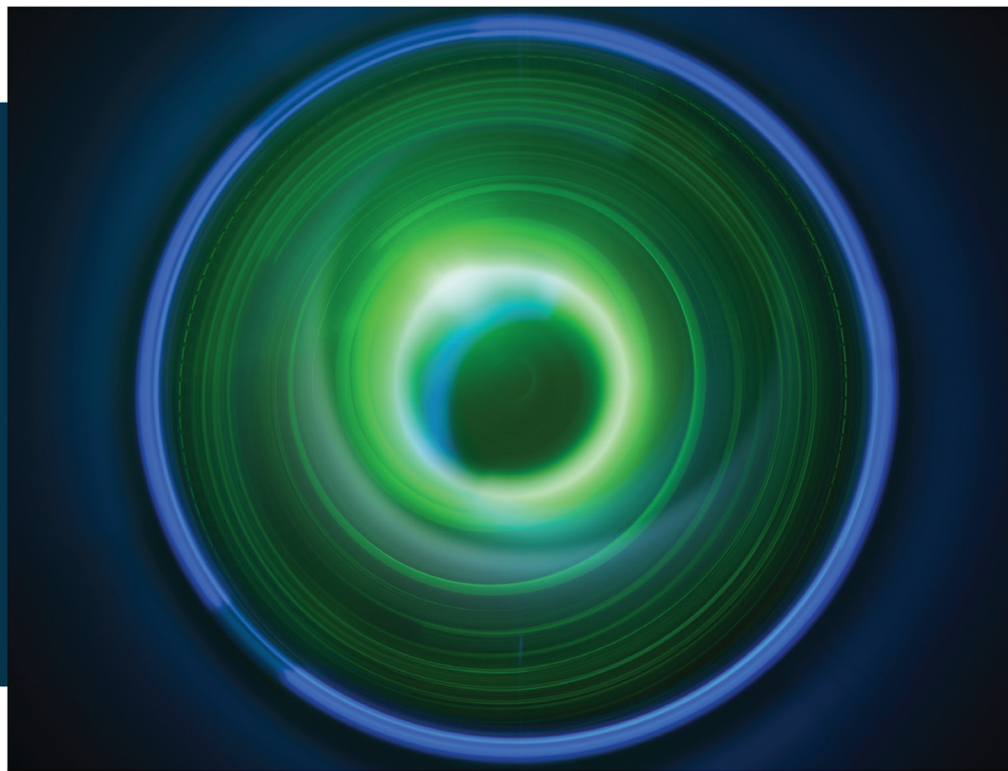


SERIES IN OPTICS AND OPTOELECTRONICS

Thin-Film Optical Filters

FIFTH EDITION



H. Angus Macleod

Thin-Film Optical Filters

Fifth Edition

SERIES IN OPTICS AND OPTOELECTRONICS

Series Editors: **Robert G W Brown**, University of California, Irvine, USA
E Roy Pike, Kings College, London, UK

Recent titles in the series

Optical MEMS, Nanophotonics, and Their Applications

Guangya Zhou and Chengkuo Lee (Eds.)

Optical Microring Resonators: Theory, Techniques, and Applications

Vien Van

Thin-Film Optical Filters, Fifth Edition

H. Angus Macleod

**Handbook of Optoelectronics, Second Edition:
Concepts, Devices, and Techniques – Volume One**

John P. Dakin and Robert Brown (Eds.)

**Handbook of Optoelectronics, Second Edition:
Enabling Technologies – Volume Two**

John P. Dakin and Robert Brown (Eds.)

**Handbook of Optoelectronics, Second Edition:
Applied Optical Electronics – Volume Three**

John P. Dakin and Robert Brown (Eds.)

Handbook of GaN Semiconductor Materials and Devices

Wengang (Wayne) Bi, Hao-chung (Henry) Kuo, Pei-Cheng Ku,
and Bo Shen (Eds.)

**Handbook of Optoelectronic Device Modeling and Simulation:
Fundamentals, Materials, Nanostructures, LEDs, and Amplifiers
– Volume One**

Joachim Piprek (Ed.)

**Handbook of Optoelectronic Device Modeling and Simulation:
Lasers, Modulators, Photodetectors, Solar Cells, and Numerical Methods
– Volume Two**

Joachim Piprek (Ed.)

Nanophotonics and Plasmonics: An Integrated View

Dr. Ching Eng (Jason) Png and Dr. Yuriy Akimov

Handbook of Solid-State Lighting and LEDs

Zhe Chuan Feng (Ed.)

Optical Compressive Imaging

Adrian Stern

Thin-Film Optical Filters

Fifth Edition

H. Angus Macleod

Thin Film Center Inc., Tucson, Arizona, USA
Professor Emeritus of Optical Sciences,
University of Arizona, Tucson, USA



CRC Press

Taylor & Francis Group
Boca Raton London New York

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

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742
© 2018 by Taylor & Francis Group, LLC

CRC Press is an imprint of Taylor & Francis Group, an Informa business
No claim to original U.S. Government works

Printed on acid-free paper

International Standard Book Number-13: 978-1-138-19824-1 (Hardback)

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

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

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

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

Library of Congress Cataloging-in-Publication Data

Names: Macleod, H. A. (Hugh Angus), author.
Title: Thin-film optical filters / H. Angus Macleod.
Other titles: Series in optics and optoelectronics (CRC Press)
Description: Fifth edition. | Boca Raton, FL : CRC Press, Taylor & Francis Group, [2018] | Series: Series in optics and optoelectronics | Includes bibliographical references and index.
Identifiers: LCCN 2017025765 | ISBN 9781138198241 (hardback ; alk. paper) | ISBN 1138198242 (hardback ; alk. paper)
Subjects: LCSH: Light filters. | Thin films—Optical properties.
Classification: LCC QC373.L5 M34 2018 | DDC 681/.42—dc23
LC record available at <https://lccn.loc.gov/2017025765>

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

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

In memory of my Mother and Father

*Agnes Donaldson Macleod
John Macleod*



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Brief Contents

Series Preface	xv
Preface to the Fifth Edition.....	xvii
Preface to the Fourth Edition.....	xix
Preface to the Third Edition	xxi
Preface to the Second Edition	xxiii
Apologia to the First Edition	xxv
Author.....	xxvii
Symbols and Abbreviations.....	xxix
1. Introduction	1
2. Basic Theory	11
3. Theoretical Techniques	65
4. Antireflection Coatings	97
5. Neutral Mirrors and Beam Splitters	161
6. Multilayer High-Reflectance Coatings	183
7. Edge Filters and Notch Filters	213
8. Bandpass Filters	251
9. Tilted Dielectric Coatings	337
10. More on Tilted Coatings	379
11. Other Topics: From Rugate Filters to Photonic Crystals	403
12. Color in Optical Coatings	457
13. Production Methods	471
14. Material Properties	535
15. Composite, Birefringent, and Metamaterials	593
16. Some Coating Properties Important in Systems	607
17. Specification of Filters and Coatings	635
18. Characteristics of Thin-Film Dielectric Materials	645
Index	



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Detailed Contents

Series Preface	xv
Preface to the Fifth Edition.....	xvii
Preface to the Fourth Edition.....	xix
Preface to the Third Edition	xxi
Preface to the Second Edition	xxiii
Apologia to the First Edition.....	xxv
Author.....	xxvii
Symbols and Abbreviations.....	xxix
1. Introduction	1
1.1 Early History	1
1.2 Thin-Film Filters.....	4
References.....	8
Bibliography	9
2. Basic Theory	11
2.1 Maxwell's Equations and Plane Electromagnetic Waves	11
2.1.1 Poynting Vector	16
2.2 Notation	18
2.3 Simple Boundary.....	19
2.3.1 Normal Incidence in Absorption-Free Media.....	21
2.3.2 Oblique Incidence in Absorption-Free Media	23
2.3.3 Optical Admittance for Oblique Incidence	27
2.3.4 Normal Incidence in Absorbing Media	28
2.3.5 Oblique Incidence in Absorbing Media.....	33
2.4 Critical Angle and Beyond.....	36
2.5 Reflectance of a Thin Film.....	38
2.6 Reflectance of an Assembly of Thin Films	42
2.7 Reflectance, Transmittance, and Absorptance	44
2.8 Units	46
2.9 Summary of Important Results	47
2.10 Potential Transmittance	51
2.11 Theorem on the Transmittance of a Thin-Film Assembly	53
2.12 Coherence.....	54
2.13 Mixed Poynting Vector.....	58
2.14 Incoherent Reflection at Two or More Surfaces.....	60
References.....	64
3. Theoretical Techniques	65
3.1 Quarter- and Half-Wave Optical Thicknesses	65
3.2 Admittance Loci.....	66
3.2.1 Electric Field and Losses in the Admittance Diagram	70
3.3 Vector Method	75
3.4 Other Techniques	77
3.4.1 Herpin Index	77
3.4.2 Alternative Method of Calculation	78

3.4.3	Smith's Method of Multilayer Design	80
3.4.4	Smith Chart	82
3.4.5	Reflection Circle Diagrams	84
3.4.6	Automatic Design	88
	References.....	96
4.	Antireflection Coatings	97
4.1	Single Layer	98
4.2	Two-Layer Antireflection Coatings	104
4.3	Multilayer Antireflection Coatings	117
4.4	Equivalent Layers	135
4.5	Antireflection Coatings for Two Zeros	145
4.6	Antireflection Coatings for the Visible and the Infrared Regions.....	148
4.7	Inhomogeneous Layers	154
4.8	Further Information.....	157
	References.....	157
5.	Neutral Mirrors and Beam Splitters	161
5.1	High-Reflectance Mirror Coatings	161
5.1.1	Metallic Layers	161
5.1.2	Protection of Metal Films.....	163
5.1.3	Overall System Performance, Enhanced Reflectance	168
5.1.4	Reflecting Coatings for the Ultraviolet Region.....	170
5.2	Neutral Beam Splitters.....	172
5.2.1	Beam Splitters Using Metallic Layers.....	173
5.2.2	Beam Splitters Using Dielectric Layers.....	177
5.3	Neutral-Density Filters	180
	References.....	181
6.	Multilayer High-Reflectance Coatings	183
6.1	Fabry-Perot Interferometer	183
6.2	Multilayer Dielectric Coatings.....	188
6.2.1	All-Dielectric Multilayers with Extended High-Reflectance Zones	195
6.2.2	Coating Uniformity Requirements.....	199
6.3	Losses.....	203
6.4	Reflectors with Multiple Peaks	208
	References.....	212
7.	Edge Filters and Notch Filters	213
7.1	Thin-Film Absorption Filters	213
7.2	Interference Edge Filters	214
7.2.1	Quarter-Wave Stack	215
7.2.2	Symmetrical Multilayers and the Herpin Index	216
7.2.3	Application of the Herpin Index to the Quarter-Wave Stack	219
7.2.4	Application to Edge Filters	220
7.2.5	Application of the Herpin Index to Multilayers of Other than Quarter Waves.....	222
7.2.6	Reduction of Passband Ripple.....	224
7.2.7	Blocking	229

7.2.8	Extending the Transmission Zone	235
7.2.9	Reducing the Transmission Zone	241
7.2.10	Edge Steepness	242
7.3	Notch Filters	243
	References.....	249
8.	Bandpass Filters	251
8.1	Broad Bandpass Filters	251
8.2	Narrowband Single-Cavity Filters	254
8.2.1	Metal–Dielectric Single-Cavity Filter	254
8.2.2	All-Dielectric Single-Cavity Filter	260
8.2.3	Losses in Single-Cavity Filters.....	268
8.2.4	Solid-Etalon Filter	273
8.2.5	Effect of Varying the Angle of Incidence	276
8.2.6	Sideband Blocking.....	285
8.3	Multiple-Cavity Filters.....	286
8.3.1	Smith’s Method.....	286
8.3.2	Thelen’s Method	291
8.4	Higher Performance in Multiple-Cavity Filters	297
8.4.1	Effect of Tilting	304
8.4.2	Losses in Multiple-Cavity Filters	307
8.4.3	Further Information	310
8.5	Phase-Dispersion Filter	310
8.6	Multiple-Cavity Metal–Dielectric Filters	316
8.6.1	Induced Transmission Filter.....	317
8.6.2	Examples of Filter Designs.....	322
8.7	Measured Filter Performance	331
	References.....	334
9.	Tilted Dielectric Coatings	337
9.1	Introduction.....	337
9.2	Tilted Thicknesses	337
9.3	Modified Admittances	338
9.4	Polarizers and Analyzers	342
9.4.1	Plate Polarizer.....	343
9.4.2	Cube Polarizers.....	344
9.4.3	Brewster Angle Polarizing Beam Splitter	344
9.5	Nonpolarizing Coatings.....	350
9.5.1	Edge Filters at Intermediate Angle of Incidence	350
9.5.2	Reflecting Coatings at Very High Angles of Incidence	354
9.5.3	Edge Filters at Very High Angles of Incidence	356
9.6	Antireflection Coatings.....	358
9.6.1	<i>p</i> -Polarization Only	358
9.6.2	<i>s</i> -Polarization Only.....	359
9.6.3	<i>s</i> - and <i>p</i> -Polarizations Together	360
9.7	Retarders	363
9.7.1	Ellipsometric Parameters and Relative Retardation.....	363
9.7.2	Series of Coated Surfaces.....	364
9.7.3	Retarders.....	365

9.7.4	Simple Retarders.....	366
9.7.5	Multilayer Retarders at One Wavelength.....	369
9.7.6	Multilayer Retarders for a Range of Wavelengths.....	371
9.8	Optical Tunnel Filters.....	374
	References.....	377
10.	More on Tilted Coatings	379
10.1	Introduction.....	379
10.2	Modified Admittances and the Tilted Admittance Diagram	379
10.3	Application of the Admittance Diagram	382
10.4	Alternative Approach.....	394
10.5	Calculated Values for Silver.....	398
	References.....	401
11.	Other Topics: From Rugate Filters to Photonic Crystals	403
11.1	Rugate Filters.....	403
11.1.1	Apodization	406
11.1.2	Discrete Layer Replacements	406
11.1.3	Fourier Technique.....	413
11.2	Ultrafast Coatings	414
11.3	Glare Suppression Filters and Coatings.....	425
11.4	Some Coatings Involving Metal Layers.....	428
11.4.1	Electrode Films for Schottky Barrier Photodiodes.....	428
11.4.2	Spectrally Selective Coatings for Photothermal Solar Energy Conversion.....	430
11.4.3	Heat-Reflecting Metal–Dielectric Coatings	434
11.5	Gain in Optical Coatings.....	436
11.5.1	Oblique Incidence including Beyond Critical	440
11.5.2	Evanescient Gain	443
11.6	Perfect Absorbers	445
11.7	Photonic Crystals	449
11.7.1	What Is a Photonic Crystal?.....	449
11.7.2	Two-Dimensional Photonic Crystals.....	450
11.7.3	One-Dimensional Photonic Crystals.....	450
	References.....	454
12.	Color in Optical Coatings	457
12.1	Introduction.....	457
12.2	Color Definition	457
12.3	1964 Supplementary Colorimetric Observer	463
12.4	Metamerism.....	463
12.5	Other Color Spaces.....	464
12.6	Hue and Chroma	465
12.7	Brightness and Optimal Stimuli.....	466
12.8	Colored Fringes.....	467
	References.....	470
13.	Production Methods	471
13.1	Deposition of Thin Films	471
13.1.1	A Word about Pressure Units.....	472

13.1.2 Thermal Evaporation	472
13.1.3 Energetic Processes	480
13.1.4 Other Processes	490
13.2 Baking	493
13.3 A Word about Materials	494
13.4 Uniformity	494
13.4.1 Flat Plate	495
13.4.2 Spherical Surface	496
13.4.3 Rotating Substrates	497
13.4.4 Use of Masks	500
13.5 Substrate Preparation	501
13.6 Thickness Monitoring and Control	503
13.6.1 Optical Monitoring Techniques	504
13.6.2 Quartz Crystal Monitor	511
13.6.3 Monitoring by Deposition Time	513
13.7 Tolerances	514
13.8 Performance Envelopes	524
13.9 Reverse Engineering	527
References	529
14. Material Properties	535
14.1 Properties of Common Materials	535
14.2 Measurement of the Optical Properties	544
14.3 Pitfalls in Optical Constant Extraction	560
14.4 Measurement of the Mechanical Properties	564
14.5 Annealing	570
14.6 Toxicity	573
14.7 Microstructure and Thin-Film Behavior	574
References	586
15. Composite, Birefringent, and Metamaterials	593
15.1 Packing Density	593
15.2 Composite Material Models	594
15.3 Birefringent Materials	597
15.4 Metallic Grid Polarizers	598
15.5 Metamaterials	599
References	604
16. Some Coating Properties Important in Systems	607
16.1 Measurements and Calculations	607
16.2 Oblique Incidence and Polarization	608
16.2.1 Polarization Maintenance	610
16.2.2 Roof Prism Problems	610
16.2.3 Cone Response at Oblique Incidence	613
16.2.4 Cone Response of Thin-Film Polarizers	613
16.2.5 Small Spot Illumination	614
16.3 Surface Figure and Uniformity	618
16.4 Contamination Sensitivity	618
16.5 Scattering	623

16.6 Temperature Shifts.....	628
16.7 Stray Light	629
16.8 Laser Damage.....	630
References.....	632
17. Specification of Filters and Coatings	635
17.1 Optical Properties.....	635
17.1.1 Performance Specification.....	635
17.1.2 Manufacturing Specification	637
17.1.3 Test Specification	638
17.2 Physical Properties	640
17.2.1 Abrasion Resistance	640
17.2.2 Adhesion	642
17.2.3 Environmental Resistance.....	643
References.....	644
18. Characteristics of Thin-Film Dielectric Materials	645
References.....	652
Index	655

Series Preface

This international series covers all aspects of theoretical and applied optics and optoelectronics. Active since 1986, eminent authors have long been choosing to publish with this series, and it is now established as a premier forum for high-impact monographs and textbooks. The editors are proud of the breadth and depth showcased by published works, with levels ranging from advanced undergraduate and graduate student texts to professional references. Topics addressed are both cutting edge and fundamental, basic science and applications-oriented, on subject matter that includes lasers, photonic devices, nonlinear optics, interferometry, waves, crystals, optical materials, biomedical optics, optical tweezers, optical metrology, solid-state lighting, nanophotonics, and silicon photonics. Readers of the series are students, scientists, and engineers working in optics, optoelectronics, and related fields in the industry.

Proposals for new volumes in the series may be directed to Lu Han, senior publishing editor at CRC Press, Taylor & Francis Group (lu.han@taylorandfrancis.com).



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Preface to the Fifth Edition

Although I did eventually write this fifth edition, I never would have done it without the encouragement and support of many friends and colleagues. It was Lu Han of Taylor & Francis who eventually managed to push me over the edge—in the nicest possible way. And then there are four people, whose identities are hidden from me, but, since they are all clearly expert in thin-film optics, must certainly be people I know well. Their constructive comments based on my proposal and initial efforts were enormously helpful. I owe them and Lu a special debt of grateful thanks.

