters and performance characteristics of this tube.

15.7.4   Other Considerations

The analysis and design methodology described above involve a number of simplifying assumptions which may not always be valid. It is found that magnetrons do not always behave as expected, for reasons which are not understood. The phenomenon of twinning (see Section 15.4.4) is just one example. These problems can occur as a result of tiny differences in manufacture between different batches of tubes. They may also depend upon the history of operation, and have a big effect on the satisfactory operation of the tube and its lifetime. Possible causes include: variations in the primary and secondary emission properties of the cathode; changes in the resistance of the heater; higher-order resonances of the anode and the end spaces; the external impedance presented by the power supply; and the presence of residual gas. These problems are generally solved by empirical 'fixes' and further research is needed to explain and understand them. Cooker magnetrons, for example, require a filter box connected to the cathode, to control the impedance between the anode and the cathode, if they are to work correctly. It has been found difficult to scale CW magnetrons to work at frequencies above 2.45 GHz [69].

Table 15.6: Leading parameters and performance characteristics of a 500 kW magnetron

Parameter	Calculated	Calculated [66]	Measured [66]
V_a (kV)	44.0	44	44
I_a (A)	15.2	15	16
B_z (T)	0.121	0.10	0.10
V_0 (kV)	3.7	3.7	
B_0 (T)	0.019	0.0153	
V_T/V_0	11.9	12.0	
B_z/B_0	6.4	6.5	
N_v	14	14	14
r_a/r_c	1.41	1.59	
r_c (mm)	31	27.7	
r_a (mm)	44	44	
L_a (mm)	53	(50)	
t (mm)	12	(12.1)	
Y_0 (1/Ω)	0.47		
Q_U	1000		
P_{rf} (kW)	510 kW		421
η_o	81%		72.5%

References

[1] E. Okress, Ed., *Crossed-Field Microwave Devices*. New York: Academic Press, 1961.

[2] A. S. Gilmour, Jr., *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-Field Amplifiers and Gyrotrons*. Norwood, MA: Artech House, 2011.

[3] L. Sivan, *Microwave Tube Transmitters*. Kluwer Academic Publishers, 1994.

[4] W. C. Brown, 'The microwave magnetron and its derivatives', *IEEE Transactions on Electron Devices,* vol. 31, pp. 1595–1605, 1984.

[5] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer-Verlag, pp. 1–84, 2008.

[6] S. H. Gold and G. S. Nusinovich, 'Review of high-power microwave source research', *Review of Scientific Instruments,* vol. 68, p. 3945–3974, 1997.

[7] J. Benford *et al.*, *High Power Microwaves*. CRC Press, 2015.

[8] J. F. Hull, 'Inverted Magnetron', *Proceedings of the IRE,* vol. 40, pp. 1038–1041, 1952.

[9] J. C. Slater, *Microwave Electronics*. New York: D. van Nostrand, 1950.

[10] J. F. Skowron, 'The continuous-cathode (emitting-sole) crossed-field amplifier', *Proceedings of the IEEE,* vol. 61, pp. 330–356, 1973.

[11] W. A. Smith, 'A wave treatment of the continuous cathode crossed-field amplifier', *IRE Transactions on Electron Devices,* vol. 9, pp. 379–387, 1962.

[12] R. G. E. Hutter, *Beam and Wave Electronics in Microwave Tubes*. Princeton, NJ: D. van Nostrand, 1960.

[13] J. R. M. Vaughan, 'A model for calculation of magnetron performance', *IEEE Transactions on Electron Devices,* vol. 20, pp. 818–826, 1973.

[14] G. B. Collins, *Microwave Magnetrons*. New York: McGraw-Hill, 1948.

[15] N. Kroll, 'The unstrapped resonant system', in G. B. Collins, ed., *Microwave Magnetrons*. New York: McGraw-Hill, pp. 49–82, 1948.

[16] S. Millman and W. V. Smith, 'The resonant system', in G. B. Collins, ed., *Microwave Magnetrons*. New York: McGraw-Hill, pp. 460–502, 1948.

[17] R. G. Carter, *Electromagnetism for Electronic Engineers*. Fredriksberg, Denmark: Ventus Publishing, 2010.

[18] R. M. Gilgenbach *et al.*, 'Cathode priming of magnetrons for rapid startup and mode-locking', in *The Joint 30th International Conference on Infrared and Millimeter Waves and 13th International Conference on Terahertz Electronics*, vol. 2, pp. 535–536, 2005.

[19] V. B. Neculaes *et al.*, 'Magnetic perturbation effects on noise and startup in DC-operating oven magnetrons', *IEEE Transactions on Electron Devices,* vol. 52, pp. 864–871, 2005.

[20] V. B. Neculaes *et al.*, 'Magnetic priming effects on noise, startup, and mode competition in magnetrons', *IEEE Transactions on Plasma Science,* vol. 33, pp. 94–102, 2005.

[21] J. I. Kim *et al.*, 'Reduction of noise in strapped magnetron by electric priming using anode shape modification', *Applied Physics Letters,* vol. 88, p. 221501, 2006.

[22] W. V. Smith, 'Mechanical tuning', in G. B. Collins, ed., *Microwave Magnetrons*. New York: McGraw-Hill, pp. 561–591, 1948.

[23] M. J. Bernstein and N. M. Kroll, 'Mechanically tuned rising-sun magnetrons', in E. Okress, ed., *Crossed-Field Microwave Devices,* vol. 2. New York: Academic Press, pp. 149–153, 1961.

[24] R. B. Nelson, 'Methods of tuning multiple-cavity magnetrons', *Proceedings of the IRE,* vol. 36, pp. 53–56, 1948.

[25] A. Pickering *et al.*, 'Electronically tuned pulse magnetron', in *IEEE International Electron Devices Meeting*, pp. 145–148, 1975.

[26] S. P. Yu and P. N. Hess, 'Slow-wave structures for M-type devices', *IRE Transactions on Electron Devices,* vol. 9, pp. 51–57, 1962.

[27] J. F. Hull, 'Crossed field electron interaction in space charge limited beams', Doctor of Electrical Engineering, Polytechnic Institute of Brooklyn, New York, 1958.

[28] J. Feinstein and R. J. Collier, 'The circular electric mode magnetron', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 123–134, 1961.

[29] A. J. Bamford *et al.*, 'A 1-MW L-band coaxial magnetron with separate cavity', *IEEE Transactions on Electron Devices,* vol. 14, pp. 844–851, 1967.

[30] T. E. Ruden, 'Design and performance of a one megawatt, 3.1–3.5 GHz coaxial magnetron', in *9th European Microwave Conference*, pp. 731–735, 1979.

[31] A. H. Pickering, 'Further developments of long anode magnetrons', in E. Okress, ed., *Crossed Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 275–290, 1961.

[32] H. A. H. Boot, 'Long anode magnetrons', in E. Okress, ed., *Crossed Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 261–274, 1961.

[33] H. A. H. Boot *et al.*, 'A new design of high-power S-band magnetron', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 105, pp. 419–425, 1958.

[34] J.-J. Feng *et al.*, 'Simulation of a long anode magnetron resonant system using MAFIA', in *International Conference on Microwave and Millimeter Wave Technology*, pp. 748–751, 1998.

[35] A. G. Smith, 'Typical magnetrons', in G. B. Collins, ed., *Microwave Magnetrons*. New York: McGraw-Hill, pp. 739–796, 1948.

[36] C. R. Schumacher, 'Frequency pushing', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 401–422, 1961.

[37] H. W. Welch, 'Prediction of traveling wave magnetron frequency characteristics: frequency pushing and voltage tuning', *Proceedings of the IRE,* vol. 41, pp. 1631–1653, 1953.

[38] W. L. Pritchard, 'Long-line effect and pulsed magnetrons', *IRE Transactions on Microwave Theory and Techniques,* vol. 4, pp. 97–110, 1956.

[39] R. Yokoyama and A. Yamada, 'Development status of magnetrons for microwave ovens', in *Microwave Power Symposium*, pp. 132–135, 1996.

[40] J. M. Osepchuk, 'The cooker magnetron as a standard in crossed-field device research', in *1st Int. Workshop on Crossed-Field Devices*, University of Michigan, USA, pp. 159–177, 1995.

[41] C. R. Schumacher, 'Spectrum shape', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 457–471, 1961.

[42] R. G. Carter *et al.*, 'Magnetron frequency twinning', *IEEE Transactions on Plasma Science*, vol. 28, pp. 905–909, June 2000.

[43] 'Preamble: Magnetrons', Chelmsford, UK: EEV Ltd., 1974.

[44] Varian, 'Technical Manual: Installation, Operation, Maintenance, Care and Handling Instructions, General: Microwave Tubes, Magnetron Tubes, Electron Tubes', 1 October 1979.

[45] R. Adler, 'A study of locking phenomena in oscillators', *Proceedings of the IRE,* vol. 34, pp. 351–357, 1946.

[46] H. L. Thal and R. G. Lock, 'Locking of magnetrons by an injected RF signal', *IEEE Transactions on Microwave Theory and Techniques,* vol. 13, pp. 836–846, 1965.

[47] I. Tahir *et al.*, 'Noise performance of frequency- and phase-locked CW magnetrons operated as current-controlled oscillators', *IEEE Transactions on Electron Devices,* vol. 52, pp. 2096–2103, September 2005.

[48] I. Tahir *et al.*, 'Frequency and phase modulation performance of an injection-locked CW magnetron', *IEEE Transactions on Electron Devices,* vol. 53, pp. 1721–1729, July 2006.

[49] W. C. Brown, 'The sophisticated properties of the microwave oven magnetron', in *IEEE MTT-S International Microwave Symposium Digest*, vol. 3, pp. 871–874, 1989.

[50] E. E. David, Jr., 'Phasing by RF signals', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 375–399, 1961.

[51] H. Kim and J. Choi, 'Characterization of a 16-vane strapped magnetron oscillator by three-dimensional particle-in-cell code simulations', *Current Applied Physics,* vol. 6, pp. 66–70, 2006.

[52] A. D. Andreev and K. J. Hendricks, 'ICEPIC simulation of a strapped nonrelativistic high-power CW UHF magnetron with a solid cathode operating in the space-charge limited regime', *IEEE Transactions on Plasma Science,* vol. 40, pp. 1551–1562, 2012.

[53] Y. Alfadhl *et al.*, 'Advanced computer modelling of magnetrons', in *2010 International Conference on the Origins and Evolution of the Cavity Magnetron (CAVMAG)*, pp. 67–70, 2010.

[54] G. E. Dombrowski, 'Simulation of magnetrons and crossed-field amplifiers', *IEEE Transactions on Electron Devices,* vol. 35, pp. 2060–2067, 1988.

[55] G. E. Dombrowski, 'Computer simulation study of primary and secondary anode loading in magnetrons', *IEEE Transactions on Electron Devices,* vol. 38, pp. 2234–2238, 1991.

[56] H. L. McDowell, 'Magnetron simulations using a moving wavelength computer code', *IEEE Transactions on Plasma Science,* vol. 26, pp. 733–754, 1998.

[57] H. L. McDowell, 'Magnetron simulations using a multiple wavelength computer code', *IEEE Transactions on Plasma Science,* vol. 32, pp. 1160–1170, 2004.

[58] H. W. Welch and W. G. Dow, 'Analysis of synchronous conditions in the cylindrical magnetron space charge', *Journal of Applied Physics,* vol. 22, pp. 433–438, 1951.

[59] J. F. Hull, 'Crossed-field electron interaction of the distributed-emission space-charge-limited type', *IRE Transactions on Electron Devices,* vol. 8, pp. 309–323, 1961.

[60] S. Riyopoulos, 'Magnetron theory', *Physics of Plasmas,* vol. 3, pp. 1137–1161, 1996.

[61] S. Riyopoulos, 'New improved formulas for magnetron characteristic curves', *IEEE Transactions on Plasma Science,* vol. 26, pp. 755–766, 1998.

[62] E.-Q. Zhang, 'On the magnetron cathode', *IEEE Transactions on Electron Devices,* vol. 33, pp. 1383–1384, 1986.

[63] J. Feinstein, 'Planar magnetron theory and applications', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, 1961.

[64] A. M. Clogston, 'Principles of design', in G. B. Collins, ed., *Microwave Magnetrons*. New York: McGraw-Hill, pp. 401–459, 1948.

[65] J. R. G. Twisleton, 'Twenty-kilowatt 890 Mc/s continuous-wave magnetron', *Proceedings of the Institution of Electrical Engineers,* vol. 111, pp. 51–56, 1964.

[66] C. Shibata *et al.*, 'High-power (500 kW) c.w. magnetron for industrial heating', *Electrical Engineering in Japan,* vol. 111, pp. 94–100, 1991.

[67] A. M. Cripps and P. A. Jerram, 'X-band linear accelerator magnetrons', in *19th European Microwave Conference*, pp. 1086–1090, 1989.

[68] G. R. Feaster, 'The cathode', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 1. New York: Academic Press, pp. 113–140, 1961.

[69] J. M. Osepchuk, 'Private communication', 2016.

16   Crossed-Field Amplifiers

16.1   Introduction

Broad-band crossed-field amplifiers (CFAs), based on electron flow in crossed fields, are analogous to travelling-wave tubes. In essence, a CFA is a magnetron in which the slow-wave structure is not a closed loop but has been broken so that the two ends can be connected to external waveguides. CFAs share with the magnetron the advantages of small size, low weight, and high efficiency. However they have the disadvantages of low gain, high noise levels and non-linearity compared with TWTs. The applications of CFAs are mainly in military radar systems. For this reason the available literature is sparse.

In theory many possible types of crossed-field tube exist, but not all of these have been found to be practical devices. The different types can be classified by considering the options for the tube geometry and for the type of beam and slow-wave structure used. The slow-wave structure fulfils the same role as the anode in a magnetron[1]. The geometry may be either linear or circular, and the electron beam may be emitted from the whole area of the cathode, as in a magnetron, or injected from an electron gun (see Figure 8.1). Cylindrical geometry has the advantage of compactness and better use of the permanent magnet material. Planar geometry offers the possibility of greater interaction length and gain, but it has only been used in experimental tubes [1–4]. Figure 16.1 shows the arrangements of circular emitting cathode, and injected beam, CFAs.

The slow-wave structure, like that in a TWT, can have either forward-wave or backward-wave characteristics, as shown in Figure 16.2. With a forward-wave structure the interaction takes place instantaneously over a band of frequencies. With a backward-wave structure the interaction is narrow band, and can be tuned over a range of frequencies by changing the electron drift velocity. It should be noted that the drift velocity corresponds to the voltage on the surface of the electron hub, which is normally around an order of magnitude less than the anode voltage. Because the slow-wave structure is periodic it has forward and backward-wave characteristics in successive Brillouin zones. Thus, the same tube can be operated as either a forward-wave or a backward-wave amplifier by changing the anode voltage, and the magnetic field, to achieve synchronism with the appropriate

---

[1] In the older literature the cathode and the slow-wave structure are called the *sole*, and the *line*, respectively.
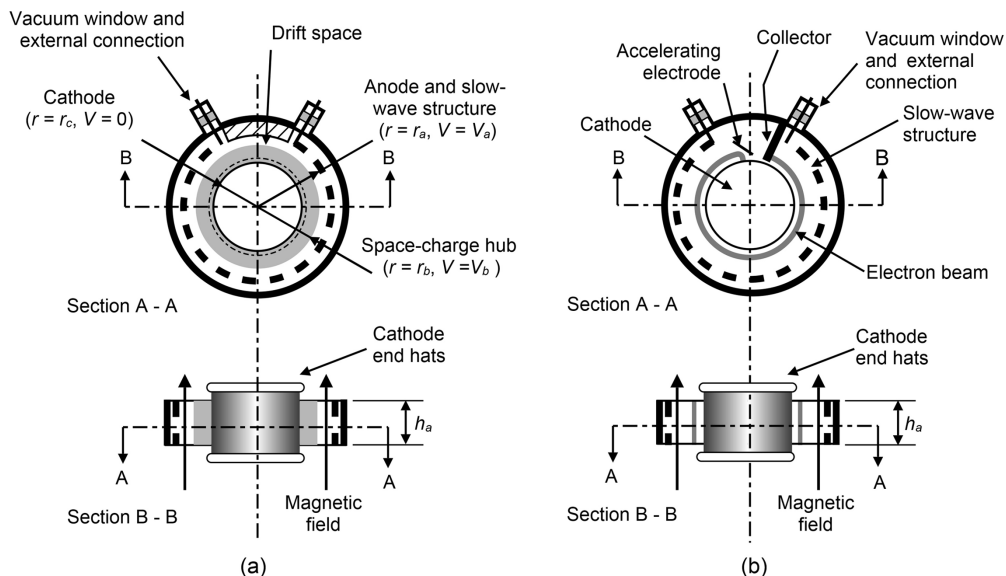
Figure 16.1: Arrangement of: (a) emitting cathode CFA, and (b) injected beam crossed-field amplifiers
(copyright 1973, IEEE, reproduced, with permission, from [5]).
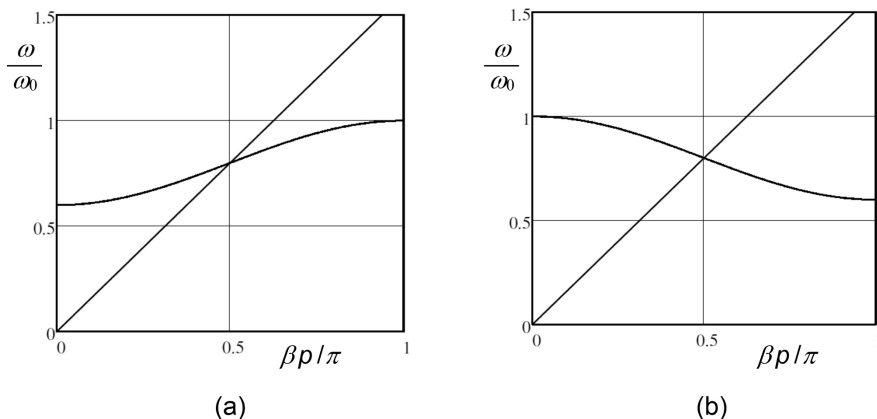
Figure 16.2: CFA dispersion diagrams, normalised to the upper cut-off frequency, showing the electron velocity line intersecting at $\pi/2$ phase shift per cell: (a) forward wave, and (b) backward wave.

space-harmonic [6]. The circular geometry means that care must be taken to avoid positive feedback through RF coupling between the ends of the slow-wave structure, by allowing sufficient separation between them.

The interaction in a CFA resembles that in a magnetron. The electrons retarded by the RF electric field move outwards towards the slow-wave structure, while those accelerated move inwards towards the cathode. The resemblance is especially close

in an emitting cathode CFA where spokes of charge are formed, and all the electrons are collected on either the slow-wave structure or the cathode.

16.1.1 Emitting Cathode CFAs

All crossed-field amplifiers which are of continuing importance are of the circular, re-entrant, emitting cathode type and the abbreviation CFA is now generally taken to refer to them [5, 7–11]. The detailed discussions, in sections from 16.2 onwards, refer exclusively to tubes of this kind. The slow wave structure can be either forward- or backward-wave, but both have the arrangement shown in Figure 16.1(a), and it is convenient to discuss them together. The input and output connections are made according to the structure type and the direction of rotation of the electrons. A drift region may be included to control the feedback caused by the use of a re-entrant electron beam. In backward-wave tubes the drift region is sometimes omitted, though the result is to produce frequencies within the band at which the output power is low. Backward-wave emitting-cathode CFAs are sometimes known by the proprietary name 'Amplitron'. The electron beam is demodulated to some extent as it passes through the drift space but some modulation is carried forward to the start of the interaction.

Emitting cathode CFAs typically provide average powers ranging from 400 kW at 2.8 GHz to 100 W at 30 GHz, with efficiency in the range 35–60% and bandwidth 10–15% [12]. The mean output power is limited by heat dissipation in the slow-wave structure. However, peak powers from hundreds of kilowatts to megawatts, with duty cycles around 0.001, may be obtained with good efficiencies. The gain is low (8–15 dB) because of the limited length of the slow-wave structure, and the need to lock out noise-generating mechanisms. Where high gain is required a chain of tubes is used. Comparisons with systems employing TWTs should be made with whole systems, and not just the power output tubes. The final CFA of a chain can be made 'transparent' so that the system has both high and low output power capabilities. Low power output is obtained by leaving the output tube switched off so that the power of the penultimate tube passes through it with little attenuation. The CFA is particularly suited for use as the final amplifier for mobile radar applications because it is smaller, lighter, more efficient, and operates at a lower voltage, than an equivalent coupled-cavity TWT. CFAs are non-linear amplifiers at all power levels. They are therefore unsuited to multi-carrier operation because of the generation of intermodulation products. In radar applications the bandwidth can be used to achieve frequency agility. For this purpose either forward- or backward-wave interactions may be used though the latter requires a more complicated modulator to adjust the anode voltage for each frequency.

When making comparisons between the efficiencies of different tube types it is important to be careful about the way in which the efficiency is defined. In linear beam tubes the heater power, and any solenoid power, need to be included. In CFAs the low gain means that the RF input power is important. The best approach is to compare the efficiencies of complete amplifier chains having the same overall gain
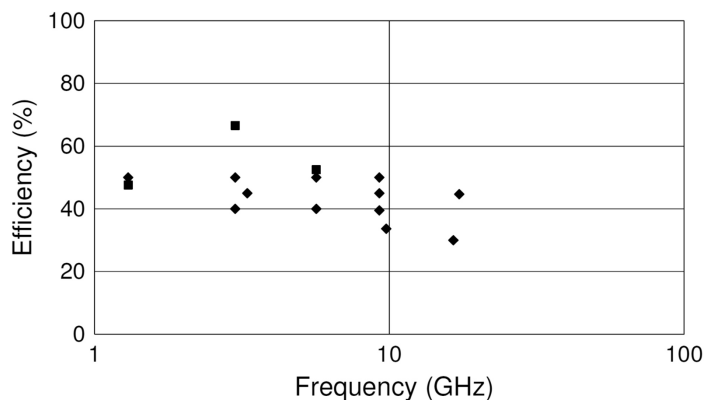
Figure 16.3: Efficiencies of emitting cathode CFAs.
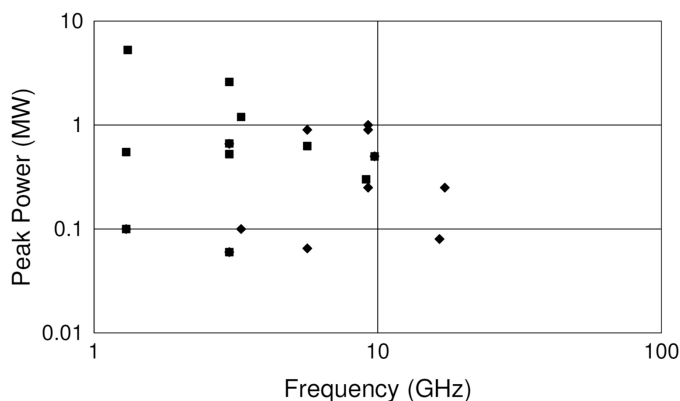
Figure 16.4: Peak power output of emitting cathode CFAs.

and output power. Figure 16.3 shows the efficiencies of emitting cathode CFAs, as far as they can be ascertained from published data. For this purpose the efficiency has been defined as the ratio of the RF output power to the DC input power. The state of the art of the peak power of emitting cathode CFAs is shown in Figure 16.4.

The signal to noise ratio of a CFA is typically −45 to −55 dBc/MHz, which is comparable with that of a TWT [12]. This noise is thought to arise from turbulence within the space-charge hub. It has been reported that the noise level can be reduced by shaping the magnetic field, or by incorporating a slow-wave structure in the cathode (see Section 16.2.3), but no modern CFAs use these techniques [13].

An emitting-cathode CFA in which the electrons are collected at the end of the interaction region, in either a linear or a circular tube, is known by the proprietary name DEMATRON [1, 14]. Although such an amplifier has theoretical attractions, it has been found that gain is lower than expected because of the variation in the electron velocity caused by the space-charge in the spokes. This can be compensated

for by tapering the DC electric field, the magnetic field, or the phase velocity of the slow-wave structure. It has not been found possible to build practical non-re-entrant CFAs of this kind whose performance is competitive with other types of tube.

16.1.2   Injected Beam CFAs

Injected-beam CFAs have been found to be inferior to linear-beam tubes in terms of power and bandwidth, and they are now chiefly of historical interest [7]. Figure 16.1(b) shows the arrangement of an injected beam CFA. Forward-wave amplifiers and backward-wave oscillators have essentially the same construction. The beam is in the form of a flat strip generated by a Kino gun (Figure 9.20). The electron velocity can be controlled by the accelerator voltage, and by the potential difference between the slow-wave structure and the cathode. As a result of the RF interaction some of the electrons are collected on the slow-wave structure. The remainder are collected on a collector electrode which may be depressed to increase the overall efficiency of the tube. It is important to ensure that secondary electrons are not emitted from the cathode. This is achieved by the use of grooves, or low secondary emission coatings. The geometry of the interaction region must be controlled very accurately for satisfactory operation.

The crossed-field backward-wave oscillator (M-type Carcinotron) resembles the injected beam backward-wave amplifier shown in Figure 16.1(b). The input port is replaced by a matched termination inside the vacuum envelope [15–20]. The attenuator provides a good RF match at the end of the slow-wave structure remote from the output. Tubes of this kind have been used in electronic counter-measures systems with CW output powers up to a kilowatt and efficiencies around 50%. Peak powers up to 230 kW at S-band with 50% efficiency, and 50 kW at X-band with 15% efficiency, and tuning ranges around 15% have been reported [17].

Circular injected beam crossed-field amplifiers can provide average powers of 1 kW and pulsed power of 5 kW at X-band with octave bandwidth and 20 dB gain [21]. The feasibility of generating 10 MW peak power at L-band has been demonstrated, but the efficiency of injected beam CFAs is not high enough to make them competitive with tubes of other types [5]. The theory of injected beam CFAs is discussed in [22–25]. They can have high noise output, and considerable research has been undertaken to find ways of reducing this [8]. Linear injected-beam CFAs have been used for experimental purposes [2, 3, 26, 27].

16.2   CFA Construction

The review of emitting-cathode CFAs by Skowron [5] remains the best available introduction to the subject, and much of this chapter is based on that paper. Further information is given in [28]. The close resemblance of a CFA to a magnetron means that the principles of operation and the characteristics of the two types of device are very similar. The electron beam is assumed to be drawn from a

space-charge hub where the electrons on the surface of the hub have a tangential velocity approximately equal to the phase velocity of the wave on the slow-wave structure.

16.2.1   Slow-Wave Structures

The slow-wave structure of a CFA, like the anode of a magnetron, must combine suitable RF properties with the ability to dissipate the heat of the electrons intercepted by it. The heat dissipation is not uniform, as in the magnetron, but is greater close to the output. It has been found by experimental observation that a few vanes close to, but not at, the output receive power around twice the average. Allowing for variation in the axial direction, a 3:1 variation in power density may occur. The structure is normally designed so that the phase shift is around $\pi/2$ per cell. The pitch of the structure is then approximately half that of a comparable magnetron. The first practical tubes used a strapped-vane structure in which the input and output waveguides were connected to the straps [6]. This structure has a backward-wave characteristic when the phase shift per cell is around $\pi/2$ (see Figure 15.19).

Figure 16.5 shows some typical structures used in emitting-cathode CFAs. Forward-wave structures having good bandwidth, and flat dispersion, are often based on the helix or the meander line [5]. Examples of helix-derived structures are the helix-coupled vane structure (Figure 16.6(a) and (b)) and the stub-mounted helix. In the latter a flattened helix is supported by quarter-wave stubs whose loading effects are small. Ceramic-supported meander lines are suitable for use at higher frequencies [5, 21]. Backward-wave structures include the strapped ladder line (Figure 16.6(d)), the inter-digital line (Figure 16.6(c)), and the split folded waveguide [14, 29]. Strapped-vane, ring-bar and dielectric-supported helix structures are described in [30].

16.2.2   Cathodes

CFAs designed for peak power of less than about 10 kW employ thermionic cathodes. Back-bombardment liberates secondary electrons so that the total current is a combination of both primary and secondary electrons in a proportion which is self-adjusting. The emission is space-charge limited, as in the magnetron. At higher power levels cold cathodes may be used so that all the electrons are produced by secondary emission. This is possible because the high RF electric fields produced by the input signal are sufficient to initiate emission using seed electrons produced by field emission, or by ionisation of the residual gas in the tube. The process requires the simultaneous application of the DC voltage and the RF pulse and the emission builds up in less than 10 ns. The current obtained from a secondary emitting cathode is constant during the pulse, whereas that of a primary emitter may decay during the pulse.

Forward wave structures   Backward wave structures

(a) Helix-coupled vane   (c) Inter-digital line

(b) Double helix-coupled vane   (d) Strapped ladder line
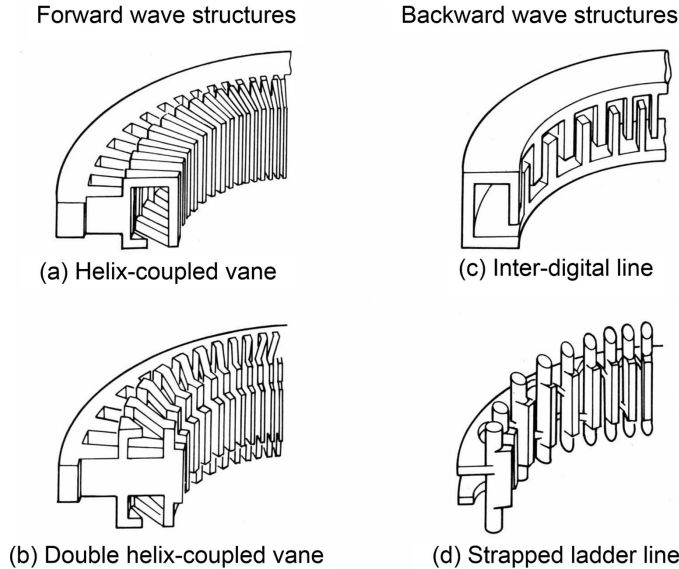
Figure 16.5: Slow-wave structures for emitting-cathode CFAs: forward wave (a) helix-coupled vane and (b) double helix-coupled vane; backward wave (c) inter-digital line and (d) strapped ladder line (copyright Thales; courtesy of Thales).

The cathode in a cold-cathode tube may be either an oxide cathode or a pure metal. The oxides of elements such as barium, aluminium, thorium, magnesium, and beryllium have secondary electron emission coefficients greater than 2.5. Pure metals have coefficients below 1.9, with platinum having the highest value at 1.85. Pure metal cathodes do not exhibit any wear-out mechanism, apart from erosion, so that very long tube lifetimes are possible. Oxide cathodes wear out because of depletion of the oxide caused by back-bombardment. However, it is possible to replenish it by, for example, including an oxygen dispenser within the vacuum envelope [31]. CFAs with oxide cathodes can have lives greater than 40,000 hours.

16.2.3   Cathode-Driven CFAs

One of the chief disadvantages of the CFA is its low gain which results from the fall-off of the RF electric field between the slow-wave structure and the surface of the electron hub. The gain can be increased appreciably by incorporating a slow-wave structure in the surface of the cathode and applying the input signal to it. This structure is much closer to the surface of the hub and the gain can be increased by 10 to 20 dB. The power output of such a tube is comparable with that of a conventional CFA and the signal to noise ratio is reportedly improved by at least 20 dB [32]. If the RF input is applied to the anode slow-wave structure of a tube of this kind the signal to noise ratio is improved by around 10 dB compared with a conventional CFA. The design of a millimetre wave high gain CFA is described in [33].

16.3   Basic Principles

The resemblance of a CFA to a magnetron means that much of the theory of magnetrons can be used, or adapted, to understand them. For the present we shall assume that the electrons are drawn from a uniform electron hub described by the equations in Section 15.2.1. We shall see later that detailed numerical modelling shows that this model is not strictly correct for cold cathode tubes because the secondary electron emission from the cathode is non-uniform.

Because the electron flow in a CFA is re-entrant the angular velocity of the spokes must be given by

$$\omega_s = \frac{\omega}{M}, \tag{16.1}$$

where ω is the frequency of operation and M is the number of spokes. Since the velocity ω_s is determined by the anode voltage it follows that a CFA may be expected to exhibit a number of modes of operation corresponding to different values of M. There is evidence from both experiment and simulation that (16.1) is satisfied for synchronous operation of both forward- and backward-wave CFA's [34–39]. It has been suggested that all the spokes are then identical [19] and the limited evidence available from the papers cited supports this, even in the presence of a drift region. Simulations using a moving wavelength code have shown that debunching effects in the drift region are not sufficient to eliminate the spokes completely.

When the operating point is moved away from synchronism, by changing either the frequency or the anode voltage, the residual spokes emerging from the drift region are no longer in phase with the RF voltage on the slow-wave structure as they re-enter it. Figure 16.6 shows the results of simulation of a forward-wave CFA when the operating voltage is slightly below that for synchronism. It can be seen that the spokes are only well-formed for the second half of the interaction region. Calculations based on the angles of these spokes show that M in (16.1) is no longer an integer. Thus the view expressed in early work [6, 34] that M remains an integer, even when the operation is not synchronous, is seen to be incorrect. When the operation of the CFA is less efficient, the spokes may be erratic and less well defined [37, 40].

If a uniform periodic slow-wave structure is assumed then each section subtends an angle $2\pi/N_a$ on the axis where N_a is an integer. The phase shift per section is θ for synchronous operation, and the structure may have either forward- or backward-wave characteristics. In a tube without a drift region N_a is the number of cells in the structure. In general, the number of active cells is reduced by those which are replaced by the drift region. At synchronism the angular velocity of the wave on the structure is

$$\omega_s = \frac{2\pi\omega}{N_a\theta}. \tag{16.2}$$
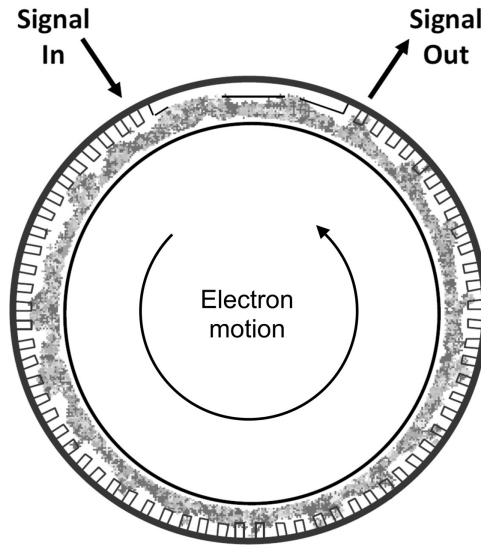
Figure 16.6: Spoke formation in a typical forward wave CFA (courtesy of Dr D.P. Chernin [41]).

Equating (16.1) and (16.2) gives

$$\theta = 2\pi \frac{M}{N_a}. \tag{16.3}$$

Note that this differs from the expression given in [6] because of a difference of π in the definition of θ. As a consequence of (16.3) the possible synchronous values of θ are fixed by the values of M and N_a, as shown in Figure 16.7, for a structure with $N_a = 11$ and a backward-wave characteristic in the first Brillouin zone. The synchronous spoke velocity line is shown for $M = 4$. It is evident from Figure 16.7 that the choice of an odd integer for N_a means that no interaction is possible at the π mode cut-off of the structure where the interaction impedance is high, and the tube would be likely to oscillate. The threshold voltage at synchronism in a CFA is given by (15.39) where ω_s is determined from (16.1). Now suppose that the tube whose dispersion diagram is shown in Figure 16.7 is operating in the $M = 4$ mode. If the anode voltage is increased then, at some point, it will exceed the threshold voltage for the $M = 3$ mode. The operation of the tube may then move discontinuously to the new mode [35]. Since the anode current increases with increasing anode voltage, in the same way as a magnetron, it may be limited by the onset of a different mode [5]. The difference between the velocity of the spokes and the phase velocity on the slow-wave structure for non-synchronous operation means that the synchronous relation (16.3) is not strictly enforced. Some phase slippage is permissible at the input end of the structure [6].
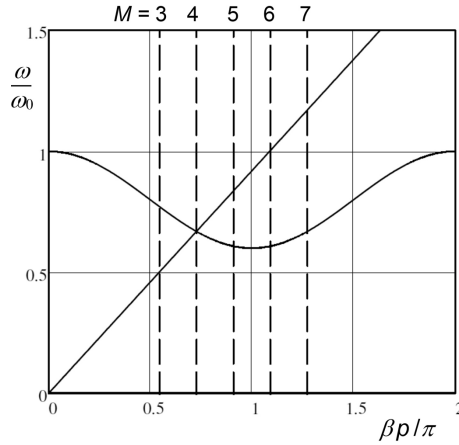
Figure 16.7: Dispersion diagram of a backward-wave CFA $(N_a = 11)$ showing synchronous points corresponding to different numbers of spokes.

Although the interactions in forward-wave and backward-wave tubes are similar there are some differences in detail. Consideration of Figure 16.2 shows that, if two tubes have slow-wave structures with the same cold bandwidth, then the change in the synchronous electron velocity for a given change in frequency is greater for the backward-wave tube. The anode voltage increases with increasing frequency in a backward-wave CFA, whereas the opposite is true for a forward-wave tube. The properties of each spoke are chiefly determined by the RF field on the slow-wave structure at the tip of the spoke. Thus the spokes are most strongly formed close to the output of the tube. This means that in a forward-wave tube the charges carried by the spokes entering the drift region, and the debunching in the drift region, are both greater than in an equivalent backward-wave tube. This difference can be seen in figures 7 and 3 in [37]. A further difference between the two types of tube arises because of the effects of the transit time of the electrons from the surface of the hub to the anode. The current flowing in a spoke is determined by the RF field experienced by electrons at its base. At high signal levels, close to the tube output, the transit time is typically small compared with the synchronous period. The current in a spoke is then determined by the strength of the local RF field. However, at low signal levels, near to the tube input, the transit time is typically greater than the synchronous period. Thus the current reaching the anode is determined by the fields one or more synchronous periods away [42]. In a forward-wave amplifier the spokes are moving in the direction of increasing RF field strength and the currents in the spokes close to the input are smaller than would be determined by the RF field at the spoke tip. The opposite is true in a backward-wave amplifier.

16.4   CFA Characteristics

16.4.1   Performance Chart

The performance chart of an emitting cathode CFA operated at a fixed frequency and RF drive level is like that of a magnetron (see Figure 16.8). For a given magnetic field no current flows until the anode voltage exceeds the threshold voltage. Thereafter the current increases, the anode voltage remains approximately constant, and the output power increases with the current. The magnetic field is normally provided by a permanent magnet and is therefore fixed by the manufacturer. The operating curve varies slightly with frequency as shown in Figure 16.9. At constant current the voltage increases with frequency in a backward-wave CFA and decreases with frequency in a forward-wave tube, as discussed in the previous section. Figure 16.10 shows how the performance of a backward-wave amplifier tube depends upon the load line of the modulator for constant voltage, constant current, and constant impedance operation. Constant voltage operation produces an unacceptably large variation in power output and efficiency across the frequency band of the tube. Constant current operation gives essentially constant output power. These two conditions are approximated by hard tube and line type modulators respectively (see Section 20.3) [5]. The change in the operating curve with frequency is smaller in a forward-wave amplifier and constant voltage operation is possible.
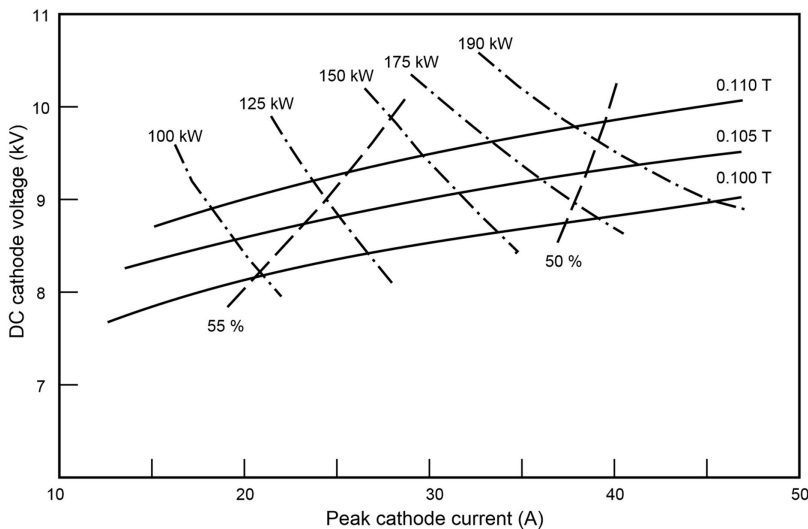
Figure 16.8: Typical performance chart of a CFA at 1190 MHz with 5 kW peak input power (copyright 1967, SFD, reproduced, with the permission of CPI, from [28]).
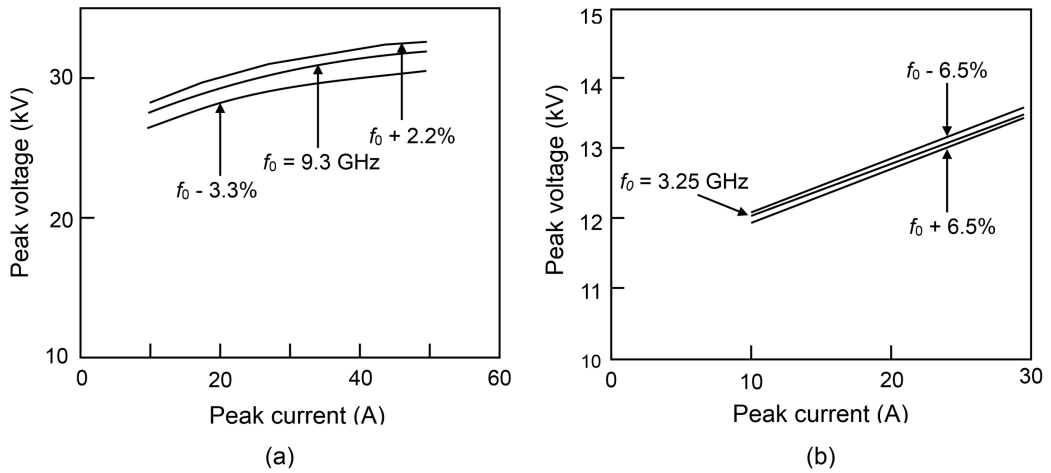
Figure 16.9: Variation in the operating curve of a CFA with frequency: (a) backward-wave amplifier, and (b) forward-wave amplifier
(copyright 1967, SFD. Reproduced, with the permission of CPI, from [28]).
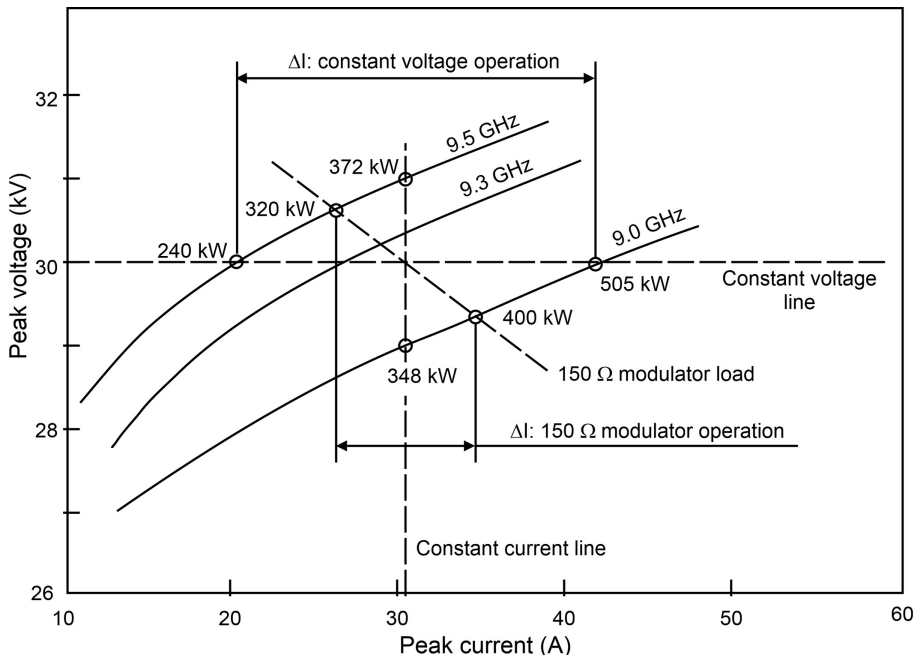
Figure 16.10: Backward-wave CFA characteristics showing the effect of different modulator load lines
(copyright 1967, SFD, reproduced, with the permission of CPI, from [28]).

16.4.2   Modulation

Figure 16.11 shows typical pulses for operation of a pulsed CFA. If a cold cathode tube is operated from a line type modulator the RF power must be applied before the voltage pulse. This ensures that current is drawn from the modulator as soon as the voltage is applied to the tube. Otherwise, the tube would present an open circuit to the modulator, resulting in a doubling of the voltage, leading to internal arcing and possible damage to the tube. If the tube has a hot cathode with appreciable thermionic emission, the application of the voltage pulse before the RF pulse leads to noise generation from the uncontrolled space-charge cloud. As a rule of thumb, the voltage pulse is designed to reach 50% of its peak value when the RF input power is 90% of the full drive level. A similar situation arises with any type of CFA if the RF input is removed before the end of the voltage pulse. It is desirable for the rise-and fall-times of the voltage pulse to be as short as possible to reduce the risk of oscillations at band edges, or in other modes where the gain is high.

Some CFAs are designed to be operated with cold cathodes and with DC voltages applied to the electrodes. The current pulse is then triggered by the RF pulse. When the RF pulse is removed the electrons continue to circulate and would, if unchecked, lead to the generation of noise. For this reason a control electrode is incorporated, as shown in Figure 16.12. At the end of the RF pulse a positive pulse is applied to the control electrode to remove the circulating electrons. Figure 16.13
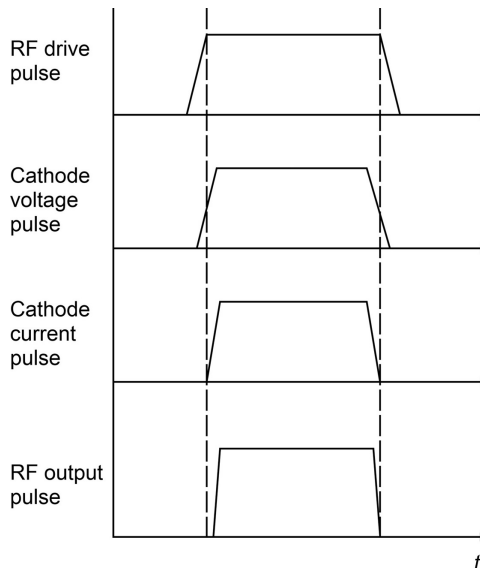
Figure 16.11: Typical pulse nesting for a cold cathode CFA (copyright 1967, SFD, reproduced, with the permission of CPI, from [28]).
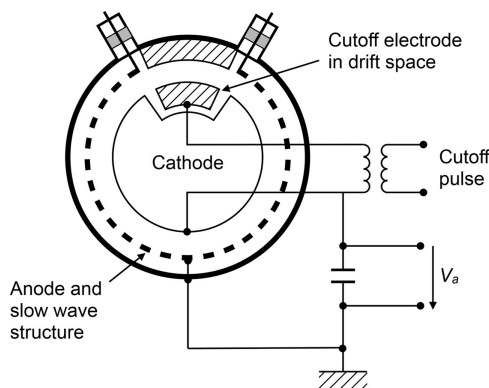
Figure 16.12: Arrangement of a CFA with a control electrode
(copyright 1973, IEEE. Reproduced, with permission, from [5]).
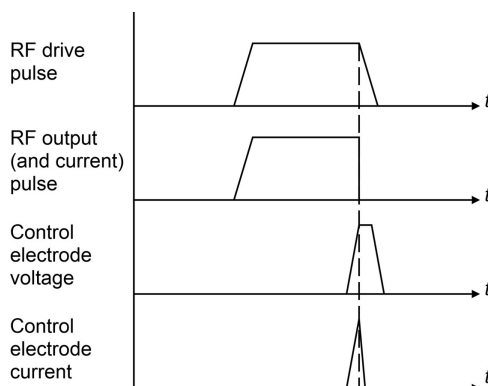
Figure 16.13: Typical pulse nesting for a CFA with a DC potential between the cathode and the anode
(copyright 1967, SFD. Reproduced, with the permission of CPI, from [28]).

shows the pulses for a cold-cathode CFA with a DC voltage applied to the cathode. Electron emission commences within a few nanoseconds of the start of the RF pulse. The proportion of electrons collected by the control electrode varies from zero, when it is unbiased, to 100% with sufficient positive bias. Thus, it is possible to find an intermediate level of DC bias where sufficient electrons continue to circulate for the tube to operate in the presence of RF drive but not when it is absent. It has been shown that RF keying can be achieved in this manner [5].

16.4.3   Transfer Characteristics

The amplitude transfer characteristics of a CFA may be displayed on either linear or logarithmic scales, as shown in Figure 16.14. It can be seen that the tube does not have a linear region of operation, nor does it exhibit saturation. The low gain
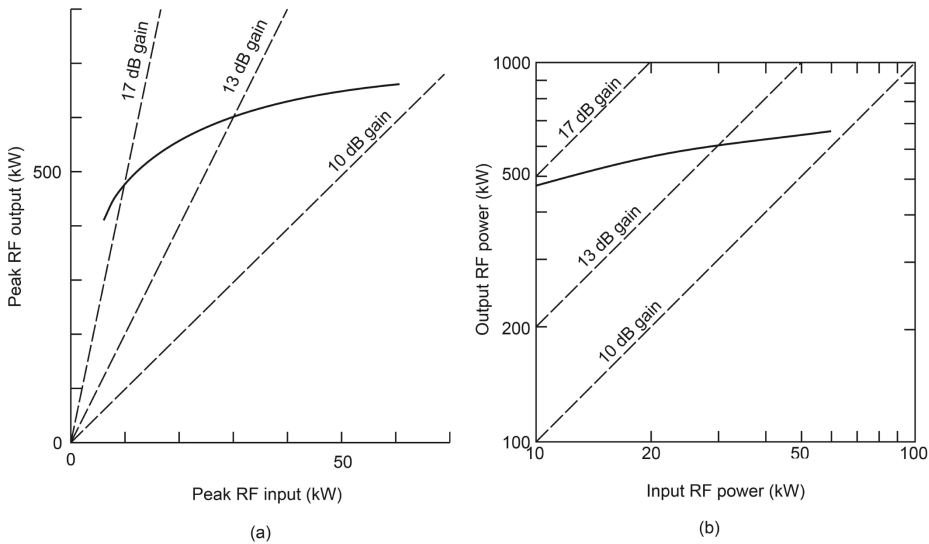
Figure 16.14: Typical CFA amplitude transfer curves: (a) linear scales, and (b) logarithmic scales
(copyright 1967, SFD, reproduced, with the permission of CPI, from [28]).

of the tube means that the input RF power is a significant fraction of the total power input to the tube. Thus the RF input power is added to the (approximately constant) RF power generated by the interaction within the tube. If the RF drive of a tube with a primary emitting cathode is increased from low levels it is found that, initially, the output at the drive frequency is proportional to the input and accompanied by noise. At higher drive levels the input signal suppresses the noise and the output is determined by the DC operating conditions. For this reason it is best to regard the CFA as a locked amplifier which is always in a non-linear condition. The minimum drive level for stable operation may be indicated on the AM/AM curves.

The bandwidth of a CFA is determined by the range of frequencies over which stable operation is possible. It may also be specified as the range of frequencies for which the output is constant within some limits (e.g. ± 0.5 dB). Because, as we have seen, the output power is affected by the load line of the modulator, it is important to specify the conditions under which the bandwidth is defined. For a given frequency, and given DC and RF input powers, it is found that a tube can operate over a range of currents. The upper limit is set by the maximum current available from the cathode, the onset of a competing oscillation, or the inability of the RF input power to suppress noise. The lower limit is set by the start of oscillation in a mode with a lower threshold voltage [5].

The phase of the output of a CFA is determined by the DC voltage and current, the RF drive level, and the variation of the phase length of the tube with frequency. Figure 16.15 illustrates the dependence on the anode voltage, and the RF drive level, of the phase of the output of an X-band backward-wave CFA, relative to the
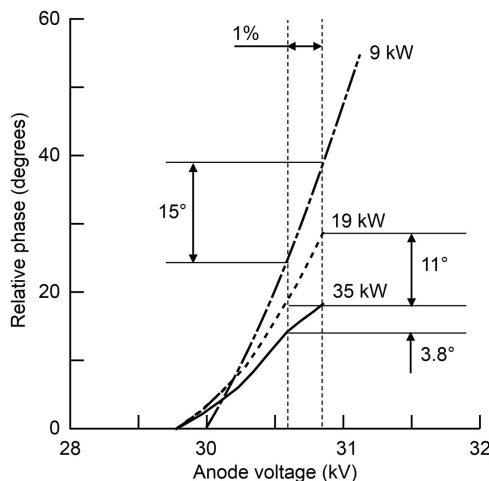
Figure 16.15: Dependence of the relative phase of the output of a CFA on the anode voltage and the RF drive leve l
(copyright 1967, SFD, reproduced, with the permission of CPI, from [28]).

input. The figure shows typical phase variation for a 1% variation in anode voltage, and three different drive levels. The phase sensitivity to voltage is proportional to the gain of the tube. It is, therefore, least at high drive levels where the gain is least (see Figure 16.14). The sensitivity of the phase of a CFA to changes in anode voltage is about an order of magnitude less than that of a TWT. The variation of phase with drive power at constant voltage is usually measured with small changes in the input power around the normal operating point. It is typically of the order of 1 or 2 degrees/dB.

