Ellipsometry at the Nanoscale

Maria Losurdo · Kurt Hingerl
Editors

Ellipsometry at the Nanoscale

Springer

Editors

Maria Losurdo
CNR Bari, Ist. Metodologie Inorganiche
Università Bari
Bari
Italy

Kurt Hingerl
Zentrum für Oberflächen- und
     Nanoanalytik (ZONA)
Universität Linz
Linz
Austria

Preamble

Dear Reader,

Probably you are thinking: "Why another book about spectroscopic ellipsometry and polarimetry, if there are classical books from Azzam and Bashara (*Ellipsometry and Polarized light*) [1], Fujiwara (*Spectroscopic Ellipsometry*: *Principles and Applications*) [2], Schubert (*Infrared Ellipsometry on Semiconductor Layer Structures*: *Phonons*, *Plasmons*, *and Polaritons*) [3], and others, including the "*Handbook on Ellipsometry*" [4], by Tompkins and Irene?"

   The main reason is that scientists, including ourselves, start in exploiting ellipsometry for characterizing structured, plasmonic, stochastic, self-assembled, and engineered nanoscale systems, where the Fresnel equations are not strictly applicable any more.

   This raises a list of puzzling questions: how accurate is the classical description representing nanomaterials? What is the limit of applying an effective medium theory to nanostructures? What is the minimum size for which nonlocal and quantum size effects predominate in the nanostructures? What are the effects of size and shape of nanostructures on their dielectric function? How do physics and chemistry change at the nanoscale?

   The opportunity to gather experts to share opinions and find answers seemed too good to miss.

   Indeed, it turns out experimentally that ellipsometry and polarimetry are extremely powerful for monitoring in situ and *in real time* nanoscale processes and characterizing ex situ nanostructured materials. Therefore, to provide the community with different aspects and opinions on this rapidly evolving field, we have decided to edit this book to critically present and review recent applications of polarimetric techniques to nanomaterials.

In order to catch the spirit of this book, let us recall the international accepted definition of nanotechnology[1]:

> **Definition of Nanotechnology (Set Out in Nanotechnology Research Directions, 1999) (see footnote 1)**
>
> Nanotechnology is the ability to control and restructure the matter at the atomic and molecular levels in the range of approximately 1–100 nm, and exploiting the distinct properties and phenomena at that scale as compared to those associated with single atoms or molecules or bulk behaviour. The aim is to create materials, devices, and systems with fundamentally new properties and functions by engineering their small structure. This is the ultimate frontier to economically change materials properties, and the most efficient length scale for manufacturing and molecular medicine. The same principles and tools are applicable to different area of relevance and may help establish a unifying platform for science, engineering, and technology at the nanoscale. The transition from single atoms or molecules behaviour to collective behaviour of atomic and molecular assemblies is encountered in nature, and nanotechnology exploits this natural threshold.

Establishing deepened knowledge at the nanoscale has been the main focus of the nanotechnology research community in the last decade. In this context, in 2008, the European Commission funded, within the 7th Framework Programme, a cross-disciplinary project entitled "*NanoCharM*."

The NanoCharM coordination action brought together a critical mass of experts and committed and passionate people who believe in a future with better understanding and control of nanomaterials and related processes achievable by exploiting the potential of noninvasive polarization-based optical techniques of ellipsometry and polarimetry. NanoCharM moved ideas from the laboratory into practice and industry through the commitment to transform knowledge and expertise in materials and diagnostics into new products, processes, and devices, according to nanotechnology definition.

This book evolved from the NanoCharM extensive program of training courses, workshops, roadmapping, and networking events that created broad awareness of the many advantages of ellipsometry characterization. We were glad to see that students and colleagues from many different areas found interest in ellipsometry. This book also contains and reflects the spirit advanced and matured of learning

---

[1] International Standards Organization (ISO), TC 229: Nanotechnologies (2010), http://www.iso.org/iso/iso_technical_committee.html?commid=381983
M.C. Roco, R.S. Williams, P. Alivisatos (eds.), *Nanotechnology Research Directions: Vision for the Next Decade. IWGN Workshop Report 1999* (National Science and Technology Council, Washington, DC, 1999), Available online: http://www.wtec.org/loyola/nano/IWGN, Research. Directions/. Published by Kluwer, currently Springer, 2000.

from the experience of growing and characterizing nanomaterials and thin films, to gain the capacity to face new challenges in material characterization at the nanoscale.

In essence, it keeps alive the aim of the NanoCharM project to develop ellipsometry and polarimetry measurements and control procedures required to mass produce nanomaterials, which are critical factors for the successful transition from ideas to commercial products.

This book cannot be comprehensive and does not reprint information already included in the classical textbooks on ellipsometry; indeed, it gives an up-to-date introduction to and an overview about the ongoing research and need in nanomaterials characterization.

Who knows what new concepts might emerge in our "*nano-minds*" and where the important applications will be? We hope that this book inspires you to apply polarimetric techniques for your research approach and that its contents facilitate the interpretation of the measurements in terms of chemical and physics phenomena.

Writing this book would not have been possible without the very helpful input and inspiration that we received from many students and colleagues, especially those who lively contributed to the NanoCharM schools and workshops. We are especially grateful to Marcin Sadowski, a colleague and the motivating scientific officer of the European Commission (Directorate General for Research), who followed the NanoCharM project also with scientist′s spirit asking us intriguing and challenging questions and giving us many enlightening suggestions. We thank all scientists who inspired the attendants of the NanoCharM schools and workshops with their lectures and contributions. Just to name a few, far from being comprehensive, Prof. Manuel Cardona (Max-Planck Institute, Suttgart), Prof. Wolfgang Richter (University of Rome), Prof. Dr. Friedhelm Bechstedt (University Jena), Prof. Bernard Drevillon (Ecole Polytechnique, LPICM), Dr. Ramdane Benferhat (Horiba-JY), Prof. Dr. Hans Arwin (University Linkoping), Prof. Dr. Joachim Krenn (University of Graz), Dr. Fritz Keilmann (Munich-Centre for Advanced Photonics), Prof. Norbert Esser (ISAS-Berlin). We are also very thankful to all authors and really appreciate the dedication in providing insights into current and future trends in ellipsometry. Finally, we thank the publisher, Springer-Verlag, especially Mayra Castro for her invaluable and unwavering commitment.

Maria Losurdo
Kurt Hingerl

References

1. R.M.A. Azzam, N.M. Bashara, *Ellipsometry and Polarized Light*. (Elsevier Science Publishers B.V., Amsterdam, 1989)
2. H. Fujiwara, *Spectroscopic Ellipsometry*: *Principles and Applications*. (Wiley, West Sussex, 2007)
3. M. Schubert, in *Infrared Ellipsometry on Semiconductor Layer Structures*: Phonons, *Plasmons*, *and Polaritons*. Springer Tracts in Modern Physics. (Springer, Berlin, 2010)
4. H.G. Tompkins, E.A. Irene, *Handbook of Ellipsometry*. (William Andrew Inc.- Springer Verlag GmbH, Norway, 2005)

Preface

"*Light has a relation to the matter which it meets with in its course, and is affected by it, being reflected, deflected, transmitted, refracted, and absorbed by particles very minute in their dimensions. At present the most instructed persons are, I suppose, very far from perceiving the full and close coincidence between all the facts of light interacting with particles and the physical account of them which the theory supplies.*" This is how Michael Faraday started his "Bakerian Lecture: Experimental Relations of Gold (and Other Metals) to Light" published in 1857 in the Philosophical Transactions of the Royal Society. More than a century and a half has passed, and despite the enormous advances in understanding the interaction of light with nanostructured materials, this sentence still expresses a way of thinking common to those who enter the world of physical investigations of the nature and action of a ray of light with nanostructures that were of the order of "*1/282000th of an inch*", which we say 35 nm nowadays!

Therefore, the term nanotechnology is new, but research at the nanometer scale is not new at all!

Richard Feynman's renowned lecture "There's plenty of room at the bottom" (1959 at Caltech) is seen as the milestone of nanoscience. Noteworthily, Feynman also presciently recognized the synergy between the ability to make things smaller and the ability to see and probe what has been made as the key to progress in the field of nanoscience and nanotechnology.

Indeed, almost 15 years earlier, in 1945, Alexandre Rothen made the pioneering statement that "*during the course of an investigation of the reaction between films of antigens and antibodies, it was found that an apparatus was needed which would measure film thickness rapidly and with an accuracy of at least* 1 Å" (which is a fraction of a nanometer!). The apparatus that was devised was given the name "ellipsometer" [1]. Already in 1945, the ellipsometer was capable of measuring a film thickness to within ±0.3 Å!

Therefore, we could say that ellipsometry was pioneering nanoscience, where both light and matter play leading roles.

In the past few years, entirely new classes of materials have been discovered and developed. These include one-dimensional nanowires and quantum dots of

various compositions, polyvalent noble metal nanostructures, graphene, metamaterials, superlattices, and a wide variety of other nanoparticle assemblies. A toolbox of nanostructures is being engineered with tailored, functional optical properties and colors for optics, photonics, and biomedical applications ranging from therapeutics to diagnostics. The need for a wider understanding of nanoparticle and nanostructures exploiting characterization techniques down to the level of 0.1–100 nm is the driving force of significant changes in optical metrology, since, from the optical point-of-view, the interaction of materials with photons is strongly dependent not only on the chemistry but also on structure, size, and shape, which can then be used to control light propagation.

The versatile nature of ellipsometry as a functional, nanoscale sensitive, and nondestructive technique, is paving the way for the application of these new nanostructures in a widening field of technologies and for breakthroughs in knowledge of thin film multilayer surfaces, composite and smart materials, and materials engineering at the nanoscale.

Scope of the Book

The primary aim of this book is to present and introduce ellipsometry in nanoscience and nanotechnology making a bridge between the classical and nanoscale optical behavior of materials. The progress in the current understanding of optical properties of nanomaterials is an important driving force for developing a variety of applications.

This book helps to delineate the role of the nondestructive and noninvasive optical diagnostics of ellipsometry in improving science and technology of nanomaterials and related processes by illustrating its exploitation ranging from fundamental studies of the physics and chemistry of nanostructures to the ultimate goal of turnkey manufacturing control.

This book is written for a broad readership: materials scientists, researchers, engineers, as well as students, and nanotechnology operators who want to deepen their knowledge about both basics and applications of ellipsometry to nanoscale phenomena. Readers might have quite different backgrounds, interests, and understanding of optics, physics, materials, and of their properties and technologies, and, despite the difficulty of having a single book addressing a varied audience, we have at least one chapter of interest to everyone!

Therefore, this book goes through different levels. It starts as a general introduction for people curious to enter the fields of ellipsometry and polarimetry applied to nanomaterials and progresses to articles by experts on specific fields that span from plasmonics, optics, to semiconductors and flexible electronics. The core belief reflected in this book is that ellipsometry applied at the nanoscale offers new ways of addressing many current needs.

The book also explores forward-looking potential applications. The potential of ellipsometry is not yet fully used, and it is currently the only optical technique

which can reliably provide phase information. Especially, the phase is sensitive to e.g., minute changes in the thickness of overlayers, or adsorbates or structural changes in nanomaterials. The competing interferometric measurements still lack solutions to the problems of being constrained to normal incidence, of managing the dispersion of the optical components and also vibrational control.

The question is: how to move forward?

The intrinsic difficulty of ellipsometry in nanotechnology is that it is a multi-disciplinary field involving optics; therefore, chemists, physicists, material scientists, engineers, molecular biologists, pharmacologists, etc., should all be familiar with the basic optics concepts in those heterogeneous fields, which is sometime not straightforward.

Therefore, the goal was not to write a monograph style book, and even less a textbook, but a book with contributions from multidisciplinary fields, with different approaches and languages that different readers with optic vocabularies imbibed in their scientific infancy can become acquainted with. The principles of ellipsometry are not always seen as easy. For many researchers, the lack of knowledge on polarized light and on the Maxwell equations—with specific boundary conditions, especially for nanomaterials—make spectroscopic ellipsometers seem mystic devices, providing (pseudo-)dielectric functions. Furthermore, the meaning of the ellipsometric measurables Ψ and Δ may not be straightforward and modeling procedures are required to derive the dimensional and optical properties of the materials. Therefore, the primary goal of this book is to provide a common ground for a better understanding of how to use ellipsometry effectively. The position reflected in this book is that this goal can only be accomplished by materials scientists, optics scientists, process engineers, and nanotechnology analysts exchanging background and expertise and working together on a found basis.

Excellent books on the fundamental of ellipsometry already exist and for more detailed description of the principles of polarized lights and ellipsometry the reader is referred to the books "*Ellipsometry and Polarized light*" edited by R. M. A. Azzam and N. M. Bashara, "*Handbook of ellipsometry*" edited by H. G. Tompkins and E. A. Irene.

Organization of the Book

This book captures the interdisciplinary nature of nanoscience and provides a balanced approach to introduce the principles of ellipsometry and exploit them in various fields serving as both an education and training text and as a reference point for research and development providing the following unique features:

- A unifying vision of ellipsometry merging nanotechnologies, including ellipsometry instrumentation and modeling best practice, fabrication processes, nanomaterials, applications, and technologies

- A full perspective about the various information that can be gained depending on the spectral range and ellipsometry configuration (from the Terahertz to near-infrared, visible and far-UV)
- A multifaceted study of novel chemical, electrochemical, and optical phenomena in nanostructures
- A coverage of inorganic and organic semiconductor nanomaterials, superconductors, biomaterials, and nanocomposites, as well as graphene optically characterised by ellipsometry
- A focus on new technologies based on the interaction of light with nanomaterials such as *plasmonics and metamaterials*
- A coverage of ellipsometry in bioresearch and medicine
- A critical and comprehensive assessment of ellipsometry in the industry and in the market place, with future forecasts.

We took care of organizing this book and sorted and clustered the contributions with respect to topics without trying to squeeze all authors into a common frame. The story is not only about the things; it is also about the people! It is peoples' background and language that make them communicative. Therefore, because of multidisciplinarity of nanoscience, each author wrote each chapter in a self-standing way, certainly referring to other chapters, but the nomenclature has not been unified all over the book.

Therefore, as the Editors of this book, we are deeply grateful to all contributing authors for their efforts and their willingness to share recent results within the framework of "nanoscience and nanotechnology". We are especially proud that the authorship includes pioneers and newcomers to this intriguing and fertile field of research. With chapters addressing fundamental and practical questions of physics, chemistry, quantum theories, and real-time monitoring of fabrication processes related to nanostructures, this book shows the reader how ellipsometry can help to achieve a better orientation in nanoscale optical phenomena.

Most readers will use the book to get a solid grasp of the fundamentals, so that they can move on to more complex topics. Some of the chapters can be read independently of the others, on the assumption that the fundamental in Chaps. 1–5 have been fully assimilated. The reader is just left to be driven by his/her curiosity and interests!

Therefore, Chap. 1 has been devoted to introducing the main concepts of the ellipsometry technique and its historical context, also related to nanomaterials. This book does not presuppose that the reader has a working knowledge of ellipsometry, therefore a beginner can grasp the fundamentals needed in this chapter.

Chapter 2 takes the fundamentals a bit further, letting the reader become more confident with the polarimetric properties of a sample and, consequently, moving from standard ellipsometry to generalized ellipsometry and Mueller polarimetry; it also deals with the instrumental aspects of ellipsometry and polarimetry, giving a perspective on "how" ellipsometry and polarimetry measure what they measure. Recent developments exploiting Mueller polarimetry in fields as diverse as sub-wavelength grating metrology and cancer detection in biomedicine are discussed.

The book then turns to fundamental optical properties and modeling advances needed to explore the nanoscale and to application-oriented considerations. Therefore, Chap. 3 is dedicated to the optical properties of materials consisting of a matrix with inclusions and layered materials. Various *effective medium theories* (EMAs) are available, the result of which is the so-called "effective dielectric function" that describes the macroscopic optical response of a heterogeneous system. Both strengths and weaknesses of the various EMAs are described in this chapter.

One of the most common applications of EMAs is the analysis of surface roughness from ellipsometric spectra. Surfaces and interfaces are also important in explaining nanomaterials' behavior. In bulk materials, only a relatively small percentage of the atoms will be at or near a surface or interface. In nanomaterials, half or more of the atoms are near interfaces and at the surface. Therefore, a critical review of major results in describing surface roughness with effective medium theories is presented in Chap. 4.

Chapter 5 extends the discussion on EMAs to plasmonic materials and introduces the "plasmonics" cluster of chapters that extends from Chaps. 5–9. The fundamental relationship between the dielectric function, ellipsometry, and plasmonic materials is given in Chap. 5, while Chap. 6 enters into the details of the optical characterization of substrate-supported nanostructured noble metal nanoparticles, extending the discussion on the validity of EMAs to the Thin Island Film theory. With Chap. 7, which presents a review of fabrication, modeling and characterization aspects related to the fascinating field of metallic periodic nanostructures, we extend the discussion also to metamaterials. Chapter 8 overviews the merging of Mueller polarimetry and rigorous coupled wave analysis, described in details in the previous chapter, for the analysis of periodic nanostructures. Chapter 9 deals with magnetic plasmonic nanocomposites giving some examples of applications of magneto-optical Kerr spectroscopy.

Chapter 10 interrelates the discussions of the different aspects of standard ellipsometry, generalized ellipsometry, and Mueller polarimetry to the measurement and analysis of exotic and fascinating shapes of biaxial nanostructures. A complete discussion on the appropriateness of generalized ellipsometry for the determination of principal optical constants of chiral and achiral multifold and helical sculptured thin films is given.

Extending the spectral range to Terahertz ellipsometry is the innovative subject of Chap. 11, since there exists a wealth of fascinating excitation mechanisms with eigenfrequencies in the THz domain in condensed and soft matter, such as spin transitions, collective modes of biological molecules, local free charge carrier oscillations, dynamic motion of magnetic domains, ferroelectric domains, or collective charge phenomena, which are discussed here as examples.

Chapter 12 focuses on infrared ellipsometry, which is proven to be a powerful technique for the studies of electronic excitations and lattice vibrations in both the normal and superconducting states and, therefore, for highlighting superconducting phenomena in nanomaterials.

Exploiting light in the NIR–VIS–UV spectral range, which is common to all commercial ellipsometers, to probe in real-time various nanoscale phenomena at surfaces and interfaces is the subject of Chaps. 13 and 14. Many nanoscale characterization methods cannot probe samples in native or desired operating (*operando*) environments. Enabling nanoscale materials real-time analysis under realistic conditions is a critical need. Chapter 13 reviews the application of real-time ellipsometry to probe charge transfer processes in surface-nanoparticle—molecule coupled systems of interest for photonics, molecular electronics and sensing, while the detection of phenomena and kinetics occurring at the solid–liquid interface, which is an important field in electrochemistry, is overviewed in Chap. 14.

Having learned that ellipsometry has all the hardware, instrumental setups, and modeling capabilities to address the process-nanostructure—optical properties interrelationship, a series of examples of the exploitation of ellipsometry in research and industry of both organic and inorganic semiconductors is given in the cluster of Chaps. 15–19.

One possible strategy toward functional optoelectronic composites for OLEDs and organic flexible solar cells is to use an organic material that can efficiently harness photons from light and convert them to useful energy. Therefore, issues related to the optical characterization of organic, polymeric layers combined with inorganic materials are reviewed in Chap. 15.

Another domain that emerged two decades ago and which is another hot topic, concerns the optical properties of quantum structures to make biological tags, efficient light emitting diodes (LEDs), efficient solar cells, or low-consumption flat panel displays. The semiconductor industry is also working on materials for "beyond CMOS" devices. Graphene is the most prevalent example of this. Ellipsometry applied to the analysis of graphene is discussed in Chap. 16 while Chap. 17 focuses on semiconductor nanocrystals.

We then turn to industrial quality control: Spectroscopic ellipsometry is the only method that can be used to measure inline multiple thicknesses on fully fabricated CMOS chips on a test area of 50 x 100 μm. One of the advantages of this technique is that standard uncertainty values of the thickness found using SE are typically between 0.01 and 0.05 nm and, therefore, considerably less than those found with the other electrophysical techniques. Therefore, inline applications of ellipsometry to the semiconductor industry are presented in Chap. 18.

Chapter 19 looks at industry and market perspectives, concluding with presenting capabilities and ideas for exploiting ellipsometry in several industries.

Chapter 20 offers a broad perspective of the concept that "the ideal" characterization tool for nanomaterials does not exist, but corroborating techniques must be used. Therefore, this chapter shows how ellipsometry can corroborate, and be corroborated by, other characterization techniques. Most commonly used structural and chemical characterization methods are introduced to corroborate ellipsometry. Structural characterization methods include scanning or transmission electron microscopies (SEM/TEM), and atomic force or scanning tunneling probe

microscopies (AFM/STM). Chemical characterization methods include electron spectroscopies (XPS).

We conclude with Chap. 21 analyzing the influence and the role that nanomaterials and nanotechnologies might or might not effectively play in the *fuzzy* future. This chapter reflects the perspective of nanotechnologists and market-product developers—a perspective that it would be well worth for scientists and engineers interested in fundamental knowledge as well as in applications of nanomaterials to know about.

This book is timely in proposing the state-of-the-art ellipsometry applications to nanomaterials and pointing the way to further exciting developments. We are just at the transition from the first foundational phase of nanotechnology (2001–2010), which was focused on interdisciplinary research at the nanoscale, on the discovery of new phenomena, properties, and functions at the nanoscale, and on the synthesis of a library of components as building blocks for potential future applications, tool advancement, to the second phase (2011–2020), which will be focused on nano-scale science and engineering integration, projected toward direct measurements with good time resolution and science-based design of fundamentally new products.

The transition from the Nano-1 to the Nano-2 phase is focused on achieving direct measurements at the nanoscale, and science-based design of nanomaterials and nanosystems. It is in this context that ellipsometry, with its nondestructive real-time capability of monitoring processes and tailor materials characteristics can play a role in the general purpose of science and technology integration!

Therefore, the Editors hope to contribute with this book to a wider use of ellipsometry in the nanomaterials community, since it is our common vision and experience as a chemist (*ML*) and as a physicist (*KH*) that many questions/problems/issues can be tackled with this technique.

<div align="right">

Maria Losurdo
Institute of Inorganic Methodology and of Plasmas
Italian National Council of Research
Italy

Kurt Hingerl
Center for Surface and Nanoanalytics
University of Linz
Austria

</div>

Reference

1. A. Rothen, *The review of Scientific Instruments*, **16**, 26 (1945)

Contents

Contributors

David E. Aspnes Department of Physics, North Carolina State University, Raleigh, NC 27695-8202, USA; Department of Physics, Kyung Hee University, Seoul 130-701, Republic of Korea, e-mail: aspnes@unity.ncsu.edu

Michael Bergmair ZONA-Center for Surface and Nanoanalytics, Johannes Kepler University, Linz, Austria, e-mail: michael.bergmair@jku.at

April S. Brown Electrical and Computer Engineering Department, Duke University, Durham, NC 27708, USA, e-mail: abrown@ee.duke.edu

Giovanni Bruno Institute of Inorganic Methodologies and of Plasmas, Italian National Council of Research, Via Orabona 4, 70126 Bari, Italy, e-mail: giovanni.bruno@ba.imip.cnr.it

Denis Cattelan HORIBA Jobin Yvon S.A.S., Avenue de la Vauve, Passage Jobin Yvon CS 45002, 91120 Palaiseau, France, e-mail: denis.cattelan@horiba.com

Jiří Chaloupka Central European Institute of Technology, Masaryk University, Brno, Czech Republic, e-mail: chaloupka@physics.muni.cz

Alain C. Diebold College of Nanoscale Science and Engineering, University at Albany, Albany, NY, USA, e-mail: adiebold@albany.edu

Bernard Drevillon Ecole Polytechnique, Palaiseau, Paris, France, e-mail: bernard.drevillon@polytechnique.edu

Céline Eypert HORIBA Jobin Yvon S.A.S., Avenue de la Vauve, Passage Jobin Yvon CS 45002, 91120 Palaiseau, France, e-mail: celine.eypert@horiba.com

Martin Foldyna LPICM, CNRS-Ecole Polytechnique, Palaiseau, France, e-mail: martin.foldyna@polytechnique.edu

Miklos Fried Research Institute for Technical Physics and Materials Science (MTA-MFA), Hungarian Academy of Sciences, Konkoly Thege Rd. 29-33, Budapest 1121, Hungary; Budapest 1525, Hungary, e-mail: fried.miklos@ttk.mta.hu

Melanie Gaillet HORIBA Jobin Yvon S.A.S., Avenue de la Vauve, Passage Jobin Yvon CS 45002, 91120 Palaiseau, France, e-mail: melanie.gaillet@ horiba.com

Rados Gajic Institute of Physics, University of Belgrade, Pregrevica 118, P.O. Box 68, 11080 Belgrade, Serbia, e-mail: rgajic@ipb.ac.rs

Enric Garcia-Caurel Ecole Polytechnique, Palaiseau, Paris, France, e-mail: enric.garcia-caurel@polytechnique.edu

Jean-Paul Gaston HORIBA Jobin Yvon S.A.S., Avenue de la Vauve, Passage Jobin Yvon CS 45002, 91120 Palaiseau, France, e-mail: jean-paul.gaston@ horiba.com

Kurt Hingerl Centre for Surface and Nano-Analytics, Johannes Kepler University, Altenbergerstr. 69, 4040 Linz, Austria, e-mail: kurt.hingerl@jku.at

Tino Hofmann Center for Nanohybrid Functional Materials, University of Nebraska-Lincoln, Lincoln, NE 68588-0511, USA, e-mail: thofmann@engr.unl.edu

Josef Humlicek Central European Institute of Technology, Masaryk University, Kamenice 735, 62500 Brno, Czech Republic; Department of Condensed Matter Physics, Faculty of Science, Masaryk University, Address Kotláská 2, 61137 Brno, Czech Republic, e-mail: humlicek@physics.muni.cz

Eugene A. Irene Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, USA, e-mail: gene_irene@unc.edu

Milka Jakovljevic Institute of Physics, University of Belgrade, Pregrevica 118, P.O. Box 68, 11080 Belgrade, Serbia, e-mail: mjakovljevic@ipb.ac.rs

Vimal K. Kamineni Technology Research Group, GLOBALFOUNDRIES, Albany, NY 12203, USA, e-mail: vimal.kamineni@globalfoundries.com

Argiris Laskarakis Lab for Thin Films-Nanosystems & Nanometrology, Department of Physics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, e-mail: alask@physics.auth.gr

Stergios Logothetidis Lab for Thin Films-Nanosystems & Nanometrology, Department of Physics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, e-mail: logot@auth.gr

Maria Losurdo Institute of Inorganic Methodologies and of Plasmas, Italian National Council of Research, Via Orabona 4, 70126 Bari, Italy, e-mail: maria.losurdo@ba.imip.cnr.it

Antonello De Martino Ecole Polytechnique, Palaiseau, Paris, France, e-mail: antonello.demartino@polytechnique.edu

Yasuhiro Mizutani Institute of Technology and Science, The University of Tokushima, 2-1 Minami-Josanjima-cho, Tokushima 770-8506, Japan, e-mail: mizutani@tokushima-u.ac.jp

Dominik Munzar Central European Institute of Technology, Masaryk University, Brno, Czech Republic, e-mail: munzar@physics.muni.cz

Florence J. Nelson College of Nanoscale Science and Engineering, University at Albany, Albany, NY, USA, e-mail: florencenel@gmail.com

Thomas W. H. Oates Forschungsbereich Material- und Grenzflächenanalytik, Leibniz-Institut für Analytische Wissenschaften-ISAS, Berlin, Albert-Einstein-Str. 9, 12489 Berlin, Germany, e-mail: tom.oates@isas.de

Yukitoshi Otani Center for Optical Research and Education, Utsunomiya University, 7-1-2 Yoto, Utsunomiya, Tochigi 321-8585, Japan, e-mail: otani@cc.utsunomiya-u.ac.jp

Razvigor Ossikovski Ecole Polytechnique, Palaiseau, Paris, France, e-mail: razvigor.ossikovski@polytechnique.edu

Peter Petrik Research Institute for Technical Physics and Materials Science (MTA-MFA), Hungarian Academy of Sciences, Konkoly Thege Rd. 29-33, Budapest 1121, Hungary; Budapest 1525, Hungary, e-mail: petrik.peter@ttk.mta.hu

Angelo Pierangelo LPICM, CNRS-Ecole Polytechnique, Palaiseau, France, e-mail: angelo.pierangelo@polytechnique.edu

Ottilia Saxl Director at NANO Magazine, ION Publishing Ltd, Glasgow, UK, e-mail: ottilia.saxl@nanomagazine.co.uk

Daniel Schmidt Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0511, USA, e-mail: schmidt@engr.unl.edu

Eva Schubert Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0511, USA, e-mail: evaschub@engr.unl.edu

Mathias Schubert Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0511, USA, e-mail: schubert@engr.unl.edu

Roland Seitz HORIBA Jobin Yvon GmbH, Hauptstrasse 1, 82008 Unterhaching, Germany, e-mail: roland.seitz@horiba.com

Assia Shagaleeva HORIBA Jobin Yvon S.A.S., Avenue de la Vauve, Passage Jobin Yvon CS 45002, 91120 Palaiseau, France, e-mail: assia.shagaleeva@horiba.com

Michel Stchakovsky HORIBA Jobin Yvon S.A.S., Avenue de la Vauve, Passage Jobin Yvon CS 45002, 91120 Palaiseau, France, e-mail: michel.stchakovsky@horiba.com

Satoshi Tomita Graduate School of Materials Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan, e-mail: tomita@ms.naist.jp

Herbert Wormeester Physics of Interfaces and Nanomaterials, MESA+ Institute for Nanotechnolgy, University of Twente, Enschede, The Netherlands, e-mail: h.wormeester@utwente.nl

Angel Yanguas-Gil Argonne National Laboratory, Argonne, IL 60439, USA, e-mail: ayg@anl.gov

Peter Zeppenfeld Atom-und Oberflaechenphysik, Johannes Kepler Universtiy, Linz, Austria, e-mail: Peter.Zeppenfeld@jku.at

Stefan Zollner Department of Physics, New Mexico State University, MSC 3D, P.O. Box 30001, Las Cruces, NM 88003-8001, USA, e-mail: zollner@nmsu.edu

Chapter 1
A Brief History and State of the Art of Ellipsometry

Eugene A. Irene

Abstract Ellipsometry is a nonperturbing optical technique that uses the change in the state of polarization of light upon reflection for the ex-situ and in-situ and real-time characterization of surfaces, interfaces, thin films and nanostructured materials. This chapter identifies the crucial ideas from which the ellipsometry technique derives, provides the key relationships for ellipsometry and presents a brief introduction to the essential ellipsometry configurations and hardware.

1.1 Introduction

As discussed in depth in the following chapters on optical theory and practice related to ellipsometry of nanomaterials, ellipsometry is first of all an optical technique. Ellipsometry is characterized as a polarized light technique and as it is commonly practiced it is a reflection technique. Therefore, the history of ellipsometry derives from the development of an understanding of the reflection of polarized light. As a reflection optical technique, ellipsometry finds application at surfaces where the optical properties of surfaces and surface films whether continuous or discontinuous can be determined. Thus the state of the art of ellipsometry developed largely as a result of demands made by surface and thin film scientists and engineers.

The history of any particular idea in physics is often impossible to trace precisely. The reason for this is that scientists rely on the existent scientific literature to find and to update their knowledge and then to contribute incrementally to the field. In essence new ideas derive from older ideas and the directions of research are often guided by what has gone before. Therefore it is often difficult to implicate only one or two key scientists related to a particular discovery. Rather there is a trail of publications with

E. A. Irene (✉)
Department of Chemistry, University of North Carolina,
Chapel Hill, NC 27599, USA
e-mail: gene_irene@unc.edu

perhaps one or two well known scientists implicated as well as scores of less well known scientists.

In this chapter one goal is to identify the crucial ideas from which the ellipsometry technique derives rather than to achieve historical accuracy. In addition the key relationships for ellipsometry will be identified and a brief introduction to the essential ellipsometry configurations and hardware will be presented. This chapter is intended to be a brief introduction to ellipsometry so as to provide an introduction to the more detailed and specialized chapters that follow.

The material for this chapter can be gleaned from various optical texts and ellipsometry books and the specific references for this chapter are Refs. [1–5] and the references therein.

1.2 Brief Historical Perspective

1.2.1 Light Waves, Polarized Light and Ellipsometry

Robert Hooke published a wave theory of light in 1660 and Christian Huygens followed soon after with the book entitled *Treatise on light* in 1690. These scientists and other early workers mentioned below are depicted in Fig. 1.1. Huygens proposed that light was emitted from a point source in all directions as a series of waves. Later Thomas Young using experiments with slits confirmed that light consists of waves. Inherent in the wave theory is the concept of interference that is crucial to understanding the reflection of light from surfaces and film covered surfaces.

In the 1820s Augustin-Jean Fresnel showed that the polarization of light waves was possible, if the light had a transverse rather than a longitudinal wave nature.

However it wasn't until the mid 1840s that Michael Faraday proposed that light was related to magnetism from the so-called Faraday effect in which the plane of polarization of light could be changed using a magnetic field and further proposed that light was an electromagnetic disturbance and therefore could propagate without a medium.

James Clerk Maxwell elegantly quantified these educated speculations made by Faraday and in 1873 published *A Treatise on Electricity and Magnetism* which contained the famous Maxwell's Equations that comprise the basic ideas of electromagnetism and electromagnetic waves.