The field of thin-film optical coatings is still expanding and will continue to do so as long as there is optics. But the term *expansion* is really an oversimplification because there are substantial and significant changes also taking place. We are getting better at what we do is probably the best way to summarize them. There is a great deal of automation in our design, processes, and testing. Computers are everywhere. But automation does not decrease the need for understanding, particularly when there is a problem, and I have still tried to make understanding the primary thrust of this book.

The fundamentals have not changed, and so these are where we start. Then we have our well-known antireflection, high reflectance, bandpass, and related coatings. Although the subjects of these early chapters are still the same, their contents have been reordered and revised. The later chapters are almost entirely reorganized. Of course, the subject has not changed to the extent that everything needs to be completely rewritten, but I hope you will feel that the new arrangement is more logical. It has allowed me to expand on a good number of the important topics. There are also some older coatings that, although perhaps not often used today, are nevertheless instructive in their designs. Then more and more, we are seeing what can be considered as optical thin-film developments in other areas. I am thinking of metamaterials, coherent perfect absorbers, and the like. They tend to appear in journals other than the core ones for our community, and the technical language is often unfamiliar. Some are simply translations of what we already know well, but others could have important implications. I have tried to explain some of these in more familiar optical coating terms. These descriptions should not be thought of as in any way complete. They are simply intended to show how certain aspects might be understood in the context of optical coatings where they could eventually play some role. That there are more than 150 new figures may give some idea of the extent of the changes.

All the additional material has to be accommodated somehow and that implies that some material has to go. That has proved to be the most difficult task of all. Which old friends do I remove? I am helped a little in this by the fact that old books do not completely die but migrate to electronic versions, perhaps in the clouds. Previous editions are not lost forever. The topics I have eliminated will still exist, although they may be a little less readily accessible.

I continue to marvel at the incredibly good fortune that so long ago led me into this wonderfully welcoming, open, generous international optical thin-film community that has supported me all these years. Once again, I thank you, all of you, my readers, publishers, friends, colleagues, family, and especially my wife, Ann.

H. Angus Macleod
Tucson, Arizona



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Preface to the Fourth Edition

In some ways first editions are easier, or perhaps I should say less difficult, to prepare than subsequent editions. By the time a fourth edition is required, there is a strong expectation among readers of the character and content of the book. Thus, the author must somehow try to maintain the style at the same time as bringing the book up to date. What to omit and what to include are very difficult questions. Modern optical coating design is virtually entirely performed by computer, frequently using automatic techniques. However, computers do not remove the need for understanding, and I think it is understanding that readers look for in the book. Also, I am conscious that a reader, having perhaps rejected an earlier edition in favor of a later and remembering something important in the earlier, might well expect to find it in the later. I made the decision, therefore, to retain most of the descriptions of the earlier design techniques because of their importance in understanding how designs work. Then, although some of the applications that I describe are rather old, nevertheless they do illustrate how optical coatings are incorporated into a system, and so I retained them. I have tried to incorporate a reasonable amount of new material throughout the book. I added a chapter on color because it is increasing in importance in optical coatings, and, although it is of largely academic interest, I could not resist a section on the effects of gain in optical coatings, because I find it a fascinating topic. Then I struggled with coatings for the soft x-ray region and, with some regret, decided not to include them at this time. It is the old design synthesis problem: one has to stop somewhere.

I am fortunate in my friends and colleagues who have helped me immeasurably with suggestions, advice, and, I have to admit it, corrections. The field of optical thin films has been very good to me. I cannot imagine a more friendly, supportive, and open group of people than the international optical thin-film community. It sets an example the rest of the world would do well to follow.

Thank you, all of you, my readers, publishers, friends, colleagues, family, and especially my wife, Ann.

H. Angus Macleod



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Preface to the Third Edition

The foreword to the second edition of this book identified increasing computer power and availability as especially significant influences in optical coating design. This has continued to the point where any description I might give of current computing speed and capacity would be completely out of date by the time this work is in print. Software for coating design (and for other tasks) is now so advanced that commercial packages have almost completely replaced individually written programs. I have often heard it suggested that this removes all need for skill or even knowledge from the act of coating design. I firmly believe that the need for skill and understanding is actually increased by the availability of such powerful tools. The designer who knows very well what he or she is doing is always able to achieve better results than the individual who does not. Coating design still contains compromises. Some aspects of performance are impossible to attain. The results offered by an automatic process that is attempting to reach impossible goals are usually substantially poorer than those when the goals are realistic. The aim of the book, therefore, is still to improve understanding.

During the years since publication of the second edition, the energetic processes, and particularly ion-assisted deposition, have been widely adopted. There are several consequences. The improved stability of optical constants of the materials has enabled the reliable production of coatings of continuously increasing complexity. We even see coatings produced now purely for their aesthetic appeal. Then the enormous improvement in environmental stability has opened up new applications, especially in communications. Unprecedented temperature stability of optical coatings can now be achieved. Specially designed coatings have simplified the construction of ultrafast lasers. Banknotes of many countries inhibit counterfeiting by carrying patches exhibiting the typical iridescence of optical coatings. Coatings to inhibit the effects of glare are now integral parts of visual display units.

I mentioned in my previous foreword the difficulty I experienced in bringing the earlier edition up to date. This time the task has been even more difficult. The volume of literature has expanded to the extent that it is almost impossible to keep up with all of it. The pressure on workers to publish has in many cases reached almost intolerable levels. I regret I do not remember exactly who introduced the idea of the half-life of a publication after which it sinks into obscurity but it is clear that the half-life has become quite short. Comprehensively to review this vast volume of material that has appeared and continues to appear would have changed completely the style of the book. The continuing demand for the now out-of-print second edition of the book suggests that it is used much more as a learning tool than a research reference and so my aim has been to try to keep it so. There have been few fundamental changes that affect our basic understanding of optical coatings and so this third edition reflects that.

I appreciate very much the help of various organizations and individuals who provided material. Many are named in the foreword to the second edition and in the apology to the first. Additional names include Shincron Company Ltd, Ion-Tech Inc, Applied Vision Ltd, Professor Frank Placido of the University of Paisley, and Roger Hunneman of the Department of Cybernetics, University of Reading.

Again I am grateful for all the helpful comments and suggestions from all my friends and colleagues. The enormous list of names is beyond what can be reproduced here but I must mention my debt to my old friend Professor Lee Cheng-Chung, who took the trouble to work completely through the book and provided me with what has to be the most detailed list of misprints and

mistakes, and Professor Shigetaro Ogura, who was instrumental in the translation of the second edition into Japanese. The people at Adam Hilger must be the most patient people on earth. I think finally it was my shame at so trying the endurance of Kathryn Cantley who simply responded with encouragement and understanding that drove me to complete the work.

My eternal and grateful thanks to my wife. She did not write the book but she made sure that I did.

H. Angus Macleod
Tucson, Arizona

Preface to the Second Edition

A great deal has happened in the subject of optical coatings since the first edition of this book. This is especially true of facilities for thin-film calculations. In 1969 my thin-film computing was performed on an IBM 1130 computer that had a random access memory of 10 kbytes. Time had to be booked in advance, sometimes days in advance. Calculations remote from this computer were performed either by slide rule, log tables or electromechanical calculator. Nowadays my students scarcely know what a slide rule is, my pocket calculator accommodates programs that can calculate the properties of thin-film multilayers, and I have on my desk a microcomputer with a random access memory of 0.5 Mbytes, which I can use as and when I like. The earlier parts of this revision were written on a mechanical typewriter. The final parts were completed on my own word processor. These advances in data processing and computing are without precedent and, of course, have had a profound and irreversible effect on many aspects of everyday life as well as on the whole field of science and technology.

There have been major developments, too, in the deposition of thin-film coatings, and although these lack the spectacular, almost explosive, character of computing programs, nevertheless important and significant advances have been made. Electron-beam sources have become the norm rather than the exception, with performance and reliability beyond anything available in 1969. Pumping systems are enormously improved, and the box-coater is now standard rather than unusual. Microprocessors control the entire operation of the pumping system and, frequently, even the deposition process. We have come to understand that many of our problems are inherent in the properties of our thin films rather than in the complexity of our designs. Microstructure and its influence on material properties is especially important. Ultimate coating performance is determined by the losses and instabilities of our films rather than the accuracy and precision of our monitoring systems.

My own circumstances have changed too. I wrote the first edition in industry. I finish the second as a university professor in a different country.

All this change has presented me with difficult problems in the revision of this book. I want to bring it up to date but do not want to lose what was useful in the first edition. I believe that in spite of the great advances in computers, there is still an important place for the appreciation of the fundamentals of thin-film coating design. Powerful synthesis and refinement techniques are available and are enormously useful, but an understanding of thin-film coating performance and the important design parameters is still an essential ingredient of success. The computer frees us from much of the previous drudgery and puts in our hands more powerful tools for improving our understanding. The availability of programmable calculators and of microcomputers implies easy handling of more complex expressions and formulae in design and performance calculations. The book, therefore, contains many more of these than did the first edition. I hope they are found useful. I have included a great deal of detail on the admittance diagram and admittance loci. I use them in my teaching and research and have taken this opportunity to write them up. SI units, rather than Gaussian, have been adopted, and I think Chapter 2 is much the better for the change. There is more on coatings for oblique incidence including the admittance diagram beyond the critical angle, which explains and predicts many of the resonant effects that are observed in connection with surface plasmons, effects used by Greenland and Billington (Chapter 8, reference 12) in the late 1940s and early 1950s for monitoring thin-film deposition.

Inevitably, the first edition contained a number of mistakes and misprints and I apologise for them. Many were picked up by friends and colleagues who kindly pointed them out to me. Perhaps the worse mistake was in Figure 9.4 on uniformity. The results were quoted as for a flat plate but, in fact, referred to a spherical work holder. These errors have been corrected in this edition

and I hope that I have avoided making too many fresh ones. I am immensely grateful to all the people who helped in this correction process. I hope they will forgive me for not including the huge list of their names here. My thanks are also due to J. J. Apfel, G. DeBell, E. Pelletier, and W. T. Welford, who read and commented on various parts of the manuscript.

To the list in the foreword of the first edition of organisations kindly providing material should be added the names Leybold-Heraeus GmbH, and Optical Coating Laboratory Inc. Aircot-Temescal is now known as Temescal, a Division of the BOC Group Inc., and the British Scientific Instrument Research Association as Sira Institute.

My publisher is still the same Adam Hilger, but now part of the Institute of Physics. I owe a very great debt to Neville Goodman, who was responsible for the first edition and who also persuaded and encouraged me into the second. He retired while it was still in preparation, and the task of extracting the final manuscript from me became Jim Reville's. Ian Kingston and Brian McMahon did a tremendous job on the manuscript at a distance of 3000 miles. Their patience with me in the delays I have caused them has been amazing.

My wife and family have once again been a great source of support and encouragement.

H. Angus Macleod

Newcastle upon Tyne and Tucson, Arizona

Apologia to the First Edition

When I first became involved with the manufacture of thin-film optical filters, I was particularly fortunate to be closely associated with Oliver Heavens, who gave me invaluable help and guidance. Although I had not at that time met him, Dr. L. Holland also helped me through his book, *The Vacuum Deposition of Thin Films*. Lacking, however, was a book devoted to the design and production of multilayer thin-film optical filters, a lack which I have since felt especially when introducing others to the field. Like many others in similar situations I produced from time to time notes on the subject purely for my own use. Then in 1967, I met Neville Goodman of Adam Hilger, who had apparently long been hoping for a book on optical filters in general. I was certainly not competent to write a book on this wide subject, but, in the course of conversation, the possibility of a book solely on thin-film optical filters arose. Neville Goodman's enthusiasm was infectious, and with his considerable encouragement, I dug out my notes and began writing. This, some two years and much labour later, is the result. I have tried to make it the book that I would like to have had myself when I first started in the field, and I hope it may help to satisfy also the needs of others. It is not in any way intended to compete with the existing works on optical thin films, but rather to supplement them, by dealing with one aspect of the subject which seems to be only lightly covered elsewhere.

It will be immediately obvious to even the most casual of readers that a very large proportion of the book is a review of the work of others. I have tried to acknowledge this fully throughout the text. Many of the results have been recast to fit in with the unified approach which I have attempted to adopt throughout the book. Some of the work is, I fondly imagine, completely my own, but at least a proportion of it may, unknown to me, have been anticipated elsewhere. To any authors concerned I humbly apologise, my only excuse being that I also thought of it. I promise, as far as I can, to correct the situation if ever there is a second edition. I can, however, say with complete confidence that any shortcomings of the book are entirely my own work.

Even the mere writing of the book would have been impossible without the willing help, so freely given, of a large number of friends and colleagues. Neville Goodman started the whole thing off and has always been ready with just the right sort of encouragement. David Tomlinson, also of Adam Hilger, edited the work and adjusted it where necessary so that all sounded just as I had meant it to, but had not quite managed to achieve. The drawings were the work of Mrs Jacobi. At Grubb Parsons, Jim Mills performed all the calculations, using an IBM 1130 (he appears in the frontispiece for which I am also grateful), Fred Ritchie kindly gave me permission to quote many of his results and helped considerably by reading the manuscript, and Helen Davis transformed my almost illegible first manuscript into one which could be read without considerable strain. Stimulating discussions with John Little and other colleagues over the years have also been invaluable. Desmond Smith of Reading University kindly gave me much material especially connected with the section on atmospheric temperature sounding which he was good enough to read and correct. John Seeley and Alan Thetford, both of Reading University, helped me by amplifying and explaining their methods of design. Jim Ring, of Imperial College, read and commented on the section on astronomical applications and Dr. J. Meaburn kindly provided the photographs for it. Dr. A. F. Turner gave me much information on the early history of multiple halfwave filters. It is impossible to mention by name all those others who have helped but they include: M. J. Shadbolt, S. W. Warren, A. J. N. Hope, H. Bucher, and all the authors who led the way and whose work I have used and quoted.

Journals, publishers, and organisations which provided and gave permission for the reproduction of material were

- *Journal of the Optical Society of America* (The Optical Society of America)
- *Applied Optics* (The Optical Society of America)
- *Optica Acta* (Taylor and Francis Limited)
- *Proceedings of the Physical Society* (The Institute of Physics and the Physical Society)
- *IEEE Transactions on Aerospace* (The Institute of Electrical and Electronics Engineers, Inc.)
- *Zeitschrift für Physik* (Springer-Verlag)
- *Bell System Technical Journal* (The American Telephone and Telegraph Co.)
- *Philips Engineering Technical Journal* (Philips Research Laboratories)
- Methuen & Co. Ltd
- OCLI Optical Coatings Limited
- Standard Telephones and Cables Limited
- Balzers Aktiengesellschaft für Hochvacuumtechnik und dünne Schichten
- Edwards High Vacuum Limited
- Aircro Temescal (A Division of Air Reduction Company Inc.)
- Hawker Siddeley Dynamics Limited
- System Computers Limited
- Ferranti Limited
- British Scientific Instrument Research Association
- And lastly, but far from least, the management of Sir Howard Grubb, Parsons & Co. Ltd, particularly Mr G M Sisson and Mr G E Manville, for much material, for facilities, and for permission to write this book.

To all these and to all the others, who are too numerous to name and who I hope will excuse me for not attempting to name them, I am truly grateful.

I should add that my wife and children have been particularly patient with me during the long writing process, which has taken up so much of the time that would normally have been theirs. Indeed, my children eventually began to worry if ever I appeared to be slacking and, by their comments, prodded me into redoubled efforts.

H. Angus Macleod
Newcastle upon Tyne

Author

H. Angus Macleod is president of Thin Film Center Inc., in Tucson, Arizona, and professor emeritus of Optical Sciences at the University of Arizona, Tucson, Arizona. Dr. Macleod is a graduate of Glasgow University. He holds the degree of Doctor of Technology from the Council for National Academic Awards (London) and an honorary doctorate from the University of Aix-Marseille in France. He is the author of over 200 academic publications and has taught courses on optical topics all over the world, with class audiences ranging from 1 person to over 200 people. He specializes in teaching techniques for understanding and logical thinking that avoid complicated theory without oversimplification.

Dr. Macleod is the recipient of numerous professional honors, including the Gold Medal (1987) from the International Society for Optics and Photonics, the Esther Hoffman Beller Medal (1997) from the Optical Society of America, the Nathaniel H. Sugerman Memorial Award (2002) from the Society of Vacuum Coaters, and the Senator Award (2008) from the European Vacuum Coaters.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Symbols and Abbreviations

The following list gives those more important symbols used in at least several places in the text. We have tried as far as possible to create a consistent set of symbols, but there are several well-known and accepted symbols that are universally used in the field for certain quantities, and changing them would probably lead to even greater confusion than would retaining them. This has meant that in some cases, the same symbol is used in different places for different quantities. We hope the table will make it clear. Less important symbols, defined and used only in very short sections, have been omitted.

<i>A</i>	Absorptance. The ratio of the power absorbed in the structure to the power incident on it.
<i>A</i>	Potential absorptance. A quantity used in the calculation of the absorptance of coatings. It is equivalent to $(1 - \psi)$, where ψ is the potential transmittance.
<i>B</i>	The normalized total tangential electric field at an interface, usually the front interface of an assembly of layers. It is also very briefly used at the beginning of Chapter 2 as the magnetic induction.
<i>C</i>	The normalized total tangential magnetic field at an interface, usually the front interface of an assembly of layers.
<i>d_q</i>	The physical thickness of the <i>q</i> th layer in a thin film coating.
<i>E</i>	The electric vector in the electromagnetic field.
\mathcal{E}	The total tangential electric field amplitude, that is, the field parallel to the thin film boundaries.
\mathcal{E}	The electric field amplitude.
$\bar{\mathcal{E}}$	The equivalent admittance of a symmetrical arrangement of layers.
<i>F</i>	A function used in the theory of the Fabry–Perot interferometer.
\mathcal{F}	Finesse. The ratio of the separation of adjacent fringes to the halfwidth of the fringe in the Fabry–Perot interferometer.
<i>g</i>	$g = \lambda_0/\lambda = v/v_0$, sometimes called the relative wavelength, or the relative wave-number, or the wavelength ratio. λ_0 and v_0 are the reference wavelength and reference wavenumber, respectively. The optical thicknesses of the layers in a coating are defined with respect to these quantities that are usually chosen to make the more important layers in the coating as close to quarter waves as possible.
<i>H</i>	The magnetic vector in the electromagnetic field.
\mathcal{H}	The total tangential magnetic field amplitude, that is, the field parallel to the thin film boundaries.
\mathcal{H}	The magnetic field amplitude.
<i>H</i>	Represents a quarter wave of high index in shorthand notation.
<i>I</i>	The irradiance of the wave, that is, power per unit area. Unfortunately, the standard International System of Units (SI) symbol for irradiance is <i>E</i> , but to use <i>E</i> would cause great confusion between irradiance and electric field. It is even more unfortunate that <i>I</i> is the SI symbol for intensity that is the power per unit solid angle from a point source. Doubly unfortunate is that the older definition of intensity is identical to the current definition of irradiance.
<i>k</i>	The extinction coefficient. The extinction coefficient denotes the presence of absorption. The complex refractive index <i>N</i> is given by $N = n - ik$.
<i>L</i>	Represents a quarter wave of low index in shorthand notation.