The phase length of a CFA is 1 to 10 wavelengths for a backward-wave tube, and 15–20 wavelengths for a forward-wave tube. These are less than those of a linear-beam tube, and the resulting variation of phase with frequency typically less than ±5°. The phase length of a CFA is affected by the reactive loading of the electrons (compare frequency pushing in a magnetron). The hot phase length is around 20–40° greater then the cold phase length, and decreases with increasing anode current. Because the CFA is a transparent amplifier, the phase variation in practical cases is strongly affected by the RF match presented to the input and output of the tube. If the matches of the source and the load are perfect then the deviations from linearity are small, and determined by mismatches within the tube. The phase variation caused by external mismatches can be estimated from the circuit in Figure 14.1 as

$$\Delta\phi = \rho_S \rho_L A L \qquad (16.4)$$

where ρ_S and ρ_L are the input and output voltage reflection coefficients combining internal and external mismatches, and A and L are the amplification and cold loss of the tube expressed as voltage ratios.

16.4.4   Signal Growth and Anode Dissipation

Additional information about the properties of CFAs is provided by a detailed experimental study of an S-band forward-wave tube [43]. Experimental versions of this tube were fitted with RF probes and thermal sensors so that the RF power, and the vane temperatures, could be measured for a range of operating conditions. Typically, the temperature of the slow-wave structure varied approximately linearly from 20–30 °C, at the input, to 70–80 °C, at the middle of the interaction region. It then remained approximately constant until the output. The RF power on the slow-wave structure increased slowly over the first 20–30% of the length of the interaction region. Thereafter it increased approximately linearly until the output. It was concluded that the initial growth in the RF power was based on the recirculating space-charge above the electron hub, and that the linear growth made use of current extracted directly from the hub.

16.5   Theoretical Models of CFAs

16.5.1   PIC Codes

CFAs can be modelled by PIC codes in much the same way as magnetrons and, in some cases, the same code has been used for both types of tube. The properties of the slow-wave structure are usually modelled by an equivalent network. The accurate modelling of secondary electron emission is essential for cold-cathode tubes. When the whole tube is modelled large computing resources and long CPU times are required [37–40, 44–46]. Moving window codes model the interaction with a single spoke as it develops through time [47–53]. Where electrons move out of the moving wavelength they are re-injected at the opposite boundary, as in the TWT simulations described by Vaughan [54]. This assumption is not strictly valid because the electron motion is not periodic in space. Additional guard wavelengths to overcome this problem do not appear to have been used in CFA models. The computer codes commonly employ rectangular Cartesian coordinates with a correction factor to account for the curvature of the interaction space [48]. In these codes it is necessary to track the interaction over multiple passes through the tube. It is found that the computations do not converge to a stable solution but show pass-to-pass fluctuations which may be an indication of the level of noise generation within the tube [49]. Stable results are obtained by taking the average over several passes, and the main features of the tube performance can be modelled with good accuracy. The models also show the details of spoke formation, including the debunching in the drift region. When secondary emission is the dominant source of electrons it is found that regions of low density propagate through the space-charge hub. The method has been extended to tubes with a slow-wave structure on the cathode [51]. Whole tube simulations have been used to give additional information about spoke formation in CFAs, and to study the

conditions for low noise operation that have been observed experimentally [38, 44, 45].

16.5.2 Soliton Theory

We have seen that considerable progress can be made in understanding the properties of linear-beam tubes by using models in which the non-linear equations describing the interaction are linearised by the small-signal approximation. This approach has been extended into the non-linear region of operation by the addition of more terms in the expansions. An alternative approach is to find non-linear solutions to weakly non-linear equations using soliton theory [55, 56]. A soliton is a solitary wave that maintains its shape while it propagates at a constant velocity. Solitons occur when the non-linearity, and dispersion, of a medium counteract each other to prevent the break-up of the wave. The application of soliton theory to CFAs shows that growing solitons can only exist for a limited range of anode voltages. For a particular tube this range was found to be from the threshold voltage (V_T) to $1.13\,V_T$. This behaviour was found to be in good agreement with experimental results. At the upper limit the amplification ends abruptly. It was shown that the true synchronous voltage is slightly greater than the threshold voltage because of differences between the electron motion and the idealised Brillouin hub. However, the difference between the synchronous voltage and the threshold voltage was small in the typical operating range of the tube studied. The model was also able to predict the dependence of the phase shift on anode voltage [56].

16.5.3 Guiding Centre Theory

The guiding centre theory, whose application to magnetrons is described in Section 15.6.3, was applied to modelling CFAs in [42, 57, 58]. The electrons were assumed to be drawn from a uniform hub whose surface was taken to be approximately that of the theoretical Brillouin hub. The electron density in the hub was not constant but was found by considering the balance between secondary emission of electrons from the cathode and their motion out of the hub into the spokes. Thus the electron density and the spoke current varied according to the local RF field. The model included the effect of the finite transit time of the electrons from the hub to the anode. The evolution of the spokes with time was computed by successive passes through the interaction region. The electron density at the start of a pass was taken to be a fixed fraction of that at the end of the previous pass. The electron stream was assumed to be unmodulated at the start of each pass. The results converged after only a few iterations and the Gauss lines which were computed showed some agreement with experimental results, including their dependence on the frequency. It is not clear whether the parameters of the model were adjusted to achieve agreement with measurements. The model also showed the spatial growth of the electron density in the hub and of the RF voltage on the slow-wave structure. The electron density in the hub was typically about 0.8 of the theoretical hub density. When the

secondary electron emission coefficient of the cathode was increased above a certain level it was found that the output power did not settle to a steady value but oscillated from pass to pass indicating the generation of noise [59].

16.5.4   Non-Linear Fluid Mechanics

The application of the methods of fluid mechanics to the modelling of crossed-field devices has been described in a large number of papers including [60–63]. In this approach the non-linear, cold-fluid, equations have been expanded into static and first-order time-varying parts, corresponding to the electron hub and spokes assumed in the guiding centre model. It has been shown that the two parts of the solution are not independent of one another, as in simpler models, but that the evolution of the spokes has an effect on the static electron density distribution. In particular, the electron density at the position of the theoretical hub surface is around half of that at the cathode, when space-charge limited emission is assumed. Although this work provides some insights into the behaviour of CFAs it does not appear to be useful for engineering modelling of their performance.

16.5.5   Rigid Spoke Model

We have seen that the rigid spoke model of the magnetron gives good agreement with experimental results, despite problems with the validity of some of the assumptions on which it is based. It is therefore of interest to consider whether a similar model can be constructed for CFAs. A simple model of this kind, described in [19], assumed that fully formed spokes were present throughout the interaction region, and that the current induced in the slow-wave structure was nearly constant. The induced current was divided equally between forward and backward travelling waves. The difference in the phase relationships means that the induced currents add constructively in one direction, and cancel in the other. Since each spoke induces the same power in the circuit it follows that the signal growth with distance is linear. Experimental measurements have shown that this is true over most of the interaction region by (see Section 16.4.4). Increasing the gain of a CFA by a factor of 10 therefore requires a tenfold increase in the length of the interaction region. In contrast the exponential growth in a TWT only requires the length of the interaction region to be doubled. The gain of a CFA can be estimated by assuming that the power induced in unit length of the slow-wave structure is equal to that of a magnetron having the same parameters and output power. The input power can then be calculated from the length of the structure, making allowance for the cold loss. The efficiency of the tube is found from the total power delivered to the structure. The application of this approach to an experimental tube gave good agreement for the efficiency. Further information on the parameters and performance of that tube can be found in [64–66].

A straightforward improvement to the model described above uses a method developed for modelling non-reentrant CFAs [14]. The assumption that all spokes

deliver equal power to the slow-wave structure is replaced by one in which the power delivered to each section is calculated from the properties of a magnetron equivalent to that section. The rigid-spoke magnetron model described in the previous chapter can be used for this purpose. The synchronous angular velocity and the number of spokes are calculated from the frequency and the cold phase shift per cell of the slow-wave structure. The threshold voltage and the properties of the electron hub are calculated in exactly the same manner as for a magnetron. These values do not depend on the RF field of the slow-wave structure and they are therefore the same throughout the interaction region. The electric field acting on the electrons, and the conduction angle of the spokes, can then be calculated as functions of the anode voltage, the RF wave voltage on the slow-wave structure, and the magnetic field. The RF wave voltage (V_1) is related to the RF power flow in the structure (P_1) by

$$V_1 = \sqrt{2Z_c P_1},\tag{16.5}$$

where Z_c is the coupling impedance of the circuit. The energies of the electrons striking the anode and the cathode are computed in the same way as for a magnetron. The DC current flowing to the anode is computed from the charge density in the hub, and the electric and magnetic fields at the base of the spoke. The impact energies of the electrons, and the current in each spoke, are functions of V_1. They therefore vary with position in the CFA. The RF power delivered to the anode of the equivalent magnetron is given by

$$P_{rf} = \left(V_a - V_{ia} - V_{ic}\right)I_0,\tag{16.6}$$

where V_a and I_0 are the DC anode voltage and current and V_{ia} and V_{ic} are the impact energies of the electrons on the anode and the cathode. Note that V_{ic} is the actual impact energy of electrons on the cathode, multiplied by the ratio of the fraction of electrons striking the cathode to the fraction striking the anode. In (16.6) all the terms on the right-hand side, apart from V_a are functions of the anode voltage, the magnetic field, and the RF power flow on the structure. The RF power delivered to the nth cell of the structure is then

$$\Delta P_n = P_{rf}\left(P_n\right)/N_a,\tag{16.7}$$

where P_n is the power in the growing wave in the nth cell and N_a is the perimeter of the anode divided by the length of a cell. The anode voltage and the magnetic field are the same for all the cells. The power in the next cell is given by

$$P_{n+1} = \left(P_n + \Delta P_n\right)\left(1 - A_c\right),\tag{16.8}$$

where A_c is the fraction of the power lost per cell. This is related to the cold loss of the slow-wave structure in decibels by

$$L_{dB} = 10 N_c \log\left(A_c\right),\tag{16.9}$$

where N_c is the number of active cells. If the RF input power is known then (16.8) can be applied repeatedly to find the power flow in each cell, and the output power of the tube, at the specified anode voltage and magnetic field. If we assume that the power in the CFA grows linearly to this level over N_c cells, whilst that in the magnetron is constant over N_a cells, we find that the power output of the CFA is approximately

$$P_{CFA} \approx \frac{N_c}{2N_a} P_{EM},$$
(16.10)

where P_{EM} is the output power of the equivalent magnetron. The total anode current is given by

$$I_a = \sum_{n=1}^{N_c} (I_n / N_a),$$
(16.11)

where I_n is the anode current of the equivalent magnetron for the nth cell. The total heat dissipation on the anode and the cathode can be found in a similar manner by summing the contributions of the equivalent magnetrons (see Worksheet 16.1).

This model was applied to the data for an S-band forward-wave CFA given in [49] and there was good qualitative agreement with the typical behaviour of a CFA, described in Section 16.4. This tube was identical, or very similar, to the experimental tube studied in [43]. Figure 16.16 shows the Gauss lines computed for three different magnetic fields at the synchronous frequency (3.3 GHz) of the tube. These curves are similar to those in Figure 16.8. Figure 16.17 shows the Gauss lines at three different frequencies, when $B = 0.3 \, \text{T}$, with the experimental curves given in [49] for comparison. The properties of the slow-wave structure, which are not given in that paper, were estimated using the data for a similar tube in [47]. It can
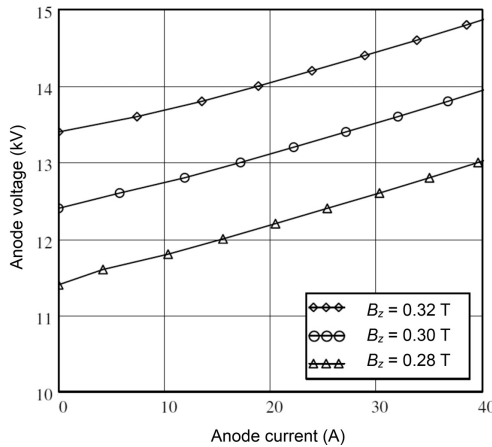
Figure 16.16: Gauss lines computed for the baseline tube at 3.3 GHz and 7 kW RF input power.
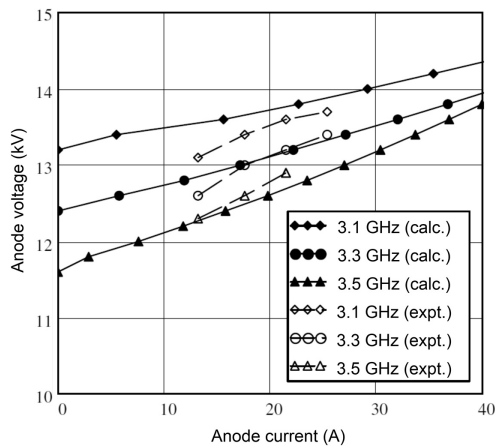
Figure 16.17: Gauss lines computed for the baseline tube at $B_z = 0.3$ T and 7 kW RF input power, with experimental data from [49] for comparison.
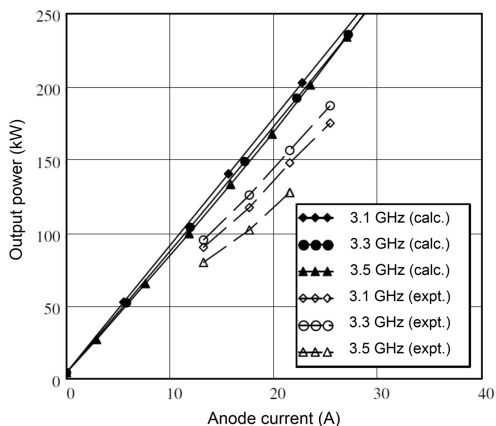
Figure 16.18: RF output power plotted against anode current for the baseline tube at $B_z = 0.3$ T and 7 kW RF input power, with experimental data from [49] for comparison.

be seen that the anode voltage increases with decreasing frequency, as shown in Figure 16.9(b). For frequencies other than the synchronous frequency the number of spokes calculated from (16.3) was not an integer. This is not inconsistent, because the interaction is assumed to be restarted from an unmodulated hub after the drift region.

Figure 16.18 shows the dependence of the output power on the anode current for three different frequencies compared with experimental data from [49]. It can be seen that the output power and efficiency predicted by the model are rather greater than the measured values. Possible reasons for this are explored below.
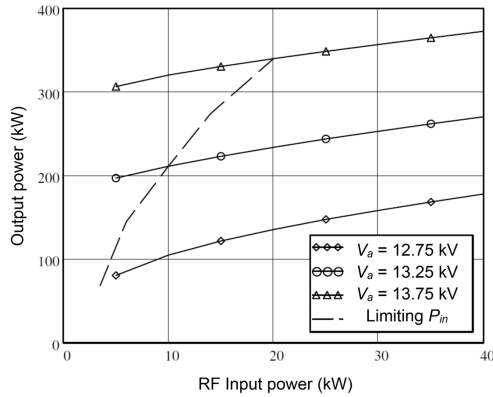
Figure 16.19: RF output power plotted against RF input power for the baseline tube at 3.3 GHz and $B_z = 0.3\,\text{T}$. The dashed line shows the minimum RF input power for which spokes are formed.
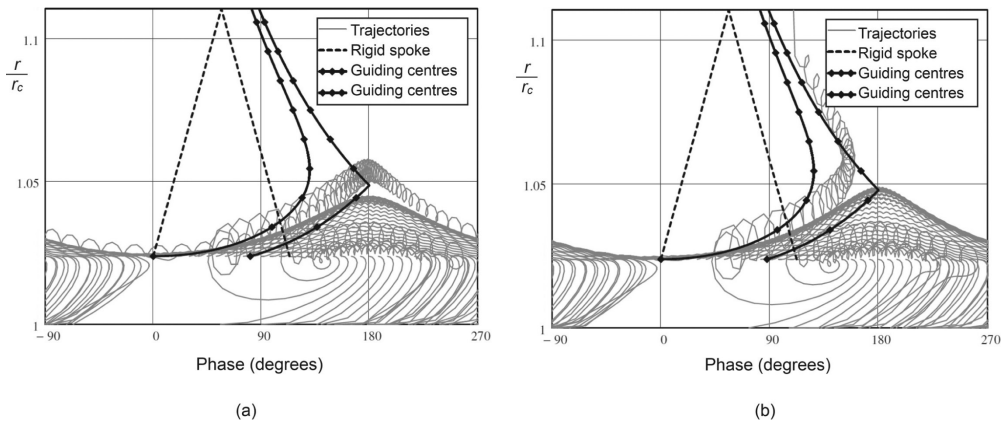
Figure 16.20: Electron trajectories for RF input powers: (a) just below the minimum for spoke formation, and (b) just above the minimum for spoke formation.

Figure 16.19 shows the AM/AM curves computed for three different anode voltages. These may be compared with Figure 16.14(a). At low values of the input power the RF field of the slow-wave structure is not sufficient to enable spokes to form, and the only output is noise. Figure 16.20 shows examples of the electron trajectories just below and just above the limiting input power. Below the limiting power no trajectories reach the anode and the modulation of the hub does not induce sufficient current in the slow-wave structure to enable the wave to grow. The figure also shows the spoke profiles for the rigid spoke and guiding centre approximations. Below the limit for spoke formation these models are incorrect because they predict non-zero anode currents.

Table 16.1 shows a detailed comparison between the results of calculations using the model described with those in [49]. It can be seen that the anode current,

Table 16.1: Comparison between calculated and experimental results for McDowell's baseline forward wave CFA (copyright 2002, IEEE, reproduced, with permission, from [49])

	Moving wavelength	Rigid spoke	Rigid spoke (corrected)	Experiment
Voltage (kV)	13.25	13.25	13.25	13.25
Current (A)	22.0	23.9	20.0	22.0
Power (kW)	156.9	203	168	155
Power added efficiency (%)	54.7	63	61	52
Gain (dB)	13.5	14.6	13.8	13.5
RF power added / DC input (%)	51.4	63.2	60.9	
Anode dissipation / DC input (%)	25.9	15.1	20.1	
Cathode dissipation / DC input (%)	10.6	5.3	7.3	
Circuit loss / d.c input (%)	10.5	13.9	12.0	

output power, efficiency and gain are over-estimated by the rigid spoke model while the anode and cathode dissipation are underestimated. The differences in the figures for the dissipation may be a result of the different assumptions about electron emission from the cathode. The rigid spoke model assumes that the current flow in the hub is space-charge limited whilst the moving wavelength code models secondary electron emission from a cold cathode. Table 15.2 shows that, in a magnetron, this can have a large effect on the anode and cathode dissipation. Because the rigid spoke model uses the dissipation to calculate the RF output power, underestimation of the former leads automatically to over-estimation of the latter. The results are, in any case, very sensitive to changes in the magnetic field. McDowell adjusted the magnetic field to compensate for the use of a linear, rather than a circular model of the tube. It is not clear whether the value quoted in Table II in [49] was adjusted in this way. He observed that the agreement between his results and those of experiment was, in any case, somewhat fortuitous because the calculations ignored the effects of penetration of electrons into the space between the vanes.

Further understanding of the difference between the two models is obtained by plotting the forward power as a function of distance along the slow-wave structure, as shown in Figure 16.21. The low rate of growth close to the input calculated by the moving wavelength model is confirmed by experimental results [43]. It can be seen that both models predict very similar linear rates of growth and that better agreement could be obtained, quite simply, by moving the curve for the rigid spoke model to the right by around eight cells. This can be explained in terms of the finite time taken for electrons to move from the base to the tip of the spoke. A magnetron operates in a steady state, and the transit time can be ignored. In a CFA, on the other hand, the RF fields experienced by the spokes change in amplitude as they move through the tube.
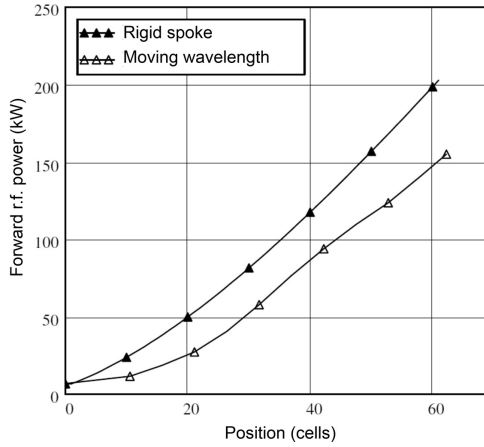
Figure 16.21: Growth in the forward RF power with distance.

The transit time can be estimated as a function of the wave voltage in the equivalent magnetron by integrating (15.113), and making use of the expression for the tangential electric field in (15.19), to give

$$\Delta t = \int_{r_b}^{r_a} \frac{B_z}{E_\theta} \, dr = \frac{B_z r_a}{n V_1} \int_{r_b}^{r_a} \left(\frac{r_a^n - r_a^{-n}}{r^n - r^{-n}} \right) \frac{1}{\sin(n\theta_s)} \, dr. \qquad (16.12)$$

This is the transit time for an electron at the centre of the spoke, assuming that its angular position, relative to the wave, is approximately constant during its motion from the hub to the anode. It can be seen that Δt is inversely proportional to the wave voltage V_1. The transit time can be expressed in terms of the number of cells traversed by the electron in its motion given by $n_s = \omega_s \Delta t / \phi_c$. For the tube under consideration this was found to vary from 2 at the input to around 1 at the output. However, this calculation underestimates the transit time close to the input. Figure 16.20(b) shows that, when the spokes are only just formed, the angular position of the electrons varies considerably. For part of the motion they are in a region of very small tangential electric field. Thus the transit time is increased, so that the electrons move through a greater number of cells. It is difficult to determine the number of cells precisely because the electron dynamics at the input are very dependent on the properties of the electron hub as it emerges from the drift region. But this argument suggests that most of the difference between the two models can be explained in terms of transit time effects.

The rigid spoke model is based on some rather crude approximations, so the effort involved in including transit time effects correctly is not justified (though they have been included in a guiding centre model [42]). However, the shift of the curve in Figure 16.21 can be achieved using a piecewise constant model for the transit time, so that no current reaches the anode in the first eight cells. Current is drawn

from the surface of the hub along the whole length of the slow-wave structure, but it is concentrated along a shorter length of the anode because of transit time effects. The results obtained when this correction is made are shown in the fourth column of Table 16.1. Because the distribution of the current along the anode is not known the power dissipated on the anode and the cathode has been calculated by scaling to achieve an overall power balance. Overall it may be concluded that, with this modification, the rigid spoke model gives a good qualitative understanding of the properties of CFAs.

16.6   CFA Design

The very limited information on CFAs which is available in the literature, the complexity of their behaviour, and the scarcity of good theoretical models means that they must usually be designed by scaling from existing tubes, using dimensionless parameters similar to those used for magnetrons (see Section 15.7.1) [49].

An initial estimate of the performance of a CFA may be made using the simplifying assumption that the power generated per unit length is constant [19]. Figure 16.21 shows that this is true for the greater part of the slow-wave structure. The gain can then be written

$$G = \frac{P_{in} + P_{gen}}{P_{in}} = 1 + \frac{P_{gen}}{P_{in}}, \tag{16.13}$$

where P_{gen} is the RF power generated within the tube. The power added efficiency is given by

$$\eta_a = \frac{P_{out} - P_{in}}{P_{dc}} = \left(\frac{P_{out} - P_{in}}{P_{gen}} \right) \cdot \left(\frac{P_{gen}}{P_{dc}} \right) = \eta_c \eta_e, \tag{16.14}$$

where η_c is the circuit efficiency and η_e is the electronic efficiency.

The electronic efficiency of the equivalent magnetron, given approximately by (15.50), depends only on the dimensions of the tube and on the magnetic field. Thus, to this level of approximation, it is constant for the growing wave throughout the CFA. The normalised magnetic field of the baseline tube in [49] was $B_z/B_0 = 4.6$, giving the efficiency of the equivalent magnetron as approximately 75%. Gewartowski and Watson multiplied this figure by the fraction of the perimeter of the tube occupied by the slow-wave structure. That adjustment is not correct because the RF power and the anode current are both proportional to the length of the slow-wave structure. However, these quantities are affected differently by the electron transit time. Thus the electronic efficiency is reduced by the fraction of the slow-wave structure that contributes to the signal growth. We saw, above, that this could be represented by reducing the active length of the slow-wave structure by 8 cells, in a total of 62, giving a reduction in efficiency of 0.87. Thus the estimated electronic efficiency of the CFA is 65%.

To find an expression for the circuit efficiency we note that the output power is given by

$$P_{out} = P_{in} \exp(-2\alpha L_c) + \frac{P_{gen}}{L_c} \int_0^{L_c} \exp(-2\alpha(L_c - x))\, dx$$

$$= P_{in} \exp(-2\alpha L_c) + \frac{P_{gen}}{2\alpha L_c}(1 - \exp(-2\alpha L_c)), \qquad (16.15)$$

where L_c is the length of the slow-wave structure and α is the attenuation coefficient which is related to the cold loss of the structure by

$$L_{dB} = -20\,\alpha L_c \log(e). \qquad (16.16)$$

Thus the circuit efficiency is

$$\eta_c = \left(\frac{1}{2\alpha L} - \frac{P_{in}}{P_{gen}} \right)(1 - \exp(-2\alpha L)). \qquad (16.17)$$

The baseline tube in [49] had 13 dB gain and 2 dB cold loss so that $2\alpha L_c = 0.46$, $P_{in}/P_{gen} = 0.05$ and $\eta_c = 0.78$. Thus, the estimated overall efficiency of the tube is 51%, which is close to the experimental figure in Table 16.1. Figure 16.3 shows the efficiencies of CFAs for which data is available.

Under synchronous conditions this tube had 19 spokes. An equivalent magnetron would therefore have 34 cavities, which is much too great for satisfactory separation of the modes. That problem does not occur in a CFA, and the anode and cathode radii can be much greater than in an equivalent magnetron (see Table 15.5). The ratio of the anode length to the free-space wavelength is 0.13 which is very similar to that in a magnetron. The modified Slater ratio is 1.54 because the gap between the anode and cathode is smaller than it would be in a typical magnetron. This could be to ensure that the RF electric field at the desired input power was strong enough to ensure satisfactory cold-cathode operation.

In this tube the cathode was offset from the mechanical centre because that had been found experimentally to give improved performance. It was also shown that similar improvements could be made by tapering the pitch of the slow-wave structure [43, 47]. This is probably a result of the reduction in the tangential velocity of the spokes, caused by space-charge potential depression close to the output, where the spoke charge is greatest [14]. Some additional information about the design of CFAs can be found in [8, 33, 67].

References

[1] J. R. M. Vaughan, 'Beam buildup in a DEMATRON amplifier', *IEEE Transactions on Electron Devices,* vol. 18, pp. 365–373, 1971.

[2] R. G. Lock, 'A non-reentrant crossed-field amplifier with cycloiding injected beam', *IEEE Transactions on Electron Devices,* vol. 16, pp. 986–995, 1969.

[3] J. Z. Ye *et al.*, 'An experimental investigation of the end-hat effects in a crossed-field amplifier via three-dimensional electron density measurements', *IEEE Transactions on Electron Devices,* vol. 41, pp. 258–265, 1994.

[4] D. Chen, 'An emitting-sole linear magnetron amplifier', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 223–246, 1961.

[5] J. F. Skowron, 'The continuous-cathode (emitting-sole) crossed-field amplifier', *Proceedings of the IEEE,* vol. 61, pp. 330–356, 1973.

[6] W. C. Brown, 'The Platinotron: Amplitron and Stabilotron', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 165–209, 1961.

[7] W. C. Brown, 'The microwave magnetron and its derivatives', *IEEE Transactions on Electron Devices,* vol. 31, pp. 1595–1605, 1984.

[8] R. M. Gilgenbach *et al.*, 'Crossed-field devices', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 289–342, 2005.

[9] A. S. Gilmour, Jr., *Microwave Tubes*. Dedham, MA: Artech House, 1986.

[10] L. Sivan, *Microwave Tube Transmitters*. London: Chapman and Hall, 1994.

[11] B. Epsztein, 'Cross-field tubes', in B. L. Smith and M.-H. Carpentier, eds, *The Microwave Engineering Handbook*, vol. 1. London: Chapman & Hall, pp. 65–79, 1993.

[12] V. Granatstein *et al.*, 'Vacuum electronics at the dawn of the twenty-first century', *Proceedings of the IEEE,* vol. 87, pp. 702–716, 1999.

[13] N. J. Dionne *et al.*, 'Methods for enhancing low-noise CFA performance (crossed-field amplifiers)', in *International Electron Devices Meeting*, pp. 529–532, 1990.

[14] G. E. Pokorny *et al.*, 'The DEMATRON – a new crossed-field amplifier', *IRE Transactions on Electron Devices,* vol. 9, pp. 337–345, 1962.

[15] R. R. Warnecke *et al.*, 'The M-type carcinotron tube', *Proceedings of the IRE,* vol. 43, pp. 413–424, 1955.

[16] G. Boucher, 'Technology of carcinotrons for short and long wavelengths', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 105, pp. 897–899, 1958.

[17] M. Favre, 'Results obtained on cross-field carcinotrons under pulsed operation', *Proceedings of the IEE – Part B: Radio and Electronic Engineering,* vol. 105, pp. 533–537, 1958.

[18] O. Doehler, 'The M-Carcinotron', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 11–34, 1961.

[19] J. W. Gewartowski and H. A. Watson, *Principles of Electron Tubes*. Princeton, NJ: D. van Nostrand, 1965.

[20] G. Faillon *et al.*, 'Microwave Tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer-Verlag, pp. 1–84, 2008.

[21] R. J. Espinosa and R. R. Moats, 'Broad-band injected-beam crossed-field amplifiers', *IEEE Transactions on Electron Devices,* vol. 24, pp. 13–21, 1977.

[22] W. Sobotka, 'Computation of nonreentrant CFA characteristics', *IEEE Transactions on Electron Devices,* vol. 17, pp. 622–632, 1970.

[23] R. H. Levy, 'Diocotron instability in a cylindrical geometry', *Physics of Fluids (1958–1988),* vol. 8, pp. 1288–1295, 1965.

[24] R. W. Gould, 'Space charge effects in beam-type magnetrons', *Journal of Applied Physics,* vol. 28, pp. 599–605, 1957.

[25] J. Feinstein and G. S. Kino, 'The large-signal behavior of crossed-field traveling-wave devices', *Proceedings of the IRE,* vol. 45, pp. 1364–1373, 1957.

[26] J. Browning *et al.*, 'A low-frequency crossed-field amplifier for experimental investigations of electron-radio frequency wave interactions', *IEEE Transactions on Plasma Science,* vol. 19, pp. 598–606, 1991.

[27] J. Browning *et al.*, 'Electron plasma and wave measurements in a crossed-field amplifier and comparison with numerical simulation', *IEEE Transactions on Electron Devices,* vol. 39, pp. 2401–2407, 1992.

[28] SFD, *Introduction to Pulsed Crossed-Field Amplifiers*. Union, NJ: SFD Laboratories Inc., 1967.

[29] H. Steyskal, 'Studies on C.W. magnetron amplifiers with continuous cathode and re-entrant beam', in *4th International Congress on Microwave Tubes*, Scheveningen, The Netherlands, pp. 168–172, 1962.

[30] M. J. Clark, 'Slow wave circuits for miniature crossed field amplifiers', in *11th European Microwave Conference*, pp. 645–649, 1981.

[31] A. Dallos *et al.*, 'Effects of ionized oxygen on primary and secondary emission, and total current of a CFA', *IEEE Transactions on Electron Devices,* vol. 34, pp. 1201–1208, 1987.

[32] G. H. MacMaster, 'Current status of crossed-field devices', in *International Electron Devices Meeting*, pp. 358–361, 1988.

[33] G. H. MacMaster and L. J. Nichols, 'Millimeter-wave high gain crossed-field amplifier', in *International Electron Devices Meeting*, pp. 963–966, 1990.

[34] G. E. Dombrowski, 'Theory of the Amplitron', *IRE Transactions on Electron Devices,* vol. 6, pp. 419–428, 1959.

[35] W. A. Smith, 'A wave treatment of the continuous cathode crossed-field amplifier', *IRE Transactions on Electron Devices,* vol. 9, pp. 379–387, 1962.

[36] H. Steyskal, 'Continuous cathode crossed-field amplifiers', *IEEE Transactions on Electron Devices,* vol. 10, pp. 95–96, 1963.

[37] D. MacGregor, 'Computer modeling of crossed-field tubes', *Applications of Surface Science,* vol. 8, pp. 213–224, 1981.

[38] D. P. Chernin, 'Computer simulations of low noise states in a high-power crossed-field amplifier', *IEEE Transactions on Electron Devices,* vol. 43, pp. 2004–2010, 1996.

[39] L. D. Ludeking *et al.*, 'An examination of the performance of backward wave CFA's in simulation and experiment', in *Pulsed Power Plasma Science*, vol. 1, pp. 229–231, 2001.

[40] D. M. MacGregor, 'Computer simulation of the backward-wave distributed-emission crossed-field amplifier', *IEEE Electron Device Letters,* vol. 1, pp. 134–136, 1980.

[41] D. Chernin, 'Private communication', ed, 2014.

[42] S. A. Riyopoulos *et al.*, 'Guiding center fluid model of the crossed-field amplifier', *IEEE Transactions on Electron Devices,* vol. 39, pp. 1529–1542, 1992.

[43] G. K. Farney, 'CFA design improvement program: Final Report. Volume 1: Instrumented CFA studies', Varian Associates, Beverly, MA, 1978.

[44] D. Chernin and A. Drobot, 'Computer simulations of re-entrant crossed-field amplifiers', in *International Electron Devices Meeting*, pp. 521–524, 1990.

[45] D. Chernin *et al.*, 'Computer studies of noise generation in crossed-field amplifiers', in *International Electron Devices Meeting*, pp. 593–596, 1991.

[46] G. E. Dombrowski, 'Simulation of magnetrons and crossed-field amplifiers', *IEEE Transactions on Electron Devices,* vol. 35, pp. 2060–2067, 1988.

[47] H. L. McDowell, 'CFA design improvement program: Final Report. Volume 2: Computer modeling studies', Varian Associates, Beverly, MA, 1978.

[48] H. L. McDowell, 'CFA computer modeling using a moving wavelength code', in *Proc. 1st Int. Workshop Crossed-Field Devices*, pp. 15–16, 1995.

[49] H. L. McDowell, 'Crossed-field amplifier simulations using a moving wavelength computer code', *IEEE Transactions on Plasma Science,* vol. 30, pp. 962–979, 2002.

[50] S. Yu *et al.*, 'Time-dependent computer analysis of electron-wave interaction in crossed fields', *Journal of Applied Physics,* vol. 36, pp. 2550–2559, 1965.

[51] G. I. Churyumov and G. I. Sergeev, 'Simulation and modeling of self-modulated re-entrant beam crossed-field amplifier', *IEEE Transactions on Electron Devices,* vol. 46, pp. 1063–1069, 1999.

[52] K. Eppley, 'Numerical simulation of cross field amplifiers SLAC-PUB-5183', Stanford Linear Accelerator Center, Menlo Park, CA (USA), 1990.

[53] K. Eppley and K. Ko, 'Design of a high power cross field amplifier at X band with an internally coupled waveguide', presented at the SPIE Symposium on High Power Lasers, Los Angeles, CA, 1991.

[54] J. R. M. Vaughan, 'Calculation of coupled-cavity TWT performance', *IEEE Transactions on Electron Devices,* vol. 22, pp. 880–890, 1975.

[55] G. E. Thomas, 'The nonlinear operation of a microwave crossed-field amplifier', *IEEE Transactions on Electron Devices,* vol. 28, pp. 27–36, 1981.

[56] G. E. Thomas, 'Soliton-voltage and phase characteristics of a microwave crossed-field amplifier', *IEEE Transactions on Electron Devices,* vol. 29, pp. 1210–1218, 1982.

[57] S. Riyopoulos, 'Simulations of crossed-field amplifier operation using guiding center dynamics', in *International Electron Devices Meeting*, pp. 525–528, 1990.

[58] S. Riyopoulos, 'Guiding center theory and simulations of the crossed-field amplifier', *Physics of Fluids B: Plasma Physics (1989–1993),* vol. 3, pp. 3505–3516, 1991.

[59] S. A. Riyopoulos, 'Feedback-induced noise in crossed field devices', *IEEE Transactions on Plasma Science,* vol. 20, pp. 360–369, 1992.

[60] D. Kaup and G. E. Thomas, 'Density profile and current flow in a crossed-field amplifier', *Journal of Plasma Physics,* vol. 58, pp. 145–161, 1997.

[61] D. Kaup and G. E. Thomas, 'Stationary operating density profiles in a crossed-field amplifier', *Journal of Plasma Physics,* vol. 59, pp. 259–276, 1998.

[62] D. J. Kaup, 'Theoretical modeling of crossed-field electron vacuum devices', *Physics of Plasmas (1994-present),* vol. 8, pp. 2473–2480, 2001.

[63] D. J. Kaup, 'Initiation and stationary operating states in a crossed-field vacuum electron device', *Proceedings of SPIE,* vol. 4720, pp. 67–74, 2002.

[64] R. J. Collier, 'Gain measurements on a forward wave crossed-field amplifier', *Proceedings of the IRE,* vol. 49, p. 372, 1961.

[65] R. J. Collier, 'Bi-signal amplification by a forward wave crossed-field amplifier', *Proceedings of the IRE,* vol. 49, p. 646, 1961.

[66] J. Feinstein and R. J. Collier, 'Waveguide-coupled crossed-field amplifier', in E. Okress, ed., *Crossed-Field Microwave Devices*, vol. 2. New York: Academic Press, pp. 211–222, 1961.

[67] T. E. Ruden, 'The Amplitron as a high power, efficient, RF power source for long pulse, high resolution linear accelerators', *IEEE Transactions on Nuclear Science,* vol. 12, pp. 169–173, 1965.

17 Fast-Wave Devices

17.1 Introduction

The purpose of this chapter is to provide an introduction to the fast-wave tubes, especially to gyrotron oscillators and amplifiers which are important sources of RF power at millimetre and sub-millimetre wavelengths. The scope of the chapter is limited to those devices up to a frequency of around 1 THz which can be classed as vacuum tubes. Many fast-wave devices are described as masers or lasers (from the acronym Microwave/Light Amplification by Stimulated Emission of Radiation) and some theoretical treatments regard them as quantum devices [1]. However, at least for the frequency range considered in this book, they can be regarded as classical devices in which the energy is transferred from a bunched electron beam to a synchronous electromagnetic wave. We shall see that, in many cases, the theory is analogous to that of the TWT. Many devices employ the electron cyclotron maser (ECM) interaction which is reviewed in Section 17.2. Examples, including gyrotron oscillators, gyro-klystrons, and gyro-TWAs, are discussed in Sections 17.3 and 17.4. Peniotrons and ubitrons are considered at the end of the chapter.

We have seen in the preceding chapters that, above some frequency, the output power and efficiency of linear-beam and crossed-field tubes decreases very rapidly with increasing frequency. This is caused chiefly by the very rapid fall-off in the RF electric field from the surface of the electromagnetic structure as discussed in Section 4.1.1. The transverse variation of the interaction field is as $I_0(\gamma_e r)$ in cylindrical geometry so that, as a tube is scaled to higher frequencies the parameters $\gamma_e b$ and $\gamma_e a$ must be kept constant to maintain the strength of the interaction. The beam plasma frequency is related to the strength of the magnetic focusing field by (7.55) which may be written

$$\omega_p = \frac{1}{\sqrt{2}} \cdot \frac{e}{m_0} \cdot \frac{B_z}{m},$$ (17.1)

where B_z is the axial magnetic field and m is the ratio of the magnetic field to the Brillouin field. If a solenoid with iron pole-pieces is used to provide the magnetic field, then the maximum value of B_z is limited to around 2T by saturation of the iron. Higher fields can be provided by normal, or superconducting, solenoids which do not incorporate any iron. But it is then more difficult to achieve the desired profile of B_z as a function of the axial position. The maximum value of ω_p is also likely

to be reduced as the frequency increases by the need to increase the beam stiffness, as the beam and tunnel radii are reduced. We have also seen (in (13.51)) that the beam perveance is approximately proportional to $\left(\omega_p/\omega\right)^2$ when $\gamma_e b$ is constant. Thus the maximum possible perveance of a magnetically focused beam decreases rapidly with increasing frequency. If the beam voltage is kept constant, then the beam current and power are inversely proportional to the square of the frequency. The RF output power is further reduced by the decreases in electronic efficiency and circuit efficiency as the frequency increases. Additional limitations may be set by the need to reduce the beam filling factor to ensure good beam transmission. Hence, the limitation in performance of a TWT or crossed-field tube at millimetre and sub-millimetre wavelengths arises directly from the need to reduce the phase velocity of the electromagnetic wave to achieve synchronism with the velocity of the electrons. Similar considerations apply to klystrons because of the need to keep the normalised interaction gap lengths $\left(\beta_e g\right)$ small to achieve useful coupling between the beam and the cavities. These limitations on performance can be mitigated to some extent by the use of sheet or annular electron beams, and of higher beam voltages. But the essential problem, of keeping an electron beam of sufficient power very close to the surface of the electromagnetic structure, remains [2].

We saw in Chapter 2 that, for frequencies above cut-off, the phase velocity of the waves in a smooth metallic waveguide circular waveguide is greater than the velocity of light, and strong RF electric fields can exist throughout the guide. Thus, if an electron beam can interact with the electromagnetic field of a uniform waveguide, it does not have to be close to the surface of the guide. In addition, the volume of space which the beam can occupy is much greater so that the possible beam current and power are increased. It is, of course, still necessary to achieve synchronism between the electrons and the electromagnetic wave if a useful interaction is to occur. In a folded-waveguide TWT synchronism is achieved by making the structure periodic so that the interaction with a uniform beam is also periodic. That suggests the alternative strategy of using a periodic electron beam and a uniform structure. In principle this could be achieved in a number of ways, but practical devices have linear beams controlled by either a uniform axial magnetic field or a periodic magnetic field. These devices are classed as fast-wave devices because the phase velocity of the wave in the electromagnetic structure is greater than the velocity of light [3]. Fast-wave devices have been discussed by many authors, and the reader is referred to the literature for detailed information [4–14]. Much of the pioneering work was done in the former Soviet Union, and a summary can be found in [15]. The sources cited in this chapter are English because they are most readily accessible, but work in Russia continues to be important.

17.2   Electron Cyclotron Masers

The electron cyclotron maser (ECM) interaction is the basis for a number of types of tube of which gyrotrons (see Section 17.3) are an important example. The

principle of operation of the ECM interaction can be understood by considering Figure 17.1 which shows an electron moving under the influence of a uniform longitudinal magnetic field in the presence of a uniform transverse RF electric field of frequency ω. The electron moves in a circular orbit with an angular velocity equal to the relativistic cyclotron frequency

$$\omega_c = \frac{eB_z}{m_0 \gamma},$$ (17.2)

where

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}}$$ (17.3)

is the relativistic factor and

$$u^2 = u_\theta^2 + u_z^2,$$ (17.4)

where u_θ and u_z are the tangential and axial components of the electron velocity and it is assumed that the radial component of velocity is negligible. It can be seen that the cyclotron frequency decreases as the electron velocity increases.

If the axial motion of the electrons is ignored, for the moment, then the motion of an electron is in synchronism with the RF field when $\omega = \omega_c$. A synchronous electron at position 1 experiences a retarding field. When it reaches position 2 the direction of the RF electric field has reversed and the motion of the electron is further retarded. Similarly, an electron starting at position 2 is accelerated and then accelerated again when it reaches position 1. The accelerated electrons fall back in phase relative to the RF field, because of the change in the cyclotron frequency with velocity. Similarly, retarded electrons move forward in phase so that electron bunches are formed. If these bunches fall in the retarding phase of the RF field then there is a net transfer of energy from the electrons to the wave, as required. When relativistic effects are ignored the cyclotron frequency is constant and no bunching
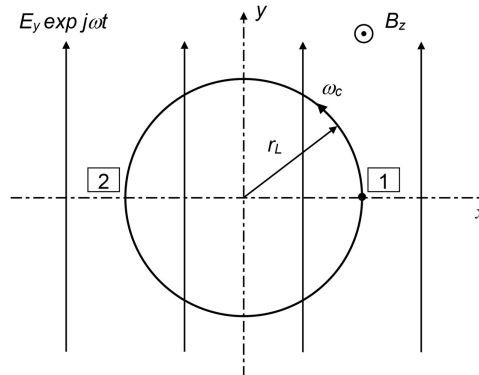
Figure 17.1: Principle of the electron cyclotron maser interaction.

takes place. Hence a necessary condition for the ECM interaction is that the electron velocity should be high enough for relativistic effects to be important.

The radius of the orbit is the Larmor radius given by

$$r_L = \frac{u_\theta}{\omega_c},$$

(17.5)

which increases as the electron velocity increases. Because the RF electric field is uniform in space, the change in the Larmor radius does not affect the field acting on the electrons. This situation changes if the electric field is non-uniform because a net transfer of energy to the wave is possible, even when the velocities are low. This is the peniotron interaction which is discussed in Section 17.5. It may also occur in ECM devices if the RF electric field is non-uniform in the region of the electron beam.

The electron beam is normally generated by a magnetron injection gun so that the energy of the electrons can be written using (1.4) as

$$u_\theta^2 + u_z^2 = c^2 \left[1 - \frac{1}{\left[1 + \left(e V_a / m_0 c^2 \right) \right]^2} \right],$$

(17.6)

where V_a is the potential of the anode relative to the cathode. The ratio of the transverse velocity to the axial velocity is denoted by the pitch factor

$$\alpha = \frac{u_\theta}{u_z}.$$

(17.7)

This parameter is normally greater than unity. Because only the azimuthal velocity is changed by the interaction it follows that only the azimuthal energy can be converted into RF power. Thus the electronic efficiency cannot exceed

$$\eta_e = \frac{\left(u_\theta^2 + u_z^2 \right) - u_z^2}{u_\theta^2 + u_z^2} = \frac{\alpha^2}{1 + \alpha^2}.$$

(17.8)

In practice it is not possible to reduce the tangential velocities of all the electrons to zero and the electronic efficiency is smaller that given by (17.8).

When the axial velocity of the electrons is taken into account the synchronous condition is slightly altered. The time taken for an electron to make one circuit of its orbit is

$$t = \frac{2\pi}{\omega_c}.$$

(17.9)

In that time the electron moves through an axial distance

$$z = u_z t = \frac{2\pi u_z}{\omega_c}.$$

(17.10)

For synchronism, the phase of the electromagnetic wave as seen by the electron must be

$$\omega t - \beta z = 2s\pi, \tag{17.11}$$

where the wave propagates as $\exp j(\omega t - \beta z)$ and s is a positive integer. Substituting for t and z in (17.11) from (17.9) and (17.10) the condition for synchronism is found to be

$$\omega = \beta u_z \pm s\omega_c, \tag{17.12}$$

where the first term on the right-hand side represents the Doppler shifting of the frequency caused by the longitudinal motion of the electron. Now the dispersion curve of any mode in a uniform waveguide is given by (2.11), which may be written

$$\beta = \frac{1}{c}\sqrt{\omega^2 - \omega_{mn}^2}, \tag{17.13}$$

where ω_{mn} is the cut-off frequency of the mode with mode numbers (m, n). Note that this notation has been used to avoid confusion with the use of ω_c for the cyclotron frequency. Equation (17.13) can be written in normalised form as

$$\frac{\omega}{\omega_{mn}} = \sqrt{1 + \frac{\beta^2}{\beta_{mn}^2}}, \tag{17.14}$$

where $\beta_{mn} = \omega_{mn}/c$. Using the same normalisations the dispersion equation for waves on the beam (17.12) can be written

$$\frac{\omega}{\omega_{mn}} = \frac{\beta}{\beta_{mn}} \cdot \frac{u_z}{c} \quad \pm \quad \frac{s\omega_c}{\omega_{mn}}. \tag{17.15}$$

Figure 17.2 shows a typical uncoupled dispersion diagram plotted using equations (17.14) and (17.15). It can be seen that there are normally two points of intersection between the waveguide mode and the electron cyclotron wave (marked A and B in the diagram). The different types of ECM device correspond to different choices of these points [10]:

- Gyrotron oscillator and gyro-klystron amplifier: The two points of intersection coincide so that the cyclotron wave line is tangential to the waveguide line close to the waveguide cut-off frequency on the right-hand side of the diagram.
- Gyro-TWT: The two points of intersection lie close to one another on the right-hand side of the diagram so that a travelling-wave interaction is possible over a band of frequencies.
- Gyro-BWO: Point A lies on the left-hand side of the diagram so that the beam interacts with the backward wave in the waveguide.
- Cyclotron auto-resonance maser (CARM): Point B lies well to the right-hand side of the diagram so that the phase velocity is close to the velocity of light. When energy is extracted from the beam, the electron velocity decreases but the relativistic cyclotron frequency increases so that the resonant condition (17.12) is maintained.

These types of tube are discussed in greater detail below.
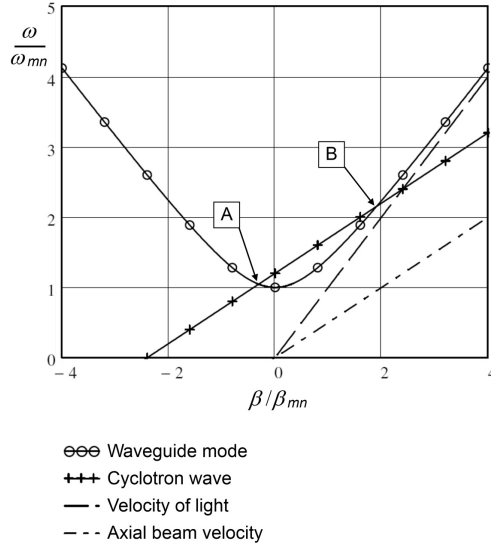
Figure 17.2: Uncoupled dispersion diagram for electron cyclotron mode interactions $\left(u_z/c = 0.264, \ \alpha = 1.5, \ s\omega_c/\omega_{mn} = 1.2\right)$.

17.2.1   Small-Signal Theory of ECM Interactions

From coupled-mode theory we expect the small-signal dispersion equation, when the modes shown in Figure 17.2 are coupled, to be of the form

$$\left(\omega^2 - c^2\left(\beta^2 + \beta_{mn}^2\right)\right)\left(\omega - \beta u_z - s\omega_c\right)^2 = K f\left(\omega, \beta\right), \tag{17.16}$$

where K is a coupling constant which depends upon the properties of the waveguide and of the electron beam, and f is a function of ω and β. The squaring of the second term on the left-hand side of the equation is suggested by considering (11.137) in the limit of low space-charge. When $K \rightarrow 0$ the solutions of (17.16) are the uncoupled modes shown in Figure 17.2, as expected. Equation (17.16) is a dispersion equation having the general form

$$D\left(\omega, \beta\right) = 0, \tag{17.17}$$

which is very similar to the dispersion equation for the travelling-wave tube (see (11.137)). The properties of a device represented by (17.17) are investigated by searching for the roots of this equation (compare Sections 11.5 and 11.7) [16, 17]. In general both ω and β may be complex.

In an ECM the RF electric field is provided by a guided electromagnetic wave. The simplest way of doing this is to use a TE_{0n} mode of a circular waveguide in which the electric field lines form concentric circles (see Figure 2.14). The tangential component of the electric field varies with radius as

$$E_\theta\left(r\right) = E_0 J_1\left(\beta r\right), \tag{17.18}$$
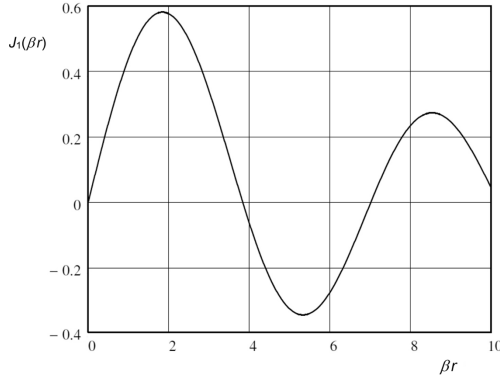
Figure 17.3: Radial variation of the tangential electric field in the TE_{0n} modes of a circular waveguide.

as shown in Figure 17.3. The tangential electric field must be zero on the wall of the waveguide so that

$$J_1(\beta_{0n}a) = 0, \tag{17.19}$$

where a is the radius of the waveguide and $n = 1, 2, 3 \ldots$. When $n = 1$ $\beta_{01}a = 3.832$ and the maximum of the curve is $\beta_{01}r = 1.841$ so that $r/a = 0.48$. Thus, for the best interaction the electrons should be in a thin annular beam whose mean radius is centred on the peak of the field as shown in Figure 17.4. The Larmor radius (r_L) should be small compared with the mean radius r_0.