Now the description of polarized light as was developed by the early workers is discussed as applied to ellipsometry. Figure 1.2 shows linearly or plane polarized light as an electromagnetic wave with orthogonal electric **E** and magnetic **B** fields that are characteristic of light as a periodic transverse electromagnetic disturbance. The linear polarization refers to the existence of one orientation of the electric field as the wave propagates in the z direction, but the **E** field (and **B** field) oscillates in x (and the **B** field in y), hence the transverse nature of the light waves. Figure 1.3a also displays the **E** field of a linearly polarized electromagnetic wave with the projection of **E**

Fig. 1.1 Major early scientists who contributed to the understanding of light, polarized light and ellipsometry. Form *top left* Robert Hooke, Christian Huygens, Thomas Young, Augustin-Jean Fresnel, Michael Faraday, James Clerk Maxwell, Etienne-Louis Malus, David Brewster, and Paul Drude

upon a screen in Fig. 1.3b as the wave propagates. Figure 1.3c shows the projection of circularly polarized light which indicates that as the wave propagates the **E** vector rotates with the endpoint of the **E** vector tracing a circle. The general case depicted in Fig. 1.3d is elliptically polarized light that shows an ellipse as the projected locus of the endpoints of **E** as the wave propagates, but **E** also changes in magnitude and hence traces an ellipse. Figure 1.4 shows an incident beam of linearly polarized light with **E** resolved into two orthogonal components $\mathbf{E_{is}}$ and $\mathbf{E_{ip}}$. The plane defined by the incident beam and the surface normal is called the plane of incidence (POI) and as seen has one incident beam reflecting from a surface with the two components of the incident polarized beam: $\mathbf{E_{ip}}$ parallel to POI and $\mathbf{E_{is}}$ orthogonal to the POI (the subscript i refers to the incident beam, r to the reflected beam and the subscript p indicates parallel to the POI and s from the German senkrecht for orthogonal to POI).
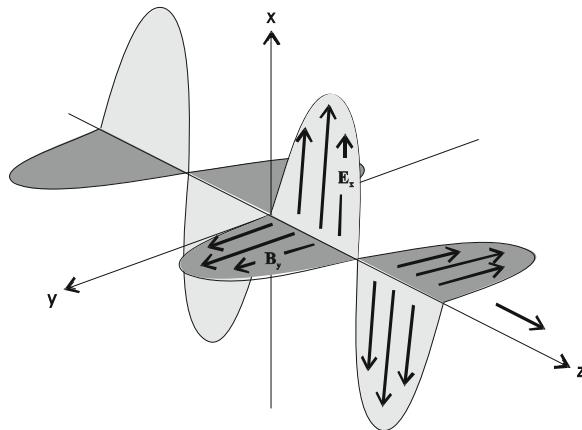
Fig. 1.2 Transverse electromagnetic wave with electric **E** and magnetic **B** fields perpendicular to the propagation direction z
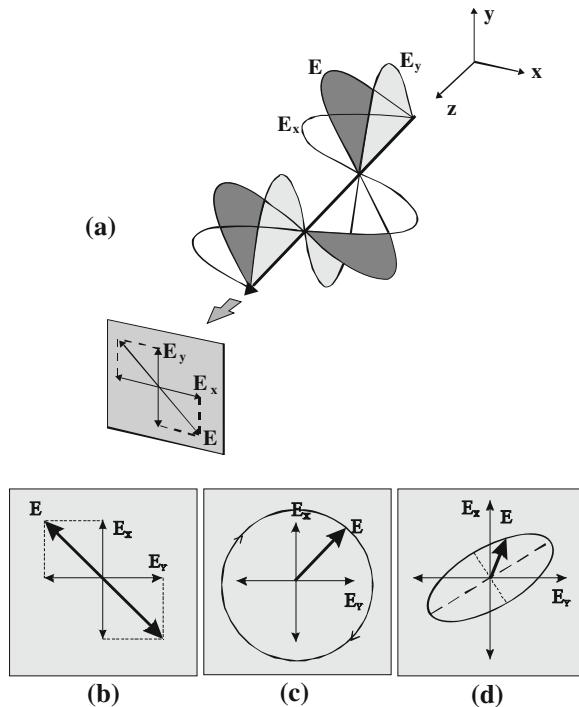
Fig. 1.3 Various polarization states of a monochromatic light wave with **E** field projections in the x and y planes yielding E_x and E_y: **a** linearly or plane polarized light; **b** projected linearly polarized light; **c** projected right circularly polarized light; **d** projected elliptically polarized light
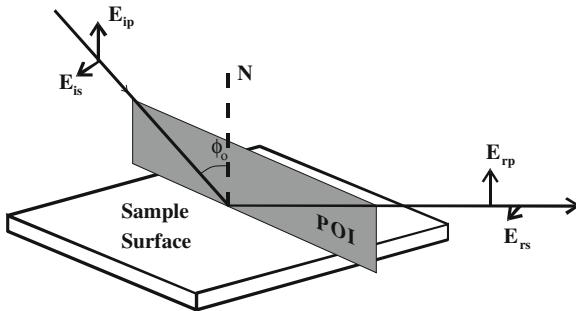
Fig. 1.4 Incident linearly polarized light with orthogonal electric field components $\mathbf{E}_{ip}$ and $\mathbf{E}_{is}$ reflected at incidence angle ϕ_o yielding reflected components $\mathbf{E}_{rp}$ and $\mathbf{E}_{rs}$. The plane of incidence POI contains the surface normal N and the incident ray

It is noted that ellipsometry as presently practiced and as discussed below involves only the **E** field of light.

Etienne-Louis Malus contributed to the understanding of polarized light in the early 1800s where he established that reflection can cause polarization and that crossed polarizers can be used to identify polarization and determine the polarization state of reflected light. After Malus, David Brewster around 1815 quantified Malus' discoveries from a surface and greatly extended the understanding of polarizing materials and methods. Now we more closely examine the specifics of all of these discoveries.

With Maxwell's work being accepted worldwide in the latter part of the 1800s elucidating the electromagnetic nature of light waves, Paul Drude (1863–1906) came onto the scene with major contributions to many areas of physics including optics. Near the end of his life Drude published "The Theory of Optics" in 1902. Part II, Sect. 1, Chap. 5 entitled "Polarization" of this book with the Table of Contents of this chapter included as Fig. 1.5 could easily serve as a modern introduction to polarized light and ellipsometry. Figure 81 from Drude's book displayed as Fig. 1.6 herein depicts all the necessary optical components needed to perform reflection ellipsometry but without mention of the term "ellipsometry". In fact the Drude ellipsometer depicted looks remarkably like a single wavelength ellipsometer marketed in the 1970–1990s by one manufacturer. There is no doubt that Drude contributed much to what would become the field of ellipsometry, but it is also clear that there are many both famous and unsung scientists who preceded Drude and contributed not only to the underlying physics but also to the specific optics for the measurement of reflected polarized light. Yet in many scientific quarters Drude is considered the father of ellipsometry.

After Drude, the development of ellipsometry appears driven by need as are most measurement techniques. For example many scientists, particularly electrochemists, became interested in thin films and therefore techniques were sought and developed that had sensitivity to thin films. In the late 1920s Leif Tronstad did research on

CHAPTER V

POLARIZATION

Fig. 1.5 Chapter 5 table of contents from Ref. [1]
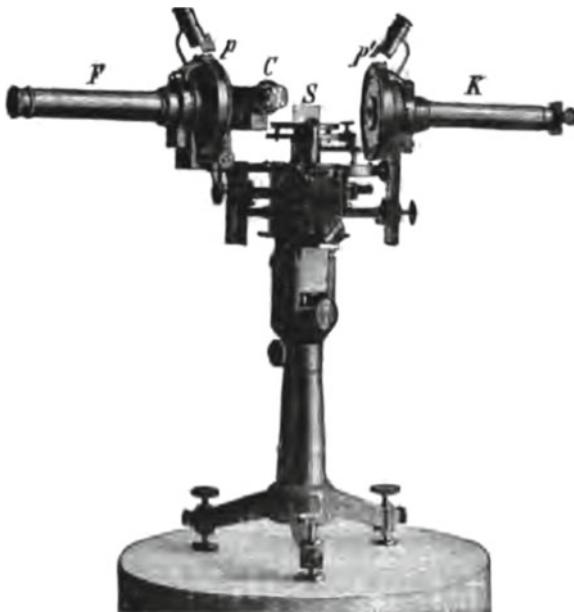
FIG. 81.

Fig. 1.6 Figure 81 from Ref. [1] depicts all the necessary optical components needed to perform reflection ellipsometry

thin films in Herbert Freundlich's laboratory in Berlin where many of the practical aspects of thin film measurement were being worked out along with G. Patcheke and H. Zocher. In 1932 Tronstad published a paper [6] entitled "The investigation of thin surface films on metals by means of reflected polarized light" that directly broached the ellipsometric measurement of the thickness of thin films that were prepared using electrochemical methods. Figure 1.7 depicts Fig. 1.1 from the Tronstad paper which is a drawing of the ellipsometer that was used in the research. This ellipsometer includes a monochromator which could indicate that some form of spectroscopic ellipsometry was also contemplated. Tronstad studied passive oxide films on polished
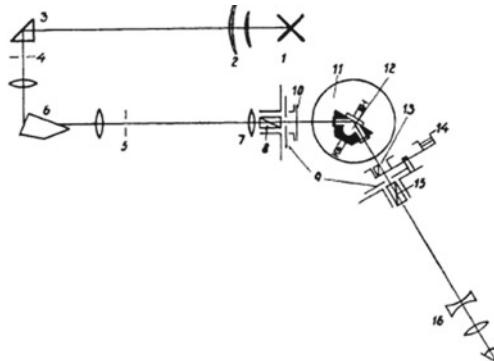
Fig. 1.—The optical system. 1. Light source. 2. Condenser. 3. Total reflecting prism. 4 and 5 Monochromator slits. 6. Lenses and constant-deviation prism of the monochromator. 7. Collimator lens. 8. Polarizer and its scale. 9. Diaphragms. 10. ¼-wave plate of mica (Sénarmont compensator). 11. Central table. 12. Metal mirror (treated as anode or cathode in an electrolytic cell). 13 and 14. Half-shade systems. 15. Analyser and its scale. 16. Telescope.

Fig. 1.7 Figure 1.1 from Ref. [6] is a drawing of the ellipsometer that was used in the research by Tronstad in 1932

iron and steel surfaces and was able to determine a film thickness around 3 nm and even earlier around 1920 similarly thin films were reported from Freundlich. Thus nanometer (nm) dimensions obtained from ellipsometry were reported at least as early as around 1920. It wasn't until 1945 in a paper by Rothen [7], that the actual term "ellipsometry" was used to describe the technique already in use.

1.2.2 The Specifics of the Reflection of Polarized Light

As seen in Fig. 1.4 the reflection of a linearly polarized incident beam (subscript i) from a surface can be described in terms of the projected components parallel (subscript p) and perpendicular (subscript s from the German senkrecht for perpendicular) to the plane of incidence, POI as:

$$\mathbf{E}_i = \left(\mathbf{E}_{ip}^2 + \mathbf{E}_{is}^2 \right)^{1/2}$$

where the POI is defined as the plane comprised of the incident beam and the surface normal [8]. Upon reflection from a surface as displayed in Fig. 1.8a, the amplitude of the reflected electric field vector components, $\mathbf{E}_{rp}$ and $\mathbf{E}_{rs}$ are in general different from the incident components and the phase is also changed. The change in amplitude and phase upon reflection comprise the change in the polarization of the light and indeed this change in polarization is measured in the ellipsometric technique and is related to the optical properties of the system that caused the reflection changes.
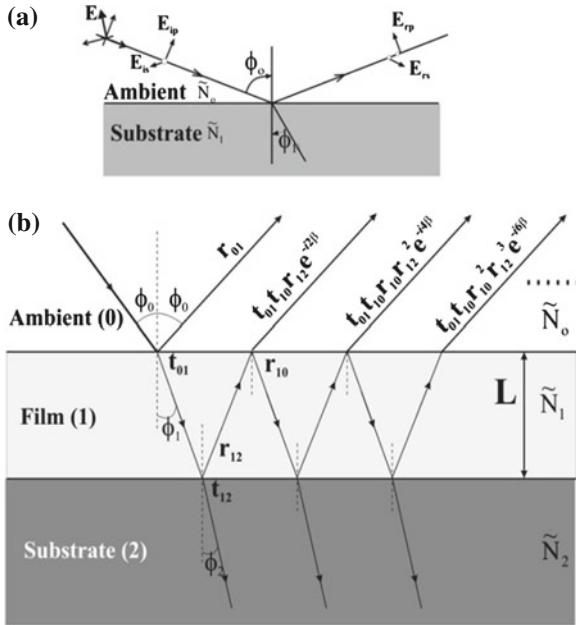
Fig. 1.8 Reflection at **a** bare substrate and **b** film covered substrate showing incident and reflected electric field components E, angles (φ) and optical properties (Ñ). r_{ij} and t_{ij} are the Fresnel reflection and transmission coefficients, respectively

These relationships are discussed further below. The Fresnel equations discussed below describe the reflection and transmission of electromagnetic waves at an interface. The Fresnel reflection coefficients for wave components parallel and perpendicular to the plane of incidence are derived for a dielectric medium by considering continuity of the wave components across the reflection boundary and then using Snell's law which embodies the relationship between angles of incidence and refraction for a transverse electromagnetic wave impinging on an interface between two media with different real parts for the refractive indices n. Snell's law follows from the boundary condition that a wave be continuous across a boundary, which requires that the phase of the wave be constant on any given plane, resulting in the following relationship:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

where n's are the real parts of the refractive indices for the two media 1 and 2 at the interface and θ's are the corresponding angles from the normal of the incident and refracted waves. From Fig. 1.8a for reflection from a single surface the Fresnel's reflection coefficients relate the electric fields and from Snell's law the incident and refraction angles as follows:

$$\mathbf{r_p} = \frac{\mathbf{E_{rp}}}{\mathbf{E_{ip}}} = \frac{\widetilde{N}_1 \cos \phi_0 - \widetilde{N}_0 \cos \phi_1}{\widetilde{N}_1 \cos \phi_0 + \widetilde{N}_0 \cos \phi_1}$$

and

$$\mathbf{r_s} = \frac{\mathbf{E_{rs}}}{\mathbf{E_{is}}} = \frac{\widetilde{N}_0 \cos \phi_0 - \widetilde{N}_1 \cos \phi_1}{\widetilde{N}_0 \cos \phi_0 + \widetilde{N}_1 \cos \phi_1}$$

where the $\widetilde{N}'$s are the complex indices of refraction given as:

$$\widetilde{N} = n + ik$$

where n is the real part of the complex refractive index and is the ratio of the speed of light in vacuum compared to the material in question, and k is the absorption index that is related to the absorption constant α. α and k are related by the formula:

$$\alpha = \frac{4\pi k}{\lambda_0}$$

where λ_0 is the wavelength of the probing light in vacuum.

The dielectric response function, or simply the dielectric function ε is a measure of the response of a material to the interaction with electromagnetic radiation and is given as:

$$\varepsilon = \varepsilon_1 + i\varepsilon_2$$

$$\varepsilon_1 = n^2 - k^2 \text{ and } \varepsilon_2 = 2nk$$

It is important to realize that the dielectric function is defined for a pure homogenous material. The dielectric function for a simple homogenous surface or a pseudodielectric function for a complex film covered surface, and denoted by brackets, can be written in terms of the ellipsometric variable ρ as follows:

$$\varepsilon \text{ or } \langle \varepsilon \rangle = \sin^2 \phi_0 + \sin^2 \phi_0 \tan^2 \phi_0 \left[\frac{1 - \rho}{1 + \rho} \right]^2$$

The use of this formula for ε is straightforward for a pure material: Δ and Ψ are measured at some angle of incidence ϕ_0 and wavelength λ and then ε is calculated using equations above. However in many cases Δ, Ψ is measured for a complicated system such as a film covered surface, possibly with many films and/or roughness and/or impurities. In all of these cases only a single Δ, Ψ is measured at ϕ_0 and λ. The measured Δ, Ψ corresponds to a complex system in which ε is undefined, and thus the use of equation will not yield a value for ε, but rather simply a number. In order to indicate that this was the case for a complex system, angular brackets are used for the calculated quantity, $\langle \varepsilon \rangle$ and it is called the pseudodielectric function in that $\langle \varepsilon \rangle$ is obtained using the relationship for ε.

For film covered surfaces Fig. 1.8b shows that additional reflections and refractions occur at the film-film and film-substrate interfaces and in the different media and as mentioned above the reflected output at the detector must be a sum of all reflected beams. The result, R, as obtained from a converging geometric series is as follows:

$$R_p = \frac{r_{01p} + r_{12p} \exp(-i2\beta)}{1 + r_{01p}r_{12p} \exp(-i2\beta)} \quad \text{and} \quad R_s = \frac{r_{01s} + r_{12s} \exp(-i2\beta)}{1 + r_{01s}r_{12s} \exp(-i2\beta)}$$

where

$$\beta = 2\pi \left(\frac{L \left(\widetilde{N}_1^2 - \widetilde{N}_0^2 \sin^2 \phi_0 \right)^{1/2}}{\lambda} \right)$$

and where the Fresnel coefficients r_{mnl} where mn identify the interface and l is either the p or s component of E are as follows:

$$r_{01p} = \frac{\widetilde{N}_1 \cos \phi_0 - \widetilde{N}_0 \cos \phi_1}{\widetilde{N}_1 \cos \phi_0 + \widetilde{N}_0 \cos \phi_1} \quad r_{12p} = \frac{\widetilde{N}_2 \cos \phi_1 - \widetilde{N}_1 \cos \phi_2}{\widetilde{N}_2 \cos \phi_1 + \widetilde{N}_1 \cos \phi_2}$$

$$r_{01s} = \frac{\widetilde{N}_0 \cos \phi_0 - \widetilde{N}_1 \cos \phi_1}{\widetilde{N}_0 \cos \phi_0 + \widetilde{N}_1 \cos \phi_1} \quad r_{12s} = \frac{\widetilde{N}_1 \cos \phi_1 - \widetilde{N}_2 \cos \phi_2}{\widetilde{N}_1 \cos \phi_1 + \widetilde{N}_2 \cos \phi_2}$$

The exponential factor β often referred to as the phase factor is particularly important since it contains the usually desired information from an ellipsometry measurement viz. film thickness, L, and the complex refractive indices, $\widetilde{N}$.

1.2.3 Fundamental Parameters of Ellipsometry

The measurables of ellipsometry are the phase change upon reflection and the light amplitude change upon reflection which are denoted as Δ and Ψ, respectively. These measurables comprise the ellipsometric experiment and for interpretation require recasting in terms of the physically relevant parameters of the optical system under study. Using the Fresnel coefficients defined above that contain optical parameters of the reflection problem, it is usual in ellipsometry to define a complex reflection coefficient, ρ, for both bare and film covered substrates:

$$\rho = \frac{r_p}{r_s} \text{ or } \frac{R_p}{R_s}$$

where the reflection coefficients, the r's and R's can also be expressed as complex exponentials in order to keep track of both phase and amplitude:

$$r_p = |r_p| \exp(i\delta_p) \text{ and } r_s = |r_s| \exp(i\delta_s)$$

$$R_p = |R_p| \exp(i\Delta_p)$$

$$\text{and } R_s = |R_s| \exp(i\Delta_s)$$

Combining the results above the following is obtained for ρ:

$$\rho = \frac{E_{rp}E_{is}}{E_{ip}E_{rs}} = \frac{|r_p| \exp(i\,\delta_p)}{|r_s| \exp(i\,\delta_s)} \quad \text{or} \quad \frac{|R_p| \exp(i\Delta_p)}{|R_s| \exp(i\Delta_s)} = \tan\psi \exp(i\Delta)$$

where:

$$\tan\psi = \frac{|r_p|}{|r_s|} \quad \text{or} \quad \frac{|R_p|}{|R_s|}$$

and $\Delta = \delta_p - \delta_s$ or $\Delta_p - \Delta_s$. Thus the measurable Δ and Ψ are related to the physical parameters n, k and L embedded within the Fresnel coefficients. From the development above the general formula for ellipsometry is obtained that relates measurables to properties as follows:

$$\rho = \tan\psi \exp(i\Delta) = \rho\left(\tilde{N}_0, \tilde{N}_1, \tilde{N}_2, ..L_1, \ldots \phi_0, \lambda\right)$$

where the last term on the right contains all the physical parameters of the measurement, viz. film thicknesses, optical properties, the wavelength of light and the angle of incidence. It is usual that the last two parameters are known a priori and the others are sought from the measurement of Δ and Ψ. It must be remembered that however complex is the film(s) upon a substrate, there are always only the two same measurables in ellipsometry: Δ and Ψ. Therefore, from a typical measurement of Δ and Ψ only two properties from the general ellipsometry equation can be obtained. Either the other unknowns need to be supplied from independent measurements or additional experimentally controlled hardware variables can be systematically varied in order to increase the number of independent equations. Often the angle of incidence and the wavelength are selected. For many in-situ measurements the experimental situation precludes changes in angle of incidence, but scanning wavelength ellipsometry or spectroscopic ellipsometry is comparatively straightforward to implement.

Now with the relevant physics identified from historical origins along with the important equations revealed that relate the ellipsometric measurables Δ and Ψ with relevant physical parameters, the next three issues to be addressed herein are the accuracy and precision of ellipsometry measurements, the hardware used to perform ellipsometry and finally the present state of ellipsometric science and technology.

1.3 State of the Art of Ellipsometry

1.3.1 Accuracy and Precision of Ellipsometry

The great sensitivity of ellipsometry is derived from the fact that the measurement of Δ, Ψ is a relative measurement of the change in polarization imparted by the sample upon the incident polarized light, as was evidenced by equations for ρ above. It is important to note that because ellipsometry is a relative measurement (see the ratio equations above) unlike absolute light measurement spectroscopies, ellipsometry is not especially sensitive to long term drift in the light source or detector. Hence ellipsometry is sensitive to small surface changes. Indeed fractions of a monolayer are readily sensed by reflection ellipsometry. However this level of sensitivity is achievable only after careful hardware alignment.

   One way to quantify the sensitivity of ellipsometry is to first calculate the effect of the presence of overlayers or films on a substrate on the measurables Δ, Ψ, and then to compare that result with the capability of an ellipsometer to measure Δ, Ψ. Table 1.1 shows calculated results for a Si surface (with n = 3.085 and k = 0.018 at λ = 632.8 nm) coated with a transparent film with n = 1.5 and k = 0. Under these conditions of the calculation for an imaginary film that is similar to SiO_2, it is seen that Δ changes by about 0.3° per 0.1 nm of film and Ψ by about 0.001° for 0.1 nm. Considering that a properly aligned ellipsometer with high quality optics is capable of precision of about 0.01–0.02° in Δ and Ψ, sensitivity approaching 0.01 nm or sub monolayer sensitivity (considering that atomic diameters are of the order of 0.1 nm) is achievable with the determination of Δ. For other film thicknesses and other measurement conditions there may be more Ψ sensitivity.

Automated ellipsometers often have lower sensitivity, but even with an order of magnitude worse sensitivity than calculated above, sensitivity to the presence of a film of the order of 0.1 nm thick is achievable. It was mentioned above that while ellipsometry has great sensitivity, it is a precision optical technique and to achieve the sensitivity, each optical component as well as the angle of incidence ϕ_0 must be carefully calibrated. In order to illustrate the level to which calibration must be done, Table 1.2 is used to show the effect of errors in ϕ_0 on film thickness L and refractive index n. The first column has film thickness of 10 and 100 nm and column

Table 1.1 Calculated Δ and Ψ for various transparent film thickness with conditions: n = 1.5 on Si, λ = 632.8 nm, ϕ_0 = 70°

Δ (°)	Ψ (°)	Film thickness (nm)
179.257	10.448	0.0
178.957	10.448	0.1
178.657	10.449	0.2
178.356	10.450	0.3
178.056	10.451	0.4
177.756	10.453	0.5
176.257	10.462	1.0

Table 1.2 Calculation of errors in film thickness L and refractive index n due to errors in ϕ_0

Film thickness L (nm)	ϕ_0 (°)	Δ (°)	Ψ (°)	L@ 70° (nm)	n@ 70°
10	70.00	150.815	11.404	9.84	1.52
10	70.01	150.770	11.390	9.58	1.56
10	70.02	150.726	11.376	9.36	1.60
10	70.03	150.682	11.362	9.18	1.64
10	70.05	150.593	11.334	8.93	1.71
100	70.00	76.026	43.541	100.01	1.500
100	70.01	75.989	43.540	99.97	1.500
100	70.02	75.952	43.539	99.93	1.501
100	70.03	75.915	43.538	99.89	1.501
100	70.05	75.841	43.536	99.82	1.502

Calculation parameters are L = 10 and 100 nm, n = 1.5 on Si, λ = 632.8 nm, various ϕ_0, Δ and Ψ calculated results in columns 3 and 4, and L and n are calculated from Δ and Ψ assuming $\phi_0 = 70°$

2 shows the variation in ϕ_0 used for the calculation. Columns 3 and 4 display the corresponding calculated Δ and Ψ values. These Δ and Ψ values are then used to calculate L and n values but assuming that ϕ_0 is fixed at 70° rather than the value of ϕ_0 actually used and given in Column 2. In all cases except for the first lines for the 10 and 100 nm films, this introduces an error due to ϕ_0 used.

The first entries in column 5 and 6 are the values for L and n that were recalculated from the columns 3 and 4 Δ and Ψ values. One might expect exact agreement between columns 1 and 5 and for n in column 6 to be exactly 1.5, since all the parameters and input values are the same. Rather for the 10 nm film there is a 1.6 % difference in L and 1.3 % in n that is due to the truncation of the Δ and Ψ values in the thousandths decimal place. This error is barely noticed for the 100 nm film. Since there is a limit of about 0.01° in the measurement of Δ and Ψ, errors of this order can be expected as minimum errors in the 10 nm film thickness range. When erroneous ϕ_0 values are used for the remaining 4 entries for the 10 and 4 entries for the 100 nm films larger errors are seen that are always larger for the thinner films. For example an error of 0.05° in ϕ_0 yield an error of almost 11 % in L and 14 % in n for the 10 nm film, and about 2 % in L and 0.1 % in n for the 100 nm film. It is clear that ellipsometry is a very sensitive technique, sensitive to fractions of a monolayer on a surface so long as the ellipsometers must be carefully aligned and calibrated. In addition there are other factors such as truncation errors and thickness regimes that also can affect the accuracy of the measurement.

1.3.2 Components

In order to briefly describe the optical components needed to perform an ellipsometric measurement of Δ and Ψ Fig. 1.9 shows a simple manual ellipsometer in the Polarizer(P)—Compensator(C)—Sample(S)—Analyzer(A), or PCSA,
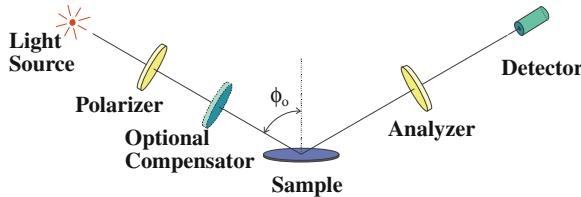
Fig. 1.9 Polarizer, compensator, sample, analyzer PCSA ellipsometer configuration

configuration. The light source for single wavelength ellipsometry is typically a low power laser such as a He-Ne laser at 632.8 nm or a wide band source such as a Xe high pressure lamp with filters to pass only the desired wavelength. For spectroscopic ellipsometry the use of a wide bandwidth light source is required. For the visible spectrum (200–900 nm) it is usual to use Xe short arc high pressure lamps with appropriate collimating optics. In the configuration shown in Fig. 1.9 the usual measurement is the "null" measurement where for a given sample the light intensity at the detector is adjusted to zero by adjusting the P, C, A azimuths with a sample S in place. This null condition plus knowledge of λ, ϕ_0, and the P, C, and A azimuths enable a deduction of what polarization the sample must have imparted to the light which in combination with the optical elements yield zero light intensity at the detector, I_D. With knowledge of how much polarization has been imparted by the sample, and with an optical model for the sample to be discussed below, the optical properties of the sample can be obtained. For null ellipsometry the following condition is sought:

$$I_D = 0 = G\mathbf{E_{AO}}\mathbf{E_{AO}^*}$$

where G is a constant and $\mathbf{E_{AO}}$ is the electric field after the analyzer and $\mathbf{E_{AO}^*}$ is the complex conjugate of $\mathbf{E_{AO}}$. In order to find expressions for the intensity we need to follow the light as it propagates from the source through P, C, and reflects from S and then through A.

When light interacts with an optical element, the polarization state of the light changes. If we are only interested in the change in polarization state of the light before and after it interacts with an ideal optical element (P, C, S or A), the effect of an optical element on the polarization of light can be represented by a 2 × 2 matrix **T** called a Jones matrix that can express the change as follows:

$$\mathbf{E_o} = \mathbf{TE_i}$$

where $\mathbf{E_i}$ and $\mathbf{E_o}$ are the field vectors of the input and output waves respectively, and the Jones matrix of the optical element is given as:

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$
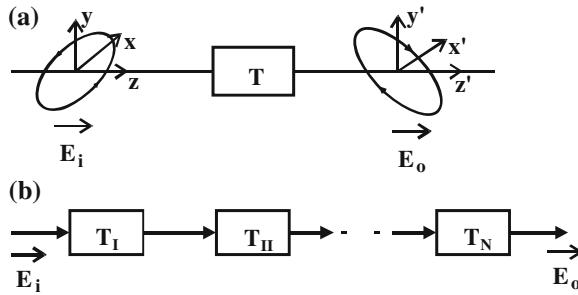
Fig. 1.10 The effect of **a** one optical element, or **b** several elements in series as represented by *boxes* (with T inside), on the polarization state of light where each box is represented by a characteristic Jones matrix and with E_i and E_o the incident light and out light electric field, respectively

In those instances where the optical elements are not ideal and the samples are not uniform more competent mathematical descriptions such as Mueller matrices and Stokes parameters are necessary.

Figure 1.10a shows the interaction using a x, y reference frame. Figure 1.10b shows several optical elements each represented by a different T and for a light wave propagating through all the elements and by using matrix algebra a combination matrix T_{comb} is generated and expressed with the following relationship:

$$\mathbf{E_o} = \mathbf{T_N T_{N-1} \cdots T_{II} T_I E_i} = \mathbf{T_{comb} E_i}$$

The Jones matrices for some most frequently encountered optical elements in ellipsometry systems are for an optically isotropic sample, polarizer or analyzer which are the same optical element differing only by function and a compensator. For an istropic medium with incident light of wavelength λ propagating through the isotropic medium of thickness L and refractive index n, the exiting **E** field, $\mathbf{E_o}$ is given by a Jones matrix operating on the incident $\mathbf{E_i}$ as follows:

$$\mathbf{E_o} = \begin{pmatrix} e^{-i2\pi nL/\lambda} & 0 \\ 0 & e^{-i2\pi nL/\lambda} \end{pmatrix} \mathbf{E_i}$$

Notice that the product nL appears as in the previous formula for β and this product is called the optical path for a material. If the medium is isotropically absorbing, the effect can be represented simply by replacing the refractive index n in the above matrix with the complex refractive index $\widetilde{N}$ of the medium.

A linear polarizer is the main element in an ellipsometer system and it converts incident light of any polarization state into linearly polarized light at its output. Additionally a polarizer is used to resolve the polarization state of light reflected from a sample and before the detector, and in this position the polarizer is called an analyzer. A linear polarizer has two orthogonal axes, i.e., a transmission axis and an extinction axis as indicated by t and e, respectively. When unpolarized or

in general elliptically polarized light passes through a linear polarizer, the light is transformed into linearly polarized light with an electric field vector that is parallel to the transmission axis of the polarizer. The effect of a linear polarizer (of thickness L and index n) can be represented by the following Jones matrix:

$$\mathbf{T} = K \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{where } K = e^{-i2\pi nL/\lambda}$$

A linear polarizer is characterized by a parameter called the extinction ratio, which is the ratio of intensity of light along the extinction and transmission directions. For ellipsometry high extinction ratios are required and suitable polarizers are commercially available with extinction ratios of 10^{-6}.

A compensator is an anisotropic element in which light travels at different speeds in different directions thereby causing a phase change for the light exiting the compensator. The light traversing the compensator travels at different speeds due to different refractive indices along two axes, a fast (f) and a slow (s) axis that are orthogonal. Therefore, when light passes through a linear compensator the phase of the electric vector that is parallel to the slow axis is retarded by δ_c and the amplitude attenuated by T_c with respect to the component parallel to the fast axis. The Jones matrix for a compensator can thus be written as:

$$\mathbf{T} = K_c \begin{pmatrix} 1 & 0 \\ 0 & \rho_c \end{pmatrix}, \quad \rho_c = T_c e^{i\delta_c}$$

where K_c is a constant that accounts for the attenuation and phase shift along the fast and slow axes of the compensator.

A compensator also called a retarder is an optical component that introduces a phase shift between orthogonally polarized components without affecting their relative amplitude. Retarders used in ellipsometry are linear retarders, that have two light propagation directions: a fast axis and a slow axis. The component of incident light parallel to the slow axis is retarded in phase relative to the component along the fast axis. When the phase retardation is $\pi/2$ the retarder is called a quarter-wave retarder that is often used in ellipsometry. There are two types of retarders based on two different mechanisms: birefringent retarders and reflection retarders (the Fresnel rhomb). The more common birefringent retarder has two refractive indices n_o the ordinary and n_e the extraordinary refractive index. Light propagating through the component will have speeds of c/n_o and c/n_e, along the two directions, respectively. For a component of thickness L, this difference in propagation speed will result in a phase shift δ given as:

$$\delta = \frac{2\pi L \, (n_o - n_e)}{\lambda}$$

Then for a material with known n_o and n_e, L can be varied to produce a specific desired retardation. With the Jones matrices for these elements we can now formulate

the propagation of light through an ellipsometer and keep track of the changing polarization state.