<i>M</i>	Represents a quarter wave of intermediate index in shorthand notation. Also used for a matrix element or to indicate an array of matrix elements.
<i>N</i>	Denotes the complex refractive index $n - ik$.
<i>n</i>	The refractive index or, sometimes, the real part of refractive index.
<i>n*</i>	The effective index of a narrowband filter, that is, the index of an equivalent layer that yields a shift of its fringes in wavelength, by the same amount as the peak of the narrowband filter, when tilted with respect to the direction of incidence.
<i>p</i>	Packing density, that is, the ratio of the solid volume of a film to its total volume.
<i>p</i>	<i>p</i> -Polarization, that is, the polarization where the electric field direction is in the plane of incidence. It is sometimes known as TM for transverse magnetic.
<i>R</i>	The reflectance. The ratio at a boundary of the normal components of reflected and incident irradiance or, alternatively, the ratio of the total reflected beam power to the total incident beam power.
<i>s</i>	<i>s</i> -Polarization, that is, the polarization where the electric field direction is normal to the plane of incidence. It is sometimes known as TE for transverse electric.
<i>T</i>	The transmittance. The ratio of the normal components of transmitted and incident irradiance or, alternatively, the ratio of the total transmitted beam power to the total incident beam power.
TE	See <i>s</i> for <i>s</i> -polarization.
TM	See <i>p</i> for <i>p</i> -polarization.
<i>x, y, z</i>	The coordinate axes. In the case of a thin film or surface, the <i>z</i> -axis is usually taken positive into the surface in the direction of incidence. The <i>x</i> -axis is usually arranged in the plane of incidence, and the <i>x</i> -, <i>y</i> -, and <i>z</i> -axes, in that order, make a right-handed set.
$\bar{x}, \bar{y}, \bar{z}$	The three color matching functions that define the CIE 1931 Standard Colorimetric Observer.
<i>X, Y, Z</i>	The tristimulus values. They are the three basic responses defining a color.
<i>x, y, z</i>	The chromaticity coordinates, $X/(X+Y+Z)$, $Y/(X+Y+Z)$, and $Z/(X+Y+Z)$. Usually, <i>z</i> is omitted because they are normalized to add to unity.
<i>X + iZ</i>	The complex surface admittance.
<i>y</i>	The characteristic admittance of a material given in SI units (siemens) by $N\eta$, that is, $(n - ik)\eta$ and in units of the admittance of free space η by N or $n - ik$.
<i>Y</i>	The surface admittance, that is, the ratio of the total tangential components of magnetic and electric field at any surface parallel to the film boundaries. $Y = C/B$.
η	The admittance of free space (2.6544×10^{-3} S).
<i>y₀</i>	The characteristic admittance of the incident medium.
<i>y_m</i> or <i>y_{sub}</i>	The characteristic admittance of the emergent medium or substrate.
α	The absorption coefficient, given by $4\pi k/\lambda$, usually in units of cm^{-1} .
α, β, γ	The three direction cosines, that is, the cosines of the angle the direction makes with the three coordinate axes.
β	Symbol for $2\pi kd/\lambda$, usually with reference to a metal.
γ	The equivalent phase thickness of a symmetrical arrangement of layers.
Δ	The relative retardation. It is given by $\phi_p - \phi_s \pm 180^\circ$ in reflection and $\phi_p - \phi_s$ in transmission, where the normal thin-film sign convention for ϕ_p is used.
Δ	η_p/η_s , where η is the modified tilted admittance. The quantity is used in the design of polarization-free coatings.
δ	The phase thickness of a coating, given by $2\pi(n - ik)d/\lambda$.
ϵ	Indicates a small error in the discussion of tolerances etc.
ϵ	The permittivity of a medium.
η	The tilted optical admittance.
ϑ	The angle of incidence.

κ	Sometimes called the wavenumber, κ is given by $2\pi(n - ik)/\lambda$, where λ is the free space wavelength. Note the confusing use of the term <i>wavenumber</i> . It is also applied to ν .
λ	The wavelength of light. In the book, except at the very beginning of Chapter 2, it always indicates the wavelength in free space.
λ_0	The reference wavelength. The optical thicknesses of the layers in a coating are defined with respect to the reference wavelength that is usually chosen to make the more important layers in the coating as close to quarter waves as possible.
ν	The wavenumber. $\nu = 1/\lambda$ and is frequently expressed in units of inverse centimetres (also sometimes known as kayser. The SI unit is strictly inverse meters or m^{-1}).
ν_0	The reference wavenumber, $1/\lambda_0$.
μ	Permeability. Used in the early part of Chapter 2.
ρ	The amplitude reflection coefficient; also used as electric charge density in the early part of Chapter 2.
τ	The amplitude transmission coefficient.
ϕ	Phase difference, often in reflection or transmission.
ψ	The potential transmittance $T/(1 - R)$ or the ratio of the quantities $\text{Re}(BC^*)$, evaluated at two different interfaces. It represents the net power emerging from a system divided by the net power entering and is unity if there is no loss.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

1

Introduction

When this book was first written, the question of the title arose. *Optical Thin Films* was the obvious choice, but the publisher feared that it might then be confused with some other existing titles, and so, eventually to avoid confusion, *Thin-Film Optical Filters* was chosen and has remained the name through four editions and, now, into a fifth. It was never intended that the subject should be limited to the narrow designation of filters but should encompass as much of the field of optical thin films as possible. This is still the intention, but in a work of this size, it is not possible to cover the entire field of thin-film optical devices in the detail that some of them may deserve. The selection of topics is due, at least in part, to the author's own preferences and knowledge.

The intention of the book has always been to form an introduction to thin-film optical coatings for both the manufacturer and the user. The topics covered are a mixture of design, manufacture, performance, and features important in applications. It begins with enough of the basic mathematics of optical thin films for the reader to carry out thin-film calculations. The aim has been to present, as far as possible, a unified treatment, and there are some alternative methods of analysis that are not discussed.

When the book first appeared, there were just a few books available that covered aspects of the field. Now the situation has changed somewhat, and there is an array of relevant books. Some of these are listed in the bibliography at the end of this chapter. However, the half-life of a work these days is so short that knowledge can actually disappear. It is well worthwhile to take the time to go back to some of the earlier books. Heavens [1], Holland [2], Anders [3], and Knittl [4] are just some of those that will repay study, and they are listed in the bibliography along with some more recent volumes.

1.1 Early History

History is impossibly complicated, and we can have only an imperfect view of it, told generally through the medium of an historian who will, because of culture changes, attempt to interpret it in a way the intended audience can understand. The history of technology is no exception. To simplify the telling, we will usually pick certain events and individuals and connect them as a kind of series, one depending on the other. Technology, however, develops over a very broad front, rather like the advance of the tide. It depends on a network of effort. If one individual does not make a required advance, another certainly will. Technology adopts the advances it needs at the time and ignores those that it does not. Over and over again, we find that discoveries credited to a particular individual were actually anticipated by others, but the time was not right, and so little or no notice was taken of them. The abbreviated account of the history of the subject that follows is no exception. A true account is beyond us, and so we pick a few events and a few individuals and connect them, but there are many other routes through history. This one is based on the preferences and limited knowledge of your author.

Thin metal layers were known from very early times, but if we consider interference as the hallmark of modern thin-film optics, then the earliest of what might be called modern thin-film optics was the work of Robert Boyle and Robert Hooke on colors exhibited by materials in thin-film form. Sir Isaac Newton [5] related the colors to exact measurements of film thickness and placed the subject on a firm quantitative base with his brilliant technique now known as Newton's rings. The explanation of the colors is nowadays thought to be a very simple matter, being due to interference in a single thin film of varying thickness. However, at that time, the theory of the nature of light was not sufficiently far advanced. Newton struggled with his concept of the interval of fits that as we now understand is a half wavelength, but it was a further 100 years before the idea of light as a wave would be accepted. On November 12, 1801, in a Bakerian Lecture to the Royal Society, Thomas Young [6] enunciated the principle of the interference of light and produced a satisfactory explanation of the effect. As Henry Crew [7] has put it, "This simple but tremendously important fact that two rays of light incident upon a single point can be added together to produce darkness at that point is, as I see it, the one outstanding discovery which the world owes to Thomas Young."

Young's theory was far from achieving universal acceptance. Indeed Young became the victim of a bitter personal attack, against which he had the greatest difficulty defending himself. Recognition came slowly and depended much on the work of Augustin Jean Fresnel [8], who, quite independently, also arrived at a wave theory of light. Fresnel's discovery, in 1816, that two beams of light that are polarized at right angles could never interfere established the transverse nature of light waves. Then Fresnel combined Young's interference principle and Huygens's ideas of light propagation into an elegant theory of diffraction. It was Fresnel who put the wave theory of light on such a firm foundation that it has never been shaken. For the thin-film worker, Fresnel's laws, governing the amplitude and phase of light reflected and transmitted at a single boundary, are of major importance. Knittl [9] reminded us that Fresnel already knew that the sum of an infinite series of rays is necessary to determine the transmittance of a thick sheet of glass and that it was Simeon Denis Poisson, in correspondence with Fresnel, who included interference effects in the summation to arrive at the important results that a half-wave-thick film does not change the reflectance of a surface and that a quarter-wave-thick film of index $(n_0 n_1)^{1/2}$ will reduce to zero the reflectance of a surface between two media of indices n_1 and n_0 , results extended to oblique incidence by Fresnel. Fresnel died in 1827, at the early age of 39.

In 1873, the great work of James Clerk Maxwell, *A Treatise on Electricity and Magnetism* [10], was published, and in his system of equations, we have all the basic theory for the analysis of thin-film optical problems.

Meanwhile, in 1817, Joseph Fraunhofer [11] made what are probably the first ever antireflection coatings. It is worth quoting his observations at some length because they show the considerable insight that he had, even at that early date, into the physical causes of the effects that were produced. The following is a translation of part of the paper as it appears in the collected works:

Before I quote the experiments which I have made on this I will give the method which I have made use of to tell in a short time whether the glass will withstand the influence of the atmosphere. If one grinds and then polishes, as finely as possible, one surface of glass which has become etched through long exposure to the atmosphere, then wets one part of the surface, for example half, with concentrated sulfuric or nitric acid and lets it work on the surface for twenty-four hours, one finds after cleaning away the acid that that part of the surface on which the acid was, reflects much less light than the other half, that is it shines less although it is not in the least etched and still transmits as much light as the other half, so that one can detect no difference on looking through. The difference in the amount of reflected light will be most easily detected if one lets the light strike approximately vertically. It is the greater the more the glass is liable to tarnish and become etched. If the polish on the glass is not very good, this difference will be less noticeable. On glass which is not liable to tarnish, the sulfuric and nitric acid does not work. . . . Through this treatment with sulfuric or nitric acid some types of glasses get on their surfaces beautiful vivid colors which alter like soap bubbles if one lets the light strike at different angles.

Then, in an appendix to the paper added in 1819:

Colors on reflection always occur with all transparent media if they are very thin. If for example, one spreads polished glass thinly with alcohol and lets it gradually evaporate, towards the end of the evaporation, colors appear as with tarnished glass. If one spreads a solution of gum-lac in a comparatively large quantity of alcohol very thinly over polished warmed metal the alcohol will very quickly evaporate, and the gum-lac remains behind as a transparent hard varnish which shows colors if it is thinly enough laid on. Since the colors, in glasses which have been colored through tarnishing, alter themselves if the inclination of the incident light becomes greater or smaller, there is no doubt that these colors are quite of the same nature as those of soap bubbles, and those which occur through the contact of two polished flat glass surfaces, or generally as thin transparent flakes of material. Thus, there must be on the surface of tarnished glass that shows colors, a thin layer of glass that is different in refractive power from the underlying. Such a situation must occur if a component is partly removed from the surface of the glass or if a component of the glass combines at the surface with a related material into a new transparent product.

It seems that Fraunhofer did not follow up this particular line into the development of an antireflection coating for glass, perhaps because optical components were not, at that time, sufficiently complicated for the need for antireflection coatings to be obvious. Possibly the important point that not only was the reflectance less but the transmittance also greater had escaped him.

In 1886, Lord Rayleigh [12] reported to the Royal Society an experimental verification of Fresnel's reflection law at off-normal incidence. In order to attain a sufficiently satisfactory agreement between measurement and prediction, he had found it necessary to use freshly polished glass because the reflectance of older material, even without any visible signs of tarnish, was too low. One possible explanation, which he suggested, was the formation, on the surface, of a thin layer of different refractive indices from the underlying material. He was apparently unaware of the earlier work of Fraunhofer and of the identical difficulties experienced by Malus and by Brewster in determining the polarizing angle, or Brewster angle, of glass.

Then, in 1891, Dennis Taylor [13,14] published the first edition of his famous book *On the Adjustment and Testing of Telescopic Objectives* and mentioned that, "As regards the tarnish which we have above alluded to as being noticeable upon the flint lens of an ordinary objective after a few years of use, we are very glad to be able to reassure the owner of such a flint that this film of tarnish, generally looked upon with suspicion, is really a very good friend to the observer, inasmuch as it increases the transparency of his objective."

In fact, Taylor [15] went on to develop a method of artificially producing the tarnish by chemical etching. This work was followed up by Kollmorgen [16], who developed the chemical process still further for different types of glasses.

At the same time, in the nineteenth century, a great deal of progress was being made in the field of interferometry. The most significant development, from the thin-film point of view, was the Fabry-Perot [17] interferometer described in 1899, which has become one of the basic structures for thin-film filters.

Developments became much more rapid in the 1930s, and indeed, it is in this period that we can recognize the beginnings of modern thin-film optical coating. In 1932, Rouard [18] observed that a very thin metallic film reduced the internal reflectance of a glass plate, although the external reflectance was increased. In 1934, Bauer [19], in the course of fundamental investigations of the optical properties of halides, produced reflection-reducing coatings, and Pfund [20] evaporated zinc sulfide layers to make low-loss beam splitters for Michelson interferometers, noting, incidentally, that titanium dioxide could be a better material. In 1936, John Strong [21] produced antireflection coatings by the evaporation of fluorite to give inhomogeneous films, which reduced the reflectance of glass to visible light by as much as 89%, a most impressive figure. At the same time, Alexander Smakula at the Carl Zeiss company, in Jena, developed antireflection coatings that were kept secret because of their military implications. Then, in 1939, Geffcken [22] constructed the first thin-film metal-dielectric interference filters. A fascinating account of Geffcken's work is given by Thelen [23], who described Geffcken's search for improved antireflection coatings and his creation of the famous quarter-half-quarter design.

Several factors were probably responsible for this sudden expansion of the field. Optical systems, particularly photographic objectives were becoming more complex, bringing a need for antireflection coatings. Telescopes and binoculars, especially for military applications, were also much improved by antireflection coatings. Then the manufacturing process was also becoming more reliable. Although sputtering was discovered about the middle of the nineteenth century, and vacuum evaporation around the beginning of the twentieth century, they had not yet been adopted as useful manufacturing processes. One difficulty was the lack of really suitable pumps, and it was not until the early 1930s that the work of C. R. Burch on diffusion pump oils introduced the oil diffusion pump. This enormously helped, although, particularly in Germany, mercury diffusion pumps were still used very effectively for some time. World War II saw a great expansion in the production of antireflection coatings. This certainly accelerated developments, but the expansion would have taken place without any war, because optics had now reached the stage where coatings were necessary. Since then, tremendous strides have been made. Modern optics without coating is unthinkable. It is almost impossible to imagine an optical instrument that would not rely on optical coatings to assure its performance. Filters with greater than 100 layers are not uncommon, and uses have been found for them in almost every branch of science and technology.

1.2 Thin-Film Filters

First of all, we assume for the purposes of this section that the materials in thin-film form are free from absorption or other loss. Then to understand in a qualitative way the performance of thin-film optical devices, it is necessary to accept several simple statements. The first is that the amplitude reflectance of light at any boundary between two media is given by $(1 - \rho)/(1 + \rho)$, where ρ is the ratio of the optical admittances at the boundary, which, in the optical region, is also the ratio of the refractive indices. The reflectance (the ratio of irradiances) is the square of this quantity. The second is that there is a phase shift of 180° when the reflectance takes place in a medium of lower refractive index than the adjoining medium and zero if the medium has a higher index than the one adjoining it. The third is that if light is split into two components by reflection at the top and bottom surfaces of a thin film, then the beams will recombine in such a way that the resultant amplitude will be the difference of the amplitudes of the two components, if the relative phase shift is 180° , or the sum of the amplitudes, if the relative phase shift is either zero or a multiple of 360° . In the former case, we say that the beams interfere destructively, and in the latter, constructively. Other cases where the phase shift is different will be intermediate between these two possibilities.

The antireflection coating depends on the more or less complete cancellation of the light reflected at the upper and lower of the two surfaces of the thin film for its operation (Figure 1.1). Let the index of the substrate be n_m ; that of the film, n_1 ; and that of the incident medium, which will in almost all cases be air, n_0 . For a completely accurate calculation, we should consider multiple beams, as in the subsequent chapters of this book, but for the moment, we adopt an approximation. We assume that although the reflection at the front surface has diminished the transmitted light a little, we shall completely neglect that loss. Then for complete cancellation of the two beams of light, they should be 180° out of phase and their amplitudes should be equal, which implies that the ratios of the refractive indices at each boundary should be equal, i.e., $n_0/n_1 = n_1/n_m$, or $n_1 = (n_0 n_m)^{1/2}$. This shows that the index of the thin film should be intermediate between the indices of air, which may be taken as unity, and of the substrate, which may be taken as at least 1.52. At both the upper and lower boundaries of the antireflection film, the reflection takes place in a medium of lower refractive index than the adjoining medium. Thus, to ensure that

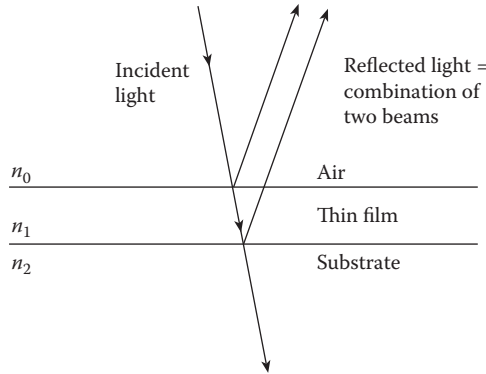


FIGURE 1.1
Single thin film.

the relative phase shift is 180° so that the beams cancel, the optical thickness of the film should be made one-quarter wavelength.

A simple antireflection coating should, therefore, consist of a single film of refractive index equal to the square root of that of the substrate and of optical thickness one-quarter of a wavelength. As will be explained in the chapter on antireflection coatings (Chapter 4), there are other improved coatings covering wider wavelength ranges involving greater numbers of layers.