The dispersion equation is then found to be [18, 19]

$$\omega^2 - c^2\left(\beta^2 + \beta_{0n}^2\right) = \frac{-vc^2}{\pi\gamma a^2 J_0^2(\beta_{0n}a)} \left\{ \frac{\left(\omega^2 - \beta^2 c^2\right)}{\left(\omega - \beta u_z - s\omega_c\right)^2} \cdot \frac{u_\theta^2}{c^2} R_s - \frac{\omega - \beta u_z}{\omega - \beta u_z - s\omega_c} Q_s \right\},$$

$$\tag{17.20}$$

where J_0 is the Bessel function of the first kind,

$$v = \frac{Ne^2}{\varepsilon_0 m_0 c^2} \tag{17.21}$$

is a dimensionless beam density parameter in which N is the number of electrons per unit length of the beam,

$$R_s = \left[J_s(\beta_{0n}r_0) J_s'(\beta_{0n}r_L)\right]^2, \tag{17.22}$$

and $Q_s(\beta_{0n}r_0, \beta_{0n}r_L)$ is defined in [18]. The second term in the brackets in (17.20) is small, except when the transverse velocity u_θ is small and it may, therefore, be neglected so that the equation becomes

$$\left[\omega^2 - c^2\left(\beta^2 + \beta_{0n}^2\right)\right]\left[\omega - \beta u_z - s\omega_c\right]^2 = \frac{-vu_\theta^2}{\pi\gamma} \cdot \left[\frac{J_s(\beta_{0n}r_0) J_s'(\beta_{0n}r_L)}{aJ_0(\beta_{0n}a)}\right]^2 \left(\omega^2 - \beta^2 c^2\right).$$

$$\tag{17.23}$$
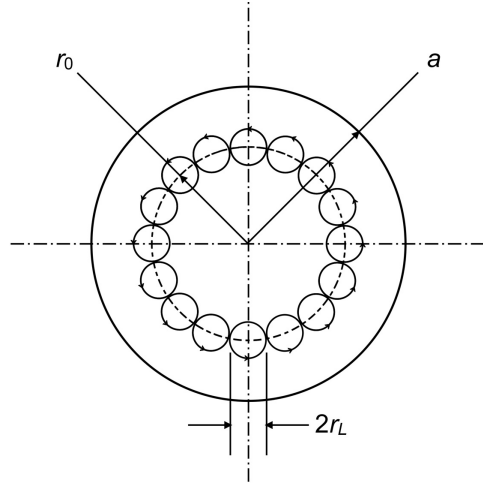
Figure 17.4: Arrangement of the electron beam in an ECM.

We note that this equation has the form expected in (17.16). The parameter v can be written

$$v = (\gamma - 1)\left(\frac{c}{u_z}\right)\frac{I_0 Z_0}{V_a},\qquad(17.24)$$

where $Z_0 = \sqrt{\mu_0/\varepsilon_0}$. The analogy with (11.137) is made more explicit if we write $\beta_0 = \omega/c$, $\beta_c = \omega_c/c$ and rearrange the right-hand side to give

$$[\beta_0^2 - \beta^2 - \beta_{0n}^2][\beta_0 - \beta(u_z/c) - s\beta_c]^2 = -K(\beta_0^2 - \beta^2)\beta_{0n}^2,\qquad(17.25)$$

where

$$K = \frac{\alpha u_z (\gamma - 1)}{\pi \gamma c} \cdot \frac{I_0 Z_0}{V_a} \cdot \left[\frac{J_s(\beta_{0n}r_0)J_s{}'(\beta_{0n}r_L)}{\beta_{0n}a J_0(\beta_{0n}a)}\right]^2.\qquad(17.26)$$

It can be seen that the coupling between the beam and the waveguide increases as the beam current and the parameter α increase. Equation (17.25) is derived using the Vlasov equation which is more familiar to plasma physicists than it is to electronic engineers. The derivation requires some major assumptions including the representation of the beam by an infinitely thin annulus in which all the electrons have the same initial velocity [20]. An alternative derivation based on an azimuthal perturbation of the charge density in a thin annular beam gives essentially the same result [21]. It should be noted that both derivations neglect the azimuthal component of the space-charge field. This can be justified on the grounds that the dominant component of the space-charge field is radial. It is found that RF space-charge effects are not important in gyrotron oscillators, but they can have an important influence of the bunching in drift regions of gyro-amplifiers [12]. Because space-charge effects have been neglected, the two cyclotron modes are
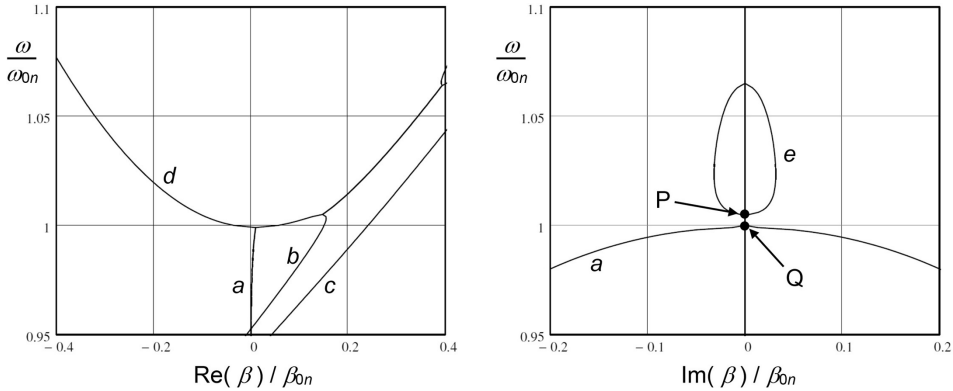
Figure 17.5: The real and imaginary parts of the propagation constants of an ECM computed for a range of real frequencies. $\left(V_a = 70\,\text{kV}, I_0 = 2\,\text{A}, \alpha = 1.5, \omega_c/\omega_{0n} = 0.945\right)$.

degenerate (i.e. they have the same propagation constants at a given frequency) like space-charge waves in the limit $\omega_p \to 0$. The resemblance between (17.25) and (11.137) has led some authors to follow Pierce [22] in approximating the equation as a cubic so that the solutions could be expressed in terms of Pierce parameters familiar to TWT engineers [19, 23]. This approach is now unnecessary because it is straightforward to find the solutions to the full quartic equation, as shown in Worksheet 17.1.

The roots of (17.25) can be computed for a given set of parameters. Figure 17.5 shows the real and imaginary parts of the propagation constant computed for real frequencies, using data from [23] and a beam current of 2 A. The axes are normalised to the cut-off frequency of the waveguide and the free-space propagation constant at that frequency. Below the cut-off frequency the forward and backward waveguide modes (a) are evanescent. The propagation constants of the negative and positive energy cyclotron modes (b and c) are real, but perturbed by coupling with the waveguide, so that the degeneracy of the uncoupled modes is removed. It should be noted that the negative energy mode (b) has a higher phase velocity than the positive energy mode (c). This is the reverse of the behaviour of longitudinal space-charge waves. Above the cut-off frequency, the propagation constants of the backward waveguide mode (d), and the positive energy cyclotron mode (c), are real. The coupling between the negative energy cyclotron mode, and the forward waveguide mode, yields a complex conjugate pair of roots (e) over a range of frequencies. Similarly, when the equation is solved for real values of the propagation constant, the result is a complex conjugate pair of roots for the frequency over the same range. Thus the interaction satisfies the criteria given by Briggs for a convective instability analogous to that in a TWT [17]. This can be used in a gyro-TWT to produce gain over a band of frequencies (see Section 17.4).

When the beam current is increased, the frequency range over which a complex conjugate pair of roots exists increases. The point P at which β is real moves towards
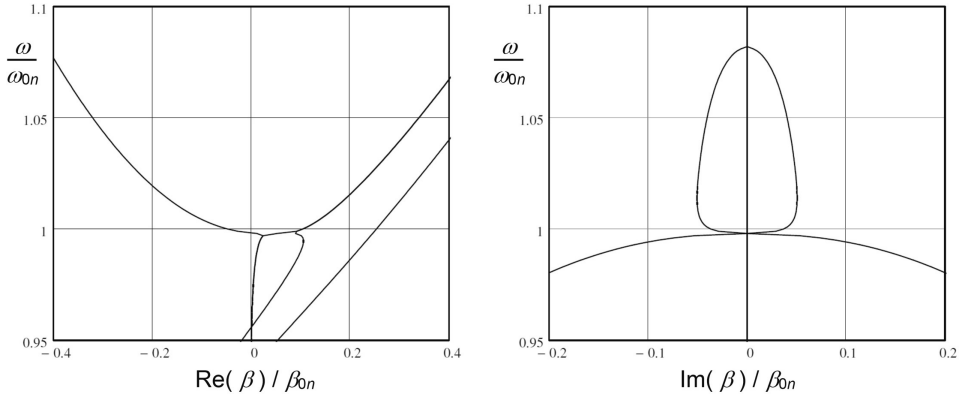
Figure 17.6: The real and imaginary parts of the propagation constants of an ECM computed for a range of real frequencies. $\left(V_a = 70\,\text{kV}, I_0 = 4\,\text{A}, \alpha = 1.5, \omega_c/\omega_{0n} = 0.945\right)$

Q. At a current of approximately 4A (for the example chosen) these points coincide, as shown in Figure 17.6 so that all the roots are real at that point and a double root exists [23]. This satisfies Briggs' criterion for the onset of a non-convective instability [17, 23]. If the beam current is increased still further the tube may oscillate under conditions determined by (17.17) and by

$$\frac{d}{d\beta} D(\omega, \beta) = 0, \tag{17.27}$$

where both ω and β are complex (compare Section 11.7). The term $Q_s\left(\beta_{0n}r_0, \beta_{0n}r_L\right)$ in (17.20) has the opposite sign to R_s and tends to increase the threshold at which a non-convective instability can occur. The threshold current is found to be increased by the wall losses of the waveguide but not by velocity spread in the electron beam [23]. The onset of oscillations is also affected by the length of the interaction region, as in the case of a backward-wave oscillator. Unlike the linear-beam BWO an ECM can also oscillate when the instability is convective because of reflections at the ends of the interaction region. From this discussion it can be concluded that a ECM oscillator may be based on either a convective, or a non-convective, instability according to the design and the operating conditions. In an ECM amplifier, however, it is necessary to ensure that the tube is stable against both the non-convective instability and feedback oscillations.

17.3   Gyrotron Oscillators

A gyrotron is an ECM oscillator. This type of tube has been the subject of much research and development effort, and nearly all the ECMs manufactured
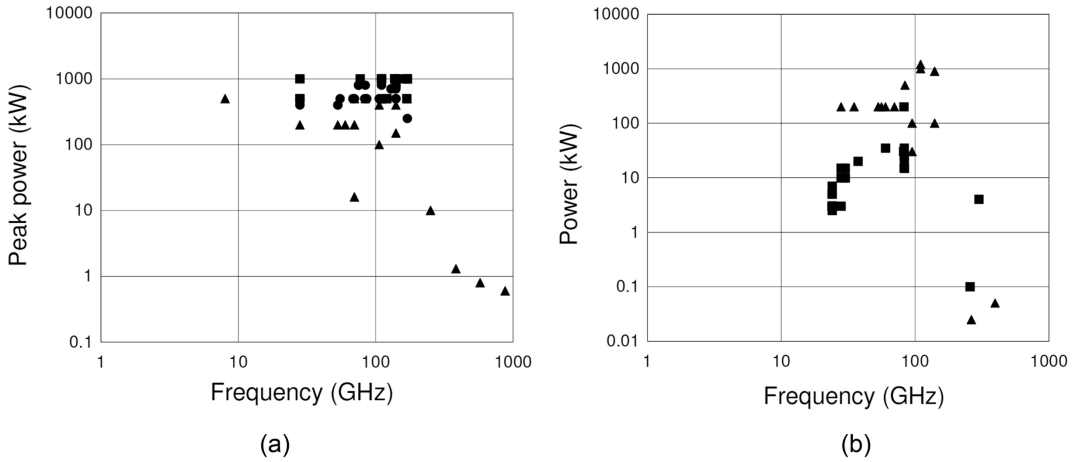
Figure 17.7: Output power of commercial gyrotron oscillators: (a) pulsed tubes, and (b) CW tubes.
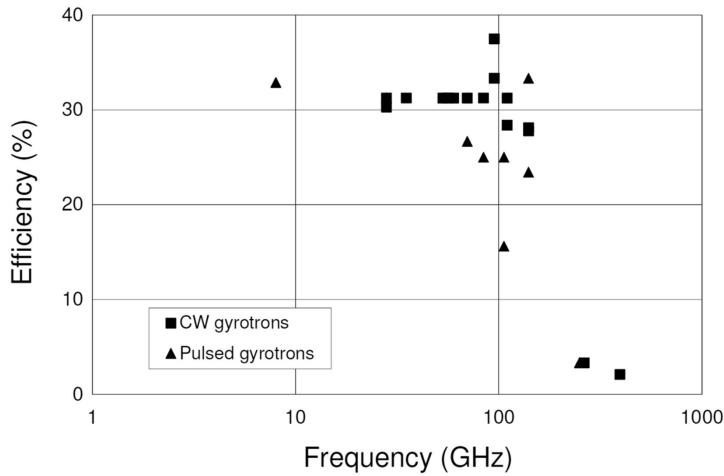
Figure 17.8: Efficiency of commercial gyrotron oscillators.

commercially are gyrotrons. Figure 17.7 shows the power output as a function of frequency for both pulsed and CW tubes based on manufacturers' data sheets. Figure 17.8 summarises the efficiencies of both pulsed and CW tubes as a function of frequency. From these figures it can be seen that the majority of tubes operate at frequencies above about 20 GHz where they are superior to linear-beam tubes. Pulsed and CW powers up to 1 MW are achieved with efficiencies greater than 30% up to a frequency around 200 GHz. Tubes exist with lower output power and efficiency up to 1.3 THz. Further information is available in [4, 24].
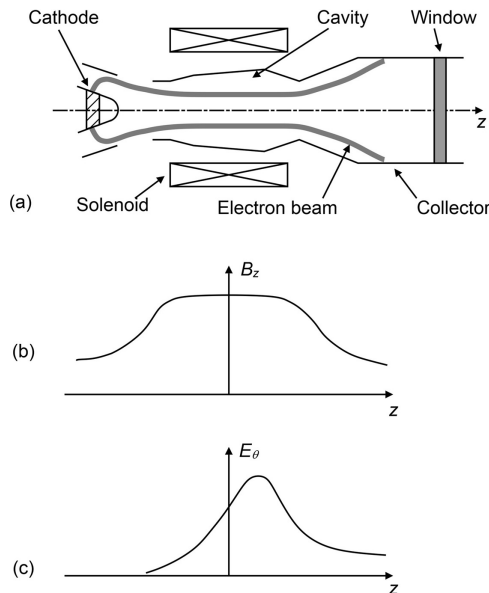
Figure 17.9: (a) Arrangement of a gyrotron oscillator, (b) Magnetic field profile, and (c) RF electric field profile
(copyright 1977, IEEE, reproduced, with permission from [15]).

Figure 17.9(a) shows the basic arrangement of a gyrotron oscillator. A hollow electron beam generated by a magnetron injection gun is injected into a cylindrical cavity resonator which is operated close to the cut-off frequency for longitudinal travelling waves. The beam is controlled by the axial magnetic field of a solenoid which also produces the desired orbits at the cyclotron frequency. For an 80 kV beam the relativistic cyclotron frequency is 44 GHz at a field of 1.8 T, which is around the maximum field which can readily be produced by a normal conducting magnet [25]. This limit is set by the saturation of iron at around 2.1 T. At higher frequencies it is necessary to use a superconducting solenoid. As the beam leaves the cavity it enters a circular waveguide of larger diameter which allows electromagnetic waves to propagate out of the cavity. The magnetic field decreases in this region, as shown in Figure 17.9(b), so that the spent electron beam can be collected on the walls of the waveguide. The RF output power passes through a circular output window into an external waveguide. The resonant modes employed are normally either TE_{0n} circular waveguide modes, or $TE_{m,p}$ whispering gallery modes (where $m \gg 1$ and $m > p$), as shown in Figure 17.10. The longitudinal variation of the azimuthal component of the RF electric field in the cavity is a standing wave whose profile is approximately Gaussian as shown in Figure 17.9(c). Methods for determining the eigenmodes of open cavity resonators are discussed in [10]. Variations from this basic arrangement are described briefly in Section 17.4.
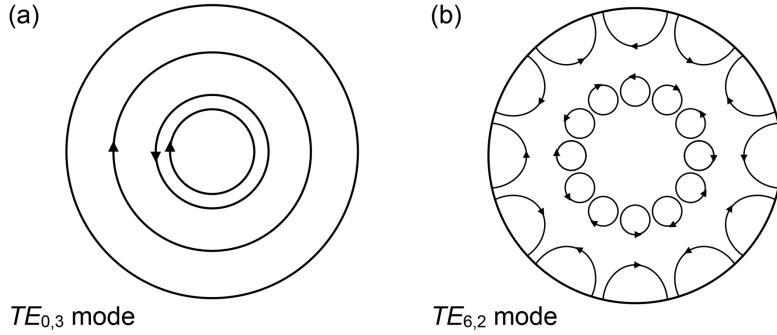
Figure 17.10: Examples of the electric fields of cavity modes used in gyrotron oscillators: (a) symmetric $\left(TE_{0n}\right)$ modes, and (b) whispering gallery $\left(TE_{m,p}\right)$ modes ($m \gg 1$ and $m > p$) (copyright 1997, IEEE, reproduced, with permission, from [12]).

17.3.1   Large-Signal Interaction Model

The small-signal theory of the gyrotron oscillator has been discussed in Section 17.2.1. A different approach is required to find the properties of a tube under large-signal conditions. Particle in Cell (PIC) codes can be used for this purpose but large computer resources and long run-times are needed [26]. Results can be obtained more quickly using codes written specifically for modelling gyrotrons, such as MAGY [27]. A code of this kind finds self-consistent solutions of the electromagnetic fields and the dynamics of the electrons. Since the purpose of this chapter is to provide basic understanding of fast-wave tubes it is convenient to use a simplified approach in which the electromagnetic fields are represented by an approximate function. The problem is then reduced to the solution of the equations of motion of the electrons in the prescribed field [28]. The principles of the method are described below and it is implemented in Worksheet 17.1.

The relativistic equations of motion of an electron in electromagnetic fields are

$$\frac{d}{dt}\left(\gamma m_0 c^2\right) = -e\,\mathbf{u}\cdot\mathbf{E}, \tag{17.28}$$

where γ is defined by (17.3), and

$$\frac{d\mathbf{p}}{dt} = -e\mathbf{E} - \frac{e}{c}\mathbf{v}\times\mathbf{B}, \tag{17.29}$$

where

$$\mathbf{p} = \gamma m_0 \mathbf{u} \tag{17.30}$$

is the momentum of an electron. These equations can be transformed using normalised variables to give

$$\frac{dw}{dZ} = \frac{E_\phi}{sB_z}\,p_\perp \sin\theta \tag{17.31}$$

and

$$\frac{d\theta}{dZ} = \delta_0 - w + \frac{E_r}{B_z}\frac{(1-w)}{p_\perp}\cos\theta,$$

(17.32)

where θ is the angular position of the electron about its gyro-centre,

$$w = 1 - \frac{\gamma}{\gamma_0},$$

(17.33)

$$Z = \frac{\omega}{u_z}z,$$

(17.34)

$$p_\perp = \left(\beta_{\perp0}^2 - 2w + w^2\right)^{\frac{1}{2}},$$

(17.35)

and the detuning parameter is given by

$$\delta_0 = 1 - \frac{s\omega_{c0}}{\omega}.$$

(17.36)

In these equations the subscript 0 is used to indicate the value of the parameter at the entrance to the resonator, $\beta_\perp = u_\theta/c$ and $\beta_\parallel = u_z/c$. A further transformation introduces the independent variable

$$\zeta \equiv \frac{\beta_{\perp0}^2}{2}Z = \pi\frac{\beta_{\perp0}^2}{\beta_{\parallel0}}\cdot\frac{z}{\lambda},$$

(17.37)

where $\lambda = 2\pi c/\omega$, and the dependent variable

$$\upsilon \equiv \frac{2}{\beta_{\perp0}^2}w = \frac{2c^2}{u_{\theta0}^2}\left(1 - \frac{\gamma}{\gamma_0}\right).$$

(17.38)

Note that the symbol υ has been used here for the normalised electron energy parameter to avoid confusion with the use of u for velocities. The normalised interaction length is defined as

$$\mu \equiv \pi\frac{\beta_{\perp0}^2}{\beta_{\parallel0}}\cdot\frac{L}{\lambda},$$

(17.39)

where L is the physical length of the cavity and the detuning parameter is transformed to

$$\Delta \equiv \frac{2\delta_0}{\beta_{\perp0}^2} = \frac{2}{\beta_{\perp0}^2}\left(1 - \frac{s\omega_{c0}}{\omega}\right).$$

(17.40)

Then equations (17.31) and (17.32) become

$$\frac{d\upsilon}{d\zeta} = 2\left(\frac{2^s s!}{s^s\beta_{\perp0}^{s-1}}\right)Ff(\zeta)\frac{p_\perp}{\beta_{\perp0}}J_s'(sp_\perp)\sin\theta$$

(17.41)

and

$$\frac{d\theta}{d\zeta} = \Delta - \upsilon - s\left(\frac{2^s s!}{s^s \beta_{\perp 0}^{s-1}}\right) F f(\zeta) \frac{\beta_{\perp 0}\left(1 - \beta_{\perp 0}^2 \upsilon/2\right)}{p_\perp^2} J_s\left(sp_\perp\right)\cos\theta. \qquad (17.42)$$

Here F is the normalised amplitude of the RF electric field given by

$$F \equiv \frac{E_0}{B_z}\beta_{\perp 0}^{s-4}\left(\frac{s^{s-1}}{s!\,2^{s-1}}\right) J_{m\pm s}\left(x_{m,p}r_0/a\right), \qquad (17.43)$$

where $x_{m,p}$ is the pth non-vanishing zero of $J_m'(x)$ and the azimuthal component of the RF electric field is

$$E_\phi = E_0 f(z) J_m'\left(x_{m,p}r/a\right)\exp\left(-jm\phi\right). \qquad (17.44)$$

The field used is that of a rotating mode and the plus and minus signs in the Bessel function subscript in (17.43) correspond to the two possible directions of rotation.

   The working equations (17.41) and (17.42) can be integrated numerically for a number of electrons with the initial conditions $\upsilon = 0$, and θ distributed uniformly over $(0, 2\pi)$. Most actual axial field profiles are approximated quite closely by a Gaussian distribution. It is therefore assumed that

$$f(\zeta) = \exp-\left(2\zeta/\mu\right)^2. \qquad (17.45)$$

The radius of the orbit of an electron is found from (17.5) to be

$$\frac{r_L}{r_{L0}} = \frac{u_\theta \gamma}{u_{\theta 0}\gamma_0}. \qquad (17.46)$$

Rearranging (17.3) we find that

$$\frac{u_\theta}{c} = \sqrt{1 - \beta_{\|0}^2 - \frac{1}{\gamma^2}}, \qquad (17.47)$$

since it is assumed that the axial component of the electron velocity is unchanged by the interaction. Substitution into (17.46) gives

$$\frac{r_L}{r_{L0}} = \sqrt{\frac{\left(1 - \beta_{\|0}^2\right)\gamma^2 - 1}{\left(1 - \beta_{\|0}^2\right)\gamma_0^2 - 1}}. \qquad (17.48)$$

But, from (17.38)

$$\gamma = \gamma_0\left(1 - \frac{\beta_{\perp 0}^2 \upsilon}{2}\right) \qquad (17.49)$$

so that the radius of the orbit of each electron as the interaction proceeds can be calculated from the values of υ which have been computed.

The electronic efficiency is computed from the change in the kinetic energy of the electrons as they move through the cavity so that

$$\eta_e \equiv \frac{\gamma_0 - \gamma}{\gamma_0 - 1} = \frac{\beta_{\perp 0}^2}{2(1 - 1/\gamma_0)} \eta_{\perp}, \qquad (17.50)$$

where $\eta_{\perp}$ is the value of υ at the output plane averaged over all the electrons. The beam current can be found by making use of the energy balance equation

$$F^2 = \eta_{\perp} I, \qquad (17.51)$$

where I is a normalised current parameter defined by

$$I = \frac{I_0}{I_R}, \qquad (17.52)$$

in which I_0 is the beam current in amps and I_R is a reference current given (in amps) by

$$I_R = \left[0.238 \times 10^{-3} \times \left(\frac{\lambda}{L} \right) \left(\frac{s^s}{s^s s!} \right)^2 \frac{J_{m \pm s}^2 \left(x_{m,p} r_0 / a \right)}{\left(x_{m,p}^2 - m^2 \right) J_m^2 \left(x_{m,p} \right)} \cdot \frac{\beta_{\perp 0}^{2(s-3)}}{\gamma_0} Q_L \right]^{-1}, \qquad (17.53)$$

where Q_L is the loaded Q of the resonator. It can be seen that I_R is a constant, independent of the operating point, for given tube parameters and mode numbers. The overall efficiency is obtained by multiplying the electronic efficiency by the circuit efficiency given by

$$\eta_c = \frac{Q_L}{Q_E}, \qquad (17.54)$$

where Q_L and Q_E are the loaded and external Q factors of the cavity. In the gyrotron literature it is usual to call these the total (Q_T) and diffractive (Q_D) Q factors, and the unloaded (Q_U) is known as the ohmic Q. For consistency the terminology used in the rest of this book is used here. The ohmic losses are typically less than 10% of the output power. The model described above provides a convenient way of investigating the properties, and design trade-offs, of gyrotrons (see Worksheet 17.1).

It can be seen that equations (17.41) and (17.42) which describe the motion of an electron depend upon the three parameters F, μ and Δ. If the first two of these are held constant then it is found that there is a value of Δ for which the transverse efficiency $(\eta_{\perp})$ is maximum. If, in addition, the beam is only weakly relativistic $(s\beta_{\perp 0}^2/2 \ll 1)$ then the transverse efficiency does not depend upon the electron energy. Thus the optimised transverse efficiency depends only on F, μ and the cyclotron harmonic s [28–30]. Figure 17.11 shows the plot for the fundamental $(s = 1)$ cyclotron interaction. When the data from Figure 17.11 is substituted into (17.51) then the result is a plot showing the dependence of transverse efficiency on the normalised current and the interaction length, as shown in Figure 17.12. These plots are useful tools for gyrotron design. Corresponding plots for cyclotron harmonics up to $s = 5$ are given in [28, 30].

We saw in Section 17.2 that a gyrotron will not oscillate until the beam current exceeds a starting value. That analysis ignored the effects of the length of the cavity,
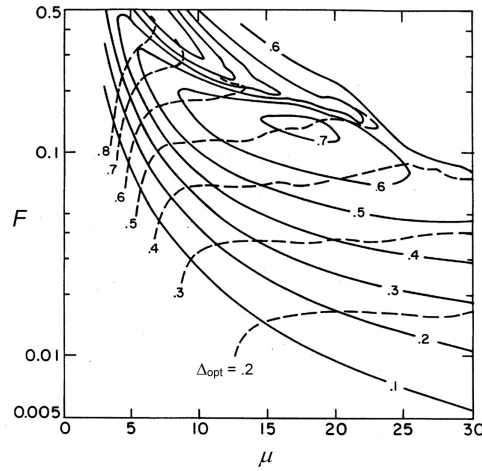
Figure 17.11: Contours of the transverse efficiency of a gyrotron oscillator optimised with respect to Δ plotted against the normalised field strength F and normalised interaction length μ for $s = 1$
(copyright 1985, MIT. Reproduced, with permission, from [30]).
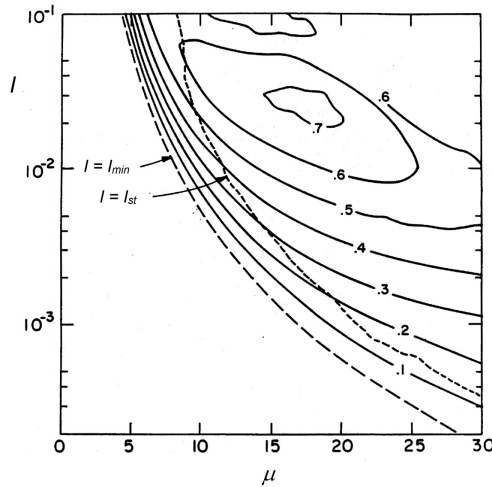
Figure 17.12: Contours of the transverse efficiency of a gyrotron oscillator optimised with respect to Δ plotted against the normalised current I and normalised interaction length μ for $s = 1$
(copyright 1986, AIP, reproduced, with permission, from [28]).

and of the variation of the interaction field along it, so it cannot provide an accurate value for the starting current. An alternative approach, using the large-signal theory above, seeks the limiting current when the field strength parameter F tends to zero. Thus, from (17.51), the normalised starting current is given by

$$I_{st} = \lim_{F \to 0} \left(\frac{F^2}{\eta_\perp} \right). \tag{17.55}$$

This can be calculated for any value of Δ by finding the solution for a sufficiently small value of F. For a given mode, and fixed beam parameters, the value of μ is constant. Then the least possible starting current is given by the value of $I_{\min}$ that corresponds to the value of Δ for which $\eta_{\perp}$ is maximum. The dashed line marked $I = I_{\min}$ in Figure 17.12 shows the locus of this current as μ is varied. Below this line it is not possible for oscillations to start.

If Δ is varied while μ is held constant it can be seen from (17.55) that the value of I_{st} must be greater than $I_{\min}$. The same result is obtained directly using the equation for the starting current given in [28]

$$I_{st} = \left(\frac{4}{\pi\mu} \right)^2 \cdot \frac{\exp\left(2x^2\right)}{\mu x - s},$$

(17.56)

where

$$x \equiv \frac{\mu\Delta}{4}.$$

(17.57)

Figure 17.13 shows a graph of I_{st} against Δ, for three values of μ, for the operating mode of a 1 MW, 140 GHz CW gyrotron [31]. This graph is independent of the cavity mode and of the properties of the beam. The starting current in amps for each mode can be plotted against the magnetic field using (17.40) and (17.52). This plot is useful for determining which mode is excited first for a given magnetic field. It should be noted that there is normally a difference between the starting currents of co- and contra-rotating modes having the same mode numbers. In general the co-rotating mode (negative sign in the Bessel function subscript in (17.53)) is found to have the best coupling to the beam [10].

The normalised starting current can be computed from (17.56) for any point in Figure 17.12. The dashed line marked $I = I_{st}$ shows the contour along which that condition is satisfied. In the region below this line the operating current I is greater
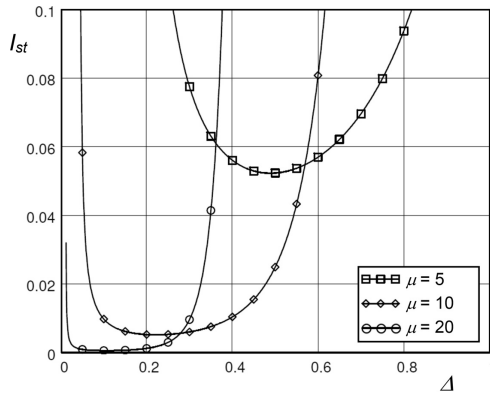
Figure 17.13: Normalised starting current of a gyrotron as a function of normalised detuning parameter.

than the starting current. This region is referred to as the *soft-excitation* region where oscillations will start if the beam current exceeds the starting current. Above the line $I = I_{st}$ lies the hard-excitation region. A tube whose intended operating point lies in this region must be started at a point in the soft-excitation region. The intended operating point is then reached by changing the current, voltage or magnetic field [10, 12, 28].

17.3.2   Case Study: A 140 GHz, 1 MW CW, Gyrotron

To illustrate the use of the theory outlined above let us consider the 140 GHz, 1 MW CW, gyrotron whose design parameters are given in Table 17.1 [31, 32]

The values of the normalised parameters of this tube for start-up and normal operating conditions are given in Table 17.2. In both cases the value of *F* was adjusted to give a beam current of 40 A. Note that the tube is not operating at optimum efficiency at start-up. The computed output power is 895 kW at start-up and 1 MW under normal operating conditions. When the operating point is plotted on Figure 17.12 it can be seen that it lies in the hard excitation region. Potential depression in the beam is important in short-pulse tubes, and in the start-up of long-pulse and CW tubes (see Section 17.3.3). In CW and long pulse operation it is found that the beam space-charge is compensated by the accumulation of ions, after around 100 ms [31]. The reduction of the effective beam voltage at start-up in this tube, caused by space-charge potential depression, ensures that the start-up point lies in the soft excitation region.

Figure 17.14 shows the phases of the electrons relative to the RF field along the length of the interaction region calculated using the normalised parameters in Table 17.2. The beam is modelled by 60 sample electrons, having the same initial

Table 17.1: Design parameters of a 140 GHz, 1 MW CW, gyrotron [31]
(* corrected values [32])

Frequency	140 GHz
Cavity mode	$TE_{28,8}$
Power, pulse length	1 MW, CW (1800 s)
Electron beam radius	10.1 mm*
Beam voltage	81 kV (74.74 kV at start-up)
Beam current	40 A
Velocity ratio α	1.3 (1.5 at start-up *)
Magnetic field	5.56 T
Cavity radius	20.48 mm
Cavity length (L_c)	14.5 mm
Q_0	48650*
Q_L	1150*

Table 17.2: Normalised operating parameters of a 140 GHz CW gyrotron oscillator [31]

	Start-up	Operating
Δ	0.35	0.49
μ	13.0	11.1
F	0.090	0.103
I	0.019	0.020
I_{st}	0.007	0.028
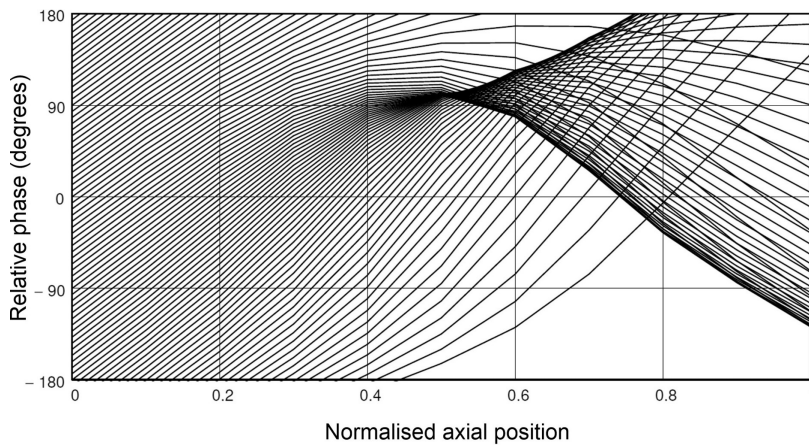
Figure 17.14: Phase of sample electrons plotted against axial position, normalised to the length of the cavity, for a 1 MW, 140 GHz CW gyrotron.

angular velocity, that are uniformly distributed in phase at the input to the cavity. At the centre of the cavity the beam is strongly bunched, but the absence of azimuthal space-charge forces means that there is a large velocity spread so that the bunch disperses by the end of the cavity. A more detailed view of the bunching is shown in Figure 17.15, which shows the azimuthal and radial positions of the electrons at six equally-spaced planes along the length of the cavity. Figure 17.16 shows the evolution of the transverse and electronic efficiencies along the cavity. Note that the maximum transverse efficiency is close to the values at the normal operating point in Figures 17.11 and 17.12.

The non-linear model described above gives results which are useful for exploring the initial design of a gyrotron. Its chief limitation lies in the assumption of a Gaussian standing-wave field in the cavity. In practice the field usually differs appreciably from the Gaussian profile at the output end of the cavity. In this region the wave becomes a travelling wave so that the phase is no longer constant [10, 33]. To get more accurate results it is necessary to find a self-consistent solution for the RF fields and the electron dynamics [12, 27].
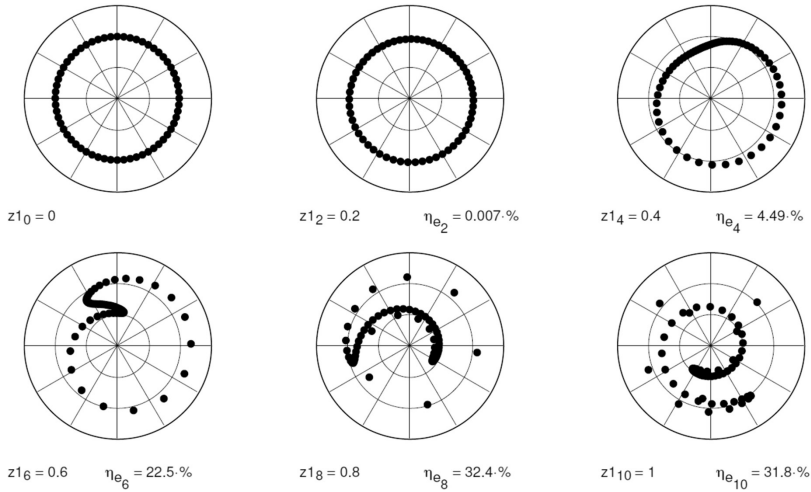
Figure 17.15: Phases and radial positions of sample electrons at equally-spaced planes ($z1 = z/L_c$) within the cavity of a 1 MW, 140 GHz CW gyrotron.
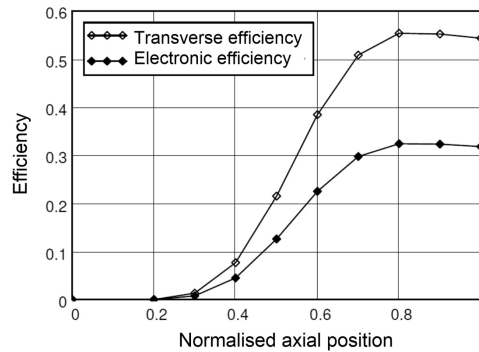
Figure 17.16: Evolution of the transverse and electronic efficiencies plotted against axial position, normalised to the length of the cavity, for a 1 MW, 140 GHz CW gyrotron.

17.3.3 Design of Gyrotron Oscillators

The design of gyrotrons is discussed in detail in the literature and only the principal points are considered here [10, 12, 29, 34, 35]. An example of a design study for a high power gyrotron can be found in [36] and a detailed description of the development of a 1 MW CW, 140 GHz, gyrotron is given in [31, 37, 38].

Factors which are important in the design of a gyrotron are ohmic heating of the cavity walls and stability of the operating mode. The power handling capability of the output window, which used to be a constraint, is of less concern following the adoption of CVD diamond for windows. Gyrotrons are typically designed and operated with parameters in the high efficiency region. It can be seen, from Figures 17.11 and 17.12 that this corresponds to a detuning parameter $\Delta \sim 0.45 - 0.55$ and a cavity length parameter $\mu \sim 15 - 20$. This lies in the hard

excitation region and programmed voltage rise during start-up must be used to access the operating point. The choice of operating point is constrained by a number of factors, especially the average ohmic power density on the cavity walls and the external (diffractive) Q [29]. The average power density should be less than $2\,\mathrm{kW\,cm^{-2}}$. The external Q has been shown to be proportional to $(L/\lambda)^2$ for power extraction from the end of the cavity. Thus, reducing the power density on the walls by increasing the length of the cavity leads to a reduction in the circuit efficiency as more power is trapped within the cavity.

Because the operation of a gyrotron depends upon the resonance between the cyclotron frequency and a resonant frequency of the cavity it is possible to design tubes which employ high-order cavity modes, while still achieving stable single-mode operation. It is found that a large-amplitude mode suppresses its neighbours so that single-frequency operation is possible, even when mode density is high. Symmetric modes (TE_{0n}) have low wall losses and require small cavity radii. Whispering-gallery modes $(TE_{m,p}\,;m \gg 1 \text{ and } m > p)$ have superior mode stability and are preferred for the highest power devices. It can be shown that $(v_{mp}^2 - m^2)$ is a measure of the stored energy in a cavity, and it is desirable for this parameter to be as large as possible. However, as the mode numbers are increased, the density of modes increases, leading to problems of mode selection. It is also found that the external Q increases, producing a decrease in the circuit efficiency. For the 1 MW tube considered above $\sqrt{v_{mp}^2 - m^2} = 53$ [31]. The excitation of possible competing modes can be examined by plotting curves of their starting currents against magnetic field. It is desirable for the operating point to be in a region relatively free from competing modes and for it to have the lowest starting current.

The choice of beam voltage and current depends upon the overall efficiency anticipated. The optimum beam radius for operation in the first cyclotron harmonic is given from (17.44) by

$$r_0 = \frac{x_{m\pm1,1}}{x_{m,p}}a = x_{m\pm1,1}\frac{\lambda}{2\pi}, \tag{17.58}$$

where the plus and minus signs refer to the direction of rotation of the wave [36]. In short-pulse tubes it is necessary to consider the space-charge potential depression in the beam and the limit which that sets on the maximum beam current. The voltage depression for a thin hollow beam is given by [36]

$$\Delta V \approx 60\,\Omega \times \frac{I_0}{\beta_\parallel}\ln\left(\frac{a}{r_0}\right). \tag{17.59}$$

If the beam thickness is too big, the variation in space-charge potential produces variations in γ, in the detuning parameter, and in F, so that $\eta_\perp$ cannot be optimised for all the electrons. Finally, it also increases the risk of multimode operation. Since energy is only extracted from the azimuthal motion of the electrons the transverse efficiency is reduced as shown by (17.50). The ratio of transverse to longitudinal velocity is generally chosen to lie in the range $1 < \alpha < 1.5$ to achieve good electronic
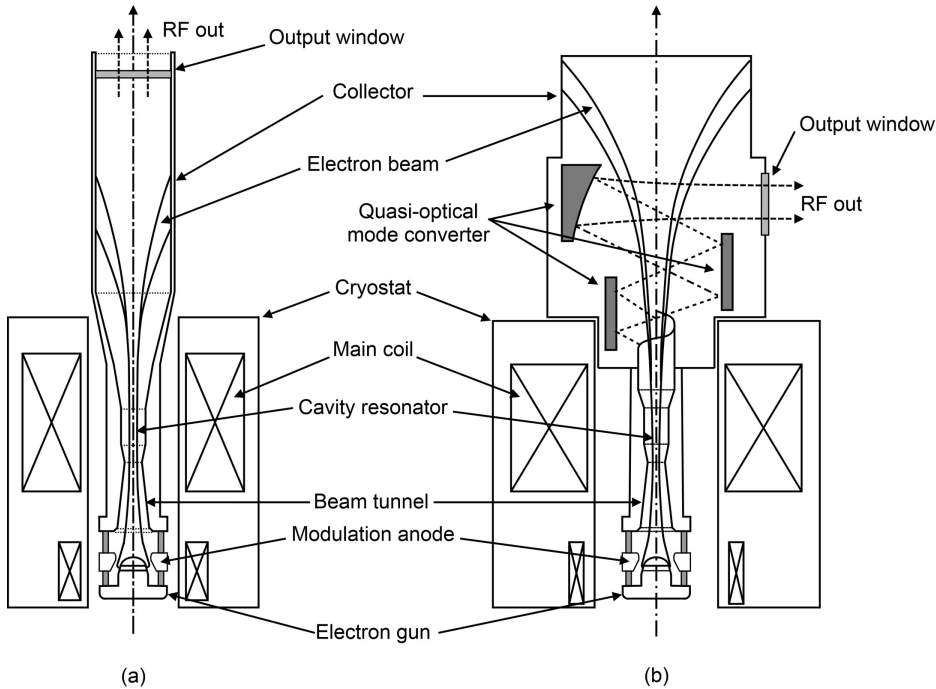
Figure 17.17: Gyrotron output coupling arrangements: (a) over-moded circular waveguide, and (b) quasi-optical mode converter
(copyright 2014, IEEE, reproduced, with permission, from [34]).

efficiency. It is desirable for the velocity spread in the electron beam to be as small as possible. However, though an RMS spread of 6% in the transverse velocity has little effect on efficiency, a10% spread causes an efficiency drop of the order of 20%. A 10% spread in the electron energy was found to cause a 50% drop in efficiency [12]. The electron beam in a gyrotron is usually provided by a magnetron injection gun with a temperature-limited cathode. This permits independent control of the beam current. The development of the gun relies heavily on computer modelling to achieve the required performance.

Figure 17.17 shows alternative output coupling methods used in gyrotrons. In both cases the electron beam is compressed as it leaves the electron gun until it reaches a point where the surrounding waveguide is cut off at the intended operating frequency. Care is taken to suppress parasitic oscillations in this region [34]. The diameter of the waveguide is increased until it reaches the start of the cavity. It is normally constant along the length of the cavity but variations are sometimes used to increase mode stability. At the end of the cavity there is an increasing taper so that the higher-order mode employed can propagate downstream. Because discontinuities in over-moded waveguides can cause unwanted mode conversion, the junctions between the cavity and the tapers at each end are made gradual to minimise mode conversion [10, 34, 39]. Early gyrotrons operating in TE_{0n} modes

employed the output coupling scheme, shown in Figure 17.17(a), in which the collector is a section of over-moded waveguide long enough for all the electrons to be collected before the output window is reached. External converters are then used to convert the unpolarised output mode into a linearly polarised plane wave. Modern high power gyrotrons employ high-order whispering-gallery modes which cannot be converted using conventional waveguide mode converters. Instead, the mode is converted into the fundamental, Gaussian, free space mode $\left(TEM_{0,0}\right)$ by an asymmetrical waveguide launcher and an arrangement of mirrors, as shown in Figure 17.17(b). This is known as a quasi-optical mode converter. The RF window is then on the side of the tube, and the spent beam expands into a collector, which can have a large surface area. This is possible because the conflict between the requirements of the electron collector, and of the output coupler, has been removed. The overall efficiency of the tube can be increased by collector depression [34]. In either case it is important that the spent electrons are collected on the side of the cylindrical collector and do not impinge on the window. A detailed description of the design of the collector for a high power gyrotron is to be found in [40]. For information about the design of depressed collectors for gyrotrons see [41–44].

It is difficult to extend the operation of conventional whispering-gallery mode gyrotrons to higher powers and frequencies because the maximum average power dissipated in the walls of the cavity is determined by the cooling technology available. In theory the average thermal loading can be reduced by going to even higher radial mode indices, but there are then severe problems with mode competition. The use of a coaxial cavity offers a solution to this problem because some of the possible modes are short-circuited by the inner conductor so that the total number of modes is reduced. The modes used are $TE_{m,p}$ modes, with $m \approx 2p$, that are found to give lower ohmic losses. Further information about coaxial gyrotrons is to be found in [45–47].

The need to provide a higher magnetic field as the frequency increases can be offset by working at harmonics of the cyclotron frequency. Theoretical and experimental studies have shown that useful efficiencies should be attainable even at quite high harmonics [28, 33]. The difficulty is, again, that of mode selection because of the possibility that modes involving lower-order harmonics will be excited [35]. This means that, in practice, it is difficult to operate a conventional gyrotron above the second harmonic. One solution to this problem is to employ an electron beam in which the guiding centres lie on the axis of the resonator so that electron orbits encircle the axis. The result is known as a *Large Orbit Gyrotron* in which only co-rotating modes whose azimuthal indices are equal to the harmonic number can be excited $\left(m = s\right)$ [35, 48, 49]. The number of possible modes is then greatly reduced. In this way it has been possible to generate 10 to 20 kW at 410 GHz using the fifth cyclotron harmonic [14]. A further possibility is to replace the cylindrical cavity by one in which the wave is trapped between a pair of opposed, curved mirrors [50, 51]. The result is a $TE_{1,p}$ mode which can be considered as a generalisation of the modes in a cylindrical resonator which has been divided into two halves by axial cuts. This resembles the kind of open resonator used in lasers [52–54]. Tubes of this
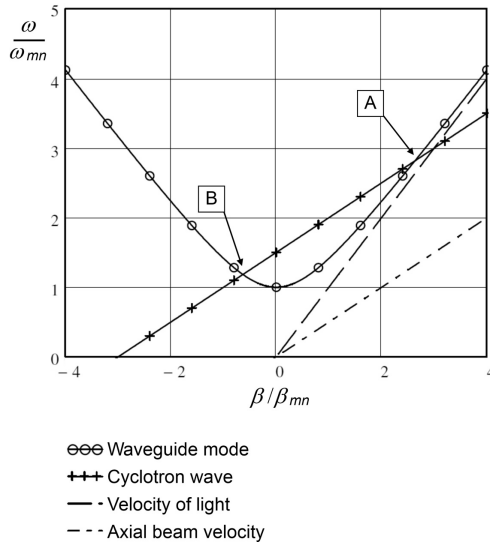
Figure 17.18: Operating points for cyclotron auto-resonance masers (A), and gyro-backward-wave oscillators (B).

kind, known as *Quasi-Optical Gyrotrons*, do not appear to have progressed beyond the experimental stage.

17.3.4   Cyclotron Auto-Resonance Masers

A way of achieving high frequency operation without a very high magnetic field is the Cyclotron Auto-resonance Maser (CARM) which operates at a point such as *A* in Figure 17.18. The frequency at this point is well above both the cut-off frequency of the guide and the cyclotron frequency. The interaction is convective and high efficiency can be achieved. However it is very sensitive to the velocity spread in the beam if the full theoretical efficiency is to be obtained [10, 55].

17.3.5   Tuneable Gyrotrons

The frequency of oscillation of a gyrotron oscillator is very close to the cut-off frequency of the cavity mode selected. Some tuning is possible by changing the magnetic field to excite different cavity modes. Step-tuned gyrotrons with power outputs up to 100 W in the frequency range 75–600 GHz have been developed for plasma diagnostics and the study of nonlinear effects in materials research [56–59]. These tubes employ both first and second harmonic cyclotron resonances. Smooth tuning over a frequency range has been achieved by excitation of successive axial modes. Power of the order of a few watts has been generated over a range of more than 1 GHz at 460 GHz [33, 60]. If the operating point is moved into the backward-wave region, as shown at *B* in Figure 17.18, the interaction is non-convective as in
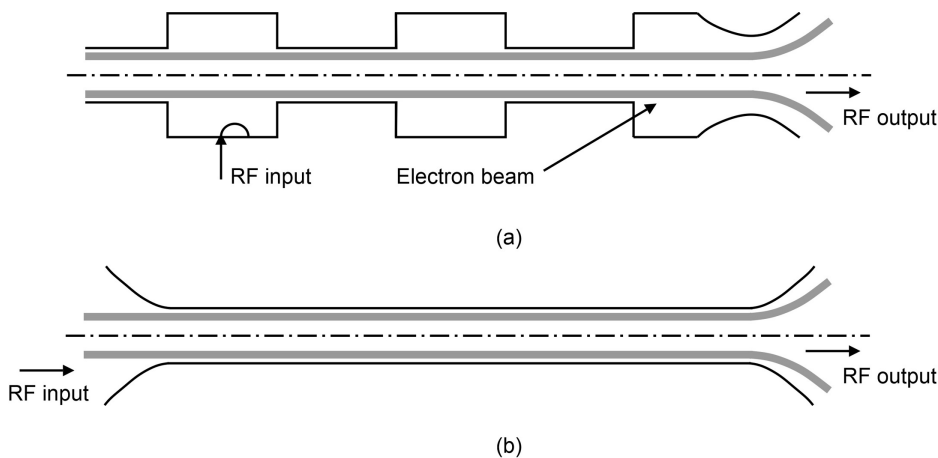
Figure 17.19: Arrangement of gyro-amplifiers: (a) gyro-klystron and (b) gyro-TWA (copyright1997, IEEE, reproduced, with permission, from [12]).

a conventional backward-wave oscillator. Smooth variation of the frequency of operation can be achieved by changing either the beam velocity, or the magnetic field. The frequency of oscillation is below the cyclotron frequency so that a high magnetic field is required for high frequency operation [10, 61–63].

17.4   Gyro-Amplifiers

We saw in Section 17.2 that the interaction between the electron beam and the electromagnetic wave takes the form of a convective instability when the beam current is below the threshold value for oscillation. The dispersion diagram in Figure 17.5 bears a strong resemblance to that for a TWT (see Figure 11.19). It is therefore possible to build gyrotron amplifiers which are analogous to klystrons and TWTs. The development of these tubes is more difficult than the development of oscillators and only a few commercial devices exist. It is essential that they are stable in the absence of RF input and the development of overmoded input couplers is challenging. They must also achieve performance defined by factors such as bandwidth, gain, phase stability, and noise which are not relevant for oscillators [7, 12]. The principal application of these tubes is in radar at 35 and 94 GHz where the attenuation of waves passing through the atmosphere is low.

In a gyro-klystron the electron beam passes through a series of cavities separated by cut-off drift tubes as shown in Figure 17.19(a). A practical, five-cavity, tube produced a peak output power of 100 kW at 94 GHz with a 3 dB bandwidth of 700 MHz, saturated gain of 33 dB and efficiency of about 25% at 10% duty cycle [64–66]. In a gyro-TWA the operating point is moved further above the cut-off frequency of the waveguide so that the growing wave has a group velocity which is not close to zero.
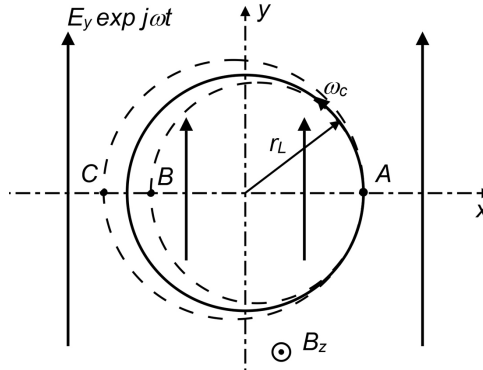
Figure 17.20: Principle of operation of the peniotron.