For null single wavelength ellipsometry the propagation of light from source to detector the transmitted intensity $\mathbf{E_t}$ can be expressed as follows:

$$\mathbf{E_t} = R_p \cos A \left[\cos C \cos (P - C) - \rho_c \sin C \sin (P - C)\right]$$
$$+ \; R_s \sin A \left[\sin A \cos (P - C) + \rho_c \cos C \sin (P - C)\right]$$

The intensity of the light collected is given by the square of the electric field:

$$I = \left|E_{AO}^{te}\right|^2$$

which is a function of P, C, A, ρ_c, R_p and R_s. For a null ellipsometry system, usually a quarter wave plate is chosen as the compensator, so ρ_c is known (because T_c and ρ_c are known from calibration). During the measurements, P, C, and A are arranged so that the light intensity detected becomes zero (null), which means that $\mathbf{E_t} = 0$. With this condition, we finally obtain the following relationship in terms of the complex reflection coefficient ρ:

$$\rho \equiv \frac{R_p}{R_s} = -\tan A \left[\frac{\tan C + \rho_c \tan (P - C)}{1 - \rho_c \tan C \tan (P - C)}\right]$$

A major ellipsometry advance that occurred in the 1970s is the automation of the measurement system. Several methods are available but all follow a similar rubric of modulating one of the optical components, P, C, or A. Historically the first modulation consisted of rotating the analyzer A called rotating analyzer ellipsometry, RAE. For RAE A is given as:

$$A = \omega t + \delta$$

where δ is a phase constant offset as mentioned above and here t is time. As before the transmitted electric field $\mathbf{E_t}$ is obtained:

$$\mathbf{E_t} = E_0 \left(R_p \cos A \cos P + R_s \sin A \sin P\right) \begin{pmatrix} \cos A \\ \sin A \end{pmatrix}$$

The measured intensity is be expressed as:

$$I = I_0 \left[1 + \left(\frac{\tan^2 \Psi - \tan^2 P}{\tan^2 \Psi + \tan^2 P}\right) \cos 2A + \left(\frac{2 \tan P \cos \Delta \tan \Psi}{\tan^2 \Psi + \tan^2 P}\right) \sin 2A\right]$$

or

$$I = I_0 \left(1 + \alpha \cos 2A + \beta \sin 2A\right)$$
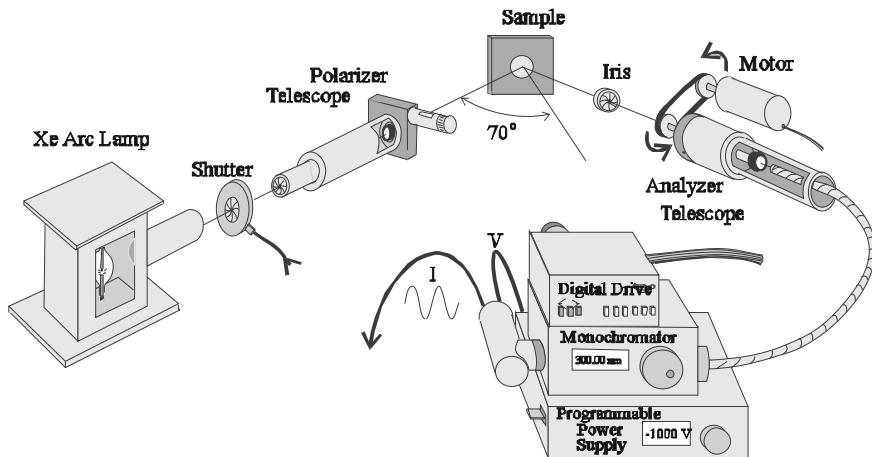
Fig. 1.11 The components for a rotating analyzer ellipsometer

The constants I_0 (the DC signal component), α and β (2nd Fourier coefficients) can be obtained from a Fourier analysis of the detected signal, and thus the two ellipsometric parameters Δ and Ψ are obtained as:

$$\tan \Psi = |\tan P| \sqrt{\frac{1+\alpha}{1-\alpha}}$$

$$\cos \Delta = \frac{\beta}{\sqrt{1-\alpha^2}}$$

The RAE system is readily automated with a sketch of such a system shown in Fig. 1.11 and usually in such a way that data acquisition and analysis are done simultaneously using a computer. There are presently available high quality commercial ellipsometer systems that modulate the polarizer (RPE), analyzer (RAE) and the compensator (RCE) via rotation and the compensator by modulating phase (PME).

The major advantage of PME is a factor of 10^3 in speed relative to RAE or RPE. In the usual polarizer—compensator—sample—analyzer configuration (PCSA) the compensator is replaced with a photoacoustic modulator (PM). The PM is essentially a fused silica window that is cut to resonate and it is put into a resonant vibration via a quartz crystal to which the window is glued. An AC potential applied to the crystal sets up stress in one direction. Hence there is a strain induced time dependence in the refractive index (n) of the window:

$$n_i(V, \lambda, t, l) = n_i^o(\lambda) + V c_i(\lambda) \sin(\omega t) \sin(\pi l/l_o)$$

where c is a piezo-optic coefficient, l is the distance from the interface of c-SiO$_2$ crystal to the light entrance point in the a-SiO$_2$. The same form would apply to the orthogonal j direction. The periodic phase retardation in a-SiO$_2$ is given as:

$$\delta(\lambda, t) = \frac{2\pi d(n_i - n_j)}{\lambda}$$

where d is the thickness of the a-SiO$_2$ part. These equations when combined yield:

$$\delta(V, \lambda, t) = \delta_o(\lambda) + A \sin(\omega t)$$
$$where$$
$$A = \frac{2\pi dV(c_i(\lambda) - c_j(\lambda))}{\lambda}$$

For P $= 45°$, PM $= 0°$, A $= 45°$ the following formulas for intensity apply:

$$I_\omega = I_o \sin(2\Psi) \sin \Delta$$
$$and$$
$$I_{2\omega} = I_o \sin(2\Psi) \cos \Delta$$

where I$_o$, I$_\omega$, I$_{2\omega}$ are the respective intensities at DC, and for the first and second harmonics. Thus Δ and Ψ are obtained at high speed. There is the issue of a spectroscopic measurement as for the RA case and the calibration of the PM.

Another major advance was made in the 1970s, largely through the efforts of Aspnes (9) namely the implementation of spectroscopic ellipsometry (SE). As was pointed out above, several early workers anticipated using wavelength selective optical components for ellipsometry. However the modern implementation made full use of the spectrally enhanced results as pointed out below. SE is often accomplished using a wide band light source that can produce the desired spectrum (a Xe high pressure lamp for the near uv-visible spectrum), a polarizer, sample stage and a rotating analyzer (discussed above) and then an automated monochromator followed by a detector. The monochromator drive is controlled by a computer that also controls the rotating analyzer, and whose angular position is sensed by an optical angular encoder, to allow data to be acquired and analyzed at the same time at each wavelength. Alternatively the light dispersed by the monochromator can be impinged upon a linear detector array (photo diode or charge coupled device CCD) and the entire spectrum accessed in one measurement. This hardware, the detector array and associated electronics to scan the array, is called an optical multichannel analyzer or OMA and is presently in wide use in commercial SE systems.

Automation is especially important for spectroscopic in situ real-time monitoring of thin film formation and/or surface evolution during reactions at surfaces. Typical automation hardware for an RAE system includes an optical angular encoder synchronized with the rotating analyzer that is used to digitize the signal after the analyzer. The digitized signal is then collected by a computer which also performs the

Fourier analysis to obtain α and β (and thus Δ and Ψ), and then the modeling regression analysis discussed briefly below. The overall speed of RAE measurements is appropriate for many processes interesting in microelectronics such as film growth or removal (etching) in real time using PC class computers. The usual scanning monochromators may take minutes to scan the visible spectrum that is used for most SE with shorter times for more limited spectral scans. Within the last 10 years detector arrays (both 1D and 2D) have been used to obtain complete spectra in seconds and less thereby increasing the kinds of processes monitored using ellipsometry.

1.3.3 Optical Models

In order to extract useful optical properties for films and surfaces from the two ellipsometric measurable Δ and Ψ, the optical system (substrate, film, ambient) being investigated needs to be approximated and this approximation of reality is called a model and for ellipsometry an optical model. The kind of modeling now discussed was found to be very useful for conventional systems of thin film materials on specular surfaces. However with the advent of nano structures and nano materials new modeling approaches are required that consider small structures in various media. Here the tried and true methods are briefly reviewed and following chapters inform about the new approaches to modeling that are being developed for nano science and technology applications.

   The extent to which an optical model is correct will determine the physical meaningfulness of the ellipsometrically determined properties. For example in order to accurately extract optical properties from an ellipsometric measurement, it must first be determined whether the system under study is a bare or film covered surface with sharp interfaces as shown in Fig. 1.12a or whether the single film is inhomogeneous as shown in Fig. 1.12b where the single inhomogeneous film on a substrate has sharp interfaces, and/or whether the films and/or interfaces have significant roughness as is shown in Fig. 1.12c or multiple films each with different characteristics as shown in Fig. 1.12d. In some instances the interfaces between films and substrate may not be sharp due to interaction or diffusion which further complicates the modeling. Once a model for the system is obtained, an algorithm can be formulated that considers reflection and refraction at each interface with different optical properties for each film. The model based algorithm is then used to invert the general ellipsometry equation:

$$\rho = \tan\psi \exp(i\Delta) = \rho\left(\widetilde{N}_0, \widetilde{N}_1, \widetilde{N}_2, ..L_1, \ldots \phi_0, \lambda\right)$$

to obtain the desired optical properties that are in the model. It should be noted that film thickness is also considered to be a film optical property in that thickness determines the path length for the optical wave in the material and the optical thickness is the product of thickness and refractive index, n·L and this product is found in the equation above for the exponent β in the Fresnel equations.

One successful strategy is to deduce an optical model based on facts obtained from an often lengthy materials science study that typically involves a number of independent measurements using a variety of techniques. Another strategy for studying a new film-substrate system is to commence with one of the established models that appears to fit the situation under investigation, and then try to determine how well the physical parameters obtained using the model agree with independently determined values. In essence this method is to make an educated guess based on previous experience, and then provide an independent test. For example in microelectronics where devices are fabricated using various films on semiconductor surfaces the dielectric film thickness, L, on a semiconductor surface is usually of great interest, particularly for SiO_2 films on Si, since many device electronic properties depend on dielectric film thickness. For this system a single film model is used because it is known that there is little interaction between Si and SiO_2 and for films greater than 10 nm the sharp interface model is applicable and yields good results. L can not only be measured by ellipsometry, but also by angle resolved x-ray photoelectron spectroscopy, ARXPS, and transmission electron microscopy, on cross-sectioned samples, XTEM. One could use the results from these independent techniques to compare with the ellipsometric optical model results or as input to the model so as to extract other parameters. The models that are shown in Fig. 1.12 are ones that have already been used successfully for a number of thin film systems of interest and will be referred to below.

Figure 1.12a is a simple single film on substrate model in which both the film and substrate are discrete, planar and homogenous. The film thickness, L, can vary upwards from 0 nm that is indicative of a smooth bare substrate. This model works very well for grown and deposited dielectric films on Si and for many other film substrate situations where nearly perfect smooth substrates are used, and where uniform stoichiometric films are possible, i.e. where there is little interaction between film and substrate during film formation, and for thicker films ($>10\,nm$) where interfacial effects are relatively small. In this single film model there are possibly seven model parameters to be determined or supplied: n and k for ambient, film and substrate and the film thickness L. With ambient air or vacuum the optical properties are known with $n = 1$ and $k = 0$ and can therefore be provided as input to the single film model algorithm. For other ambients for example liquids the properties of the medium (n and k) would add another two unknowns. Typically, n and k for the substrate can be measured separately for a bare surface and then input the algorithm. With known substrate and ambient optical properties, a single ellipsometric measurement of Δ and Ψ would not yield remaining three film parameters (n, k, L). In order to overcome this problem sometimes λ and a range of λ can be chosen so that k for the film is 0, i.e. the film is optically transparent. This is possible and convenient for many dielectric films in the visible photon energy range. Another way around this problem of too many unknowns is to measure thickness by another technique and then use it as input to obtain n and k. Once found, n and k can be used as input to find L. Regression analyses can also be used that enable the best fit for underspecified systems and this will be discussed below.

Fig. 1.12 Optical models:
a uniform single film on
substrate; **b** inhomogeneous
film on substrate; **c** rough film
on substrate; **d** multiple films
on substrate
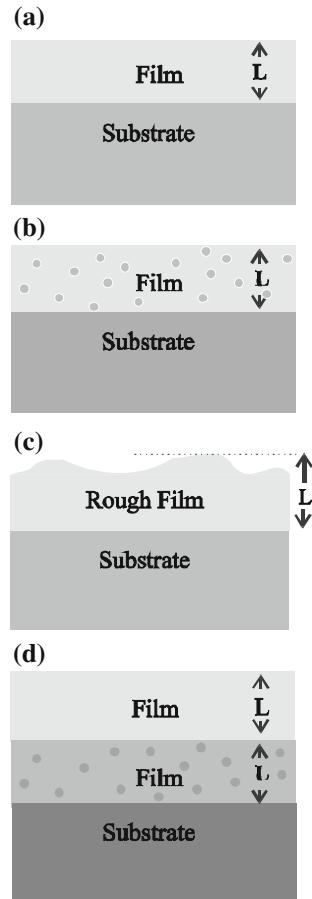
Figure 1.12b–d show the more complicated cases where the film(s) and/or the substrate are inhomogeneous. For Fig. 1.12b there are two different materials and therefore two different film compositions with commensurate properties. In this case there will be composition unknowns in addition to the unknowns in Fig. 1.12a. For many situations similar to Fig. 1.12b it has been shown that the Bruggeman effective medium approximation, BEMA, can yield excellent results for a variety of inhomogeneous film-substrate situations. Essentially this model that is discussed further below considers that an inhomogeneous layer is discrete, i.e. with sharp interfaces as for the simpler models, but with a dielectric response that is the composite of the dielectric responses of the individual components. The manner in which the individual components with their respective contributions to the total dielectric response, $\langle \varepsilon \rangle$ are summed, varies according to assumptions made based about the state of aggregation. In order to understand this more clearly and without lengthy derivations we proceed quickly to the final forms used for optical response calculations. Justification
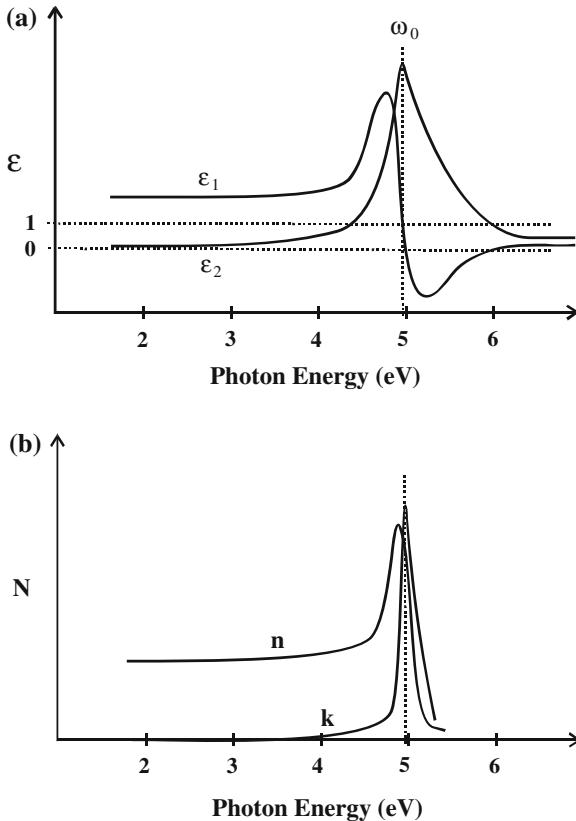
Fig. 1.13 Optical properties of a typical dielectric in the visible light range: **a** in terms of the optical dielectric function ε and real and imaginary components ε_1 and ε_2 , respectively, and **b** in terms of the refractive index n and absorption constant k

for the application of the various models is in the original literature as are derivations. The various effective medium approximations that can be used for a variety of applications are discussed after other useful approximations are introduced.

It is often useful to approximate the dispersion of ε (the changes in ε with photon energy or ν or λ) or changes in $\widetilde{N}$ for a film or substrate. Typical spectra for a material in terms of ε and $\widetilde{N}$ are shown in Fig. 1.13 for the visible photon energy range. Figure 1.13a shows a strong absorption near 5 eV with both ε_1 and ε_2 slowly varying at lower energies away from the absorption. Figure 1.13b show the same spectrum but in terms of n and k. At energies lower than the adsorption k is close to zero and n is slowly varying. For the sub-absorption photon energies where k can be safely approximated as 0 the Cauchy formula is often used to approximate n(λ) as follows:

$$n(\lambda) = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^3}$$

where A, B and C are Cauchy parameters. For materials or at photon energies where absorption is in evidence the Cauchy formula is not a good approximation and a useful formula that includes optical absorption, the Lorentz Oscillator formula is used and for a single oscillator is given as:

$$\varepsilon = 1 + \frac{4\pi e^2}{m\left(\omega_0^2 - \omega^2 - j\Gamma\omega\right)}$$

where m is the oscillator reduced mass, ω_0 is a oscillator resonance frequency, ω is the probe frequency and Γ is a broadening parameter for the resonance. For more than one oscillator the Lorentz formula is a sum as:

$$\varepsilon = 1 + \frac{4\pi e^2}{m} \sum_i \frac{N_i}{\omega_i^2 - \omega^2 - j\Gamma_i\,\omega}$$

where N_i is the number of oscillators per volume with resonance at ω_i. For $\sum_i N_i = N$ and ω_o is the resonance frequency, ε_1 and ε_2 are as follows:

$$\varepsilon_1 = 1 + \frac{4\pi Ne^2\left(\omega_0^2 - \omega^2\right)}{m\left((\omega_0^2 - \omega^2)^2 + \Gamma^2\,\omega^2\right)} \quad \text{and} \quad \varepsilon_2 = \frac{4\,\pi\,Ne^2\Gamma\,\omega}{m\left((\omega_0^2 - \omega^2)^2 + \Gamma^2\,\omega^2\right)}$$

Figure 1.13a shows a single oscillator with Γ given by the width of the resonance at ω_0 at half maximum. The region near ω_0 where ε_1 decreases rapidly and even becomes negative and n goes through a maximum is called the anomalous dispersion region. The Cauchy and/or Lorentz formulas are used to approximate various films in single or multiple film models for non-absorbing or absorbing films, respectively. The remaining issue is the modeling of inhomogeneous films and rough interfaces and this can be done using effective medium approximations, EMA's.

The effective medium approximations formulas that are used for approximating inhomogeneous films are now discussed and are derived by considering how the incident electromagnetic light waves interact with the inhomogeneous material. Thus different formulas are obtained for different materials and different materials circumstances such as mixtures. The starting point is the Clausius-Massotti equation which connects a microscopic material property, the polarizability, α, to the macroscopic dielectric response, ε:

$$\frac{\varepsilon - 1}{\varepsilon + 2} = \frac{4\pi}{3} n\alpha$$

In this formula n is the number of polarizable species in the volume of material probed or the density. This equation obtains for a pure substance and is derived from a consideration of the local electric fields. For a material that is approximated as a heterogeneous mixture of polarizable points (atoms/molecules) a and b in vacuum and each with a different polarizability, α_a and α_b, then the Lorentz-Lorenz equation applies:

$$\frac{\varepsilon - 1}{\varepsilon + 2} = f_a \frac{\varepsilon_a - 1}{\varepsilon_a + 2} + f_b \frac{\varepsilon_b - 1}{\varepsilon_b + 2}$$

where the f's are the volume fractions of constituents.

If a and b are not points in a vacuum, but rather are in a host with a dielectric response ε_h then the Maxwell-Garnet equation is obtained:

$$\frac{\varepsilon - \varepsilon_h}{\varepsilon + 2\,\varepsilon_h} = f_a \frac{\varepsilon_a - \varepsilon_h}{\varepsilon_a + 2\,\varepsilon_h} + f_b \frac{\varepsilon_b - \varepsilon_h}{\varepsilon_b + 2\,\varepsilon_h}$$

If a is considered as the host and thus we have a mixture of points of b in a then one obtains:

$$\frac{\varepsilon - \varepsilon_a}{\varepsilon + 2\,\varepsilon_a} = f_b \frac{\varepsilon_b - \varepsilon_a}{\varepsilon_b + 2\,\varepsilon_a}$$

The Maxwell-Garnet relationship has found application in the field of cermets which are ceramic composites composed of hard brittle ceramic particles in a connected ductile phase. If $f_a \approx f_b$, i.e. there are ample amounts of both materials present, and we let $\varepsilon_h = \varepsilon$ with a as host the result is:

$$f_a \frac{\varepsilon_a - \varepsilon}{\varepsilon_a + 2\,\varepsilon} + f_b \frac{\varepsilon_b - \varepsilon}{\varepsilon_b + 2\,\varepsilon} = 0$$

which is called the Bruggeman effective medium approximation, BEMA, as was mentioned above and this formula is generalized for i constituents as follows:

$$\sum_i f_i \frac{\varepsilon_i - \varepsilon}{\varepsilon_i + 2\,\varepsilon} = 0$$

The BEMA assumes mixtures on a scale smaller than the wavelength of light, but that each constituent retains its original dielectric response. One can imagine that this model might be appropriate for mixed phase films, large amounts of impurities in substrates and damage, and roughness and indeed applications to these cases have been successful in many instances.

Now with these approximation tools in hand a regression analysis can be used to extract desired optical properties from Δ and Ψ measurements at numerous photon energies (spectroscopic ellipsometry SE will be discussed further below). One recipe for this analysis is shown in Fig. 1.14 and is described as follows:

1. Measure Δ, Ψ at various λ's and obtain ρ_{exp}. This experimentally determined quantity, ρ_{exp}, provides one input to the regression analysis and is displayed in the top box in Fig. 1.14.
2. As was discussed above the optical model is deduced from independent experiments and measurements, and/or good guesses and this is seen as the second box in Fig. 1.14. The identity of each film, its constituents and state of aggregation (homogeneous, inhomogeneous, rough etc.) is determined.
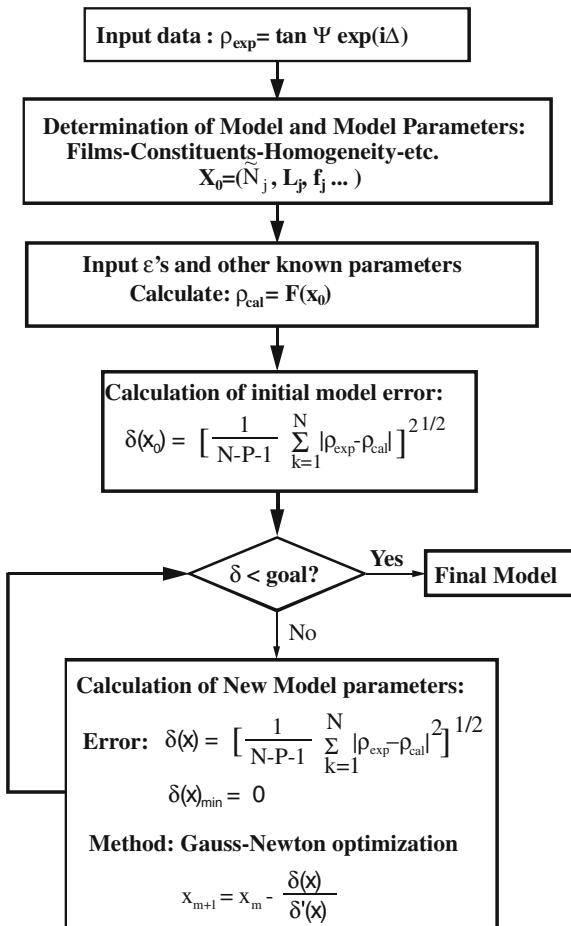
Fig. 1.14 Regression analysis flow chart. Δ and Ψ are measured variables and ρ is the complex reflection coefficient. $\widetilde{N}_j$ is the complex index for the j component, L_i is film thicknesses and f_j are the volume fractions of constituents, N is the number of measurements and P the number of parameters and δ the tolerable error

3. Measure or find in the literature the ε's ($\widetilde{N}$'s) for all the constituents. Steps 2 and 3 are summarized in the second box from the top in Fig. 1.14.
4. The third box from the top indicates that from items 2 and 3 above ρ_{calc} is calculated.
5. The values for ρ_{exp} and ρ_{calc} are compared as shown in the fourth box from the top where N is the number of measurements (the number of λ's) and P is the number of parameters.

6. The difference is compared in the fifth box with a tolerable difference between the model and experimental results. If this result is satisfactory the calculation ends and the parameter values obtained are deemed to be correct.

7. If not then the process repeats with a change of parameters as shown in the sixth box This first comparison derives from the initial values for the parameters in the model. There are many mathematical methods that could be used for the minimization routines with one shown in the last box. Also it should be remembered that the parameter values obtained are only as good as the model. It is possible to get a good fit to a physically or chemically incorrect model and correlation among parameters is also a source for error and often determined. With a well substantiated optical model, a regression analysis has been found to be useful in obtaining desired materials parameters and properties.

Figure 1.15 shows a summary of a typical thin film ellipsometry study on the plasma oxidation of a single crystal Si surface where optical modeling was performed. Various optical models were proposed based on transmission electron microscopy (TEM) and x-ray photoelectron spectroscopy (XPS) studies of the interface between the oxide film and the Si substrate combined with a knowledge of the plasma oxidation mechanism. Figure 1.15a shows four models that fit some of the observed facts. Model 1 considers only a homogenous amorphous SiO_2 film on the Si surface and is included only for completeness, Model 2 considers that the oxide film is a homogenous mixture of amorphous Si and SiO_2 and is based on the higher refractive index observed for thin oxide films. Models 3 and 4 consider a separate mixed interface layer coated with pure amorphous SiO_2 where either the other ingredient in the interface layer is either crystalline or amorphous Si, respectively. It is seen that the best fit model is Model 4 based on the mathematical quality of the fit measured by an unbiased estimator δ. While Model 4 is the best mathematical fit it cannot be argued
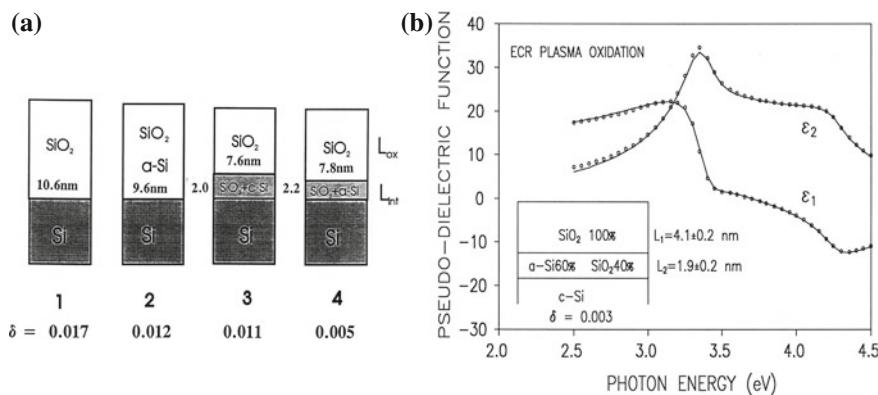

Fig. 1.15 a Various optical models considered for ECR plasma oxidation of Si with typical values for the unbiased estimator δ. **b** Shows a typical data fit to model 4 that consistently displayed the best fit to the model (inset). (Adapted from Hu et al. [10] Fig. 1.1b)

that this Model is the true Model based on the fit alone, because there are always errors associated with the measurements that could influence the fit. However not only is the mathematical fit best for Model 4 but the excess Si observed by XPS near the interface but in the oxide film is also found to be amorphous by TEM. So the good fit merely confirms the physical measurements to further strengthen the conclusion that Model 4 is likely the true model for the plasma oxidation of Si.

1.4 The Past, Present and Future of Ellipsometry

As was briefly discussed above, historically ellipsometry was developed from the emerging understanding of light in the 1600s and the employment of classical optics with the intent to study thin films into the early 1900s. The term "ellipsometry" was coined in the mid 1900s. Ellipsometry as it is presently practiced has been developed as one of the most surface sensitive surface optical techniques and as such has been found useful to determine surface optical properties of materials where typically only bulk optical properties are available from other techniques. Furthermore since ellipsometry requires only conventional light sources and the required hardware can be placed at a distance from a sample and outside typical thin film process chambers, ellipsometry has been found very useful for in situ and/or real-time thin film process monitoring. Many of the earliest applications were derived from electrochemical research, since that field produced thin films on electrode surfaces. The formation of some films reduced the electrochemical chemical reactivity of the surfaces and consequently the surfaces were passivated. The formation of passive films was and is an important field of study in corrosion science. Even greater progress was made in the 1970s due mainly to the demands of the microelectronics industry that utilizes thin films for computer chips. The example summarized in Fig. 1.15 of Si oxidation is a classical example in microelectronics where ellipsometry can follow important processes very accurately and in real time. The microelectronics industry resulted in large part from the ability to control the electric potential at a semiconductor which when accomplished via the thermal oxidation of Si led to metal oxide semiconductor (MOS) devices that comprise a major fraction of modern microelectronics applications.

It is interesting to consider the cyclic nature of the advances in ellipsometry. While it is clear that the demands of the emerging microelectronics industry in the late 1960s and 1970s for high quality thin films and surface properties obtainable from ellipsometry led to increased interest and development in ellipsometry, the developments within microelectronics such as microprocessors leading to powerful PC's and digital signal processing chips such as CCD's and the analysis algorithms to analyze and reduce data have lead to further developments of ellipsometry automation, speed and accuracy. Ellipsometer manufacturers were encouraged to develop better hardware with automation and sophisticated software so as to both improve the quality of the ellipsometric measurements and increase the accessibility to users who are not deeply trained in optical techniques and data analysis.

Now after a hundred plus years of progress in ellipsometry due in large part to the requirements and demands of the now mature microelectronics industry what lies ahead? The answer is general terms is somewhat obvious. Recent research and product development has focused on small structures, even smaller than those in microelectronics. The new prefix "Nano" is used to describe the many new terms that reflect new research paths and potentially new products. From a materials science point of view nanomaterials or materials based on nanostructures offer unlimited commercial potential. Consequently, numerous large and small research programs worldwide have been dedicated to some form of nanoscience or nanotechnology and this trend continues ever rapidly. While some products derived from nano studies are in the marketplace, the potential is far from realized. One key issue in both nanoscience and nanotechnology is the availability of methods to characterize the nano-materials. Ellipsometry has been shown to be a competent technique for determining film dimensions well into the lower end of the nano size range. Efforts are well underway (the NanoCharm program for example) to develop ellipsometry for nano-materials characterization. The following chapters are dedicated to summarizing the understanding developed and progress made with ellipsometry and associated polarized light techniques Here only some of the issues and objectives for ellipsometry nanostructure measurements are summarized.

The predominant issue is size. It will be seen that the consequences of size leads to most of the other difficult issues with characterizing nanomaterials. From the voluminous literature on thin films, ellipsometry has been proven useful to obtain accurate film thicknesses, L, to 0.1 nm and the materials optical dielectric function ε that includes n and k. However there are some limitations for conventional thin films. First for homogenous films less than 10 nm thick it is often difficult to obtain both L and ε. Often either ε or L is independently determined and used as input to an ellipsometric calculation where measured Δ and Ψ are to obtain the other. A usual approach is to determine the film thickness using electron microscopy, electron spectroscopy or scanning probe microscopy and then use the thickness to determine ε for a given material of nm thickness. Once established, the ε for the nm range thick films can then be used as input for the further ellipsometric characterization of the material. It is unwise to use tabulated values for ε obtained from bulk materials or even derived from thick films that are well outside of the nm range, since it is known that for many films the optical properties are a function of thickness and of the exact method of preparation. These notions are well known for thin films and apply equally to all small structures.

Next is the question of homogeneity. Many nanostructures are composed of nm sized particles, columns, wires etc that must be supported on a surface or included as a composite in a medium. If for example nanoparticles are included in a film of a different material like a pure substance as depicted in Fig. 1.12b then there is a possibility that effective medium approximations can be used to formulate the dielectric response. If the occluded nanoparticles can be described by a dielectric function then a Bruggeman Effective Medium Approximation (BEMA) has been found to be useful. However, it should be recalled that a dielectric function is strictly only defined for a pure material of sufficient extent to exhibit a uniform dielectric

response. For very small particles composed of relatively small numbers of atoms or molecules, the addition or subtraction of a few atoms or molecules could change the dielectric response. Thus a single dielectric function is not defined for that material and the BEMA would fail. The argument persists in physics just how large a material must be to be described by a single dielectric function. This issue is related to the similarly persistent argument about how large a material must be to exhibit an electron energy band structure.

Small structures can have many shapes other than spherical. Shape gives rise to directionality which combined with size can potentially result in quantum confinement and then ultimately affect the optical response of the nanostructure. Also small structures can scatter incident light. Non-specular reflection can be interpreted at the detector as absorption if the loss is not otherwise accounted for. This issue arises when performing ellipsometry on rough surfaces when the rough feature size relative to the wavelength of the incident light are in the range of $\lambda/2$ and therefore scattering is maximum.

These issues of size, shape and state of aggregation are simultaneously manifest in nanostructures. The resolution of these issues are subjects for study and progress is discussed in the following chapters.

References

 1. P. Drude, *The Theory of Optics* (Longmans Green and Co., New York, 1902)
 2. F.A. Jenkins, H.E. White, *Fundamentals of Optics* (McGraw-Hill, New York, 1950)
 3. L. Pedrotti, F.L. Pedrotti, *Introduction to Optics* (Prentice Hall, Englewood Cliffs, 1993)
 4. R.M.A. Azzam, N.M. Bashara, *Ellipsometry and Polarized Light* (North-Holland Publishing Company, Amsterdam, 1977)
 5. H.G. Tompkins, E.A. Irene, *Handbook of Ellipsometry* (William Andrews, New York, 2005)
 6. L. Tronstad, Trans. Faraday Soc. **29**, 504 (1932)
 7. A. Rothen, Rev. Sci. Instrum. **16**, 26–30 (1945)
 8. E.A. Irene, *Surfaces, Interfaces, and Thin Films for Microelectronics*, Chapter 9 (Wiley-Interscience, Hoboken, 2008)
 9. D.E. Aspnes, A.A. Studna, Appl. Opt. **14**, 220 (1975)
10. Y.Z. Hu, J. Joseph, E.A. Irene, Appl. Phys. Lett. **59**, 1353 (1991)

Chapter 2
Advanced Mueller Ellipsometry Instrumentation and Data Analysis

Enric Garcia-Caurel, Razvigor Ossikovski , Martin Foldyna, Angelo Pierangelo, Bernard Drévillon and Antonello De Martino

Abstract The main object of this chapter is to give an overview the possibilities offered by instruments capable of measuring full Mueller matrices in the field of optical characterization. We have chosen to call these instruments Mueller ellipsometers in order to highlight their close relation with instruments traditionally used in ellipsometry. We want to make clear to the reader the place that Mueller ellipsometry takes with respect to standard ellipsometry by showing the similarities but also the differences among these techniques, both in instrumentation and data treatment. To do so the chapter starts by a review of the optical formalisms used in standard and Mueller ellipsometry respectively. In order to highlight the particularities and the advantages brought by Mueller ellipsometry, a special section is devoted to the algebraic properties of Mueller matrices and to the description of different ways to decompose them. Matrix decompositions are used to unveil the basic polarimetric properties of a the sample when a precise model is not available. Then follows a description of the most common optical configurations used to build standard ellipsometers. Special attention is paid to show what can and what cannot be measured with them. On the basis of this knowledge it is shown the interest of measuring whole

E. Garcia-Caurel (✉) · R. Ossikovski · M. Foldyna · A. Pierangelo · B. Drévillon · A. De Martino
LPICM, CNRS-Ecole Polytechnique,
Palaiseau, France
e-mail: enric.garcia-caurel@polytechnique.edu

R. Ossikovski
e-mail: razvigor.ossikovski@polytechnique.edu

M. Foldyna
e-mail: martin.foldyna@polytechnique.edu

A. Pierangelo
e-mail: angelo.pierangelo@polytechnique.edu

B. Drévillon
e-mail: bernard.drevillon@polytechnique.edu

A. De Martino
e-mail: antonello.demartino@polytechnique.edu

Mueller matrices, in particular for samples characterized by complex anisotropy and/or depolarization. Among the numerous optical assemblies able to measure full Mueller matrices, most of them are laboratory prototypes, and only very few have been industrialized so far. Because an extensive and comparative review of all the Mueller ellipsometric instruments developed to date is clearly out of the scope of this chapter, we limit our description to four Mueller ellipsometers, two imaging and two spectroscopic systems that have been developed by us in the past years. The technical description of the Mueller ellipsometers is accompanied by some examples of applications which, without being exhaustive, are representative of the type of analyses performed in ellipsometry, and also illustrate the advantages that can be brought by modern Mueller ellipsometers to optical metrology, materials science and biomedicine.