Another basic type of thin-film structure is a stack of alternate high- and low-index films, all one-quarter wavelength thick (see Figure 1.2). Light reflected within the high-index layers will not suffer any phase shift on reflection, while that reflected within the low-index layers will suffer a change of 180° . It is fairly easy to see that the various components of the incident light produced by reflection at successive boundaries throughout the assembly will reappear at the front surface all in phase so that they will constructively recombine. This implies that the effective reflectance of the assembly can be made very high indeed, as high as may be desired, merely by increasing the number of layers. This is the basic form of the high-reflectance coating. When such a coating is constructed, it is found that the reflectance remains high over only a limited range of wavelengths, depending on the ratio of high and low refractive indices. Outside this zone, the reflectance abruptly changes to a low value. Because of this behavior, the quarter-wave stack, as it is called, is used as a basic building block for many types of thin-film filters. It can be used as a longwave-pass

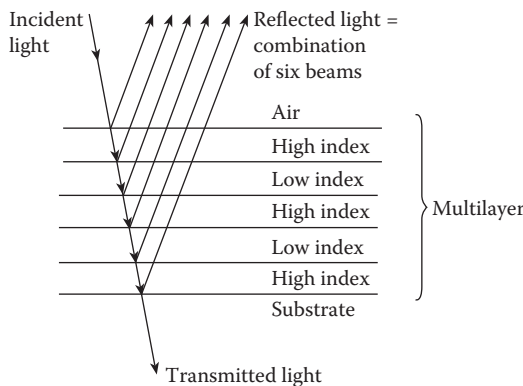


FIGURE 1.2
Multilayer.

filter; a shortwave-pass filter; a bandstop filter; a straightforward high-reflectance coating, for example, in laser mirrors; and a reflector in a thin-film Fabry–Perot interferometer (Figure 1.3), which is another basic filter type described in some detail in Chapters 6 and 8. Here, it is sufficient to say that it consists of a cavity layer, sometimes called a spacer layer, that is usually half a wavelength thick, bounded by two high-reflectance coatings. Multiple-beam interference in the cavity layer causes the transmission of the filter to be extremely high over a narrow band of wavelengths around that for which the cavity is a multiple of one-half wavelength thick. It is possible, as with lumped electric circuits, to couple two or more Fabry–Perot filters in series to give a more rectangular shape to the pass band.

Our assumption of vanishingly small absorption and other losses so that the films are completely transparent is true in the great majority of cases. Since no energy is lost, the filter characteristic in reflection is the complement of that in transmission. This fact is used in the construction of such devices as dichroic beam splitters for color separation in, for example, color projection engines.

This brief description has neglected the effect of multiple reflections in most of the layers, and for an accurate evaluation of the performance of a filter, these extra reflections must be taken into account. This involves extremely complex calculations, and an alternative, and more effective, approach has been found in the development of entirely new forms of solution of Maxwell's equations in stratified media. This is, in fact, the principal method used in Chapter 2 where basic theory is considered. The solution appears as a very elegant product of 2×2 matrices, each matrix representing a single film. Unfortunately, in spite of the apparent simplicity of the matrices, the calculation by hand of the properties of a given multilayer, particularly if there are absorbing layers present and a wide spectral region is involved, is an extremely tedious and time-consuming task. The preferred method of calculation is to use a computer. This makes calculation so rapid and straightforward that it makes little sense to use anything else. Even pocket calculators, especially the programmable kind, can be used to great effect. However, despite the enormous power of the modern computer, it is still true that skill and experience play a major part in successful coating design. The computer brings little in the way of understanding. Understanding is the emphasis in the bulk of this book. There are many techniques that date back to times when computers were expensive, cumbersome, and scarce, and alternatives, usually approximate, were required. These

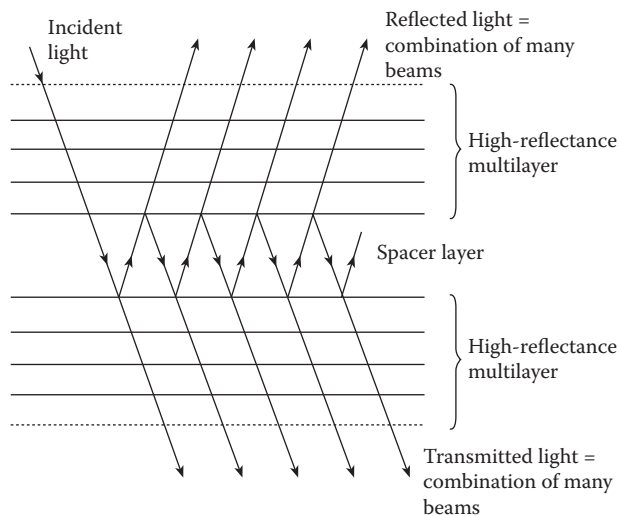


FIGURE 1.3

Fabry–Perot filter showing multiple reflections in the spacer or cavity layer.

would not be used for calculation today, but they bring an insight that straightforward calculation cannot deliver, even if it is very fast. Thus, we include many such techniques, and it is convenient to introduce them often in an historical context. The matrix method itself brings many advantages. For example, it has made possible the development of exceedingly powerful design techniques based on the algebraic manipulation of the matrices. These are also included. Graphical techniques are of considerable usefulness in the visualization of the properties of coatings. There are many such techniques, but in this book, we pay particular attention to one such method known as the admittance diagram. This is one that your author has found of considerable assistance over the years. It is an accurate technique in the sense that it contains no approximations other than those involved perhaps in sketching it, but it is normally used as an aid to understanding rather than as a calculation tool.

In the design of a thin-film multilayer, we are required to find an arrangement of layers to give a performance specified in advance, and this is much more difficult than the straightforward calculation of the properties of a given multilayer. There is no precise analytical solution to the general problem. The normal method of design is to arrive at a possible structure for a filter, using techniques to be described, which consist of a mixture of analysis, experience, and use of well-known building blocks. The evaluation is then completed by calculating the performance on a computer. Depending on the results of the computations, adjustments to the proposed design may be made and then recomputed, until a satisfactory solution is found. This adjustment process can itself be undertaken by a computer and is usually known by the term *refinement*. A related term is *synthesis*, which implies an element of construction as well as adjustment. The ultimate in synthesis is the complete construction of a design with no starting information beyond the performance specification, but it is normal to provide some starting information, such as materials to be used and, possibly, total thickness of coating or a very rough starting design.

The successful application of refinement techniques largely depends on a starting solution that has a performance close to that required. Under these conditions, it has been made to work exceedingly well. The operation of a refinement process involves the adjustment of the parameters of the system to minimize a merit coefficient (in some less common versions, a measure of merit may be maximized) representing the gap between the performance achieved by the design at any stage and the desired performance. The main difference between the various techniques is in the details of the rules used in adjusting the design. A major problem is the enormous number of parameters that can potentially be involved. Refinement is usually kept within bounds by limiting the search to small changes in an almost acceptable starting design. In synthesis with no starting design, the possibilities are virtually infinite, and so the rules governing the search procedure have to be very carefully organized. The most effective techniques incorporate two elements, an effective refinement technique that operates until it reaches a limit and a procedure for complicating the design that is then applied. These two elements alternate as the design is gradually constructed. Automatic design synthesis is undoubtedly increasing in importance in step with developments in computers, but it is still true that in the hands of a skilled practitioner, the achievements of both refinement and synthesis are much more impressive than when no skill is involved. Someone who knows well what he or she is doing will always succeed much better than someone who does not. This branch of the subject is much more a matter of computing techniques rather than fundamental to the understanding of thin-film filters, and so it is largely outside the scope of this book. The book by Liddell [24] and the more recent text by Furman and Tikhonravov [25] give good introductory accounts of various methods. The real limitation to what is, at the present time, possible in optical thin-film filters and coatings is the capability of the manufacturing process to produce layers of precisely the correct optical constants and thickness, rather than any deficiency in design techniques.

The common techniques for the construction of thin-film optical coatings can be classified as physical vapor deposition. They are vacuum processes where a solid film condenses from the vapor phase. The most straightforward, and the traditional method, is known as thermal evaporation,

and this is still much used. Because of the defects of solidity possessed by thermally evaporated films, there has, in recent years, been a shift, now accelerating, toward what are described as the energetic processes. Here, mechanical momentum is transferred to the growing film, either by deliberate bombardment or by an increase in the momentum of the arriving film material, and this added momentum drives the outermost material deeper into the film, increasing its solidity. These processes are briefly described in the later chapters of the book, but much more information will be found in the books listed in the bibliography at the end of this chapter.

Then some words of explanation might be useful. Except for some deliberately simplified, and therefore approximate, techniques that will be positively identified, the theory that will be presented of our thin-film interference effects is exact and of perfect precision so that any numerical results can reflect the precision of the numerical data that are entered. However, in almost the entire field of thin-film activity, we do not have perfect precision. For example, characteristic values of material parameters fluctuate with deposition conditions and are difficult to measure with extreme accuracy. Reflectance and transmittance measurements are similarly limited. Thus, although the theory can accommodate any degree of precision that we might wish, in practice, we are rather limited by our imperfect knowledge and the behavior of our real samples. We reflect this in our calculations in this book where we will frequently suggest numerical values for our demonstrations that will rarely be of greater precision than two places of decimal for refractive indices. In particular, we will assume a value of the refractive index of air as unity, although in reality, air has a very slightly higher index, at around 1.00029, that varies with factors such as humidity, composition and temperature, and wavelength. Thus, our air will be indistinguishable from vacuum in our numerical results. This is common practice in the thin-film field. The theory will certainly support whatever precision the user should require.

In Chapter 2, which deals with theory, it will become clear that an optical material is characterized by two different physical parameters, its refractive index and its characteristic admittance, both of which may be complex. In the optical region, from soft X-rays to the far infrared, in other words, the region that interests us throughout this book, the characteristic admittance is proportional to the refractive index, the constant of proportionality being the characteristic admittance of free space. The refractive index is unitless, while the SI unit of admittance is the siemens. By changing the units of the characteristic admittance to units of the free space admittance, the numerical value of characteristic admittance becomes equal to the refractive index. Therefore, a specification of refractive index is also a specification of characteristic admittance. When discussing the properties of a thin film, therefore, there is no need slavishly to state both quantities. The theoretical expressions will indicate the correct choice. Unless it is important that we distinguish between the two parameters, we shall usually follow normal practice in tending to use refractive index when referring to a material.

References

1. O. S. Heavens. 1955. *Optical Properties of Thin Solid Films*. London: Butterworth Scientific Publications.
2. L. Holland. 1956. *Vacuum Deposition of Thin Films*. London: Chapman & Hall.
3. H. Anders. 1965. *Dünne Schichten für die Optik* (English translation: *Thin Films in Optics*. Waltham, MA: Focal Press, 1967). Stuttgart: Wissenschaftliche Verlagsgesellschaft mbH.
4. Z. Knittl. 1976. *Optics of Thin Films*. Hoboken, NJ: John Wiley & Sons; Berlin: SNTL.
5. I. Newton. 1704. *Opticks or a Treatise of the Reflections, Refractions, Inflections and Colours of Light*. London: The Royal Society.
6. T. Young. 1802. On the theory of light and colours (The 1801 Bakerian Lecture). *Philosophical Transactions of the Royal Society of London* 92:12–48.

7. H. Crew. 1930. Thomas Young's place in the history of the wave theory of light. *Journal of the Optical Society of America* 20:3–10.
8. H. de Sénarmont, E. Verdet, and L. Fresnel, eds. 1866–1870. *Oeuvres complètes d'Augustin Fresnel*. Paris: Imprimerie Impériale.
9. Z. Knittl. 1978. Fresnel historique et actuel. *Optica Acta* 25:167–173.
10. J. C. Maxwell. 1873. *A Treatise on Electricity and Magnetism*. Wotton-under-Edge: Clarendon Press.
11. J. von Fraunhofer. 1817. Versuche über die Ursachen des Anlaufens und Mattwerdens des Glases und die Mittel, denselben zuvorzukommen. In J. von Fraunhofer. 1888. *Gesammelte Schriften*, pp. 33–49. Munich: Verlag der Königlich Bayerischen Akademie der Wissenschaften.
12. Lord Rayleigh. 1886. On the intensity of light reflected from certain surfaces at nearly perpendicular incidence. *Proceedings of the Royal Society of London* 41:275–294.
13. H. D. Taylor. 1891. *On the Adjustment and Testing of Telescopic Objectives*. York: T. Cooke & Sons.
14. H. D. Taylor. 1983. *On the Adjustment and Testing of Telescopic Objectives*. Fifth ed. Bristol: Adam Hilger.
15. H. D. Taylor. 1904. *Lenses*. UK Patent, 29561.
16. F. Kollmogoren. 1916. Light transmission through telescopes. *Transactions of the American Illumination Engineering Society* 11:220–228.
17. C. Fabry and A. Perot. 1899. Théorie et applications d'une nouvelle méthode de spectroscopie interférentielle. *Annales de Chimie et de Physique, Paris, 7th series* 16:115–144.
18. P. Rouard. 1932. Sur le pouvoir réflecteur des métaux en lames très minces. *Comptes Rendus de l'Académie de Science* 195:869–872.
19. G. Bauer. 1934. Absolutwerte der optischen Absorptionskonstanten von Alkalihalogenidkristallen im Gebiet ihrer ultravioletten Eigenfrequenzen. *Annalen der Physik (Leipzig), 5th series* 19:434–464.
20. A. H. Pfund. 1934. Highly reflecting films of zinc sulphide. *Journal of the Optical Society of America* 24:99–102.
21. J. Strong. 1936. On a method of decreasing the reflection from non-metallic substances. *Journal of the Optical Society of America* 26:73–74.
22. W. Geffcken. 1939. Interferenzlichtfilter. *German Patent, DE 716 153*.
23. A. Thelen. 1997. The pioneering contributions of W. Geffcken. In *Thin Films on Glass*, H. Bach and D. Krause (eds), pp. 227–239. Berlin: Springer-Verlag.
24. H. M. Liddell. 1981. *Computer-Aided Techniques for the Design of Multilayer Filters*. Bristol: Adam Hilger.
25. S. A. Furman and A. V. Tikhonravov. 1992. *Basics of Optics of Multilayer Systems*. First ed. Gif-sur-Yvette: Editions Frontières.

Bibliography

A complete bibliography of primary references would stretch to an enormous length. This list is, therefore, primarily one of secondary references.

1. H. Anders. 1965. *Dünne Schichten für die Optik* (English translation: *Thin Films in Optics*. Waltham, MA: Focal Press, 1967). Stuttgart: Wissenschaftliche Verlagsgesellschaft mbH.
2. H. Bach and D. Krause, eds. 1997. *Thin Films on Glass*. Berlin, Heidelberg: Springer-Verlag.
3. P. W. Baumeister. 2004. *Optical Coating Technology*. Bellingham, WA: SPIE Press.
4. D. H. Cushing. 2011. *Enhanced Optical Filter Design*. Bellingham, WA: SPIE Press.
5. J. A. Dobrowolski. 1995. Optical properties of films and coatings. In *Handbook of Optics*, M. Bass, E. W. V. Stryland, D. R. Williams et al. (eds), pp. 42.1–42.130. New York: McGraw-Hill.
6. F. R. Flory, ed. 1995. *Thin Films for Optical Systems*. New York: Marcel Dekker.
7. H. Frey and G. Kienel, eds. 1987. *Dünnschicht Technologie*. Düsseldorf: VDI-Verlag GmbH.
8. S. A. Furman and A. V. Tikhonravov. 1992. *Basics of Optics of Multilayer Systems*. First ed. Gif-sur-Yvette: Editions Frontières.
9. H. L. Hartnagel, A. L. Dawar, A. K. Jain et al. 1995. *Semiconducting Transparent Thin Films*. Bristol: Institute of Physics.
10. O. S. Heavens. 1955. *Optical Properties of Thin Solid Films*. London: Butterworth Scientific Publications.

11. I. J. Hodgkinson and Q. H. Wu. 1997. *Birefringent Thin Films and Polarizing Elements*. First ed. Singapore: World Scientific.
12. L. Holland. 1956. *Vacuum Deposition of Thin Films*. London: Chapman & Hall.
13. R. E. Hummel and K. H. Guenther, eds. 1995. *Thin Films for Optical Coatings*. Boca Raton, FL: CRC Press.
14. M. R. Jacobson, ed. 1988. *Modeling of Optical Thin Films*, vol. 821. Bellingham: SPIE Press.
15. M. R. Jacobson, ed. 1989. *Deposition of Optical Coatings*. Bellingham: SPIE Press.
16. M. R. Jacobson, ed. 1990. *Design of Optical Coatings*. Bellingham: SPIE Press.
17. M. R. Jacobson, ed. 1992. *Characterization of Optical Coatings*. Bellingham: SPIE Press.
18. N. Kaiser and H. K. Pulker, eds. 2003. *Optical Interference Coatings*. Berlin: Springer-Verlag.
19. Z. Knittl. 1976. *Optics of Thin Films*. Hoboken, NJ: John Wiley & Sons; Berlin: SNTL.
20. H. M. Liddell. 1981. *Computer-Aided Techniques for the Design of Multilayer Filters*. Bristol: Adam Hilger.
21. P. H. Lissberger. 1970. Optical applications of dielectric thin films. *Reports of Progress in Physics* 33:197–268.
22. B. E. Perilloux. 2002. *Thin Film Design: Modulated Thickness and Other Stopband Design Methods*, vol. TT57. Bellingham: SPIE Optical Engineering Press.
23. A. Piegari and F. Flory, eds. 2013. *Optical Thin Films and Coatings. From Materials to Applications*. Oxford: Woodhead.
24. H. K. Pulker. 1999. *Coatings on Glass*. Second ed. Amsterdam: Elsevier.
25. J. D. Rancourt. 1987. *Optical Thin Films: Users' Handbook*. New York: Macmillan.
26. D. Ristau, ed. 2015. *Laser-Induced Damage in Optical Materials*. Boca Raton, FL: CRC Press.
27. O. Stenzel. 2014. *Optical Coatings: Material Aspects in Theory and Practice*. Berlin: Springer-Verlag.
28. A. Thelen. 1988. *Design of Optical Interference Coatings*. First ed. New York: McGraw-Hill.
29. A. Vasicek. 1960. *Optics of Thin Films*. Amsterdam: North Holland.
30. R. R. Willey. 2002. *Practical Design and Production of Optical Thin Films*. Second ed. New York: Marcel Dekker.

2

Basic Theory

The next part of the book is a long and rather tedious account of some basic theory that is necessary in order to make calculations of the properties of multilayer thin-film coatings. It is perhaps worth reading just once or when some deeper insight into thin-film calculations is required. In order to make it easier for those who have read it to find the basic results or for those who do not wish to read it at all, to proceed with the remainder of the book, the principal results are summarized in Section 2.9.