Figure 17.19(b) shows a schematic view of a gyro-TWA. Practical tubes employing a helically corrugated waveguide have been developed [67–69]. A recent example generated more than 1 kW of continuous power in the range 30–31 GHz, and efficiency greater than 5% [70]. In general, like their linear-beam counterparts, gyro-klystrons have the highest efficiency and gyro-TWAs the greatest bandwidth [12].

17.5   Peniotrons

A different interaction between a fast wave and a periodic electron beam can exist where the RF electric field is not uniform, as shown in Figure 17.20. It is not necessary for the electrons to have relativistic velocities, and the cyclotron frequency is constant. Let us suppose that the signal frequency is twice the cyclotron frequency. An electron which experiences a retarding field at A moves in an orbit with smaller radius, as shown by the inner dashed circle. After half a cyclotron period it reaches B where it encounters an accelerating field. However, because the orbit is smaller than before, this field is weaker than the retarding field at A. There is therefore a net transfer of energy from the electron to the field. Similarly, an electron which is accelerated by the field at A moves in a larger orbit and encounters a retarding field at C that is stronger than the field at A. Once again there is a net transfer of energy to the RF wave. This process continues in successive cyclotron periods so that the guiding centres drift away from the original centre of the beam. A device operating on this principle is known as a *Peniotron*.

   The peniotron is a cyclotron resonance device, like the gyrotron, but with the important difference that energy is taken from all the electrons, and no phase bunching takes place. Thus, it is possible for the interaction to be highly efficient. Also, the phase of the RF field seen by an electron varies by 2π as it moves once around its orbit. Thus the resonant condition becomes

$$\omega t - \beta z = (s \pm 1)\,2\pi, \tag{17.60}$$

which may be compared with the corresponding equation for the gyrotron (17.11). Hence, substituting for t and z using (17.9) and (17.10) we obtain

$$\omega - \beta u_z = (s \pm 1)\omega_c. \tag{17.61}$$

Then, by analogy with (17.16) we expect the small-signal dispersion relation to take the form

$$\left(\omega^2 - c^2\left(\beta^2 + \beta_{m,n}^2\right)\right)\left(\omega - \beta u_z - (s \pm 1)\omega_c\right)^2 = K\, f\left(\omega, \beta\right). \tag{17.62}$$

Hence the dispersion diagram for a peniotron is identical to that for a gyrotron, shown in Figure 17.2 and the gyrotron and peniotron mechanisms may compete with one another in the same device [71].

The original experimental peniotron employed an axis-encircling hollow beam within a double-ridged waveguide propagating the TE_{10} mode [72–75]. More recent tubes have employed cavities with longitudinal vanes which resemble unstrapped magnetron cavities. This type of cavity is more suited to high harmonic peniotron operation because it is smaller, permitting lower voltage operation, and because it provides better mode selection. Large signal modelling of both gyrotron and peniotron interactions in a magnetron type cavity show that the fast peniotron interaction (the positive sign in (17.61)) has very poor efficiency. Very good efficiency is possible for the slow peniotron harmonics $\left((s-1), (s-2)\ldots\right)$ but it decreases rapidly with increasing spread in the positions of the guiding centres [76, 77]. The peniotron is therefore chiefly of interest because of its potential for delivering high efficiency at high harmonic numbers and, therefore, for high-frequency operation without the need for a superconducting magnet. It has been the subject of a number of theoretical and experimental studies (see for example: [78–84]) but the experimental results were generally disappointing. A 35 GHz third harmonic oscillator, which generated 6.9 kW of RF output power, demonstrated the achievement of a theoretical electronic efficiency of 75% but the conversion efficiency was only 39% because of high circuit losses [83]. The same paper reported results for operation in the tenth cyclotron harmonic at 100 GHz. The highest electronic conversion efficiency was 2.3% at an output power of 160 W, corresponding to an electronic efficiency of 6%. At the time of writing these seem to have been the best results obtained.

17.6   Ubitrons (Free Electron Lasers)

The properties of an electron beam can also be made periodic by passing it through a periodic magnetic field, as shown in Figure 17.21 [85]. Figure 17.21(a) shows a sheet electron beam within a rectangular waveguide. The periodic transverse magnetic field causes the electron trajectories to follow undulating paths in the y-z plane. This motion can be coupled to the RF electric field of the TE_{10} mode of the waveguide. A microwave tube employing this principle is known as an *Ubitron* (undulating beam interaction) [86]. Figures 17.21(b) and (c) show alternative arrangements
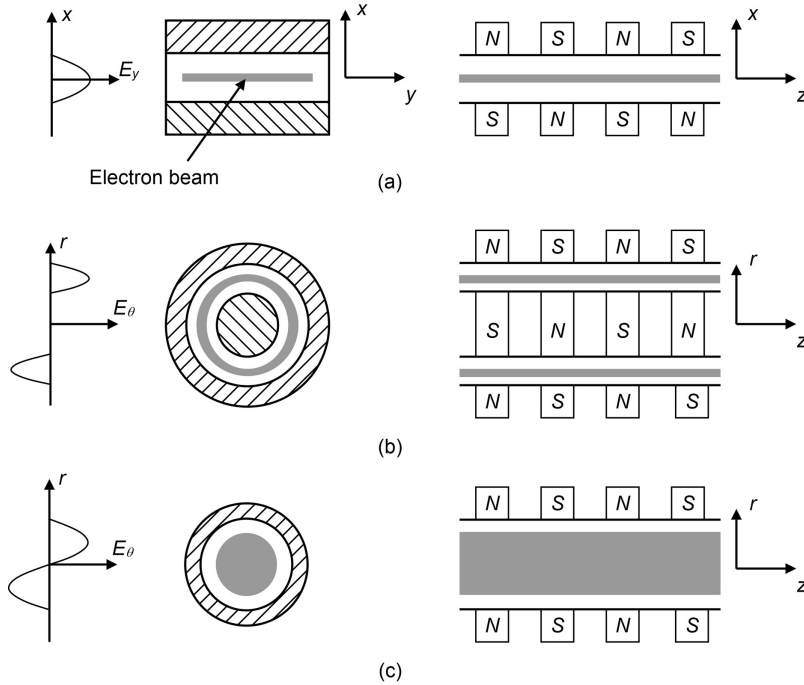
Figure 17.21: Arrangements of ubitrons: (a) planar, (b) coaxial and (c) cylindrical (copyright 1960, IRE, reproduced, with permission, from [86]).

of such a tube where the beam is focused by a periodic permanent magnet field. The arrangement of electro- or permanent magnets to produce an undulating beam is known as a *wiggler* (or *undulator*).

If space-charge forces are ignored, the propagation constants of the space-harmonics of the waves on the beam are given by

$$\beta_v = \frac{\omega}{u_z} \pm \frac{2v\pi}{\lambda_w}, \tag{17.63}$$

where λ_w is the wavelength of the wiggler and $v = 0, 1, 2$, etc. The strongest interaction with the field of the waveguide is that of the $v = -1$ space harmonic for which the dispersion equation takes the form

$$\left(\omega^2 - c^2\left(\beta^2 + \beta_{m,n}^2\right)\right)\left(\omega - \left(\beta + \beta_w\right)u_z\right)^2 = K f\left(\omega, \beta\right), \tag{17.64}$$

where $\beta_w = 2\pi/\lambda_w$. The dispersion diagram is then, as shown in Figure 17.22, where β_w and u_z have been chosen so that the space-harmonic line is tangential to the waveguide dispersion curve. It can be shown that the condition for tangential operation is

$$\beta_w = \omega_{m,n} \sqrt{\frac{1}{u_z^2} - \frac{1}{c^2}}. \tag{17.65}$$
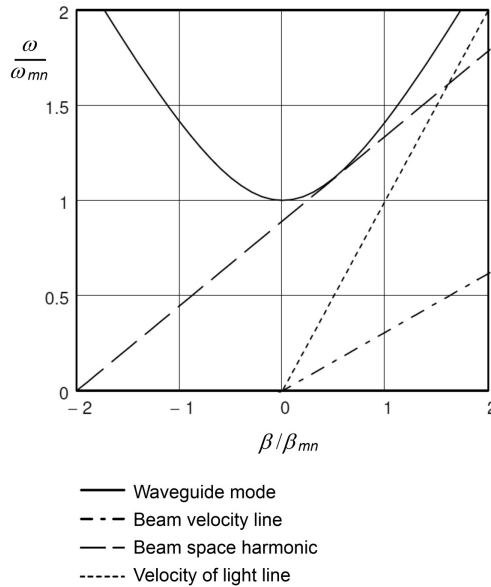
Figure 17.22: Uncoupled dispersion diagram for an ubitron.

The magnitude of the term under the square root is set by the beam velocity. The frequency of operation is slightly above the cut-off frequency of the waveguide mode $(\omega_{m,n})$, so the wavelength of the wiggler decreases as the frequency increases. The maximum frequency of operation at a given voltage is determined by the minimum practical period of the wiggler. Since the field of the wiggler is essentially governed by Laplace's equation, the transverse distance between the pole-pieces needs to be less than their axial separation, in order to maintain a high transverse component of the magnetic field. Thus, the transverse space available for the beam decreases as the frequency increases. For example, at a beam voltage of 80 kV $(u_z/c \sim 0.5)$, and a cut-off frequency of 50 GHz, it is found that $\lambda_w \sim 3.5$ mm. This is probably around the minimum practical value. Operation at higher frequencies requires the use of higher beam voltages.

The first experimental ubitron used the arrangement in Figure 17.21(a) and achieved a maximum gain of 13 dB at about 2.6 GHz with a beam voltage of 135 kV [86]. The peak output power was 1.2 MW at a beam voltage of 170 kV with an efficiency of 10%. Subsequent tubes, using the arrangement in Figure 17.21(c), delivered 1.65 MW pk at 15.75 GHz with a solid beam, and 150 kW pk at 54 GHz with a hollow beam [85]. Another tube gave a peak output power of 4.2 MW at 16.6 GHz with 17.5% efficiency, 29 dB gain and 22% instantaneous bandwidth [87]. Ubitrons tend to suffer from poor beam transmission, problems with excitation of lower modes, and high sensitivity to changes in the beam voltage and the magnetic field. Information about the design, and in some cases, the performance of a number of other tubes can be found in [88–92].

Ubitrons work at relatively low velocities where the interaction between the electrons and the waves is analogous to that in a TWT. It is interesting to note that computer modelling of this interaction shows axial bunching of the beam similar to that in a TWT. If the velocities of the electrons approach the speed of light, then they lose appreciable energy by radiation, as a result of transverse acceleration within the wiggler. When this process takes place in a strong RF field then the radiation emitted is phase locked to that field and the device becomes a coherent power source. A device which works in this way is known as a Free Electron Laser (FEL). However, despite its name, the FEL is not a quantum device because the electrons are free, and therefore do not have quantised energies. The operation of an FEL can be explained entirely by classical theory. In an FEL the beam is provided by an accelerator, and coherent radiation can be generated at wavelengths from the infra-red to the X-ray region of the electromagnetic spectrum. These devices lie outside the scope of this book but further information can be found in [4, 6, 8].

References

[1] M. I. Petelin, 'One century of cyclotron radiation', *IEEE Transactions on Plasma Science,* vol. 27, pp. 294–302, 1999.

[2] J. H. Booske *et al.*, 'Vacuum electronic high power terahertz sources', *IEEE Transactions on Terahertz Science and Technology,* vol. 1, pp. 54–75, 2011.

[3] J. Feinstein and K. Felch, 'Status review of research on millimeter-wave tubes', *IEEE Transactions on Electron Devices,* vol. 34, pp. 461–467, 1987.

[4] G. Faillon *et al.*, 'Microwave tubes', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer-Verlag, pp. 1–84, 2008.

[5] V. L. Granatstein and I. Alexeff, *High-Power Microwave Sources*. Boston, MA: Artech House, 1987.

[6] J. A. Pasour, 'Free-electron lasers', *IEEE Circuits and Devices Magazine,* vol. 3, pp. 55–64, 1987.

[7] K. L. Felch *et al.*, 'Characteristics and applications of fast-wave gyrodevices', *Proceedings of the IEEE,* vol. 87, pp. 752–781, 1999.

[8] H. P. Freund and G. R. Neil, 'Free-electron lasers: vacuum electronic generators of coherent radiation', *Proceedings of the IEEE,* vol. 87, pp. 782–803, 1999.

[9] B. G. Danly *et al.*, 'Gyro-amplifiers', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 247–287, 2005.

[10] M. V. Kartikeyan *et al.*, *Gyrotrons: High-Power Microwave and Millimeter Wave Technology*. Berlin: Springer-Verlag, 2010.

[11] K. Chu, 'The electron cyclotron maser', *Reviews of Modern Physics,* vol. 76, p. 489, 2004.

[12] V. L. Granatstein *et al.*, 'A quarter century of gyrotron research and development', *IEEE Transactions on Plasma Science,* vol. 25, pp. 1322–1335, 1997.

[13] G. Nusinovich *et al.*, 'The gyrotron at 50: historical overview', *Journal of Infrared, Millimeter, and Terahertz Waves,* vol. 35, pp. 325–381, 2014.

[14] M. Y. Glyavin *et al.*, 'Terahertz gyrotrons: state of the art and prospects', *Journal of Communications Technology and Electronics,* vol. 59, pp. 792–797, 2014.

[15] V. A. Flyagin *et al.*, 'The gyrotron', *IEEE Transactions on Microwave Theory and Techniques,* vol. 25, pp. 514–521, 1977.

[16] P. A. Sturrock, 'Kinematics of growing waves', *Physical Review,* vol. 112, pp. 1488–1503, 1958.

[17] R. J. Briggs, *Electron-Stream Interaction with Plasmas*. Cambridge, MA: MIT Press, 1964.

[18] K. R. Chu *et al.*, 'Characteristics and optimum operating parameters of a gyrotron traveling wave amplifier', *IEEE Transactions on Microwave Theory and Techniques,* vol. 27, pp. 178–187, 1979.

[19] A. J. Sangster, 'Small-signal analysis of the travelling-wave gyrotron using Pierce parameters', *IEE Proceedings I: Solid-State and Electron Devices,* vol. 127, pp. 45–52, 1980.

[20] P. Lindsay, 'Gyrotrons (electron cyclotron masers): different mathematical models', *IEEE Journal of Quantum Electronics,* vol. 17, pp. 1327–1333, 1981.

[21] Y. Y. Lau, 'Simple macroscopic theory of cyclotron maser instabilities', *IEEE Transactions on Electron Devices,* vol. 29, pp. 320–335, 1982.

[22] J. R. Pierce, *Traveling-Wave Tubes*. Princeton, NJ: D. van Nostrand, 1950.

[23] K. R. Chu and A. T. Lin, 'Gain and bandwidth of the gyro-TWT and CARM amplifiers', *IEEE Transactions on Plasma Science,* vol. 16, pp. 90–104, 1988.

[24] M. Thumm, *State-of-the-Art of High Power Gyro-Devices and Free Electron Masers. Update 2015 (KIT Scientific Reports; 7717)*, vol. 7717. KIT Scientific Publishing, 2016.

[25] D. Borodin and M. Einat, 'Copper solenoid design for the continuous operation of a second harmonic 95-GHz gyrotron', *IEEE Transactions on Electron Devices,* vol. 61, pp. 3309–3316, 2014.

[26] J. Neudorfer *et al.*, 'Efficient parallelization of a three-dimensional high-order particle-in-cell method for the simulation of a 170 GHz gyrotron resonator', *IEEE Transactions on Plasma Science,* vol. 41, pp. 87–98, 2013.

[27] M. Botton *et al.*, 'MAGY: a time-dependent code for simulation of slow and fast microwave sources', *IEEE Transactions on Plasma Science,* vol. 26, pp. 882–892, 1998.

[28] B. Danly and R. J. Temkin, 'Generalized nonlinear harmonic gyrotron theory', *Physics of Fluids,* vol. 29, pp. 561–567, 1986.

[29] K. E. Kreischer *et al.*, 'The design of megawatt gyrotrons', *IEEE Transactions on Plasma Science,* vol. 13, pp. 364–373, 1985.

[30] B. Danly and R. J. Temkin, *Generalized Nonlinear Harmonic Gyrotron Theory*, MIT, Cambridge MA 1070–6631, 1985.

[31] M. Thumm *et al.*, 'EU megawatt-class 140-GHz CW gyrotron', *IEEE Transactions on Plasma Science,* vol. 35, pp. 143–153, 2007.

[32] M. Thumm, 'Private communication', 2015.

[33] M. K. Hornstein *et al.*, 'Second harmonic operation at 460 GHz and broadband continuous frequency tuning of a gyrotron oscillator', *IEEE Transactions on Electron Devices,* vol. 52, pp. 798–807, 2005.

[34] M. Thumm, 'Recent advances in the worldwide fusion gyrotron development', *IEEE Transactions on Plasma Science,* vol. 42, pp. 590–599, 2014.

[35] V. Bratman *et al.*, 'Review of Subterahertz and Terahertz Gyrodevices at IAP RAS and FIR FU', *IEEE Transactions on Plasma Science,* vol. 37, pp. 36–43, 2009.

[36] M. V. Kartikeyan *et al.*, 'Possible operation of a 1.5-2-MW, CW conventional cavity gyrotron at 140 GHz', *IEEE Transactions on Plasma Science,* vol. 28, pp. 645–651, 2000.

[37] G. Dammertz *et al.*, 'Development of a 140-GHz 1-MW continuous wave gyrotron for the W7-X stellarator', *IEEE Transactions on Plasma Science,* vol. 30, pp. 808–818, 2002.

[38] S. Alberti *et al.*, 'European high-power CW gyrotron development for ECRH systems', *Fusion Engineering and Design,* vol. 53, pp. 387–397, 2001.

[39] M. K. Thumm and W. Kasparek, 'Passive high-power microwave components', *IEEE Transactions on Plasma Science,* vol. 30, pp. 755–786, 2002.

[40] K. T. Nguyen *et al.*, 'Electron gun and collector design for 94-GHz gyro-amplifiers', *IEEE Transactions on Plasma Science,* vol. 26, pp. 799–813, 1998.

[41] M. E. Read *et al.*, 'Depressed collectors for high-power gyrotrons', *IEEE Transactions on Electron Devices,* vol. 37, pp. 1579–1589, 1990.

[42] G. P. Saraph *et al.*, 'Design of a single-stage depressed collector for high-power, pulsed gyroklystron amplifiers', *IEEE Transactions on Electron Devices,* vol. 45, pp. 986–990, 1998.

[43] R. L. Ives *et al.*, 'Design of a multistage depressed collector system for 1 MW CW gyrotrons. II. System consideration', *IEEE Transactions on Plasma Science,* vol. 27, pp. 503–511, 1999.

[44] A. Singh *et al.*, 'Design of a multistage depressed collector system for 1-MW CW gyrotrons. I. Trajectory control of primary and secondary electrons in a two-stage depressed collector', *IEEE Transactions on Plasma Science,* vol. 27, pp. 490–502, 1999.

[45] M. E. Read *et al.*, 'Design of a 3-MW 140-GHz gyrotron with a coaxial cavity', *IEEE Transactions on Plasma Science,* vol. 24, pp. 586–595, 1996.

[46] R. Advani *et al.*, 'Experimental investigation of a 140-GHz coaxial gyrotron oscillator', *IEEE Transactions on Plasma Science,* vol. 29, pp. 943–950, 2001.

[47] O. Dumbrajs and G. S. Nusinovich, 'Coaxial gyrotrons: past, present, and future (review) ', *IEEE Transactions on Plasma Science,* vol. 32, pp. 934–946, 2004.

[48] V. L. Bratman *et al.*, 'Large-orbit gyrotron operation in the terahertz frequency range', *Physical Review Letters,* vol. 102, p. 245101, 2009.

[49] V. L. Bratman *et al.*, 'Moderately relativistic high-harmonic gyrotrons for millimeter/submillimeter wavelength band', *IEEE Transactions on Plasma Science,* vol. 27, pp. 456–461, 1999.

[50] A. W. Fliflet *et al.*, 'Review of quasi-optical gyrotron development', *Journal of Fusion Energy,* vol. 9, pp. 31–58, 1990.

[51] B. Levush and W. M. Manheim, 'Generation of high-frequency radiation by quasi-optical gyrotron at harmonics of the cyclotron frequency', *IEEE Transactions on Microwave Theory and Techniques,* vol. 32, pp. 1398–1401, 1984.

[52] R. G. Carter, 'Synthesis of the fields of barrel open resonators', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 657, pp. 1–5, 2011.

[53] R. N. Clarke and C. B. Ronseberg, 'Fabry-Perot and open resonators at microwave and millimetre wave frequencies, 2–300 GHz', *J. Phys. E.: Sci. Instrum.,* vol. 15, pp. 9–24, 1982.

[54] H. Kogelnik and T. Li, 'Laser beams and resonators', *Applied Optics,* vol. 5, pp. 1550–1567, 1966.

[55] V. Bratman *et al.*, 'FEL's with Bragg reflection resonators: cyclotron autoresonance masers versus ubitrons', *IEEE Journal of Quantum Electronics,* vol. 19, pp. 282–296, 1983.

[56] G. F. Brand and M. Gross, 'A tunable source of linearly-polarized, near-millimeter wave radiation', *International Journal of Infrared and Millimeter Waves,* vol. 10, pp. 121–136, 1989.

[57] K. D. Hong *et al.*, 'Submillimeter wave generation by second harmonic operation of tunable gyrotrons', *International Journal of Infrared and Millimeter Waves,* vol. 13, pp. 215–227, 1992.

[58] K. Hong *et al.*, 'A 150–600 GHz step-tunable gyrotron', *Journal of Applied Physics,* vol. 74, pp. 5250–5258, 1993.

[59] T. Idehara *et al.*, 'Development of a high-frequency, second-harmonic gyrotron tunable up to 636 GHz', *Physics of Fluids B: Plasma Physics (1989–1993),* vol. 5, pp. 1377–1379, 1993.

[60] A. C. Torrezan *et al.*, 'Operation of a continuously frequency-tunable second-harmonic CW 330-GHz gyrotron for dynamic nuclear polarization', *IEEE Transactions on Electron Devices,* vol. 58, pp. 2777–2783, 2011.

[61] S. Y. Park *et al.*, 'Experimental study of a Ka-band gyrotron backward-wave oscillator', *IEEE Transactions on Plasma Science,* vol. 18, pp. 321–325, 1990.

[62] S. V. Samsonov *et al.*, 'Frequency-tunable CW gyro-BWO with a helically rippled operating waveguide', *IEEE Transactions on Plasma Science,* vol. 32, pp. 884–889, 2004.

[63] B.-T. Liu *et al.*, 'Experimental study of a Ku-band gyrotron backward-wave oscillator with a single stage depressed collector', *IEEE Transactions on Plasma Science,* vol. 35, pp. 1065–1069, 2007.

[64] B. G. Danly *et al.*, 'Development and testing of a high-average power, 94-GHz gyroklystron', *IEEE Transactions on Plasma Science,* vol. 28, pp. 713–726, 2000.

[65] M. Garven *et al.*, 'Experimental studies of a four-cavity, 35 GHz gyroklystron amplifier', *IEEE Transactions on Plasma Science,* vol. 28, pp. 672–680, 2000.

[66] B. G. Danly, 'Gyro-amplifiers for high power millimeter wave radar', in *Third IEEE International Vacuum Electronics Conference*, Monterey, CA, pp. 361–362, 2002.

[67] A. Cross *et al.*, 'Helically corrugated waveguide gyrotron traveling wave amplifier using a thermionic cathode electron gun', *Applied Physics Letters,* vol. 90, p. 253501, 2007.

[68] V. Bratman *et al.*, 'High-gain wide-band gyrotron traveling wave amplifier with a helically corrugated waveguide', *Physical Review Letters,* vol. 84, p. 2746, 2000.

[69] G. Denisov *et al.*, 'Gyrotron traveling wave amplifier with a helical interaction waveguide', *Physical Review Letters,* vol. 81, p. 5680, 1998.

[70] S. V. Samsonov *et al.*, 'CW Ka-band kilowatt-level helical-waveguide gyro-TWT', *IEEE Transactions on Electron Devices,* vol. 59, pp. 2250–2255, 2012.

[71] G. Dohler and D. Gallagher, 'The small-signal theory of the cyclotron maser and other gyrotron-type devices', *IEEE Transactions on Electron Devices,* vol. 35, pp. 1730–1745, 1988.

[72] G. Dohler *et al.*, 'The peniotron: fast wave device for efficient high power mm-wave generation', in *International Electron Devices Meeting*, pp. 400–403, 1978.

[73] G. Dohler and B. Wilson, 'A small signal theory of the peniotron', in *International Electron Devices Meeting*, pp. 810–813, 1980.

[74] G. Dohler *et al.*, 'Peniotron oscillator operating performance', in *International Electron Devices Meeting*, pp. 328–331, 1981.

[75] G. Dohler *et al.*, 'Peniotron amplifier results', in *International Electron Devices Meeting*, pp. 845–848, 1984.

[76] P. S. Rha *et al.*, 'Self-consistent large theory and simulation of high harmonic gyrotron and peniotron oscillators operating in a magnetron-type cavity', in *International Electron Devices Meeting*, pp. 535–538, 1985.

[77] P. S. Rha *et al.*, 'Self-consistent simulation of harmonic gyrotron and peniotron oscillators operating in a magnetron-type cavity', *IEEE Transactions on Electron Devices,* vol. 36, pp. 789–801, 1989.

[78] H. Shimawaki *et al.*, '2nd cyclotron harmonic peniotron experiments', in *International Electron Devices Meeting*, pp. 787–790, 1991.

[79] G. Dohler *et al.*, 'Harmonic high power 95 GHz peniotron', in *International Electron Devices Meeting*, pp. 363–366, 1993.

[80] T. Ishihara *et al.*, 'Experiments of 10th cyclotron harmonic peniotron oscillator', in *International Electron Devices Meeting*, pp. 367–370, 1993.

[81] A. K. Ganguly *et al.*, 'Nonlinear theory of harmonic peniotron and gyrotron interactions in a rising-sun slotted waveguide', *IEEE Transactions on Plasma Science,* vol. 22, pp. 902–912, 1994.

[82] T. Ishihara *et al.*, 'Space harmonic peniotron in a magnetron waveguide resonator', *IEEE Transactions on Electron Devices,* vol. 43, pp. 827–833, 1996.

[83] T. Ishihara *et al.*, 'Highly efficient operation of space harmonic peniotron at cyclotron high harmonics', *IEEE Transactions on Electron Devices,* vol. 46, pp. 798–802, 1999.

[84] D. B. McDermott *et al.*, 'Efficient Ka-band second-harmonic slotted peniotron', *IEEE Transactions on Plasma Science,* vol. 28, pp. 953–958, 2000.

[85] R. M. Phillips, 'History of the ubitron', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 272, pp. 1–9, 1988.

[86] R. M. Phillips, 'The ubitron, a high-power traveling-wave tube based on a periodic beam interaction in unloaded waveguide', *IRE Transactions on Electron Devices,* vol. 7, pp. 231–241, 1960.

[87] D. E. Pershing *et al.*, 'Amplifier performance of the NRL ubitron', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 358, pp. 104–107, 1995.

[88] H. Bluem *et al.*, 'Demonstration of a new free-electron laser harmonic interaction', in *International Electron Devices Meeting*, pp. 791–794, 1991.

[89] A. J. Balkcum *et al.*, '250-MW X-band TE_{01} ubitron using a coaxial PPM wiggler', *IEEE Transactions on Plasma Science,* vol. 24, pp. 802–807, 1996.

[90] A. J. Balkcum *et al.*, 'High-power coaxial ubitron oscillator: theory and design', *IEEE Transactions on Plasma Science,* vol. 26, pp. 548–555, 1998.

[91] H. P. Freund *et al.*, 'Designs for W-band free-electron masers', *IEEE Transactions on Plasma Science,* vol. 27, pp. 243–253, 1999.

[92] H. P. Bluem *et al.*, 'A compact, high-power THz source', in *IEEE International Conference on Plasma Science*, p. 7B-7, 2012.

18 Emission and Breakdown Phenomena

18.1 Introduction

This chapter reviews the variety of emission and breakdown phenomena which occur, intentionally or unintentionally, in vacuum tubes. The underlying physics of each process is summarised but more specialised texts (e.g. [1]) should be consulted for the details. In order for electrons to be released from a solid surface into a vacuum they must acquire sufficient energy to overcome a potential barrier. This energy can be provided by heat, by the impact of electrons or ions, or by irradiation by short-wavelength light. The nature of the potential barrier depends upon the material and also upon the external electric field. These processes are important for the cathode, which is the source of the electrons employed in a tube, and also because of their role in initiating DC and RF voltage breakdown in vacuum. Secondary electrons, liberated by the impact of primary electrons, affect the performance of multi-element depressed collectors (see Chapter 10). They also play an important role in crossed-field tubes (see Chapters 15 and 16), and in sustaining multipactor discharges (Section 18.8). Outside a tube it is important to avoid breakdown across high voltage insulators, and in waveguides. Bombardment of solid surfaces by high energy electrons leads to the production of X-rays which can be a hazard in the operation of high power tubes.

18.2 Emission of Electrons from Metal Surfaces

The discussion of the emission of electrons from metals in a vacuum begins with a review of the free electron model of a metal [1]. The conduction electrons, which are free to move within the metal, are considered to occupy a large number of closely spaced, quantised energy states, within a potential well. According to the Pauli Exclusion Principle no two electrons can have the same quantum numbers (including spin). Thus, each state can contain at most two electrons with opposing spin. At the absolute zero of temperature $(T = 0 \text{ K})$ the electrons fill the lowest possible states up to a maximum energy known as the *Fermi Energy* (E_F). All the states above the Fermi Energy are then empty. If the electron energy of the bottom of the potential well is taken to be zero, then the Fermi Energy is of the order of a
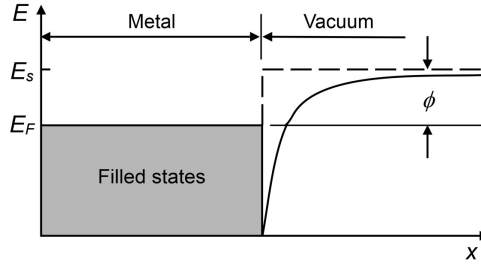
Figure 18.1: Free electron model of a metal.

few electron volts for typical metals. The Fermi Energy lies somewhat below that of stationary free electrons outside the metal (E_s), as shown in Figure 18.1. This energy difference, known as the *Work Function* (ϕ), depends upon the metal. It is typically in the range 2 to 6 eV but is strongly dependent upon the condition of the surface, including any contamination. Tables of work functions for metals can be found in [2, 3].

In the free electron model the electrons are assumed to be trapped in a rectangular potential well, as shown by the dashed line in Figure 18.1. If an electron escapes from the well, and its distance from the surface (x) is large compared with the atomic spacing, it moves in the potential of a positive image charge. This potential, referred to the potential at the surface of the metal in the absence of emission, is given by

$$V = \frac{e}{16\pi\varepsilon_0 x}. \tag{18.1}$$

The energy of the electron in this potential is shown by the solid curve in Figure 18.1. At a distance of 1 μm (3000–4000 atomic spacings) from the surface the potential is 0.36 mV below the external potential. This model does not hold very close to the surface where the atomic structure becomes important.

In general, the probability that a quantum state with energy E is occupied is given by the Fermi-Dirac distribution function

$$p(E) = \frac{1}{1 + \exp\left(\dfrac{E - E_F}{kT}\right)}, \tag{18.2}$$

where T is the absolute temperature and k is Boltzmann's constant $(k = 1.3806 \times 10^{-23} \text{ J K}^{-1})$. At $T = 0$ the function defined by (18.2) has the property that $p(E) = 1$ when $E < E_F$ and $p(E) = 0$ when $E > E_F$ so that all the energy states below the Fermi energy are filled, and all those above it are empty, as required. The density of states is of the order of 10^{23} cm^{-3} so that, at normal temperatures, the energy difference between adjacent states is small compared with thermal energies. As the temperature rises, some of the electrons have sufficient energy to move from
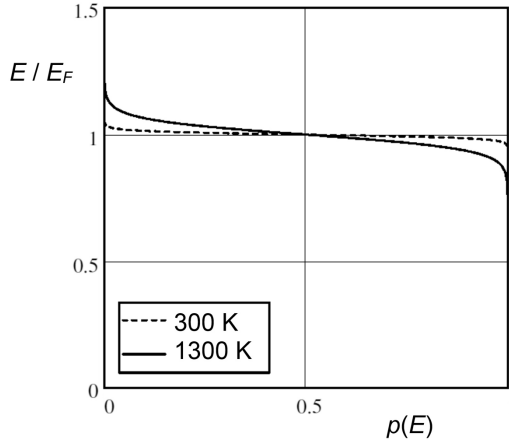
Figure 18.2: Probability of occupation of states given by the Fermi-Dirac distribution function at 300 °K and 1300 °K when $E_F = 3.8\,\mathrm{eV}$.

states just below the Fermi energy to states just above it. Figure 18.2 shows the probability of occupation of states at 300 °K and 1300 °K for barium $(E_F = 3.8\,\mathrm{eV})$.

If electrons are to escape from a metal, then the difference between their energies and the Fermi energy must exceed the work function. The energy required to enable electrons to escape may be supplied thermally, by photons, or by electron or ion bombardment. The height of the potential barrier can be reduced by a strong external electric field. These processes are discussed in the sections which follow.

18.2.1   Thermionic Emission

We have seen that work functions are typically of the order of a few electron volts. It is therefore possible to approximate (18.2) for energies of the order of $E_F + \phi$ by

$$p(E) = \exp\left(-\frac{E - E_F}{kT}\right), \tag{18.3}$$

because $E - E_F \gg kT$. At 17 C for example $kT = 0.025\,\mathrm{eV}$. The probability that an electron has just enough energy to escape from the metal is $\exp(-\phi/kT)$. For example, the probability that an electron can escape from the surface of tungsten $(\phi = 4.54\,\mathrm{eV})$ is 2×10^{-76} at 300 °K but is 4×10^{-17} at 1300 °K. Since the density of conduction electrons is of the order of 10^{23} cm^{-3} it can be seen that no emission is possible at 300 °K but that appreciable emission may occur at the higher temperature.

The current density of thermionic emission is governed by the Richardson-Dushman equation [1]

$$J = A_0 T^2 \exp(-\phi/kT). \tag{18.4}$$

The theoretical value of the constant A_0 is

$$A_0 = \frac{4\pi m_0 e k^2}{h^3} = 120 \text{ A cm}^{-2}\text{K}^{-2} \tag{18.5}$$

where h is Planck's constant $(h = 6.626 \times 10^{-34} \text{ J s})$. The probability that an electron will be reflected at the surface is negligible because of the smooth variation of potential outside the metal (see Figure 18.1). The work functions of metals depend upon temperature, and this can be modelled by rewriting (18.4) as

$$J = A_0 T^2 \exp-(\phi + \alpha T)/kT, \tag{18.6}$$

where α is the temperature coefficient of the work function. Taking the logarithm of (18.6) we obtain

$$\ln\left(\frac{J}{T^2}\right) = \ln A_0 - \frac{\alpha}{k} - \frac{\phi}{kT} = \ln A_R - \frac{\phi}{kT}, \tag{18.7}$$

so that a plot of $\ln(J/T^2)$ against $1/T$ is a straight line from which the constants A_R and ϕ can be determined. The current density measured is affected by a number of factors:

- The apparent work function increases if there is negative space-charge close to the surface. This is the case with space-charge limited emission when a virtual cathode is formed (see Section 5.3).
- The apparent work function decreases with increasing external field strength (see Section 18.2.2).
- The value of ϕ depends of the crystallographic plane of the emitting surface
- The work function is strongly affected by small quantities of adsorbed gases [4].
- The effective area of the emitting surface is not normally equal to the actual surface area.

Thus considerable care must be exercised in making experimental measurements of the properties of thermionic emitting surfaces. Nevertheless it is found that the properties of nearly all surfaces can be represented by (18.4) when A_0 is replaced by values of A_R determined by experiment [2, 5]. Practical thermionic cathodes are discussed in Section 18.5.

18.2.2   Field-Enhanced Emission (the Schottky Effect)

When the emitting surface forms the cathode of a plane-parallel diode then the potential outside the metal is given by

$$V = \frac{e}{16\pi\varepsilon_0 x} + \frac{x}{d}V_a, \tag{18.8}$$
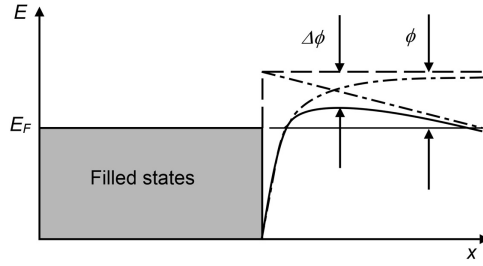
Figure 18.3: Reduction in the effective work function by an external electric field (the Schottky effect).

where d is the diode spacing and V_a is the anode voltage relative to the cathode, as shown by the solid line in Figure 18.3. The work function is then reduced by

$$\Delta\phi = \sqrt{\frac{eV_a}{4\pi\varepsilon_0 d}}. \tag{18.9}$$

It should be noted that the change in the work function has been exaggerated in Figure 18.3 for clarity. In practice the change is small. If the electric field is 1 MV m^{-1}, for example, then $\Delta\phi = 0.038$ V. For field-enhanced emission (18.7) becomes

$$\ln\left(\frac{J}{T^2}\right) = \ln A_R - \frac{\phi}{kT} + \frac{e}{kT}\sqrt{\frac{eV_a}{4\pi\varepsilon_0 d}}. \tag{18.10}$$

A plot of $\ln(J)$ against $\sqrt{V_a}$ at constant temperature gives a straight line for voltages greater than those for which the emission is space-charge limited. Extrapolation of the line back to $V_a = 0$ gives the zero field current (the *saturation current*) at that temperature. If the measurements are repeated at different temperatures then the plot of (18.7) gives the constants A_R and ϕ [6].

18.2.3   Field Emission

When the external electric field is greater than about 5 MV mm^{-1} it is possible for electrons to pass through the potential barrier by quantum mechanical tunnelling without the need for the surface to be heated (see Figure 18.4) [7, 8]. The process of tunnelling is analogous to the attenuation of waves as they pass through a short section of cut-off waveguide. Field emission, typically, takes place at sharp points, edges, and surface irregularities, where the electric field is enhanced by the geometry. A special case which is important in vacuum tubes is *triple point emission*. This occurs at the intersection between a metallic surface, a dielectric surface and vacuum where strong enhancement of the electric field is produced by the geometry [9].
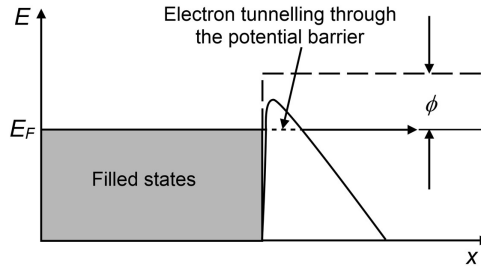
Figure 18.4: Cold field emission of electrons by tunnelling through the energy barrier.

The emitted current density is given by the Fowler-Nordheim equation in the form [8]

$$J = aV^2 \exp(-b/V),$$ (18.11)

where J is the current density in A cm^{-2}, V is the potential difference between the emitting surface and a distant anode, and a is defined by

$$a \cong 1.5 \times 10^{-6} \frac{\beta^2}{\phi} \exp\left(\frac{10.4}{\sqrt{\phi}}\right),$$ (18.12)

where ϕ is the work function in eV, b is defined by

$$b \cong 6.44 \times 10^7 \frac{\phi^{\frac{3}{2}}}{\beta}.$$ (18.13)

β is a geometric factor which depends on the electrode configuration, such that the electric field over the emitting area is $E = \beta V$. For a given material and geometry the parameters a and b are constants and a plot of $\ln(J/V^2)$ against $1/V$ (a Fowler-Nordheim plot) should be a straight line. It is found that this is generally the case over many orders of magnitude of current. Changes in the operating conditions (temperature, residual gas composition and pressure, and the current being drawn) can produce large changes in the plot. These changes are usually due to changes in the surface conditions and in the shape of the emitter. They may, or may not, be reversible.

18.2.4   Photo-Electric Emission

The energy of a photon is hc/λ, where λ is its wavelength and h is Planck's constant. If $hc/\lambda > e\phi$ then an electron may be ejected from the surface with kinetic energy given by

$$\frac{1}{2} m_0 u^2 = \frac{hc}{\lambda} - e\phi.$$ (18.14)

Thus, for a given material, the maximum wavelength of light for which photo-electric emission is possible is determined by the work function [2, 10]. The energies of photons in the visible region of the spectrum are in the range 1.6 to 3.3 eV. The work functions of metals are mostly greater than 4 eV so it is necessary to irradiate them with ultra-violet light to obtain photo-electric emission. The optical power incident on the surface is

$$P_{opt} = \frac{n_p hc}{\lambda},$$
(18.15)

where n_p is the number of photons per second. The ratio of the number of electrons emitted to the number of incident photons is known as the *quantum efficiency* $\left(\eta_q\right)$ of the surface. The electron current is then

$$I = n_p \, \eta_q \, e = P_{opt} \, \eta_q \left(\frac{\lambda e}{hc}\right).$$
(18.16)

The quantum efficiencies of clean metal surfaces are of the order of 10^{-4} to 10^{-5} but values as high as 0.3 have been obtained from compound surfaces.

In order to get an electron current high enough to be useful in an electron tube it is necessary to illuminate the surface with a high power laser. Since the light from the laser can be in very short pulses it is possible to produce a bunched stream of electrons. Photo-cathodes are used as electron sources in particle accelerators but are not yet suitable for use in microwave tubes. Unwanted photo-electric emission can occur in a microwave tube as a result of the bombardment of surfaces by X-rays produced within the tube (see Section 18.4). Further information on photo-emission can be found in [10–13].

18.3   Secondary Electron Emission

When electrons strike a solid surface in a vacuum they may be reflected elastically, or penetrate the surface and exchange energy with the electrons in the material [1]. An incident electron entering a solid material makes a series of collisions losing energy of about 30 eV each time. The electrons which gain energy through collisions may make further collisions so that a cascade of secondary electrons moves through the material as a result of the impact of a primary electron. One or more secondary electrons may reach the surface and escape. In addition the primary electron may re-emerge with reduced energy. Secondary electrons can also be generated by ion bombardment [14]. Secondary electron emission is possible from metals, and from insulating and semi-conducting materials. The ways in which the phenomenon is important in vacuum tubes include:

• the influence on the performance of a magnetron of the relative proportions of primary and secondary emission from the cathode (see Chapter 15);

- the use of secondary electron emission from cold cathodes in some crossed-field amplifiers (CFAs) (see Section 16.2.2);
- the effect of secondary electron emission on the performance of multi-element depressed collectors (see Section 10.4);
- voltage breakdown initiated by secondary emission from the surface of insulators (see Section 18.7.3);
- resonant RF discharges (multipactor) are sustained by the emission of secondary electrons (see Section 18.8).

Figure 18.5 shows the energy distribution of electrons emitted from a gold surface for three different primary energies [15]. This behaviour is typical of that found with other surfaces [16–18]. For incident energies greater than 100 eV it is convenient to divide the electrons emitted into three categories:

- elastically reflected primaries;
- inelastically reflected (re-diffused) primaries (energies > 50 eV);
- true secondary electrons (energies < 50 eV).

The energy of the boundary between the second and third groups is arbitrary, but 50 eV is commonly used in the literature [18–20]. The total number of true secondary electrons emitted may exceed the number of incident primary electrons. For primary energies greater than 100 eV the energy distribution of the true secondary electrons is essentially independent of the primary energy [20]. It is characterised by the most probable energy (E_{sm}), and full width at half maximum (HW), of the distribution, as shown in Figure 18.5. Typically, for metals, $1\,\text{eV} < E_{sm} < 5\,\text{eV}$ and $3\,\text{eV} < HW < 15\,\text{eV}$. The full width at half maximum is much smaller for insulators than for metals, and it depends strongly on very thin surface layers. It is found that the angular distribution of true secondary electrons from polycrystalline metals is independent of the angle of incidence of the primary electron, to a good approximation, and that it is proportional to the cosine of the angle between the direction of emission and the normal.

The fraction of elastically reflected electrons falls rapidly with increasing primary energy, as shown, for oxidised beryllium, in Figure 18.6. However, at primary energies greater than 50 keV the fraction of all (elastic and inelastic) reflected primaries starts to increase again [1]. Note that the absence of re-diffused electrons below a primary energy of 50 eV in Figure 18.6 is a consequence of the definition of secondary electrons. For primary energies greater than 100 eV the true secondaries are the dominant group [15, 17, 21, 22]. The secondary electrons that are able to escape from the surface are generated in a very thin surface layer which is of the order of 5 nm in metals, and 75 nm in insulators. Beyond that depth the secondary emission no longer depends upon the properties of the bulk material [20].

The *secondary electron emission coefficient* (*secondary electron yield*) of a material (δ) is the ratio of the number of secondary electrons emitted to the number of incident primary electrons. For all materials it is found that the dependence of δ on the energy of normally incident primary electrons can be represented approximately
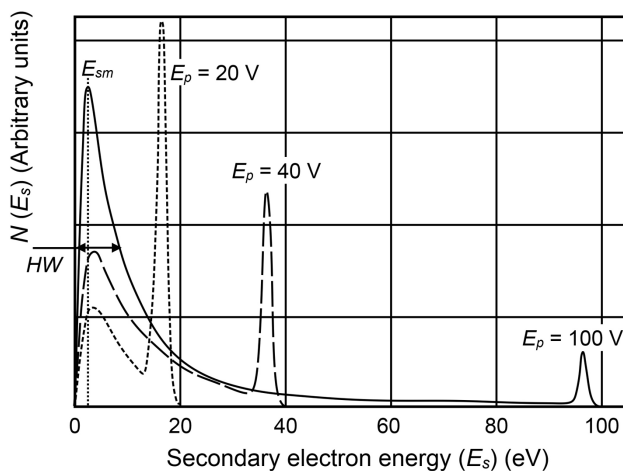
Figure 18.5: Typical distributions of the energies of secondary and reflected primary electrons from a gold surface for various primary electron energies (copyright1996, Elsevier, reproduced, with permission, from [15]).
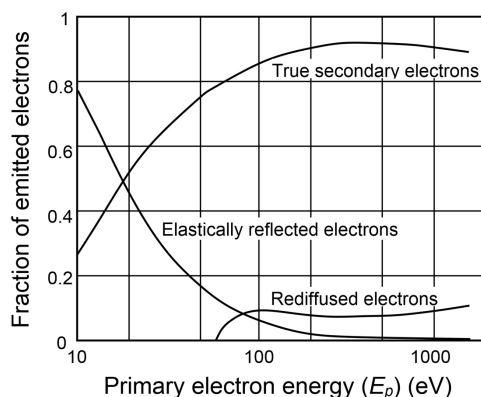
Figure 18.6: Fraction of elastically-reflected, rediffused and secondary electrons as a function of primary electron energy for oxidised beryllium (copyright 1994, IEEE, reproduced, with permission, from [17]).

by a universal curve having the shape shown in Figure 18.7 for $\theta_i = 0$. This curve is defined by the maximum value of the secondary emission coefficient (δ_m) and the primary electron energy at which it occurs (E_{pm}). The value of E_{pm} usually lies between 200 and 1000 V [23, 24]. When the angle of incidence of the primary electrons is not zero it is found that the coefficient of secondary electron emission increases, and the peak of the curve moves to higher primary energies, as shown in Figure 18.7. The effect of increasing the angle of incidence is that the primary electrons stay closer to the surface for longer so that the secondary electrons generated can escape more readily.
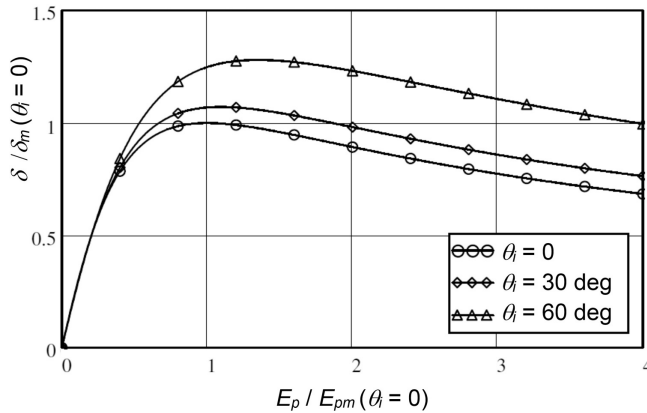
Figure 18.7: Variation of the normalised secondary electron emission coefficient with normalised primary electron energy for three different angles of incidence.

The secondary electron emission properties of materials depend upon the preparation and conditioning of the surfaces, and upon any surface layers present [25, 26]. The preparation may include cleaning, coating, and texturing of the surface before assembly. Conditioning may include heating, and exposure to ion, or electron, bombardment after evacuation of the tube. In addition to surface layers added deliberately, the surface may be contaminated or oxidised. The contamination of the surface may occur from exposure to air before the tube is evacuated. After evacuation there may be contamination from residual gases or evaporation of materials such as barium (from the cathode). For this reason it is necessary to treat the extensive data in the literature with some caution. For example, a comparison of measured data for aluminium from 13 different sources showed consistent values of the E_{pm} (~ 400 eV), but values of δ_m ranging from 0.5 to 3.3 [23]. The first figure may be the coefficient for a clean surface, while higher figures probably represent the effects of various thicknesses of oxide layer. A study of the properties of technical copper found $\delta_m \sim 2.3$ for material as received from the supplier. This figure was reduced to about 1.8 by baking the sample to 300 °C, and reduced to about 1.4 after exposure to an argon glow discharge [25]. The secondary electron emission coefficients of insulators are usually greater than those of metals. However, the measurement of secondary emission from insulators is made more complicated by the build-up of surface charge. Tables of the secondary electron emission properties of pure materials are given in [20, 23, 24, 27], but it should be noted that the figures given show considerable differences. The properties of some compounds and insulators are given in [20, 28]. The conditions under which the measurements of secondary electron yield are made usually differ from those existing in a vacuum tube. It is therefore difficult to have reliable knowledge of the properties of a surface within a tube.

The secondary electron emission coefficient of a surface can be adjusted by the choice of material, by the application of surface coatings, and by changing the

texture of the surface. Where a high secondary electron yield is required, as for the cathodes of some crossed-field amplifiers, the surface of a metal may be deliberately oxidised [17, 29]. In most other cases the aim is to reduce secondary emission as far as possible. In particular, it is desirable to reduce the possibility of avalanche multiplication of the number of electrons by ensuring that the range of primary energies for which $\delta > 1$ is as small as possible. This is often achieved by the addition of surface coatings such as titanium nitride and diamond-like carbon [30–32]. An alternative approach modifies the surface of the material so that the ability of secondary electrons to escape from it is reduced. The modification can take the form of grooves in the surface which tend to trap the electrons emitted [33]. The emission of secondary electrons can also be reduced by adding microscopic textures to surfaces [34–36]. Further details of these techniques, and their applications, can be found by searching for papers citing those referenced here.

18.3.1   Modelling Secondary Electron Emission

Models of secondary electron emission are needed to represent its effects in magnetrons, in multi-element depressed collectors, and in multipactor discharges (see Section 18.8). Theoretical studies of secondary electron emission give values of δ_m which bear some resemblance to measured values [37]. But the large scatter in the experimental data available makes the validation of the theory uncertain. The strong dependence of δ_m on the nature of the emitting surface means that measured values should be used in modelling, if possible. We have noted that E_{pm} is subject to smaller variations so that typical values can be obtained from the literature.

The shape of the curve for normal incidence in Figure 18.7 is much the same for all materials and empirical functions have been suggested by many authors [1, 15, 18, 20, 38, 39]. One approach is based on simple models of the penetration of the primary electrons into the material, and the subsequent escape of the secondary electrons generated. The results, for normal incidence, can be expressed in the form [24]

$$\frac{\delta}{\delta_m} = \frac{1}{1 - e^{-x_m}} \left(\frac{E_p}{E_{pm}} \right)^{1-k} \left\{ 1 - \exp\left[-x_m \left(\frac{E_p}{E_{pm}} \right)^k \right] \right\}, \tag{18.17}$$

where E_p is the energy of the primary electrons, k is adjusted to fit the curve to experimental data and x_m is the solution of

$$x_m = \left(1 - \frac{1}{k} \right)\left(e^{x_m} - 1 \right). \tag{18.18}$$

Equation (18.17) is identical to those given by other authors who assumed fixed values for k in the range 1.35 to 1.67 [23, 40–43]. However, the fit to experimental data can be considerably improved by using k as an adjustable parameter. The curves generated by (18.17) in this way give an excellent fit to the best available
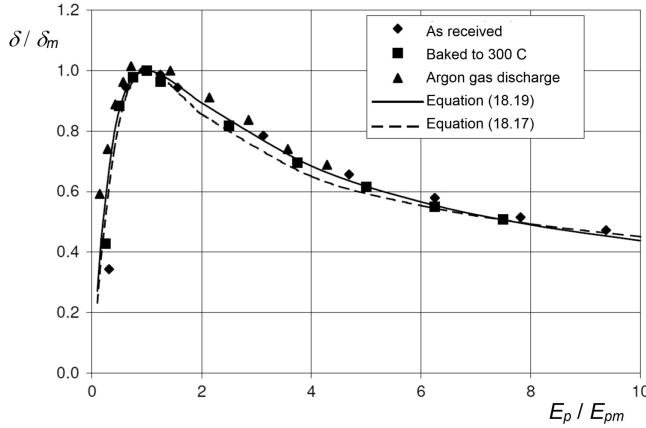
Figure 18.8: Normalised secondary electron emission curves for technical copper compared with the curves calculated using equations (18.17) and (18.19).

experimental data for 18 different elements from beryllium to lead [24]. The values of k ranged from 1.25 to 1.69. The mean absolute relative error of the fit to experimental data was better than 5% in nearly every case. It seems, therefore, that the shape of the curve is not truly universal, as was assumed in earlier work.