2.1 Introduction

The use of polarized light to characterize the optical properties of materials, either in bulk or thin film format, is generally called Ellipsometry. The technique finds its roots in the pioneering work by Paul Drude in the nineteenth century when he used polarized light in a reflection configuration to study the optical properties and thickness of very thin metallic films. Since then, many methods have been successfully used to generate and analyze the polarization properties of light, and this technique has enjoyed a great success over the past decades. Hundreds of studies and industrial applications have emerged, which are either directly based on ellipsometry, or profit from its sensitivity.

Classical ellipsometric measurements require that the light beam remains completely polarized during the measurement process. Let us recall that a light beam is said to be fully polarized, when the relative phase between the different components of the electromagnetic field along two orthogonal directions remains constant. If for some reason, this relative phase varies (spatially, spectrally and/or temporally), the light will become partially polarized. If so, single ellipsometric measurements lose their physical meaning. To correctly measure and physically interpret the properties of partially polarized light, it is necessary to use the more general technique called Mueller Ellipsometry or Polarimetry. Mueller ellipsometry is thus needed for complete and accurate characterization of the anisotropic and/or depolarizing samples of interest in many instances, both in academic research and "real life" activities. We prefer the term "Mueller Ellipsometry" instead of "Polarimetry", to emphasize the close relationship of this technique with standard ellipsometry. Several excellent monographs [1–3] have been published covering different aspects of both standard and Mueller ellipsometry, such as the theory of polarization, the optical response of solids, the instrumentation and innovative applications.

This chapter is aimed at accounting for the novel Muller ellipsometric data analysis, instrumentation, and their applications in the context of standard ellipsometry. Our purpose is not only to give a simple technical description of these aspects but also to show their interest for non-experienced and experienced users of standard

ellipsometry. The chapter is organized in different sections but it can be divided in three conceptual parts.

The first one starts by recalling the Jones and Stokes formalisms used in standard and Mueller ellipsometry respectively. Next we introduce the basic polarimetric effects such as dichroism, retardance, depolarization and polarizance. Those are the basic "building blocks" which allow to interpret the information brought by Mueller matrices in physical terms. Next we show different ways to decompose Mueller matrices into sums or products of simpler matrices which can be helpful to understand the physics and/or the structure of the samples. The next section consists of an overview of the basics of an ellipsometric measurement, and the usual process of treating ellipsometric data (inversion problem). This section emphasizes the features common to standard and Mueller ellipsometry.

The next section opens the second conceptual part of the chapter, devoted to instrumentation. The section gives an overview of the four most widely used types of ellipsometers. A special attention will be paid to show what can be measured with these ellipsometric configurations, or in other words, which elements of the Mueller matrix can be accessed. The purpose is to show the limitations of standard ellipsometers and the advantages brought by modern Mueller ellipsometers. In the next section we compare Generalized and Mueller ellipsometry, to show their equivalence for the study of complex but non-depolarizing samples, and the interest of measuring the full Mueller matrix when depolarization is present. The next section is devoted to various instrumentation issues of Mueller ellipsometers. We start recalling the basic theoretical background needed to design optimal polarimeters. Special emphasis is put on the concept of condition number, which has been used as a figure of merit to predict the optical performance of the designed instrument. Calibration is also an important issue because it strongly influences the quality of the measurements. We provide a concise description of a particularly useful, robust and versatile calibration procedure called "eigenvalue calibration method" (ECM). In the following section, we review different Mueller ellipsometers developed by the authors of the present chapter: (i) a spectroscopic system based on liquid crystal variable retarders working in the visible and near infrared, (ii) a broadband spectroscopic Mueller ellipsometer which has been adapted to work in the mid-infrared range, (iii) an imaging Mueller ellipsometer based on liquid crystal retarders to work with macroscopic samples with a characteristic size of few centimeters, (iv) an angle-resolved imaging Mueller ellipsometer coupled to a microscope in order to analyze tiny parts samples with a high degree of magnification.

The third part is devoted some examples of application of these Mueller ellipsometers. These applications, namely optical metrology and determination of dielectric functions of materials in broad spectral ranges, are representative of very common uses of Mueller ellipsometers. Optical metrology means the determination of physical dimensions such as thin film thicknesses, profile reconstruction of one-dimensional or two-dimensional diffraction gratings, or the overlay (misalignment) between stacked patterned structures. we show that even if we use only visible light we are able to push the accuracy of the optical metrology to the nanometer scale. Finally, we present some very recent and promising results of Mueller imaging for the detection and staging of
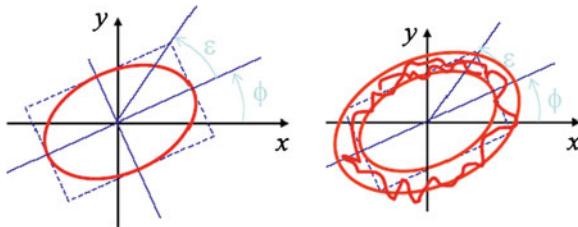
Fig. 2.1 Examples of electric field trajectories in the plane perpendicular to the propagation direction for fully polarized (*left*) or partially polarized (*right*) light waves

cancer both ex-vivo, and in vivo. In the conclusion, we try to summarize our vision of the status and the possibilities of Mueller ellipsometry, both from the instrumental point of view and for "real world" applications.

2.2 The Polarization of Light

In this section, we briefly review the most widely used theoretical descriptions of the light polarization properties, namely the Jones formalism for totally polarized light and the Stokes-Mueller formalism, which is the most general representation and can adequately account for any polarization states. The polarimetric properties of any sample are then defined from the changes this sample introduces in the polarization state of a probe light beam. In turn, these properties may be used for various purposes, from very well established applications (such as material and thin film characterizations) to more advanced ones, such as remote sensing and/or medical diagnosis.

As described in textbooks on electromagnetism [4], when a light ray propagates (through an isotropic or weakly birefringent medium) along the z direction, the electric field vector $\mathbf{E}$ is confined to vibrate in a x–y plane perpendicular to z, as illustrated in Fig. 2.1.

For totally polarized states, the electric field $\mathbf{E}$ describes an ellipse, characterized by its ellipticity ε and the azimuth of its major axis ϕ. The particular cases of linear and circular polarizations respectively correspond to $\varepsilon = 0$, and $\phi = 45°$. In contrast, partially polarized states correspond to more disordered motions of the electric field, which can be properly described only statistically, from cross-correlation functions, as discussed below.

2.2.1 The Jones Formalism

As mentioned above, the Jones formalism is well adapted to the description of totally polarized states. Any elliptical motion of $\mathbf{E}$ can be decomposed along the x and y axes,

with real amplitudes A_i and phases ϕ_i, $(i = x, y)$ which can be lumped into complex numbers E_i. Because in ellipsometry measurements are in general performed either in reflection or transmission configurations, it is common to choose as reference for the coordinates x and y, the directions parallel and perpendicular with respect to the plane of incidence, called p and s respectively. Then for transversal electromagnetic waves the Jones vector for the electric field is given by:

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} E_p \\ E_s \end{pmatrix} = \begin{pmatrix} A_p e^{i\varphi_p} \\ A_s e^{i\varphi_s} \end{pmatrix} \tag{2.1}$$

Stated in the form (2.1), the Jones vector contains also an overall phase factor, which may be important in some cases, when the polarized beam under study interferes with another beam. However, as long as only single-beam ellipsometry is concerned, this overall phase can be removed, for example by setting $\phi_p = 0$. In the absence of depolarization, the interaction with a sample transforms the Jones vector of the incident beam into another Jones vector, by a linear transformation:

$$\begin{pmatrix} E_p^{out} \\ E_s^{out} \end{pmatrix} = \begin{pmatrix} J_{pp} & J_{ps} \\ J_{sp} & J_{ss} \end{pmatrix} \begin{pmatrix} E_p^{in} \\ E_s^{in} \end{pmatrix} \tag{2.2}$$

where the J_{ij} are the elements of the Jones matrix. In a similar way as for Jones vectors, if one is interested only in the polarimetric properties of the sample and not its overall optical path (or phase shift) then one element can be taken as a real number (phase set arbitrarily to zero), and the Jones matrix depends on seven real parameters. This dependence can be further reduced to six if the overall amplitude transmission (or reflectivity) is also neglected. For plane and isotropic samples the Jones matrix in (2.2) takes on a special simple form: diagonal. It turns out that in practice the majority of substrates and thin films produced in research or industrial laboratories are isotropic, which makes the study by ellipsometry particularly simple.

$$\begin{pmatrix} E_x^{in} \\ E_y^{in} \end{pmatrix} = \begin{pmatrix} r_p & 0 \\ 0 & r_s \end{pmatrix} \begin{pmatrix} E_x^{in} \\ E_y^{in} \end{pmatrix} \tag{2.3}$$

Jones matrix elements can be interpreted in terms of the Fresnel reflection coefficients in polarization parallel, r_p, and perpendicular, r_s, to the plane of incidence. If the measurement is performed in transmission configuration, r_p and r_s must be substituted by the respective equivalent Fresnel transmission coefficients t_p and t_s (Fig. 2.2).

Ellipsometry measures the change on the polarization state of a beam after reflection or transmission by a sample. In the simplest case of isotropic samples, standard ellipsometry measures a couple of values Ψ and Δ called the ellipsometric angles. The angles Ψ and Δ are usually defined from the ratio ρ,
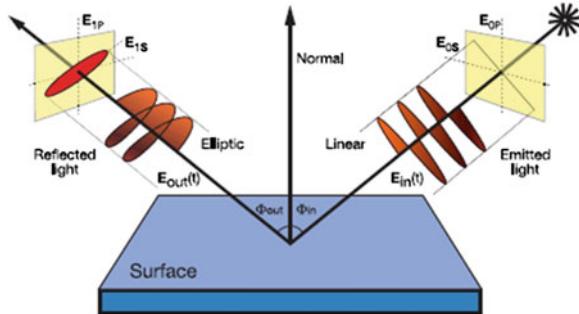
Fig. 2.2 Schematic view of a ellipsometric measurement in reflection configuration. The polarized beam is incident on the sample from the *right side*. After reflection the polarization state of the beam has changed and pursuits its propagation to the *left side*

$$\rho = \frac{r_p}{r_s} = \tan \Psi e^{i\Delta} \tag{2.4}$$

where

$$\tan \Psi = \frac{|r_p|}{|r_s|} \quad \text{and} \quad \Delta = \delta_p - \delta_s \tag{2.5}$$

Thus, $\tan \Psi$ is the amplitude ratio upon reflection, and Δ is the difference in phase shift. As it will be shown in a subsequent section of this chapter, Sect. 2.5.3, standard ellipsometers do not measure directly Ψ and Δ, but functions of them.

2.2.2 The Stokes-Mueller Formalism

2.2.2.1 General Polarization States: Coherence and Stokes Vectors

For partially depolarized states, the disordered motion of the electric field in the (x, y) plane schematically shown in Fig. 2.1, can be properly described only by its *statistical properties* instead of its instantaneous values. For this reason it is preferable to use field intensities instead of amplitudes. At first sight, one might think that a full probability distribution of the electric field **E** would be needed to fully characterize such states. In fact, as long as only intensity measurements can be performed with state-of-the art detectors at optical frequencies, all that is needed to predict the result of any classical measurement are the *second moments* (statistical averages of quadratic functions) of the electric field distributions. As a result, in the framework of linear optics, any possible polarization state of partially polarized field can be fully characterized by a four dimensional vector, called the field coherence vector **C**, defined for any set of a orthogonal axes (p, s). In the context of ellipsometry, the p and s axes are traditionally chosen to be oriented along the directions parallel and

perpendicular to the plane of incidence. Accordingly C can be written as:

$$\mathbf{C}^T = (C_1, C_2, C_3, C_4) = \left(\left\langle E_p E_p^* \right\rangle, \left\langle E_p E_s^* \right\rangle, \left\langle E_s E_p^* \right\rangle, \left\langle E_s E_s^* \right\rangle \right) \qquad (2.6)$$

where the first and last components are real while the other two are complex conjugates of each other. However, in practice the most widely used vector to characterize arbitrary polarization states is the Stokes vector :

$$\mathbf{S} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I_p + I_s \\ I_p - I_s \\ I_{45°} - I_{-45°} \\ I_L - I_R \end{pmatrix} = \begin{pmatrix} \left\langle E_p E_p^* + E_s E_s^* \right\rangle \\ \left\langle E_p E_p^* - E_s E_s^* \right\rangle \\ \left\langle E_p E_s^* + E_s E_p^* \right\rangle \\ i \left\langle E_p E_s^* - E_s E_p^* \right\rangle \end{pmatrix} = \mathbf{AC} \quad \text{with}$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{pmatrix} \qquad (2.7)$$

The popularity of the four dimensional real Stokes vector is certainly due to its immediate relationship with the directly measurable quantities $I_p, I_s, I_{+45}, I_{-45}$ i.e. the intensities which would be measured through ideal linear polarizers oriented along the p, s, $p + 45°$ and $p-45°$ in the plane perpendicular to the direction of propagation, while I_L and I_R would be the intensities transmitted by left and right circular polarizers [1]. Contrarily to the Stokes vectors, the Jones vector, which is defined in terms of electric field amplitudes, cannot be measured at optical frequencies.

In the most general case of partially polarized light the brackets at the right hand sides of (2.6) and (2.7) stand for all possible ways to take averages, e.g. spatially, spectrally or temporally, depending on the sample and measurement conditions. Thus, partially polarized states can be viewed as *incoherent superpositions of fully polarized states with different polarizations*, with simple addition of intensities and no interference effect. Conversely, for fully polarized states the field amplitudes are well defined and there is no need of averaging whatsoever. The corresponding brackets can thus be removed from (2.6) and (2.7).

Within the Stokes formalism, the *degree of polarization ρ_S* related to a given Stokes vector **S** is defined as:

$$\rho_S = \frac{\sqrt{Q^2 + U^2 + V^2}}{I} \qquad (2.8)$$

This parameter varies between 0, for totally depolarized (fully disordered) states; and 1, for totally polarized states. This is an important difference between Jones and

Stokes vectors: *while Jones vectors may have any complex components, without any limitation whatsoever, this is no longer true for Stokes vectors*, for which ρ_s must be comprised between 0 and 1 to actually represent a physically realizable polarization state.

2.2.2.2 Interaction with a Sample: Mueller Matrices

Let us first consider the transformation of a fully polarized light by interaction with a non-depolarizing sample. From the transformation described within the Jones formalism by (2.2) we immediately obtain

$$(E_i E_j^*)^{out} = \sum_{k,l} J_{ik} J_{jl}^* \ (E_k E_l^*)^{in} \tag{2.9}$$

If we now consider the general case of a partially depolarized state interacting with a partially depolarizing sample, then both sides of (2.9) must be averaged. Now, in linear optics there is absolutely no reason to expect any statistical correlations between the fluctuations of the Jones matrix elements characterizing the sample and those of the incoming field amplitudes, as the light field cannot affect the sample properties in any way at the intensities typical of ellipsometric measurements. As a result, the quadratic functions of the Jones matrix and of the field amplitudes can be averaged separately,

$$\left\langle E_i E_j^* \right\rangle^{out} = \sum_{k,l} \left\langle J_{ik} J_{jl}^* \right\rangle \left\langle E_k E_l^* \right\rangle^{in} \quad \text{or} \quad \mathbf{C}^{out} = \mathbf{F} \ \mathbf{C}^{in} \tag{2.10}$$

where the matrix $\mathbf{F}$ is obtained by renumbering the couples of indices *ik* and *jl* in the same way as it was done for the coherence vector $\mathbf{C}$ in (2.6):

$$\mathbf{F} = \begin{bmatrix} \left\langle J_{pp} J_{pp}^* \right\rangle & \left\langle J_{pp} J_{ps}^* \right\rangle & \left\langle J_{ps} J_{pp}^* \right\rangle & \left\langle J_{ps} J_{ps}^* \right\rangle \\ \left\langle J_{pp} J_{sp}^* \right\rangle & \left\langle J_{pp} J_{ss}^* \right\rangle & \left\langle J_{ps} J_{sp}^* \right\rangle & \left\langle J_{ps} J_{ss}^* \right\rangle \\ \left\langle J_{sp} J_{pp}^* \right\rangle & \left\langle J_{sp} J_{ps}^* \right\rangle & \left\langle J_{ss} J_{pp}^* \right\rangle & \left\langle J_{ss} J_{ps}^* \right\rangle \\ \left\langle J_{sp} J_{sp}^* \right\rangle & \left\langle J_{sp} J_{ss}^* \right\rangle & \left\langle J_{ss} J_{sp}^* \right\rangle & \left\langle J_{ss} J_{ss}^* \right\rangle \end{bmatrix} = \left\langle \mathbf{J} \otimes \mathbf{J}^* \right\rangle \tag{2.11}$$

where $\otimes$ stands for a Kronecker product. While the coherence vector $\mathbf{C}$ is transformed by $\mathbf{F}$, upon interaction with a sample the Stokes vector is transformed by the well-known Mueller matrix $\mathbf{M}$ [5–7]:

$$\mathbf{S}^{out} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}^{out} = \mathbf{M}\mathbf{S}^{in} = \begin{pmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{pmatrix} \cdot \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}^{in} \tag{2.12}$$

By combining (2.7), (2.10) and (2.11) we immediately obtain

$$\mathbf{M} = \mathbf{A}\, \mathbf{F}\, \mathbf{A}^{-1} = \mathbf{A}\, \langle \mathbf{J} \otimes \mathbf{J}^* \rangle\, \mathbf{A}^{-1} \tag{2.13}$$

which can be written in an expanded notation for a general case as:

$$\mathbf{M} = \begin{pmatrix} \frac{1}{2}\langle |J_{pp}|^2 + |J_{ss}|^2 + |J_{sp}|^2 + |J_{ps}|^2 \rangle & \frac{1}{2}\langle |J_{pp}|^2 - |J_{ss}|^2 + |J_{sp}|^2 - |J_{ps}|^2 \rangle & \langle \mathrm{Re}\left(J_{ps}J_{pp}^* + J_{ss}J_{sp}^*\right)\rangle & -\langle \mathrm{Im}\left(J_{ps}J_{pp}^* + J_{ss}J_{sp}^*\right)\rangle \\ \frac{1}{2}\langle |J_{pp}|^2 - |J_{ss}|^2 - |J_{sp}|^2 + |J_{ps}|^2 \rangle & \frac{1}{2}\langle |J_{pp}|^2 + |J_{ss}|^2 - |J_{sp}|^2 - |J_{ps}|^2 \rangle & \langle \mathrm{Re}\left(J_{ps}J_{pp}^* - J_{ss}J_{sp}^*\right)\rangle & \langle \mathrm{Im}\left(-J_{ps}J_{pp}^* + J_{ss}J_{sp}^*\right)\rangle \\ \langle \mathrm{Re}\left(J_{sp}J_{pp}^* + J_{ss}J_{ps}^*\right)\rangle & \langle \mathrm{Re}\left(J_{sp}J_{pp}^* - J_{ss}J_{ps}^*\right)\rangle & \langle \mathrm{Re}\left(J_{ss}J_{pp}^* + J_{sp}J_{ps}^*\right)\rangle & \langle \mathrm{Im}\left(-J_{ss}J_{pp}^* + J_{sp}J_{ps}^*\right)\rangle \\ \langle \mathrm{Im}\left(J_{sp}J_{pp}^* + J_{ss}J_{ps}^*\right)\rangle & \langle \mathrm{Im}\left(J_{sp}J_{pp}^* - J_{ss}J_{ps}^*\right)\rangle & \langle \mathrm{Im}\left(J_{ss}J_{pp}^* + J_{sp}J_{ps}^*\right)\rangle & \langle \mathrm{Re}\left(J_{ss}J_{pp}^* - J_{sp}J_{ps}^*\right)\rangle \end{pmatrix} \tag{2.14}$$

Of course, in the absence of depolarization, i.e. when the Jones matrix is well defined, the brackets meaning averages can be removed in (2.14). Furthermore, if the Jones matrix is diagonal, as it is the case for standard ellipsometry, only the upper left and lower right 2×2 sub-matrices do not vanish, and the Mueller matrix can be recast in terms of the ellipsometric angles Ψ and Δ as:

$$\mathbf{M}\,(\tau, \Psi, \Delta) = \tau \begin{pmatrix} 1 & -\cos(2\Psi) & 0 & 0 \\ -\cos(2\Psi) & 1 & 0 & 0 \\ 0 & 0 & \sin(2\Psi)\cos\Delta & \sin(2\Psi)\sin\Delta \\ 0 & 0 & -\sin(2\Psi)\sin\Delta & \sin(2\Psi)\cos\Delta \end{pmatrix} \tag{2.15}$$

Due to the capability of the Stokes vectors to describe *any* polarization state, the Mueller matrix can fully describe the polarimetric properties of *any* sample, be it depolarizing or not. In other words, Mueller polarimetry is the only technique able to fully characterize the polarization responses of any sample, in any measurement conditions.

In contrast with the Jones matrix, the Mueller matrix does not carry any information about the overall optical phase shift introduced by the sample. So, depending on whether the overall transmission (or reflectivity) of the sample is of interest or not, the Mueller matrix may be considered in its original or in its *normalized* form: in the latter case, all its elements are divided by the upper left element M_{11} which is set equal 1.

2.2.2.3 Coherence Matrix: Physical Realizability and Depolarizing Character of Mueller Matrices

The coherence matrix, $\mathbf{N}$, is an interesting object, obtained from a smart rearrangement of the matrix $\mathbf{F}$. To obtain the coherence matrix it is necessary to redefine the

Jones matrix in a "vector" form $\mathbf{J}^V$ as:

$$\left[\mathbf{J}^V\right]^T = \left(J_1^V, J_2^V, J_3^V, J_4^V\right) = \left(J_{pp}, J_{ps}, J_{sp}, J_{ss}\right) \tag{2.16}$$

then the matrix $\mathbf{N}$ can be expressed in terms of matrix $\mathbf{F}$ as:

$$\mathbf{N} = \left\langle \left[\mathbf{J}^V\right] \otimes \left(\left[\mathbf{J}^V\right]^T\right)^* \right\rangle = \begin{bmatrix} \langle J_{pp}J_{pp}^*\rangle & \langle J_{pp}J_{ps}^*\rangle & \langle J_{pp}J_{sp}^*\rangle & \langle J_{pp}J_{ss}^*\rangle \\ \langle J_{ps}J_{pp}^*\rangle & \langle J_{ps}J_{ps}^*\rangle & \langle J_{ps}J_{sp}^*\rangle & \langle J_{ps}J_{ss}^*\rangle \\ \langle J_{sp}J_{pp}^*\rangle & \langle J_{sp}J_{ps}^*\rangle & \langle J_{sp}J_{sp}^*\rangle & \langle J_{sp}J_{ss}^*\rangle \\ \langle J_{ss}J_{pp}^*\rangle & \langle J_{ss}J_{ps}^*\rangle & \langle J_{ss}J_{sp}^*\rangle & \langle J_{ss}J_{ss}^*\rangle \end{bmatrix}$$

$$= \begin{bmatrix} F_{11} & F_{12} & F_{21} & F_{22} \\ F_{13} & F_{14} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{41} & F_{42} \\ F_{33} & F_{34} & F_{43} & F_{44} \end{bmatrix} \tag{2.17}$$

where the superscripts, T and $*$ stand for transposed and complex conjugated. When applied successively, they are equivalent to the Hermitian conjugate. The relation between the $\mathbf{N}$ and $\mathbf{F}$ matrices allows to combine expressions (2.17), (2.11) and (2.14) to write $\mathbf{N}$ in terms of the elements of the Mueller matrix $\mathbf{M}$ as:

$$\mathbf{N} = \frac{1}{2} \begin{bmatrix} M_{11} + M_{22} + & M_{13} + M_{23} + & M_{31} + M_{32} - & M_{33} + M_{44} + \\ M_{12} + M_{21} & i(M_{14} + M_{24}) & i(M_{41} + M_{42}) & i(M_{34} - M_{43}) \\ \\ M_{13} + M_{23} - & M_{11} - M_{22} - & M_{33} - M_{44} - & M_{31} - M_{32} - \\ i(M_{14} + M_{24}) & M_{12} + M_{21} & i(M_{34} + M_{43}) & i(M_{41} - M_{42}) \\ \\ M_{31} + M_{32} + & M_{33} - M_{44} + & M_{11} - M_{22} + & M_{13} - M_{23} + i \\ i(M_{41} + M_{42}) & i(M_{34} + M_{43}) & M_{12} - M_{21} & (M_{14} - M_{24}) \\ \\ M_{33} + M_{44} - & M_{31} - M_{32} & M_{13} - M_{23} - & M_{11} + M_{22} - \\ i(M_{34} - M_{43}) & +i(M_{41} - M_{42}) & i(M_{14} - M_{24}) & M_{12} - M_{21} \end{bmatrix} \tag{2.18}$$

An alternative definition of the matrix $\mathbf{N}$, giving the same result as (2.17), has been proposed by [8]. As it can be seen, the coherence matrix has the same elements as the matrix $\mathbf{F}$, hence it carries the same information. By construction, the matrix $\mathbf{N}$ is Hermitian, implying its eigenvalues are real. This algebraic property will be used to define some of the Mueller matrix decompositions detailed in a forthcoming section, Sect. 2.4. Moreover by construction it can be easily seen that the coherence matrix has the form of a variance-covariance matrix. In the following, in order to keep the same terminology as other authors, we will refer to matrix $\mathbf{N}$ as coherence matrix instead of variance-covariance matrix.

2.2.2.4 Physical Realizability of Mueller Matrices

We already pointed out in paragraph 2.2.2.1 that any two dimensional complex vector may represent the Jones vector of a physically realizable polarization state, while a four dimensional real vector is not necessarily an acceptable Stokes vector, as its degree of polarization ρ_s must be comprised between 0 and 1.

The situation is quite analogous for Jones and Mueller matrices. Any 2×2 complex matrix is a physically acceptable Jones matrix (provided we consider the polarization behavior of optical amplifiers, to include matrices with coefficients having moduli larger than 1), while *any* 4×4 *real matrix* **M** *is not necessarily a physically realizable Mueller matrix*. An obvious necessary condition is that any acceptable Stokes vector must be transformed by **M** into another acceptable Stokes vector, with ρ_s between 0 and 1. However, this condition is not sufficient.

The real criterion of physical realizability of **M** as a Mueller matrix is directly related to our previous definition of a Mueller matrix as a linear combination of the second moments of a probabilistic (and not deterministic) Jones matrix. A necessary and sufficient condition for the existence of such matrix, and thus of the acceptability of **M** as a Mueller matrix is that *the matrix* **N** *calculated from* **M** *by* (2.18) *is an acceptable coherence matrix, i.e. that this Hermitian matrix is definite positive* [9, 10] (all eigenvalues are non-negative, and at least one is strictly positive).

#### 2.2.2.5 Depolarizing and Non-Depolarizing Mueller Matrices
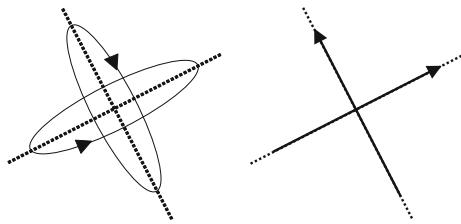
From the above definitions, it should be quite clear now that a non-depolarizing sample is characterized by a well-defined (deterministic) Jones matrix. As a result, all the averaging introduced in Eqs. (2.10), (2.11), (2.14) and (2.17) can be safely eliminated. If so, Eq. (2.17) shows that **N** can be seen as a "projector" of the C^4 space onto the vector $\mathbf{J}^V$.

Actually, a necessary and sufficient condition for a Mueller matrix **M** *to be non-depolarizing and equivalent to a Jones matrix* **J** *is that its coherence matrix* **N** *has only one strictly positive eigenvalue The eigenvector associated with this eigenvalue is the vector* $\mathbf{J}^V$ *related to* **J** *by* $\mathbf{J}^V$ (2.16).

As discussed later in Sect. 2.3.3 devoted to pure depolarizers, there are many ways to quantify the depolarizing character of a given sample, after more or less complex treatments of the Mueller matrix M. However, a very simple quantifier, the quadratic depolarization index P_q proposed by Gil [11] deserves a special mention here due to its ease of implementation:

$$P_q = \sqrt{\frac{\sum\limits_{ij} M_{ij}^2 - M_{11}^2}{3M_{11}^2}} = \sqrt{\frac{trace(\mathbf{M}^T\mathbf{M}) - M_{11}^2}{3M_{11}^2}} \qquad (2.19)$$

Fig. 2.3 *Left* a pair of ellip-
tically polarized orthogonal
states. *Right* a pair of linearly
polarized orthogonal states

where the superscript *T* stands for the transpose operation and, *trace*, indicates the algebraic trace operator. The quadratic depolarization index P_q varies from 0, for a perfect depolarizer (only M_{11} is nonzero) to 1, for non-depolarizing matrices.

2.3 The Essential Polarimetric Properties of Any Sample

Generally speaking, the polarimetric response of a given sample describes how the incident light polarization is changed due to the interaction with the sample. In spite of its apparent complexity, this response can be rationalized in terms of three fundamental properties, namely the sample diattenuation, retardation, and depolarization. In many cases of practical interest, among which the usual ellipsometric characterization of isotropic materials or thin films, all these properties can be unambiguously defined from the measured data. The fundamental polarimetric properties used as "building blocks" to characterize more complex systems are given by pure diattenuators, pure retarders and depolarizers.

To understand these properties it is useful to use the concept of pairs of fully polarized orthogonal eigenstates, represented in Fig. 2.3. Each eigenstate is characterized by its length and ellipticity. An ellipticity equal to zero corresponds to linearly polarized light, an ellipticity equal to ± 1 corresponds to circularly polarized light, while other values of ellipticity correspond to elliptically polarized light. The sign of the ellipticity states the difference between clockwise or counterclockwise rotation. For the vast majority of usual polarization optical components, such as retardation plates or polarizers, these eigenstates are actually linearly polarized, (also shown in Fig. 2.3).

Linear diattenuators, which can be seen as partial linear polarizers, transmit (or reflect) each of their eigenstates without altering their ellipticity nor azimuth, but may change their intensities. Simple linear retarders, which can be assimilated to wave plates, transmit their eigenstates without changing their respective ellipticity, azimuth and intensity but modify their relative phases (or optical paths). These effects are schematically represented in Fig. 2.4.

In contrast with diattenuators and retarders, ideal depolarizers do not leave any polarization state invariant, excepted the totally depolarized one. Actually these components reduce the light degree of polarization ρ_s defined in (2.8). In particular, these
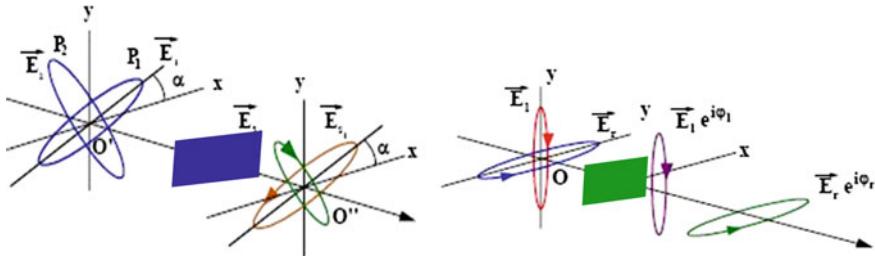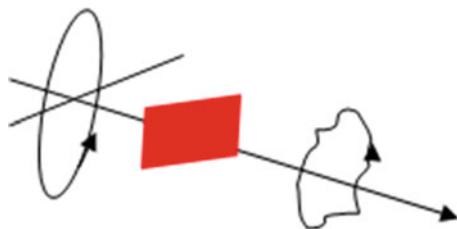
Fig. 2.4 Illustration of the transformation of polarization eigenstates by simple diattenuators (*left*) or simple retarders (*right*)

Fig. 2.5 Analogous to Fig. 2.4, for a typical depolarizer

components transform totally polarized states into partially polarized ones, as illustrated in Fig. 2.5.

As mentioned above, partially polarized states can be physically interpreted as incoherent superpositions of totally polarized states with different polarizations. As a result, a depolarizer can be described as an optical system with different non-depolarizing responses which add up incoherently. In practice, this occurs for samples featuring spatial, spectral or temporal inhomogeneities resulting in different output polarization states leading to intensity signals which are integrated by the detector. If retarders and diattenuators sharing the same eigenvectors are present in a sample or in a given medium, their combined effect on the final state of a polarized beam does not depend on the order on which each one appears; in other words, their polarization effects are commutative. The same is not true for a depolarizer. When combined with retardance and diattenuation, the position where depolarization takes place in the sample matters, and determines the final polarization state of the beam.

In the following we provide a description of each one of the elementary polarimetric properties together with a representation of the corresponding Mueller matrices. More details can be found in [12]. The matrix representation is the basis of a powerful method of interpretation of polarimetric measurements. The method consists on decomposing the measured Mueller matrix into basic polarimetric effects, and it will be detailed in a forthcoming section, Sect. 2.4.