2.1 Maxwell's Equations and Plane Electromagnetic Waves

For those readers who are still with us, we begin our attack on thin-film problems by solving Maxwell's equations together with the appropriate material equations. In isotropic media, these are

$$\mathbf{curl} \mathbf{H} = \nabla \times \mathbf{H} = \mathbf{j} + \partial \mathbf{D} / \partial t, \quad (2.1)$$

$$\mathbf{curl} \mathbf{E} = \nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t, \quad (2.2)$$

$$\mathbf{div} \mathbf{D} = \nabla \cdot \mathbf{D} = \rho, \quad (2.3)$$

$$\mathbf{div} \mathbf{B} = \nabla \cdot \mathbf{B} = 0, \quad (2.4)$$

$$\mathbf{j} = \boldsymbol{\sigma} \mathbf{E}, \quad (2.5)$$

$$\mathbf{D} = \boldsymbol{\epsilon} \mathbf{E}, \quad (2.6)$$

$$\mathbf{B} = \boldsymbol{\mu} \mathbf{H}, \quad (2.7)$$

where the symbols in bold are vector quantities. In anisotropic media, Equations 2.1 through 2.7 become much more complicated with $\boldsymbol{\sigma}$, $\boldsymbol{\epsilon}$, and $\boldsymbol{\mu}$ being tensor rather than scalar quantities.

Anisotropic media are covered by Yeh [1] and Hodgkinson and Wu [2]. They are discussed quite briefly in Chapter 15.

The International System of Units (SI) is used as far as possible throughout this book. Table 2.1 shows the definitions of the quantities in the equations together with the appropriate SI units.

To the equations, we can add

$$\epsilon = \epsilon_r \epsilon_0, \quad (2.8)$$

$$\mu = \mu_r \mu_0, \quad (2.9)$$

$$\epsilon_0 = 1/(\mu_0 c^2), \quad (2.10)$$

where ϵ_0 and μ_0 are the permittivity and permeability of free space, respectively. ϵ_r and μ_r are the relative permittivity and permeability, respectively, and c is a constant that can be identified as the velocity of light in free space. ϵ_0 , μ_0 , and c are important constants, the values of which are given in Table 2.2.

In the normal way the parameters in Equations 2.8 through 2.10 do not depend on either E or H , and so the phenomena are linear. Note that there is a branch of optics dealing with non-linear effects, but the electromagnetic power density necessary for the production of such effects is usually enormous, and the materials, rather special. Such effects are outside the scope of this book.

Linear implies that the response to the sum of a set of stimuli is the sum of the responses to each stimulus separately. Thus, we can divide any arbitrary electromagnetic wave into components that can be separately considered. The usefulness of this approach lies in the fact that what we call a harmonic wave, that is a wave with a sine or cosine profile, propagates through any dispersive

TABLE 2.1

Electromagnetic Parameters

Symbol	Physical Quantity	SI Unit	Symbol for SI Unit
E	Electric field strength	Volts per meter	V/m
D	Electric displacement	Coulombs per square meter	C/m ²
H	Magnetic field strength	Amperes per meter	A/m
j	Electric current density	Amperes per square meter	A/m ²
B	Magnetic flux density or magnetic induction	Tesla	T
ρ	Electric charge density	Coulombs per cubic meter	C/m ³
σ	Electric conductivity	Siemens per meter	S/m
μ	Permeability	Henrys per meter	H/m
ϵ	Permittivity	Farads per meter	F/m

TABLE 2.2

Physical Constants

Symbol	Physical Quantity	Value
c	Velocity of light in free space	2.99792458×10^8 m/s
μ_0	Permeability of free space	$4\pi \times 10^{-7}$ H/m
ϵ_0	Permittivity of free space = $1/(\mu_0 c^2)$	$8.854187817 \times 10^{-12}$ F/m

medium with no change of frequency and, therefore, retains its shape and has a precise velocity. There is a complete body of theory, known as Fourier, that permits a profile to be broken down into a set of sine and/or cosine functions. Thus, we use the harmonic wave as our basic component, and the collection of harmonic components that makes up our primary wave is known as its spectrum. We are quite used to breaking any light input into its spectrum and following the spectral components through the system separately, and this is the way we will normally operate in this book. In our theoretical analysis, therefore, we will concentrate on a single, general, spectral component, that is a harmonic wave, and we will usually derive what is known as the spectral performance of our coatings. Also we will tend to use the simplest type of harmonic wave, the linearly polarized, plane, harmonic wave.

The following analysis is brief and incomplete. For a full, rigorous treatment of the electromagnetic field equations, the reader is referred to Born and Wolf [3].

First, we assume an absence of space charge so that ρ is zero. This implies

$$\operatorname{div} \mathbf{D} = \epsilon(\nabla \cdot \mathbf{E}) = 0, \quad (2.11)$$

and solving for \mathbf{E} ,

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) = -\mu \sigma \frac{\partial \mathbf{E}}{\partial t} - \mu \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (2.12)$$

i.e.,

$$\nabla^2 \mathbf{E} = \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu \sigma \frac{\partial \mathbf{E}}{\partial t}. \quad (2.13)$$

A similar expression holds for \mathbf{H} .

First of all, we look for a solution of Equation 2.13 in the form of a linearly polarized plane harmonic wave (or *plane polarized*, a term meaning the same as linearly polarized), and we choose the complex form of this wave, the physical meaning being associated with either the real or the imaginary part of the expression:

$$\mathbf{E} = \mathcal{E} \exp[i\omega(t - z/v)] \quad (2.14)$$

represents such a wave propagating along the z -axis with velocity v . \mathcal{E} is the vector amplitude and ω is the angular frequency of this wave. Note that since we are dealing with linear phenomena, ω is invariant as the wave propagates through media with differing properties. The advantage of the complex form of the wave is that phase changes can be dealt with very readily by including them in a complex amplitude. If we include a relative phase ϕ in Equation 2.14, then it becomes

$$\mathbf{E} = \mathcal{E} \exp[i\{\omega(t - z/v) + \phi\}] = \mathcal{E} \exp(i\phi) \exp[i\omega(t - z/v)], \quad (2.15)$$

where $\mathcal{E} \exp(i\phi)$ is the complex vector amplitude. The complex scalar amplitude is given by $\mathcal{E} \exp(i\phi)$, where $\mathcal{E} = |\mathcal{E}|$. Equation 2.15, which has phase ϕ relative to Equation 2.14, is simply Equation 2.14 with the amplitude replaced by the complex amplitude.

In Equation 2.14, we chose to place the time variable first and the spatial variable second in the argument of the exponential. This is a convention, because we could have chosen the alternative of the spatial variable first. However, to reverse the direction of the wave in this convention, we simply change the minus sign to a plus, reversing the spatial direction. In the alternative convention, it is tempting to reverse the wave once again by changing the sign from minus to plus, but that would reverse the time axis, not the spatial direction. We shall stick to the convention in Equation 2.14 throughout this book.

For Equation 2.14 to be a solution of Equation 2.13, it is necessary that

$$\omega^2/v^2 = \omega^2\epsilon\mu - i\omega\mu\sigma. \quad (2.16)$$

In a vacuum, we have $\sigma = 0$ and $v = c$, so that from Equation 2.16,

$$c^2 = 1/\epsilon_0\mu_0, \quad (2.17)$$

which is identical to Equation 2.10. By multiplying Equation 2.15 by Equation 2.17 and dividing through by ω^2 , we obtain

$$\frac{c^2}{v^2} = \frac{\epsilon\mu}{\epsilon_0\mu_0} - i\frac{\mu\sigma}{\omega\epsilon_0\mu_0},$$

where c/v is clearly a dimensionless parameter of the medium, which we denote by N :

$$N^2 = \epsilon_r\mu_r - i\frac{\mu_r\sigma}{\omega\epsilon_0}. \quad (2.18)$$

This implies that N is of the form

$$N = c/v = n - ik. \quad (2.19)$$

There are two possible values of N from Equation 2.18, but for physical reasons, we choose that which gives a positive value of n . N is known as the complex refractive index; n , as the real part of the refractive index (or, often simply, as the refractive index, because N is real in an ideal dielectric material), and k is known as the extinction coefficient.

If the various parameters are real (not always the case), then from Equations 2.18 and 2.19,

$$n^2 - k^2 = \epsilon_r\mu_r, \quad (2.20)$$

$$2nk = \frac{\mu_r\sigma}{\omega\epsilon_0}. \quad (2.21)$$

Equation 2.14 can now be written as

$$E = \mathcal{E} \exp\{i[\omega t - (2\pi N/\lambda)z]\}, \quad (2.22)$$

where we have introduced the wavelength in free space λ ($= 2\pi c/\omega$).

Substituting $n - ik$ for N in Equation 2.22 gives

$$E = \mathcal{E} \exp[-(2\pi k/\lambda)z] \exp\{i[\omega t - (2\pi n/\lambda)z]\}, \quad (2.23)$$

and the significance of k emerges as being a measure of loss in the medium. The distance $\lambda/(2\pi k)$ is that in which the amplitude of the wave falls to $1/e$ of its original value. The way in which the power carried by the wave falls off will be considered shortly.

In passing, we also note that, provided we continue dealing with a harmonic wave, Equation 2.18 can also be written as

$$N^2 = \mu_r \left(\epsilon_r - i\frac{\sigma}{\omega\epsilon_0} \right) = \hat{\epsilon}_r\mu_r, \quad (2.24)$$

where $\hat{\epsilon}_r$ is not the same as ϵ_r but contains the entire contents of the bracketed quantity and, of course, is complex. This complex permittivity is a function of frequency. It is much used because it avoids the complication in theoretical studies of the conductivity, which it simply contains. Unfortunately, it is normally consistent with the convention, opposite to our current one, which treats the complex index as $n + ik$.

The change in phase produced by a traversal of distance z in the medium is the same as that produced by a distance nz in a vacuum. Because of this, nz is known as the optical distance, as distinct from the physical or geometrical distance. Generally, in thin-film optics, one is more interested in optical distances and optical thicknesses than in physical ones.

Since \mathcal{E} is constant, Equation 2.22 represents a linearly polarized plane wave propagating along the z -axis. For a similar wave propagating in a direction given by direction cosines (α, β, γ) , the expression becomes

$$E = \mathcal{E} \exp\{i[\omega t - (2\pi N/\lambda)(\alpha x + \beta y + \gamma z)]\}. \quad (2.25)$$

This is the simplest type of wave in an absorbing medium. In an assembly of absorbing thin films, we shall see that we are occasionally forced to adopt a slightly more complicated expression for the wave.

There are some important relationships for this type of wave which can be derived from Maxwell's equations. Let the direction of propagation of the wave be given by unit vector \hat{s} where

$$\hat{s} = \alpha i + \beta j + \gamma k$$

and where i, j , and k are unit vectors along the x, y , and z axes, respectively. From Equation 2.25, we have

$$\partial E / \partial t = i\omega E,$$

and from Equations 2.1, 2.5, and 2.6,

$$\begin{aligned} \text{curl} \mathbf{H} &= \sigma \mathbf{E} + \epsilon \partial \mathbf{E} / \partial t \\ &= (\sigma + i\omega\epsilon) \mathbf{E} \\ &= i \frac{\omega N^2}{c^2 \mu} \mathbf{E}. \end{aligned}$$

Now

$$\text{curl} = \left(\frac{\partial}{\partial x} i + \frac{\partial}{\partial y} j + \frac{\partial}{\partial z} k \right) \times,$$

where \times denotes the vector product. But

$$\frac{\partial}{\partial x} = -i \frac{2\pi N}{\lambda} \alpha = -i \frac{\omega N}{c} \alpha,$$

$$\frac{\partial}{\partial y} = -i \frac{\omega N}{c} \beta, \quad \frac{\partial}{\partial z} = -i \frac{\omega N}{c} \gamma,$$

so that

$$\text{curl} \mathbf{H} = -i \frac{\omega N}{c} (\hat{s} \times \mathbf{H}).$$

Then,

$$-i \frac{\omega N}{c} (\hat{s} \times \mathbf{H}) = i \frac{\omega N^2}{c^2 \mu} \mathbf{E},$$

i.e.,

$$(\hat{s} \times \mathbf{H}) = -\frac{N}{c\mu} \mathbf{E}, \quad (2.26)$$

and similarly

$$\frac{N}{c\mu} (\hat{s} \times \mathbf{E}) = \mathbf{H}. \quad (2.27)$$

For this type of wave, therefore, \mathbf{E} , \mathbf{H} , and \hat{s} are mutually perpendicular and form a right-handed set. The quantity $N/c\mu$ has the dimensions of an admittance and is known as the characteristic optical admittance of the medium, written as y . In free space, it can be readily shown that the optical admittance is given by

$$y = (\epsilon_0/\mu_0)^{1/2} = 2.6544 \times 10^{-3} \text{ S}. \quad (2.28)$$

Now

$$\mu = \mu_r \mu_0. \quad (2.29)$$

Direct magnetic interactions at optical frequencies are vanishingly small so that μ_r is effectively unity. Thus we can write

$$y = N y_0, \quad (2.30)$$

and

$$\mathbf{H} = y(\hat{s} \times \mathbf{E}) = N y_0 (\hat{s} \times \mathbf{E}). \quad (2.31)$$

We recall that since μ_r is unity, we can write Equation 2.8 as

$$N^2 = \epsilon_r - i \frac{\sigma}{\omega \epsilon_0} = \hat{\epsilon}_r, \quad (2.32)$$

where $\hat{\epsilon}_r$ is a complex relative permittivity. We have already mentioned this in Equation 2.24 and noted that it can usefully simplify related analytical expressions and so is frequently employed. We emphasize once again that, in the literature, the sign convention commonly used puts a plus sign in Equation 2.32 rather than the minus. The complex permittivity is also called the dielectric function, and the complex relative permittivity, the relative dielectric function.

2.1.1 Poynting Vector

An important feature of electromagnetic radiation is that it is a form of energy transport, and it is the energy associated with the wave that is normally observed. The instantaneous rate of flow of energy across a unit area is given by the Poynting vector:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}. \quad (2.33)$$

The direction of the vector is the direction of energy flow.

When we add or subtract complex numbers, or multiply them by a real number, the real parts and imaginary parts remain independent. Such operations are known as linear. Interference calculations involve adding the waves, and so we can happily use the complex wave with all its advantages in such calculations. The multiplication of two complex numbers, however, mixes the real and imaginary parts in the result. Such operations are known as nonlinear, and we are unable to directly use the complex form of the wave in them. The Poynting expression is a nonlinear one (E is multiplied by H), and so we have a problem with the complex form of the wave. Either the real or the imaginary part of the wave expression should be used. The real sine or cosine form of the wave implies its square in the result, and so the instantaneous value of the Poynting vector must oscillate at twice the frequency of the wave. We turn our attention to the mean value because it is the mean that is significant in our measurements. This is defined as the irradiance or, in the older systems of units, intensity. (Beware. Intensity is defined differently in the SI system as the power per unit solid angle from a point source.) In the SI system of units, irradiance is measured in watts per square meter. An unfortunate feature of the SI system, for our purposes, is that the symbol for irradiance is E . The use of this symbol would make it very difficult for us to distinguish between irradiance and electric field. Since both are extremely important in almost everything we do, we must be able to differentiate between them, and so we adopt a nonstandard symbol I for irradiance (which, unfortunately, is the SI symbol for intensity). The mean of the Poynting vector involves integrating the real expression over a cycle, but the complex form of the wave actually comes to our rescue. For a harmonic wave, we find that we can derive a very attractive and simple expression for the irradiance using the complex form of the wave and thus avoiding the integration. This is

$$I = \frac{1}{2} \operatorname{Re}(E \times H^*), \quad (2.34)$$

where $*$ denotes complex conjugate. It should be emphasized that the complex form *must* be used in Equation 2.34. The irradiance I is written in Equation 2.34 as a vector quantity, when it has the same direction as the flow of energy of the wave. The more usual scalar irradiance I is simply the magnitude of I . Since E and H are perpendicular, Equation 2.34 can be written as

$$I = \frac{1}{2} \operatorname{Re}(EH^*), \quad (2.35)$$

where E and H are the scalar magnitudes.

It is important to note that for the net irradiance, the electric and magnetic vectors in Equation 2.34 should be the total resultant fields due to all the waves involved. This is implicit in the derivation of the Poynting vector expression. We will return to this point when calculating reflectance and transmittance.

For a single, homogeneous, harmonic wave of the form shown in Equation 2.25,

$$H = y(\hat{s} \times E),$$

so that

$$I = \operatorname{Re} \left(\frac{1}{2} y E E^* \hat{s} \right) = \frac{1}{2} y E E^* \hat{s}. \quad (2.36)$$

Now, from Equation 2.25, the magnitude of E is given by

$$\begin{aligned} E &= \mathcal{E} \exp\{i[\omega t - (2\pi[n - ik]/\lambda)(\alpha x + \beta y + \gamma z)]\} \\ &= \mathcal{E} \exp[-(2\pi k/\lambda)(\alpha x + \beta y + \gamma z)] \exp\{i[\omega t - (2\pi n/\lambda)(\alpha x + \beta y + \gamma z)]\}, \end{aligned}$$

implying that

$$EE^* = \mathcal{E}\mathcal{E}^* \exp[-(4\pi k/\lambda)(\alpha x + \beta y + \gamma z)]$$

and

$$I = \frac{1}{2} n \mathcal{I} |\mathcal{E}|^2 \exp[-(4\pi k/\lambda)(\alpha x + \beta y + \gamma z)].$$

The expression $(\alpha x + \beta y + \gamma z)$ is simply the distance along the direction of propagation, and thus, the irradiance drops to $1/e$ of its initial value in a distance given by $\lambda/4\pi k$. The inverse of this distance is defined as the absorption coefficient α ; that is,

$$\alpha = 4\pi k/\lambda. \quad (2.37)$$

The absorption coefficient α should not be confused with the direction cosine. However,

$$|\mathcal{E}| \exp[-(2\pi k/\lambda)(\alpha x + \beta y + \gamma z)]$$

is really the amplitude of the wave at the point (x, y, z) so that a much simpler way of writing the expression for irradiance is

$$I = \frac{1}{2} n \mathcal{I} (\text{amplitude})^2 \quad (2.38)$$

or

$$I \propto n \times (\text{amplitude})^2. \quad (2.39)$$

This expression is a better form than the more usual

$$I \propto (\text{amplitude})^2. \quad (2.40)$$

The expression will frequently be used for comparing irradiances, in calculating reflectance or transmittance, for example, and if the media in which the two waves are propagating are of different index, then errors will occur unless n is included as mentioned earlier.

2.2 Notation

Throughout the book, we will be dealing with assemblies of elements involving different materials and where their order is important. In many cases, these will consist of an incident medium and a substrate separated by a number of thin films. To identify these various entities and to make their order unambiguously clear, we shall endeavor to use a consistent notation involving suffices. The incident medium will have the subscript 0 as in y_0 . The substrate, or emergent medium, will usually have the subscript m , as in y_m , although occasionally the subscript sub , as in y_{sub} . Layers will be numbered sequentially from the incident medium to the substrate or emergent medium so that the layer next to the incident medium has the subscript 1; that next to it, 2; and so on, as in y_1 , y_2 , and so on. Layer $q - 1$ will be next to layer q and will be situated on the side toward the incident medium (Figure 2.1).