An alternative, simpler, equation adopted by Furman and Pivi that appears to give a better fit to data for technical materials is [18]

$$\frac{\delta}{\delta_m} = \frac{s\left(E_p/E_{pm}\right)}{s-1+\left(E_p/E_{pm}\right)^s},$$
(18.19)

where s is a parameter which is chosen to give the best fit to experimental data. The value of this parameter was 1.54 for copper, and 1.81 for stainless steel. Figure 18.8 shows a comparison between experimental results for technical copper taken from [25] and curves calculated from (18.17) with $k = 1.4$, and (18.19) with $s = 1.54$.

When the angle of incidence of the primary electrons is not zero it is found that the secondary electron emission curve can be derived from the curve for normal incidence. The maximum value of δ, and the primary energy at which it occurs, can be expressed in terms of empirical functions of the angle of incidence. Furman and Pivi [18] proposed the relations

$$\delta_m\left(\theta_i\right) = \delta_m\left(0\right)\left[1+0.66\left(1-\cos^{0.8}\theta_i\right)\right]$$
(18.20)

and

$$E_{pm}\left(\theta_i\right) = E_{pm}\left(0\right)\left[1+0.7\left(1-\cos\theta_i\right)\right].$$
(18.21)

These functions, which apply to the unpolished, polycrystalline surfaces commonly found in vacuum tubes, are substituted into (18.19) to give the secondary electron

emission coefficient as a function of the energy of the primary electrons and the angle of incidence

$$\frac{\delta}{\delta_m(\theta_i)} = \frac{s\left(E_p/E_{pm}(\theta_i)\right)}{s-1+\left(E_p/E_{pm}(\theta_i)\right)^s}.$$ (18.22)

Alternative expressions, given by Vaughan [44], are

$$\delta_m(\theta_i) = \delta_m(0)\left[1+k_{s\delta}\frac{\theta_i^2}{2\pi}\right]$$ (18.23)

and

$$E_{pm}(\theta_i) = E_{pm}(0)\left[1+k_{sV}\frac{\theta_i^2}{2\pi}\right],$$ (18.24)

where $k_{s\delta}$ and k_{sV} are constants which depend upon the surface roughness. The default value of both constants is unity for normal machined surfaces. The range is from zero, for deliberately roughened surfaces, such as textured carbon, up to 2 for exceptionally smooth, clean, surfaces [44, 45]. The curves given by these two equations are close to those from (18.20) and (18.21) if the constants are set to 1.6 and 1.9, respectively. The curves in Figure 18.7 were computed using (18.19) and (18.22). The effect of surface layers was studied by Yu et al. who were able to show agreement with experimental results for layers of silicon dioxide of different thicknesses on silicon substrates [46]. In any particular case the model used should, if possible, be validated by reference to experimental measurements of the properties of the surface. It is also desirable that it should correctly model all three groups of electrons leaving the surface in both energy and angular distribution [18, 19].

18.4   X-ray Emission

X-rays are generated whenever electrons with energies greater than about 5 keV strike solid surfaces. Tubes which are designed to generate X-rays are outside the scope of this book, but X-ray generation is of more general importance because of its implications for safety [47]. The rapid deceleration of the electrons as they enter the material generates a broad spectrum of *brehmsstrahlung* (brake radiation) as shown in Figure 18.9. The spectrum has a sharp cut-off at the photon energy (eV_p) equal to that of the incident electrons. In addition there are spectral lines, arising from transitions between the electron energy levels of atoms within the material. The wavelength of X-rays is related to the photon energy (V_p) by

$$\lambda = \frac{hc}{eV_p}.$$ (18.25)
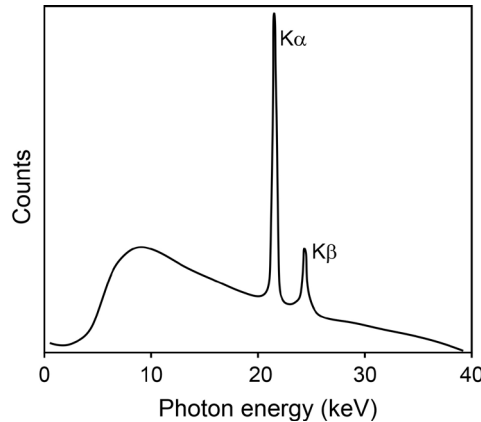
Figure 18.9: X-ray emission spectrum for a silver target with 40 keV electrons (copyright 2017, AMPTEK Inc., reproduced with permission, from [48]).

where h is Planck's constant. The fraction of the power carried by the electrons that is converted into X-rays is given approximately by

$$F = 7 \times 10^{-4} Z V_a, \tag{18.26}$$

where Z is the atomic number of the target and V_a is the accelerating voltage. The remainder of the power is dissipated as heat. Thus the X-ray intensity increases with increasing electron energy, and with increasing current striking the surface.

When X-rays with energies less than 1 MeV pass through materials they experience coherent, and incoherent, scattering and photo-electric absorption. Additional loss mechanisms apply at higher energies. The intensity of the radiation (I) as a function of depth into the material is given by

$$I = I_0 \exp(-\mu x), \tag{18.27}$$

where I_0 is the initial intensity, x is the thickness of the material, and μ is the *attenuation coefficient*. Equation (18.27) applies to the attenuation of monochromatic X-rays incident normally on a plane sheet of material. The attenuation coefficient increases rapidly with atomic number. It decreases with increasing photon energy, as shown in Figure 18.10, which is based on data from [49]. The absorption edges, at which the value of μ is discontinuous, correspond to the lines in the spectra of the elements. It can be seen that, at a photon energy of 50 keV, the thickness of copper required to produce a given attenuation is about four times that of lead. Similarly the thickness of aluminium must be 24 times that of copper. Equation (18.27) can also be written

$$I = I_0 \exp\left(-\left(\frac{\mu}{\rho}\right)(\rho x)\right), \tag{18.28}$$
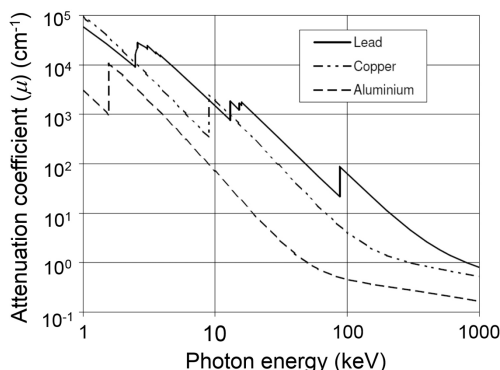
Figure 18.10: Comparison between the X-ray attenuation coefficients of lead, copper, and aluminium.

where ρ is the density of the material. The parameter (μ/ρ) is the *mass attenuation coefficient* and (ρx) is the mass per unit area of the sheet of material. This form is more convenient for calculations of X-ray shielding. Tables of the mass attenuation coefficient as a function of photon energy for many materials are given in [49]. It should be noted that in practical cases, it is necessary to consider the spectrum of the X-rays, and the geometry of the material, when computing the attenuation [50]. The implications of X-ray generation for the safe operation of tubes are discussed in Section 20.8.

18.5   Thermionic Cathodes

All the tubes discussed in this book require sources of electrons for their operation. In the great majority of cases this source is a thermionic cathode. The cathode is a critical component because failure of emission is one of the chief causes of tube failure. According to (18.6) the current density depends upon the temperature of the cathode and the work function of the surface. The active material of the surface of the cathode is gradually evaporated at a rate which increases with temperature. Thus there is a trade-off between the current density and the life of the cathode. The normal end of life of a tube is determined by the failure of the cathode emission, which is commonly taken to be a reduction of 10% in the current density [51]. The temperature of a cathode is normally measured by optical pyrometry in which the brightness of the surface is compared with that of the filament of a standardised lamp. The temperature measured is the *brightness temperature.* This is the temperature at which the radiation from a black body is as bright as the cathode. The brightness temperature is always lower than the true temperature which must be obtained by applying a conversion formula [6].

The current which can be obtained from a cathode surface is reduced by:

- Bombardment by positive ions generated by ionisation of the residual gas in the tube. Material is removed from the surface of the cathode by sputtering leading to an increase in the work function. This problem arises if the residual gas pressure in a tube is too high. It is more likely to occur in continuously operated tubes, where the ions accumulate in the potential well caused by the space-charge in the beam. In pulsed tubes there is less time for the ions to accumulate, and they are dispersed by their space-charge in the intervals between the pulses.
- Absorption of oxidising gases such as oxygen, carbon dioxide, and water vapour that *poison* the cathode and increase its work function.

The emitted current density from a damaged cathode can sometimes be restored if the source of damage is removed, but the final current density does not always return to the previous level [4, 52, 53].

The life expectancy of the cathode is a major factor in tube design, and manufacturers are understandably wary of the premature introduction of new technologies until they have been subjected to intensive life testing. An introduction to the main types of thermionic cathode is given below. These have been established for many years but they are still the subject of research and of an extensive literature. Further information about thermionic cathodes can be found in [2, 6, 13, 51, 54–56].

18.5.1   Metal Emitters

Pure metals have relatively high work functions and they must be operated at high temperatures to give useful emission. The metal, in the form of a wire, is heated by the passage of an electric current through it. The cathode is, typically, constructed in the form of a cage, or a helix. The choice of metals is limited to those which have high melting points of which tungsten $\left(\phi = 4.54 \text{ eV, MP} = 3640 \text{ K} \right)$ is the one most commonly used. The high operating temperature means that surface contaminations are rapidly evaporated and a tungsten emitter can operate in a poor vacuum, such as that in an unbaked demountable system. The current density is typically 0.4 A cm^{-2} at a temperature around 2,500 °K. Over time the wire becomes thinner because of evaporation of the tungsten. The life of a tungsten emitter is assumed to end when 10% of the material has evaporated because hot spots then tend to form [57].

The addition of a monolayer of thorium to a tungsten wire reduces the work function to 2.7 eV. Thus, a directly heated thoriated tungsten wire can deliver higher current densities, at a lower temperature, than pure tungsten. At a typical operating temperature of 2000 °K such an emitter can deliver up to 4 A cm^{-2} with a life in excess of 10,000 hours. The thorium layer is readily evaporated if the wire temperature is too high, leading to premature failure of emission. It has been found that this problem can be reduced, to some extent, by carburising the top layer of

the tungsten to form tungsten carbide [2, 57]. The tungsten is doped with small grains of thorium oxide, which react with the tungsten carbide to replenish the free thorium on the surface of the filament. This type of emitter is used in triodes and tetrodes and in magnetrons for industrial heating and microwave ovens [58].

18.5.2   Oxide Cathodes

Cathodes having lower work functions than metallic emitters and, therefore, lower operating temperatures can be made using oxides of certain elements. Oxide cathodes are made by painting, or spraying, a coating onto a metal base. The coating is usually a mixture of the carbonates of barium, strontium, and calcium with a molar ratio (50–58%):(38–45%):(4–8%) [2, 59]. The surface is activated during the evacuation of the tube by heating it to convert the carbonates to oxides, giving a surface whose work function is between 1.2 and 1.8 eV. The DC current density of a conventional oxide cathode, with a life in excess of 10,000 h, is up to 1 A cm^{-2} but higher currents can be obtained for short-pulse, low duty cycle operation [13]. The DC current density is limited by the high resistivity of the oxide layer. Oxide cathodes are easily poisoned and must be kept in high vacuum conditions once they have been activated. Problems may also be caused by poor adhesion of the oxide layer to the metal base. Because oxide cathodes are simple and inexpensive to manufacture they are a continued subject of research [2, 13, 59].

18.5.3   Dispenser Cathodes

The cathodes generally used in microwave tubes are dispenser cathodes in which the active materials are held within a porous metal matrix [2, 6]. In these cathodes the active material stored in the matrix can diffuse to the surface and replace that lost by evaporation. Dispenser cathodes have work functions of 1.8–2.1 eV but do not suffer from the problem of high resistivity which limits the usefulness of oxide cathodes. Many different types of dispenser cathode have been invented and details can be found in the literature. The three main arrangements are shown in Figure 18.11. The essential component in all three is a porous metal matrix which is commonly made by sintering tungsten powder. The simplest arrangement is the impregnated cathode in Figure 18.11(a) in which the matrix is impregnated with a $BaO:CaO:Al_2O_3$ mixture in proportions such as 5:3:2 (the Phillips B cathode), or 4:1:1 (the S cathode). The life of the cathode can be greatly increased by adding a reservoir of BaO behind the tungsten plug to form the reservoir cathode shown in Figure 18.11(b). This arrangement is especially useful for space tubes where very long life is essential. In the third arrangement a surface layer having a thickness of about 0.5 μm is added to the tungsten plug as shown in Figure 18.12(c). This layer, which is commonly an osmium–ruthenium alloy, reduces the work function of the surface allowing operation at a lower temperature and thus extending the cathode life (the M cathode).
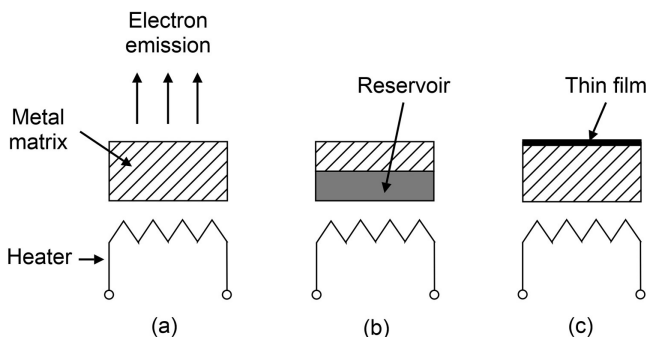
Figure 18.11: Arrangements of dispenser cathodes: (a) impregnated cathode, (b) reservoir cathode, and (c) impregnated cathode with a thin surface film.

The way in which the saturated current available from a dispenser cathode changes with life can be modelled by supposing that it depends upon the relative proportions of the surface in which the emission is limited by space-charge, and by temperature (see (5.9)). Where there is space-charge limited emission, a low work function is maintained by a thin layer of the active material on the surface of the matrix. In an impregnated cathode this layer can be thought of as a balance between the rate of evolution of active material from the matrix and the rate of evaporation. It has been shown that this model gives a good fit to experimental data [60]. The balance is strongly affected by the geometry of the diode in which the measurements are made. Lifetime measurements made using a closely spaced parallel diode are not representative of the lifetime achieved when the same cathode is incorporated in an electron gun. This is because evaporated barium is refluxed to the cathode in a closely spaced parallel diode so that the active layer is maintained [61]. In an M-type cathode the diffusion of tungsten, from the matrix, into the surface layer changes the properties of the layer over time. With this addition, the model is a good fit to experimental data. An interesting feature of the model is that it predicts that at temperatures exceeding 1400 °K the saturated current falls with increasing temperature. This is because the rate of evaporation exceeds that at which the surface layer can be replenished from the matrix. The shape of the curve showing the transition between space-charge limited and temperature limited emission depends upon the parameter α in (5.9). This parameter, which is related to the uniformity of the emitting surface, decreases with time so that knee of the curve in Figure 5.2 becomes more rounded. The rate at which α decreases is increased as the temperature of the cathode is raised [61].

Dispenser cathodes can deliver DC current densities up to 4 A cm^{-2} at around 1300 °K with lifetimes of over 100,000 hours. The addition of Sc_2O_3 to the barium mixture in a dispenser cathode has been found to decrease the work function and increase the current density (the scandate cathode). Although these cathodes have delivered current densities of 60 A cm^{-2} at around 1250 °K it has been found that they are sensitive to ion bombardment, from which they do not readily recover [52].

Because the thin surface layer in Figure 18.12(c) is also readily eroded by ion bombardment, cathodes have been developed in which the coating material is added to the tungsten matrix. The result is a mixed-matrix (MM) cathode [62]. Fuller information about all these types of cathode, including details of their manufacture, can be found in the references cited.

18.6   Field Emission Cathodes

A thermionic cathode requires a heater which has a finite warm-up time and adds to the power consumption of a tube. Hence there is continuing interest in the possibility of using cold cathodes employing field emission. The most successful embodiment of this concept is the Spindt cathode shown in Figure 18.12 [13, 63, 64]. This cathode comprises an array of field emitter tips, made using micro-electronic technology, arranged in holes in a gate electrode. The emitting tips are made from metals with high melting points, such as tungsten and molybdenum, from semiconductors, or from carbon nanotubes [13]. The presence of a gate electrode allows the electron beam to be pulsed, and gives the potential for modulation of the beam at microwave frequencies. This has been demonstrated in experimental tubes, but the cathode lifetime is short [65–67]. In one example a cathode with 50,000 tips delivered a current of 120 mA for several days at a current density of 13 A cm^{-2}. The reliability is limited by local overheating and voltage breakdown leading to short-circuiting of the structure. Research into improved reliability is continuing [68, 69]. This type
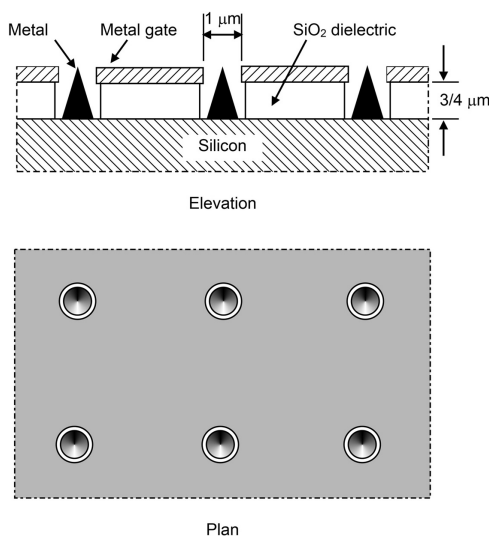
Figure 18.12: The Spindt cathode
(copyright 1991, IEEE, reproduced, with permission, from [63]).

of cathode also lends itself to incorporation in vacuum micro-electronic devices [8, 64]. For further information on this and other types of cold cathode see [12, 13].

18.7   Voltage Breakdown

Voltage breakdown sets limits to the design of tubes. It may take place either inside, or outside, the vacuum envelope, or through insulators, with both DC and RF voltages. The physical processes involved are discussed below. A voltage breakdown dissipates stored energy which may cause damage to the tube, and alter its performance. In extreme cases the damage may be so severe that the tube can no longer operate. For this reason high power tubes are provided with protection circuits which are designed to turn the tube off when a breakdown takes place (see Chapter 20). Whenever a tube 'trips out' it ceases operation for the time taken to restore normal operating conditions. Thus the objective of the designer is to eliminate voltage breakdown as far as possible to increase the reliability of a tube.

18.7.1   Voltage Breakdown in Vacuum

The primary cause of voltage breakdown between metal electrodes in vacuum is field emission from whiskers and from microscopic dielectric inclusions in the surfaces. The latter are sites for triple point emission as described in Section 18.2.3. However, it is found that breakdown occurs at lower voltages than can be explained by field emission alone. The additional emission is believed to be in the form of secondary electrons produced by ion bombardment. If it is assumed that the secondary electron emission is linearly proportional to the ion energy then the threshold for breakdown is given by

$$WE^2 \exp\left(-\frac{K_1}{E}\right) \leq K_2, \tag{18.29}$$

where W is the maximum possible ion energy, E is the electric field on the surface of the negative electrode and $K_1 = 17 \, \mathrm{MV \, m^{-1}}$ and $K_2 = 1.8 \, \mathrm{MV^3 \, m^{-2}}$ are empirical constants. This is Kilpatrick's criterion, which is commonly used as a guide to DC, pulsed DC, and RF vacuum breakdown [70, 71].

For DC voltages W is equal to the potential difference between the electrodes. Thus, for example, the breakdown voltage between plane electrodes with a spacing of 10 mm given by (18.29) is 100 kV. When the electrodes are not flat the electric field on the surface of the negative electrode may be increased and the breakdown voltage reduced. The effect of departures of the field from uniformity can be represented by field enhancement factors defined as the ratio of the maximum electric field to that between plane electrodes with the same spacing and potential difference [72]. The field enhancement factor may be as high as 10 for typical geometries, but figures up to 100 occur for tiny protuberances on the surfaces. The breakdown

voltage also depends upon surface finish, the metal from which the electrodes are made, and on any contamination present. It is usual for a high-voltage system to be conditioned by gradually raising the voltage and allowing breakdown events to remove the sources of emission. The criterion in (18.29) applies when the voltage is pulsed with a pulse length is 1 ms or more. For shorter pulses the breakdown voltage increases by a factor of 3 when the pulse length is reduced to 1 μs [73]. When the ion energy is greater than 100 keV there is a marked decline in secondary electron yield and the breakdown voltage is greater than that predicted by (18.29) [71].

For an RF voltage (V_1) the maximum ion energy (W) in (18.29) is reduced by transit time effects. A graph for the ion transit time correction for higher voltages is given in [70] as a plot of W/V_1 against V_1/V^* where the normalising voltage is defined by

$$V^* = \left(\frac{2\pi g}{\lambda}\right)^2 \left(\frac{M_0 c^2}{\pi e}\right),$$ (18.30)

where g is the gap between plane electrodes, λ is the free space wavelength and M_0 is the rest mass of a hydrogen atom. If $V_1/V^* \leq 1$ the correction can be expressed as

$$\frac{W}{V_1} \simeq 0.63 \frac{V_1}{V^*}.$$ (18.31)

Substituting for W in (18.29) and rearranging gives

$$f(MHz) = 1.63 E^2 \exp(-8.5/E),$$ (18.32)

where the electric field E is in MV m^{-1} [71]. The maximum electric field is found to be proportional to the square root of the frequency up to about 1 GHz [74]. At higher frequencies the increase is more gradual. In the range 20 GHz to 40 GHz the maximum attainable surface field has been found to be independent of frequency [75].

The breakdown process described above may lead to the generation of a full vacuum arc [72, 76, 77]. An arc may be initiated in various ways including the explosion of a microscopic protrusion on the negative electrode, and the presence of loose particles within the gap. The material of the negative electrode is vaporised and ionised so that a plasma cloud is formed, which provides a conducting path between the electrodes. On the negative electrode the arc terminates at one, or more, small spots a few micrometres in diameter. The intense local heating causes thermionic emission of electrons, liberates adsorbed gases, and releases metal vapour, all of which sustain the plasma in the arc. At arc currents up to a few hundred amps the anode spot is the passive collector of the cathode plasma jet, and is therefore fixed.

The cathode spots, however, move rapidly over the surface producing a characteristic pattern of erosion. The arc voltage is usually 10–30 volts and it varies little with the current. The extent to which material is removed from the negative electrode by the arc depends upon the current, the duration of the arc, and the

material from which the electrode is made. The erosion is generally greater in materials having low boiling points and low thermal conductivity. Tube manufacturers sometimes specify the maximum permissible arc energies for their products beyond which irreversible damage may be caused.

A DC arc commonly has a lifetime of between 1 msec. and 1 sec. after which it extinguishes spontaneously. The lifetime depends upon the material of the cathode and increases rapidly with the current. For example, the duration of an arc on copper is about 1 ms at a current of 5 A but becomes less than 1 μs if the current is reduced below about 3 A [76]. If the current in the arc is reduced by the action of protection circuitry (see Chapter 20) then it will extinguish very fast. The high rate of change of current with time is sufficient to produce dangerously high voltages on the wires connected to the tube. It is therefore very important to ensure that their self-inductances are as low as possible. This applies especially to the connection between the tube body and ground.

18.7.2   Voltage Breakdown in Gases

DC voltage breakdown in gases depends upon the type of gas, and its pressure, and also upon the shapes and surface condition of the electrodes [11, 78–80]. The processes are complex and only a brief introduction is given here. Figure 18.13 shows a typical graph of current against the voltage applied between a pair of electrodes surrounded by a gas. This graph illustrates the properties of the Townsend discharge [81]. Some of the gas molecules are ionised by cosmic radiation and radioactivity in the ground. When a voltage is applied to the electrodes the free electrons and ions can drift towards them so that a small current flows. The energy gained by the electrons is not sufficient to cause further ionisation of the gas. The current saturates when the formation of new electron-ion pairs balances the rate at which they are removed by the applied field (region T_0 in Figure 18.13). When the voltage is increased, the electrons can gain enough energy from the electric field to produce some additional ionisation and the current starts to rise (region T_1). The positive ions drift towards the cathode. A further increase in the voltage gives the positive ions sufficient energy to liberate secondary electrons from the negative electrode (cathode). The electron current reaching the positive electrode (anode) then increases more rapidly (region T_2). The electric discharge through the gas in these three regions relies on an external source of electrons, and the flow of current ceases if that source is removed. As the current increases, a point is reached where the generation of secondary electrons, by ions striking the cathode, is sufficient to sustain the flow of current without any external source of electrons. At the point of transition from a non-self sustaining discharge to a self-sustaining discharge the voltage reaches a peak. This is the breakdown (or sparking) voltage. Once electrical breakdown has occurred the current is determined by the impedance of the external circuit.

It has been found that experimental values of the breakdown voltage, for a given gas and arrangement of electrodes, depend only upon the product of the
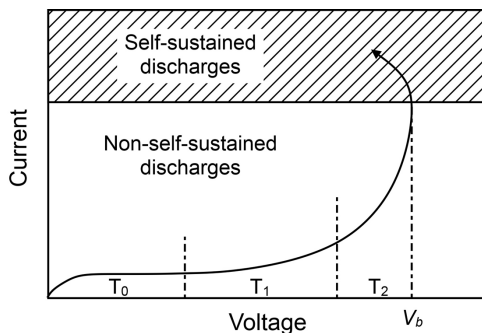
Figure 18.13: Current – voltage characteristics of the Townsend discharge (copyright 1941, Dover Publications, reproduced, with permission, from [78]).
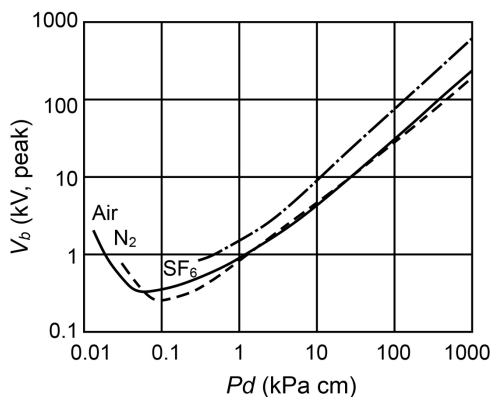
Figure 18.14: The Paschen curves for air, nitrogen (N_2) and sulphur hexafluoride (SF_6) (copyright 1982, IEEE, reproduced, with permission, from [82]).

gas pressure and the separation of the electrodes. This result is known as Paschen's Law. Figure 18.14 shows measured Paschen curves for air, nitrogen (N_2) and sulphur hexafluoride (SF_6) for plane electrodes [82]. The results for plane electrodes are a useful guide to other cases, if the field enhancement factor is not too great. For example, the curve for air in Figure 18.14 gives a breakdown voltage of 30 kV for a 1 cm gap at standard pressure (100 kPa). This is close to that commonly quoted for breakdown between spheres in dry air. The breakdown voltage of air is reduced by humidity, and by the presence of dust particles, and it is good practice to use a safety factor of three for design purposes. The DC breakdown voltage may be used as an estimate of the RF breakdown voltage to calculate the maximum power which can be passed down coaxial lines and waveguides (see Section 2.3).

For typical electrode spacings the minimum of the Paschen curve normally occurs well below atmospheric pressure. In the region of the minimum the breakdown voltage is particularly affected by the material of the electrodes. To the right of the minimum the breakdown voltage of a fixed gap increases with pressure. This

result is sometimes used in high voltage systems by pressurising them to reduce the risk of voltage breakdown. High power waveguide systems may also be pressurised for the same reason. The breakdown voltage can also be increased by using a gas such as SF_6 which readily attaches electrons to form negative ions, so reducing the build-up of the discharge. This gas is inert and non-toxic but it breaks down into components which are corrosive and highly toxic when subjected to sparking. It is also expensive and it has been found that results which are nearly as good can be obtained by adding a small fraction of SF_6 to N_2 [79]. Moreover, it is a greenhouse gas that is extremely long-lived in the atmosphere and its use is therefore subject to restrictions. It should be noted that the breakdown voltage decreases as the pressure is reduced, and this is important in the design of systems for use at high altitudes. For further information on voltage breakdown in gases see [78, 83, 84].

To develop an expression for the Paschen curve we consider a pair of plane parallel electrodes with separation d, and potential difference V between them, surrounded by a uniform gas at pressure P. Let us suppose that electrons are liberated from the surface of the negative electrode (cathode) at a steady rate n_0 per second. The actual cause of emission (themionic emission, photo-electric emission etc.) is not important to the argument. These electrons are accelerated towards the anode and collide with the gas molecules. Some of these collisions will result in positive ionisation of the molecules so that the number of electrons moving through the gas increases with distance towards the anode. If the rate of electron flow is n at a plane distant x from the cathode then the number of additional electrons produced in a distance dx is proportional to n so that

$$\frac{dn}{dx} = \alpha n, \tag{18.33}$$

where α is the fraction of electrons which make ionising collisions in unit distance. Integrating (18.33) with respect to x from the cathode to the anode gives the number of electrons per second reaching the anode as

$$n_a = n_0 \, e^{\alpha d}. \tag{18.34}$$

Thus the current reaching the anode is

$$I_a = I_0 \, e^{\alpha d}, \tag{18.35}$$

where I_0 is the initial electron current leaving the cathode. The positive ions formed by the collisions drift towards the cathode under the influence of the electric field. In the steady state the total current must be the same at the cathode and the anode. Thus the number of positive ions striking the cathode per second must be $n_0 \left(e^{\alpha d} - 1 \right)$. The impact of the ions on the cathode may liberate secondary electrons, which are added to the initial electron current. Thus, the total electron current leaving the cathode is

$$I_c = I_0 + \gamma I_c \left(e^{\alpha d} - 1 \right), \tag{18.36}$$

where γ is the secondary electron emission coefficient for impact by the positive ions. The dependence of γ on the energy of the incident ions is neglected. The electron current leaving the cathode is increased by ionisation so that the anode current becomes

$$I_a = I_c e^{\alpha d} = \frac{I_0 e^{\alpha d}}{1 - \gamma(e^{\alpha d} - 1)}. \tag{18.37}$$

The parameters α and γ are the first and third Townsend ionisation coefficients. The second coefficient (β), representing ionisation caused by the impact of ions on neutral molecules, is effectively zero for discharges of interest here.

Now, let us consider the effect of the collisions between electrons and gas molecules on the energies of the electrons. For the moment it will be assumed that that these collisions do not result in ionisation, and that an electron loses all its energy as a result of a collision. If the fraction of electrons making a collision in unit length is $1/\lambda$ then the number of electrons which have yet to make a collision decreases with position so that

$$\frac{dn}{dx} = -\frac{n}{\lambda} \tag{18.38}$$

and therefore

$$n = n_0 e^{-x/\lambda}. \tag{18.39}$$

The mean distance travelled by an electron before making a collision is

$$\bar{x} = \frac{1}{n_0} \int_0^\infty \frac{nx}{\lambda} \, dx. \tag{18.40}$$

Substituting for n from (18.39) and performing the integration we find that $\bar{x} = \lambda$. Thus λ is the *mean free path* of the electrons.

If an electron starts from rest, and moves in a uniform electric field E_x, then it has sufficient energy to ionise a molecule if

$$E_x x \geq V_i \text{ or } x \geq V_i / E_x, \tag{18.41}$$

where V_i is the ionisation potential of the molecule. Thus, electrons which have travelled less than this distance do not have sufficient energy to cause ionisation. From (18.39) the fraction of the electrons which are able to produce ionisation is

$$\frac{n}{n_0} = \exp\left(-\frac{V_i}{\lambda E_x}\right). \tag{18.42}$$

The fraction of electrons per unit length which cause ionisation (α) is then the product of the fraction which make collisions in unit length $(1/\lambda)$ and the fraction which have sufficient energy to cause ionisation from (18.42). Hence

$$\alpha = \frac{1}{\lambda}\exp\left(-\frac{V_i}{\lambda E_x}\right). \tag{18.43}$$

Now the mean free path is inversely proportional to the density of the gas and, therefore, inversely proportional to the pressure. Thus we can write (18.43) as

$$\alpha = AP\exp\left(-\frac{BP}{E_x}\right) \tag{18.44}$$

where P is the gas pressure and A and $B = AV_i$ are constants for the gas. This equation gives results which agree with measurements for values of E_x/P from 450 to 7500 V kPa^{-1} cm^{-1} [82].

   The condition for electric breakdown is found by setting the denominator in (18.37) to zero so that

$$\alpha d = \ln\left(1+\frac{1}{\gamma}\right). \tag{18.45}$$

We note that this condition does not depend upon the initial current I_0. Also, from (18.44),

$$E_x = \frac{BP}{\ln(AP/\alpha)}. \tag{18.46}$$

Setting $E_x = V_b/d$, where V_b is the breakdown voltage, and substituting for α from (18.45) we find that the breakdown voltage is given by

$$V_b = \frac{BPd}{\ln\left[APd/\ln(1+1/\gamma)\right]}. \tag{18.47}$$

In this equation the constants A and B depend only on the properties of the gas while γ depends also on the properties of the cathode surface. We note that, if all these are constant, then the breakdown voltage depends only on the product Pd, in conformance with Paschen's Law. The graph of V_b against Pd has a minimum when

$$Pd_{\min} = \frac{e}{A}\ln(1+1/\gamma) \tag{18.48}$$

and

$$V_{\min} = \frac{eB}{A}\ln(1+1/\gamma). \tag{18.49}$$

Then (18.47) can be expressed in normalised form as

$$\frac{V_b}{V_{\min}} = \frac{Pd/(Pd)_{\min}}{\ln\left[Pd/(Pd)_{\min}\right]+1}. \tag{18.50}$$
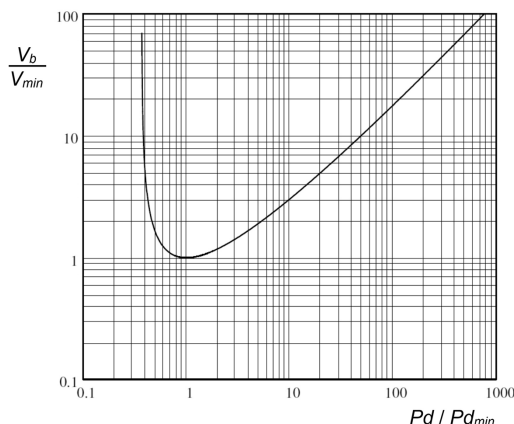
Figure 18.15: Theoretical Paschen curve.

Figure 18.15 shows the curve defined by (18.50). This curve has the same general form as the curves in Figure 18.14, but the prediction that all cases can be represented by the same normalised curve is not confirmed. This is a consequence of the simplifying assumptions which were made in the derivation of (18.50). However, the derivation does give insight into the physics of voltage breakdown. To the right of the minimum of the Paschen curve the breakdown voltage increases because the mean free path decreases, and fewer electrons gain sufficient energy to cause ionisation. To the left of the minimum the breakdown voltage increases because the number of molecules available for ionisation decreases. It should be noted that the equations derived above do not agree well with experimental results for breakdown in narrow gaps. This is because the electric field can be high enough to cause field emission, which has been ignored in the derivation.

18.7.3   Voltage Breakdown on Insulators

Any vacuum tube has one or more insulators which form part of the vacuum envelope. These insulators, which are usually ceramic, include those separating the cathode from the anode, and collector elements from one another and from the tube body. Voltage breakdown on the vacuum side of these insulators occurs at a lower voltage than for an equivalent vacuum gap [79]. This breakdown is initiated by the impact of electrons on the surface of the ceramic. These may come from any source but a particular problem is caused by triple junction emission. If the ceramic is attached to the metal of the tube body by brazing, then some of the brazing alloy usually spreads a little way onto the surface of the ceramic, providing a sharp metallic edge. The risk of triple junction emission at that edge can be reduced by shaping the electrodes at the ends of the insulator to shield the brazed joints from the electric field. The aim is to ensure that any breakdown

takes place directly between the electrodes, rather than along the surface of the insulator.

Ceramics have high secondary electron emission coefficients. Thus the impact of an electron on the surface of a ceramic may generate a number of secondary electrons and leave behind a positive surface charge. The secondary electrons are accelerated across the surface of the ceramic by the external field, and drawn back to it by the surface charge, so that they make further collisions and an electron avalanche develops. The growing discharge liberates adsorbed gas molecules from the surface of the ceramic. These may become ionised, increasing the risk of breakdown. A breakdown can damage the surface of the insulator so that the breakdown voltage is reduced. In extreme cases the insulator may be punctured or may fracture, leading to loss of vacuum. It is important that the surface of the insulator should be free from contamination, and that adsorbed gases are driven off, as far as possible, by baking when the tube is evacuated. The breakdown voltage is strongly dependent on the shape of the insulator. It can be increased by adding corrugations to increase the path length. A detailed discussion of surface breakdown on insulators in vacuum is given in [85].

To reduce the risk of breakdown on the air side of an insulator it is important to keep the surface clean. The surface may be corrugated to increase the path length. The risk may be also be reduced by pressurisation, or by immersing the insulator in a dielectric fluid, such as transformer oil [79].

18.8   Multipactor Discharges

Multipactor is resonant radiofrequency vacuum discharge which is sustained by secondary electron emission [86, 87]. The basic principle can be explained by considering a pair of parallel metal plates in vacuum with an RF voltage between them, as shown in Figure 18.16. Note that the horizontal axis is time. A free electron between the plates at A at a moment when the potential of the upper plate is positive is accelerated towards the upper plate. When it strikes the plate at B it may liberate secondary electrons. If the potential of the upper plate is now negative these electrons are accelerated downwards so that they strike the lower plate at C. There they may liberate further secondary electrons. If the time taken for the electrons to cross from one plate to the other is just half of an RF cycle then the process is repeated and the number of electrons increases each time. In order for a two-surface multipactor discharge to be sustained the RF voltage between the electrodes must give impact energies in the range for which the secondary emission coefficient (δ) of at least one of the surfaces is greater than unity (see Figure 18.7). It is found that electrons, which are emitted from a surface before the optimum phase, experience a smaller accelerating force than those which are emitted after the optimum phase. Thus, the electrons are grouped into a smaller phase range when they arrive at the opposite plate. This phenomenon is known as *phase focusing*. As the intensity of
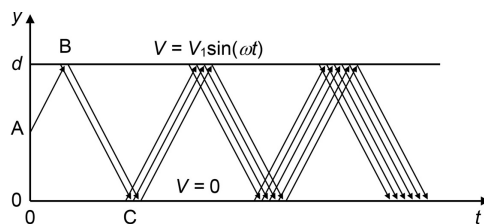
Figure 18.16: Principle of the two-surface multipactor discharge.

the discharge increases the process is limited by space-charge debunching so that the current saturates [88]. In resonant RF structures saturation may also be caused by detuning of the resonator [87]. A two-surface multipactor discharge typically involves currents of less than 1 A and voltages of a few hundred volts. The theory of two-surface multipactor is discussed in Section 18.8.1.

When a magnetic field is present the electron trajectories are bent by the field so that the impacts made on the surfaces are oblique. This type of discharge can involve impacts on either one or two surfaces. When the primary electrons strike a surface obliquely the peak value of δ is greater than for normal incidence and the range of energies over which δ is greater than unity is increased (see Figure 18.7). Thus crossed-field multipactor discharges can occur at much higher energies than the simple multipactor and they are, in consequence, potentially much more damaging. The theory of multipactor in crossed fields is discussed in Section 18.8.3. Because strong magnetic fields are used to focus linear-beam tubes it is quite possible for the conditions for crossed-field multipactor to exist somewhere within the tube. The manufacturer will normally have taken steps to ensure that this is not the case. However, if the magnetic field around the tube is disturbed in any way, for example by the field of a circulator, then it is possible for a destructive discharge to occur. It is also possible for multipactor discharges to occur on ceramic surfaces such as vacuum windows where surface charging produces a static electric field [87]. A discharge of this kind changes the match of the window and the energy transferred may be sufficient to crack the ceramic.

It is probable that multipactor discharges occur in most vacuum tubes. They have been observed in klystrons [86, 89] and magnetrons [90]. They also occur in transmission lines of many kinds including rectangular waveguides [91, 92] and coaxial lines [93, 94], and in components such as windows [95, 96] and filters [97]. Signs of multipactor discharges include heating, and the emission of light and X-rays. They are sources of noise, and can produce appreciable changes in the RF performance of a tube by loading the RF circuit. The energy transferred to the surfaces by the impact of electrons can cause window failure, and local outgassing, which may lead to the formation of arcs. Multipactor discharges can sometimes be suppressed by changing the shapes of surfaces, by surface coatings to reduce the secondary electron yield, and by the imposition of static electric and magnetic fields [86, 87, 98].

18.8.1   Theory of Multipactor Discharges between Parallel Plates

The theory of the multipactor discharge between parallel metal plates has been discussed by a number of authors. The analysis presented here is based on that given in [86, 99][1]. It provides insight into the factors contributing to the existence of multipactor discharges.

Consider two plates located at $y = 0$ and $y = d$ with an RF voltage $V_1 \sin(\omega t)$ between them. The equation of motion for an electron is

$$\frac{d^2 y}{dt^2} = \frac{e V_1}{m_0 d} \sin(\omega t).\tag{18.51}$$

If the electron is injected at $x = 0$ with velocity v_0 when $\omega t = \theta_1$ then its velocity at time t is

$$\frac{dy}{dt} = \frac{e V_1}{m_0 \omega d} (\cos \theta_1 - \cos \omega t) + v_0.\tag{18.52}$$

A second integration gives the position of the electron as a function of time

$$\frac{y}{d} = \frac{e V_1}{m_0 \omega^2 d^2} \{(\omega t - \theta_1)\cos \theta_1 - \sin \omega t + \sin \theta_1\} + \frac{v_0}{\omega d}(\omega t - \theta_1).\tag{18.53}$$

If the electron reaches the second plate when $\omega t = \theta_2$ then

$$1 = \frac{e V_1}{m_0 \omega^2 d^2} \{(\theta_2 - \theta_1)\cos \theta_1 - \sin \theta_2 + \sin \theta_1\} + \frac{v_0}{\omega d}(\theta_2 - \theta_1).\tag{18.54}$$

In order for the discharge to be resonant, the phase of the RF voltage when the electron arrives at the second plate (θ_2) must differ from the initial phase by an odd multiple of π. Thus

$$\theta_2 = \theta_1 + n\pi, \quad n = 1,3,5\dots\tag{18.55}$$

Substituting this condition into (18.54) gives the RF voltage as a function of the initial phase

$$V_1 = \frac{m_0}{e} \cdot \frac{\omega d (\omega d - n\pi v_0)}{n\pi \cos \theta_1 + 2 \sin \theta_1}.\tag{18.56}$$

We note that v_0 only appears in the numerator of (18.56) and θ_1 only in the denominator. Thus, the minimum RF voltage for which multipactor can occur is given by the value of θ_1 which maximises $n\pi\cos \theta_1 + 2\sin \theta_1$ regardless of the value of v_0. This occurs when

$$\tan \theta_1 = \frac{2}{n\pi}.\tag{18.57}$$

---

[1] I am indebted to my colleague Dr Amos Dexter for making his lecture notes on this subject available to me.

If the RF voltage is greater than the minimum, electrons can reach the second electrode with a range of initial phases. An electron whose initial velocity (v_0) is zero can only reach the second plate if it is emitted in the positive half-cycle of the RF voltage. If $v_0 > 0$ then the electron can still reach the second plate if it is emitted into a slightly retarding field, and the limiting phase, beyond which the electron cannot escape from the surface, is negative.

   A sustained discharge can only exist if there is phase focusing as described above. Thus two electrons which leave the first plate at slightly different times arrive at the second plate with a smaller time interval between them. This is expressed mathematically by

$$\left| \frac{d\theta_2}{d\theta_1} \right| < 1. \tag{18.58}$$

If this condition is not satisfied then the electrons drift away from the resonant condition over a number of cycles. Differentiating (18.54) with respect to θ_1 gives

$$\frac{eV_1}{m_0 \omega d} \left\{ \left(\frac{d\theta_2}{d\theta_1} - 1 \right) \cos \theta_1 - (\theta_2 - \theta_1) \sin \theta_1 - \frac{d\theta_2}{d\theta_1} \cos \theta_2 + \cos \theta_1 \right\} + v_0 \left(\frac{d\theta_2}{d\theta_1} - 1 \right) = 0. \tag{18.59}$$

When $d\theta_2 / d\theta_1 = 1$ the limiting phase satisfies

$$\cos \theta_1 = (\theta_2 - \theta_1) \sin \theta_1 + \cos \theta_2. \tag{18.60}$$

Substituting the resonant condition from (18.55) we find that this is identical to the phase given by (18.57). The second limit for the phase is found by setting $d\theta_2 / d\theta_1 = -1$ to give

$$\frac{eV_1}{m_0 \omega d} \left\{ -\cos \theta_1 - (\theta_2 - \theta_1) \sin \theta_1 + \cos \theta_2 \right\} = 2v_0. \tag{18.61}$$

Applying the condition for resonance we obtain

$$2 \cos \theta_1 + n\pi \sin \theta_1 = -\frac{2m_0 \omega d v_0}{eV_1}. \tag{18.62}$$

To eliminate V_1 (18.56) is rewritten as

$$n\pi \cos \theta_1 + 2 \sin \theta_1 = \frac{m_0 \omega d (\omega d - n\pi v_0)}{eV_1}. \tag{18.63}$$

Dividing (18.63) by (18.62), and rearranging gives

$$\tan \theta_1 = -\frac{4n\pi v_0 + 2\omega d}{4v_0 + n\pi \omega d - (n\pi)^2 v_0}. \tag{18.64}$$
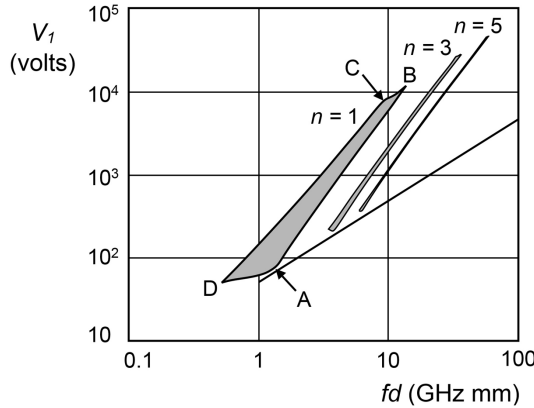
Figure 18.17: Hatch diagram showing the regions in which multipactor is theoretically possible.

Thus, the minimum value of θ_1 is the greater of the two values determined by the condition that the electron can escape from the surface, and by the second condition for phase focusing.

If the RF voltage is increased, at a fixed value of ωd, from the minimum for which a discharge can occur, then θ_1 decreases until it reaches the minimum value described above. The lower and upper limits of the RF voltage between which a discharge is possible are obtained by substituting the two values of θ_1 obtained from (18.57) and (18.64) into (18.56). We may therefore plot V_1 against $fd = \omega d/2\pi$ to show the boundaries of possible regions of multipactor for each mode, as shown in Figure 18.17. The lines representing the limiting phase conditions for the lowest mode are AB and DC.

In order that an electron avalanche can develop the product of the secondary electron emission coefficients of the two plates must be greater than unity. The impact velocity is found by substitution of the resonant condition into (18.52) to give

$$v_i = \frac{2eV_1}{m_0\omega d}\cos\theta_1 + v_0 \tag{18.65}$$

which may be rearranged, eliminating θ_1 using (18.56), to give

$$V_1 = \frac{m_0\omega d}{2e}\sqrt{\left\{\omega d - \frac{n\pi}{2}(v_i + v_0)\right\}^2 + (v_i - v_0)^2}. \tag{18.66}$$

The limits set by this condition are found by inserting the two values of v_i for which the secondary electron emission coefficient is unity into (18.66). These boundaries are shown as DA and CD in Figure 18.18. This plot is known as a Hatch diagram [100, 101].

The boundaries of the three lowest multipactor regions in Figure 18.17 were computed using the theory developed above for comparison with the experimental results in Figure 3 of [101] (see Worksheet 18.1). The lowest experimental breakdown voltages were obtained when the aluminium electrodes had been left at a gas pressure of 2 microns Hg for 24 hours. It is probable that the surfaces were then fully oxidised. In the calculations it was assumed that $\delta_m = 3.5$ and $E_{pm} = 400$ V [23]. The electron emission energy was fixed at 2 eV and the secondary electron emission curve was defined by (18.19) with $s = 1.65$. The calculations predict that multipactor can exist in well defined regions with clear gaps between them. In the experiments it was found that discharges existed at all points above the dashed line in Figure 18.17. This discrepancy can be explained by a number of factors which have been ignored in the simple theory: the spread in the secondary electron energies; the effect of elastically and inelastically reflected electrons; and the possibility of asymmetric modes in which the value of n differs for the outward and backward motion [102–106]. All these effects serve to broaden the bands so that they tend to merge. The boundary below which multipactor cannot occur is then a line which is tangential to the bottoms of the theoretical regions shown in Figure 18.17. The points on that line can be found by eliminating V_1 between (18.56) and (18.65) and using θ_1 from (18.57) to give

$$\omega d(n) = (v_i + v_0)\frac{n\pi}{2} + (v_i - v_0)\frac{2}{n\pi} \qquad (18.67)$$

where v_i is the minimum impact velocity for which $\delta = 1$. The corresponding values of $V_1(n)$ are found by substitution into (18.65).

In the experiments described in [101] the voltage at which the discharge commenced increased when the electrodes were first subjected to a discharge at a higher pressure for half an hour. This presumably reduced the thickness of the oxide layer, and increased the minimum impact voltage for which the secondary electron emission coefficient was unity. The effect can be modelled by reducing the value of δ_m used in the calculations. This raises the boundary line described above. There is then a distinction between those multipactor discharges which can be eliminated by conditioning, and those which cannot. The properties of two-surface multipactor discharges between parallel plates can be explored using Worksheet 18.1.

The assumption of a uniform RF electric field between plane electrodes is necessarily idealised. The experiments of Hatch and Williams used a pair of circular plates without any kind of compensation for the field concentration at the edges. The assumption of a uniform field is a useful approximation when the discharge is in a rectangular waveguide or in a re-entrant cavity resonator [92, 107, 108]. It can be seen from Figure 18.17 that, theoretically, the discharge cannot occur if the RF voltage is great enough. That limit is represented by a line which is tangential to the tops of the multipactor regions. However, even if the maximum RF voltage is greater than the upper limit, its spatial variation ensures there is always some point in the waveguide or cavity where multipactor can occur. Thus, in a non-uniform field the discharge is located at a position where the conditions are favourable.

18.8.2   Multipactor Discharges between Coaxial Cylinders

The other simple geometry in which multipactor can be studied is the vacuum-spaced coaxial line [94, 109, 110]. This case has practical importance in the coaxial waveguides and couplers used with many vacuum tubes. The symmetry of the field allows the problem to be treated as one-dimensional, but it is not possible to find a simple analytical solution. An important difference from the parallel plate case is that the electron transit times are different for inward and outward motion. It is therefore necessary to find trajectories for which the sum of the transit times is an integral number of RF periods. Both classical, and hybrid, two-surface modes are possible, in a manner very similar to those for parallel plates. But, it is found that the modes gradually disappear, as the ratio of the radius of the outer conductor to that of the inner conductor increases, until two-surface multipactor is no longer possible. However, the velocity of an electron starting from the outer electrode can be reversed before it reaches the inner electrode. If it returns to the outer electrode in an integral number of RF periods then a single-surface discharge is possible. Hatch diagrams similar to Figure 18.17 are obtained for both two-surface and single surface discharges [110]. Further information is given in the references cited above.

18.8.3   Multipactor Discharges in Crossed Fields

Static magnetic fields are used to control the motion of the electrons in most vacuum tubes. It is important to consider how multipactor discharges are affected by them. The simplest case is the parallel plate geometry considered above, with the addition of a uniform static magnetic field parallel to the planes of the plates [111]. The derivations of the equations are lengthy and will not be reproduced here. It turns out that both single surface, and double surface, discharges are possible. Figure 18.18 shows examples of single and double-surface trajectories starting from $x = 0$. The impact of the single-surface trajectory in Figure 18.18(a) on the plane at $x = 0$ is normal so that the energy involved is similar to that in a two-surface multipactor without a magnetic field. On the other hand, the trajectories in Figure 18.18(b) make oblique impacts so that the electron energy can be much higher, while the secondary electron yield remains greater than unity (see Figure 18.8). In consequence the energy dissipated can be much greater, and more damaging, than in the other kinds of multipactor [86]. The properties of crossed-field discharges of both kinds can be explored using Worksheet 18.2.