2.3.1 Diattenuation and Linear Diattenuators (Polarizers)

Diattenuation, the polarimetric property of the diattenuators, is defined in practice by a scalar, called D, characterizing the maximum variation of transmitted (or reflected) light intensity with as a function of the incident polarization state. Diattenuation is defined as follows

$$D = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{2.20}$$

where I_{max} and I_{min} correspond to the intensities of the two transmitted (or reflected) eigenstates. This definition recalls the definition of the ellipsometric angle Ψ. Indeed $\tan^2 \Psi$ can be written as I_{max}/I_{min}. The square is justified because D is defined for intensities and Ψ for amplitudes of the electromagnetic field. From this relation it is easy to derive:

$$D = \frac{1 - \tan^2 \Psi}{1 + \tan^2 \Psi} = \cos 2\Psi. \tag{2.21}$$

For ideal linear polarizers, I_{min} is close to 0, (in practice the ratio I_{min}/I_{max} is typically of the order of 10^{-3} to 10^{-6} I_{max}), thus D is almost 1 and Ψ almost 0° or 90°. The notion of diattenuation can be further extended to a vector, which gives information about the orientation if diattenuator eigenstates. More specifically, if the vector diattenuation is defined as:

$$\mathbf{D} = D \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} D_{horizontal} \\ D_{45°} \\ D_{circular} \end{pmatrix} \tag{2.22}$$

with $d_1^2 + d_2^2 + d_3^2 = 1$, then the polarization eigenstates Stokes vectors $\mathbf{S}_{min}$, $\mathbf{S}_{max}$ are given by

$$\mathbf{S}_{max}^T = (1, d_1, d_2, d_3,), \quad \mathbf{S}_{min}^T = (1, -d_1, -d_2, -d_3,) \tag{2.23}$$

The three components of the diattenuation vector $\mathbf{D}$ define respectively the horizontal, the 45°, and the circular diattenuation. The vector diattenuation of any sample is a very simple function of the first row of the sample Mueller matrix:

$$\mathbf{D} = \frac{1}{M_{11}} \begin{pmatrix} M_{12} \\ M_{13} \\ M_{14} \end{pmatrix} \tag{2.24}$$

The Mueller matrix of a pure diattenuator can be expressed in terms of the scalar and the vector diattenuation as follows:

$$\mathbf{M} = \tau \begin{pmatrix} 1 & \mathbf{D}^T \\ \mathbf{D} & (\mathbf{m}_d) \end{pmatrix}; \quad \text{where} \quad \mathbf{m}_d = \sqrt{1 - D^2}\mathbf{I}_3 + \left(1 - \sqrt{1 - D^2}\right)\mathbf{D}\,\mathbf{D}^T \quad (2.25)$$

This Mueller matrix has been written in a contracted form. The elements of the first row and column are represented by the diattenuation vector. $\mathbf{m_d}$ is a 3×3 symmetric sub-matrix which is function of the identity (3×3) matrix, $\mathbf{I}_3$, the vector and the scalar diattenuation. Finally, τ represents the overall transmission or reflectivity or the sample when the incident light is totally depolarized.

For the practically very interesting case of linear diattenuators (e.g. whose eigenstates are linearly polarized) the Mueller matrix $\mathbf{M}_D$ reads:

$$\mathbf{P} = \frac{\tau_P}{2}$$

$$\begin{pmatrix} 1 & \cos(2\theta)\cos 2\Psi & \sin(2\theta)\cos 2\Psi & 0 \\ \cos(2\theta)\cos 2\Psi & \cos^2(2\theta) + \sin^2(2\theta)\sin 2\Psi & \cos(2\theta)\sin(2\theta)(1 - \sin 2\Psi) & 0 \\ \sin(2\theta)\cos 2\Psi & \cos(2\theta)\sin(2\theta)(1 - \sin 2\Psi) & \sin^2(2\theta) + \cos^2(2\theta)\sin 2\Psi & 0 \\ 0 & 0 & 0 & \sin 2\Psi \end{pmatrix}$$

$$(2.26)$$

where θ is the azimuth of the high transmission polarization direction with respect to the x axis and Ψ is the ellipsometric angle. The diattenuator vector takes the particularly simple form

$$\mathbf{D} = \cos 2\Psi \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \\ 0 \end{pmatrix} \quad (2.27)$$

2.3.2 Retardance and Linear Retarders

The time delay generated between two eigenstates after propagation through a pure retarder is represented by the scalar retardance, R. In a frequency representation of the electromagnetic fields, the time delay is represented by the phase difference between the two eigenstates. Similarly to the diattenuation, it is possible to define a vector retardance, $\mathbf{R}$, as

$$\mathbf{R} = R \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} R_H \\ R_{45°} \\ R_C \end{pmatrix} \quad (2.28)$$

with $r_1^2 + r_2^2 + r_3^2 = 1$. Again, the Stokes vectors of the fast and slow eigenstates $\mathbf{S}_f$, $\mathbf{S}_s$ are given respectively by

$$\mathbf{S}_f^T = (1, r_1, r_2, r_3,), \quad \mathbf{S}_s^T = (1, -r_1, -r_2, -r_3,) \quad (2.29)$$

A pure retarder can be described geometrically as rotation in the space of Stokes vectors. Mathematically the Mueller matrix $\mathbf{M_R}$ of the retarder can be written in compact notation as:

$$\mathbf{M}_R = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{m_R} \end{pmatrix}, \quad \text{and} \quad (\mathbf{m_R})_{ij} = \delta_{ij} \cos(R) + r_i r_j \, (1 - \cos R) \sum_{k=1}^{3} \varepsilon_{ijk} r_k \sin R$$

$$(2.30)$$

where the 3×3 sub-matrix, $\mathbf{m_R}$, is orthogonal and has a unit determinant, $\det(\mathbf{m_R}) = +1$. $\mathbf{0}$ represents the null vector. δ_{ij} is the Kronecker symbol, and ε_{ijk} is the Levi-Civita permutation sign. Scalar and vector retardances can be easily determined from the measured Mueller matrices as follows:

$$R = \cos^{-1} \left(\frac{trace\,(M_R)}{2} - 1 \right) \tag{2.31}$$

$$r_i = \frac{1}{2 \sin R} \sum_{j,k=1}^{3} \varepsilon_{ijk} (\mathbf{m_R})_{jk} \tag{2.32}$$

Again, the case of linear retarders is of particular interest. For such a component, with its fast axis oriented at an azimuth, θ, with respect to the x axis and a scalar retardation Δ, the Mueller matrix reads :

$$\mathbf{M}_R(\theta, \Delta) =$$
$$\tau_R \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\theta) + \sin^2(2\theta)\cos(\Delta) & \cos(2\theta)\sin(2\theta)\,(1 - \cos(\Delta)) & -\sin(2\theta)\sin(\Delta) \\ 0 & \cos(2\theta)\sin(2\theta)\,(1 - \cos(\Delta)) & \sin^2(2\theta) + \cos^2(2\theta)\cos(\Delta) & \cos(2\theta)\sin(\Delta) \\ 0 & \sin(2\theta)\sin(\Delta) & -\cos(2\theta)\sin(\Delta) & \cos(\Delta) \end{pmatrix}$$

$$(2.33)$$

It is straightforward to check that the scalar retardation R given by (2.31) is nothing else but the ellipsometric angle Δ, and the retardance vector is

$$\mathbf{R} = \Delta \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \\ 0 \end{pmatrix} \tag{2.34}$$

2.3.3 Depolarization and Pure Depolarizers

The Mueller matrix of a general depolarizer, $\mathbf{M}_\delta$, is given in compact notation by:

$$\mathbf{M}_\delta = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{m}_\delta \end{pmatrix} \qquad (2.35)$$

where $\mathbf{m}_\delta$ is a 3×3 real symmetric matrix. Thus this matrix can be reduced to a diagonal form in a suitable orthonormal basis formed by three (3-dimensional) eigenvectors $\mathbf{v}_i$. As a result, in a basis formed by the four Stokes vectors

$$\mathbf{S}_0^T = (1, 0, 0, 0,), \quad \text{and} \quad \mathbf{S}_i^T = \left(1, \mathbf{v}_i^T\right), (1 \le i \le 3) \qquad (2.36)$$

the Mueller matrix $\mathbf{M}_\delta$ of the depolarizer becomes diagonal

$$\mathbf{M}_\delta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & c \end{pmatrix} \qquad (2.37)$$

where the eigenvalues a, b, c of $\mathbf{m}_\delta$ are real numbers comprised between -1 and 1. The Eq. (2.37) above shows that for the three Stokes vectors $\mathbf{S}_i$ defined in (2.36) the degree of polarization is reduced by a factor equal to the corresponding eigenvalue a, b or c, while the totally depolarized state remains unchanged. In other words, a general depolarizer features only one eigenpolarization, corresponding ot a totally depolarized state.

Due to the symmetry of $\mathbf{m}_\delta$, the Mueller matrices of pure depolarizers defined in (2.35), clearly depend on six parameters, i.e. the independent components of $\mathbf{m}_\delta$. Another possible choice, which may be more physically relevant, are the three eigenvalues a, b, and c, and the three Euler angles defining the directions of the normalized vectors $\mathbf{v}_i$ in the three dimensional space of the (Q, U, V) coordinates of Stokes vectors.

Due to this dependence on six parameters, in the most general case, depolarizers are more mathematically complex than retarders or diattenuators, which involve only three parameters each. Fortunately, in many situations of physical interest the symmetry properties of the sample greatly reduce the number of independent parameters. For example, when observed in forward or backward scattering geometries a suspension of spherical (or statistically isotropic) scatterers behaves as a pure depolarizer with different depolarization powers for linearly and circularly polarized incident states. Moreover, for the particular case of a suspension of spheres, the depolarization power for linear states is independent of the orientation of the incident polarization. Mathematically, these intuitive properties can be reformulated as

$$a = b \ne c \quad \text{and} \quad v_1^T = (\cos\alpha, \sin\alpha, 0), \quad v_2^T = (-\sin\alpha, \cos\alpha, 0), \quad v_3^T = (0, 0, 1) \qquad (2.38)$$

where α can be chosen arbitrarily.

However, as depolarizers are studied not only for their own sake but also to characterize the depolarizing properties of more general Mueller matrices by mathematical treatments presented in the next section, it is desirable to define a single numeri-

cal function defining the "overall" depolarizing power of a depolarizer, and which
would vary from 0 to 1, these extreme values being reached for nondepolarizing
matrices and for total depolarizers respectively. A first example of such function
is the quadratic depolarization index P_q defined in Eq. (2.19) for a general Mueller
matrix, and obviously applies also to pure depolarizers. Another possible definition
has been given specifically for depolarizers by Lu and Chipman [12].

$$\delta = 1 - \frac{1}{3} (|a| + |b| + |c|) \tag{2.39}$$

Finally, we point out that the best definition of an "overall" depolarizing power
depends on the system under consideration. This is an open field, with many new
definitions being proposed [13, 14].

2.3.4 Polarizance: Homogeneous and Inhomogeneous Systems

The concept of polarizance is linked to the ability of a sample to increase the degree
of polarization of an initially non-polarized beam. This increase can be done either by
a selective reorientation or by a selective elimination of certain vibration directions
of the electric field. Mathematically this concept is expressed as follows:

$$\mathbf{S}_{out} = \begin{pmatrix} M_{11} \\ M_{21} \\ M_{31} \\ M_{41} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.40}$$

The degree of polarization of the Stokes vector representing the final state of the beam
is called scalar polarizance P. Moreover, in analogy with the vector diattenuation and
the vector retardance the vector polarizance $\mathbf{P}$ can be defined from the Mueller matrix
elements of the polarizing element as:

$$P = \frac{\sqrt{M_{21}^2 + M_{31}^2 + M_{41}^2}}{M_{11}^2} \quad \text{and} \quad \mathbf{P} = \frac{1}{M_{11}} \begin{pmatrix} M_{21} \\ M_{31} \\ M_{41} \end{pmatrix} \tag{2.41}$$

For many systems, the diattenuation and polarizance vectors are equal: $\mathbf{P} = \mathbf{D}$. Such
systems are said to be *homogeneous*. The diattenuators and the retarders respectively
defined in Sects. 2.3.1 and 2.3.2 are indeed homogeneous.

At first glance a diattenuating system may seem necessary to partially polarize
an initially unpolarized beam. The emerging polarization being directly determined
by the diattenuation vector $\mathbf{D}$, one might expect that in all cases $\mathbf{P} = \mathbf{D}$, and thus
the notion of polarizance would eventually be pointless. This is not true, as we now
show with very simple examples of inhomogeneous systems.

- Let us first consider a setup consisting of a *polarizer followed by a perfect depolarizer*. Such a system clearly exhibits diattenuation, as the intensity transmitted by the polarizer depends on the incoming polarization. However, the beam emerging from the depolarizer is, by definition, totally depolarized, implying that an initially depolarized beam would remain totally depolarized. As a result, such a system exhibits a strong diattenuation but no polarizance.
- Let us now consider the same elementary components, but *in reverse order* (depolarizer first, polarizer afterwards). In this case, any incoming polarized beam is transformed by the depolarizer into a totally depolarized beam whose intensity does not depend on the incoming polarization. Then, the polarizer transforms this beam into a polarized one, with always the same polarization. In contrast with the previous case, now the system exhibits zero diattenuation but a strong polarizance ($P = 1$ for a perfect polarizer).

2.3.5 Summary

In this part, we have introduced the "elementary" polarimetric properties, namely diattenuation, retardance, depolarization and polarizance. As diattenuation, retardance and polarizance are defined by 3D non-normalized vectors, each of them depends on three independent parameters. On the other hand, we have seen that depolarization depends on another six parameters. So, all these polarimetric properties imply 15 independent parameters, as expected for normalized Mueller matrices (M_{11}, which is an overall transmission factor, is irrelevant for polarimetry).

However, the elementary polarimetric properties of a given Mueller matrix cannot be defined unambiguously. As shown in the next section, several decomposition procedures may be used to retrieve these properties, the final results depending on the chosen decomposition. For a correct interpretation of the data, it is therefore essential to choose properly the decomposition best adapted to the system under study, if available in the "toolbox".

2.4 Mueller Matrix Algebra: Decomposition of Mueller Matrices

Ideally, any polarimetric measurement should be interpreted by fitting to the measured Mueller matrices numerical simulations based on a relevant model, as it is very commonly done with standard ellipsometry studies of samples like stacks of isotropic thin films on plane substrates. Unfortunately this is far from being always possible with the "complex" and/or "disordered" samples such as biological tissues, which exhibit depolarization and justify the use of Mueller polarimetry. For such samples accurate models of their polarimetric responses are very difficult to elaborate.

In the absence of adequate physical models, experimental Mueller matrices can still be phenomenologically interpreted by decomposing them into simpler components with well-defined polarimetric properties. Basically, Mueller matrix decompositions can be classified into two groups:

- *sum decompositions*: Sum decompositions treat a depolarizing Mueller matrix as an incoherent addition of non-depolarizing matrices. Sum decompositions are primarily used to assess the physical realizability of the measured Mueller matrix, a condition which is not necessarily satisfied due to experimental errors, and to "filter" out the contribution these errors if needed. In other cases sum decompositions allow to isolate the individual non-depolarizing Mueller matrices which contribute to a measured depolarizing matrix.
- *product decompositions*, which describe the sample as a stack of elementary samples traversed sequentially by the light beam. These decompositions are mostly used to evaluate the diattenuation, retardation, depolarization and polarizance of the input Muller matrix, and, in some cases, to locate the various elementary polarization properties inside the sample.

2.4.1 Sum Decompositions

Sum decompositions treat the depolarizing Mueller matrix as an incoherent addition of non-depolarizing matrices. The physical image behind these decompositions is that of a beam which does not shine a single sample, but on N different ones at the same time. Consequently each sample transmits of reflects, simultaneously and independently of the other samples, a portion of the beam. To complete the image we consider that a unique detector integrates incoherently (without interferences) and simultaneously the light coming from all the N samples. An electrical analogue would be a circuit comprising N resistances connected in parallel. A current flow, arriving to the circuit, splits, and a portion of the total intensity current goes through each resistance. Once the individual current flows have gone through the resistances, they sum up again at the output of the circuit.

Sum decompositions are very useful because they may represent physical situations frequently encountered in ellipsometric measurements. Incoherent superpositions of differently polarized contributions may be caused by the sample or by the measurement system itself. Typical examples of such situations are

(a) *multiple reflections* by the double face of substrates thicker than the light coherence length, typically of the order of 0.2 mm for most spectroscopic ellipsometers.
(b) *Spatially inhomogenous samples*. Many examples can be found in this category which have in common the fact that the probe beam is much larger than the characteristic size of homogeneous zones. Thus the beam illuminates a region of the sample with different optical responses. Once all of these responses arrive at the detector they add-up incoherently and generate depolarization.

(c) *Tightly focused beams*. The reduction of beam spot sizes sometimes involves the creation of highly divergent (convergent) beams incident on the sample. A highly divergent or convergent can be understood as a bundle of well collimated beams with a specific angle of incidence each. As a result, the reflected beam is also divergent and contains a "bundle" of optical responses, which add-up incoherently when they are detected.

Many other examples can still be found which shows the interest for a decomposition which simplifies the analysis of data.

2.4.1.1 Cloude Decomposition

The most popular sum decomposition is known as the Cloude decomposition [15, 16]. Accordingly, any depolarizing matrix $\mathbf{M}$ can be represented as a weighted sum of up to four non-depolarizing Mueller matrices $\mathbf{M_i}$ in the following way:

$$\mathbf{M} = \lambda_1 \mathbf{M}_1 + \lambda_2 \mathbf{M}_2 + \lambda_3 \mathbf{M}_3 + \lambda_4 \mathbf{M}_4, \qquad (2.42)$$

with the weight factors λ_i being positive. To form the decomposition, it is necessary to evaluate the coherence matrix $\mathbf{N}$ from the original Mueller matrix $\mathbf{M}$ via the $\mathbf{F}$ matrix by Eqs. (2.11) and (2.17). Then, let the eigenvalues and the normalized eigenvectors of $\mathbf{N}$ be λ_i and $\mathbf{e_i}$, respectively. The coherence matrices $\mathbf{N_i}$ of the *non-depolarizing* components $\mathbf{M_i}$ are given by

$$\mathbf{N_i} = \mathbf{e_i} \mathbf{e}_i^{\dagger} \qquad (2.43)$$

where the symbol † stands for Hermitian conjugate Comparison of (2.43) and (2.18) immediately shows that $\mathbf{e_i}$ is nothing else but the vector form $\mathbf{J}_i^V$ defined in (2.16) of the Jones matrix $\mathbf{J_i}$ associated with the nondepolarizing matrix $\mathbf{M_i}$. Finally, the matrices $\mathbf{M_i}$ can be derived from J_i by using (2.14).

As shown previously, for a physically realizable Mueller matrix, the corresponding coherence matrix $\mathbf{N}$ is positive semi-definite and therefore, has non-negative eigenvalues λ_i that are usually sorted according to $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. Conversely, if $\lambda_4 < 0$ for instance, the matrix $\mathbf{M}$ is unphysical, i.e. $\mathbf{M}$ may transform a valid input Stokes vector into an invalid output one. When the Mueller matrix to be measured is non- or very weakly depolarizing, measurement errors may easily make it nonphysical. In such cases, Cloude decomposition provides a convenient tool to "filter out" the effect of such errors and make the measured matrix physical: to this end, once the eigenvalues λ_i and non-depolarizing matrices $\mathbf{M_i}$ are obtained, the (presumably small) negative eigenvalues are set equal to zero, and the relationship (2.42) is used one more to reconstruct a physically acceptable matrix close to the initially measured one.

In general *for a spatially inhomogeneous system the non-depolarizing components* $\mathbf{M_i}$ *do not correspond to actual physical Mueller matrices corresponding to different parts of the system* ! A simple reason for that is that the eigenvectors $\mathbf{e_i}$, of $\mathbf{N}$ form

an orthonormal basis. Then the matrices $\mathbf{M}_i = \mathbf{e}_i \mathbf{e}_i^\dagger$ have special properties which are not necessarily verified by the actual physical Mueller matrices of different parts of the system. To retrieve such matrices, and thus fully characterize such systems in spite of their inhomogeneity, more information than the simple knowledge of $\mathbf{M}$ is needed. An example of such characterization is described in Sect. 2.7.2.

2.4.1.2 Le Roy-Bréhonnet Decomposition

A special case of Cloude decomposition is that proposed by Le Roy-Bréhonnet et al. [17] representing $\mathbf{M}$ as the sum of a non-depolarizing Mueller matrix $\mathbf{M}_{nd}$ and an ideal diagonal depolarizer $\mathbf{M}_{id}$ (i.e., the kind of depolarizer introduced in the preceding section with null diagonal elements $a = b = c = 0$):

$$\mathbf{M} = \mathbf{M}_{nd} + \mathbf{M}_{id} \tag{2.44}$$

Unlike Cloude decomposition which is valid for an arbitrary depolarizing Mueller matrix, Le Roy-Bréhonnet expression is valid only if the condition $\lambda_2 = \lambda_3 = \lambda_4$ ($\neq \lambda_1$) is fulfilled. Le Roy-Bréhonnet decomposition offers another way to filter noisy experimental Mueller matrices of non-depolarizing samples. The procedure attributes to $\mathbf{M}_{nd}$ the whished matrix and to $\mathbf{M}_{id}$ the noise.

2.4.2 Product Decompositions

Product decompositions represent an arbitrary Mueller matrix as a product of elementary Mueller matrices—diattenuators, retarders and depolarizers. These decompositions are characterized by the number of elementary components and their respective positions in the multiplication. The order of the components is important since depolarizer matrices do not commute with diattenuator nor with retarder matrices. In principle, product decompositions are adequate to describe physical situations in which the beam interacts sequentially with different parts of the sample, each of which being characterized by a well-defined fundamental polarization property.

An *ad-hoc* example of such situation is the propagation of a beam through a wedge made of an anisotropic material with a rough output face. We assume that the input face is tilted with respect to the propagation direction, whereas the rough surface is perpendicular to the output beam. Then the tilted plane surface introduces diattenuation due the difference in transmission coefficients for *s* and *p* polarizations, the propagation though the bulk crystal introduces retardation, and the scattering on the rough output surface may depolarize. Accordingly, the Mueller matrix $\mathbf{M}$ of the wedge can be represented by a matrix multiplication of the form $\mathbf{M} = \mathbf{M}_{\delta P} \mathbf{M}_R \mathbf{M}_D$ with elementary Mueller matrices in this order. $\mathbf{M}_{\delta P}$, $\mathbf{M}_R$, $\mathbf{M}_D$ are the Mueller matrices of the depolarizer, the retarder and the diattenuator respectively.

Of course, situations in which one knows very well a priori the right order of the elementary components are very rare in "real life" applications. Thus various product decompositions have been (and are still being) developed to better cope with complex situations.

2.4.2.1 Forward and Reverse Decompositions into Three Factors

All these decompositions describe the input matrix $\mathbf{M}$ as a product of a diattenuator, a retarder, and a depolarizer. Actually with three elementary component types, there are six different possible orders. Among these, the most widely used choice is that chosen by Lu and Chipman [12], namely:

$$\mathbf{M} = \mathbf{M}_{\delta\mathbf{P}} \, \mathbf{M}_{\mathrm{R}} \, \mathbf{M}_{\mathrm{D}} \qquad (2.45)$$

where the "special" symbol $\mathbf{M}_{\delta\,\mathbf{P}}$ had been used for the depolarizer. Actually, for this decomposition to be quite general, if the diattenuator and the retarder are of the forms defined in Sects. 2.3.1 and 2.3.2, then the depolarizer cannot be a "pure" depolarizer of the form defined in Sect. 2.3.3, as the product matrix $\mathbf{M}$ would exhibit no polarizance (three parameters are missing). As a result, the "depolarizer" has nonzero polarizance and its matrix is of the form:

$$\mathbf{M}_{\delta\mathbf{P}} = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{P} & \mathbf{m}_\delta \end{pmatrix} \qquad (2.46)$$

With these assumptions the procedure is numerically stable and always provides physically realizable elementary matrices $\mathbf{M}_{\delta\,\mathbf{P}}$, $\mathbf{M}_{\mathrm{R}}$ and $\mathbf{M}_{\mathrm{D}}$. This procedure is thus very convenient and is widely used for the phenomenological interpretation of experimental (or even simulated) Mueller matrices.

What happens if the order of the elementary components is changed? A very simple calculation shows that the above results are easily generalized to the other two cases in which the diattenuator precedes the depolarizer, namely

$$\mathbf{M} = \mathbf{M}'_{\mathrm{R}} \, \mathbf{M}'_{\delta\mathbf{P}} \, \mathbf{M}'_{\mathrm{D}} \quad \text{or} \quad \mathbf{M} = \mathbf{M}''_{\delta\mathbf{P}} \, \mathbf{M}''_{\mathrm{D}} \, \mathbf{M}''_{\mathrm{R}} \qquad (2.47)$$

More precisely, the depolarizer matrices keep the form defined in (2.46) and the $\mathbf{M}'$ and $\mathbf{M}''$ matrices are deduced from those provided by the standard decomposition (2.45) by unitary transformations.

This kind of simple generalization is no longer valid for the three cases in which the depolarizer precedes the diattenuator. Morio and Goudail [18] introduced a "reverse" decomposition procedure for these three cases with the same definition of the depolarizer, but this procedure could lead to unstable or even unphysical results in case of very strong depolarizations. This issue has been solved by Ossikovski et al. [19] assuming that when the depolarizer precedes the diattenuator, the former features zero polarizance. The "standard" reverse decomposition takes then the form:

$$\mathbf{M} = \mathbf{M}_D \ \mathbf{M}_R \ \mathbf{M}_{D'\delta} \tag{2.48}$$

with a depolarizer matrix of the form:

$$\mathbf{M}_{D'\delta} = \begin{pmatrix} 1 & \mathbf{D}'^T \\ \mathbf{0} & \mathbf{m}_\delta \end{pmatrix} \tag{2.49}$$

As in the case of "direct" decompositions, the matrices of the three possible "reverse" cases (i.e. when the depolarizer precedes the diattenuator) are deduced from one another by simple orthogonal transformations. This procedure too is stable and always provides physically realizable Mueller matrices for the elementary components.

2.4.2.2 Symmetric Decompositions

The symmetric decomposition was first introduced for non-depolarizing Mueller matrices [20] in the following form:

$$\mathbf{M} = \mathbf{M}_{LR2} \ \mathbf{M}_{\Psi\Delta} \ \mathbf{M}_{LR1} \tag{2.50}$$

where $\mathbf{M}_{LR1}$ and $\mathbf{M}_{LR2}$ represent *linear* retarders, and $\mathbf{M}_{\Psi\Delta}$ is a *linear* retarding diattenuator with known orientation of its common diattenuation and retardation axes (like in standard ellipsometry, where these axes are along *s* and *p* directions). The procedure allowing to retrieve the three matrices $\mathbf{M}_{LR1}$, $\mathbf{M}_{\Psi\Delta}$ and $\mathbf{M}_{LR2}$ from $\mathbf{M}$ has been experimentally validated by measuring the Mueller spectra of a standard ellipsometric sample (10 nm of SiO2 on a c-Si substrate) with and without inserting retardation plates in the input and output beams [21]. A possible application of this procedure could be the elimination of the effect of birefringence of strained windows for in situ ellipsometric measurements.

This procedure was subsequently generalized to depolarizing Mueller matrices, which could be decomposed according to [22]

$$\mathbf{M} = \mathbf{M}_{D2} \ \mathbf{M}_{R2}\mathbf{M}_{d\,\delta} \ \mathbf{M}_{R1} \ \mathbf{M}_{D1} \tag{2.51}$$

where $\mathbf{M}_{D1}$ and $\mathbf{M}_{D2}$ represent generic diattenuators of the form defined in Sect. 2.2.1, $\mathbf{M}_{R1}$ and $\mathbf{M}_{R2}$ generic retarders, and $\mathbf{M}_{d\,\delta}$ a diagonal depolarizer. The central position of the depolarizer in the symmetric decomposition can be very useful for samples which can be viewed as purely depolarizing media limited by tilted input and output interfaces: in this case the diattenuation and retardation effects are likely to occur at the output interfaces and the depolarization in between. Moreover, in many cases of practical interest the Mueller matrix of the depolarizer is indeed diagonal.

However, this decomposition has two essential limits:

- $\mathbf{M}_{d\,\delta}$ an take a diagonal form if, and only if, the eigenvector related to the largest eigenvalue of the matrix product $\mathbf{M}' = \mathbf{G}\,\mathbf{M}^T\,\mathbf{G}\,\mathbf{M}$ is *not* fully polarized, which is quite generally, but not always the case. If so, $\mathbf{M}$ can be termed "*Stokes diagonalizable*" and the procedure may be used
- Assuming $\mathbf{M}$ Stokes diagonalizable the depolarizer should be *nondegenerate*, i.e. its three diagonal elements a, b, c (other than M_{11}, set equal to 1) must be different from each other. Otherwise, $\mathbf{M}_{d\,\delta}$ may commute with $\mathbf{M}_{R1}$ and $\mathbf{M}_{R2}$, and if so only the product $\mathbf{M}_{R1}\mathbf{M}_{R2}$ can be determined unambiguously.

This decomposition has been thoroughly studied experimentally. First, its validity has been demonstrated on an *ad hoc* system including nondegenerate and degenerate depolarizers set between retarders and diattenuators [23]. In both cases, the achieved accuracy was better than that of forward or reverse decompositions. Then a system with a non-Stokes diagonalizable Mueller matrix was implemented and studied [24], and finally it was shown that such matrices may occur in natural photonic systems, such as the cuticles of beetles [25].

2.4.2.3 Logarithmic Decomposition

The logarithmic decomposition was proposed recently as a complementary alternative to the standard product decompositions. This decomposition is a natural generalization [26] of the classic differential matrix formalism [27] to the depolarizing case. The approach, based on the physical picture of a continuously distributed depolarization, parallels and complements the product decomposition approach whereby depolarization is modeled as a spatially localized "lump" phenomenon. In particular, the differential matrix methodology appears as particularly well adapted to the phenomenological description of the continuous scattering in turbid media. According to this decomposition, the space derivative along the propagation direction z of a Mueller matrix $\mathbf{M}$ can be expressed as:

$$\frac{d\mathbf{M}}{dz} = \mathbf{m}\mathbf{M} \tag{2.52}$$

For a non-depolarizing medium, the (4×4) differential matrix $\mathbf{m}$ contains all the seven elementary properties of the medium and is given by:

$$\mathbf{m} = \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \mu & \nu \\ \gamma & -\mu & \alpha & \eta \\ \delta & -\nu & -\eta & \alpha \end{pmatrix} \tag{2.53}$$

in which α is the isotropic absorption, β is the linear dichroism along the x–y laboratory axes, γ is the linear dichroism along the $\pm 45°$ axes, δ is the circular dichroism,

η is the linear birefringence along the x–y axes, ν is the linear birefringence along the $\pm 45°$ axes and μ is the circular birefringence [27]. Note that $\mathbf{m}$ is z (z the direction of light propagation) dependent in the general case of longitudinally inhomogeneous medium. If the medium can be considered homogeneous in the z direction over a distance, d, expression (2.52) can be easily integrated giving:

$$\mathbf{M} = \exp(\mathbf{m}d) = \exp(\mathbf{L}) \tag{2.54}$$

which can be also written as:

$$\ln(\mathbf{M}) = \mathbf{m}d = \mathbf{L} \tag{2.55}$$

This expression indicates that the fundamental properties of the medium in consideration can be easily deduced from the simple logarithm of the related Mueller matrix if the total thickness d is known.

When the medium shows depolarization, the matrix $\mathbf{m}$ must be substituted by matrix $\mathbf{m}'$:

$$\mathbf{m}' = \begin{pmatrix} \alpha & \beta'' & \gamma'' & \delta'' \\ \beta' & \alpha_1 & \mu'' & \nu'' \\ \gamma' & -\mu' & \alpha_2 & \eta'' \\ \delta' & -\nu' & -\eta' & \alpha_3 \end{pmatrix} \tag{2.56}$$

According to this representation, the primed and the double primed betas, β' and β'', are still related to the linear dichroism in the x–y laboratory axis. Similarly, γ' and γ'' are related to the linear dichroism along the $\pm 45°$ axes, δ' and δ'' are related to the circular dichroism, η' and η'' are related to the linear birefringence along the x–y axes, ν' and ν'' are related to the linear birefringence along the $\pm 45°$ axes and μ' and μ'' are related to the circular birefringence.

As discussed in Sect. 2.2.2.2, a depolarizing Mueller matrix can be seen as a statistical superposition of non-depolarizing Mueller matrices. As a result, the matrix $\mathbf{m}'$ can also be interpreted as a statistical superposition of matrices of the form given by (2.53), which are the only differential ones for which diattenuation and birefringence may be unambiguously defined. Then, for $\mathbf{m}'$ all we can define are estimates of the polarimetric fundamental parameters, together with the uncertainties affecting these estimates.

Following the idea behind the Le Roy-Bréhonnet decomposition, it is possible to decompose a matrix $\mathbf{m}'$ into the sum of two components, a non-depolarizing component and a purely depolarizing one, represented by the matrices $\mathbf{m_m}$ and $\mathbf{m_u}$ respectively. The first matrix, $\mathbf{m_m}$, has the same shape and symmetry properties as $\mathbf{m}$ in (2.53), therefore it represents a non-depolarizing component. Accordingly, $\mathbf{m_u}$ represents the depolarizing component. Matrices $\mathbf{m_m}$ and $\mathbf{m_u}$ can be easily deduced from:

$$\text{If } \mathbf{m}' = \mathbf{m_m} + \mathbf{m_u} \quad \text{then,} \quad \mathbf{m_m} = \frac{1}{2}\left(\mathbf{m}' - \mathbf{G}\mathbf{m}'^{T}\mathbf{G}\right), \quad \text{and} \quad \mathbf{m}_u = \frac{1}{2}\left(\mathbf{m}' + \mathbf{G}\mathbf{m}'^{T}\mathbf{G}\right) \tag{2.57}$$

in which $\mathbf{G}$ is a diagonal matrix, $\mathbf{G} = \mathrm{diag}(1, -1, -1, -1)$ and the superscript T stands for the transpose. The latter expression written in an extended notation leads to:

$$\mathbf{m_m} = \frac{1}{2} \begin{pmatrix} 2\alpha & \beta' + \beta'' & \gamma' + \gamma'' & \delta' + \delta'' \\ \beta' + \beta'' & 2\alpha_1 & \mu' + \mu'' & \nu' + \nu'' \\ \gamma' + \gamma'' & \mu' + \mu'' & 2\alpha_2 & \eta' + \eta'' \\ \delta' + \delta'' & \nu' + \nu'' & \eta' + \eta'' & 2\alpha_3 \end{pmatrix} \quad \text{and}$$

$$\mathbf{m_u} = \frac{1}{2} \begin{pmatrix} 0 & \beta' - \beta'' & \gamma' - \gamma'' & \delta' - \delta'' \\ \beta' - \beta' & 0 & \mu' + \mu'' & \nu' - \nu'' \\ \gamma'' - \gamma' & \mu'' - \mu' & 0 & \eta' - \eta'' \\ \delta'' - \delta' & \nu'' - \nu' & \eta'' - \eta' & 0 \end{pmatrix} \quad (2.58)$$

What is the physical meaning of the elements of the matrices $\mathbf{m_m}$ and $\mathbf{m_u}$? The random character of depolarization leads to a statistical interpretation of these two matrices. According to (2.17) all the off-diagonal terms of $\mathbf{m_m}$ represent the average of the prime and double prime estimations of each one of the fundamental properties. Therefore, the off-diagonal terms of $\mathbf{m_m}$ are interpreted to be the most probable statistical estimates of each fundamental property. The main diagonal of $\mathbf{m_m}$ has four entries, the isotropic absorption α, and the anisotropic absorptions α_1, α_2 and α_3 along the x–y, the $\pm 45°$ and the circular axes respectively, which characterize the type of depolarization affecting the sample. The off-diagonal elements of the matrix $\mathbf{m_u}$ are the semi-differences between the prime and the double prime estimates of each one of the fundamental properties. These elements are interpreted as the statistical uncertainties associated by the degree of randomness on each fundamental property because of depolarization.