There will be occasions where we will want to identify interfaces rather than layers. To differentiate between an interface and a layer, we will often use letters, usually in the order of the alphabet, to refer to interfaces, as a, b, c , etc., but numbers will be used to refer to layers.

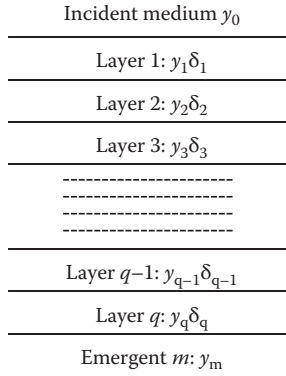


FIGURE 2.1

Numbering system for designs will normally follow the arrangement shown. A lower number will normally mean that the element is closer to the incident medium. The quantity δ is the phase thickness of the appropriate layer and will be defined shortly.

2.3 Simple Boundary

Thin-film filters usually consist of a number of boundaries between various homogeneous media, and it is the effect of these boundaries on an incident wave that we will wish to calculate. A single boundary is the simplest case. First of all, we consider absorption-free media, i.e., $k = 0$. The arrangement is sketched in Figure 2.2. A plane harmonic wave is incident on a plane surface, separating the incident medium from a second, or emergent, medium. The plane containing the normal to the surface and the direction of propagation of the incident wave is known as the plane of incidence, and the plane of the sketch corresponds to this plane. We take the z -axis as the normal into the surface in the sense of the incident wave and the x -axis as normal to it and on the plane of incidence. At a boundary, the tangential components of E and H , that is, the components along the

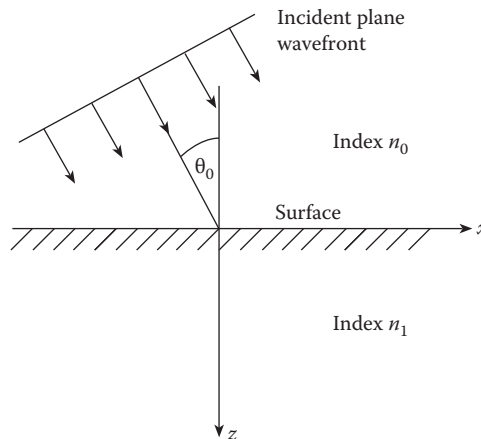


FIGURE 2.2

Plane wavefront incident on a single surface.

boundary, are continuous across it because there is no mechanism that will change them. This boundary condition is fundamental in our thin-film theory.

The first problem we have is that the boundary conditions are incompatible with a simple traversal of the boundary by the incident wave. The discontinuity in the characteristic admittance implies a power discontinuity impossible if the wave simply crosses the boundary with no other consequence. This difficulty is immediately solved by introducing a reflected wave in the incident medium, and this, of course, is directly in line with our experience. Our objective then becomes the calculation of the relative parameters of the three waves, incident, reflected, and transmitted. However, this introduces a further complication. We will use the boundary conditions to construct a set of equations from which we will extract the required relations. The complication is that the reflected wave will certainly be traveling in a different sense from the others so that there will be differences in the phase factors that will considerably complicate the calculations. We can enormously help ourselves by defining the boundary by $z = 0$, eliminating the z term from the phase factors at the boundary. Then the tangential components must be continuous for all values of x , y , and t .

We therefore have three harmonic waves, an incident, a reflected, and a transmitted wave. The incident wave, in the plane of incidence, has direction cosines $(\cos \vartheta_0, 0, \sin \vartheta_0)$. Let the direction cosines of the \hat{s} vectors of the transmitted and reflected waves be given by $(\alpha_t, \beta_t, \gamma_t)$ and $(\alpha_r, \beta_r, \gamma_r)$, respectively. We can therefore write the phase factors in the following forms:

- Incident wave: $\exp\{i[\omega t - (2\pi n_0/\lambda_i)(x \sin \vartheta_0 + z \cos \vartheta_0)]\}$
- Reflected wave: $\exp\{i[\omega t - (2\pi n_0/\lambda_r)(\alpha_r x + \beta_r y + \gamma_r z)]\}$
- Transmitted wave: $\exp\{i[\omega t - (2\pi n_1/\lambda_t)(\alpha_t x + \beta_t y + \gamma_t z)]\}$

The relative phases of these waves are included in the complex amplitudes. For waves with these phase factors to satisfy the boundary conditions for all x , y , and t at $z = 0$ implies that the coefficients of these variables must be separately identically equal. Had we not already known that there would be no change in frequency, this would have confirmed it. Since the frequencies are constant, so too will be the free space wavelengths. Next,

$$0 \equiv n_0 \beta_r \equiv n_1 \beta_t; \quad (2.41)$$

that is, the directions of the reflected and transmitted or refracted beams are confined to the plane of incidence. This, in turn, means that the direction cosines of the reflected and transmitted waves are of the forms

$$\alpha = \sin \vartheta, \quad \gamma = \cos \vartheta. \quad (2.42)$$

Also,

$$n_0 \sin \vartheta_0 \equiv n_0 \alpha_r \equiv n_1 \alpha_t,$$

so that if the angles of reflection and refraction are ϑ_r and ϑ_t , respectively, then

$$\vartheta_0 = \vartheta_r, \quad (2.43)$$

that is, the angle of reflection equals the angle of incidence, and

$$n_0 \sin \vartheta_0 = n_1 \sin \vartheta_t.$$

The result appears more symmetrical if we replace ϑ_t with ϑ_1 , giving

$$n_0 \sin \vartheta_0 = n_1 \sin \vartheta_1. \quad (2.44)$$

This is the familiar relationship known as Snell’s law. γ_r and γ_t are then given either by Equation 2.42 or by

$$\alpha_r^2 + \gamma_r^2 = 1 \quad \text{and} \quad \alpha_t^2 + \gamma_t^2 = 1. \tag{2.45}$$

Note that for the reflected beam, we must choose the *negative* root of Equation 2.45 so that the beam will propagate in the correct direction.

2.3.1 Normal Incidence in Absorption-Free Media

Let us limit our initial discussion to normal incidence, and let the incident wave be a linearly polarized plane harmonic wave. The coordinate axes are shown in Figure 2.3. The xy plane is the plane of the boundary. We can take the incident as propagating along the z -axis with the positive direction of the E vector along the x -axis. Then the positive direction of the H vector will be the y -axis. It is clear that the only waves that satisfy the boundary conditions are linearly polarized in the same sense as the incident wave.

A quoted phase difference between two waves travelling in the same direction is immediately meaningful. A phase difference between two waves travelling in opposite directions is absolutely meaningless, unless a reference plane at which the phase difference is measured is first defined. This is simply because the phase difference between oppositely propagating waves of the same frequency has a term $(\pm 4\pi ns/\lambda)$ in it, where s is a distance measured along the direction of propagation. Before proceeding further, therefore, we need to define the reference point for measurements of relative phase between the oppositely propagating beams. Since we have already used the device of defining the boundary as $z = 0$, we can continue this idea and define the boundary as that plane where the reflected phase shift should be defined.

Then there is another problem. The waves have electric and magnetic fields that with the direction of propagation form right-handed sets. Since the direction of propagation is reversed in the reflected beam, the orientation of electric and magnetic fields cannot remain the same as that in the incident beam; otherwise, we would no longer have a right-handed set. We need to decide on

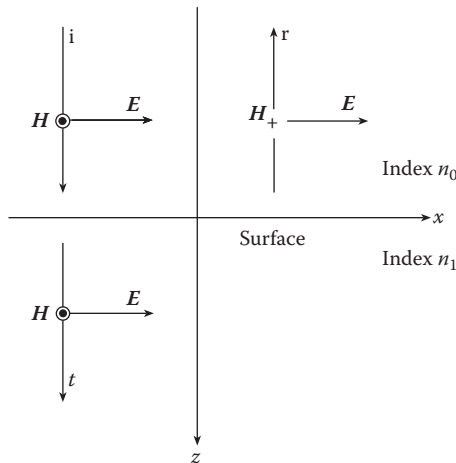


FIGURE 2.3 Convention defining positive directions of the electric and magnetic vectors for reflection and transmission at an interface at normal incidence.

how we are going to handle this. Since the electric field is the one that is most important from the point of view of interaction with matter, we will define our directions with respect to it.

The matter of phase references and electric field directions are what we call conventions because we do have complete freedom of choice, and any self-consistent arrangement is possible. We must simply ensure that once we have made our choice, we adhere to it. A good rule, however, is to never make things difficult when we can make them easy, and so we will normally choose the rule that is most convenient and least complicated. We define the positive direction of E along the x -axis for all the beams that are involved. Because of this choice, the positive direction of the magnetic vector will be along the y -axis for the incident and transmitted waves, but along the negative direction of the y -axis for the reflected wave.

We now consider the boundary conditions. Since we have already made sure that the phase factors are satisfactory, we can cancel out the $i\omega t$ terms and have only to consider the amplitudes, and we will be including any phase changes in these:

1. Electric vector continuous across the boundary

$$\mathcal{E}_i + \mathcal{E}_r = \mathcal{E}_t. \quad (2.46)$$

2. Magnetic vector continuous across the boundary

$$\mathcal{H}_i - \mathcal{H}_r = \mathcal{H}_t,$$

where we must use a minus sign because of our convention for positive directions. The relationship between magnetic and electric fields through the characteristic admittance, gives

$$y_0 \mathcal{E}_i - y_0 \mathcal{E}_r = y_1 \mathcal{E}_t. \quad (2.47)$$

This can also be derived using the vector relationship in Equations 2.31 and 2.46. We can eliminate \mathcal{E}_t to give

$$y_1(\mathcal{E}_i + \mathcal{E}_r) = y_0(\mathcal{E}_i - \mathcal{E}_r),$$

i.e.,

$$\frac{\mathcal{E}_r}{\mathcal{E}_i} = \frac{y_0 - y_1}{y_0 + y_1} = \frac{n_0 - n_1}{n_0 + n_1}, \quad (2.48)$$

the second part of the relationship being correct only because at optical frequencies, we can write

$$y = n\mathcal{Y}.$$

Similarly, eliminating \mathcal{E}_r ,

$$\frac{\mathcal{E}_t}{\mathcal{E}_i} = \frac{2y_0}{y_0 + y_1} = \frac{2n_0}{n_0 + n_1}. \quad (2.49)$$

These quantities are called the amplitude reflection and transmission coefficients and are denoted by ρ and τ , respectively. Thus

$$\rho = \frac{y_0 - y_1}{y_0 + y_1} = \frac{n_0 - n_1}{n_0 + n_1}, \quad (2.50)$$

$$\tau = \frac{2y_0}{y_0 + y_1} = \frac{2n_0}{n_0 + n_1}. \quad (2.51)$$

We are still assuming zero for k , and so in this particular case, all y are real, and these two derived quantities are therefore real. τ is always a positive real number, indicating that according to our phase convention, there is no phase shift between the incident and transmitted beams at the interface. The behavior of ρ indicates that there will be no phase shift between the incident and reflected beams at the interface provided $n_0 > n_1$, but that if $n_0 < n_1$, there will be a phase change of π because the value of ρ becomes negative.

We now examine the energy balance at the boundary. The total tangential components of electric and magnetic field are not only continuous across the boundary, but also, since the boundary is of zero thickness, it can neither supply energy to nor extract energy from the various waves. On both counts, the Poynting vector, that is the net irradiance, will be continuous across the boundary, so that we can write

$$\begin{aligned}\text{Net irradiance} &= \text{Re} \left[\frac{1}{2} (\mathcal{E}_i + \mathcal{E}_r)(y_0 \mathcal{E}_i - y_0 \mathcal{E}_r)^* \right] \\ &= \text{Re} \left[\frac{1}{2} \mathcal{E}_i (y_1 \mathcal{E}_t)^* \right]\end{aligned}$$

(using $\text{Re}(\frac{1}{2} \mathbf{E} \times \mathbf{H}^*)$ and Equations 2.46 and 2.47). Now

$$\mathcal{E}_r = \rho \mathcal{E}_i \quad \text{and} \quad \mathcal{E}_t = \tau \mathcal{E}_i,$$

i.e.,

$$\text{Net irradiance} = \frac{1}{2} y_0 \mathcal{E}_i \mathcal{E}_i^* (1 - \rho^2) = \frac{1}{2} y_0 \mathcal{E}_i \mathcal{E}_i^* (y_1/y_0) \tau^2. \quad (2.52)$$

We recognize $(1/2)y_0 \mathcal{E}_i \mathcal{E}_i^*$ as the irradiance of the incident beam I_i . We can identify $\rho^2(1/2)y_0 \mathcal{E}_i \mathcal{E}_i^* = \rho^2 I_i$ as the irradiance of the reflected beam I_r and $(y_1/y_0) \times \tau^2(1/2)y_0 \mathcal{E}_i \mathcal{E}_i^* = (y_1/y_0)\tau^2 I_i$ as the irradiance of the transmitted beam I_t . We define the reflectance R as the ratio of the reflected and incident irradiances and the transmittance T as the ratio of the transmitted and incident irradiances. Then,

$$\begin{aligned}T = \frac{I_t}{I_i} &= \frac{y_1}{y_0} \tau^2 = \frac{4y_0 y_1}{(y_0 + y_1)^2} = \frac{4n_0 n_1}{(n_0 + n_1)^2}, \\ R = \frac{I_r}{I_i} &= \rho^2 = \left(\frac{y_0 - y_1}{y_0 + y_1} \right)^2 = \left(\frac{n_0 - n_1}{n_0 + n_1} \right)^2.\end{aligned} \quad (2.53)$$

From Equation 2.52, we have, using Equation 2.53,

$$(1 - R) = T. \quad (2.54)$$

Equations 2.52 through 2.54 are therefore consistent with our ideas of splitting the irradiances into incident, reflected, and transmitted irradiances which can be treated as separate waves, the energy flow into the second medium being simply the difference of the incident and reflected irradiances. Remember that all this, so far, assumes that there is no absorption. We shall shortly see that the situation changes when absorption is present.

2.3.2 Oblique Incidence in Absorption-Free Media

Now let us consider oblique incidence, still retaining our absorption-free media. For any general direction of the vector amplitude of the incident wave, we quickly find that the application of the boundary conditions leads us into complicated and difficult expressions for the vector amplitudes of the reflected and transmitted waves. Fortunately, there are two orientations of the incident

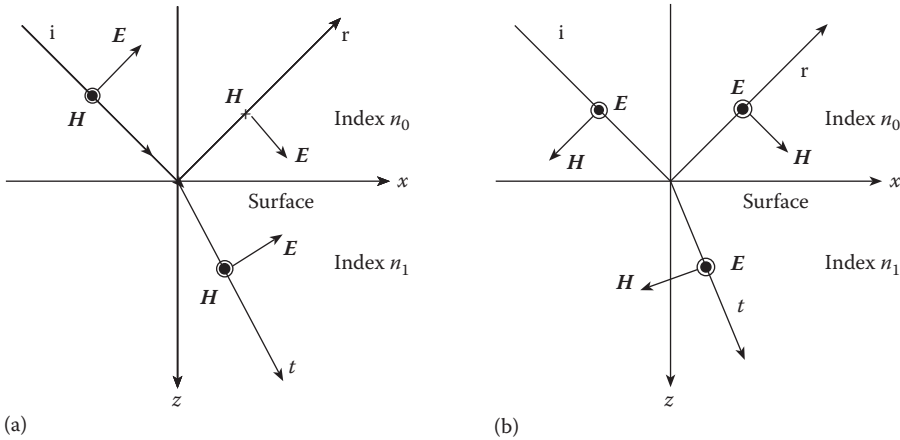


FIGURE 2.4

(a) Convention defining the positive directions of the electric and magnetic vectors for p -polarized light (TM waves).
 (b) Convention defining the positive directions of the electric and magnetic vectors for s -polarized light (TE waves).

wave, which lead to reasonably straightforward calculations, the vector electrical amplitudes aligned in the plane of incidence (i.e., the xy plane in Figure 2.2) and the vector electrical amplitudes aligned normal to the plane of incidence (i.e., parallel to the y -axis in Figure 2.2). In each of these cases, the orientations of the transmitted and reflected vector amplitudes are the same as for the incident wave. Any incident wave of arbitrary polarization can therefore be split into two components having these simple orientations. The transmitted and reflected components can be separately calculated for each orientation and then combined to yield the resultant. Since, therefore, it is necessary to consider two orientations only, they have been given special names. A wave with the electric vector in the plane of incidence is known as p -polarized or, sometimes, as TM (for transverse magnetic), and a wave with the electric vector normal to the plane of incidence, as s -polarized or, sometimes, TE (for transverse electric). p and s are derived from the German *parallel* and *senkrecht* (perpendicular). Before we can actually proceed to the calculation of the reflected and transmitted amplitudes, we must choose the various reference directions of the vectors from which any phase differences will be calculated. We have, once again, complete freedom of choice, but once we have established the convention, we must adhere to it, just as in the normal incidence case. The conventions which we will use in this book are illustrated in Figure 2.4. They have been chosen to be compatible with those for normal incidence already established. In some works, an opposite convention for the p -polarized reflected beam has been adopted, but this leads to an incompatibility with results derived for normal incidence, and we prefer to avoid this situation. Note that for reasons connected with consistency of reference directions for elliptically polarized light, the convention normal in ellipsometric calculations is opposite to that of Figure 2.4 for reflected p -polarized light. When ellipsometric parameters (in reflection) are compared with the results of the expressions we shall use, it will usually be necessary to introduce a shift of 180° in the p -polarized reflected results.

We can now apply the boundary conditions. Since we have already ensured that the phase factors will be correct, we need only consider the vector amplitudes.

2.3.2.1 p -Polarized Light

1. Electric component parallel to the boundary; continuous across it:

$$\mathcal{E}_i \cos \vartheta_0 + \mathcal{E}_r \cos \vartheta_0 = \mathcal{E}_t \cos \vartheta_1. \quad (2.55)$$

2. Magnetic component parallel to the boundary; continuous across it: Here we need to calculate the magnetic vector amplitudes, and we can do this by either using Equation 2.31 to operate on Equation 2.55 directly, or, since the magnetic vectors are already parallel to the boundary, using Figure 2.4 and then converting, since $\mathcal{H} = y\mathcal{E}$:

$$y_0\mathcal{E}_i - y_0\mathcal{E}_r = y_1\mathcal{E}_t. \quad (2.56)$$

At first sight, it seems logical just to eliminate first \mathcal{E}_r and then \mathcal{E}_t from these two equations to obtain $\mathcal{E}_r/\mathcal{E}_i$ and $\mathcal{E}_t/\mathcal{E}_i$:

$$\begin{aligned} \frac{\mathcal{E}_r}{\mathcal{E}_i} &= \frac{y_0 \cos \vartheta_1 - y_1 \cos \vartheta_0}{y_0 \cos \vartheta_1 + y_1 \cos \vartheta_0}, \\ \frac{\mathcal{E}_t}{\mathcal{E}_i} &= \frac{2y_0 \cos \vartheta_0}{y_0 \cos \vartheta_1 + y_1 \cos \vartheta_0}, \end{aligned} \quad (2.57)$$

and then simply to set

$$R = \left(\frac{\mathcal{E}_r}{\mathcal{E}_i}\right)^2 \quad \text{and} \quad T = \frac{y_1}{y_0} \left(\frac{\mathcal{E}_t}{\mathcal{E}_i}\right)^2,$$

but when we calculate the expressions which result, we find that $R + T \neq 1$. In fact, there is no mistake in the calculations. We have computed the irradiances measured along the direction of propagation of the waves, and the transmitted wave is inclined at an angle which differs from that of the incident wave. This leaves us with the problem that adopting these definitions will involve the rejection of the ($R + T = 1$) rule.