18.8.4   Modelling Multipactor Discharges

In general, the possibility that a multipactor discharge will occur can be explored for any geometry for which the RF electric field can be computed. If the saturated discharge is to be computed then it is necessary to use a full particle in cell (PIC) code so that the modification of the RF fields by the electronic space
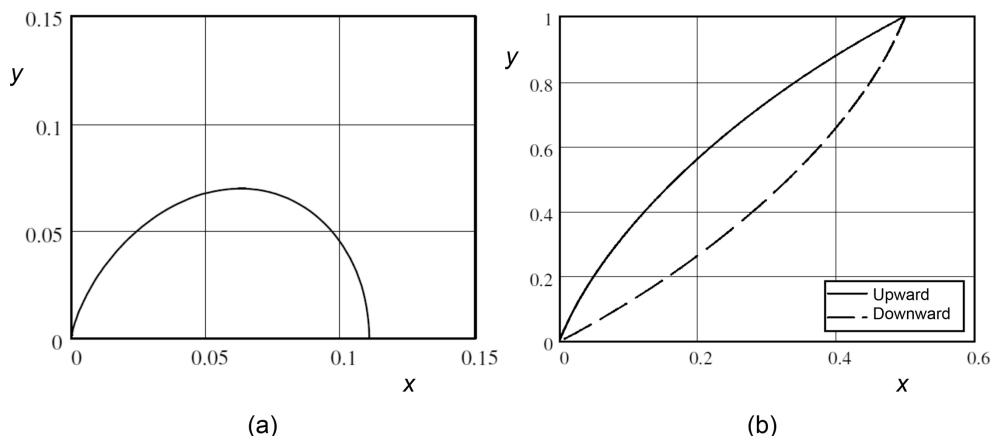
Figure 18.18: Examples of multipactor trajectories in crossed fields: (a) single-surface, and (b) double-surface.

charge can be included. However, if the objective is only to determine whether multipactor is possible, then space-charge effects can be ignored. The electron trajectories can be computed using RF electric field components which have been calculated in advance and stored. The conditions under which the discharge can happen may be quite specific so it is necessary to track sample electrons from a large number of points on the surfaces with a range of starting phases and energies. The spatial and energy distributions of the secondary electrons, and the time delay following the impact of the primary electron, must be represented correctly. Elastically and inelastically reflected primary electrons must also be included. It is evident that a large number of particles are needed at the start of the simulation, and that the number will grow very rapidly with successive impacts. The growth in the number of electrons can be used as an indication of the existence of a discharge. Examples of the simulation of multipactor discharges are given in [94, 108, 112–115].

References

[1] A. J. Dekker, *Solid State Physics*. London: Macmillan, 1963.

[2] G. Gaertner and W. W. P. Koops, 'Vacuum electron sources and their materials and technologies', in J. A. Eichmeier and M. K. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer, pp. 429–481, 2008.

[3] G. W. C. Kaye and T. H. Laby, *Tables of Physical and Chemical Constants.* Available at: www.kayelaby.npl.co.uk/ (accessed 5 October 2017), 2014.

[4] A. Sharma *et al.*, 'Emission poisoning studies on impregnated tungsten dispenser cathode under CO_2 and O_2 environment', *Applied Surface Science,* vol. 40, pp. 97–101, 1989.

[5] W. H. Kohl, *Handbook of Materials and Techniques for Vacuum Devices*. New York: American Institute of Physics, 1995.

[6] J. L. Cronin, 'Modern dispenser cathodes', *IEE Proceedings I: Solid-State and Electron Devices,* vol. 128, pp. 19–32, 1981.

[7] R. H. Fowler and L. Nordheim, 'Electron emission in intense electric fields', *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character,* vol. 119, pp. 173–181, 1928.

[8] I. Brodie and P. R. Schwoebel, 'Vacuum microelectronic devices', *Proceedings of the IEEE,* vol. 82, pp. 1006–1034, 1994.

[9] N. M. Jordan *et al.*, 'Electric field and electron orbits near a triple point', *Journal of Applied Physics,* vol. 102, p. 033301, 2007.

[10] J. A. Eichmeier, 'Radiation sensitive vacuum electronic components and devices', in J. A. Eichmeier and M. Thumm, eds, *Vacuum Electronics: Components and Devices*. Berlin: Springer, pp. 127–154, 2008.

[11] J. W. Gewartowski and H. A. Watson, *Principles of Electron Tubes*. Princeton, NJ: D. van Nostrand, 1965.

[12] J. A. Nation *et al.*, 'Advances in cold cathode physics and technology', *Proceedings of the IEEE*, vol. 87, pp. 865–889, 1999.

[13] R. J. Umstattd, 'Advanced electron-beam sources', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimetre-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 393–443, 2005.

[14] E. Sternglass, 'Theory of secondary electron emission by high-speed ions', *Physical Review,* vol. 108, pp. 1–12, 1957.

[15] J. J. Scholtz *et al.*, 'Secondary electron emission properties', *Philips Journal of Research,* vol. 50, pp. 375–389, 1996.

[16] E. Rudberg, 'Inelastic scattering of electrons from solids', *Physical Review,* vol. 50, pp. 138–150, 1936.

[17] A. Shih *et al.*, 'Secondary electron emission properties of oxidized beryllium CFA cathodes', *IEEE Transactions on Electron Devices,* vol. 41, pp. 2448–2454, 1994.

[18] M. A. Furman and M. T. F. Pivi, 'Probabilistic model for the simulation of secondary electron emission', *Physical Review Special Topics – Accelerators and Beams,* vol. 5, p. 124404, 2002.

[19] D. Chernin *et al.*, 'A model of secondary emission for use in computer simulation of vacuum electronic devices', in *International Electron Devices Meeting*, pp. 773–776, 1993.

[20] H. Seiler, 'Secondary electron emission in the scanning electron microscope', *Journal of Applied Physics,* vol. 54, pp. R1–R18, 1983.

[21] V. Baglin *et al.*, 'A summary of main experimental results concerning the secondary electron emission of copper', CERN, Geneva, 2001.

[22] B. Henrist *et al.*, 'Secondary electron emission data for the simulation of electron cloud', in *Proc. Mini Workshop on Electron Cloud Simulations for Proton and Positron Beams (ECLOUD'02)*, pp. 75–78, 2002.

[23] Y. Lin and D. C. Joy, 'A new examination of secondary electron yield data', *Surface and Interface Analysis,* vol. 37, pp. 895–900, 2005.

[24] P. Tolias, 'On secondary electron emission and its semi-empirical description', *Plasma Physics and Controlled Fusion,* vol. 56, p. 123002, 2014.

[25] V. Baglin *et al.*, 'The secondary electron yield of technical materials and its variation with surface treatments', in *EPAC 2000*, Vienna, Austria, pp. 217–221, 2000.

[26] C. Walker *et al.*, 'The secondary electron emission yield for 24 solid elements excited by primary electrons in the range 250–5000 eV: a theory/experiment comparison', *Scanning,* vol. 30, pp. 365–380, 2008.

[27] C. Walker *et al.*, 'The secondary electron emission yield for 24 solid elements excited by primary electrons in the range 250–5000 eV: a theory/experiment comparison', *Scanning,* vol. 30, pp. 365–380, 2008.

[28] K. Kanaya *et al.*, 'Secondary electron emission from insulators', *Journal of Physics D: Applied Physics,* vol. 11, p. 2425, 1978.

[29] A. Dallos *et al.*, 'Effects of ionized oxygen on primary and secondary emission, and total current of a CFA', *IEEE Transactions on Electron Devices,* vol. 34, pp. 1201–1208, 1987.

[30] J. Lorkiewicz *et al.*, 'Surface TiN coating of TESLA couplers at DESY as an antimultipactor remedy', in *The 10th Workshop on RF Superconductivity*, 2001.

[31] S. Michizono *et al.*, 'Secondary electron emission of sapphire and anti-multipactor coatings at high temperature', *Applied Surface Science,* vol. 235, pp. 227–230, 2004.

[32] G. Fuentes *et al.*, 'Spectroscopic investigations of Cr, CrN and TiCr anti-multipactor coatings grown by cathodic-arc reactive evaporation', *Applied Surface Science,* vol. 253, pp. 7627–7631, 2007.

[33] M. Pivi *et al.*, 'Sharp reduction of the secondary electron emission yield from grooved surfaces', *Journal of Applied Physics,* vol. 104, p. 104904, 2008.

[34] A. N. Curren and K. A. Jensen, 'Secondary electron emission characteristics of iontextured copper and high-purity isotropic graphite surfaces', NASA Lewis Research Center, Cleveland OH, July 1984.

[35] M. Ye *et al.*, 'Suppression of secondary electron yield by micro-porous array structure', *Journal of Applied Physics,* vol. 113, p. 074904, 2013.

[36] V. Nistor *et al.*, 'Multipactor suppression by micro-structured gold/silver coatings for space applications', *Applied Surface Science,* vol. 315, pp. 445–453, 2014.

[37] A.-G. Xie *et al.*, 'Maximum secondary electron yield and parameters of secondary electron yield of metals', *Surface Review and Letters,* vol. 23, p. 1650039, 2016.

[38] B. K. Agarwal, 'Variation of secondary emission with primary electron energy', *Proceedings of the Physical Society,* vol. 71, p. 851, 1958.

[39] J. R. M. Vaughan, 'A new formula for secondary emission yield', *IEEE Transactions on Electron Devices,* vol. 36, pp. 1963–1967, 1989.

[40] G. F. Dionne, 'Effects of secondary electron scattering on secondary emission yield curves', *Journal of Applied Physics,* vol. 44, pp. 5361–5364, 1973.

[41] G. F. Dionne, 'Origin of secondary-electron-emission yield-curve parameters', *Journal of Applied Physics,* vol. 46, pp. 3347–3351, 1975.

[42] M. Salehi and E. Flinn, 'An experimental assessment of proposed universal yield curves for secondary electron emission', *Journal of Physics D: Applied Physics,* vol. 13, p. 281, 1980.

[43] Z. Insepov *et al.*, 'Comparison of candidate secondary electron emission materials', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms,* vol. 268, pp. 3315–3320, 2010.

[44] R. Vaughan, 'Secondary emission formulas', *IEEE Transactions on Electron Devices,* vol. 40, p. 830, 1993.

[45] A. Shih and C. Hor, 'Secondary emission properties as a function of the electron incidence angle', *IEEE Transactions on Electron Devices,* vol. 40, pp. 824–829, 1993.

[46] S. Yu *et al.*, 'Secondary electron emission for layered structures', *Journal of Vacuum Science & Technology A,* vol. 20, pp. 950–952, 2002.

[47] R. Kitchen, *RF and Microwave Radiation Safety Handbook*. Newnes, 2001.

[48] AMPTEK Inc., *Mini-X X-Ray Tube System for XRF*, Available at: http://amptek.com/products/mini-x-ray-tube/#5 (accessed 5 October 2017), 25 July 2017.

[49] J. H. Hubbell and S. M. Seltzer, *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients (version 1.4)*, 2 March. Available at: http://physics.nist.gov/xaamdi (accessed 5 October 2017), 2004.

[50] J. K. Shultis and R. E. Faw, 'Radiation shielding technology', *Health Physics,* vol. 88, pp. 587–612, 2005.

[51] R. Forman, 'Update of thermionic cathode progress', in *International Electron Devices Meeting*, pp. 387–390, 1991.

[52] J. Hasker *et al.*, 'Properties and manufacture of top-layer scandate cathodes', *Applied Surface Science,* vol. 26, pp. 173–195, 1986.

[53] G. Gaertner and D. Barratt, 'Life-limiting mechanisms in Ba-oxide, Ba-dispenser and Ba-Scandate cathodes', in *The 5th International Vacuum Electron Sources Conference*, pp. 59–61, 2004.

[54] R. Tuck, 'Thermionic cathode surfaces: the state-of-the-art and outstanding problems', *Vacuum,* vol. 33, pp. 715–721, 1983.

[55] L. R. Falce, 'Dispenser cathodes: The current state of the technology', in *International Electron Devices Meeting*, pp. 448–451, 1983.

[56] A. Shroff, 'Review of dispenser cathodes', *Revue technique-Thomson-CSF,* vol. 23, pp. 947–1026, 1991.

[57] W. E. Harbaugh, 'Tungsten, thoriated tungsten and thoria emitters', in *Electron Tube Design*. Harrison, NJ: Radio Corporation of America, pp. 90–98, 1962.

[58] T. E. Yingst *et al.*, 'High-power gridded tubes -1972', *Proceedings of the IEEE,* vol. 61, pp. 357–381, 1973.

[59] M. Zhang *et al.*, 'Influence of plasma spraying on the performance of oxide cathodes', *IEEE Transactions on Electron Devices,* vol. 58, pp. 2143–2148, 2011.

[60] R. T. Longo *et al.*, 'Dispenser cathode life prediction model', in *International Electron Devices Meeting*, pp. 318–321, 1984.

[61] R. Longo, 'Physics of thermionic dispenser cathode aging', *Journal of Applied Physics,* vol. 94, pp. 6966–6975, 2003.

[62] R. S. Raju, 'Studies on W-Re mixed-matrix cathodes', *IEEE Transactions on Electron Devices,* vol. 56, pp. 786–793, 2009.

[63] C. A. Spindt *et al.*, 'Field-emitter arrays for vacuum microelectronics', *IEEE Transactions on Electron Devices,* vol. 38, pp. 2355–2363, 1991.

[64] C. Spindt, 'A brief history vacuum nanoelectronics, the IVNC, and the present status of the Spindt cathode', in *25th International Vacuum Nanoelectronics Conference*, pp. 1–2, 2012.

[65] D. R. Whaley *et al.*, 'Application of field emitter arrays to microwave power amplifiers', *IEEE Transactions on Plasma Science,* vol. 28, pp. 727–747, 2000.

[66] D. R. Whaley *et al.*, 'Experimental demonstration of an emission-gated traveling-wave tube amplifier', *IEEE Transactions on Plasma Science,* vol. 30, pp. 998–1008, 2002.

[67] D. R. Whaley *et al.*, '100 W operation of a cold cathode TWT', *IEEE Transactions on Electron Devices,* vol. 56, pp. 896–905, 2009.

[68] D. Whaley *et al.*, 'High average power field emitter cathode and testbed for X/Ku-band cold cathode TWT', in *IEEE International Vacuum Electronics Conference (IVEC)*, pp. 1–2, 2013.

[69] C. Spindt *et al.*, '11.1: A reliable improved Spindt cathode design for high currents', in *IEEE International Vacuum Electronics Conference (IVEC)*, pp. 201–202, 2010.

[70] W. D. Kilpatrick, 'Criterion for vacuum sparking designed to include both RF and DC', *Rev. Sci. Instrum.,* vol. 28, pp. 824–826, 1957.

[71] W. Peter *et al.*, 'Criteria for vacuum breakdown in RF cavities', *IEEE Transactions on Nuclear Science,* vol. 30, pp. 3454–3456, 1983.

[72] A. S. Gilmour, Jr., *Microwave Tubes*. Dedham, MA: Artech House, 1986.

[73] G. Faillon, 'Technical and industrial overview of RF and microwave tubes for fusion', *Fusion Engineering and Design,* vol. 46, pp. 371–381, 1999.

[74] S. Döbert, 'Gradient limitations for high-frequency accelerators', presented at the LINAC 2004, Lübeck, Germany, 2004.

[75] H. H. Braun *et al.*, 'Frequency and temperature dependence of electrical breakdown at 21, 30, and 39 GHz', *Physical Review Letters,* vol. 90, p. 224801, 2003.

[76] G. A. Farrall, 'Vacuum arcs and switching', *Proceedings of the IEEE,* vol. 61, pp. 1113–1136, 1973.

[77] I. I. Beilis, 'State of the theory of vacuum arcs', *IEEE Transactions on Plasma Science,* vol. 29, pp. 657–670, 2001.

[78] J. D. Cobine, *Gaseous Conductors*. New York: Dover, 1941.

[79] A. S. Denholm, 'High voltage technology', *IEEE Transactions on Nuclear Science,* vol. 12, pp. 780–791, June 1965.

[80] D. Xiao, 'Fundamental Theory of Townsend Discharge', in *Gas Discharge and Gas Insulation*. Shanghai: Shanghai Jiao Tong University Press, pp. 47–88, 2016.

[81] J. Townsend, *Electrons in Gases*. London; New York: Hutchinson's Scientific and Technical Publications, 1948.

[82] E. Husain and R. S. Nema, 'Analysis of Paschen curves for air, N_2 and SF_6 using the Townsend breakdown equation', *IEEE Transactions on Electrical Insulation,* vol. EI-17, pp. 350–353, 1982.

[83] J. Kuffel *et al.*, *High Voltage Engineering Fundamentals*. Amsterdam: Newnes, 2000.

[84] R. Arora and W. Mosch, *High Voltage and Electrical Insulation Engineering,* vol. 69. John Wiley & Sons, 2011.

[85] H. C. Miller, 'Surface flashover of insulators', *IEEE Transactions on Electrical Insulation,* vol. 24, pp. 765–786, 1989.

[86] J. R. M. Vaughan, 'Multipactor', *IEEE Transactions on Electron Devices,* vol. 35, pp. 1172–1180, 1988.

[87] R. Kishek *et al.*, 'Multipactor discharge on metals and dielectrics: historical review and recent theories', *Physics of Plasmas (1994 – present),* vol. 5, pp. 2120–2126, 1998.

[88] S. Riyopoulos, 'Multipactor saturation due to space-charge-induced debunching', *Physics of Plasmas (1994–present),* vol. 4, pp. 1448–1462, 1997.

[89] C. Hill and R. G. Carter, 'Investigation of possible multipactor discharge in a klystron input cavity', in *2006 IEEE International Vacuum Electronics Conference Held Jointly with 2006 IEEE International Vacuum Electron Sources*, Monterey, CA, pp. 81–82, 2006.

[90] J. R. M. Vaughan, 'Observations of multipactor in magnetrons', *IEEE Transactions on Electron Devices,* vol. 15, pp. 883–889, 1968.

[91] R. L. Geng *et al.*, 'Suppression of multipacting in rectangular coupler waveguides', *Nuclear Instruments & Methods in Physics Research Section A: Accelerators Spectrometers Detectors and Associated Equipment,* vol. 508, pp. 227–238, 11 August 2003.

[92] R. L. Geng *et al.*, 'Dynamical aspects of multipacting induced discharge in a rectangular waveguide', *Nuclear Instruments & Methods in Physics Research Section A: Accelerators Spectrometers Detectors and Associated Equipment,* vol. 538, pp. 189–205, 11 February 2005.

[93] P. Ylä-Oijala, 'Analysis of electron multipacting in coaxial lines with traveling and mixed waves', Deutsche Elektronen-Synchrotron DESY, MHF-SL Group, 1997.

[94] E. Somersalo *et al.*, 'Analysis of multipacting in coaxial lines', in *Proceedings of the 1995 Particle Accelerator Conference*, pp. 1500–1502, 1995.

[95] J. R. M. Vaughan, 'Some high-power window failures', *IRE Transactions on Electron Devices,* vol. 8, pp. 302–308, 1961.

[96] S. Yamaguchi *et al.*, 'Trajectory simulation of multipactoring electrons in an S-band pillbox RF window', *IEEE Transactions on Nuclear Science,* vol. 39, pp. 278–282, 1992.

[97] A. Woode and J. Petit, 'Investigations into multipactor breakdown in satellite microwave payloads', *ESA Journal,* vol. 14, pp. 467–478, 1990.

[98] C. Chang *et al.*, 'Review of recent theories and experiments for improving high-power microwave window breakdown thresholds', *Physics of Plasmas,* vol. 18, p. 055702, 2011.

[99] V. Shemelin, 'Generalized phase stability in multipacting', *Physical Review Special Topics-Accelerators and Beams,* vol. 14, p. 092002, 2011.

[100] A. J. Hatch and H. B. Williams, 'The secondary electron resonance mechanism of low-pressure high-frequency gas breakdown', *Journal of Applied Physics,* vol. 25, pp. 417–423, 1954.

[101] A. J. Hatch and H. B. Williams, 'Multipacting modes of high-frequency gaseous breakdown', *Physical Review,* vol. 112, pp. 681–685, 1958.

[102] S. Riyopoulos *et al.*, 'Effect of random secondary delay times and emission velocities in electron multipactors', *IEEE Transactions on Electron Devices,* vol. 44, pp. 489–497, 1997.

[103] A. Dexter and R. Seviour, 'Rapid generation of multipactor charts by numerical solution of the phase equation', *Journal of Physics D: Applied Physics,* vol. 38, p. 1383, 2005.

[104] R. Seviour, 'The role of elastic and inelastic electron reflection in multipactor discharges', *IEEE Transactions on Electron Devices,* vol. 52, pp. 1927–1930, 2005.

[105] S. Riyopoulos, 'Higher-order, asymmetric orbit multipactors', *Physics of Plasmas (1994–present),* vol. 14, p. 112101, 2007.

[106] V. E. Semenov *et al.*, 'Importance of reflection of low-energy electrons on multipactor susceptibility diagrams for narrow gaps', *IEEE Transactions on Plasma Science,* vol. 37, pp. 1774–1781, 2009.

[107] D. Proch *et al.*, 'Measurement of multipacting currents of metal surfaces in RF fields', in *Proceedings of the 1995 Particle Accelerator Conference*, pp. 1776–1778, 1995.

[108] V. Semenov *et al.*, 'Multipactor in rectangular waveguides', *Physics of Plasmas (1994–present),* vol. 14, p. 033501, 2007.

[109] R. Woo, 'Multipacting discharges between coaxial electrodes', *Journal of Applied Physics,* vol. 39, pp. 1528–1533, 1968.

[110] R. Udiljak *et al.*, 'Multipactor in a coaxial transmission line. I. Analytical study', *Physics of Plasmas (1994–present),* vol. 14, p. 033508, 2007.

[111] S. Riyopoulos *et al.*, 'Theory of electron multipactor in crossed fields', *Physics of Plasmas (1994–present),* vol. 2, pp. 3194–3213, 1995.

[112] V. Semenov *et al.*, 'Multipactor in a coaxial transmission line. II. Particle-in-cell simulations', *Physics of Plasmas (1994–present),* vol. 14, p. 033509, 2007.

[113] G. Burt *et al.*, 'Benchmarking simulations of multipactor in rectangular waveguides using CST-particle studio', in *SRF 2009*, Berlin, pp. 321–325, 2009.

[114] C. Lingwood *et al.*, 'Phase space analysis of multipactor saturation in rectangular waveguide', *Physics of Plasmas (1994–present),* vol. 19, p. 032106, 2012.

[115] J. W. You *et al.*, 'Highly efficient and adaptive numerical scheme to analyze multipactor in waveguide devices', *IEEE Transactions on Electron Devices,* vol. 62, pp. 1327–1333, 2015.

19   Magnets

19.1   Introduction

Virtually all microwave tubes employ a magnetic field to control the flow of electrons. This chapter reviews the theory and design of both electromagnets and permanent magnets for this purpose. The objective is to create a specified magnetic field in the volume of space in which the electron interaction takes place.

In linear beam tubes, gyrotrons, and long anode magnetrons, the volume is a long thin cylinder. The magnetic field may vary in the axial direction, for example to provide the correct entry conditions at the electron gun, to increase the beam stiffness near the output gap of a klystron, or to allow beam expansion into the collector of a gyrotron. It is important that the field should be cylindrically symmetrical about the axis of the beam. Multiple-beam tubes require a very uniform axial field to minimise the radial field components which can cause the beams to drift in the radial direction. The field is provided either by an electro-magnet, or by one or more permanent magnets [1–3]. Electro-magnets may be normal conducting, or superconducting, and operated pulsed or continuously. They have the advantage that the strength of the magnetic field can be adjusted, as required, for the optimum operation of the tube; however, the power required to energise and cool an electromagnet reduces the overall efficiency of the tube. Permanent magnets do not have these disadvantages, but heavy and expensive magnets are required to produce uniform fields. The size and weight of the magnet system can be reduced substantially by using periodic permanent magnet (PPM) focusing (see Section 7.6). This is common for TWTs and has also been used for a few klystrons.

In crossed field tubes, other than long anode magnetrons, the interaction space is a relatively short annular cylinder. The field, which is normally supplied by a permanent magnet, should be uniform to within 5% to 15%, and be symmetrical about the axis to better than 3% [2].

The chapter begins with a review of the theory of magneto-statics, including the properties of iron and other ferromagnetic materials. This leads to a discussion of the use of magnetic circuits for analysis and design. The properties of common ferromagnetic materials are reviewed in Section 19.4. The remainder of the chapter considers the properties, and design, of three types of magnet:

• In coil dominated magnets (Section 19.5) the magnetic field and its distribution in space are determined by the positions of conducting coils, and the currents in them.

- In iron dominated magnets (Section 19.6) the magnetic field is produced by currents in conducting coils but its distribution in space is determined by the shape of iron yokes and pole-pieces.
- In permanent magnet systems (Section 19.7) the magnetic field is provided by one, or more, permanent magnets, while iron yokes and pole-pieces are used to determine its distribution in space.

19.2   Review of Theory

The theory of electro-magnets and permanent magnets is discussed in textbooks on electromagnetism [4–6]. The main results are reviewed below for convenience. Paramagnetic and diamagnetic effects are too small to be of importance for the applications considered in this book. Therefore the magnetic properties of all materials, except ferromagnetic materials (iron, cobalt, nickel and their alloys), will be assumed to be those of free space. Two vectors are used to represent magnetic fields: the magnetic flux density $\mathbf{B}$ and the magnetic field $\mathbf{H}$. In a linear magnetic material they are related by

$$\mathbf{B} = \mu\mathbf{H} = \mu_0\mu_r\mathbf{H}, \tag{19.1}$$

where μ is the permeability of the material, $\mu_0 = 4\pi\times10^7 \text{ H m}^{-1}$ is the *primary magnetic constant* (or *permeability of free space*), and μ_r is the *relative permeability* of the material. The normal component of $\mathbf{B}$ and the tangential component of $\mathbf{H}$ are continuous at boundaries between regions occupied by different materials.

The magnetic flux density generated by a current element $I\,\mathbf{dl}$ at a point P, whose distance from the element is r, is given by

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi r^2}\left(\mathbf{dl} \wedge \hat{\mathbf{r}}\right), \tag{19.2}$$

where $\hat{\mathbf{r}}$ is a unit vector directed from the current element to P. The vector product ensures that the vector $d\mathbf{B}$ is normal to the plane containing $\mathbf{dl}$ and $\hat{\mathbf{r}}$. A steady electric current must flow in a closed circuit and, taking the integral of (19.2) around the circuit, we obtain the Biot-Savart Law

$$\mathbf{B} = \frac{\mu_0 I}{4\pi}\oint\frac{\mathbf{dl} \wedge \hat{\mathbf{r}}}{r^2}. \tag{19.3}$$

This can be used to compute the flux density, at any point, generated by an arrangement of current-carrying conductors in free space.

The magnetic flux and field vectors obey Maxwell's integral equations

$$\oiint\mathbf{B}\cdot\mathbf{dS} = 0, \tag{19.4}$$

where the integral is taken over a closed surface, and

$$\oint \mathbf{H} \cdot \mathbf{dl} = \iint \left(\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot \mathbf{dS}, \tag{19.5}$$

where $\mathbf{J}$ is the current density, and $\mathbf{D} = \varepsilon \mathbf{E}$ is the electric flux density. The integral on the left-hand side of (19.5) is taken around a closed path, and the integral on the right-hand side is taken over a surface which spans the path. For steady currents flowing in conductors the right-hand side is the total current enclosed by the path of integration. The equation can then be written

$$\oint \mathbf{H} \cdot \mathbf{dl} = \sum_{n=1}^{N} I_n \tag{19.6}$$

where N is the number of conductors and the current in the nth conductor is I_n. This form is known as the Magnetic Circuit Law. The application of this law to the analysis and design of magnets is discussed in Section 19.3.

The differential forms of equations (19.4) and (19.5) are

$$\nabla \cdot \mathbf{B} = 0 \tag{19.7}$$

and

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \tag{19.8}$$

It is convenient to define a magnetic scalar potential U by analogy with the electro-static potential so that

$$\mathbf{H} = -\nabla U. \tag{19.9}$$

The magnetic scalar potential difference between two points is

$$U_2 - U_1 = -\int_1^2 \mathbf{H} \cdot \mathbf{dl}, \tag{19.10}$$

where the integral is taken along any convenient path joining them. The magnetic scalar potential difference only has a unique value if the path does not encircle any current. In that case, substituting into (19.7) from (19.9) we obtain

$$\nabla \cdot \mu \nabla U = 0. \tag{19.11}$$

For a homogeneous material μ is constant and U satisfies Laplace's equation

$$\nabla^2 U = 0. \tag{19.12}$$

Then the magnetic field can be determined from the gradient of the scalar potential.

19.2.1   Ferromagnetism

Iron, cobalt, and nickel, and alloys based on these metals, are known as *ferromagnetic materials*. They have magnetic properties which make them important for

engineering applications. The flux density in a ferromagnetic material depends not only on the magnetising field to which it is subjected, but also on the history of magnetisation. This is displayed in a $B - H$ curve, as shown in Figure 19.1(a). If the specimen is initially un-magnetised, then its state is represented by the point a. When the magnetising field H_i is gradually increased the curve a-b, known as *initial magnetisation curve*, is followed. The flux density can be written

$$B_i = J + \mu_0 H_i,$$ (19.13)

where J is the magnetic polarisation of the material. At high values of H_i the polarisation reaches saturation (J_s) and thereafter an increase in H_i only serves to increase B_i by the change in $\mu_0 H_i$. If H_i is then reduced to zero, the flux density does not retrace the curve a-b. Instead, it follows a curve such as b-c, so that the flux density in the circuit does not fall to zero when the magnetising field is zero. The value of B_i at c is known as the *remanence* (B_r) of the material. In order to reduce the flux density to zero it is necessary to reverse the direction of H_i. The value of H_i at which B_i becomes zero (point d) is known as the *coercive force* (H_c). Increasing the magnitude of H_i beyond the value needed to reach d eventually produces saturation of the material at e with the direction of the flux opposite to that at b. Finally, reducing the magnitude of H_i to zero and then increasing it with positive polarity produces the curve e-f-g-b. This behaviour, in which the magnetisation lags behind the magnetising field, is known as *hysteresis*. When the magnetising field is repeatedly taken through the same cycle, the *hysteresis loop b-c-d-e-f-g-b* is traversed repeatedly in a stable manner, provided only that the maximum magnitude of H_i is the same for both polarities. If, on the other hand, the magnitude of the magnetising field is smaller, it is found that the behaviour of the material is described by a smaller loop, such as the one shown dotted in Figure 19.1(a). From a practical point of view the main division is between 'hard' and 'soft' magnetic materials, as shown in Figure 19.1(b). Soft magnetic materials are easy to magnetise and demagnetise, whereas both processes are more difficult for hard magnetic materials. Soft magnetic materials are used as conductors of magnetic flux, as described in the next section. Hard magnetic materials are used as permanent magnets. The properties of some common ferromagnetic materials are reviewed in Section 19.4.

19.2.2   Conduction of Magnetic Flux by Soft Magnetic Materials

Figure 19.2 shows a magnetic flux line as it passes from air into a material such as pure iron which has a high relative permeability. Applying the boundary conditions gives

$$B_1 \cos \alpha_1 = B_2 \cos \alpha_2$$ (19.14)

and

$$\frac{B_1}{\mu_0} \sin \alpha_1 = \frac{B_2}{\mu_0 \mu_r} \sin \alpha_2.$$ (19.15)
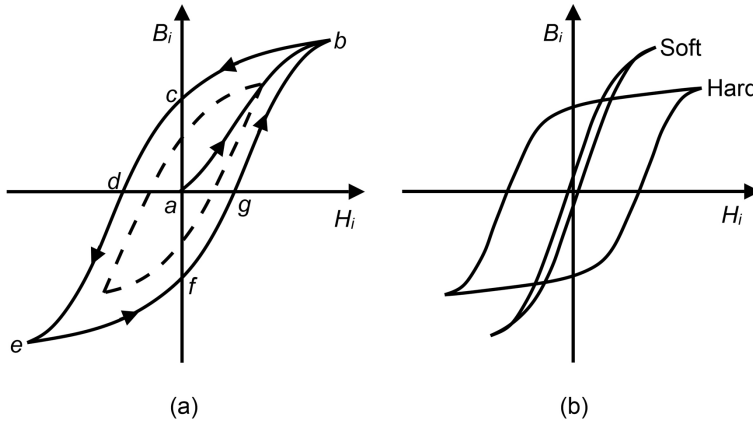
Figure 19.1: (a) A typical hysteresis loop showing the initial magnetisation curve and (b) typical hysteresis loops for hard and soft magnetic materials.
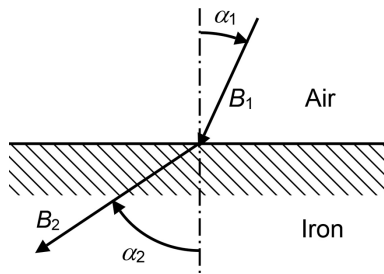
Figure 19.2: Change in the direction of a magnetic flux line at a boundary between air and iron.

Dividing (19.15) by (19.14) we obtain

$$\tan \alpha_1 = \frac{1}{\mu_r} \tan \alpha_2 . \qquad (19.16)$$

For soft magnetic materials μ_r is large so that $\alpha_1 \ll \alpha_2$. For example, if $\mu_r = 5000$ and $\alpha_2 = 89°$ then $\alpha_1 = 0.7°$. Thus the flux lines in air are very nearly normal to the magnetic surface. It is therefore possible to treat soft magnetic materials as perfect magnetic conductors, to a good approximation. Laplace's equation can then be used to find the distribution of a magnetic field in the air spaces between sections of soft magnetic material.

19.3   Magnetic Circuits

The analogy between the conduction of electricity by conductors, and the conduction of magnetic flux by iron, leads to the useful concept of the magnetic circuit.
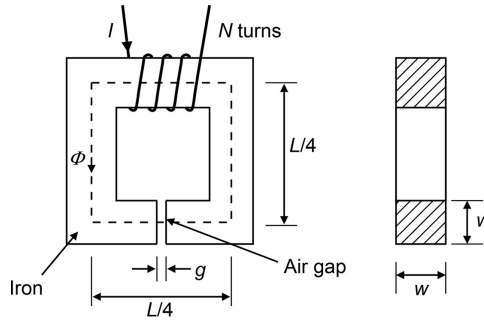
Figure 19.3: A simple magnetic circuit made up of an iron core with an air gap in it. The flux is supplied by a winding of N turns of wire.

Figure 19.3 shows a simple magnetic circuit formed by a square iron core with a narrow air gap in it. A coil of N turns of wire is wound on the core and carries current I. To make the problem easy to handle we make some simplifying assumptions. Let us suppose that the magnetic field strength H_i is the same everywhere within the iron, and the magnetic field in the air gap is H_a. Now consider a closed path around the circuit which follows the centre line (shown dotted). Applying the magnetic circuit law (19.6) to this path gives

$$LH_i + gH_a = NI \qquad (19.17)$$

where L is the length of the path in iron. If, for the moment, we assume that the cross-sectional area of the air gap is equal to that of the iron then

$$B_i = B_a = \frac{\Phi}{A}, \qquad (19.18)$$

where Φ is the total flux circulating, and A is the cross-sectional area of the iron at right angles to the direction of the magnetic field. We have therefore assumed that all the flux due to the coil is contained within the bounds of the iron core and its projection across the air gap. Making use of (19.1) we can write $B_i = \mu H_i$ and $B_a = \mu_0 H_a$. Substituting for H_i and H_a in Equation (19.17) in terms of Φ and the various constants gives

$$NI = \frac{1}{A}\left(\frac{L}{\mu} + \frac{g}{\mu_0}\right)\Phi, \qquad (19.19)$$

which can be written

$$\mathsf{M} = \mathsf{R}\Phi, \qquad (19.20)$$

where M is known as the *magneto-motive* force and R as the *reluctance*. Equation (19.20) is analogous to Ohm's law for electric circuits, but the analogy must not be pressed too far. Unlike an electric current, the magnetic flux in the circuit dissipates no energy, and there are no circulating magnetic charges. Moreover, it cannot be
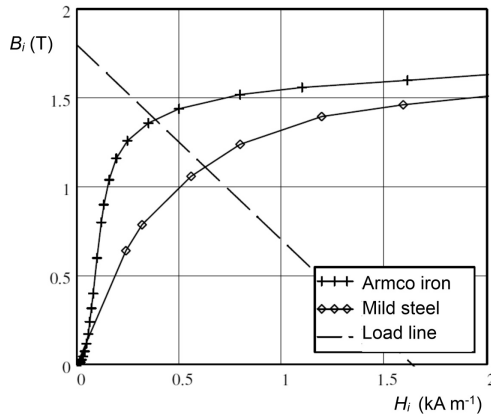
Figure 19.4: Initial magnetisation curves of Armco iron and mild steel showing a typical load line (dashed).

emphasised too often that magnetic materials do not normally behave in a linear fashion, so the reluctance of a circuit is only approximately constant, and then only under a limited range of conditions.

If the permeability of the iron is not constant, the application of the magnetic circuit law to the circuit in Figure 19.3 gives

$$NI = LH_i + gH_a = LH_i + \frac{gB_i}{\mu_0} \tag{19.21}$$

or

$$B_i = \frac{\mu_0 L}{g}\left(\frac{NI}{L} - H_i\right). \tag{19.22}$$

But we know that B_i and H_i must be related to each other by the initial magnetisation curve of the iron

$$B_i = f\left(H_i\right). \tag{19.23}$$

The non-linear simultaneous equations (19.22) and (19.23) can be solved graphically, or by using an analytical approximation to the equation of the hysteresis loop. The graphical solution is shown in Figure 19.4. The working point is at the intersection of the straight line represented by (19.22) with the initial magnetisation curve. The intercept of this load line on the horizontal axis is the magneto-motive force in the circuit divided by the path length, while its slope is determined by the relative sizes of the iron path and the air gap.

The preceding discussion of magnetic circuits made some rather crude assumptions about the distribution of the magnetic flux. A more realistic view is shown in Figure 19.5(a). The flux lines in the air gap spread out into a *fringing field* as shown. Because the magnetic flux in the circuit is constant, the fringing reduces the flux

density in the air gap and, hence, the reluctance of the gap. A second difficulty in calculating the flux in the air gap is that some of the flux simply takes a short cut such as A-A' and never reaches the gap at all. This flux is known as *leakage flux*. It exists because, although iron conducts flux better than air, the air is a poor magnetic insulator. If we wish to pursue the analogy with electric circuits it is necessary to think of the material surrounding the conductor as being a poor conductor rather than an insulator. Fortunately, in most cases, a very crude estimate of the leakage flux is adequate to provide a satisfactory value for the flux in the air gap. The flux leakage can be minimised by placing the source of the magneto-motive force as close to the air gap as possible. Thus, the magnetising coil can be replaced by a pair of coils close to the air gap, as shown in Figure 19.5(b). It can be seen that the leakage flux is then greatly reduced. Estimates of the fringing and leakage fluxes can be made using formulae to be found in the literature [7–9] (see also Section 19.7.1). These methods are only suitable for the creating an initial design and computational magnetics software must be used to obtain accurate results [10]. Figure 19.5 and similar images in this chapter were produced using Gemini 1.2 [11].

19.3.1   Circuits Including Permanent Magnets

We have already noted that hard magnetic materials can be magnetised so that they produce a substantial magnetic flux even when the magnetising field is removed. These materials are used for making permanent magnets which are used especially in magnetrons, CFAs, and PPM focused TWTs. Figure 19.6 shows a portion of a typical B–H plot for a permanent magnet material. This curve, which is the part of the hysteresis loop lying in the second quadrant, is known as the *demagnetisation curve* of the material.

Suppose that the circuit shown in Figure 19.5 is made of this material and magnetised to saturation by passing a current through the coil when the air gap is bridged by a piece of soft iron. When the current in the coil is reduced to zero the working point will be at P in Figure 19.6. If the soft iron is removed from the gap, I is set to zero in (19.22), and fringing and leakage fluxes are neglected, we have

$$B_i = -\left(\frac{\mu_0 L}{g}\right) H_i. \tag{19.24}$$

This load line passes through the origin and has negative slope. It is plotted in Figure 19.6 as the line OQ. The effect of opening the gap is therefore to demagnetise the magnet to some extent. For this reason it is common to speak of the *demagnetising field* of the air gap.

If the piece of soft iron is reinserted in the air gap, the operating point of the magnet moves along a *minor loop* (or *recoil line*) to the point P'. In this state the magnet is largely immune to the effects of external fields. For this reason it is usual to store a permanent magnet with the air gaps bridged by a piece of soft iron which is known as a 'keeper'. If the keeper is absent, and the magnet is exposed to an external demagnetising field, then the working point might move to R. When the
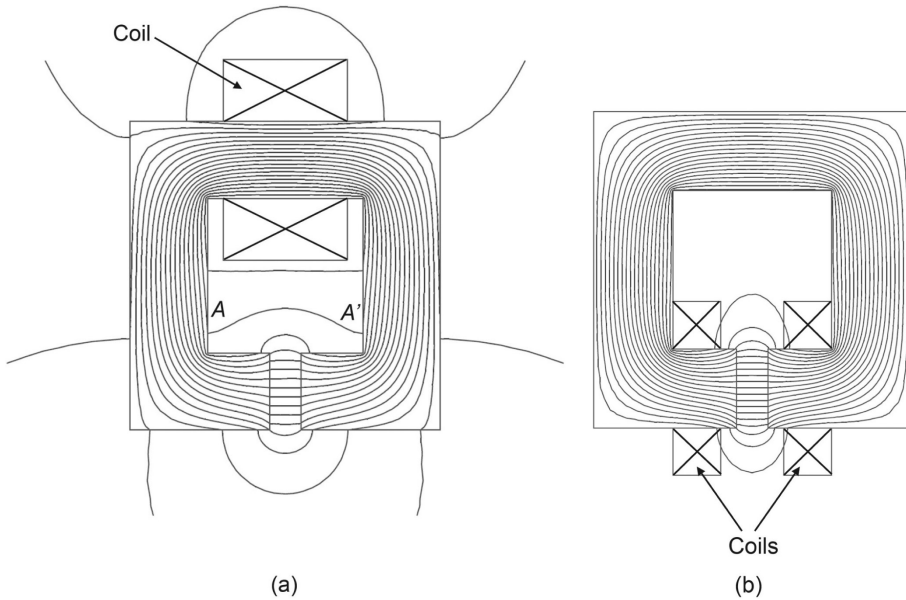
Figure 19.5: (a) An electromagnet showing the fringing field around the air gap and a typical flux leakage path $A - A'$, and (b) an alternative arrangement of the coil to minimise the leakage flux.
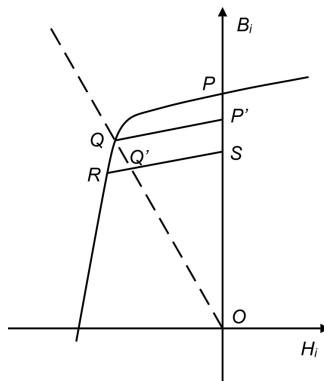
Figure 19.6: The working point for a permanent magnet circuit lies in the second quadrant of the B–H plot at a point such as Q. A permanent magnet can be stabilised against demagnetisation by external fields by operating it at a point Q' on a minor loop such as R–S.

external field is removed the working point now lies at Q' on the minor loop RS. This illustration shows that it is necessary to treat permanent magnets carefully if their properties are to be preserved. If the demagnetising field is strong enough for the point R to move into the third quadrant then the magnet may be completely

demagnetised. It is undesirable for the working point of the magnet to be sensitive to external influences in this way and permanent magnets are often stabilised by deliberately demagnetising them beyond their working points on the main hysteresis loop. The operation of a magnet which has been stabilised is then along a minor loop such as *RS* unless the external demagnetising field is very strong. Smaller external fields produce a temporary shift in the working point, but it returns to *Q'* when the perturbing field is removed.

In a circuit involving a permanent magnet, the cost of the magnet is usually a considerable part of the cost of the whole circuit. It is, therefore, desirable to use the magnet material as efficiently as possible. A formula for this purpose can be derived by considering a uniform magnetic circuit with an air gap. Let the lengths of the magnet and the air gap be l_i and l_a and let their cross-sectional areas be A_i and A_a. The magnetic circuit law gives

$$l_i H_i + l_a H_a = 0 \qquad (19.25)$$

and the conservation of flux gives

$$A_i B_i = A_a B_a \qquad (19.26)$$

so that

$$V_i B_i H_i = -V_a B_a H_a, \qquad (19.27)$$

where $V_i = l_i A_i$ and $V_a = l_a A_a$ are, respectively, the volumes of the iron and the air gap. Now the density of the stored energy in a magnetic field is $BH/2$ so (19.27) shows that magnitude of the stored energy in the magnet must equal that in the air gap. Thus the minimum magnet volume to provide a stored energy in a given air gap is achieved by making the *energy product* of the magnet $(B_i H_i)$ as great as possible. This is known as *Evershed's criterion*. Figure 19.7 shows curves of H_i and $B_i H_i$ against B_i for a typical permanent magnet material. The optimum working point of the material is then Q_{opt} where the value of $B_i H_i$ is greatest. It is usual to design the magnetic circuit so that the magnet operates near this point.

19.4   Magnetic Materials

The properties of magnetic materials can be varied by making alloys with different proportions. They may be isotropic, or anisotropic, and are strongly influenced by the ways in which the finished material is prepared, especially any heat treatments used. A very great variety of materials now exists whose properties can be looked up in the literature [7, 12, 13] and in manufacturers' catalogues.

19.4.1   Soft Magnetic Materials

Soft magnetic materials are characterised by narrow hysteresis loops, low remanence and small coercive force. The area of the hysteresis loop is proportional to
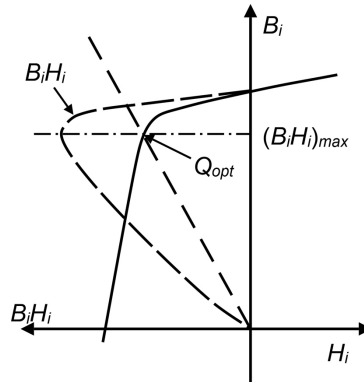
Figure 19.7: Typical curves of H_i and B_iH_i versus B_i for a permanent magnet material showing the choice of working point at the maximum value of B_iH_i to make the most efficient use of the material.

the work required to move the working point around it. Thus, narrow loops are needed to minimise the generation of heat through hysteresis loss in transformer cores and other AC applications [14]. Microwave tubes generally use static magnetic fields and hysteresis loss does not occur. The properties of soft magnetic materials can be approximated by their initial magnetisation curves. Figure 19.4 shows two examples. The curves are nearly straight for fields well below saturation, so it is possible to make the approximation that the relative permeability is constant. The relative permeability decreases with increasing temperature, falling to zero at the Curie temperature (T_c). Soft magnetic materials are used in vacuum tubes for pole-pieces, for shaping the magnetic field for electron guns, for yokes for electromagnets, and for magnetic shielding. It is useful to remember that materials such as nickel alloys, which are sometimes used for other purposes in tube construction, are also ferromagnetic and may affect the magnetic field close to them. Typical properties of soft magnetic materials are given in Table 19.1.

19.4.2 Permanent Magnet Materials

The choice of permanent magnets for use in microwave tubes is based on magnetic strength, operating temperature, corrosion resistance, stability, and cost. The properties of these materials can be changed irreversibly by vibration and shock, and by exposure to high temperatures. The available materials fall into four main classes as shown in Table 19.2 [3, 16]. Typical demagnetisation curves are shown in Figures. 19.8 and 19.9.

Ferrite magnets are made from mixed oxides of iron, and certain other metals, sintered at a high temperature to form an insulating ferromagnetic compound. They are inexpensive and have good stability but the energy product and the maximum working temperature are low. Forcing two magnets together in repulsion can cause demagnetisation of up to 10%. Below –40 °C they suffer from large irreversible

Table 19.1: Typical properties of soft magnetic materials used in microwave tubes [15]

Material	μ_r (max)	J_s (T)	H_c (A m^{-1})	B_r (T)	T_c (°C)
Armco iron (99.85% Fe)	7000	2.16	80	1.3	770
Mild steel	2000	2.15	150	–	–
Nickel	600	0.615	400	–	358
Mumetal	240,000	0.77	1.0	0.45	350

Table 19.2: Typical properties of permanent magnet materials [17]

	Ferrite	Alnico	Samarium cobalt	Neodynium iron boron
BH_{max} (kJ m^{-3})	26	40	210	280
B_{opt} (T)	0.14	0.13	0.35	0.45
B_r (T)	0.22–0.40	0.80–1.26	0.85–1.0	1.2
H_c (kA m^{-1})	110–240	40–155	700–800	820–1040
Density (g cm^{-3})	4.9	7.3	8.4	7.5
Working temperature (°C)	250	550	300	120
Reversible change in B_r (C^{-1})	0.19%	0.02%	0.03%	0.12%
Corrosion resistance	Excellent	Fair	Excellent	Poor
Ease of demagnetisation	Low	High	Very low	Very low
Cost comparison	1	5	20	10
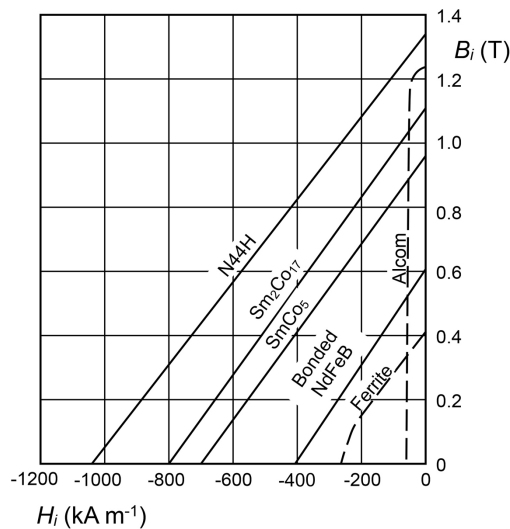
Figure 19.8: Typical demagnetisation curves of permanent magnet materials (copyright 2010 Delta Magnets, reproduced, with permission, from [16]).
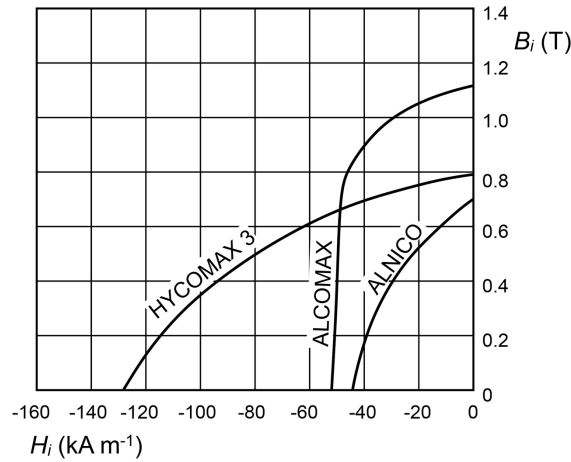
Figure 19.9: Typical demagnetisation curves of Alnico magnets (copyright 2010 Delta Magnets, reproduced, with permission, from [16]).

losses in magnetisation. These materials are brittle and must be shaped by grinding, which is an expensive process.

Alnico materials are a group of cast alloys of iron with aluminium, cobalt, nickel and other metals. The preferred magnetic axis is determined during manufacture. They are relatively cheap, can be machined and have a high working temperature. Because the coercive force is low they are easily demagnetised. Forcing two magnets together in repulsion can cause demagnetisation by up to 50%. For this reason it is usual to stabilise them. They should be stored with a keeper in place to avoid demagnetisation, and are best magnetised after assembly [18].

Samarium cobalt, and neodymium iron, magnets are known collectively as rare earth magnets. Samarium cobalt magnets have a high energy product, excellent stability, and a moderate working temperature. They are hard, brittle, materials which are manufactured by sintering and must be shaped by grinding or abrasive cutting. Their corrosion resistance is excellent but they will break down if exposed to hydrogen. The cost of these magnets is high but they are very useful when it is important to minimise the volume of the magnet. Neodymium iron boron magnets have the highest energy product of all and their cost is lower than that of samarium cobalt magnets. However, the maximum working temperature is too low for most vacuum tube applications and they break down when exposed to hydrogen.

19.5   Coil Dominated Magnets

The magnetic field generated by a system of coils without any iron pole-pieces is sometimes used for experimental purposes [19]. It is also relevant to the design of superconducting magnets where iron cannot be used because the high flux densities would saturate it [20, 21]. The simplest example is a circular current loop of radius
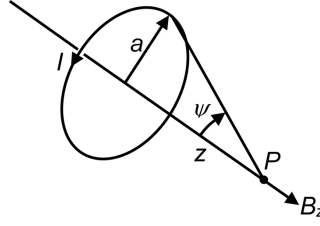
Figure 19.10: Calculation of the flux density on the axis of a circular wire loop carrying current I.
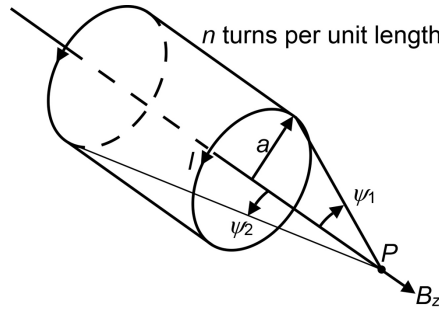
Figure 19.11: Calculation of the flux density on the axis of a single-layer solenoid having n turns per unit length carrying current I.

a carrying current I, as shown in Figure 19.10. It can be shown that the flux density at a point P on the axis distant z from the plane of the loop is given by

$$B_z = \frac{\mu_0 I}{2a} \sin^3 \psi = \frac{\mu_0 I}{2a} \cdot \frac{1}{\left(1 + z^2/a^2\right)^{\frac{3}{2}}}. \tag{19.28}$$

The calculation of the flux density at points which do not lie on the axis is more difficult because it requires the use of elliptic integrals [22].