In summary, the logarithmic decomposition provides an easy way to determine the fundamental polarization properties characterizing a medium, considered as homogenous along the direction of propagation of light over a distance d. If the medium is depolarizing, then the decomposition provides a matrix $\mathbf{m'}$ whose elements are related to the fundamental properties. A simple sum decomposition of matrix $\mathbf{m'}$ allows to obtain an estimation of the most probable values of the fundamental properties together with their related uncertainty created by the depolarization. Moreover, the matrix $\mathbf{m'}$ provide three values allowing to characterize the depolarization affecting the sample unambiguously.

2.4.2.4 Experimental Validation of the Reverse Decomposition

The reverse decomposition procedure has been validated [28] by an *ad hoc* setup schematized in Fig. 2.6, used in conjunction with an imaging Mueller polarimeter described in more detail later. This setup comprised three samples: a depolarizer, a diattenuator and a retarder, traversed by the light beam in this order, and realized as follows:
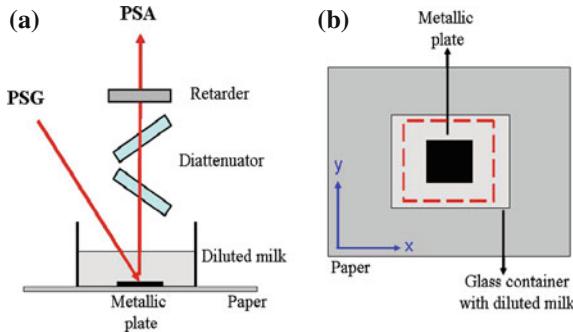
Fig. 2.6 Experimental set-up used in conjunction with an imaging Mueller polarimeter for the experimental validation of the "reverse" decomposition. **a** Side view of the whole setup. **b** View of the depolarizer as seen by the imager, from the *top*. The *red dotted line* on panel **b** defines the field imaged on the CCD

- *The depolarizer*. As a controllable depolarizer we used a transparent glass container, resting on a white piece of paper, with a small square metallic plate at the center of the field of view. This container was filled with milk diluted in water at variable concentrations. The paper always appeared as highly depolarizing due to its bulk scattering properties while the depolarization power of the plate could be varied from 0 to 1 by increasing the milk concentration.
- *The diattenuator*. This element consisted of two high index (n = 1.8) glass plates, tilted as shown on the figure at about 45°, providing a uniform diattenuation D close to 0.3 all over the field of view. This diattenuation was kept constant throughout this investigation. Moreover, both glass plates were tilted around the y axis of the image, implying that in the Mueller matrix of the diattenuator alone the only non vanishing elements were M_{11}, $M_{12} = M_{21} < M_{11}$ and $M_{22} = M_{11}$, as it can be easily checked from (2.26) with $\theta = 0$.
- *The retarder*. The retarder was a commercially available 50 mm clear aperture mica quarter wave-plate (Melles Griot 02WRM009). This element also can be inserted and removed without any displacement of the image of the depolarizer. The orientation of the plate was also kept constant but arbitrary, so that the lower right 3 × 3 sub-matrix in (2.33) was "full".

Both the diattenuator and the retarder could be inserted or removed from the beam path without any displacement or deformation of the image of the depolarizer. The measurements were performed as follows: the container was filled with 30 ml of water first, and milk concentration was gradually increased, from 0 to 5 %. For each concentration, three Mueller images were taken, the first one with the depolarizer alone, the second one with the depolarizer followed by the diattenuator and the third one with the retarder added to the other two components. Figure 2.7 shows typical experimental Mueller images, taken at 0.6 % milk concentration. Moreover:

- The image taken with the depolarizer alone exhibits the expected shape for a suspension of spherical scatterers, i.e. a diagonal matrix with $M_{22} = M_{33}$. As
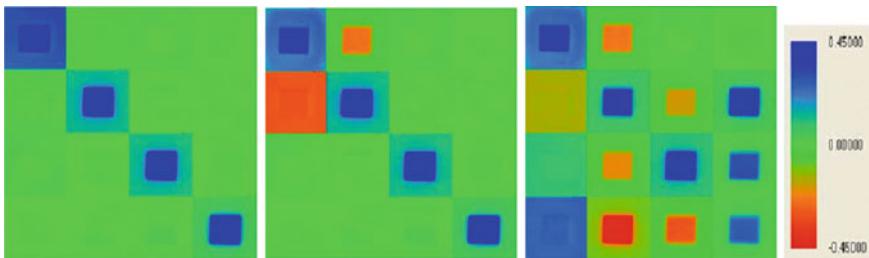
Fig. 2.7 Mueller matrix images taken with 0.6 % of milk concentration on water. *Left* depolarizer alone. *Middle left* depolarizer followed by the diattenuator. *Middle right* depolarizer, diattenuator and retarder. All elements are normalized by M$_{11}$, the only one shown without normalization and displayed according to the color scale shown at *right* of the figure. Images taken from Ref. [28] with authorization of the publisher

expected, the central part of the image, corresponding to the metal plate, features higher values of the diagonal terms than the more depolarizing surrounding part, corresponding to the paper.

- The second image was taken with the diattenuator after the depolarizer. Again, the measured Mueller images exhibit the expected trends: with respect to the previous image, only the M$_{12}$ and M$_{21}$ are affected. Moreover, the observed diattenuation (M_{12}) is significant only on the metal plate, and practically vanishes in the peripheral part due the strong depolarization characteristic of the paper. In contrast, the polarizance (M$_{21}$) is uniform, as expected for a diattenuator set after the depolarizer and covering the whole field of view. This result clearly shows how depolarization breaks the expected symmetries for a non-depolarizing Mueller matrix of a diattenuator.
- In the third image, taken with all three elements in place, we see a "mixing" of lower three lines due to the shape of the retarder Mueller matrix M$_R$ given by (2.30). All three the components of the polarizance vector are nonzero, but of course they remain spatially uniform as expected for a polarizance is introduced by the spatially uniform diattenuator.

To assess the performance of the appropriate decomposition (in the present case the "reverse" one) we applied the procedure to various images obtained with the diattenuator and/or the retarder after the depolarizer. The Fig. 2.8 shows the image of the depolarizer extracted by applying both the forward and reverse decompositions to the image acquired with all elements present (right panel of Fig. 2.7). Simple visual inspection of Fig. 2.8 clearly shows the relevance of the reverse decomposition in our case, as this decomposition retrieves a diagonal image quite similar to that taken with the depolarizer alone (left panel of Fig. 2.7). In contrast, the forward decomposition introduces an artificial polarizance, together with significant errors in the diagonal terms (M_{44} appears larger than M_{33} and M_{22}, while the opposite is true, as correctly found by the reverse decomposition).
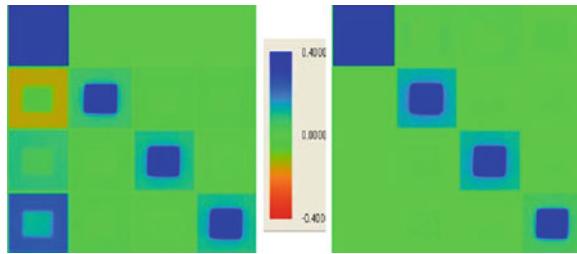
Fig. 2.8 Mueller images of the depolarizer obtained by the forward (*left*) and the reverse (*right*) decomposition of the image taken with all three elements present. Images taken from Ref. [28] with authorization of the publisher
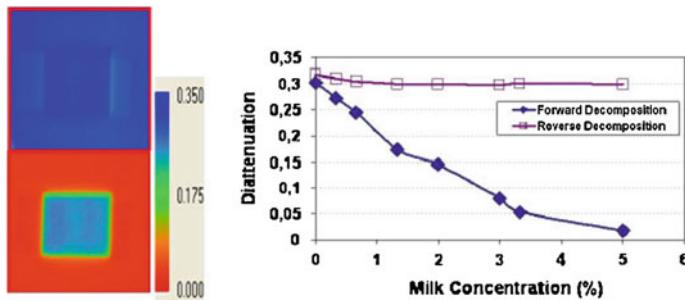

Fig. 2.9 Data relevant to the diattenuator retrieved from images taken with all components present by both forward and reverse decompositions. *Left* scalar diattenuation images provided by the reverse (*top*) and the forward (*bottom*) decompositions at 0.6 % milk concentration. *Right* evolution of the scalar diattenuation retrieved by both decompositions in the central region (where the plate is located) as a function of milk concentration. Images taken from Ref. [28] with authorization of the publisher

Another test was performed on the images of the diattenuator provided by both procedures. The results are summarized in Fig. 2.9. The left panel shows the diattenuation images provided by both procedures from measurements with all elements present. The image at the top, provided by the reverse decomposition, shows an almost spatially uniform diattenuation, as expected. In contrast, the reverse decomposition displays widely different values in the regions with the plate and the paper, with a very large error for the latter, due to its strong depolarization power. The right panel of Fig. 2.9 shows the evolution of the diattenuation retrieved in the region of the plate as a function of milk concentration, or, equivalently, of the depolarizing power δ in this region. When δ is low both decompositions provide a value close to the expected value, 0.3. Then, when δ increases, the value given by the forward decomposition decreases very significantly, while that given by the reverse decomposition remains stable.

2.4.3 Summary

In summary, in this section a number of possible ways to decompose Mueller matrices have been presented. A decomposition is always possible on a given matrix, however each decomposition has its own domain of validity and it will produce correct and results when applied to the cases for which it is adapted. Otherwise it may produce non-sense results. When it comes to decomposing Mueller matrix it is thus advisable to have a good knowledge about the physical nature of the system to evaluate the relevance of the various decompositions. It is also advisable to apply different decompositions to the same Mueller matrix and select the 'correct one' (or, rather, the 'best one') after a critical comparison of all the results. In spite of these *caveats*, Mueller matrix decompositions may prove extremely useful, as we will show for several experimental examples in, Sect. 2.7, the last section of this chapter.

2.5 Standard and Generalized Ellipsometry

Ellipsometry is a well-established and powerful optical tool for the characterization of optical substrates, thin films and multilayer samples. This technique is based on the measurement of the sample Jones matrix, followed by numerical treatments involving direct data inversion in "simple" cases, or, more frequently, data fitting by simulations with a suitable multi-parameter model.

Depending on whether the sample Mueller matrix is diagonal or not, the technique is called Standard or Generalized Ellipsometry. Non-diagonal Jones matrices are generated because the sample has the ability to transform p polarized light into s polarized light and vice-versa, which is called cross-polarization. Standard ellipsometry is, by far, the most widely used and is perfectly well suited to optically isotropic samples. On the other hand, generalized ellipsometry is needed for thorough characterization of samples such as: magnetized materials with their associated magnetization directed out of the plane of incidence, anisotropic crystals, diffraction gratings in conical configuration or roughened surfaces.

In the rest of this section, we will essentially consider standard ellipsometry, including the basic principles of operation of the most widespread ellipsometers. However, we will show that these instruments, though explicitly designed for standard ellipsometry, can also be used for generalized ellipsometry provided the sample Jones matrix obeys some specific symmetry properties. We postpone to Sect. 2.6 the presentation of Mueller ellipsometers, which are specifically designed to provide the whole set of 16 coefficients (or 15 if the matrix is measured in its normalized form).

2.5.1 Standard Ellipsometry

Standard ellipsometric measurements are commonly performed in external config-
uration, which means that a light beam propagating in air (or vacuum) is reflected
by or transmitted through, a sample, and then it propagates again in air (or vacuum)
before arriving at the detector. The interest of ellipsometry is that it can measure
simultaneously the moduli and the phases of the polarization components of the
light.

The sensitivity of phase measurements, exploited to determine thin film thickness,
has its roots in an interferometric effect. The light reflected by the first interface of a
layer present in the sample, interferes with the light reflected by the second face of
the layer. The same principle remains valid when a stack of multi-layers are present.
Therefore, the maximum film thickness that can be measured with ellipsometry has
to be less than the coherence length of the light source (otherwise, true depolarization
occurs and must be properly taken into account). Accordingly, ellipsometry is capable
of characterizing transparent or low absorbing thin films with thicknesses ranging
from less than a nanometer to several micrometers.

The information provided by ellipsometry is very rich for layer stack descrip-
tion. It enables accurate measurements of surface roughness and interfaces, while
the determination of complex refractive index gives access to fundamental physical
parameters which are related to a variety of sample properties including: morphology,
crystallinity, chemical composition and electrical conductivity, etc.

The informations extracted from ellipsometric measurements are greatly enhanced
by using wide spectral ranges, from vacuum ultraviolet to mid-infrared. The vacuum-
ultraviolet, (wavelengths between 130 and 400 nm) is the most sensitive to small
changes such as ultra-thin layers or interfaces, films with low index contrast, gra-
dient and anisotropy. Ultraviolet is also highly sensitive to the roughness of sample
surface. The near-infrared (NIR), (wavelengths between 0.8 and 3 µm), and mid-
infrared (MIR) (wavelengths between 2 and 14–50 µm), are necessary to determine
the thickness of materials which are strongly absorbing in the visible spectrum. NIR
and MIR are also used to determine the optical conductivity (typically metals and
doped oxides) because in this spectral region the optical response of samples is
dominated by free charge carriers.

In addition to high sensitivity, ellipsometry has the advantages of being non-
destructive and contactless. A spectroscopic ellipsometer is relatively easy to use
and requires no sample preparation. Standard ellipsometers can be built with light-
weight optomechanical components and they are relatively compact. They can be
mounted as stand-alone instruments or coupled to other systems such as vacuum
chambers, chemical reactors or bio-reactors, etc. In the former case, measurements
are said to be ex-situ and in the latter they are called in-situ. In-situ measurements are
interesting because they allow for the characterization of a sample in "real-time" and
at the same conditions (no alterations by the atmosphere) as it is prepared, deposited
or treated (Fig. 2.10).
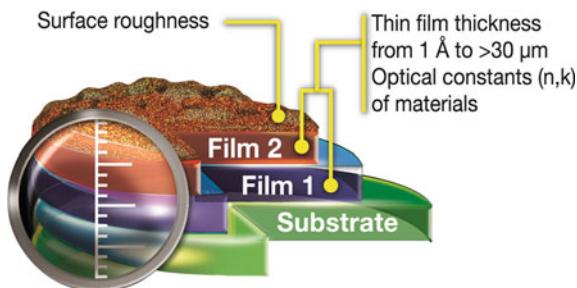
Fig. 2.10 Schematic representation summarizing the different parameters related to the sample that can be deduced using ellipsometry. These parameters include: thin film thickness, refractive index, roughness, porosity, composition and uniformity, etc.

The variety of samples that can be studied with ellipsometry opens a wide range of possibilities for this technique. A recent survey [29] of the most relevant databases of scientific publications concluded that ellipsometry has been successfully applied in many studies concerning material science (semiconductors and photovoltaics), biology (biofilms, and biosensors) and pharmacy.

2.5.2 Analysis of Ellipsometric Data

Conventional techniques used for thin film characterization (e.g., ellipsometry and reflectometry) rely on the fact that the complex reflectivity of an unknown optical interface depends on both its intrinsic characteristics (material properties and thickness of individual layers) and on three properties of the light beam that is used for the measurements: wavelength, angle of incidence, and polarization state. In practice, characterization instruments record reflectivity spectra resulting from the combined influence of these parameters. The extraction of the information concerning the physical parameters of the sample from the recorded spectra is an indirect process, in other words, from a given ensemble of experimental data, we need to build a theoretical model of the sample allowing to reproduce as closely as possible the measured data.

In general, theoretical models depend on a series of parameters characteristic of the sample, which must be adjusted to make the theoretical data "fit" the measurements. A common model for a stack of layers includes the thicknesses and the refractive indices of the layers. In many cases, the refractive index of the substrate must be considered as well [30, 31]. The quality of the fit is usually evaluated with a figure of merit and it is used during the fit process to guide the numerical algorithm which searches for the best-fitted values of the model parameters. According to [30] it is necessary to define an unbiased figure of merit in order to judge for the goodness of fit. There exist different expressions for the figure of merit, but the most popular is the one based on the mean square deviation between simulated and measured data.

$$\chi^2 = \frac{1}{N-M-1} \sum \frac{\left(\Psi^{Th} - \Psi^{Exp}\right)^2}{\sigma_\Psi^2} + \frac{\left(\Delta^{Th} - \Delta^{Exp}\right)^2}{\sigma_\Delta^2} + \frac{\left(R^{Th} - R^{Exp}\right)^2}{\sigma_{R'}^2}$$

(2.59)

N refers to the total number of data points and M is the total number of fitted para-
meters. The superscripts *Th*, and, *Exp*, refer the theoretical and experimental data
respectively. The summation is done over all the spectral data points. The sigmas
in the denominators correspond to the estimated uncertainties of the corresponding
experimental values. Typical values for sigmas of our experimental setups are around
0.5 or 0.1 % depending on the setup. The advantage of the formulation given by (2.59)
is that it allows to include non-ellipsometric data, such as total reflectivity R, in the
fitting process. Combination of ellipsometric data with information coming from
other sources can be interesting and enhances the accuracy in the determination of
fitted parameters. According to [30] the figure of merit behaves like a multivariate
mathematical function which depends on a given number of fitting parameters.

Once the figure of merit has been defined, it is possible to take advantage of
modern computers to automatize the process of parameter fitting, which is based
on the search of the minimum value of the figure of merit. The automatic process
of minimization of a multivariate function is far from being obvious. The principal
difficulty that arises almost systematically, is the fact that the figure of merit may
have either multiple minima with the same value, or multiple partial minima with
different values. In order to minimize the impact of this drawback in the final results, it
is possible to use smart or advanced minimization strategies which are based either on
systematic multiple guesses for the initial parameters, or genetic algorithms or even
on simulated annealing algorithms. In spite of the advantages of those minimization
strategies, it is important to keep in mind that at the end of a minimization process,
a supervision of the results is necessary to check their pertinence and efficiency.

A second factor that can complicate data fitting, which is inherent to the fact that
ellipsometry data analysis is an indirect process, is the correlation between fitted
parameters. We use the term parameter correlation when in a fit process it is possible
to find multiple sets of parameters that produce the same value of the figure of merit.
Correlation is said to be linear when the couples of correlated parameters follow
a linear relation. Correlation between fitting parameters happens because experi-
mental data are not sensitive to individual parameters but to combinations of them.
Correlation between two parameters may also occur if one of the two parameters
has much more impact on the data, the optical response, than the other. A statisti-
cal treatment of optical and simulated data based on the variance-covariance matrix
formalism [30] can quantify parameter correlations which are specific to the sample
and the model. It is therefore difficult to establish general rules to treat the problem.
A procedure based on common sense consists, whenever correlation appears, to keep
one of the parameters fixed at a 'reasonable' value, determined a priori or from a com-
plementary technique (microscopy, XPS...), and to fit the other parameters. Modern
commercial ellipsometers come with data modelling software packages specially
designed to help the user, professionals or beginners, to overcome the most usually
encountered difficulties in data analysis. They provide various data fitting procedures

based on multi-start or multi-guess strategies to avoid problems related to multiple relative minima and parameter correlation. In many cases, they also provide graphical user friendly interfaces which are of great help to build models of complex structures.

In spite of the great advantages of ellipsometry, this technique is limited to the analysis of samples which do not depolarize light. As stated previously, depolarization results from incoherent superposition of light beams with different polarization states. In practice, depolarization is commonly encountered when measuring inhomogeneous samples (either in composition of thickness) or very rough surfaces. Depolarization can be caused by the measurement device itself as in the case of monochromators with poor spectral resolution or instruments with high angular acceptance [3]. When depolarization is present it may introduce significant errors into the results provided by ellipsometry and the usual optical models. In such cases it is safer to make measurements with a Mueller polarimeter and use advanced optical models to take into account the sources of incoherent mixing of light which cause depolarization.

2.5.3 Instrumental Implementations of Standard Ellipsometers

Very many designs have been successfully implemented for standard ellipsometers. As an exhaustive review of all these designs is clearly beyond the scope of this chapter, in this subsection we will restrict ourselves to the configurations schematized in Fig. 2.11.

Overall, the instrument is made of two optical arms and a sample-holder in between. The first arm, the entry, is coupled to a source of light, and includes a linear polarizer set at an azimuth P with respect to the p direction in the plane perpendicular to the input beam. The second arm, or exit arm, is used to determine the polarization of the outcoming beam. It comprises a Polarization State Analyser, or PSA, and a detector which may be a single channel device (photodiode, photomultiplier...) or a multichannel one (typically a CCD coupled with a spectrometer, or, less frequently, with an imaging system). The PSA typically includes a polarizer and possibly other components. The PSA design actually defines the various types of instruments outlined in this part.
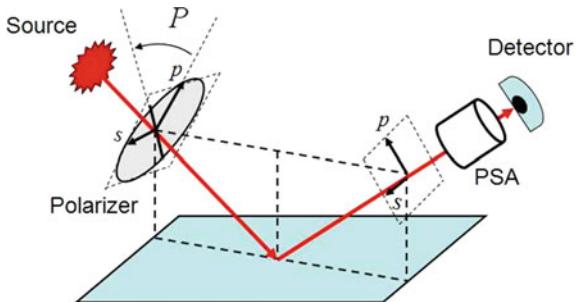
Of course, in all cases the polarization components can be inverted: all the PSAs described in the following can be placed in the input

If the instrument is designed as shown in Fig. 2.11, with only a linear polarizer (set at azimuth P) between the source and the sample, then the incident polarization is linear, with a normalized Stokes vector of the form:

$$\mathbf{S}_{inc}^T = \left(1, I_p - I_s, I_{45} - I_{-45}, 0\right) = (1, \cos(2P), \sin(2P), 0) \qquad (2.60)$$

As a result, the output Stokes vector $\mathbf{S}_{out}$ does not depend on the last column of the sample Mueller matrix $\mathbf{M}$: *only the first three columns of $\mathbf{M}$ can be determined by such an instrument*. Similarly, if the design is reversed and the *output arm includes*

Fig. 2.11 General scheme of a standard ellipsometer. The PSA is the Polarization State Analyzer, which distinguishes the various setups outlined in this subsection

only a linear polarizer before the detector, then only the first three lines of **M** *will be accessible*, as this type of analysis does not provide any information about the circular component *V* of the Stokes vector $\mathbf{S}_{out}$ of the emerging beam.

Basically, standard ellipsometers can be classified into two general families, null-ellipsometers and non-null ellipsometers. In null ellipsometers, the optical components of the system must be rotated until the detected intensity vanishes, then the ellipsometric values are deduced from the orientations of the optical elements needed to achieve the null intensity. Conversely, in non-null ellipsometers the light intensity is modulated temporally by the action of at least one of the optical components integrating the ellipsometer, then after an harmonic analysis of the signal, the ellipsometric values are deduced. The non-null ellipsometers can be classified into three groups: rotating polarizers or analyzers, rotating compensators and phase-modulated. In the following we overview some characteristics of the different types of ellipsometers. Our interest is to show which functions of the ellipsometric angles Ψ and Δ can be measured with each type of instrument, and also, how those measurements are related to the Mueller matrix elements. A deeper and more exhaustive analysis of different ellipsometric configurations can be found elsewhere [1, 3, 32].

2.5.3.1 Null Ellipsometers

Null ellipsometers were the first type of instruments developed in late nineteenth century because of their instrumental simplicity and ease of use. As outlined before, the operation of this kind of ellipsometers consists of rotating the optical elements in order to cancel the transmitted intensity. The operation is so simple that in the former systems, rotation was done manually and the null intensity was evaluated with the naked eye. During the twentieth century, thanks to the generalization of electronics, automatic rotation by motors and photodiodes substituted the human hand and eye respectively, making the measurement task much more comfortable. Two possible PSA designs for null ellipsometers are shown in Fig. 2.12.

In the first design (left panel of the figure) a variable retarder (VR) is included just after the sample with its axes parallel to the *s* and *p* directions, followed by a linear analyzer. (The VR represented as a Babinet Soleil Bravais compensator, but of course any other equivalent device, such as a nematic liquid crystal cell, may also
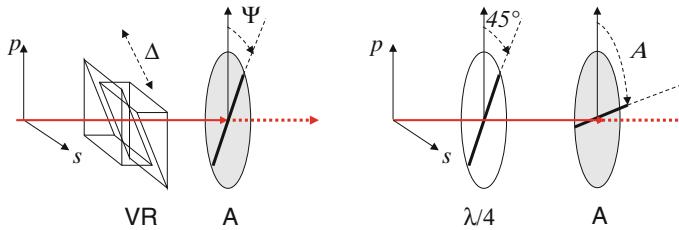
Fig. 2.12 Two possible designs of PSAs for null-ellipsometers. *Left* variable retarder (VR) with its axes aligned along the s and p directions, followed by a linear analyzer (A). *Right* the variable retarder is replaced by a quarter wave plate (λ/4). In both cases, the high transmission axis of the polarizer is shown as a black bar

be used). This PSA requires to set the polarizer in the input arm at $P = 45°$. Then, for the incoming beam $E_p = E_s$, while after the sample one has

$$E_p = E_s \tan \Psi \quad \exp \; [\mathrm{i}\Delta] \tag{2.61}$$

Now, if (and only if) the VR is set at a retardation $-\Delta$, the elliptical polarization described by (2.61) is transformed into a linear one, oriented at an angle Ψ from the s direction. Then extinction is merely obtained by setting the analyzer A at an azimuth Ψ from the p direction. The ellipsometric angles Δ and Ψ are thus nothing else but the retardation introduced by the VR and the azimuth of the analyzer that extinguish the output beam. Once the VR has been properly calibrated, such a setup can be used at various wavelengths.

In the second setup (right panel of Fig. 2.12) the VR is replaced by a quarter wave plate (QWP), which is cheaper and may also be more accurate due to its wider angular and spectral acceptances angle (at least for zero or low order plates), but will operate at a single wavelength. The fast axis of the QWP is oriented 45° with respect to respect to p direction the plane of incidence. The intensity measured by the detector is then:

$$I = \sin (2A) \sin (2\Psi) \left[\sin (2P) \cos (\Delta) - \cos (2P) \sin (\Delta) \right] - \cos (2A) \cos (2\Psi) + 1 \tag{2.62}$$

and vanishes if and only if

$$A = \Psi \quad \text{and} \quad 2P + 90° = \Delta \tag{2.63}$$

meaning that now the ellipsometric angles Ψ and Δ are retrieved from the orientations of the input polarizer P and the output analyzer A.

Null ellipsometers based on QWPs have been shown to be very accurate, and comparable to good modern instruments, but a single wavelength. As mentioned above, to circumvent this limitation VR can be used (in both the configurations shown in Fig. 2.12) to make the instrument spectroscopic, but the overall accuracy may be limited by that of the VR calibration.

2.5.3.2 Rotating Analyzer Ellipsometers

In this case, the PSA consists only of a single linear analyzer, rotating at constant angular speed ω. This configuration is often referred to as "PSRA" for "Polarizer-Sample Rotating Analyzer". The mechanical rotation of the analyzer is used to harmonically modulate the intensity of the light beam, for subsequent synchronous detection. The detected signal by a PSRA ellipsometer can be written as follows:

$$S(t) = S_0 \left[1 + \alpha \cos(2\omega t) + \beta \sin(2\omega t) \right] \tag{2.64}$$

where ω is the angular rotation speed of the analyzer. The Fourier coefficients of the modulated signal can be written as functions of the ellipsometric angles Ψ, Δ and the orientation of the polarizer with respect to the plane of incidence, P:

$$\alpha = \frac{\tan^2 \Psi - \tan^2 P}{\tan^2 \Psi + \tan^2 P}, \quad \beta = \frac{2 \tan \Psi \cos \Delta \, \tan P}{\tan^2 \Psi + \tan^2 P} \tag{2.65}$$

from which one easily gets

$$\tan \Psi = \sqrt{\frac{1 + \alpha}{1 - \alpha}} \, |\tan P| \quad \cos \Delta = \frac{\beta}{\sqrt{1 - \alpha^2}} \times \frac{\tan P}{|\tan P|} \tag{2.66}$$

As a result $\tan \Psi$, and thus Ψ itself, is determined unambiguously. In contrast, *as only* $\cos\Delta$ *is actually retrieved*, for this type of instrument:

- Only the absolute value of Δ is measured,
- This value becomes inaccurate when Δ is close to 0 or 180°, where the cosine function reaches its extrema. This situation typically occurs for thick transparent or highly absorbing samples.

However, this shortcoming may be obviated by inserting an additional known retarder, with its axes aligned with the *s* and *p* directions, to "shift" the retardation to be measured away from 0° or 180°. Another possible issue to be solved are the systematic errors which may be introduced by any residual polarization of the source and/or of the detector. On the other hand, as the technique uses only polarizers, it is possible to operate it over wide spectral ranges (from 200 nm to 30 μm), and the rotation speed may be chosen according to other requirements, such as a possible acquisition by a linear CCD after a spectrometer, which can be very convenient in many cases.

About Muller matrix elements, as stated above the fourth column cannot be measured with a linear polarization incident on the sample. Moreover, when the Polarization State Analyzer at the output consists also in a simple (rotating) linear analyzer, the fourth row is also inaccessible. Only the upper left 3×3 sub-matrix of the sample Mueller matrix can be determined, provided the measurements and data analysis outlined above are repeated with at least four different azimuths P of the input polarizer.

Alternatively, both the analyzer and the polarizer can be rotated simultaneously, at different angular frequencies, in a "RPSRA" configuration. For example, if these frequencies are set equal to 3ω for the analyzer and ω for the polarizer, then the intensity measured by a detector can be written as follows:

$$4\frac{I}{I_0} = \alpha_0 + \sum_{j=1}^{4} \left(\alpha_{2j} \cos\left(2j\omega t\right) + \beta_{2j} \sin\left(2j\omega t\right)\right) \tag{2.67}$$

where I_0 represents the light source intensity. The 3×3 upper left sub-matrix of the sample can be expressed in terms of the Fourier coefficients appearing in (2.67) as

$$\mathbf{M} = \begin{pmatrix} \alpha_0 & \alpha_2 & \beta_2 & \bullet \\ \alpha_6 & (\alpha_4 + \alpha_8) & (\beta_8 - \beta_4) & \bullet \\ \beta_6 & (\beta_8 + \beta_4) & (\alpha_4 - \alpha_8) & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{pmatrix} \tag{2.68}$$
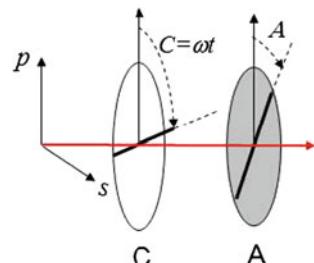
2.5.3.3 Rotating Compensator Ellipsometers

Rotating compensator ellipsometers include at least one linear retarder, usually called (somewhat improperly) compensator. The linear polarizers at the entry and exit arms are kept still and, as the name indicates, the retarder is continuously rotated in order to temporally modulate the detected signal.

Depending on whether the rotating compensator is placed at the entry or at the exit arm there are two possible configurations known as PRCSA or PSRCA where the meaning of P, S and A is the same as previously RC stands for Rotating Compensator. In the following we will consider the PSRCA configuration, which corresponds to the general scheme of Fig. 2.11, with a Polarization State Analyzer designed as shown in Fig. 2.13.

A major difference between this type of PSA and the previous one, based on a rotating analyzer, is that with a rotating compensator and a fixed linear analyzer it is

Fig. 2.13 Scheme of the Polarization State Analyzer of a rotating compensator ellipsometer in the PSRCA configuration. The linear analyzer (A) is fixed while the compensator (C) is rotated continuously, with its azimuth $C = \omega t$

possible to retrieve all four components of the Stokes vector $\mathbf{S}_{\text{out}}$, implying that more quantities are measurable, both in standard ellipsometry and for Mueller matrices.

If the compensator is a quarter wave plate (retardation equal to $90°$), the intensity recorded by the detector is:

$$\frac{I}{I_0} = 2 - \cos 2\Psi + 2 \sin 2\Psi \sin \Delta \sin 2C - \cos 2\Psi \cos 4C + 2 \sin 2\Psi \cos \Delta \sin 4C$$
(2.69)

where I_0 is again the non-modulated (DC) intensity if the light source, and the trigonometric functions of $C = \omega t$, the compensator orientation. As a result, the three different Fourier harmonics of the modulated signal directly provide the three quantities $\cos 2\Psi$, $\sin 2\Psi \sin \Delta$ and $\sin 2\Psi \cos \Delta$. In other words, rotating compensator ellipsometers provide accurate measurements of the ellipsometric Ψ and Δ angles over the complete measurement range ($\Psi = 0 - 90°$; $\Delta = 0 - 360°$). Similar results can be obtained for PRCSA ellipsometers.

However, the construction of a rotating compensator ellipsometer, with a compensator which behaves ideally providing an achromatic retardance of $90°$ over a wide spectral range, is a difficult optomechanical challenge, and it requires more complicated calibration and data reduction procedures than rotating polarizer or analyzer ellipsometers. Any deviation of the optical response of the compensator from the ideal behavior must be carefully calibrated, otherwise it will be the source of important systematic errors.