We could correct this situation by modifying the definition of T to include this angular dependence, but an alternative, preferable, and generally adopted approach is to use the components of the energy flows that are normal to the boundary. The \mathbf{E} and \mathbf{H} vectors that are involved in these calculations are then parallel to the boundary. Since these are those that directly enter into the boundary, it seems appropriate to concentrate on them when we are dealing with the amplitudes of the waves. Note that reflectance and transmittance defined for infinite plane waves in terms of normal flows of irradiance are absolutely consistent with reflectance and transmittance defined in terms of the ratios of total beam power when using confined beams such as the output from a laser.

The thin-film approach to all this, then, is to use the components of \mathbf{E} and \mathbf{H} parallel to the boundary, what are called the tangential components, in the expressions ρ and τ that involve amplitudes. Note that the normal approach in other areas of optics is to use the full components of \mathbf{E} and \mathbf{H} in amplitude expressions but to use the components of irradiance in reflectance and transmittance. The amplitude coefficients are then known as the Fresnel coefficients. The thin-film coefficients are *not* the Fresnel coefficients except at normal incidence, although the only coefficient that actually has a different value is the amplitude transmission coefficient for p -polarization.

The tangential components of \mathbf{E} and \mathbf{H} , that is, the components parallel to the boundary, have already been calculated for use in Equations 2.55 and 2.56. However, it is convenient to introduce special symbols for them, \mathcal{E} and \mathcal{H} .

Then, we can write

$$\mathcal{E}_i = \mathcal{E}_i \cos \vartheta_0, \quad \mathcal{H}_i = \mathcal{H}_i = y_0 \mathcal{E}_i = \frac{y_0}{\cos \vartheta_0} \mathcal{E}_i, \quad (2.58)$$

$$\mathcal{E}_r = \mathcal{E}_r \cos \vartheta_0, \quad \mathcal{H}_r = \frac{y_0}{\cos \vartheta_0} \mathcal{E}_r, \quad (2.59)$$

$$\mathcal{E}_t = \mathcal{E}_t \cos \vartheta_1, \quad \mathcal{H}_t = \frac{y_1}{\cos \vartheta_1} \mathcal{E}_t. \quad (2.60)$$

The orientations of these vectors are exactly the same as for normally incident light.

Equations 2.55 and 2.56 can then be written as follows:

1. Electric field parallel to the boundary

$$\mathcal{E}_i + \mathcal{E}_r = \mathcal{E}_t.$$

2. Magnetic field parallel to the boundary

$$\frac{y_0}{\cos \vartheta_0} \mathcal{H}_i - \frac{y_0}{\cos \vartheta_0} \mathcal{H}_r = \frac{y_1}{\cos \vartheta_1} \mathcal{H}_t,$$

giving us, by a process exactly similar to that we have already used for normal incidence,

$$\rho_p = \frac{\mathcal{E}_r}{\mathcal{E}_i} = \left(\frac{y_0}{\cos \vartheta_0} - \frac{y_1}{\cos \vartheta_1} \right) / \left(\frac{y_0}{\cos \vartheta_0} + \frac{y_1}{\cos \vartheta_1} \right), \quad (2.61)$$

$$\tau_p = \frac{\mathcal{E}_t}{\mathcal{E}_i} = \left(\frac{2y_0}{\cos \vartheta_0} \right) / \left(\frac{y_0}{\cos \vartheta_0} + \frac{y_1}{\cos \vartheta_1} \right), \quad (2.62)$$

$$R_p = \left[\left(\frac{y_0}{\cos \vartheta_0} - \frac{y_1}{\cos \vartheta_1} \right) / \left(\frac{y_0}{\cos \vartheta_0} + \frac{y_1}{\cos \vartheta_1} \right) \right]^2, \quad (2.63)$$

$$T_p = \left(\frac{4y_0y_1}{\cos \vartheta_0 \cos \vartheta_1} \right) / \left(\frac{y_0}{\cos \vartheta_0} + \frac{y_1}{\cos \vartheta_1} \right)^2, \quad (2.64)$$

where $y_0 = n_0 \cos \vartheta_0$ and $y_1 = n_1 \cos \vartheta_1$ and the ($R + T = 1$) rule is retained. The subscript p has been used in the expressions mentioned earlier to denote p -polarization.

It should be noted that the expression for τ_p is now different from that in Equation 2.57, the form of the Fresnel amplitude transmission coefficient. Fortunately, the reflection coefficients in Equations 2.57 and 2.63 are identical, and since much more use is made of reflection coefficients, confusion is rare.

2.3.2.2 s -Polarized Light

In the case of s -polarization, the amplitudes of the components of the waves parallel to the boundary are

$$\mathcal{E}_i = \mathcal{E}_i, \quad \mathcal{H}_i = \mathcal{H}_i \cos \vartheta_0 = (y_0 \cos \vartheta_0) \mathcal{E}_i,$$

$$\mathcal{E}_r = \mathcal{E}_r, \quad \mathcal{H}_r = \mathcal{H}_r \cos \vartheta_0 = (y_0 \cos \vartheta_0) \mathcal{E}_r,$$

$$\mathcal{E}_t = \mathcal{E}_t, \quad \mathcal{H}_t = (y_1 \cos \vartheta_1) \mathcal{E}_t,$$

and here we have again an orientation of the tangential components exactly as for normally incident light, and so a similar analysis leads to

$$\rho_s = \frac{\mathcal{E}_r}{\mathcal{E}_i} = (y_0 \cos \vartheta_0 - y_1 \cos \vartheta_1) / (y_0 \cos \vartheta_0 + y_1 \cos \vartheta_1), \quad (2.65)$$

$$\tau_s = \frac{\mathcal{E}_t}{\mathcal{E}_i} = (2y_0 \cos \vartheta_0) / (y_0 \cos \vartheta_0 + y_1 \cos \vartheta_1), \quad (2.66)$$

$$R_s = [(y_0 \cos \vartheta_0 - y_1 \cos \vartheta_1) / (y_0 \cos \vartheta_0 + y_1 \cos \vartheta_1)]^2, \quad (2.67)$$

$$T_s = (4y_0 \cos \vartheta_0 y_1 \cos \vartheta_1) / (y_0 \cos \vartheta_0 + y_1 \cos \vartheta_1)^2, \quad (2.68)$$

where once again $y_0 = n_0 \mathcal{N}$ and $y_1 = n_1 \mathcal{N}$ and the $(R + T = 1)$ rule is retained. The subscript s is used in the preceding expressions to denote s -polarization.

2.3.3 Optical Admittance for Oblique Incidence

The expressions which we have derived so far have been in their traditional form (except for the use of the tangential components rather than the full vector amplitudes), and they involve the characteristic admittances of the various media or their refractive indices together with the admittance of free space \mathcal{N} . However, the notation is becoming increasingly cumbersome and will appear even more so when we consider the behavior of thin films.

Equation 2.31 gives $\mathbf{H} = y(\hat{s} \times \mathbf{E})$, where $y = N\mathcal{N}$ is the optical admittance. We have found it convenient to deal with \mathcal{E} and \mathcal{H} , the components of \mathcal{E} and \mathcal{H} parallel to the boundary, and so we introduce a tilted optical admittance η which connects \mathcal{E} and \mathcal{H} as

$$\eta = \frac{\mathcal{H}}{\mathcal{E}}. \quad (2.69)$$

At normal incidence, $\eta = y = n\mathcal{N}$ while at oblique incidence,

$$\eta_p = \frac{y}{\cos \vartheta} = \frac{n\mathcal{N}}{\cos \vartheta}, \quad (2.70)$$

$$\eta_s = y \cos \vartheta = n\mathcal{N} \cos \vartheta, \quad (2.71)$$

where the ϑ and the y in Equations 2.70 and 2.71 are those appropriate to the particular medium. In particular, Snell's law (Equation 2.44) must be used to calculate δ .

Then, in all cases, we can write

$$\rho = \left(\frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right), \quad \tau = \left(\frac{2\eta_0}{\eta_0 + \eta_1} \right), \quad (2.72)$$

$$R = \left(\frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right)^2, \quad T = \frac{4\eta_0 \eta_1}{(\eta_0 + \eta_1)^2}. \quad (2.73)$$

These expressions can be used to compute the variation of reflectance of simple boundaries between extended media. Examples are shown in Figure 2.5 of the variation of reflectance with the angle of incidence. In this case, there is no absorption in the material, and it can be seen that the reflectance for p -polarized light (TM) falls to zero at a definite angle. This particular angle is known as the Brewster angle and is of some importance. There are many applications where the windows of a cell must have close to zero reflection loss. When it can be arranged that the light will be linearly polarized, a plate tilted at the Brewster angle will be a good solution. The light that is reflected at the Brewster angle is also linearly polarized with electric vector normal to the plane of incidence. This affords a way of identifying the absolute direction of polarizers and analyzers—very difficult in any other way.

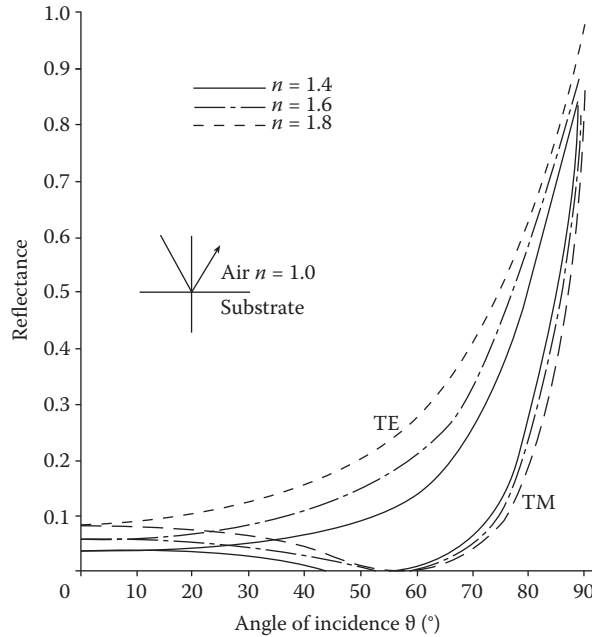


FIGURE 2.5

Variation of reflectance with angle of incidence for various values of refractive index. *TE* is *s*-polarization and *TM* is *p*-polarization.

The expression for the Brewster angle can be derived as follows. For the *p*-reflectance to be zero, from Equation 2.63,

$$\frac{y_0}{\cos \vartheta_0} = \frac{n_0 \mathcal{Y}}{\cos \vartheta_0} = \frac{y_1}{\cos \vartheta_1} = \frac{n_1 \mathcal{Y}}{\cos \vartheta_1}.$$

Snell's law gives another relationship between ϑ_0 and ϑ_1 :

$$n_0 \sin \vartheta_0 = n_1 \sin \vartheta_1.$$

Eliminating ϑ_1 from these two equations gives an expression for ϑ_0 :

$$\tan \vartheta_0 = n_1/n_0. \quad (2.74)$$

Note that this derivation depends on the relationship $y = n\mathcal{Y}$, valid at optical frequencies.

Figure 2.6 shows the variation of tilted admittance of a number of dielectric materials as a function of the angle of incidence in air. Note that the divergence of the two tilted admittances, the polarization splitting, becomes less as the index of refraction increases.

2.3.4 Normal Incidence in Absorbing Media

We must now examine the modifications necessary in our results in the presence of absorption. First, we consider the case of normal incidence and write

$$N_0 = n_0 - ik_0,$$

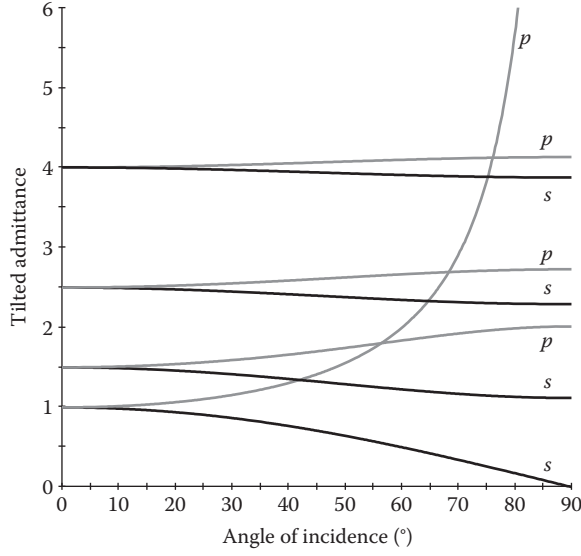


FIGURE 2.6

Tilted admittances of several dielectric (absorption-free) materials as a function of the angle of incidence in air.

$$N_1 = n_1 - ik_1,$$

$$y_0 = N_0 \mathcal{Z} = (n_0 - ik_0) \mathcal{Z},$$

$$y_1 = N_1 \mathcal{Z} = (n_1 - ik_1) \mathcal{Z}.$$

The analysis follows that for absorption-free media. The boundaries are, as before,

1. Electric vector continuous across the boundary

$$\mathcal{E}_i + \mathcal{E}_r = \mathcal{E}_t.$$

2. Magnetic vector continuous across the boundary

$$y_0 \mathcal{E}_i - y_0 \mathcal{E}_r = y_1 \mathcal{E}_t,$$

and by eliminating first \mathcal{E}_t and then \mathcal{E}_r , we obtain the expressions for the amplitude coefficients

$$\rho = \frac{\mathcal{E}_r}{\mathcal{E}_i} = \frac{y_0 - y_1}{y_0 + y_1} = \frac{(n_0 - ik_0) \mathcal{Z} - (n_1 - ik_1) \mathcal{Z}}{(n_0 - ik_0) \mathcal{Z} + (n_1 - ik_1) \mathcal{Z}} = \frac{(n_0 - n_1) - i(k_0 - k_1)}{(n_0 + n_1) - i(k_0 + k_1)}, \quad (2.75)$$

$$\tau = \frac{\mathcal{E}_t}{\mathcal{E}_i} = \frac{2y_0}{y_0 + y_1} = \frac{2(n_0 - ik_0) \mathcal{Z}}{(n_0 - ik_0) \mathcal{Z} + (n_1 - ik_1) \mathcal{Z}} = \frac{2(n_0 - ik_0)}{(n_0 + n_1) - i(k_0 + k_1)}. \quad (2.76)$$

Our troubles begin when we try to extend this to reflectance and transmittance. We remain at normal incidence. Following the method for the absorption-free case, we compute the Poynting vector at the boundary in each medium and equate the two values obtained. In the incident medium, the resultant electric and magnetic fields are

$$\mathcal{E}_i + \mathcal{E}_r = \mathcal{E}_i(1 + \rho)$$

and

$$\mathcal{H}_i - \mathcal{H}_r = y_0(1 - \rho)\mathcal{E}_i,$$

respectively, where we have used the notation for tangential components, and in the second medium, the fields are

$$\tau\mathcal{E}_i \quad \text{and} \quad y_1\tau\mathcal{E}_i,$$

respectively. Then the net irradiances on either side of the boundary are as follows:

- Medium 0: $I = \text{Re}\left\{\frac{1}{2}[\mathcal{E}_i(1 + \rho)][y_0^*(1 - \rho^*)\mathcal{E}_i^*]\right\}$
- Medium 1: $I = \text{Re}\left\{\frac{1}{2}[\tau\mathcal{E}_i][y_1^*\tau^*\mathcal{E}_i^*]\right\}$

We then equate these two values which gives, at the boundary,

$$\text{Re}\left[\frac{1}{2}y_0^*\mathcal{E}_i\mathcal{E}_i^*(1 + \rho - \rho^* - \rho\rho^*)\right] = \frac{1}{2}\text{Re}(y_1)\tau\tau^*\mathcal{E}_i\mathcal{E}_i^*,$$

$$\frac{1}{2}\text{Re}(y_0^*)\mathcal{E}_i\mathcal{E}_i^* - \frac{1}{2}\text{Re}(y_0^*)\rho\rho^*\mathcal{E}_i\mathcal{E}_i^* + \frac{1}{2}\text{Re}[y_0^*(\rho - \rho^*)]\mathcal{E}_i\mathcal{E}_i^* = \frac{1}{2}\text{Re}(y_1)\tau\tau^*\mathcal{E}_i\mathcal{E}_i^*. \quad (2.77)$$

We can replace the different parts of Equation 2.77 with their normal interpretations to give

$$I_i - RI_i + \frac{1}{2}\text{Re}[y_0^*(\rho - \rho^*)]\mathcal{E}_i\mathcal{E}_i^* = TI_i. \quad (2.78)$$

$(\rho - \rho^*)$ is imaginary. This implies that if y_0 is real, the third term in Equation 2.78 is zero. The other terms then make up the incident, the reflected, and the transmitted irradiances, and these balance. If y_0 is complex, then its imaginary part will combine with the imaginary $(\rho - \rho^*)$ to produce a real result that will imply that $T + R \neq 1$. The irradiances involved in the analysis are those actually at the boundary, which is of zero thickness, and it is impossible that it should either remove or donate energy to the waves. Our assumption that the irradiances can be divided into separate incident, reflected, and transmitted irradiances is therefore incorrect. The source of the difficulty is a coupling between the incident and reflected fields which occurs only in an absorbing medium and which must be taken into account when computing energy transport. The expressions for the amplitude coefficients are perfectly correct. The phenomenon is well understood and has been described in a number of contributions, for example, that by Berning [4]. We shall return to the problem in more general terms in Section 2.13 once we have had a brief look at coherence in Section 2.12. It is, however, convenient to look now at some consequences.

The extra term is on the order of (k^2/n^2) . For any reasonable experiment to be carried out, the incident medium must be sufficiently free of absorption for the necessary comparative measurements to be performed with acceptably small errors. In such cases, the error is vanishingly small. Although we will certainly be dealing with absorbing media in thin-film assemblies, our incident media will never be heavily absorbing, and it will not be a serious lack of generality if we assume

that our incident media are absorption free. Since our expressions for the amplitude coefficients are valid, then any calculations of amplitudes in absorbing media will be correct. We simply have to ensure that the calculations of reflectances are carried out in a transparent medium. With this restriction, then, we have

$$R = \left(\frac{y_0 - y_1}{y_0 + y_1} \right) \left(\frac{y_0 - y_1}{y_0 + y_1} \right)^*, \quad (2.79)$$

$$T = \frac{4y_0 \operatorname{Re}(y_1)}{(y_0 + y_1)(y_0 + y_1)^*}, \quad (2.80)$$

where y_0 is real.