A second simple case is the single-layer solenoid of finite length shown in Figure 19.11. The flux density on the axis is given by

$$B_z = \frac{\mu_0 n I}{2}\left(\cos \psi_2 - \cos \psi_1\right), \tag{19.29}$$

where n is the number of turns per unit length. Note that if P is inside the solenoid then the sign of $\cos \psi_1$ is reversed. The flux density at the centre of a solenoid of length L is then

$$B_z = \frac{\mu_0 n I L}{\sqrt{L^2 + \left(2a\right)^2}}. \tag{19.30}$$

If $L \to 0$ then B_z tends to the value given by (19.28) at $z = 0$. When $L \to \infty$ then $B_z \to \mu_0 n I$ which can also be derived by the use of the magnetic circuit law (19.5).
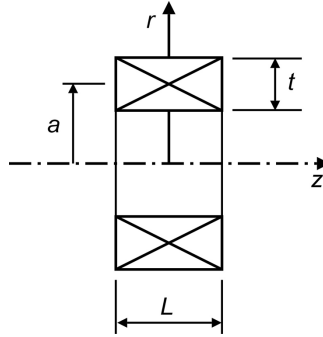
Figure 19.12: Arrangement of a coil having finite cross-sectional area.
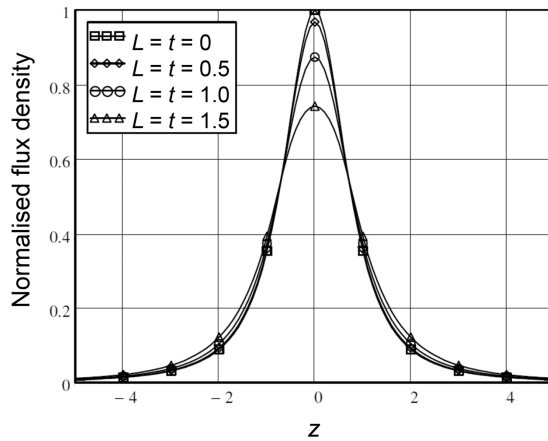
Figure 19.13: Comparison between the normalised axial flux densities of coils having square cross-sections of different sizes $(a = 1)$.

Practical coils differ from the idealised current loop and solenoid described above because the windings have finite length and thickness in all cases, as shown in Figure 19.12. The flux density of a practical coil having N turns is calculated by summing the contributions of all the turns to the total flux density (see Worksheet 19.1). To find the field on the axis of this coil it is convenient to normalise all dimensions to the mean radius (a). Then the flux density on the axis, normalised to that at the centre of a current loop, is given by

$$B_z(z) = \frac{1}{Lt} \int_{-\frac{L}{2}}^{\frac{L}{2}} \left\{ \int_{1-\frac{t}{2}}^{1+\frac{t}{2}} \left(1 + \frac{(z-x)^2}{r^2} \right)^{-\frac{3}{2}} dr \right\} dx. \qquad (19.31)$$
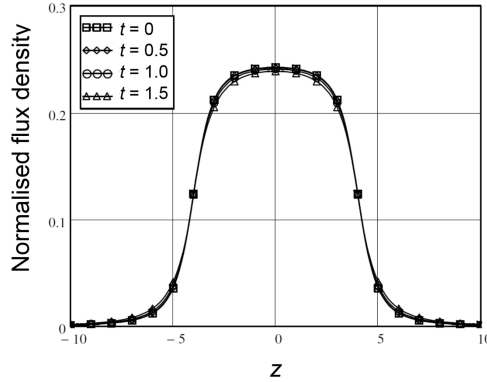
Figure 19.14: Comparison between the normalised axial flux densities of solenoids having different thicknesses $(a = 1, L = 8)$.

Figure 19.13 shows a comparison of the flux density on the axis of coils having square cross-section. It can be seen that the flux density can be approximated quite closely by that of a current loop even when the cross-sectional dimensions of the coil are comparable with the mean radius. Figure 19.14 shows similar results for a solenoid for which $L = 8a$. All the cases plotted can be modelled quite accurately by a single-layer solenoid

19.5.1   Arrays of Coils

The flux density produced by two or more coils can be calculated using the principle of superposition. The flux density on the axis of a circular current loop located at $z = d$ is obtained from (19.28) as

$$B_z = \frac{\mu_0 I}{2} \cdot \frac{a^2}{\left(a^2 + (z - d)^2\right)^{\frac{3}{2}}}. \tag{19.32}$$

This expression can be expanded as a Taylor series about the origin as

$$B_z(z) = b_0 + b_1 z + b_2 z^2 + \cdots \text{etc.,} \tag{19.33}$$

where

$$b_n = \frac{1}{n!}\left(\frac{d^n B_z}{dz^n}\right)_{z=0}. \tag{19.34}$$

When a pair of coils is arranged symmetrically with respect to the origin the odd order terms cancel. Using (19.34) it is found that

$$b_2 = \frac{3a^2 \left(a^2 - 4d^2\right)}{\left(a^2 + d^2\right)^{7/2}}. \tag{19.35}$$
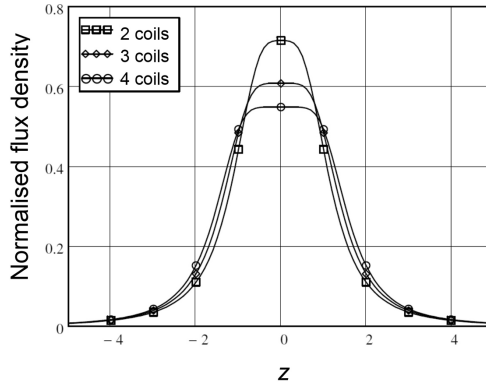
Figure 19.15: Normalised, maximally flat, axial flux density distributions for two, three and four coils when the total current is unity $(a = 1)$.
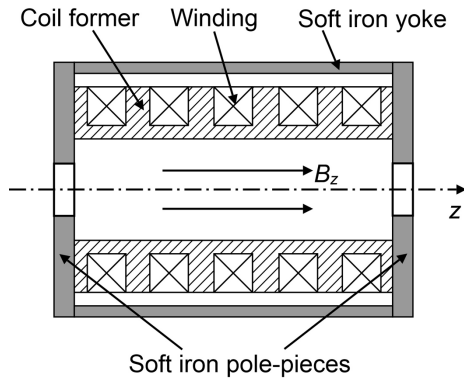
Figure 19.16: Arrangement of a solenoid for a microwave tube.

Thus setting $d = a/2$ produces a maximally flat field at the origin giving the well-known Helmholtz coils. The region of uniform field can be extended by the addition of more coils and by using equal ripple solutions [23–25]. Figure 19.15 shows the maximally flat results for two, three and four coils when the total current is unity (see Worksheet 19.1). The method has also been extended to coils of finite cross-section [26, 27]. The design of superconducting magnets to produce a long length of uniform axial magnetic field is described in [20, 21]. Detailed information about the design of coil dominated magnets can be found in [28, 29].

19.5.2 Solenoids

Figure 19.16 shows a typical arrangement of a cylindrically symmetrical solenoid for a microwave tube. The air space is bounded at each end by a soft iron pole-piece and the return path for the flux is provided by an external yoke. This may take the

form of an iron cylinder or of a number of iron bars. The magneto-motive force is provided by a winding made of copper or aluminium wire or tape [1, 3, 28, 30–32]. It is possible to wind the coils directly onto the body of the tube but this is a difficult process which requires the complete tube to be spun in a lathe and can cause distortion of the tube [30]. A wound-on solenoid limits the temperature at which the tube can be baked during evacuation and increases the cost of replacing a failed tube. Thus, it is much more common to wind the solenoid on a separate former (as shown in Figure 19.16) and arrange for the tube to be inserted into it. Part of each pole-piece is incorporated in the tube body and care must be taken to ensure that there is a good magnetic joint between the sections of each pole-piece when the tube is installed. If the tube fails it can be replaced without changing the solenoid which remains a permanent part of the equipment in which it is installed.

The basic design of a solenoid as shown in Figure 19.16 is very simple. For the time being we ignore the holes in the pole-pieces which are necessary in a linear beam tube for the passage of the electron beam. The effect of these holes is discussed in Section 9.3.1. The inner faces of the pole-pieces can be regarded as perfect magnetic conductors to a good approximation, and the problem is solved using the method of images. The region between the pole-pieces is regarded as a section of an infinitely long uniform solenoid which therefore has a uniform flux density (B_0) within it. The application of the magnetic circuit law (19.6) then gives

$$\mu_0 H_0 L = NI, \tag{19.36}$$

where the solenoid winding has N turns, each carrying current I. The shape and cross-sectional area of the conductors are chosen to give a convenient combination of voltage and current for the solenoid, bearing in mind the space needed for insulation. The current density in the winding can be chosen independently depending on the trade-off between size and cooling requirements. The current density is typically 2 to 3 A cm^{-2}. It is straightforward to calculate the resistance of the solenoid and, hence, the power dissipated in it. For pulsed operation it is also necessary to calculate the self-inductance. The design of the solenoid must include a thermal analysis, and arrangements for cooling, to ensure that the local temperature is never high enough to damage the insulation between the turns of the winding. The design of the pole-pieces and the yoke should include checks to ensure that the flux density is never sufficient to cause local magnetic saturation because that would cause the flux inside the solenoid to be non-uniform, and would increase the leakage flux outside it.

In practice a solenoid usually comprises a number of separate coils as shown in Figure 19.16 allowing the sides of the coils to be cooled. Alternatively there may be gaps between the coils to allow access the tube body. The effect of this arrangement of the winding is to cause a periodic variation in the axial magnetic field. The magnitude of the ripple can be estimated by assuming that each coil can be represented by a single current loop located at the centre of the winding and that the effect of the pole pieces can be represented by the mirror images of the coils in them. The final design of the solenoid is checked by numerical modelling using one of the many computer programs which are commercially available. The design of solenoids is discussed in [3, 33, 34].
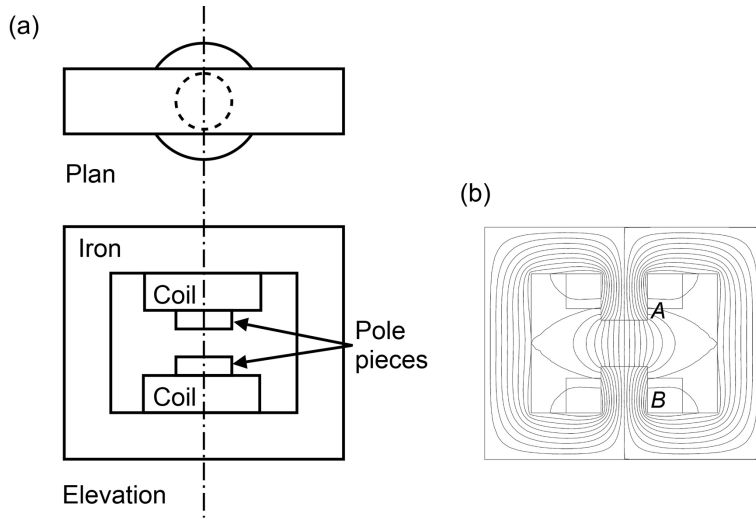
Figure 19.17: Electromagnet for a small microwave tube: (a) general arrangement, and (b) flux distribution.

19.6   Iron Dominated Magnets

A few tubes, such as small klystrons, have dipole electromagnets arranged as shown in Figure 19.17(a). It can be seen from Figure 19.17(b) that the flux density in the gap between the pole-pieces is not uniform. Also, there are two regions where the flux concentration may be great enough to saturate the iron. The flux concentration at *A* can be reduced by rounding the corners of the pole-pieces. The concentration at *B* is caused by the increase in flux towards the roots of the pole-pieces caused by the fringing field. This effect can be reduced by tapering the pole-pieces so that the cross-sectional area increases as the flux increases. It has been shown that the saturation of the pole-piece tips is almost uniform if they have Rogowksi profiles [29]. This profile was originally devised to produce a uniform electric field between electrodes for studies of voltage breakdown. The profile is such that the electric field on the surface of an electrode is greatest on the axis and decreases monotonically with increasing distance from it [35]. In practice this profile can be approximated quite well by a circular arc and a taper, as shown in Figure 19.18(a) [29]. If saturation of the pole-piece does not occur, then the region where the field is uniform can be increased by shaping the surface of the pole-piece, as shown in Figure 19.18(b). If the reluctance of the iron is neglected then the number of ampere-turns required to produce a given flux density in the gap can be calculated directly. The flux density in the iron can be obtained from the flux density in the gap multiplied by a factor to allow for the fringing field. Then the additional ampere-turns needed to overcome the reluctance of the iron can be calculated if necessary. Further information on iron dominated magnets can be found in [29, 36].
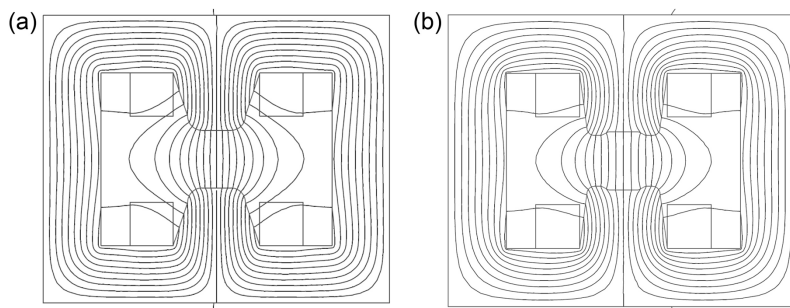
Figure 19.18: Pole-piece design: (a) for uniform field in the pole-piece, and (b) for uniform field in the air gap.

19.7   Permanent Magnet Design

A permanent magnet system comprises one, or more, permanent magnets and may include a soft iron yoke and pole-pieces. The pole-pieces may concentrate the flux so that the flux density in the air gap is greater than the maximum which can be achieved in the magnets. For initial design purposes it is possible to neglect the reluctance of the soft iron parts of the circuit. The permanent magnets are then designed to provide the magneto-motive force required to produce the desired flux in the air space, as described in Section 19.3.1. If the air space is between soft iron pole-pieces, which are not saturated, then its reluctance can be calculated using a numerical solution of Laplace's equation. It is usually important to minimise some combination of the size, cost, and weight of the magnet.

19.7.1   Permanent Magnets for Magnetrons and CFAs

Early conventional magnetrons and CFAs employed Alnico magnets. The relatively low energy products and low remanence of these materials required the use of large and heavy magnets, as shown in Figure 19.19(a) [2]. The higher energy products of ferrite, and rare earth, materials allow the use of smaller magnets, as shown in Figure 19.19(b). In some cases there may be iron pole-pieces between the permanent magnets and the gap. Alternative versions of both types of magnet exist with rotational symmetry about the vertical axis, and the geometry may be annular instead of cylindrical. The general principles of design are that the permanent magnet material should be located as close to the air gap as possible, and that it should be operated close to its maximum energy product.

To illustrate the design procedure for a magnet of this kind let us suppose that we wish to generate a uniform flux density B_g between cylindrical magnets of radius R separated by a gap g. The flux density in the magnets must be greater than that in the air gap because of the fringing field. Figure 19.20 shows one quadrant of the
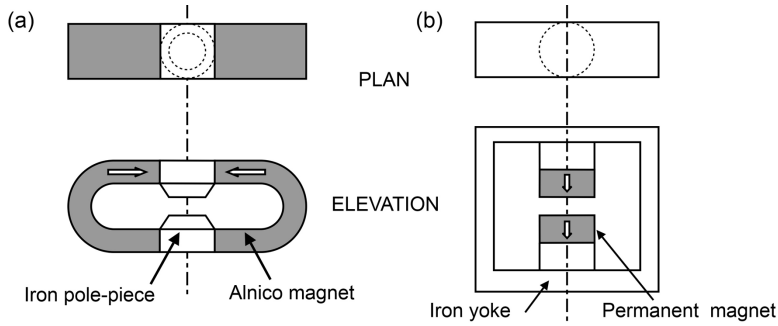
Figure 19.19: Arrangement of permanent magnets for magnetrons and CFAs using (a) Alnico, and (b) ferrite and rare earth magnets.
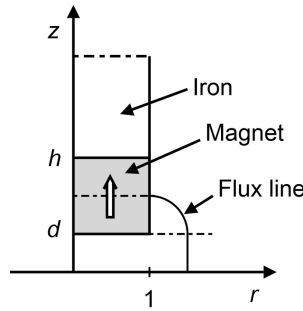
Figure 19.20: Geometry of a dipole magnet showing the form of an approximate flux line.

system with dimensions normalised to the radius of the magnets. The fringing is represented by approximate flux lines each comprising a circular arc and a straight line as shown [9]. We will assume that the magneto-motive force for a flux line is given by

$$\mathsf{M} = \left(\frac{h-z}{h-d}\right)\mathsf{M}_0 \tag{19.37}$$

so that it is constant on the flat surface of the magnet $(z = d)$ and varies linearly along the curved surface $(r = 1)$. In the absence of fringing the flux generated by the magnet is

$$\Phi = \frac{\mu_0 \pi}{d}\mathsf{M}_0. \tag{19.38}$$

The contribution to the fringing flux from an element of the surface dz is

$$d\Phi = 8\mu_0\mathsf{M}_0\left(\frac{h-z}{h-d}\right)\frac{1}{z+c_1 g}\,dz, \tag{19.39}$$

where $c_1 = (4/\pi) - 1$ and h is an initial estimate. Then the additional flux is

$$\Delta\Phi = \frac{8\mu_0 M_0}{h-d} \int_d^h \frac{h-z}{z+c_1 g} dz, \tag{19.40}$$

which can be integrated to give

$$\Delta\Phi = 8\mu_0 M_0 \left\{ \left(\frac{h+c_1 d}{h-d} \right) \ln \left(\frac{h+c_1 d}{d+c_1 d} \right) - 1 \right\}. \tag{19.41}$$

Dividing by Φ gives

$$\frac{\Delta\Phi}{\Phi} = \frac{8d}{\pi} \left\{ \left(\frac{h+c_1 d}{h-d} \right) \ln \left(\frac{h+c_1 d}{d+c_1 d} \right) - 1 \right\}. \tag{19.42}$$

Since the flux is continuous the flux density in the magnet must be

$$B_m = \left(1 + \frac{\Delta\Phi}{\Phi} \right) B_g. \tag{19.43}$$

We suppose that the operating point lies on a linear demagnetisation curve (or recoil loop) defined by the intersections on the axes $(0, B_r)$ and $(-H_c, 0)$. Then

$$H_m = \left(1 - \frac{B_m}{B_r} \right) H_c \tag{19.44}$$

and, from the magnetic circuit law, we find that

$$h = \left(1 + \frac{B_g}{\mu_0 H_m} \right) d. \tag{19.45}$$

The final value of h is found after a few iterations. Comparison between the results of this calculation and those from finite element modelling shows good qualitative agreement. Figure 19.21 shows, as an example, the flux plot and the variation of B_z with radius at the mid-plane of the gap for a typical case, using a samarium cobalt magnet, for a target flux density in the gap of 0.4 T (pole radius 15 mm; gap 10 mm; magnet thickness 4.2 mm). These calculations are implemented in Worksheet 19.2. It should be noted that the theory given above assumes that the leakage flux is negligible. It is, in fact, rather difficult to find approximate formulae for the leakage flux in practical geometries and it is best to design the circuit so that the return flux path is kept well away from the permanent magnets. The method of estimating the fringing flux can be adapted for annular magnets, and for those fitted with cylindrical or conical pole-pieces.

19.7.2   Permanent Magnets for Linear Beam Tubes

Magnet structures similar to those in Figure 19.19(b) have been used for small klystrons and TWTs but they are not well-adapted to producing a long region of uniform field. An alternative arrangement is shown in Figure 19.22. This is suitable for use with Alnico magnets where a long magnet is required to generate sufficient
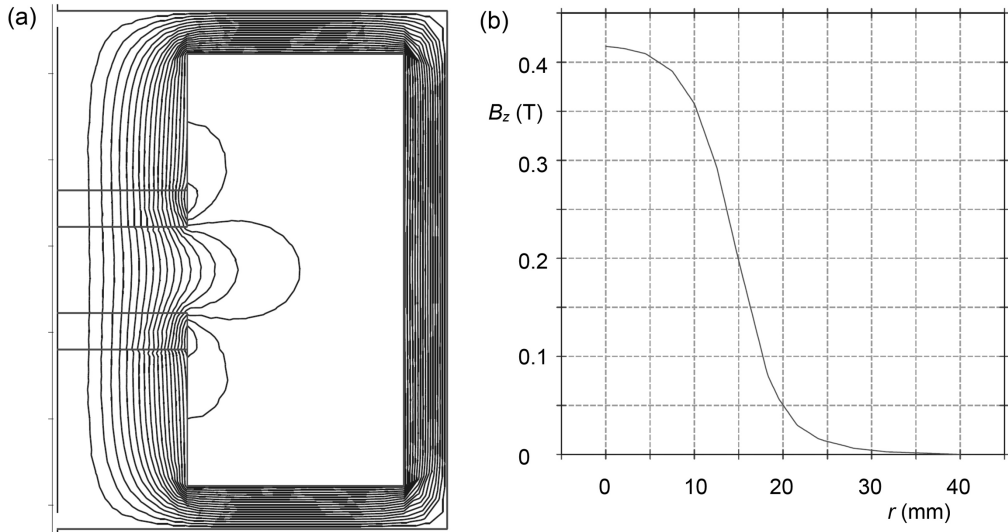
Figure 19.21: Results of finite element calculations for a dipole magnet (a) flux plot, and (b) variation of B_z with r when $z = 0$.
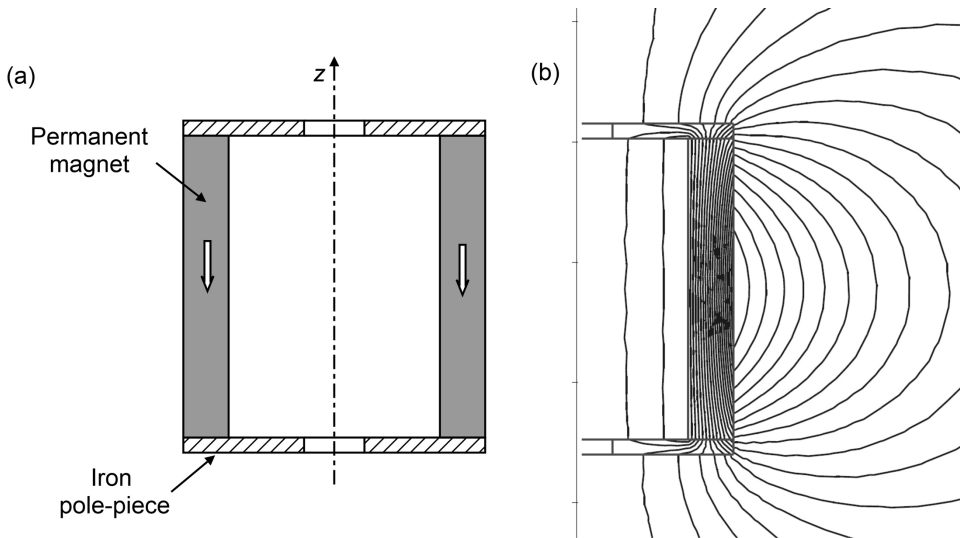
Figure 19.22: Alnico magnet for a linear-beam tube: (a) general arrangement, and (b) flux plot.

magneto-motive force. The length of the magnet relative to the air space can be increased, if necessary, by making it barrel-shaped and by making the pole-pieces re-entrant. If rare earth magnets are used then it may be necessary to create magnet stacks with alternate layers of permanent magnet and soft iron to obtain the correct combination of flux and magneto-motive force. The flux density on the axis is
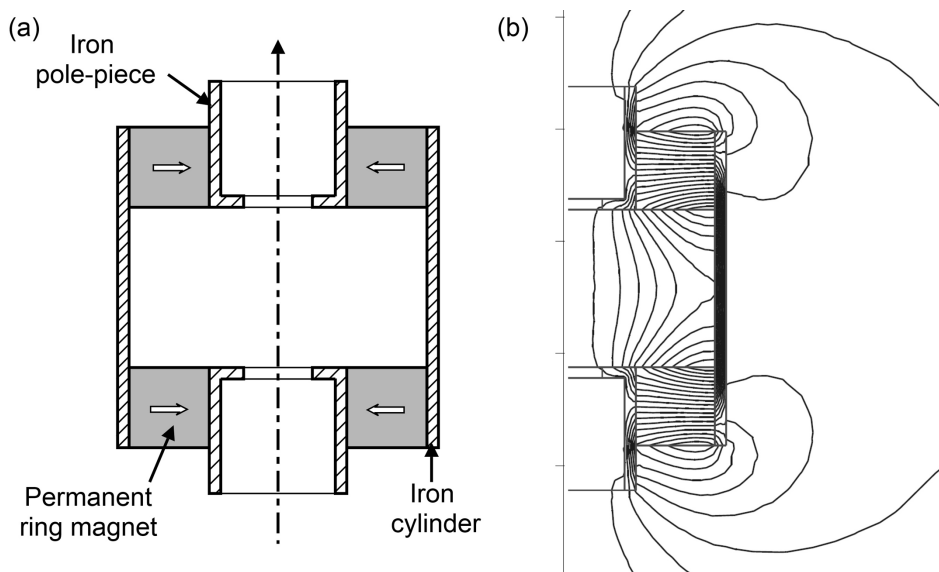
Figure 19.23: Rare earth magnet for a linear-beam tube: (a) general arrangement, and (b) flux plot.

approximately uniform within the magnet but the sign is reversed outside the magnet. There is also a substantial leakage flux outside the magnet [1, 3]. An example of a klystron focused in this way is given in [37].

Figure 19.23 shows an alternative arrangement using radially magnetised rings, which is suitable for use with rare earth magnets [3, 38]. This suffers from the same problems of flux uniformity and leakage as the previous example. The reversal of the direction of the axial flux at the collector end of the magnet can be reduced by using a single ring magnet at the gun end. The non-uniformity of the axial flux, in this and other magnets, can be reduced by including a number of transverse soft iron plates as flux straighteners.

Further details of magnet design can be found in [1–3]. The problems of flux leakage and non-uniform flux density can be overcome by cladding the magnet system with additional magnets having appropriate directions of magnetisation [39–44].

19.7.3   Periodic Permanent Magnet (PPM) Systems

The theory of electron beam focusing by a periodic magnetic field is discussed in Section 7.6. In this section we consider the problem designing the periodic permanent magnet stack to produce the field required. This subject was considered by a number of authors who developed graphical methods [45–49]. The availability of powerful personal computers means that the formulae used in the graphs can be
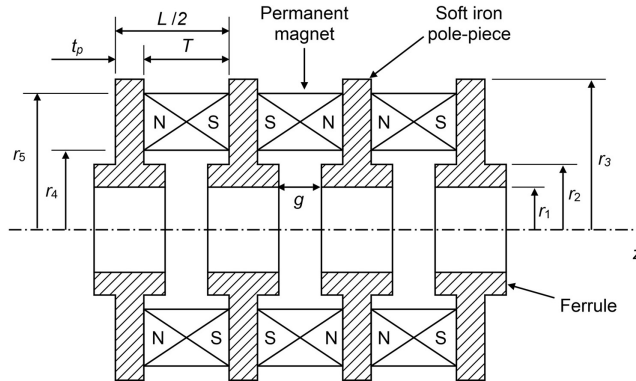
Figure 19.24: Arrangement of a periodic permanent magnet stack (copyright 2009, IEEE, reproduced, with permission, from [50]).

computed readily, so that graphical methods are no longer required. The discussion here follows that of Santra et al. which has the widest applicability and has been shown to give good results [50].

Figure 19.24 shows the arrangement and dimensions of a PPM structure. This allows a number of common variations to be considered:

- The relationship between the outer radius of the pole-pieces (r_3) and the outer radius of the magnets (r_5). If $r_3 < r_5$ then the external leakage flux is reduced, and the peak flux density on the axis is increased, for a given magnet. However, it is then necessary to introduce some non-magnetic material to increase the overall radius of the pole-piece if it is to be cooled by conduction. This problem is avoided if $r_3 > r_5$, although the peak magnetic flux density on the axis is reduced, and the leakage flux is increased. This arrangement is also better adapted to the addition of small pieces of soft iron (shunts) to the outside of the stack to adjust its properties when hot-testing. The third possibility $(r_3 = r_5)$ is said to keep the variations in the field at the ends of the stack to a minimum [48].
- The relationship between the inner radius of the magnets (r_4) and the outer radius of the ferrule (r_2). In helix TWTs it is possible to make the inner radius of the magnets only slightly greater than the outer radius of the ferrule. This corresponds to the good design practice of placing the permanent magnet as close to the gap as possible. In coupled-cavity TWTs the ferrule forms part of the RF structure (see Figure 4.29) while the magnets must be placed outside the structure.
- Some PPM structures have ferrules, whereas others do not. The choice is determined partly by the ratio of the gap length to the pitch of the structure which gives the greatest fundamental component of the flux density on the axis. However, it may be constrained by compatibility with the dimensions of the RF structure.

The model used here allows for all these possible variations.
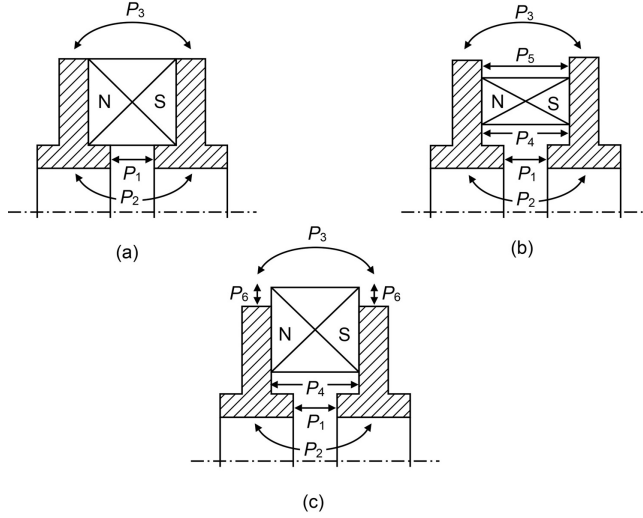
Figure 19.25: Flux paths in a periodic permanent magnet stack (copyright 2009, IEEE. reproduced, with permission, from [50]).

The possible flux paths in air around the structure are illustrated in Figure 19.25. Because these paths are mostly in parallel with one another it is convenient to work in terms of permeance (P) which is the reciprocal of the reluctance defined in (19.20). The calculation of the permeances is based on treating the surface of the pole-piece as a magnetic scalar equipotential so that the same magneto-motive force acts along each path. The case shown in Figure 19.25(a) was discussed in [46] and only the results are given here. It is assumed that the flux density is uniform in the gap between the ferrules (g). We have seen, in discussions of the RF electric field in a gap, that this is quite a good approximation if the ends of the ferrules are blunt (see Section 3.5.3). Then, in SI units,

$$P_1 = \frac{\mu_0 \pi \left(r_2^2 - r_1^2\right)}{g},$$ (19.46)

$$P_2 = \sum_{n=1,3,5\ldots}^{\infty} \frac{4\mu_0 L r_1}{\pi n^2 g} \cdot \frac{I_1\left(2n\pi r_1/L\right)}{I_0\left(2n\pi r_1/L\right)} \cdot \sin\left(\frac{n\pi g}{L}\right),$$ (19.47)

and

$$P_3 = \sum_{n=1,3,5\ldots}^{\infty} \frac{4\mu_0 L r_3}{\pi n^2 T} \cdot \frac{K_1\left(2n\pi r_3/L\right)}{K_0\left(2n\pi r_3/L\right)} \cdot \sin\left(\frac{n\pi T}{L}\right),$$ (19.48)

where I_0, I_1, K_0 and K_1 are modified Bessel functions and the dimensions of the structure are defined in Figure 19.24. The additional permeances in Figure 19.25(b) are

$$P_4 = \frac{\mu_0 \pi \left(r_4^2 - r_2^2\right)}{T}$$ (19.49)

and

$$P_5 = \frac{\mu_0 \pi \left(r_3^2 - r_5^2\right)}{T}.$$ (19.50)

Finally, the indented pole pieces in Figure 19.25(c) add permeances given by

$$P_6 = \frac{2\mu_0 \pi t_p}{\ln\left(r_3/r_5\right)}$$ (19.51)

in series with P_3 so that the modified permeance of this flux path is

$$P_{3m} = \frac{P_3 P_6}{2P_3 + P_6}.$$ (19.52)

Note that this equation is given incorrectly in [50]. The total permeance P_t is found by summing the appropriate permeances in each case. Then the slope of the load line is

$$K = \frac{T}{\pi\left(r_5^2 - r_4^2\right)} P_t.$$ (19.53)

If rare-earth magnets are used then the demagnetisation curve is linear and the operating point $\left(H_d, B_d\right)$ is given by

$$H_d = \frac{B_r}{K + \dfrac{B_r}{H_c}},$$ (19.54)

where B_r and H_c are the remanence and coercive force of the material. Ferrite and Alnico magnets must be stabilised so that the operating point is on a minor loop [48]. We have seen that minor loops are effectively linear so that (19.54) can be used in this case also if B_r and H_c are taken to be the intersections of the projections of the minor loop on the two axes. Finally, the axial component of the magnetic flux density in the beam hole of an infinite PPM stack is given by

$$B_z\left(r, z\right) = \sum_{n=1,3,5\ldots}^{\infty} \frac{4B_g \sin\left(n\pi g/L\right)}{n\pi} \cdot \frac{I_0\left(2n\pi r/L\right)}{I_0\left(2n\pi r_1/L\right)} \cdot \cos\left(\frac{2n\pi z}{L}\right),$$ (19.55)

where $B_g = \mu_0 H_d T/g$. Equations (19.46) to (19.55) are the working equations in Worksheet 19.3.

The design of a PPM stack starts from the period of the magnetic field (L) and the peak flux density on the axis $\left(B_z(0,0)\right)$. The inner radius of the ferrule (r_1) is normally fixed by the outer radius of the shield of a helix slow-wave structure, or by the beam hole radius of a coupled-cavity slow-wave structure. The flux density in the gap $\left(B_g\right)$ can be determined from (19.55) as a function of the gap length (g) which can be chosen for the best results if it is not constrained by the dimensions of the RF structure. The outer radius of the ferrule (r_2) needs to be great enough to avoid saturation of the ferrule. In practice this means that the flux density must be less than 1T (see Figure 19.4). The choice must also be compatible with the design

of the RF structure. The inner radius of the magnet (r_4) will normally be taken to be as small as possible. There then remain only three dimensions to be chosen namely the outer radius of the magnet (r_5), the outer radius of the pole-piece (r_3), and the thickness of the pole-pieces (t_p). If the two radii are set to the same value then the number of free parameters is reduced to two. The thickness of the pole-piece must be great enough to avoid saturation. Subject to these restrictions the free parameters can be chosen to minimise the volume of the magnets. Final optimisation of the design requires the use of computational electromagnetics software. That analysis can also investigate the fields at the ends of the stack, and the effects of any gaps which are required to allow for the passage of the RF input and output connections [51]. Methods similar to those described above can be used to analyse and design long-period PPM structures (see Section 7.6.1) and those where additional permanent magnets have been added to reduce the leakage flux [44, 52–54].

References

[1] J. F. Gittins, *Power Travelling-Wave Tubes*. English Universities Press, 1965.

[2] W. Harrold and W. Reid, 'Permanent magnets for microwave devices', *IEEE Transactions on Magnetics,* vol. 4, pp. 229–239, 1968.

[3] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, UK: Research Studies Press, 1995.

[4] R. G. Carter, *Electromagnetism for Electronic Engineers*. Available at: http://bookboon.com/en/electromagnetism-for-electronic-engineers-ebook (accessed 5 October 2017), 2010.

[5] G. W. Carter, *The Electromagnetic Field in its Engineering Aspects*, 2nd ed. London: Longman, 1967.

[6] B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*. London: Oxford University Press, 1957.

[7] M. MacCaig, *Permanent Magnets in Theory and Practice*. London: Pentech Press, 1987.

[8] A. Edwards, 'Magnet design and selection of material', in D. Hadfield, ed., *Permanent Magnets and Magnetism*. London: Iliffe, pp. 191–296, 1962.

[9] W. L. Upson and E. L. Furth, 'Abridgment of magnetic leakage and fringing flux calculations', *Journal of the A.I.E.E.,* vol. 47, pp. 340–344, 1928.

[10] J. Sykulski, *Computational Magnetics*. Springer Science & Business Media, 2012.

[11] Infolytica Corp., 'Gemini', Montreal, Canada, 1966.

[12] P. Campbell, *Permanent Magnet Materials and Their Application*. Cambridge University Press, 1996.

[13] H. R. Kirchmayr, 'Permanent magnets and hard magnetic materials', *Journal of Physics D: Applied Physics,* vol. 29, p. 2763, 1996.

[14] R. G. Carter, *Electromagnetism for Electronic Engineers*, 2nd ed. London: Chapman & Hall, 1992.

[15] G. W. C. Kaye and T. H. Laby, *Tables of Physical and Chemical Constants.* Available at: www.kayelaby.npl.co.uk/ (accessed 5 October 2017), 2014.

[16] Magdev Ltd, *Permanent Magnets*. Available at: www.magdev.co.uk/permanent-magnets (accessed 5 October 2017), 8 August 2017.

[17] Delta Magnets, *Permanent Magnets*. Available at: www.magdev.co.uk/permanent-magnets (accessed 5 October 2017), 2011.

[18] A. J. Bamford *et al.*, 'A 1-MW L-band coaxial magnetron with separate cavity', *IEEE Transactions on Electron Devices,* vol. 14, pp. 844–851, 1967.

[19] D. Bhasavanich and H. G. Hammon, III, 'Pulsed magnet system considerations for repetitive operation of high-power magnetrons', *IEEE Transactions on Electron Devices,* vol. 38, pp. 791–795, 1991.

[20] Y. Dai *et al.*, 'Design study on A 9.2 T NbTi superconducting magnet with long-length uniform axial field', *IEEE Transactions on Applied Superconductivity,* vol. 25 (3), June, paper no. 4101204, 2015.

[21] R. Hirose *et al.*, 'Development of 7 T cryogen-free superconducting magnet for gyrotron', *IEEE Transactions on Applied Superconductivity,* vol. 18, pp. 920–923, 2008.

[22] W. R. Smythe, *Static and Dynamic Electricity*, 3rd ed. New York: McGraw-Hill, 1950.

[23] L. McKeehan, 'Combinations of circular currents for producing uniform magnetic fields', *Review of Scientific Instruments,* vol. 7, pp. 150–153, 1936.

[24] M. W. Garrett, 'Axially symmetric systems for generating and measuring magnetic fields. Part I', *Journal of Applied Physics,* vol. 22, pp. 1091–1107, 1951.

[25] R. G. Carter, 'Coil-system design for production of uniform magnetic-fields', *Proceedings of the Institution of Electrical Engineers-London,* vol. 123, pp. 1279–1283, 1976.

[26] W. Franzen, 'Generation of uniform magnetic fields by means of air-core coils', *Review of Scientific Instruments,* vol. 33, pp. 933–938, 1962.

[27] W. Gosling and M. J. Cunningham, 'Gapped solenoid as a means of producing a highly uniform magnetic field over an extended volume', *Proceedings of the Institution of Electrical Engineers,* vol. 121, pp. 1589–1593, 1974.

[28] D. B. Montgomery, *Solenoid Magnet Design*. New York: Wiley-Interscience, 1969.

[29] S. Russenschuck, *Field Computation for Accelerator Magnets: Analytical and Numerical Methods for Electromagnetic Design and Optimization*. John Wiley & Sons, 2011.

[30] A. Staprans *et al.*, 'High-power linear-beam tubes', *Proceedings of the IEEE,* vol. 61, pp. 299–330, 1973.

[31] W. G. Worcester *et al.*, 'Light-weight aluminum foil solenoids for traveling-wave tubes', *IRE Transactions on Electron Devices,* vol. 3, pp. 70–74, 1956.

[32] R. C. Glass and P. Gottfeldt, 'Aluminum foil solenoids for traveling-wave tubes', *IRE Transactions on Electron Devices,* vol. 4, p. 186, 1957.

[33] E. A. Perigo *et al.*, 'Design of a multi-sections solenoid for power microwave tubes', in *IEEE MTT-S International Conference on Microwave and Optoelectronics*, pp. 233–236, 2005.

[34] D. Borodin and M. Einat, 'Copper solenoid design for the continuous operation of a second harmonic 95-GHz gyrotron', *IEEE Transactions on Electron Devices,* vol. 61, pp. 3309–3316, 2014.

[35] J. Harrison, 'A computer study of uniform-field electrodes', *British Journal of Applied Physics,* vol. 18, p. 1617, 1967.

[36] J. T. Tanabe, *Iron Dominated Electromagnets: Design, Fabrication, Assembly and Measurements*. Singapore: World Scientific, 2005.

[37] J. Lebacqz, 'High power klystrons', *IEEE Transactions on Nuclear Science,* vol. 12, pp. 86–95, 1965.

[38] C. Fu *et al.*, 'A type of tube-shaped permanent magnet focusing system of high-power klystron', in *Asia-Pacific Microwave Conference Proceedings*, vol. 1, pp. 333–335, 1997.

[39] J. P. Clarke and H. A. Leupold, 'Shaping of cylindrically symmetric magnetic fields with permanent magnets', *IEEE Transactions on Magnetics,* vol. 22, pp. 1063–1065, 1986.

[40] Y. Fuwa *et al.*, 'Focusing system with permanent magnets for klystrons', *IEEE Transactions on Applied Superconductivity,* vol. 24, pp. 1–5, 2014.

[41] D. K. Abe and J. P. Calame, 'Advanced materials technologies', in R. J. Barker *et al.*, eds, *Modern Microwave and Millimeter-Wave Power Electronics*. Piscataway, NJ: IEEE Press, pp. 649–689, 2005.

[42] H. Leupold *et al.*, 'Permanent magnet sources for extended interaction klystrons', *Journal of Applied Physics,* vol. 70, pp. 6624–6626, 1991.

[43] H. Leupold and E. Potenziani II, 'Augmentation of field uniformity and strength in spherical and cylindrical magnetic field sources', *Journal of Applied Physics,* vol. 70, pp. 6621–6623, 1991.

[44] Q. L. Peng *et al.*, 'Axial magnetic field produced by axially and radially magnetized permanent rings', *Journal of Magnetism and Magnetic Materials,* vol. 268, pp. 165–169, 2004.

[45] F. Sterzer and W. Siekanowicz, 'The design of periodic permanent magnets for focusing of electron beams', *RCA Review,* vol. 18, pp. 39–59, 1957.

[46] J. E. Sterrett and H. Heffner, 'The design of periodic magnetic focusing structures', *IRE Transactions on Electron Devices,* vol. 5, pp. 35–42, 1958.

[47] M. Schindler, 'The magnetic field and flux distributions in a periodic focusing stack for traveling-wave tubes', *RCA Review,* vol. 21, pp. 414–436, 1960.

[48] M. Schindler, 'An improved procedure for the design of periodic-permanent-magnet assemblies for traveling-wave tubes', *IEEE Transactions on Electron Devices,* pp. 942–949, 1966.

[49] M. Corsi *et al.*, 'Graphical method for optimum design of periodic magnetic focusing structures', *IEEE Transactions on Electron Devices,* vol. 21, pp. 2–11, 1974.

[50] M. Santra *et al.*, 'An improved analysis of PPM focusing structures including the effect of magnetic saturation in the iron pole pieces', *IEEE Transactions on Electron Devices,* vol. 56, pp. 974–980, May 2009.

[51] C. L. Kory, 'Effect of geometric azimuthal asymmetries of PPM stack on electron beam characteristics [TWTs]', *IEEE Transactions on Electron Devices,* vol. 48, pp. 38–44, 2001.

[52] H. A. Leupold and J. P. Clarke, 'Bulk reduction and field enhancement in periodic permanent-magnet structures', *IEEE Transactions on Electron Devices,* vol. 34, pp. 1868–1872, August 1987.

[53] H. A. Leupold *et al.*, 'Lightweight permanent magnet stack for traveling-wave tubes', *Journal of Applied Physics,* vol. 67, pp. 4656–4658, 1990.

[54] M. Santra *et al.*, 'Analysis of long-periodic permanent magnet structures for electron beam focusing', *IEEE Transactions on Electron Devices,* vol. 60, pp. 1776–1781, 2013.

20 System Integration

20.1 Introduction

The previous chapters of this book have provided a detailed discussion of the principles of vacuum tube amplifiers and oscillators. The purpose of this chapter is to review the factors affecting the satisfactory integration of a tube into the system of which it is a part. The principal elements of a high power RF system employing a vacuum tube amplifier are shown in Figure 20.1. The arrangement is similar if the tube is an oscillator, except that there is no low-level RF system.

The key to satisfactory operation of any system is good integration of all the parts. Usually the different elements are obtained from a number of sources and successful system integration requires good communication between all the engineers concerned. This is particularly important where vacuum tubes are concerned because the knowledge bases of the tube and the system engineers are not usually the same (see Section 1.5) [1]. Some of the interface problems can be avoided if a single organisation takes responsibility for the design and manufacture of a packaged amplifier incorporating the tube and its subsystems. A good example is provided by the microwave power module in which a solid-state driver amplifier, a miniature TWT and its power supply are combined into a single unit [2].

The sections that follow consider how each of the blocks in Figure 20.1 contributes to the successful operation of the system. A detailed analysis of the nature and causes of reliability problems in tube-based systems can be found in [3]. Although great advances have since been made in tube manufacturers' quality assurance it is to be suspected that many of the other problems identified still remain. An overview of the factors affecting the reliability of systems incorporating vacuum tubes is in given in Section 20.9. More detailed information can be found in [4–6] and in other sources cited in this chapter.

20.2 DC Power Supplies

The high-voltage supply is the source of DC power to the tube, and any instability or noise in it is likely to lead to degradation of the tube performance. There may be combinations of voltage and current which must be avoided because the tube does not work correctly (see Sections 14.3.8 and 15.4.1) [7]. Some tubes have regions with negative resistance characteristics, and the impedance of the power supply
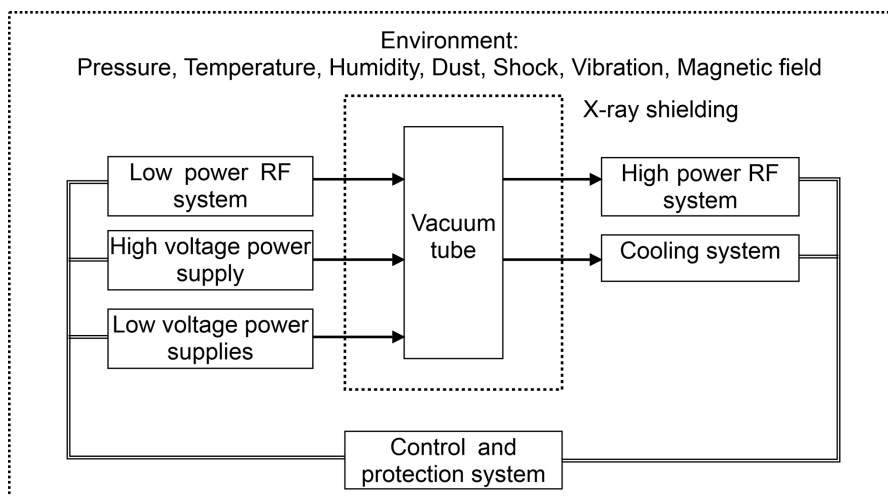
Figure 20.1: Block diagram of a typical vacuum tube amplifier.

must then be chosen correctly if oscillations are not to occur. When the tube is pulsed that may be achieved by applying a pulsed cathode voltage (see Section 20.3) or by using a DC high-voltage supply and applying a pulsed voltage to a control electrode (see Section 9.5). The design of the high-voltage supply is more complicated if the tube has a multi-element depressed collector. The collector electrodes are connected to taps on the high-voltage supply, and these intermediate voltages must be kept steady when the currents drawn change with the operating conditions of the tube. The possible voltages at the taps of the high-voltage supply may be determined by its design, and that places constraints on the design of the collector.

In addition to the main high-voltage supply there may be DC supplies at lower voltages for the cathode heater, electromagnets, control electrodes, and ion pumps. If these supplies are not accurately calibrated, and well-regulated, then malfunction, and possibly damage, of the tube may result. The heater power is important in ensuring the correct emission from the cathode. Noise in the heater power supply can cause additional noise in the output from the tube. Incorrect magnet currents and control voltages may lead to imperfect control of the electrons leading to additional power dissipation within the tube, and possible damage to internal structures. All the power supplies draw their power from an external AC or DC source which may not be well-regulated, and may exhibit voltage transients.

20.2.1   High-Voltage Switches

All power supplies rely on electronic high-voltage switches for their operation. These devices can be grouped into three classes:

- switched on and off by the applied voltage;
- switched on by a control signal, and off by the applied voltage;
- switched on and off by a control signal.

Table 20.1: Classification of electronic high-voltage switches

Switch on	Switch off	Semiconductor	Vacuum	Gas discharge
Applied voltage	Applied voltage	Diodes	Diodes	
Control signal	Applied voltage	Thyristors		Thyratrons
				Spark gaps
				Ignitrons
Control signal	Control signal	MOSFETS	Triodes	
		IGBTs	Tetrodes	
		GTOs	Hobetron [15]	
		IGCTs [14]		

Semiconductor devices switching up to 8 kV are now used in all DC supplies, but switches using vacuum or gas discharge technology are to be found in older systems [8]. Thyratrons (triggered gas discharge tubes) which can work at voltages up to 200 kV and peak currents of 100 kA are still used in many pulsed supplies [9, 10], but semiconductor-based alternatives are being developed [11–13]. Table 20.1 shows a classification of electronic high-voltage switches.

20.2.2 Load Impedance

The impedance presented to the high voltage power supply by the tube depends upon the type of tube and the operating conditions. The DC current drawn by a gridded tube operated in class B or C, or by an IOT, depends upon the RF input power (see Figure 12.21(b)). The examples in Chapter 12 show that typical values of the DC load impedance are 1 kΩ for a tetrode, and greater than 25 kΩ for an IOT. The load impedance presented by a linear-beam tube with a space-charge limited gun depends upon the perveance (K) and the anode voltage (V_a) as

$$R_L = \frac{1}{K\sqrt{V_a}}.$$ (20.1)

Figure 20.2 shows the impedance for typical values of voltage and perveance. If the current is reduced from the space-charge limited value by the voltage on a control grid, or by temperature limited operation (as in a gyrotron), then the impedance is increased. The operating points of both gridded tubes and linear-beam tubes are determined by the anode voltage so the power supply should approximate to an ideal voltage source. In models of the tube with its power supply, a linear-beam tube may be represented by a fixed resistor in series with a perfect diode.

Crossed-field tubes draw no current until the threshold voltage (V_T) is exceeded. Thereafter the voltage varies slowly for a large variation in current (see Figures 15.27 and 16.8) so that the incremental impedance is of the order of 100 Ω. The operating point is determined by the current so the power supply should approximate to an
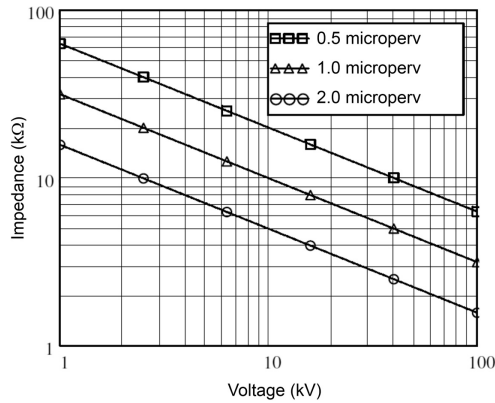
Figure 20.2: D.C. input resistance of tubes with space-charge limited cathodes.

ideal current source. A crossed-field tube can be modelled by a diode and a fixed resistance in series with a DC voltage source [16, 17]. The current drawn is then given by

$$I_0 = \frac{V_a - V_T}{R_L},$$ (20.2)

where V_a is the anode voltage, V_T is the threshold voltage and R_L is the incremental resistance.

For all tubes the regulation of the power supply must be able to handle the variation in the impedance during start-up as well as in normal operation. Figure 20.3 compares the static characteristics of typical linear-beam, and crossed-field, tubes including a linear approximation for a linear-beam tube. In practice this idealised behaviour is modified by the inter-electrode capacitance of the tube and by the parasitic inductance of the cables connecting it to the power supply.

20.2.3   Electric Power Converters

The stable high voltage supply is provided by an electric power converter [14, 18]. Converters typically have the bridge configuration shown in Figure 20.4 where the switches represent semiconductor power devices.