Rotating compensator ellipsometers can be implemented in more general configurations, among which:

(a) *The RP/RCFA configuration*, which consists of a rotating polarizer at the entry arm and a rotating compensator followed by a fixed analyzer at the exit arm.
(b) *The FPRC/RA configuration*, which consists of a fixed polarizer and rotating compensator at the entry arm and at rotating analyzer at the exit arm.

In the best operation mode of the RP/RCFA configuration, the compensator and the polarizer are rotated synchronously at different frequencies. In an optimal operation configuration the rotation frequency of the polarizer is 3 times that of the compensator. Then the detected signal can be decomposed in a Fourier series:

$$4\frac{I}{I_0} = \alpha_0 + \sum_{j=1}^{7} \left(\alpha_{2j} \cos (2jP) + \beta_{2j} \sin (2jP) \right)$$
(2.70)

where I_0 is again the light source intensity. The Fourier analysis of the modulated signal provides 15 coefficients which allow to determine the elements of the first three columns of the Mueller matrix as follows:

$$\mathbf{M} = \begin{pmatrix} (\alpha_0 - \alpha_6) & (\alpha_1 - \alpha_5 - \alpha_7) & (\beta_1 - \beta_5 + \beta_7) & \bullet \\ 2\alpha_6 & 2(\alpha_5 + \alpha_7) & 2(\beta_7 - \beta_5) & \bullet \\ 2\beta_6 & 2(\beta_7 + \beta_5) & 2(\alpha_5 - \alpha_7) & \bullet \\ -2\beta_3 & -2\beta_2 & -2\alpha_2 & \bullet \end{pmatrix}$$
(2.71)

The Fourier components are functions of the compensator properties, in particular retardation, which can be wavelength dependent. The calibration of such as system is extremely complex, especially when the ellipsometer is spectroscopic [33]. Conversely, the advantage of such as system is that a single measurement scheme allows to obtain 12 out of 16 Mueller matrix elements. If a simplified operation mode is used, in which only the compensator is rotated continuously, the 12 elements of the Mueller cannot be obtained after a single measurement. The polarizer must be placed at different azimuths, and for each position, a new measurement must be made. Once the process is finished, the combination of the Fourier coefficients extracted from all the measurements allows to obtain the first three columns of the Mueller matrix. Similar arguments can be given to illustrate the operation of the FPRC/RA configuration which then provide the first three *rows* of the Mueller matrix.

2.5.3.4 Phase-Modulation Ellipsometers

Finally we describe the phase-modulation ellipsometers, which include a photoelastic modulator, This component can be placed between the linear polarizer and the sample, either at the entry or exit optical arm, in PMSA or the PSMA configurations respectively. Here P, M S and A stand for fixed polarizer, modulator, sample, and fixed analyzer respectively.

A photoelastic modulator consists of a bar of optically isotropic material, which is made birefringent by means of an applied mechanical stress. To enhance this naturally weak effect, the applied stress is periodic and at a proper frequency to generate a stationary sound wave in the bar. The needed time dependent mechanical stress is usually applied using piezoelectric transducers attached to the end of the bar. As acoustic losses in the bar are very weak, very sharp resonances occur for ultrasonic waves, resulting in a dramatic increase of the induced birefringence with respect to a static stress. As a result, the device behaves as a retarder with a time dependent retardation $\delta(t)$, which varies sinusoidally at frequencies of several tens of kHz.

We now consider the PSMA configuration in which at the entry arm the polarizer is fixed and set at an azimuth P with respect to the plane of incidence, while in the exit arm, as represented in Fig. 2.14, the photoelastic modulator is set to an azimuth M and the linear analyzer is set at an azimuth angle A with respect to the plane of incidence. The detected signal then takes the form:
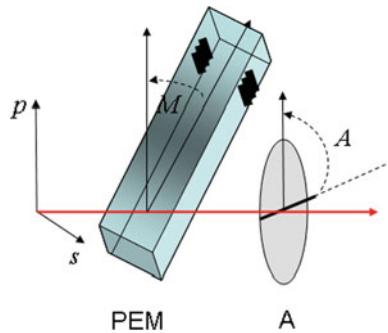
$$S(t) = S_0[1 + I_s \sin(\delta(t)) + I_c \cos(\delta(t))], \tag{2.72}$$

with:

$$I_c = \sin[2(A-M)] \, [\sin 2M(\cos 2\Psi - \cos 2P) + \sin 2P \cos 2M \sin 2\Psi \cos \Delta] \tag{2.73}$$

$$I_s = \sin[2(A-M)] \sin 2P \sin 2\Psi \sin \Delta \tag{2.74}$$

Fig. 2.14 Scheme of a PSA
comprising a photoelastic
modulator (PEM) and a linear
analyzer (A), set respectively
at azimuths M and A. The
black "patches" in the higher
part of the component are
piezoelectric transducers

In practice, as $\delta(t) = \sin(\omega t)$, the Eq. (2.72) must be developed in Fourier series (with the well-known Bessel functions as coefficients) to express I_s and I_c as functions of the directly measured quantities, actually the amplitudes of the $\sin(\omega t)$ and $\sin(2\omega t)$ components of the signal.

The expressions (2.73) and (2.74) show that the signal S (I_s, I_c) are maximized when A-$M = 45°$. Moreover it is also clear that it is not possible to unambiguously determine Ψ and Δ from a single measurement configuration. In practice, two configurations are typically used

- $M = 0°$, $A = 45°$, $P = 45°$, known as configuration II, for which we get $I_s = \sin 2\Psi \sin \Delta$, $I_c = \sin 2\Psi \cos \Delta$
- $M = 45°$, $A = 90°$, $P = 45°$, known as configuration III, for which we get $I_s = \sin 2\Psi \sin \Delta$, $I_c = \cos 2\Psi$

As a result, in configuration II, one measures Δ accurately over the full range (0–360°) but Ψ is ambiguously defined, as only $\sin 2\Psi$ is retrieved. Conversely, in configuration III one fully determines 2Ψ, but only $\cos \Delta$ is obtained. However, all that is needed to shift from one configuration to the other is to rotate the whole PSA, which can be done automatically without major difficulties, and then combine the results of the two measurements for a complete, unambiguous determination of both Ψ and Δ.

For Mueller matrix measurements, the three quantities which can be directly retrieved from the time evolution of the signal can be recast in terms of the matrix elements M_{ij} and the azimuths P, A and M as [34]

$$
\begin{aligned}
I_\alpha &= S_0 = M_{11} + M_{12} \cos (2A) + M_{13} \sin (2A) \\
I_\beta &= S_0 I_S = (M_{31} + M_{32} \cos (2A) + M_{33} \sin (2A)) \cos (2M) \\
&\quad - (M_{21} + M_{22} \cos (2A) + M_{23} \sin (2A)) \sin (2M) \\
I_\gamma &= S_0 I_C = M_{41} + M_{42} \cos (2A) + M_{43} \sin (2A)
\end{aligned}
\tag{2.75}
$$

from which it is clear that the nine elements M_{1i}, M_{2i} and M_{4i} can be retrieved with three measurements carried out with $M = 45°$ and $A = 0°$, 60° and 120° for example. Then the last four element M_{3i} can be obtained by another three measurements,

with the same A values as before but $M = 0°$. Moreover, this new set of measurements over-determines the values of M_{1i} and M_{4i}. As a result, six measurements are necessary to retrieve the full set of 12 elements of the first three columns of the Mueller matrix, with partial redundancy. In order to "spread" this redundancy over the whole set of elements, eight measurements can be realized, as described in [34]. Of course, the need to realize so many measurements to obtain the first three columns of the sample Mueller matrix makes phase modulation ellipsometers less practical than those based on rotating compensators, which provide the same set of data with a single measurement.

However, if combined with a suitable synchronous detection (which may be performed digitally) the high modulation frequency (50 kHz) specific of photoelastic phase modulation allows very efficient noise rejection and provides wide dynamic ranges. This technique is thus particularly well suited for very demanding measurements, such as the characterization of extremely thin films of example.

As it can be seen, the elements of the fourth column of the Mueller matrix cannot be addressed. To determine the elements of the sample Mueller matrix, eight measurements are necessary, corresponding to appropriate and alternative settings of the azimuths of the modulator and the analyzer. If the modulator were placed at the exit arm, the 12 elements of the matrix that would be addressed would correspond to the first three rows of the Mueller matrix. The situation is comparable to the rotating compensator ellipsometer previously described. The difference is that with the rotating compensator ellipsometer, the 12 elements can be measured in a single measuring run, whereas 8 runs are needed with an ellipsometer with only one photoelastic modulator.

2.5.3.5 Summary

In this subsection we have presented the most commonly used experimental configurations for standard ellipsometry, with particular emphasis of the quantities that actually can, or cannot, be measured by each of them. In Table 2.1, we summarize the main characteristics of these configurations, including their strengths and weaknesses.

We want to make two points absolutely clear:

- We did not try to review the many refinements are more complex systems which have been tested and developed, possibly up to commercialization. A full book would be needed for this to be done seriously.
- By listing the main advantages and weaknesses of each technique we absolutely do not mean that commercially available systems using this technique necessarily presents these strengths and weaknesses. While some basic limitations, such as those concerning the measurable Mueller matrix elements, cannot be solved in a given configuration, many other practically essential issues, among which those related to the measurements accuracy, the speed, the signal to noise ratio and the

Table 2.1 Summary of essential characteristics, advantages and weaknesses of the main ellipsometric techniques

Technology	Modulation (Hz)	Measured parameters	Strengths	Weakness
Null	0	Ψ and Δ 9 Mueller matrix elements	• Simple optical assemblies • High accuracy and precision relatively easy to achieve	• Spectroscopic extension difficult • Sensitive to residual polarization from source or detector
Rotating polarizer analyser	Several to hundreds	$\tan(\Psi)$ and $\cos(\Delta)$ 9 Mueller matrix elements	• Simple optical assemblies • Quasi-achromatic instruments • Wide spectral range of operation • Easy CCD detection	• Inaccurate measurements for Δ approaching $0°$ or $180°$ • Does not measure the V component of the Stokes vector • Sensitive to residual polarization from source or detector
Rotating compensator	Several to hundreds	$S_2 = -\cos(2\Psi)$ $S_3 = \sin(2\Psi)\cos(\Delta)$ $S_4 = \sin(2\Psi)\sin(\Delta)$ 12 Muller matrix elements	• Complete measurement of the Stokes vector from a single detection configuration • Easy CCD detection	• Complicated optical assemblies • Rotating may introduces inaccuracies because of beam wandering due to compensator imperfections

(continued)

Table 2.1 (continued)

Technology	Modulation (Hz)	Measured parameters	Strengths	Weakness
			• Generalized ellipsometry possible depolarization + 12 elements of the Mueller matrix	• Complex calibration procedures
Phase-modulation	50×10^3 to 100×10^3	$S_2 = \sin(2\Psi)\cos(\Delta)$; conf. (II) $S_3 = \sin(2\Psi)\sin(\Delta)$; confs (II) and (III) $S_4 = \cos(2\Psi)$; conf. (III) 12 Mueller matrix elements	• No rotating elements and Fast measurements • Excellent signal-to-noise ratio from VUV to NIR • Accurate measurement of (Ψ, Δ) with two detection configurations (II) and (III) • Generalized ellipsometry possible Depolarization + 12 elements of the Mueller matrix	• Temperature sensitive photoelastic modulator • No "CCD" detection • Chromatic dependence of the photoelastic-modulator • Measurements of the 12 Mueller matrix elements require multiple (6–8) runs

like greatly depends on engineering developments which are clearly beyond the scope of this contribution

We thus stress that the information presented in Table 2.1 is by no means a "buyer's guide": it might be useful only to ask the manufacturers some reasonably relevant questions !

2.5.4 Generalized Ellipsometry

In practice, when it comes to measure non-diagonal matrices characterizing non-depolarizing cross-polarizing samples, the ease of use, and the adequacy of the instrument to retrieve all the necessary information from the sample is of prime importance for the researcher. Even though none of the previously shown standard ellipsometers (rotating analyzer-polarizer, rotating compensator with a single compensator and phase-modulation with a single modulator) is able to perform a complete measurement of the 16 elements of a Mueller matrix, this does not mean that they cannot be used to perform generalized ellipsometry.

In many cases of practical interest it can be shown [31] that for non-depolarizing systems the form of their associated Jones-Mueller matrix (2.14) is sufficiently redundant that either the corresponding Jones matrix can be inferred from a partially measured Mueller matrix, or the non-measured Mueller-Jones matrices can be inferred from the measured ones. Inference procedure is valid under some hypothesis, for instance, the fact that the off-diagonal element of the Jones matrix J_{sp} is equal to the conjugate of J_{ps}.

At this point it important to point-out that the information carried by a Jones matrix and the related Jones-Mueller matrix is strictly the same, and it can be retrieved independently of the formalism, either Jones or Stokes, used to represent the data and the physical problem to which they are related.

However, it is important to recall that the accuracy of the non-measured matrix elements depends on the systematic errors of the measured matrix elements and also, on the degree of applicability of the conditions under which the non-measured elements have been inferred.

2.6 Mueller Ellipsometry: Instrumental Issues

When the redundancies of the Mueller matrix are not sufficient to reconstruct the full matrix from partial measurements, or when depolarization modifies the structure of the matrix respect to the non-depolarizing case, it is mandatory to measure the 16 elements of the Mueller matrix (15 if it is normalized). In this section we address some essential issues specific of the instruments capable of achieve such measurements.

Fig. 2.15 Schematic representation of the optical setup of a general Mueller ellipsometer. The Stokes formalisms has been used to represent the polarization properties of light. In the framework of Stokes formalism the optical properties of the PSG, sample and PSA can be represented by matrices

2.6.1 General Considerations on Mueller Ellipsometers

Throughout this section, Mueller ellipsometers will be described within the framework of the Stokes-Mueller formalism, which unifies the presentation of both techniques, and simplifies the algebra with respect to original descriptions of generalized ellipsometry based on Jones matrices.

In the following, to avoid too lengthy developments, we will limit our presentation to *sequential* ellipsometers, i.e. systems in which the polarization states are generated in the entry arm and analysed in the exit arm sequentially. The ideas and techniques presented in this section is also valid, with minor adaptations, for instruments performing sinusoidal modulation of the input and/or output polarizations with subsequent Fourier analysis of the signals (Fig. 2.15).

Within this formalism, the operation of any ellipsometric or polarimetric system can be schematized as follows:

The PSG produces a set of input Stokes vectors $\mathbf{S}_i$, which are transformed by the sample into $\mathbf{M}\,\mathbf{S}_i$ ($\mathbf{M}$ being the Mueller matrix of the sample). These output Stokes vectors are then analysed by the PSA, which delivers the raw signals B_{ij} by projecting each vector $\mathbf{M}\,\mathbf{S}_i$ onto its basis states. This scheme can be summarized by the simple matrix equation.

$$\mathbf{B} = \mathbf{A}\,\mathbf{M}\,\mathbf{W} \qquad\qquad (2.76)$$

where the modulation matrix $\mathbf{W}$, which characterizes the PSG, is formed by the $\mathbf{S}_i$ vectors in columns, while the $\mathbf{S'}_j$ are the line vectors of the analysis matrix $\mathbf{A}$ characterizing the PSA. In the most general case, $\mathbf{B}$ is rectangular, with m lines and n columns, where m and n respectively represent the numbers of states generated by the PSG and analyzed by the PSA.

To get the full Mueller matrix $\mathbf{M}$, both the PSG and the PSA must be "complete", with at least 4 basis states. Then expression (2.76) is sufficient to extract $\mathbf{M}$ from $\mathbf{B}$ by merely inverting the (in principle well known !) matrices $\mathbf{A}$ and $\mathbf{W}$, if both m and n are equal to 4, or by pseudo-inverting these matrices if the system is overdetermined. In the following for simplicity reasons, we will consider only the case of "minimal" Mueller ellipsometers, for which $m = n = 4$, but we emphasize that all the ideas exposed in the following about instrument optimization and calibration can be easily transposed to overdetermined configurations.

While their principle of operation may seem straightforward, Mueller ellipsometers are not so widespread (only two have been very recently made commercially available), because of the added technical complexity due to the simultaneous presence of *complete* PSG *and* PSA. Two issues are of paramount importance (as for any other instruments, but they are particularly critical here)

- The *optimization of the instrument design*, to get the optimal performance if all components were ideal (perfectly well described by the model). The general criterion for this optimization, namely the minimization of the condition numbers of matrices **A** and **W** is now widely accepted.
- The instrument *calibration*, in other words the determination of the actual **A** and **W**matrices, which are necessarily affected by the many imperfections of the optical components, positioning systems and the like. *This is a crucial issue*, especially for the complex setups which may be necessary to follow the optimization criterion defined above. Actually, for such complex systems, the usual approach based on a detailed modeling of the whole instrument and its non-idealities may be totally inapplicable. Conversely, the Eigenvalue Calibration Method developed and experimentally validated by Compain [35] circumvents this problem by determining both **A** and **W** matrices from a set of measurements on reference sample directly, by algebraic methods, *without any modeling of the instrument*. Moreover, the *only requirement on reference samples is that they are linear dichroic retarders*; not too close to half-wave plates, as the ECM allows to determine their precise characteristics during the calibration procedure itself. As a result, no very specific samples, such as retardation plates with accurate retardation values, are needed.

Due to its flexibility and robustness, the ECM has been a cornerstone of all the instrumental developments in Mueller ellipsometry at LPICM (and a few other laboratories as well). Its usefulness could hardly be overestimated for the development of innovative Mueller ellipsometers.

2.6.2 Design Optimization of the Polarization State Generators and Analyzers

Item 1 is probably the easiest to address. If we rewrite expression (2.76) as

$$\mathbf{M} = \mathbf{A}^{-1} \mathbf{B} \, \mathbf{W}^{-1} \tag{2.77}$$

we see that the optimization of the instrument design is equivalent to a minimization of the errors in **M** for a given value of the measurement errors in the raw matrix **B**. Due to the algebraic properties of matrices, the error propagation from **B** to **M** will be minimized if the *condition numbers* of **A** and **W** are minimized [36–39]. Without trying to be too rigorous, we now illustrate the rationale behind this criterion by considering the noise propagation from raw intensities to final results in the case of a PSA.

Let us recall that the condition number $c(\mathbf{X})$ any square matrix $\mathbf{X}$ is defined as

$$c(\mathbf{X}) = \|\mathbf{X}\| \ \|\mathbf{X}^{-1}\| \tag{2.78}$$

where the norm of the matrices (and vectors) can be defined in several ways. In our case, the most relevant choice is the Euclidean norm for the vectors while for matrices we define

$$\|\mathbf{X}\| = \sup \left[s_i(\mathbf{X}) \right] \tag{2.79}$$

where s_i are the singular values of $\mathbf{X}$. With such definitions, for any vectors $\mathbf{T}$ and $\mathbf{Z}$,

$$\mathbf{Z} = \mathbf{TX} \quad \Rightarrow \quad \|\mathbf{Z}\| \leq \|\mathbf{T}\| \ \|\mathbf{X}\| \tag{2.80}$$

If now we consider the operation of a PSA with an analysis matrix $\mathbf{A}$, the Stokes vector $\mathbf{S}$ of the light impinging on the PSA and the four dimensional vector $\mathbf{I}$ whose components are the four intensities measured at the output of the PSA are related by

$$\mathbf{I} = \mathbf{A} \, \mathbf{S} \quad \Leftrightarrow \quad \mathbf{S} = \mathbf{A}^{-1} \mathbf{I} \tag{2.81}$$

Now when the configuration of the PSA is varied, for example by changing the relative orientations of its optical components, the matrix $\mathbf{A}$ varies too. However, for typical PSA as those described below, the four line vectors of $\mathbf{A}$ are always the transposes of totally polarized Stokes vectors, and it is easy to see that in these conditions that $\|\mathbf{A}\|$ hardly changes. Conversely, $\|\mathbf{A}^{-1}\|$ may become arbitrarily large when $\mathbf{A}$ becomes close to a singular matrix, for example when two line vectors become almost equal, meaning that the same polarization state is being measured twice while for another polarization state, say $\mathbf{S}_{\min}$ the measured intensity vector $\mathbf{I}_{\min}$ is very small. If so, as the norm of $\mathbf{S}_{\min}$ is always between 1 and 2, $\|\mathbf{A}^{-1}\|$ has to be large for (2.80) to be verified.

Then, if we make the realistic assumption that any intensity measurement vector $\mathbf{I}$ is affected by an additive noise $\delta \mathbf{I}$, then the resulting error $\delta \mathbf{S}$ in the extracted $\mathbf{S}$ is simply

$$\delta \mathbf{S} = \mathbf{A}^{-1} \delta \mathbf{I} \tag{2.82}$$

but, in contrast with the signal, *the amplitude $\|\delta \mathbf{I}\|$ of the additive noise is assumed to remain constant when the* PSA *configuration is changed. As a result, the maximum error $\|\delta \mathbf{S}\|$ on the extracted Stokes vector $\mathbf{S}$ is proportional to $\|\mathbf{A}^{-1}\|$, which must then be mimimized. Due to the typically very small variation of $\|\mathbf{A}\|$, this criterion is basically equivalent to the minimization of the condition number $c(\mathbf{A})$, which has* the advantage to be dimensionless.

Moreover the condition number $c(\mathbf{X})$ is a widely used criterion to see "how well" a given square matrix $\mathbf{X}$ can be numerically inverted, between the two extremes of unitary ($c = 1$) and singular ($c = 0$). *Actually this parameter describe "how well the polarization space is sampled" by the* PSA. At this point, the generalization to complete Mueller polarimeters is intuitively straightforward: in these instruments,

both the PSG *and the* PSA *must realize an optimal sampling of their respective polarization spaces*.

Mathematically, the argumentation presented above can be easily generalized by recasting the matrices **B** and **M** defined in (2.76) as 16 dimensional vectors $\mathbf{B}^{(16)}$ and $\mathbf{M}^{(16)}$, which transforms (2.76) into [40]

$$\mathbf{B}^{(16)} = \left[\mathbf{W}^T \otimes \mathbf{A}\right]\mathbf{M}^{(16)} \tag{2.83}$$

and using the relationship

$$c\left[\mathbf{W}^T \otimes \mathbf{A}\right] = c(\mathbf{W})c(\mathbf{A}) \tag{2.84}$$

For sequential PSAs the line vectors are actually physically realizable Stokes vectors, which thus exhibit a degree of polarization between 0 and 1. Any set of such 4 vectors can never be orthogonal to each other, which prevents the matrix **A** to be unitary and imposes

$$c(\mathbf{A}) \geq \sqrt{3} \tag{2.85}$$

(a similar limitation actually occurs also for modulation-based ellipsometers, even though the basis states forming the **A** and **W** matrices are not necessarily physically realizable Stokes vectors).

Finally, we point out that the minimization of the condition numbers $c(\mathbf{A})$ and $c(\mathbf{W})$ optimizes the propagation of *additive noise*. In principle, other indicators may be found to minimize the effect of other types of noise, such as the multiplicative noise due to speckle effects in imaging with spatially coherent light.

However, in practice, this criterion provides very efficient guidelines to optimize the design of complete Mueller ellipsometer, as it has been experimentally demonstrated among others, on a double rotating compensator setup operated in discrete rotation steps [41]. In addition to the "Standard" double rotating compensator operated with continuous rotations [42], many optimized designs of complete PSA and PSG have been published in the past decade, based on photoelastic modulator in double pass [43], achromatic division of amplitude prism [44], Pockels cells [45, 46] nematic or ferroelectric. The last two types of PSGs and PSAs will be described in more detail below.

Last but not least, we conclude this subsection with the two following remarks

- Minimizing the conditions numbers of **A** and **W** not only minimizes the noise on the extracted Mueller matrix **M**, but it also "equalizes" the noise among its various components [38], and is thus recommended only for the complete Mueller ellipsometers described in this section. For more specialized instruments, or when particular attention is paid to some particular elements of **M**, other criteria may be much more adapted.

- In principle, the minimization of $c(\mathbf{A})$ and $c(\mathbf{W})$ is intended to minimize the effects of statistical noise on $\mathbf{B}$, but in practice it turns out to be also a good criterion to minimize *systematic errors* even though such errors cannot be treated by a general theory comparable to those available for statistical noises.

2.6.3 Actual Implementations of Polarization State Generators and Analyzers

In this subsection we briefly outline various widely used configurations for PSGs and PSAs, without trying to be exhaustive. We first consider those based on what we call "traditional" approaches, which make use of the elements previously described for standard ellipsometers, with, however, suitable modifications to provide full Mueller matrix measurements.

We then focus on the original systems developed at LPICM, and based on

- rotatable Fresnel rhombs for broadband spectroscopic operation in the infrared,
- nematic and ferroelectric liquid crystals. These devices are actually extremely easy to use, and typically feature wide angular and spatial acceptances, which make them particularly well suited for *imaging* applications, in the visible and near infrared range. Spectroscopic Mueller ellipsometers based on these devices have also been successfully developed and commercialized.

For all the optimized PSGs described in the following, the corresponding PSAs are nothing else but the mirror images of the PSGs.

2.6.3.1 Traditional Approaches

Two of the standard ellipsometric configurations, the rotating compensator and the photoelastic modulator have been generalized, at the expense of extensive instrumental and calibration complication, in order to access the full Mueller matrix.

Concerning the rotating analyzer configuration, the generalization consists of using at least two rotating compensators, both with an ideal retardance of 90° and rotating synchronously with different angular speeds [3, 34, 41, 42]. One compensator is placed at the entry arm between the polarizer and the sample, whereas the second compensator is placed at the exit arm between the sample and the analyzer. Following the nomenclature previously described, this configuration can be addressed as PRCSRCA, or in a shortened version just as PCSCA. The advantage of this configuration is that it allows to access the full Mueller matrix in a single measurement run. This approach has been used to develop the commercially available spectroscopic Mueller ellipsometer RC2 by Woollam [47].

The second type of generalized ellipsometer, based on photoelastic modulators, can be found in two variants. The first one, similar to the rotating compensator consists of a system with two modulators. One modulator is placed at the entry arm,

between the polarizer and the sample. The second modulator occupies a symmetric position respect to the first. It is placed at the exit arm between the sample and the analyzer. This configuration can be called, PMSMA. The two modulators can be operated synchronous or asynchronously, but they must be resonant at different frequencies. The drawback of this configuration is that in order to access the whole Mueller matrix, the modulators must be placed at different orientations, and that a complete measurement must be carried out for each orientation [48]. The second variant, consist of a system with four photoelastic modulators. According to the description given in [3, 49], two modulators are placed at the entry arm between the sample and the polarizer, and two modulators are placed at the exit arm between the sample and the analyzer. Again the modulators must vibrate at different frequencies in order to get maximum sensitivity and to avoid possible ambiguities. The advantage is that the four-modulator configuration is exempt of mechanical movements and therefore it can measure the full Mueller matrix in a single run.

2.6.3.2 Rotatable Fresnel Rhombs for Broadband Spectroscopic Operation

As it is well known [4] a prism working in total internal reflection behaves as a retardation plate, with, however, a high degree of achromaticity, as the retardation depends only on the material optical index and the angle of incidence on the face working in internal reflection, without the $1/\lambda$ factor typical of retarders cut in birefringent materials. As a result, if used in conjunction with suitable linear polarizers, this type of component is very appealing to develop broadband PSGs, PSAs, and complete Mueller ellipsometers.

Because of the bulky shape of the prisms, care must be taken to select an appropriate geometry that does not deviate the beam the element is rotated. Among the possible prism geometries [50], we selected the V-shaped retarder shown in Fig. 2.16 as it is the easiest to manufacture. This component consists of two identical Fresnel rhombs disposed symmetrically and joined by optical contact.

The PSG of the Mueller ellipsometer consists then of a linear polarizer, which is kept fixed, followed by an achromatic retarder that can rotate about the light propagation axis. The PSA is identical to the PSG, but with its optical elements in reverse order. During current operation of the Mueller ellipsometer, the PSG generates at least four polarization states corresponding to four different azimuths of the retarder and, similarly, the PSA is operated with at least four analysis configurations.

The optical configuration of the broadband Mueller ellipsometer is very similar to that of a PCSCA generalized ellipsometer, with two fixed polarizers and two mobile retarders, which explains why the broadband Mueller ellipsometer is able to measure a full Mueller matrix in a single run. However, the operation mode and the retardance created by the retarders is quite different. To work in optimal conditions, the retardance of the compensators of a PCSCA generalized ellipsometer must be 90°, 270° or a multiple of them, whereas in the broadband system, it is not the case as it will be shown later. The operation is very different too. The compensators in
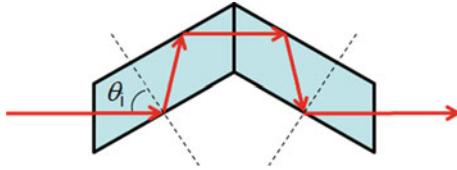
Fig. 2.16 V-shaped achromatic retarder made of two assembled Fresnel rhombs operating with four total internal reflections at incidence θ_i

the PCSCA system are continuously rotated and a double Fourier analysis, either continuous or discrete, of the time-modulated detected signal must be performed. In contrast, our broadband Mueller ellipsometer, is operated sequentially setting each retarder at four predetermined orientations and the signal is measured between two sequential rotations. During an acquisition, the retarders do not move. Simple linear algebraic operations, (two matrix multiplications), are needed to obtain the Mueller matrix from the measured intensities **B**. Finally, the system is calibrated by the ECM method described below.

As described above, the four polarization states generated in the PSG and analyzed by the PSA, have been selected in order to minimize the condition numbers $c(\mathbf{W})$ and $c(\mathbf{A})$. To this end, we simulated the optical behavior of PSG assuming the polarizer and the retarder were ideal. The Mueller matrices of a linear polarizer, **P**, and an ideal retarder, **C**, oriented respectively at azimuths 0 and θ, are:
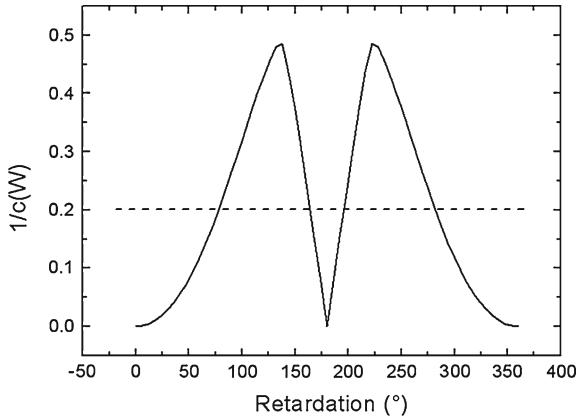
$$\mathbf{P} = \tau_P \begin{pmatrix} 1\ 1\ 0\ 0 \\ 1\ 1\ 0\ 0 \\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0 \end{pmatrix};$$

$$\mathbf{C}(\theta, \delta) =$$
$$\tau_R \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\theta) + \cos(\delta)\sin^2(2\theta) & \sin^2(\tfrac{\delta}{2})\sin(4\theta) & -\sin(\delta)\sin(2\theta) \\ 0 & \sin^2(\tfrac{\delta}{2})\sin(4\theta) & \cos^2(2\theta) + \cos(\delta)\sin^2(2\theta) & \sin(\delta)\cos(2\theta) \\ 0 & \sin(\delta)\sin(2\theta) & -\sin(\delta)\cos(2\theta) & \cos(\delta) \end{pmatrix}$$
$$(2.86)$$

where τ_P is the polarizer transmission, while τ_R and δ represent the transmission and the retardance of the retarder respectively. Assuming that the light entering the PSG is completely depolarized, the Stokes vector generated by the PSG is given by the expression:

$$\mathbf{S_{out}}(\theta, \delta) = \mathbf{C}(\theta, \delta) \cdot \mathbf{P} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (2.87)$$

Fig. 2.17 Reciprocal condition number as function of the value of the retardance of the V-shaped prism. The two maxima correspond to the selected retardances

Hence, the matrix **W** and its condition number c(**W**) can be understood as functions of five parameters. Finding the values of the parameters that minimize the condition number is a numerical problem that can be solved by means of a standard numerical optimization routines. As a result, two optimal values for the retardation and two sets of optimal azimuth angles can be found. $\delta_1 = 132° + n360°$ and $\delta_2 = 227° + n360°$, with "n" being any integer. Figure 2.17 shows the values of the condition number calculated as function of the Fresnel rhomb pair retardation for one set of optimal values of the azimuths (In this figures and all the following analogous ones, we plot the reciprocal $1/c$ of the condition number, which may vary from 0 to 0.577 and is thus easier to plot than c.)

The two maxima and correspond respectively to the optimal values for the retardation. The two sets of azimuths are:

- $\{\theta_1, \theta_2, \theta_3, \theta_4\} = \{38° + n360°, 74° + n360°, 106° + n360°, 142° + n360°\}$
- $\{\theta_1, \theta_2, \theta_3, \theta_4\} = \{218° + n360°, 254° + n360°, 286° + n360°, 322° + n360°\}$

with "n" being any integer.

2.6.3.3 Nematic Liquid Crystals

These devices behave as electrically controllable variable retarders, analogous to Babinet Soleil Bravais compensators, with fixed orientation of their slow and fast axes and retardations which may be adjusted from $1 \sim 2$ times 360° to almost 0° by applying a.c. driving voltages, typically in square wave form, with rms values from 0 to about 15 V. We used NLC variable retarders from Meadowlark; detailed information about these devices is available on their site [51]. One limitation of NLCs is their slow switching times, of the order of tens of milliseconds.