2.3.4.1 Rear Surface of Absorbing Substrate

We have avoided the problem connected with the definition of reflectance in a medium with complex y_0 simply by not defining it unless the incident medium is sufficiently free of either gain or absorption. Without a definition of reflectance, however, we have trouble with the meaning of antireflection, and there are cases such as the rear surface of an absorbing substrate where an antireflection coating would be relevant. We do need to deal with this problem, and although we have not yet discussed antireflection coatings, it is most convenient to include the discussion here where we have already the basis for the theory. The discussion was originally published in a paper of Macleod [5].

The usual purpose of an antireflection coating is the reduction of reflectance. But frequently the objective of the reflectance reduction is the corresponding increase in transmittance. Although an absorbing or amplifying medium will rarely present us with a problem in terms of a reflectance measurement, we must occasionally treat a slab of such material on both sides to increase overall transmittance. In this context, therefore, we define an antireflection coating as one that increases transmittance and, in the ideal case, maximizes it. But to accomplish that, we need to define what we mean by *transmittance*.

We have no problem with the measurement of irradiance at the emergent side of our system, even if the emergent medium is absorbing. The incident irradiance is more difficult. We can define this as the irradiance we would measure at the position of the surface if the transmitting structure were removed and replaced by an infinite extent of incident medium material. Then the transmittance will simply be the ratio of these two values.

That is,

$$I_{\text{inc}} = \frac{1}{2} \operatorname{Re}(y_0) \mathcal{E}_i \mathcal{E}_i^*,$$

and then

$$T = \frac{\frac{1}{2} \operatorname{Re}(y_1) \mathcal{E}_t \mathcal{E}_t^*}{\frac{1}{2} \operatorname{Re}(y_0) \mathcal{E}_i \mathcal{E}_i^*}.$$

This is completely consistent with Equation 2.78, that is, with a slight manipulation,

$$T = 1 - \rho\rho^* + \frac{\operatorname{Re}[y_0^*(\rho - \rho^*)]}{\operatorname{Re}(y_0)}. \quad (2.81)$$

An alternative form uses

$$\mathcal{E}_t = \frac{2y_0}{(y_0 + y_1)} \mathcal{E}_i,$$

so that

$$T = \frac{4y_0y_0^* \operatorname{Re}(y_1)}{\operatorname{Re}(y_0) \cdot [(y_0 + y_1)(y_0 + y_1)^*]}. \quad (2.82)$$

Now let the surface be coated with a dielectric system so that it presents the surface admittance Y . We have not, so far, introduced the idea of surface admittance, and we shall deal with it in more detail in Section 2.5. The ratio of the total tangential magnetic amplitude to the total tangential electric amplitude at a surface is an admittance that we can consider to be a property of the surface and call it the surface admittance. In the case of a simple boundary, the surface admittance is simply the characteristic admittance (tilted if necessary) of the emergent medium. In the case of an optical coating, it plays the same role, but its value is now a function of the interference effects in the coating. For the moment, let us accept that the surface admittance of the rear surface of our absorbing substrate, because of the coating, now presents a surface admittance of Y that will be interpreted by the incident wave as if it were a simple surface before a medium of characteristic admittance Y .

Then, since, in the absence of absorption, the net irradiance entering the thin-film system must also be the emergent irradiance,

$$T = \frac{4y_0y_0^* \operatorname{Re}(Y)}{\operatorname{Re}(y_0) \cdot [(y_0 + Y)(y_0 + Y)^*]}. \quad (2.83)$$

Let $Y = \alpha + i\beta$; then

$$T = \frac{4\alpha(n_0^2 + k_0^2)}{n_0[(n_0 + \alpha)^2 + (k_0 - \beta)^2]},$$

and T can readily be shown to be a maximum when

$$Y = \alpha + i\beta = n_0 + ik_0 = (n_0 - ik_0)^*. \quad (2.84)$$

The matching admittance should therefore be the *complex conjugate* of the incident admittance. For this perfect matching, the transmittance becomes

$$T = \left(1 + \frac{k_0^2}{n_0^2}\right),$$

and this is greater than unity. This is not a mistake but rather a consequence of the definition of transmittance. Irradiance falls by a factor of roughly $4\pi k_0$ in a distance of one wavelength, rather larger than any normal value of k_0^2/n_0^2 , so that the effect is quite small. It originates in a curious pattern in the otherwise exponentially falling irradiance. It is caused by the presence of the interface and is a cyclic fluctuation in the rate of irradiance reduction. Note that the transmittance is unity if the coating is designed to match $n_0 - ik_0$ rather than its complex conjugate.

A dielectric coating that transforms an admittance of y_1 to an admittance of y_0^* will also, when reversed, exactly transform an admittance of y_0 to y_1^* . This is dealt with in more detail later (Section 8.6) when induced transmission filters are discussed. Thus, the optimum coating to give highest transmittance will be the same in both directions. This implies that an absorbing substrate

in identical dielectric incident and emergent media should have exactly similar antireflection coatings on both front and rear surfaces.

Although also a little premature, it is convenient to mention here that the calculation of the properties of a coated slice of material involves multiple beams that are combined either coherently or incoherently. The coherent case considers the slice as an ordinary absorbing thin film and is simply the usual interference calculation, and we will return to it in considerable detail when we deal with induced transmission filters. We will see then that as the absorbing film becomes thicker, the matching rules for an induced transmission filter tend to approach Equation 2.84. The incoherent case is at first sight less obvious. An estimate of the reflected beam is necessary for a multiple beam calculation. Such calculations imply that the absorption is not sufficiently high to completely eliminate a beam that suffers two traversals of the system. This implies, in turn, a negligible absorption in the space of one wavelength, in other words, $4\pi k_0$ is very small. The upper limit on the size of the effect under discussion is k_0^2/n_0^2 , and this will still be less significant. For an incoherent calculation to be appropriate, there must be a jumbling of phase that washes out its effect. We can suppose for this discussion that the jumbling comes from a variation in the position of the reflecting surface over the aperture, although in the normal way, there will also be some variation of the incident angle. The variation of the extra term in Equation 2.84 is locked for its phase to the reflecting surface, and so at any exactly plane surface that may be chosen as a reference, an average of the extra term is appropriate, and this will be zero because ρ will have a phase that varies throughout the four quadrants. For multiple beam calculations, therefore, the reflectance can be taken simply as $\rho\rho^*$. Where k_0^2/n_0^2 is significant, the absorption will be very high and certainly enough for the influence of the multiple beams to be automatically negligible.

2.3.5 Oblique Incidence in Absorbing Media

Remembering what we said in the Section 2.3.4, we limit ourselves to a transparent incident medium and an absorbing second, or emergent, medium. Our first aim must be to ensure that the phase factors are consistent. Taking advantage of some of the earlier results, we can write the phase factors as follows:

- Incident:

$$\exp\{i[\omega t - (2\pi n_0/\lambda)(x \sin \vartheta_0 + z \cos \vartheta_0)]\}$$

- Reflected:

$$\exp\{i[\omega t - (2\pi n_0/\lambda)(x \sin \vartheta_0 - z \cos \vartheta_0)]\} \quad (2.85)$$

- Transmitted:

$$\exp\{i[\omega t - (2\pi\{n_1 - ik_1\}/\lambda)(\alpha x + \gamma z)]\},$$

where α and γ in the transmitted phase factors are the only unknowns. The phase factors must be identically equal for all x and t with $z = 0$. This implies

$$\alpha = \frac{n_0 \sin \vartheta_0}{(n_1 - ik_1)},$$

and since $\alpha^2 + \gamma^2 = 1$,

$$\gamma = (1 - \alpha^2)^{1/2}.$$

There are two solutions to this equation, and we must decide which is to be adopted. We note that it is strictly $(n_1 - ik_1)\alpha$ and $(n_1 - ik_1)\gamma$ that are required:

$$\begin{aligned} (n_1 - ik_1)\gamma &= [(n_1 - ik_1)^2 - n_0^2 \sin^2 \vartheta_0]^{1/2} \\ &= [n_1^2 - k_1^2 - n_0^2 \sin^2 \vartheta_0 - i2n_1k_1]^{1/2}. \end{aligned} \quad (2.86)$$

The quantity within the square root is in either the third or fourth quadrant, and so the square roots are in the second quadrant (of the form $-a + ib$) and in the fourth quadrant (of the form $a - ib$). If we consider what happens when these values are substituted into the phase factors, we see that the fourth quadrant solution must be correct because this leads to an exponential falloff with z of amplitude together with a change in the phase of the correct sense. The second quadrant solution would lead to an increase with z and a change in phase of the incorrect sense, which would imply a wave travelling in the opposite direction. The fourth quadrant solution is also consistent with the solution for the absorption-free case. The transmitted phase factor is therefore of the form

$$\begin{aligned} \exp\{i[\omega t - (2\pi n_0 \sin \vartheta_0 x/\lambda) - (2\pi/\lambda)(a - ib)z]\} \\ = \exp(-2\pi bz/\lambda) \exp\{i[\omega t - (2\pi n_0 \sin \vartheta_0 x/\lambda) - (2\pi az/\lambda)]\}, \end{aligned}$$

where

$$(a - ib) = [n_1^2 - k_1^2 - n_0^2 \sin^2 \vartheta_0 - i2n_1k_1]^{1/2}.$$

A wave which possesses such a phase factor is known as inhomogeneous. The exponential falloff in amplitude is along the z -axis, while the propagation direction in terms of phase is determined by the direction cosines, which can be extracted from

$$(2\pi n_0 \sin \vartheta_0 x/\lambda) + (2\pi az/\lambda).$$

The existence of such waves is another good reason for our choosing to consider the components of the fields parallel to the boundary and the flow of energy normal to the boundary.

We should note at this stage that provided we include the possibility of complex angles, the formulation of the absorption-free case applies equally well to absorbing media, and we can write

$$\begin{aligned} (n_1 - ik_1) \sin \vartheta_1 &= n_0 \sin \vartheta_0, \\ \alpha &= \sin \vartheta_1, \\ \gamma &= \cos \vartheta_1, \\ (a - ib) &= (n_1 - ik_1) \cos \vartheta_1. \end{aligned}$$

The calculation of amplitudes follows the same pattern as before. However, we have not previously examined the implications of an inhomogeneous wave. Our main concern is the calculation of the tilted admittance connected with such a wave. Since the x , y , and t variations of the wave are contained in the phase factor, we can write

$$\begin{aligned} \text{curl} &\equiv \left(\frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial x} \mathbf{j} + \frac{\partial}{\partial x} \mathbf{k} \right) \times \\ &\equiv \left(-i \frac{2\pi N}{\lambda} \alpha \mathbf{i} - i \frac{2\pi N}{\lambda} \gamma \mathbf{k} \right) \times \end{aligned}$$

and

$$\frac{\partial}{\partial t} \equiv i\omega,$$

where the k is a unit vector in the z -direction and should not be confused with the extinction coefficient k .

For p -waves, the \mathbf{H} vector is parallel to the boundary in the y -direction, so $\mathbf{H} = \mathcal{H}_y \mathbf{j}$. The component of E parallel to the boundary will then be in the x -direction $\mathcal{E}_x \mathbf{i}$. We follow the analysis leading up to Equation 2.26 and as before

$$\begin{aligned} \text{curl} \mathbf{H} &= \sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t} \\ &= (\sigma + i\omega\varepsilon) \mathbf{E} \\ &= \frac{i\omega N^2}{c^2 \mu} \mathbf{E}. \end{aligned}$$

Now the tangential component of $\text{curl} \mathbf{H}$ is in the x -direction so that

$$-i \frac{2\pi N}{\lambda} \gamma (\mathbf{k} \times \mathbf{j}) \mathcal{H}_y = i \frac{\omega N^2}{c^2 \mu} \mathcal{E}_x \mathbf{i}.$$

But

$$-(\mathbf{k} \times \mathbf{j}) = \mathbf{i},$$

so that

$$\begin{aligned} \eta_p &= \frac{\mathcal{H}_y}{\mathcal{E}_x} = \frac{\omega N \lambda}{2\pi c^2 \mu \gamma} = \frac{N}{c \mu \gamma} \\ &= \frac{N \mathcal{Y}}{\gamma} = \frac{y}{\gamma}. \end{aligned}$$

For the s -waves, we use

$$\text{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu \frac{\partial \mathbf{H}}{\partial t}.$$

E is now along the y -axis, and a similar analysis to that for p -waves yields

$$\eta_s = \frac{\mathcal{H}_x}{\mathcal{E}_y} = N \mathcal{Y} \gamma = y \gamma. \quad (2.87)$$

Now γ can be identified as $\cos \vartheta$, provided that ϑ is permitted to be complex, and so

$$\begin{aligned} \eta_p &= y / \cos \vartheta, \\ \eta_s &= y \cos \vartheta. \end{aligned} \quad (2.88)$$

Alternatively, we can use the expressions in Equations 2.86 and 2.87, together with the fact that $y = (n - ik) \mathcal{Z}$, to give

$$\eta_s = \mathcal{Z} [n_1^2 - k_1^2 - n_0^2 \sin^2 \vartheta_0 - i2n_1 k_1]^{1/2}. \quad (2.89)$$

The fourth quadrant being the correct solution, and then

$$\eta_p = \frac{y^2}{\eta_s}. \quad (2.90)$$

This second form is completely consistent with Equation 2.28 but avoids any problems with the quadrant. Then the amplitude and irradiance coefficients become as before:

$$\rho = \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1}, \quad (2.91)$$

$$\tau = \frac{2\eta_0}{\eta_0 + \eta_1}, \quad (2.92)$$

$$R = \left(\frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right) \left(\frac{\eta_0 - \eta_1}{\eta_0 + \eta_1} \right)^*, \quad (2.93)$$

$$T = \frac{4\eta_0 \operatorname{Re}(\eta_1)}{(\eta_0 + \eta_1)(\eta_0 + \eta_1)^*}, \quad (2.94)$$

And, of course, these expressions are valid for absorption-free media as well.

2.4 Critical Angle and Beyond

Let us consider the case of a simple boundary between two dielectric materials with refractive indices n_0 and n_1 and corresponding characteristic admittances y_0 and y_1 . Given ϑ_0 , Snell's law (Equation 2.44) allows us to calculate the corresponding propagation angle ϑ_1 in the second medium:

$$n_0 \sin \vartheta_0 = n_1 \sin \vartheta_1.$$

Suppose, however, that n_0 is greater than n_1 . Since $\sin \vartheta_1 = (n_0/n_1) \sin \vartheta_0$, as the angle ϑ_0 increases from 0° to 90° , there comes a point where $\sin \vartheta_1$ becomes greater than unity. The value of ϑ_0 where this first occurs is known as the critical angle. Beyond the critical angle, ϑ_1 is imaginary. The critical angle is given by

$$\vartheta_0 = \arcsin\left(\frac{n_1}{n_0}\right). \quad (2.95)$$

It is difficult to visualize an imaginary angle. Fortunately, that is not necessary. What we actually need is the cosine of the angle, and if we already know the sine, a simple relationship gives us the cosine:

$$\cos^2 \vartheta_1 = 1 - \sin^2 \vartheta_1 = 1 - \frac{n_0^2}{n_1^2} \sin^2 \vartheta_0. \quad (2.96)$$

It is not only simply $\cos \vartheta_1$ that is required, but also $n_1 \cos \vartheta_1$. Then,

$$n_1 \cos \vartheta_1 = \sqrt{n_1^2 - n_0^2 \sin^2 \vartheta_0} = i\sqrt{n_0^2 \sin^2 \vartheta_0 - n_1^2}, \quad (2.97)$$

where the quantity under the root sign is positive and where, as yet, we have not assigned a positive or negative character to the root. Equation 2.86 comes to our rescue. We are actually solving for the direction cosine γ times the refractive index, a combination that appears in the spatial part of the phase factor. We have already established that even in the case of miniscule absorption, we should take the fourth quadrant solution rather than the second in Equation 2.86 and that applies to Equation 2.97 so that

$$n_1 \cos \vartheta_1 = -i\sqrt{n_0^2 \sin^2 \vartheta_0 - n_1^2} = \gamma n_1. \quad (2.98)$$

The phase factor of the transmitted wave is then

$$\begin{aligned} & \exp \left\{ i \left[\omega t - \frac{2\pi}{\lambda} (n_0 \sin \vartheta_0) x - \frac{2\pi}{\lambda} \left(-i\sqrt{n_0^2 \sin^2 \vartheta_0 - n_1^2} \right) z \right] \right\} \\ & = \exp \left\{ -\frac{2\pi}{\lambda} \left(\sqrt{n_0^2 \sin^2 \vartheta_0 - n_1^2} \right) z \right\} \exp \left\{ i \left[\omega t - \frac{2\pi}{\lambda} (n_0 \sin \vartheta_0) x \right] \right\}. \end{aligned} \quad (2.99)$$

The wave is known as evanescent. It is pinned to the surface along which it propagates with a wavelength smaller than could be supported by a progressive wave, and it decays exponentially away from the surface in the z -direction. The tilted admittance is imaginary, and so the component of the magnetic field that also propagates in the z -direction is 90° out of phase with the electric field so that no energy actually propagates in the z -direction. It is not surprising, therefore, that since there is no loss in the emergent medium, the reflectance is total. It is usually termed *total internal reflectance* and is sometimes abbreviated to TIR.

The tilted admittances, in free space units, are given by

$$\eta_s = -i[n_0^2 \sin^2 \vartheta_0 - n_1^2]^{1/2} \quad (2.100)$$

and

$$\eta_p = \frac{y^2}{\eta_s} = +i \frac{n_0^2}{[n_0^2 \sin^2 \vartheta_0 - n_1^2]^{1/2}}. \quad (2.101)$$

It is easy to see that since both are imaginary, the reflectance for both p - and s -polarizations will be 100% (or unity in absolute terms). However, the phase change on reflection will vary with the angle of incidence. At the critical angle where η_s is zero and η_p is infinite, it is zero for s -polarization and 180° for p -polarization. The s -polarization phase shift then moves through the first into the second quadrant with increasing incidence while the p -polarization phase shift moves through the third into the fourth quadrant. The relative retardation, or delta, we recall, has to be corrected by 180° to take account of the sign convention, and it rises from zero to a maximum value that depends on the incident index and then drops back to zero at grazing incidence.

Figure 2.7 shows calculated properties for a glass incident medium (1.52 index) and air emergent medium as a function of angle of incidence.

Strictly, the critical angle is a phenomenon that is related to completely dielectric media only. However, if the k value of the emergent medium is quite small, then a reference angle, usually simply referred to as the critical angle, is sometimes defined by dropping k and using Equation 2.95. This is an especially frequent practice when gain rather than absorption is involved (gain is covered in Chapter 11), but it should be used with caution because it may be thought to imply some kind of abrupt transition when there is none. Otherwise, the behavior of absorbing materials even at high angles of incidence in a high-index incident medium is already covered in Section 2.3.5.