According to the type of switches, and the way in which they are controlled, the basic converter can be used to convert power from DC or AC sources to supply DC, or variable frequency AC, power to a load. The output voltages of circuits using devices other than diodes can be controlled by the timing of the switching. The simplest example of an AC to DC converter is the single-phase bridge rectifier in which the switches are diodes as shown in Figure 20.5(a). The output waveform is shown in Figure 20.5(b). Referring to Figure 20.4 we see that when the input voltage is positive switches A and D are closed, and when the input voltage is negative switches B and C are closed. Thus the voltage applied to the load always has the
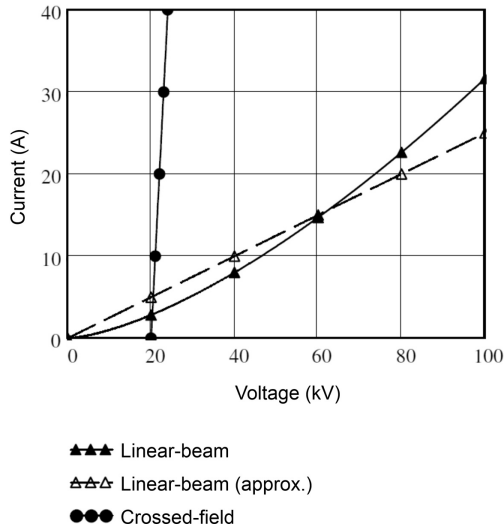
Figure 20.3: Comparison between the static characteristics of typical tubes: linear-beam $\left(K = 1\,\mu\text{perv},\ R_L \approx 4\,\text{k}\Omega\right)$ and crossed-field $\left(V_T = 20\,\text{kV},\ R_L = 100\,\Omega\right)$.
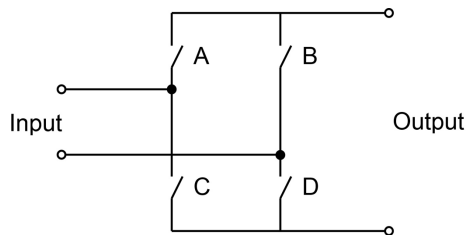
Figure 20.4: Generic single phase bridge converter.
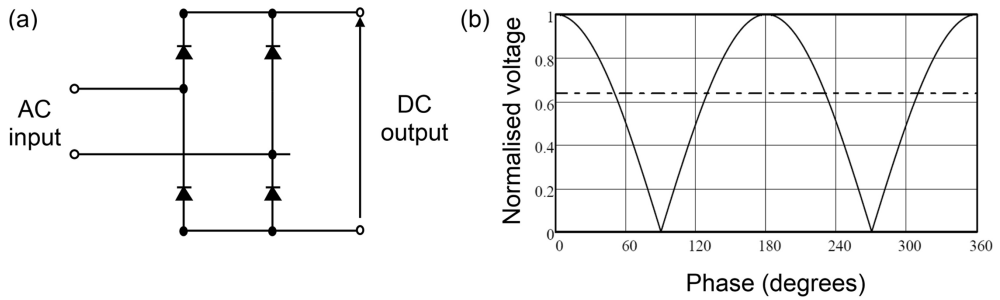
Figure 20.5: Single-phase diode bridge rectifier: (a) circuit diagram, (b) unsmoothed output waveform showing the mean voltage.
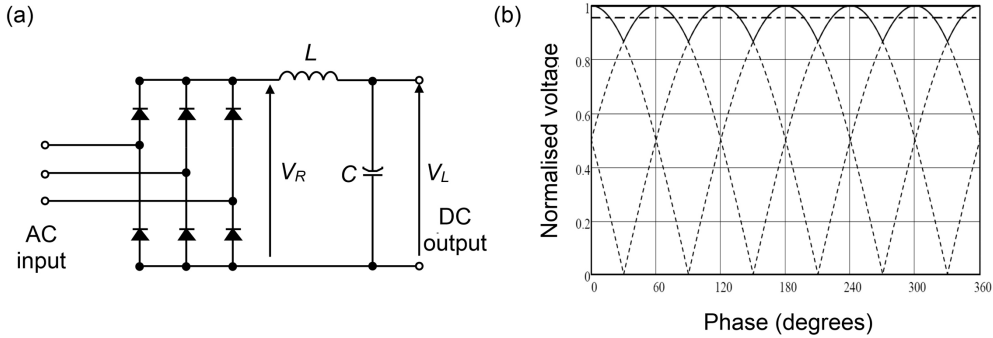
Figure 20.6: Three-phase bridge rectifier: (a) circuit diagram and (b) output waveform showing the mean voltage.

same polarity. The working voltages of individual switching devices are typically 1–5 kV depending upon the type. Higher voltages are obtained by stacking converter units in series, or by using voltage multiplier circuits. The properties of single-phase bridge converters of different kinds can be explored using Worksheet 20.1.

The output voltage provided by a power converter is modulated by the switching frequency and its harmonics. The ripple on the output of the rectifier can be reduced by employing a three-phase bridge circuit, as shown in Figure 20.6(a). This reduces the amplitude of the ripple, and increases the ripple frequency, as shown in Figure 20.6(b). In order to obtain a nearly constant voltage the output must be passed through a low pass filter, comprising a series inductor and shunt capacitor, as shown in Figure 20.6(a). Provided that the reactance of the capacitor $(1/\omega C)$ is small compared with the load resistance, the ripple in the voltage at the load is related to the ripple in the voltage at the output of the rectifier by

$$\left|V_{1,L}\right| = \frac{1}{\left|1 - \omega^2/\omega_0^2\right|} \left|V_{1,R}\right|,\tag{20.3}$$

where $\omega_0 = 2\pi f_0 = 1/\sqrt{LC}$ and $V_{1,R}$ and $V_{1,L}$ are the amplitudes of the first harmonic at the output of the rectifier, and the output of the filter, respectively.

A typical three-phase power supply of this kind, capable of delivering 16 A at 52 kV with 0.7% ripple at 300 Hz, is described in [19]. The smoothing circuit comprises a 2 H inductor and a 24 μF capacitor so that $f_0 = 23$ Hz. Substitution of these numbers into (20.3) shows that the filter reduces the amplitude of the ripple at the output of the rectifier by a factor of 0.006. The energy stored in the capacitor is 32 kJ which is sufficient to cause damage to the tube if an arc occurs. It is therefore necessary to provide an additional fast acting switch (known as a *crowbar*), in parallel with the capacitor, to divert the energy so that it can be dissipated harmlessly under fault conditions. Capacitors are available with capacitances up to 2 μF and working voltages of 100 kV [12, 20].

The sizes of the smoothing inductor and capacitor and, therefore, the magnitude of the energy stored in the capacitor, can be further reduced by converting the AC input to the rectifier to single phase AC at a higher frequency [21]. This AC to AC conversion is accomplished using a bridge converter connected to a circuit which is resonant at the higher AC frequency. An alternative topology employs a number of converters connected in such a way that their collective output has a high-frequency ripple so that only a small smoothing capacitor is required [11]. Further information on electronic power converters for vacuum tubes can be found in [5, 22].

The current drawn from the power supply may vary according to the operating conditions of the tube, causing variations in the voltage as a result of the internal impedance of the supply. This is particularly true when tubes with a DC high-voltage supply are pulsed by a voltage applied to a control electrode. The variation in voltage can be controlled by a post regulator which is a gridded tube, or a transistor, in series with the tube. The voltage drop across this component can be varied rapidly to compensate for changes in the voltage as the current changes [5].

20.3   Pulse Modulators

Many tubes, especially those for radar and for particle accelerators, are pulsed by applying voltage pulses to the cathode. The pulse duration can be from a few microseconds up to several milliseconds, and the pulse repetition frequency from a few pulses per second up to a few kilohertz. The product of the pulse duration and the pulse repetition frequency, known as the *duty cycle*, is typically of the order of 0.001. Pulse modulators work by storing energy over a period of time and then discharging it into the load [17]. The mean power delivered by the DC power supply is then the product of the peak power and the duty cycle. The energy may be stored in either a capacitance or an inductance. Comparison between capacitive and inductive energy storage shows that the energy density in the latter can be greater by a factor of about one hundred. However the specification for the switches for an inductive system is considerably more demanding, and modulators of this type are not widely used with vacuum tubes [10]. Further information about magnetic pulse modulators can be found in [23, 24]. The discussion below is restricted to capacitive energy storage.

Figure 20.7 shows two basic arrangements for a pulsed modulator with a capacitive energy store. It is usual for the anode, and the tube body, to be earthed because that makes it easier to make RF connections and provide cooling. During the storage period the switch is open and the stored energy is replenished from the DC supply. Then, while the switch is closed, all or part of the stored energy is transferred to the tube. It is not essential for the pulses to occur at regular intervals, provided that the minimum inter-pulse period is sufficient for the energy store to be recharged completely. The purpose of the isolating element is to restrict the current drawn from the power supply during the pulse.

(a)

(b)
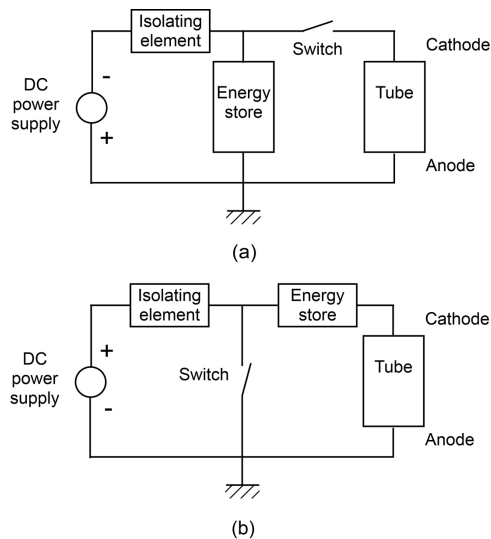
Figure 20.7: Pulsed modulators with a capacitive energy store: (a) direct coupled, and (b) capacitor coupled.

The shape of the voltage pulse can have important effects on the performance of the tube. Excessive voltages of either polarity can cause voltage breakdown. Ripple on the pulse top can produce temporal variations in power output, in the gain and phase of an amplifier, and in the frequency of an oscillator. Unwanted oscillations may occur if the rise and fall times of the pulse are too slow or if the over-swing following the rise of the pulse is too great (see Figure 20.13).

20.3.1   Active-Switch Modulator with a Resistive Load

When the energy store is a capacitor the voltage across it rises slowly during the inter-pulse period, and decays exponentially during the pulse, at a rate set by the storage capacitance and the resistance of the load. Since the voltage across the load should be nearly constant during the pulse, the switch must be opened again after a time which is short compared with the decay constant of the circuit. Therefore the switch must be of a type which can be both opened and closed by a control signal (see Table 20.1). The pulse duration is variable because the switch can be opened at any time. Modulators of this type are known as *active-switch* or *hard-tube* modulators. As only a small fraction of the energy stored in the capacitor is delivered to the load during the pulse it is necessary to provide a crowbar circuit to discharge the capacitor safely under fault conditions. The isolating element can be either a high resistance or an inductor.

The pulse shape delivered to a resistive load can be found by considering the equivalent circuit shown in Figure 20.8(a). The resistance of the load is R_L, the residual resistance of the switch in the on state is R_S, and the capacitor C_S represents

the stray capacitance of the load, the switch, and the connecting cables. During the pulse the voltage across the storage capacitor varies as

$$V_C(t) = V_0 \exp(-\alpha t), \tag{20.4}$$

where V_0 is the initial voltage,

$$\alpha = \frac{1}{(R_L + R_S)C}, \tag{20.5}$$

and the capacitance of the storage capacitor is C. If the switch is closed at $t = 0$ the voltage across the load, during the rise of the pulse, is given by

$$V_L(t) = V_0 \frac{R_L}{R_L + R_S} \exp(-\alpha t) \cdot (1 - \exp(-\alpha_1 t)), \tag{20.6}$$

where

$$\alpha_1 = \frac{R_L + R_S}{R_L R_S C_S} \tag{20.7}$$

and it has been assumed that the rise-time is small compared with the duration of the pulse. During the fall of the pulse the stray capacitance is discharged through the load and the voltage across the load is

$$V_L(t) = V_0 \frac{R_L}{R_L + R_S} \exp(-\alpha t_1) \cdot \exp(-\alpha_2(t - t_1)), \tag{20.8}$$

where

$$\alpha_2 = \frac{1}{R_L C_S}, \tag{20.9}$$

and t_1 is the time at which the switch is opened. Note that the fall of the pulse is much slower than the rise (see Figure 20.8(b)). The long fall time can be reduced by adding a 'tail-biter' switch, in parallel with the load, which is closed at the same moment that the main switch is opened [5]. It is customary to define the rise time of a pulse as the time taken for the voltage to change from 10% to 90% of its final value. The fall time is defined in the same way. The peak voltage is less than the initial voltage on the storage capacitor because of the voltage drop in the switch. It follows that some of the power input to the modulator is dissipated as heat in the switch so that the efficiency of the modulator is less than 100%. If the switching device is a vacuum triode, or a transistor, it is possible to operate in its active region to control the current drawn from the storage capacitor. But the penalty for this is greater power dissipation in the active device and, therefore, reduced modulator efficiency. The voltage on the pulse top falls slowly during the pulse as the capacitor is discharged. The extent of the fall is known as the *droop* of the pulse. The droop can be reduced by increasing the capacitance of the storage capacitor. Alternatively
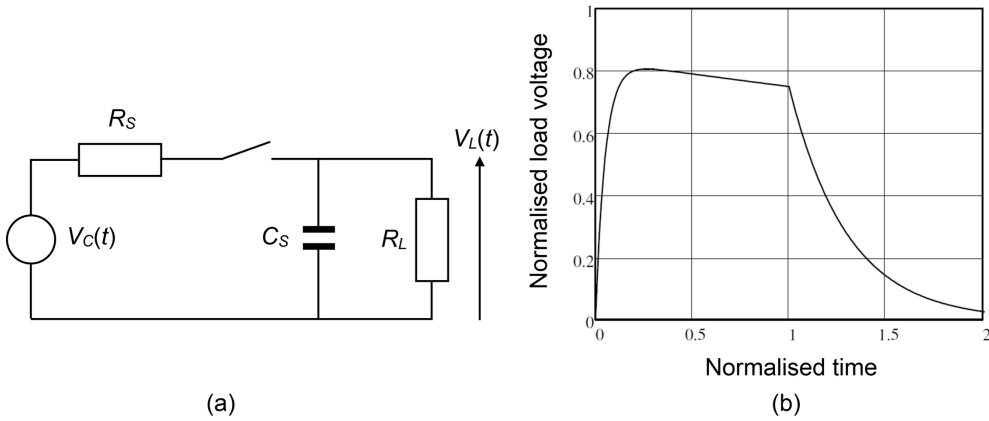
Figure 20.8: Pulse generation by capacitor discharge: (a) equivalent circuit diagram for the rise and fall of the pulse, and (b) a typical pulse shape plotted with the voltage normalised to the initial voltage on the storage capacitor and time normalised to the pulse duration.

a *bouncer circuit* which provides a compensating ramped voltage can be inserted between the load and earth [17]. This approach has been used to compensate for a 20% droop in a 2 msec pulse, leading to a considerable reduction in the energy to be stored in the capacitor [25–27].

20.3.2   Active-Switch Modulator with a Biased Diode Load

When the load is a crossed-field tube, modelled by a biased diode, the equivalent circuit for the rise and fall of the pulse is as shown in Figure 20.9(a). The switch $S1$ controls the pulse, and the biased diode model of the load is represented by $S2$ together with the voltage V_T. The additional impedance Z_R is necessary to allow the stray capacitance to be discharged at the end of the pulse because the tube stops conducting before the voltage has fallen to zero. In a capacitor-coupled modulator it also provides a recharging path for the storage capacitor. In the simplest case this component is a resistor R_R. For purposes of analysis the circuit can be redrawn in terms of current sources as in Figure 20.9(b). Switch $S2$ is closed when $V_L > V_T$ and open otherwise.

The voltage across the storage capacitor can be approximated by

$$V_C(t) = V_0 \exp(-\alpha' t), \tag{20.10}$$

where

$$\alpha' = \frac{1}{(R_1 + R_S)C}, \tag{20.11}$$

and

$$R_1 = \frac{R_L R_R}{R_L + R_R}. \tag{20.12}$$

(a)

(b)
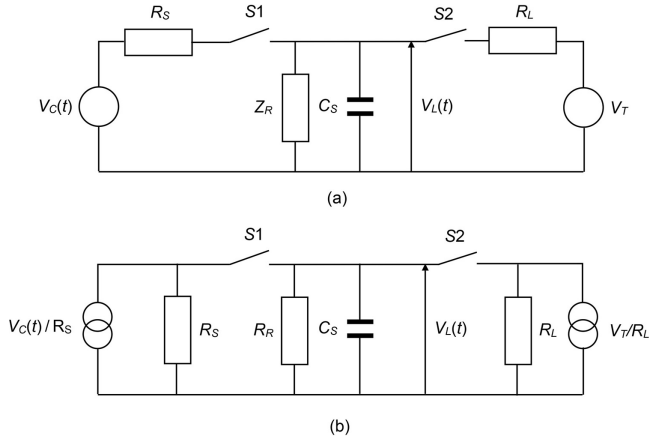
Figure 20.9: Equivalent circuits for an active-switch modulator with a biased diode load: (a) with voltage sources, and (b) with equivalent current sources.

The rise of the pulse is analysed in two steps [17]:

i) From $t = 0$ until $t = t_1$ when $V_L = V_T$

$$V_L(t) = V_0 \frac{R_R}{R_R + R_S} \exp(-\alpha' t) \cdot (1 - \exp(-\alpha_1 t)), \tag{20.13}$$

where

$$\alpha_1 = \frac{R_R + R_S}{R_R R_S C_S}, \tag{20.14}$$

and it is assumed that the rise time is short compared with the duration of the pulse.

ii) From $t = t_1$ until S1 is opened when $t = t_2$

$$V_L(t) = \left[\frac{V_0}{R_S} \exp(-\alpha' t) + \frac{V_T}{R_L} \right] R_2 \left[1 - \exp(-\alpha_2(t - t_1)) \right] + V_T \exp(-\alpha_2(t - t_1)), \tag{20.15}$$

where

$$\frac{1}{R_2} = \frac{1}{R_S} + \frac{1}{R_R} + \frac{1}{R_L}, \tag{20.16}$$

and

$$\alpha_2 = \frac{1}{R_2 C_S}. \tag{20.17}$$

If the droop is small then the voltage on the pulse top is

$$V_L(t) = \left[\frac{V_0}{R_S} + \frac{V_T}{R_L} \right] R_2 . \tag{20.18}$$

The fall of the pulse is analysed in a similarly manner:

iii) From $t = t_2$ until $t = t_3$ when $V_L = V_T$

$$V_L(t) = V_T \frac{R_R}{R_L + R_R} \left[1 - \exp\left(-\alpha_3 \left(t - t_2 \right) \right) \right] + V_L\left(t_2 \right) \exp\left(-\alpha_3 \left(t - t_2 \right) \right), \tag{20.19}$$

where

$$\alpha_3 = \frac{1}{R_1 C_S} . \tag{20.20}$$

iv) From $t = t_3$ onwards

$$V_L(t) = V_T \exp\left(-\alpha_4 \left(t - t_3 \right) \right), \tag{20.21}$$

where

$$\alpha_4 = \frac{1}{R_R C_S} . \tag{20.22}$$

Figure 20.10 shows a typical waveform calculated in this way. The pulse top lies only just above the threshold voltage and the capacitance of the storage capacitor must be large enough so that the droop is small and the voltage does not fall below V_T before the end of the pulse. The losses incurred by the use of the resistor R_R can be reduced by replacing it by an inductor [17]. However, the storage of energy in the inductor during the fall of the pulse can cause the decaying voltage to become oscillatory so that its polarity is periodically reversed. It is undesirable for a large reverse voltage to be imposed on the tube, and an 'over-swing' diode may be added in parallel with the tube to provide a path for the reverse current.

During the pulse an active switch modulator can be regarded as a voltage source whose source resistance is the forward resistance of the switch. The current through the switch is approximately equal to that through the load. Thus, to minimise the power dissipated in the switch, it is desirable for the switch resistance to be small compared with the load resistance. The usual principle of matching the load to the source for maximum energy transfer is not relevant in this case. However, there are circumstances in which it is useful to place a pulse transformer between the modulator and the load. Chief among these is the use of a transformer to provide a high voltage output pulse from a lower voltage modulator. This provides one way of obtaining high voltage pulses from a modulator employing solid state switches. The disadvantages of using a transformer are: the increased current and switch losses in the modulator; losses in the transformer itself; and degradation of the pulse shape
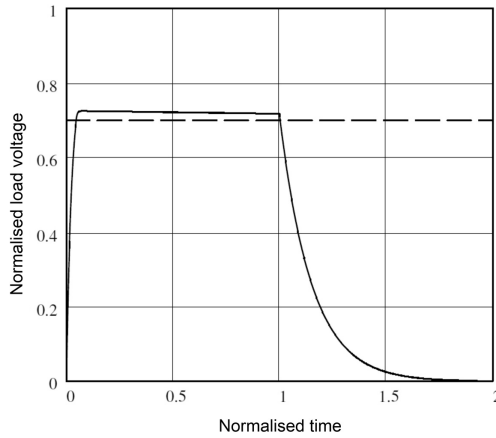
Figure 20.10: Typical voltage pulse for an active-switch modulator connected to a biased diode load.

caused by reactances in the transformer. An alternative approach is the Marx bank modulator in which a number of capacitors are charged in parallel and then discharged in series [28, 29]. A good comparison between different modulator topologies is given in [27] and examples of modulator design in [25, 30]. For discussion of the design of pulse transformers see [16, 17, 31–33].

20.3.3   Line-Type Modulators

An alternative way of generating a flat-topped pulse is to charge a length of TEM transmission line and then discharge it by connecting it to the load, as shown in Figure 20.11(a). Initially both switches are open and the line in uncharged. Switch $S1$ is closed for long enough for the line to charge to the source voltage V_0. When $S2$ is closed the line appears to the load as a voltage source V_0 with an impedance equal to the characteristic impedance of the line Z_c. The load voltage rises instantaneously to a value given by the potential divider rule. If the load is matched to the line $(R_L = Z_c)$ then the voltage across the load is $0.5V_0$. The change in voltage propagates along the line towards the source. If the line is dispersionless then the change in voltage propagates as a step having amplitude $-0.5V_0$ and phase velocity v_p. The voltage on the line is the sum of the initial voltage and this step wave, as shown in Figure 20.11(b). When the wave reaches the start of the line it encounters an open circuit and is reflected as a wave having amplitude $-0.5V_0$, as shown in Figure 20.11(c). The time taken for the reflected step wave to reach the load is $\tau = 2l/v_p$ where l is the length of the line. At that moment the wave is absorbed by the load and the voltage falls to zero. Thus, the load experiences a voltage pulse of amplitude $0.5V_0$ and a duration τ which is fixed by the electrical length of the line. If the load is not matched to the line then the reflected wave is reflected by the load with an amplitude which satisfies the boundary conditions. As a result the pulse
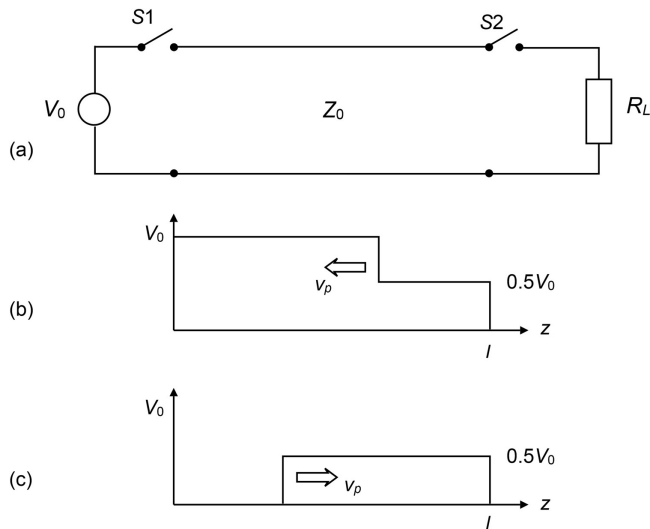
Figure 20.11: Principle of operation of a line-type modulator: (a) circuit diagram, (b) voltage distribution on the line during the first half of the pulse, and (c) voltage distribution on the line during the second half of the pulse.

does not end cleanly and the subsequent voltage may be positive, or oscillatory, depending upon the boundary conditions [17, 34].

The simple arrangement described is not practical, in most cases, because it requires an inconveniently long transmission line. For example, a line with a phase velocity half the velocity of light designed to produce a 1 μs pulse would be 75 m long. Thus a practical line-type modulator employs a pulse-forming network having a number of sections with series inductance and shunt capacitance, as shown in Figure 20.12. The switch $S1$ is a diode, while $S2$ is a thyratron or a thyristor which is switched on by a control signal and switches off automatically when the voltage across it falls to zero. The inductor in series with the DC source, and the capacitance of the pulse-forming network, form a resonant circuit. As the line is charged the voltage rises sinusoidally until it reaches a maximum of $2V_0$, at which value it is held by the diode. Thus the amplitude of the pulse is equal to the DC supply voltage. The impedance of the load is not normally matched to a practical value of characteristic impedance and modulators of this type require a pulse transformer to match the load to the line.

A lumped-element network cannot reproduce the behaviour of a uniform transmission line exactly. Figure 20.13 shows the pulse shape when a uniform five-section network is terminated by a matched resistance. It can be seen that there is an overshoot at the start of the pulse, ripples on the pulse top, and an overshoot at the end of the pulse. The properties of a uniform transmission line can be approached more exactly by increasing the number of sections in the line but the pulse remains imperfect to some extent. A detailed discussion of theoretical methods for designing pulse-forming networks is given in [17]. In one useful form of pulse-forming
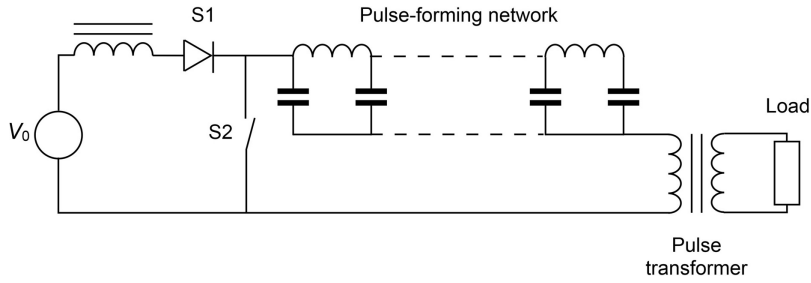
Figure 20.12: Arrangement of a line-type modulator.
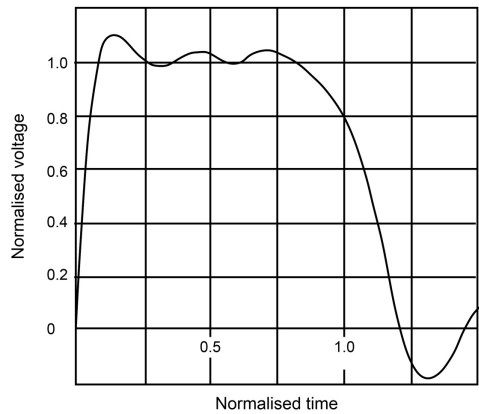
Figure 20.13: Pulse generated by a uniform five-section network terminated by a matched resistance (reproduced, with permission, from [17]).

network the series inductors are wound on a common former so that there is mutual inductance between the sections. It is then possible to use capacitors which all have the same capacitance by suitable choice of the tapping points at which they are connected to the inductors. It is important to recognise that the actual performance of a line-type modulator depends upon the impedance presented by the load, including the effects of reactances in the pulse transformer, and of parasitic reactances. The performance of modulators can be optimised by making empirical adjustments to the components in the pulse-forming network. It has been shown that the ripple in the pulse top can be reduced by using a pair of pulse-forming networks arranged so that their contributions to the ripple are in anti-phase [35]. The design of a simple line-type modulator with a pulse transformer is described in [36].

20.4   RF Systems

The RF system into which the tube is incorporated has an influence on its performance. The construction and calibration of the system are important, as also are

differences between the system, and the test stand used by the tube manufacturer. When the tube is an amplifier the correct drive level is important. Overdriving a tube increases its non-linearity and can cause oscillations and overheating. Because some RF power may emerge from the input port of the tube it is necessary to ensure that a good match is presented by the input system. Under fault conditions the returned power may be high, and at frequencies other than the normal operating frequency, so that it is necessary to protect the input system by a circulator.

When a tube is tested by the manufacturer the output is usually directed to a high power load which has a good match over a wide range of frequencies. The effects of an incorrect match may be variations in the power output, in the gain and phase of an amplifier, and in the frequency and spectrum of an oscillator. In extreme cases it may lead to RF voltage breakdown within the tube. The effects of the RF match are indicated on a Rieke diagram which may show regions where operation is forbidden (see Figures 13.12 and 15.29). The RF match presented by the system may be very different from that of the test stand and may affect the correct operation of the tube. For this reason the output of the tube is normally directed through a high-power circulator. However, the performance of a circulator is usually only specified over the working frequency band of the system. Thus, its properties at harmonic and other out-of band frequencies may not be defined. If a filter is included in the system to block power at these frequencies, the effect of reflection into the output port of the tube must be considered. The construction of the RF output system must conform to good high power practice and particular care must be taken to avoid resonances. The output power of the tube and the power reflected by the system are measured using directional couplers whose calibration is important. The measurement of power at harmonic frequencies is not straightforward because the power may be divided between several waveguide modes [37]. Reflection of output power into the tube may cause overheating within it which may be indicated by increased gas pressure. It can also cause failure of the output window if a peak of the standing wave occurs there. Abnormal causes of power reflection include arcs and multipactor discharges in the output waveguide. Reflected power, and waveguide arc, detectors connected to fast trips are used to protect the tube (see Section 20.6.2). The design and maintenance of high power RF systems is discussed in [4, 38].

20.5   Cooling System

The methods used for cooling the collectors of IOTs, linear-beam tubes and gyrotrons are discussed in Section 10.5. Similar methods are used for cooling the anodes of gridded tubes, magnetrons and CFAs. The greatest part of the heat is dissipated on these electrodes. However, it is usually necessary to cool the body of the tube, and cooling may also be required for electromagnets and output windows. The surfaces to be cooled may not be at earth potential. The density of the power to be dissipated is frequently great enough for the tube to be permanently damaged if the

cooling is inadequate. Overheating can cause the release of adsorbed gases leading to an increase in the pressure within the tube. If the background gas pressure rises then poisoning of the cathode may lead to reduced emission, and ion oscillations may add noise to the tube output. Large tubes are usually fitted with at least one ion pump whose function is to monitor, and maintain, the vacuum within the tube. Permanent damage is revealed by changes in the RF performance of the tube. Under normal operation the internal dimensions of the tube change because of the thermal expansion, leading to changes in performance with temperature. Vibration caused by cooling fans, pumps, and turbulence in the cooling system can cause internal parts of a tube to vibrate, imposing low frequency modulation on the RF output.

If the tube is to be cooled correctly it is important that air and liquid flow rates should be adequate, that the incoming coolant temperature is correct, and that the surfaces to be cooled are clean and not damaged by corrosion. In air-cooled systems corrosion can be caused by excessive humidity. In water-cooled systems galvanic corrosion may occur, leading to shortening of the life of the tube [39]. To prevent this it is important to consider the metals used in the construction of the cooling system, the purity and pH value of the water used, and the nature of any additives such as anti-freeze. Further information on the design and maintenance of cooling systems is given in [4, 5, 38, 40].

20.6   Control System

The purpose of the control system is to ensure correct and safe operation of the tube, and to protect it under fault conditions. The details of the system depend upon the tube type, and the power level at which it is operating, but the general principles are the same for all.

20.6.1   Interlocks

The parts of the system shown in Figure 20.1 must be interlocked in such a way that the different parts are turned on and off in the correct order. In addition, the interlocks provide protection if any of the correct operating conditions, including the maximum ratings specified by the tube manufacturer, are violated. The interlocks may also include conditions relating to other parts of the complete system (for example the cooling of a high-power circulator). Examples of switch-on sequences are given in [5–7]. At each stage the operating conditions are monitored so that the next stage cannot commence unless they are within prescribed limits. If any quantity lies outside the permitted limits then the system is shut down. A typical start-up sequence is:

i) The cooling and auxiliary supplies (electromagnets, ion pump) are turned on.
ii) The cathode heater is turned on with a time delay which inhibits the next step until the full heater temperature has been reached.

iii) The control grid bias voltage is applied.

iv) The anode voltage is applied. In some cases the voltage is increased gradually until the working voltage is reached.

v) The collector voltages are applied to a tube with depressed collector electrodes. The screen grid voltage is applied to a tetrode.

vi) The RF drive is applied. Note that in some linear-beam tubes the collector may not be designed to dissipate the full beam power in the absence of RF drive. It is then necessary for the anode voltage and the RF drive to be applied simultaneously and for the beam power to be reduced if the RF drive fails [41].

vii) The heater power of a magnetron or CFA is reduced to compensate for heating by back-bombardment of the cathode.

The same steps are followed in the reverse order to turn the tube off. Because parts of the tube may still be quite hot after it has been turned off it may be necessary to keep the cooling systems in operation for some time.

20.6.2   Tube Protection

The interlocks described above cannot normally act fast enough to protect the tube under fault conditions. For this purpose, trips are required which can remove power from the tube under fault conditions within tens of microseconds, before damage can occur. For tubes with DC supplies this is achieved by firing the crowbar switch of the power supply. For cathode-pulsed tubes, the triggering of the next pulse must be suppressed. Faults to be detected include:

- DC arcs inside the tube indicated by excessive cathode, or grid, current;
- excessive reflected power and RF arcs in the output waveguide;
- excessive current interception on the tube body;
- grid bias failure;
- depressed collector bias failure;
- excessive gas pressure

20.7   Care of Tubes

The performance of a tube can be affected by its environment during storage, transportation, installation, and operation. Because tubes appear to be mechanically robust they may be treated carelessly, with the result that the performance is degraded and the life of the tube reduced. Some common problems are:

- The copper body of a tube is very soft as a result of annealing during manufacture so that the precision of the alignment of internal parts may be disturbed by shocks.
- Ceramic insulators are brittle and can be fractured so that the vacuum is lost.

- The cathode assembly of a tube is often cantilevered, making it liable to be displaced by transverse shocks. Internal spacings, especially those of grids, can be very small and critical to the operation of the tube.
- The magnetisation of permanent magnets can be altered by exposure to external magnetic fields or ferromagnetic materials. The presence of the magnet of a high-power circulator close to the tube may cause problems with focusing.
- The external surfaces of insulators can become dirty, and metal surfaces can be corroded.

For all these reasons it is important that a tube should always be transported in the packing provided by the manufacturer, and the handling procedures specified should be followed with care. If ion pumps are fitted they can be used to monitor and maintain the vacuum during storage. When the manufacturer specifies a gas test this should be carried out before the tube is installed. Some tubes are very heavy and should only be handled and installed with appropriate equipment. After a tube has been installed it is usually necessary to follow a conditioning procedure before it is ready for normal use. That can include gradually raising the operating voltage over an extended period of time (see Section 18.7.1). A detailed discussion of these issues, and of the correct procedures is given in [4, 38].

20.8   Safety

Vacuum tubes are high power devices whose safe operation requires careful attention to the construction and use of the systems in which they are employed. A detailed discussion of the issues involved is given in [4, 38, 42].

Tubes operate at voltages which are potentially lethal and proper high-voltage precautions must be taken. All high-voltage connections must be enclosed, or adequately insulated, and interlocks provided so that high-voltage terminals are inaccessible when they are live. The design of the equipment must ensure that all high voltage terminals are earthed before any access to the enclosure is possible. High voltage hazard warning notices should be clearly displayed. Great care must be taken to keep earthing straps short and of very low inductance. Under fault conditions the rate of change of current with time can be very high so that high-induced voltages may exist on exposed metalwork (see Section 18.7.1).

RF power is dangerous to biological tissue especially the eyes. For this reason one should never look into the end of a waveguide carrying microwave power at any level. All joints in the high-power waveguide system must be correctly made and checked for leakage. It is important to be aware that RF power may be emitted from parts of the tube such as the collector insulator of a linear-beam tube, or the cathode insulator of a crossed-field tube. There should, therefore, be adequate screening against such emissions, including those which only exist under fault conditions. National regulations for safe levels of exposure to radio-frequency radiation should be observed.

X-rays are produced by the impact of electrons on solid surfaces at voltages of 5 kV or more, as described in Section 18.4. Low-energy X-rays are screened out by the metal body of a tube but may still escape through ceramic insulators. The energies of electron impacts in crossed-field tubes are smaller than those in other tubes, but X-rays can still be emitted under fault conditions. If X-rays are likely to be generated then the tube must only be operated in a properly screened enclosure and national regulations for the protection of personnel must be observed [43, 44].

Some tubes contain beryllium oxide ceramics or beryllium metal parts. Beryllium oxide dust is highly toxic, and great care must be taken in handling a tube in which a ceramic has been broken. Otherwise a tube containing these materials is safe to handle unless an attempt is made to dismantle it. The tube should be clearly marked with a hazard label.

Sulphur hexafluoride (SF_6) and Freon gases are sometimes used to prevent breakdown in high-voltage systems and high-power waveguides (see Section 18.7.2). If a breakdown does occur the gas may decompose, and combine with air or water vapour, to produce highly toxic and corrosive compounds. When systems use these gases it is essential to establish safe procedures for handling the consequences of a breakdown.

If an insulator is broken by a mechanical shock, the vacuum within a tube can cause fragments to fly outwards at high velocity, leading to injury. Correct handling procedures should be used to avoid this. Especial care is required if beryllium oxide ceramics are present.

Exposed surfaces on an air-cooled tube may be at temperatures of 200 °C to 300 °C during operation, and they may remain dangerously hot for some time after the tube has been turned off. The cooling fans should continue to operate until the temperature has been reduced to a safe level. Liquid coolants can also become dangerously hot, and care is needed to avoid breakage of cooling pipes, especially if they are under pressure.

20.9   Reliability

The reliability of a vacuum tube can impinge on the reliability of the system, in which it is installed, in two ways. First, there are short-term system outages caused by trips when the tube is otherwise performing correctly. These can be common when a tube is first installed, after a period of storage, and their frequency decreases if the tube is conditioned by gradually raising the anode voltage. Sometimes the frequency of tripping increases after a long period of operation, but it can be reduced by repeating the conditioning process. The duration of each outage depends upon the length of time required to restart the tube after a trip. The effect of the outages may be only a minor irritation, or it may be serious where continuous availability of

the system is important. The second way in which the reliability of the tube affects that of the system is the point at which the tube no longer delivers the performance required so that it has to be replaced. This can be a major operation during which the system is not available for hours, or even days. Both kinds of outage affect the cost of ownership of the system which needs to be minimised by careful design and operation.

The normal end of life of a tube is determined by reduction in emission of electrons from the cathode, caused by evaporation of barium. If the gas pressure in a tube is too high, then the ions generated may bombard the cathode surface. This is more likely to occur in CW tubes than in pulsed tubes. The result may be damage by sputtering of material from the surface, or by 'poisoning', leading to reduced emission. A third cause of low emission is low cathode heater power, which may be caused by an internal short circuit, or by incorrect setting of the heater supply. If the cathode becomes temperature-limited, because the cathode work function has increased, or the heater power is too low, then the cathode current varies with the heater current. In addition, the reduction in the current drawn from the high voltage supply may cause its voltage to rise and voltage breakdown may occur. In magnetrons reduced emission leads to a poor RF spectrum. The lifetime of a tube can sometimes be extended by de-rating it to extend the cathode life, or by increasing the heater current when the emission starts to fall. The tube manufacturer should be consulted before either of these courses of action is pursued.

It is found that many tube failures are caused by incorrect handling, or by system faults. The ways in which these affect the performance of the tube have been discussed in the preceding sections. The tendency to assume that the tube is faulty, and to attempt to cure system faults by replacing it, should be resisted. The tube is an expensive component and the cost of replacing it is not trivial. Therefore proper diagnostic procedures should be followed to ensure that the fault does not lie elsewhere, before the tube is replaced [4]. Table 20.2 shows some typical symptoms and causes of tube failure. Further information is to be found in [4, 5].

20.10   Conclusion

Vacuum tubes continue to be important components in systems of many kinds. Progress in design, and in manufacturing technology, continue to yield improvements in performance, reliability, and lifetime, as noted in Chapter 1 of this book. As a result, the performance of systems incorporating tubes is also improving, and new applications are becoming possible. However, if the full benefits are to be achieved, it is essential that the tube engineers, and the systems engineers, have a good understanding of each others' technologies. It is hoped that this book will contribute to that understanding by providing a detailed, and authoritative, account of the principles of vacuum electron devices.

Table 20.2: Symptoms and causes of tube failure

Symptom	Internal causes	External causes
Low cathode current	Cathode end of life Cathode poisoning Cathode erosion Heater short circuit	Low anode voltage Low heater power
Voltage breakdown	Contamination of insulators Increased gas pressure	Contamination of insulators High anode voltage
High body current	Mechanical damage Demagnetised magnets	
Overheating	Mechanical damage Demagnetised magnets	Wrong electromagnet current High RF drive Inadequate cooling High ambient temperature
High ion pump current	Mechanical damage Overheating	Wrong electromagnet current High RF drive Inadequate cooling High ambient temperature
High grid current	Short-circuit	Wrong bias voltage
Wrong RF power	Mechanical damage Demagnetised magnets	Wrong RF match Wrong grid bias Wrong electromagnet current Wrong anode voltage Wrong RF drive
Oscillation	Mechanical damage Demagnetised magnets	Wrong RF match Wrong electromagnet current Wrong anode voltage Wrong RF drive
Poor spectrum (magnetrons)	Reduced cathode emission Demagnetised magnet	Wrong anode voltage

References

[1] J. W. Hansen *et al.*, *Applications Note: System Aspects of Communications TWTAs, or How to Deal with the Tube Manufacturer to Your Best Advantage.* Torrance, CA: Hughes Aircraft Company, Electron Dynamics Division, 1982.

[2] C. R. Smith *et al.*, 'The microwave power module: a versatile RF building block for high-power transmitters', *Proceedings of the IEEE,* vol. 87, pp. 717–737, 1999.

[3] M. Katzman and R. Wills, *Microwave Tube Reliability.* San Diego, CA: Naval Electronics Lab Center, CA, 1976.

[4] Varian, *Technical Manual: Installation, Operation, Maintenance, Care and Handling Instructions, General: Microwave Tubes, Magnetron Tubes, Electron Tubes.* 1 October 1979.

[5] L. Sivan, *Microwave Tube Transmitters*. London: Chapman and Hall, 1994.

[6] M. J. Smith and G. Phillips, *Power Klystrons Today*. Taunton, UK: Research Studies Press, 1995.

[7] EEV Ltd, *A User's Guide to the Coupled-Cavity TWT*. Chelmsford, UK: EEV Ltd, 1987.

[8] E. Carroll, 'High-power active devices', in D. Brandt, ed., *Specialized Course on Power Converters*. Geneva: CERN, pp. 43–56, 2004.

[9] EEV Ltd, *Hydrogen Thyratron Preamble*. Chelmsford, UK: EEV Ltd., 2002.

[10] P. D. Pearce, *Pulsed Power for Future Linear Accelerators*. Geneva: CERN, 1999.

[11] D. Poole *et al.*, 'A crowbarless high voltage power converter for RF klystrons', in *EPAC 96*, pp. 2326–2328, 1996.

[12] D. E. Anderson, 'Recent developments in pulsed high power systems', presented at LINAC 2006, Knoxville, Tennessee, 2006.

[13] C. Burkhart *et al.*, 'Development of a solid state thyratron replacement for the LCLS klystron modulator', in *19th IEEE Pulsed Power Conference*, pp. 1–4, 2013.

[14] N. Mohan *et al.*, *Power Electronics: Converters, Applications and Design*, 3rd ed. Hoboken, NJ: Wiley & Sons Inc., 2003.

[15] R. B. True *et al.*, 'The Hobetron – a high power vacuum electronic switch', *IEEE Transactions on Electron Devices,* vol. 48, pp. 122–128, 2001.

[16] H. W. Lord, 'Pulse transformers', *IEEE Transactions on Magnetics,* vol. 7, pp. 17–28, 1971.

[17] G. N. Glasoe and J. V. Lebacz, eds, *Pulse Generators* (MIT Radiation Lab Series). New York: McGraw-Hill, 1948.

[18] D. Brandt, ed., *CAS Specialized Course on Power Convertors*. Geneva: CERN, 2004.

[19] J. E. Theed *et al.*, '52 kV power supply for energy recovery linac prototype RF' in *EPAC 2004*, Lucerne, pp. 1762–1764, 2004.

[20] Sorrento Electronics Inc., *High Energy Capacitors Overview.* San Diego, Sorrento Electronics, Inc, 2002.

[21] D. J. Cook *et al.*, 'Design of a direct converter for high power RF applications', in *PAC 07*, Albuquerque, NM, pp. 2221–2223, 2007.

[22] I. Barbi and R. Gules, 'Isolated DC-DC converters with high-output voltage for TWTA telecommunication satellite applications', *IEEE Transactions on Power Electronics,* vol. 18, pp. 975–984, 2003.

[23] E. W. Manteuffel and R. E. Cooper, 'D.C. charged magnetic pulse modulator', *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics,* vol. 78, pp. 843–850, 1960.

[24] J. Mankowski and M. Kristiansen, 'A review of short pulse generator technology', *IEEE Transactions on Plasma Science,* vol. 28, pp. 102–108, 2000.

[25] H. Pfeffer *et al.*, 'A long pulse modulator for reduced size and cost', in *21st International Power Modulator Symposium*, Costa Mesa, CA, 1994.

[26] H. Pfeffer *et al.*, 'A second long pulse modulator for TESLA using IGBTs', in *Fifth European Particle Accelerator Conference*, pp. 2585–2587, 1996.

[27] C. A. Martins and K. Rathsman, 'An assessment on klystron modulator topologies for the ESS project', in *Pulsed Power Conference*, Chicago, IL, 2011.

[28] J. Casey *et al.*, 'Marx bank technology for accelerators and colliders', in *LINAC08*, Victoria, BC, Canada, pp. 1002–1004, 2008.

[29] R. L. Cassel, 'An all solid state pulsed Marx type modulator for magnetrons and klystrons', in *IEEE Pulsed Power Conference*, pp. 836–838, 2005.

[30] R. Richardson *et al.*, 'High energy pulse modulator for microwave tube applications', in *25th International Power Modulator Symposium*, pp. 240–243, 2002.

[31] M. Akemoto *et al.*, 'Pulse transformer R & D for NLC klystron pulse modulator', in *11th IEEE International Pulsed Power Conference*, vol. 1, pp. 724–729, 1997.

[32] M. Akemoto and A. Tokuchi, 'A 600 kV, 150 ns rise-time pulse transformer for klystron pulse modulator', in *IEEE International Power Modulators and High Voltage Conference*, p. 396, 2008.

[33] D. Bortis *et al.*, 'Design procedure for compact pulse transformers with rectangular pulse shape and fast rise times', in *IEEE International Power Modulator and High Voltage Conference*, pp. 298–302, 2010.

[34] R. G. Carter, *Electromagnetism for Electronic Engineers*. Available at: http://bookboon.com/en/electromagnetism-for-electronic-engineers-ebook (accessed 5 October 2017) 2010.

[35] C. B. Figley, 'Results of a modulator pulse top ripple reduction study', *IEEE Transactions on Nuclear Science,* vol. 40, pp. 114–119, 1993.

[36] N. Carleto and C. C. Motta, 'Design, construction and characterization of a line-type pulse modulator for driving high power magnetron', in *IEEE MTT-S International Conference on Microwave and Optoelectronics*, pp. 330–333, 2005.

[37] V. Price, 'Measurement of harmonic power generated by microwave transmitters', *IRE Transactions on Microwave Theory and Techniques,* vol. 7, pp. 116–120, 1959.

[38] J. Whitaker, *Power Vacuum Tubes Handbook*. CRC Press, 2012.

[39] R. Heppinstall and G. Clayworth, 'The importance of water purity in the successful operation of vapour-cooled television klystrons', *Radio and Electronic Engineer,* vol. 45, pp. 391–399, 1975.

[40] EIMAC, *Care and Feeding of Power Grid Tubes*, 5th ed. San Carlos, CA: CPI Inc. Eimac Division, 2003.

[41] H. Frischholz, 'Generation and distribution of radio-frequency power in LEP', *IEEE Transactions on Nuclear Science,* vol. NS-32, pp. 2791–2793, 1985.

[42] R. Kitchen, *RF and Microwave Radiation Safety Handbook*. Oxford: Newnes, 2001.

[43] H. E. Johns and R. J. Cunningham, *The Physics of Radiation Protection*, 3rd ed. Springfield, IL: Charles C. Thomas, 1974.

[44] W. Meredith and J. B. Massey, *Fundamental Physics of Radiology*, 2nd ed. Baltimore: Williams and Wilkins, 1972.

Appendix: Mathcad Worksheets

Available online at www.cambridge.org/9780521198622

WS 1.1 Coupled Modes
WS 1.2 Modulation
WS 1.3 Linearity
WS 2.1 Ridged Waveguides
WS 2.2 Waveguide Capacitive Iris
WS 2.3 Waveguide Inductive Iris
WS 3.1 Resonant Circuits
WS 3.2 Cavity Resonators
WS 3.3 Re-entrant with Beam Hole
WS 3.4 Fujisawa's model
WS 3.5 Gap Field
WS 3.6 Iris-Coupled Cavity
WS 4.1 Capacitance of a Planar Grid
WS 4.2 Meander Line
WS 4.3 Sheath Helix
WS 4.4 Capacitance of a Cylindrical Grid
WS 4.5 Tape Helix Model
WS 4.6 Ring-Bar Structure
WS 4.7 Folded Waveguide
WS 4.8 Space-Harmonic Structure
WS 5.1 Planar Diode
WS 5.2 Cylindrical Diode, Relativistic Effects
WS 5.3 Spherical Diode, Relativistic Effects
WS 5.4 Planar Diode with Initial Velocity
WS 6.1 Electrostatic Triode with Island Formation
WS 6.2 Effective Diode Spacing
WS 6.3 Tetrode Fields
WS 7.1 Calculation of Beam Parameters
WS 7.2 Solenoid Focusing
WS 7.3 Beam Spreading Curves
WS 7.4 PPM Focusing (Mean Radius)
WS 7.5 Periodic Electrostatic Focusing

WS 8.1 Electron Trajectories
WS 8.2 Planar Crossed-Field Diode
WS 8.3 Cylindrical Crossed-Field Diode
WS 9.1 Pierce Electron Gun
WS 9.2 PPM Stack Entry
WS 10.1 Collector Design
WS 10.2 Collector Depression (Klystron)
WS 10.3 Collector Depression (IOT)
WS 10.4 Collector Depression (TWT)
WS 11.1 Gap Coupling
WS 11.2 Reduced Plasma Frequency
WS 11.3 Beam Gap Interaction
WS 11.4 Continuous TWT
WS 11.5 Discrete TWT
WS 11.6 Large Signal Gap Model
WS 12.1 Amplifier Classes
WS 12.2 ML-5681 Triode Characteristics
WS 12.3 RS 2058 Tetrode
WS 12.4 Tetrode Design
WS 12.5 IOT Model
WS 12.6 IOT Collector Depression
WS 13.1 Klystron Small Signal
WS 13.2 Klystron Simplified Small Signal
WS 13.3 Klystron Large Signal
WS 13.4 Klystron Efficiency
WS 13.5 Klystron Spent-Beam Characteristics
WS 13.6 Klystron Design
WS 14.1 Helix TWT Small Signal
WS 14.2 Helix TWT Large Signal
WS 14.3 Helix TWT Design
WS 15.1 Magnetron Anode Worksheet 4J50
WS 15.2 Rising Sun Anode Worksheet
WS 15.3 4J50 Magnetron
WS 15.4 Magnetron Design Sheet
WS 16.1 SFD-262 Model
WS 17.1 Gyrotron Model
WS 18.1 Two Surface Multipactor Model
WS 18.2 Crossed Field Multipactor Model
WS 19.1 Coil Field
WS 19.2 Permanent Magnet Design
WS 19.3 PPM Stack Design
WS 20.1 Power Converters
WS 20.2 Pulse Shape

Index

'I wholeheartedly recommend this book.'
John H Booske, *University of Wisconsin-Madison*

'Richard Carter's fifty years of internationally-leading expertise is distilled in this essential reading for everyone working on high power microwave and RF sources.'
Alan D. R. Phelps, *FRSE, University of Strathclyde*

'…excellent and comprehensive…'
Manfred Thumm, *Karlsruhe Institute of Technology*

'This book is a tour de force … and a "must have" for anyone in the field.'
Robert Rimmer, *JLab*

Do you design and build vacuum electron devices, or work with the systems that use them? Quickly develop a solid understanding of how these devices work with this authoritative guide, written by an author with over 50 years of experience in the field. Rigorous in its approach, it focuses on the theory and design of commercially significant types of gridded, linear-beam, crossed-field and fast-wave tubes. Essential components such as waveguides, resonators, slow-wave structures, electron guns, beams, magnets and collectors are also covered, as well as the integration and reliable operation of devices in microwave and RF systems. Complex mathematical analysis is kept to a minimum, and Mathcad worksheets supporting the book online aid understanding of key concepts and connect the theory with practice. Including coverage of primary sources and current research trends, this is essential reading for researchers, practitioners and graduate students working on vacuum electron devices.

Richard G. Carter is a Professor Emeritus in the Department of Engineering at Lancaster University, and a Fellow of the IET.

Online Resources
www.cambridge.org/carter
↘ Mathcad worksheets

Cover illustration: background image © Getty Images/traffic_analyzer