The whole PSG is composed of a linear polarizer followed by two NLCs with their fast axes set at the (fixed) azimuths θ_1 and θ_2 with respect to the polarization defined by the polarizer. Calling respectively δ_1 and δ_1 the retardations of the NLCs, a straightforward calculation provides the output Stokes vector

$$\mathbf{S}_{PSG} = \mathbf{M}_R(\theta_2, \delta_2)\mathbf{M}_R(\theta_1, \delta_1)\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{2.88}$$

where the matrices $\mathbf{M}_R(\theta_i, \delta_i)$ are those defined in (2.33) for linear retarders. Finally we get the explicitly

$$\mathbf{S}_{PSG} =$$
$$\begin{pmatrix} 1 \\ (c_1^2 + s_1^2 \cos \delta_1)(c_2^2 + s_2^2 \cos \delta_2) + c_1 c_2 s_1 s_2 (1 - \cos \delta_1)(1 - \cos \delta_2) - s_1 s_2 \sin \delta_1 \sin \delta_2 \\ c_2 s_2 (1 - \cos \delta_2)(c_1^2 + s_1^2 \cos \delta_1) + c_1 s_1 (1 - \cos \delta_1)(c_2^2 + s_2^2 \cos \delta_2) + s_1 c_2 \sin \delta_1 \sin \delta_2 \\ s_2 \sin \delta_2 (c_1^2 + s_1^2 \cos \delta_1) - c_1 s_1 c_2 \sin \delta_2 (c_1^2 + s_1^2 \cos \delta_1)(1 - \cos \delta_1) + s_1 \sin \delta_1 \cos \delta_2 \end{pmatrix} \tag{2.89}$$

where $c_i = \cos 2\theta_i$, $s_i = \sin 2\theta_i$. Now to generate the needed four Stokes vectors to be complete, we can play with 10 parameters (the fixed azimuths and the four pairs of retardations); which are far too many! Actually, among the many possibilities, the theoretical minimum of $c(\mathbf{W})$ is reached for azimuth values given by

$$\theta_1 = \varepsilon 27.4° + q90° \quad \text{and} \quad \theta_2 = \varepsilon 72.4° + r90° \tag{2.90}$$

where $\varepsilon = \pm 1$ has the same value in both equations, while q and r, are any integer numbers (not necessary equal), and retardation sequences the form

$$(\delta_1, \delta_2) = (\Delta_1, \Delta_1), (\Delta_2, \Delta_1), (\Delta_1, \Delta_2), \quad \text{and} \quad (\Delta_2, \Delta_2), \tag{2.91}$$

with
$$\Delta_1 = 315° + p90° \quad \text{and} \quad \Delta_2 = 135° + p90° \tag{2.92}$$

where, again, p is an arbitrary integer.

As retardations can be adjusted on demand, PSGs based on nematic liquid crystals can in principle reach the theoretical minimum of $c(\mathbf{W})$ for any wavelength within their spectral range. This possibility of complete optimization make them very well suited for Mueller ellipsometric measurements discrete wavelengths, provided total acquisition times of the order of 1 s for the whole set of 16 images is acceptable.

2.6.3.4 Ferroelectric Liquid Crystals

With respect to nematics, ferroelectric liquid crystals (FLCs) feature the following quite different, and complementary, characteristics

- They are also linear retarders, but with constant retardation. What is driven electrically is the orientation of their fast axis. This *orientation is actually bistable*, with two possible azimuths 45° apart from each other. The polarity of the DC driving voltage actuelly defines which of these two azimuths is actually reached.
- These devices may switch from one state to other *extremely fast*, typically in less than 100 μs.

The commutation speed of these components allow fast Mueller ellipsometry, either in spectroscopic or in imaging modes. However, due to the fixed values of retardations, PSG is built with these components will not allow a fine minimization of the condition number like that possible with nematics. This minimization can be performed only as a compromise over all the spectral range of interest. On the other hand, if acceptable values are obtained throughout this range, with c values typically less than 4 or so, then the data can be taken simultaneously over this range, allowing fast spectral ellipsometry and/or "color" Mueller imaging.

We first consider a configuration similar to that described above for nematic LCs (Fig. 2.18). A linear polarizer is followed by two FLCs, which are switched alternatively to actually generate the four needed polarization states. If we now call θ_1 and θ_2 two possible azimuths of the FLCs, when the driving voltages are switched, the resulting pairs of azimuth are

$$\{(\theta_1, \theta_2), (\theta_1 + 45°, \theta_2), (\theta_1, \theta_2 + 45°), (\theta_1 + 45, \theta_2 + 45°)\} \qquad (2.93)$$

Again, the four generated Stokes vectors can be calculated by putting these azimuths, and the constant retardations δ_1 and δ_2 into (2.89). With this configuration, the best result was obtained with

$$\delta_1 = 90° \quad \text{and} \quad \delta_2 = 180° (\text{at } 510\,\text{nm}), \quad \text{and} \quad \theta_1 = 70°, \theta_2 = 165.5° \qquad (2.94)$$

Fig. 2.18 Configuration of a PSG based on two nematic liquid crystals. The *red arrows* indicate the directions of the fast axes
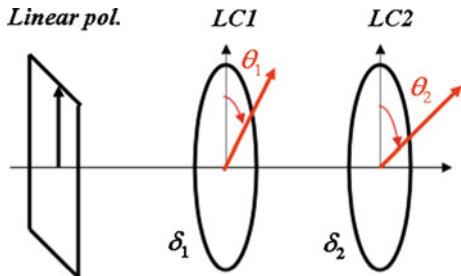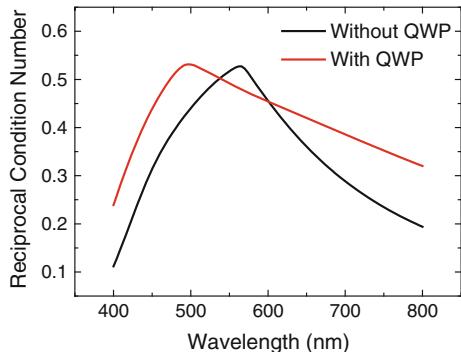
Fig. 2.19 Spectral depen-
dence of the reciprocal con-
dition number $1/c(\mathbf{W})$ of the
matrix $\mathbf{W}$ associated to the
FLC based PSG. The effect
of the insertion of a quartz
wave-plate between the FLCs
can be clearly seen. *Red line*
with the wave-plate and *Black
line* without it

The spectral dependence of the reciprocal condition number $1/c(\mathrm{W})$ obtained with
these parameters is shown as the black line on Fig. 2.19. The qualitative criterion
defined above, namely $1/c \geq 0.25$ is obeyed in a relatively narrow range, between
450 and 700 nm.

This useful spectral range can be significantly extended by adding a true zero
order quarter wave plate for 633 nm. With the same values of δ_1 δ_2 (quarter_ and
half-wave at 510 nm) the red curve is obtained for

$$\theta_1 = -10° \quad \text{and} \quad \theta_2 = 165.5° \tag{2.95}$$

meaning that the PSG can be used with the same noise propagation as before between
420 nm (limited by the transmission of the FLCs) and 1000 nm.

This configuration has been generalized to extend the spectral range towards
the near infrared (NIR). Ladstein and co-workers [52] used the same configuration
described above but with FLCs specially adapted to work in the NIR. The working
principle of this system was based in a simple scale law. They used two FLCs pro-
viding a retardance which was roughly the double of the retardance provided by the
ellipsometer working in the visible range, and as a result, they obtained a system
which worked well in the range from 700 to 1900 nm. However, the high retardance
provided by the NIR FLCs prevented the system to work in the visible range. A
second attempt made in parallel by the group of Letnes [53] in Norway and us in
France [54] showed that the addition of a third FLC to the PSG (and to the PSA as
well) allowed to work in an extended range from 450 nm in the visible to 1800 nm
in the NIR. There are two possible modes of operation of the 3FLC system:

(a) In a first mode, the orientations of the three FLCs in the PSG have to be commuted
to generate a sequence of 8 polarization states. The same sequence must to
be done for the PSA for acquiring a sequence of 64 measurements at each
wavelength to extract the Mueller matrix of the sample. This mode is referred
as "complete mode".
(b) In a second mode, the orientations of the three FLCs in the PSG have to be
commuted to generate a reduced sequence of only 6 polarization states. The
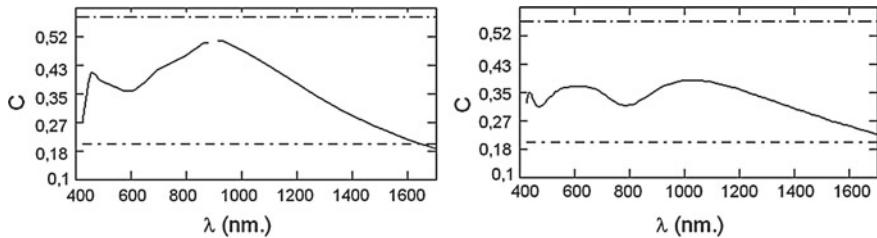
Fig. 2.20 Spectral values of (1/*c*) associated to the 3 FLCs PSG operated in the complete (*left*) mode and the reduced (*right*) mode respectively. λ represents the wavelength in nanometers

> same sequence must to be done for the PSA for acquiring a sequence of N = 36 measurements at each wavelength to extract the Mueller matrix. This mode is referred to below in the present document as "reduced mode".

Figure 2.20 shows the spectral value of the reciprocal condition number for the complete and the reduced modes of operation respectively. As it can be seen in both cases, the condition number remains acceptable. The reduced mode, despite of displaying a condition number that is slightly lower than the complete mode, allows to perform a complete Mueller matrix measurement, twice faster than the complete mode because it needs half of polarization station states. The choice between operation modes depends on the particular experimental conditions. Complete mode will always produce less noisy results than the reduced mode, therefore it is recommended to use it when measuring samples with low reflectivity.

2.6.4 The Eigenvalue Calibration Method (ECM)

We now focus on item 2, i.e. the crucial issue of instrument *calibration*. Obviously, the more complex the instrument, the more difficult the detailed modeling of all its imperfections. This modeling, however, is needed to calibrate the instrument by the usual procedure: the presumably relevant parameters are included in the model describing non-idealities, and then their values are determined by fitting the measured deviations from expected values on well-known samples (in many cases, vacuum or NIST samples made of crystalline silicon covered with various thicknesses of SiO_2) [33, 48]. While this classical procedure may be of great value as it provides in-depth knowledge of the instrument, it can be extremely time-consuming and may not be applicable to otherwise interesting optical assemblies, due to their complexity. In this respect, the optimization of **A** and **W** condition numbers, if it may enhance the instrument performance, may also render its calibration by traditional methods more difficult.

As it was outlined at the beginning of Sect. 2.6, the purpose of ECM method is to accurately determine the **W** and **A** matrices without any modeling of the instrument

nor any specific calibration samples. Provided these samples are linear dichroic retarders, their characteristics (transmission, diattenuation, retardation) are retrieved by the procedure itself, as shown in this subsection. Of course, in the following we expose the ideas presented in the original article by Compain et al. [35] but the practical implementation, and in particular the choice of the reference samples, is somewhat different.

We first address the case of a Mueller ellipsometer operated in transmission, before generalizing the method to the case of measurements in reflection.

2.6.4.1 Calibration of Mueller Ellipsometers in Transmission

The general principle. Let us first assume we have a set of reference samples, characterized by perfectly well known Mueller matrices $\mathbf{M_i}$. As the instrument is operated in transmission, it is possible to make a first measurement on vacuum, to obtain

$$\mathbf{B_0} = \mathbf{A}\,\mathbf{W}. \tag{2.96}$$

Then, putting the sample i in the beam we measure the corresponding raw matrix

$$\mathbf{B_i} = \mathbf{A}\,\mathbf{M_i}\,\mathbf{W} \tag{2.97}$$

Now, with the experimentally available matrices $\mathbf{B_0}$ and $\mathbf{B_i}$ we calculate the matrices $\mathbf{C_i}$ as

$$\mathbf{C_i} = (\mathbf{B_0})^{-1}\,\mathbf{B_i} = \mathbf{W}^{-1}\,\mathbf{M_i}\,\mathbf{W}. \tag{2.98}$$

Then, by multiplying both sides of the above equation by $\mathbf{W}^{-1}$ we finally get

$$\mathbf{M_i}\,\mathbf{W} - \mathbf{W}\,\mathbf{C_i} = \mathbf{0}. \tag{2.99}$$

For all values of I (i.e. for all samples). As $\mathbf{M_i}$ is assumed to be known and $\mathbf{C_i}$ has been obtained experimentally, (2.99) is actually a linear system of equations whose unknowns are the elements of $\mathbf{M}$. The question is therefore: how can we solve this system most efficiently? To define the best procedure, we must keep in mind two essential points

- The system may be undetermined, well determined or overdetermined, depending on how many samples have been measured, and "how different" these samples were. Thus the solving procedure must be suitable for all cases.
- Even if the $\mathbf{M_i}$ matrices are perfectly well known, the matrices $\mathbf{C_i}$ are certainly affected by experimental errors, and thus the system may very well have no *exact* solution. All we can try to do is to determine the best possible approximation of the ideal solution which would have required "perfect" measurements for $\mathbf{C_i}$.

The best way to cope with these two issues is to solve (2.99) in the "least-squares" sense, by determining the **X** matrix (or matrices, if the system is undetermined) which minimizes the sum of the squares of the 16 N elements of the N 4 × 4 matrices appearing in (2.99). Of course, is the system turns out to be undetermined, then more samples have to be included. On the other hand, inclusion of new measurements is straightforward.

To practically realize this rms minimization, we define for each measurement the linear operator **T$_i$** which associates to any matrix **X** the matrix

$$\mathbf{T_i}\,(\mathbf{X}) = \mathbf{M_i}\,\mathbf{X} - \mathbf{X}\,\mathbf{C_i} \tag{2.100}$$

Provided its elements are properly renumbered any 4 × 4 real matrix **X** can be considered as a vector $\mathbf{X}_k^{(16)}$ with 16 components. Equation (2.100) can then be rewritten in the form

$$(\mathbf{T}_i(\mathbf{X}))_k^{(16)} = \mathbf{Y}_{i,k}^{(16)} = \sum_m \mathbf{H}_{i,km}^{(16,16)}\mathbf{X}_m^{(16)} \tag{2.101}$$

where the $\mathbf{H}^{(16,16)}$ matrix is obtained by a little bit lengthy but straightforward calculation by expliciting the renumbering of the elements of **X** and **T**$_i$(**X**). Then the sum of the squares of the 16 matrix elements appearing in (2.100) for the sample i can be rewritten as

$$\left|\mathbf{Y}_i^{(16)}\right|^2 = \left[\mathbf{X}_i^{(16)}\right]^T \left[\mathbf{H}_i^{(16,16)}\right]^T \mathbf{H}_i^{(16,16)}\,\mathbf{X}_i^{(16)} = \left[\mathbf{X}^{(16)}\,\right]^T \mathbf{K}_i\,\mathbf{X}^{(16)} \tag{2.102}$$

where the 16 × 16 matrix $\mathbf{K_i} = \mathbf{H_i}^T\,\mathbf{H_i}$ is symmetrical and semi-definite positive, i.e. its eigenvalues are all positive or zero. Clearly, if Eq. (2.100) were "exact" (i.e. the matrices **C**$_i$ had no errors) *the ≪ vector ≫* **W**$^{(16)}$ *must be an eigenvector associated with a zero eigenvalue of the* **K** *matrix associated with this sample*. If we now define a matrix **K**$_{tot}$ as

$$\mathbf{K}_{tot} = \sum_i \mathbf{K}_i \tag{2.103}$$

again **W**$^{(16)}$ will be an eigenvector associated with a zero eigenvalue. In fact, the quantity $^t\mathbf{X}^{(16)}\mathbf{K}_{tot}\mathbf{X}^{(16)}$ is nothing else but the sum of the squares of the 16 N matrix elements of the **T$_i$**(**X**) appearing in (2.100) for all samples.

As all the eigenvalues of the **K**$_i$ matrices are positive or zero, when the number of samples included in the procedure (and so the number of **K**$_i$ matrices in the sum (2.103)) is sufficient, then **K**$_{tot}$ will have *only one* eigenvector **W**$^{(16)}$ associated with a zero eigenvalue, which then determines the **W** matrix unambiguously.

Again, we point out once more than due to the experimental errors affecting the matrices **C**$_i$, none of the eigenvalues of the actual **K** matrix is exactly zero. However, if the intensity measurements are reasonably accurate (and if the samples are sufficient

to (over)determine the system) *then one eigenvalue of* **K** *will be found to be much smaller than the 15 others*. In the following, we assume that the eigenvalues of **K** are sorted in decreasing order, from the largest (λ_1) to the smallest (and theoretically vanishing) λ_{16}.

The choice of sets of optimal reference samples. In addition to constitute a very efficient tool to find **W** (if the Mi are known), the above formalism also provides a simple criterion to optimize the choice of the reference samples. Our goal is to determine as precisely as possible the eigenvector of **K** associated with the smallest eigenvalue (λ_{16}). It is reasonable to assume that, the larger the other eigenvalues, the more the eigenvector $\mathbf{W}^{(16)}$ will be "isolated" from the others and thus well determined (or, if the take the argument in the other way around, if λ_{15} is very small, then the two smallest eigenvalues are almost degenerate and $\mathbf{W}^{(16)}$ will be difficult to distinguish from the eigenvector associated to λ_{15}).

In summary, *the best sets of calibration samples are those for which λ_{15} is largest, or, equivalently, those for which the ratio $R = \frac{\lambda_{15}}{\lambda_1}$ is largest*. As we will see in the following, to be characterized during the calibration procedure, these samples must also linear dichroic retarders set at various azimuths, with Mueller matrices of the form:

$$
\mathbf{M}(\tau, \Psi, \Delta, \phi) = \tau \mathbf{R}(-\phi)
\begin{pmatrix}
1 & -\cos 2\Psi & 0 & 0 \\
-\cos 2\Psi & 1 & 0 & 0 \\
0 & 0 & \sin 2\Psi \cos \Delta & \sin 2\Psi \sin \Delta \\
0 & 0 & -\sin 2\Psi \sin \Delta & \sin 2\Psi \cos \Delta
\end{pmatrix}
\mathbf{R}(\phi)
$$

(2.104)

where we recognize a matrix of the form defined in (2.15), while $\mathbf{R}(\phi)$ is matrix describing a rotation by an angle ϕ around the light propagation direction within the Stokes Mueller formalism

$$
\mathbf{R}(\phi) =
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(2\phi) & -\sin(2\phi) & 0 \\
0 & \sin(2\phi) & \cos(2\phi) & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
$$

(2.105)

It turns out that the first condition (only one vanishing eigenvalue) is met with the following set of three measurements (in addition to the measurement of $\mathbf{B}_0$, without any sample):

- *Polarizer set at two different azimuths ϕ_{P1} and ϕ_{P2}. Actually, ϕ_{P1} can be set equal to zero throughout*, without any loss of generality,
- *Retarder with retardation angle Δ, set at an azimuth ϕ_D.*

The ratio $R = \frac{\lambda_{15}}{\lambda_1}$ is then optimized by a (little bit lengthy) trial-and-error procedure. A maximum $R_{\max} = 0.1015$, which seems to be a *global* maximum, is reached for:

ϕ_{P1}	ϕ_{P2}	Δ	ϕ_D
0°	90°	109°	30.5°

Fig. 2.21 Variation of the ratio R with the retarder azimuth, for three azimuths of the 2nd polarizer
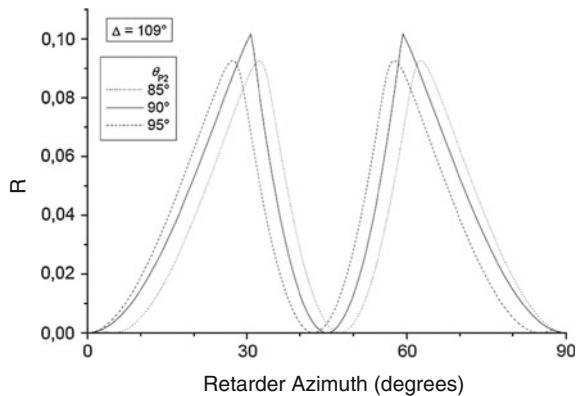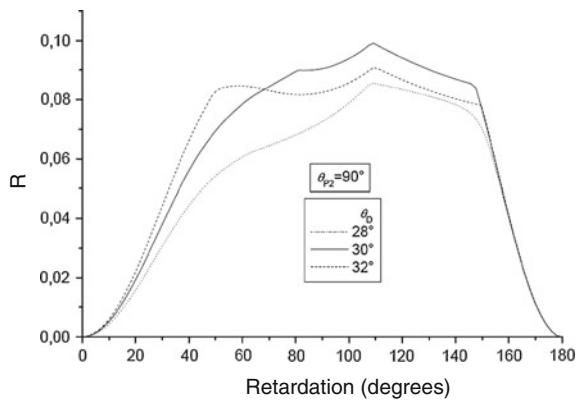

Fig. 2.22 Variation of the ratio R with the retardation Δ of the retarder, for three values of its azimuth

As for the choice of PSG and PSA configuration, this optimization must be completed by an evaluation of the «sharpness» of this maximum: a calibration procedure characterized by a local maximum $R_{loc} < R_{max}$ may be more convenient than that corresponding to R_{max}, if the local maximum is much broader, as the values of the parameters are less critical.
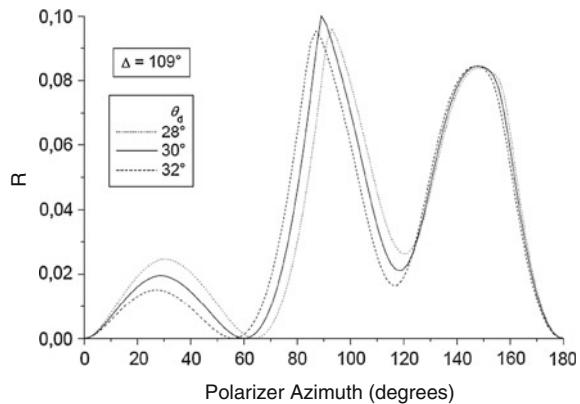
In the three following figures, the ratio, R is plotted against ϕ_D, Δ, and ϕ_{P2} respectively for values close to the optimum.

The curves reproduced in Figs. 2.21, 2.22 and 2.23 clearly show that the value of the retardation is by no means critical: this retarder must not be close to a half-wave plate, but basically any value of Δ between 30° and 150° is fine. Conversely, the angular positioning of the polarizer is more critical.

Of course, many other sets of calibration samples can be used, including at least a polarizer and a retarder. For the latter, the acceptable retardation interval is very broad, as seen above.

Characterization of reference samples. The basic idea is to use the very special form of the $\mathbf{M}(\tau_i, \Psi_i, \Delta I, \phi_i)$ matrices of the calibration samples to determine

Fig. 2.23 Variation of R with the 2nd polarizer azimuth, for three orientations of the retarder

these matrices during the calibration procedure itself. A key in this respect is that the eigenvalues of these matrices are independent of ϕ. Two of them are real and the other two are complex conjugates:

$$\lambda_{R1} = 2\tau \cos^2 \Psi, \quad \lambda_{R2} = 2\tau \sin^2 \Psi, \quad \lambda_{C\pm-} = \tau \sin(2\Psi)\exp[\pm i\Delta] \quad (2.106)$$

Now, (2.98) shows that the experimentally available matrix C_i and M_i have the same eigenvalues, from which we can determine all the parameters of M_i, except the azimuth, by inverting (2.106)

$$\tau = \frac{1}{2}\left(\lambda_{R1} + \lambda_{R2}\right), \quad \cos(2\Psi) = \frac{\lambda_{R1} - \lambda_{R2}}{\lambda_{R1} + \lambda_{R2}} \quad (2.107)$$

$$\sin(2\Psi)\cos\Delta = \frac{\lambda_{C+} + \lambda_{C-}}{\lambda_{R1} + \lambda_{R2}}, \quad \sin(2\Psi)\sin\Delta = -i\frac{\lambda_{C+} + \lambda_{C-}}{\lambda_{R1} + \lambda_{R2}} \quad (2.108)$$
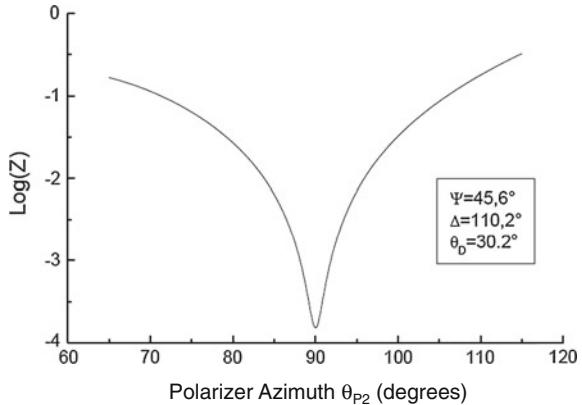
Moreover, these eigenvalues must obey the simple criterion:

$$\frac{\lambda_{R1}\lambda_{R2}}{\lambda_{C+}\lambda_{C-}} = 1 \quad (2.109)$$

which provides a simple test of the quality of the measurements involved in the determination of C_i.

At this point, the azimuths ϕ_i can be determined from the spectrum of the matrix K. If we use a polarizer set at two different azimuths (one of which is zero by definition, and the other, ϕ_{P2}, is close to 90°) and a retarder set at $\phi_D \sim 30°$ the matrix K can be calculated with ϕ_{P2} and ϕ_D "injected" into the M_i as variable parameters. Then, the system (2.99) has a solution, and l_{16} will tend to zero if, and only if, these parameters take the values corresponding to the actual azimuths of the elements during the measurements. Thus, these azimuths can be determined by minimizing the ratio $\lambda_{16}/\lambda_{15}$. A typical example of such a search is shown in Fig. 2.24. The

Fig. 2.24 Typical result of the
determination of the azimuths
of the calibration samples by
minimizing $Z = \lambda_{16}/\lambda_{15}$

minimum is quite sharp, allowing a determination of the angles ϕ_{P2} and ϕ_D with
accuracies of the order of $0.1°$.

Once **W** has been determined, the matrix A is readily obtained by inverting (2.96):

$$\mathbf{A} = \mathbf{B_0}\mathbf{W}^{-1} \tag{2.110}$$

2.6.4.2 Calibration of Mueller Ellipsometers in Reflection

We consider now cases where the vacuum cannot be measured, but an object (say a
sample holder) with a matrix $\mathbf{M_0}$ (a priori unknown) is necessarily placed between
the PSG and the PSA. Then the PSG and the PSA have to be calibrated separately.
In a first (common) step, the holder is measured alone, yielding

$$\mathbf{B_0} = \mathbf{A}\,\mathbf{M_0}\mathbf{W} \tag{2.111}$$

Then the calibration samples are inserted in the input arm, between the PSG and the
sample holder. The corresponding raw data matrices now write

$$\mathbf{B_i} = \mathbf{A}\,\mathbf{M_0}\,\mathbf{M_i}\,\mathbf{W} \tag{2.112}$$

From which we calculate the products

$$\mathbf{C_i} = \mathbf{B_0}^{-1}\mathbf{B_i} = \mathbf{W}^{-1}\mathbf{M_i}\mathbf{W} \tag{2.113}$$

And the matrix **W** is then determined in the same way as described previously. Then
the calibration elements are inserted in the output arm, between the holder and the
PSA, to measure

$$\mathbf{B}'_i = \mathbf{A} \ \mathbf{M_i} \ \mathbf{M_0} \ \mathbf{W} \tag{2.114}$$

And then calculate

$$\mathbf{C}'_i = \mathbf{B}'_i \ \mathbf{B_0}^{-1} = \mathbf{A}^{-1} \ \mathbf{M_i} \ \mathbf{A} \tag{2.115}$$

And the same procedure now provides $\mathbf{A}$. There is no need to know $\mathbf{M_0}$ this matrix can even be measured once both $\mathbf{A}$ and $\mathbf{W}$, are calibrated, the only requirement about this matrix is that it should be reasonably well conditioned to allow a safe inversion of $\mathbf{B}_0$.

2.6.4.3 Summary: Practical Implementation

We now summarize how the method is used in practice. The essential steps appear in a different order with respect to the previous (hopefully pedagogical !) presentation:

1. **Choice of the set of calibration samples**. Of course this choice is made once and for all, and may be dictated not only by the criteria shown above, but also by practical considerations about the availability of the components, and the space they need to be inserted. As we already mentioned, there are many possibilities with combinations of polarizers and retarders, the latter being really uncritical. The main issue to keep in mind is that these components should be "ideal" in the sense that they should not depolarize, for example. As a result, these components must feature wide enough spectral, angular or spatial acceptances, depending on the type of instrument.
2. **Measurement of $\mathbf{B}_0$.** This is the first step of the procedure, but if the instrument is operated in transmission (or equivalently, in reflection under normal or quasi-normal incidence) this measurement may also be used for another purpose, namely the minimization of the condition numbers of $\mathbf{A}$ and $\mathbf{W}$, if needed, for example by optimizing the driving voltages of NLCs (this step too is made once and for all for a given configuration): as $\mathbf{B}_0 = \mathbf{AW}$, the lower the condition numbers of $\mathbf{A}$ and $\mathbf{W}$, the lower that of $\mathbf{B}_0$. As a simple measurement of $\mathbf{B}_0$ is much faster than a complete calibration, the multi-parameter optimization using this criterion is also much more efficient, even though at the end a full calibration is needed to evaluate separately $c(\mathbf{A})$ and $c(\mathbf{W})$!
3. **Measurement of the $\mathbf{B_i}$, calculation of the, $\mathbf{C_i}$, matrices and of their eigenvalues**. This step offers very interesting possibilities to check that the system is working properly. Actually, the eigenvalues of the matrices $\mathbf{C_i}$ should be of the form

   a. $(\tau, 0, 0, 0)$ for polarizers,
   b. $(\lambda_{R1}, \lambda_{R2}, \lambda_C, \lambda_C^*)$ for the retarders, and these eigenvalues should verify (2.107)
   c. For any given component, the eigenvalues of its $\mathbf{C}$ matrix should be invariant when the component is rotated (even though the $\mathbf{B}$ matrix changes wildly !)

Let us emphasize that these criteria are pretty robust, as they do not involve in any way the (still unknown **A** and **W** matrices) but are based on two conditions which are mandatory for a good operation of the instrument:

- the intensity measurements are linear.
- the components, in their current use, can be considered as "good" linear dichroic retarders.

It is a very strong point of the method to offer the possibility to check these issues in a very robust way.

4. **The K matrix is evaluated** as a function of the unknown azimuths of $(n - 1)$ calibration samples (in the example showed above, these parameters were θ_{P2} and θ_D) and the ratio $Z = \lambda_{16}/\lambda_{15}$ must be minimized to determine the actual values of these parameters,
5. **W is determined** as the eigenvector associated to λ_{16}.
6. **A is determined** either from W and B0 (in transmission) or by repeating the procedure with the calibration samples in the exit arm.

Though this procedure may seem quite complex, once it is implemented it is very easy to use and provides accurate results even for systems which would be impossible to model correctly such as the ellipsometer based on FLCs described in Sect. 2.6.3.2. As a matter of fact, by using this calibration method, *full Mueller ellipsometry becomes easier to implement than more limited techniques with traditional calibration procedures* !

2.6.5 Examples of Complete Mueller Ellipsometers Developed at LPICM

2.6.5.1 Broad Band Spectroscopic Mueller Ellipsometer with Fresnel Rhombs

The motivation to develop a broadband Mueller ellipsometer was to perform complete Mueller matrix measurements in a wide spectral range. Two examples of interesting wide spectral ranges are the vacuum-ultraviolet (130 nm) to the mid infrared (7000 nm), or the visible to mid-long infrared (500–18000 nm). Vacuum-ultraviolet and infrared ranges have particular advantages. Infrared, features enhanced sensitivity to the sample chemical compositions due to molecular vibrations and phonon absorptions. Moreover, free carrier absorption in the infrared can be useful to study the dopant concentrations or electrical mobilities. On the other hand, the shortness of the vacuum-ultraviolet wavelengths make this range ideal to measure the thickness of very thin films (down to a few nm). In addition, the enhanced sensitivity of vacuum-ultraviolet to small defects and structures in the surface of samples is used for optical metrology and surface state quality control.

To date, several systems have been shown which are able to work in the vacuum ultraviolet [55–57], the infrared [33, 58–63], or even the terahertz [64–66] spectral

ranges. However, almost none of them can perform complete Mueller matrix measurements in a wide spectral range because they are either standard (incomplete) ellipsometers, or they are designed to work in a relative narrow spectral range. We developed a broadband Mueller ellipsometer to fill this lack of full polarimetric solutions.

Two main parameters actually determine the performances of a broadband system, namely the transparency and the dispersion properties of the optical components used to build it. Polarizers, based either on prisms or grids, can be used over broad spectral range, with quasi-achromatic properties. For instance, magnesium fluoride (MgF_2) used in prism polarizers is transparent from 0.12 to 7 μm and ZnSe, used for grid polarizers, transmits from 0.6 to 21 μm.

In contrast, issues appear with retarder. Liquid crystals are limited to a relative narrow spectral band covering the visible and the near infrared. These materials strongly absorb in the mid-infrared and may be irreversibly damaged by ultraviolet light. Moreover, liquid crystals show a strong dispersion of retardance at short wavelengths. Similar issues related with either absorption or dispersion appear when other optical components such Kerr optical rotators, photoelastic modulators, or wave-plate compensators are considered. We decided to use retardation induced by total internal reflection in prisms because it is almost non dispersive [67]. Limitations are only induced by the transmission of the prisms.

In accordance to these general principles we developed a prototype which is optimized to work in the mid-infrared spectral range, from 2 to 14 μm. A functional scheme of the prototype optical assembly is represented in the Fig. 2.25a. The instrument is represented in a reflection configuration even though measurements in transmission are also possible. The Mueller ellipsometer consists of an input arm, a sample holder, an exit arm and an acquisition system. The input arm includes a light source, a PSG and a retractable sample-holder for calibration samples. The exit arm includes a second retractable sample-holder for calibration samples, a PSA and a detection platform which can contain one or multiple detectors. The illumination source consists of a commercial conventional FTIR interferometer providing an collimated beam. The PSG is of the type described in Sect. 2.6.3.2, with a fixed grid type linear polarizer, and an V-shaped achromatic retarder. The retarder is mounted on a motorized rotating platform that allows it to be rotated at four consecutive orientations following a predetermined sequence. Both the polarizer and the V-shaped retarder are made on ZnSe. The PSA is identical to the PSG as discussed above.

As shown in Fig. 2.25 it is possible to design the optical assembly either to minimize either the spot size or the divergence of the beam arriving to the sample. In the focused configuration (Fig. 2.25c), the spot size can be reduced to around 300 μm. This configuration is ideal to measure small or inhomogeneous samples. If a motorized X–Y table is used to move the sample, it is possible to record mappings of the sample with a good lateral resolution. In the collimated beam configuration (Fig. 2.25b), the divergence of the beam can be reduced to values of the order of several milliradians. This configuration is ideal for measuring strongly anisotropic samples such diffraction gratings or photonic structures. In this configuration, the spot size is about 0.7 mm. In this configuration, it is advisable to use an automatic
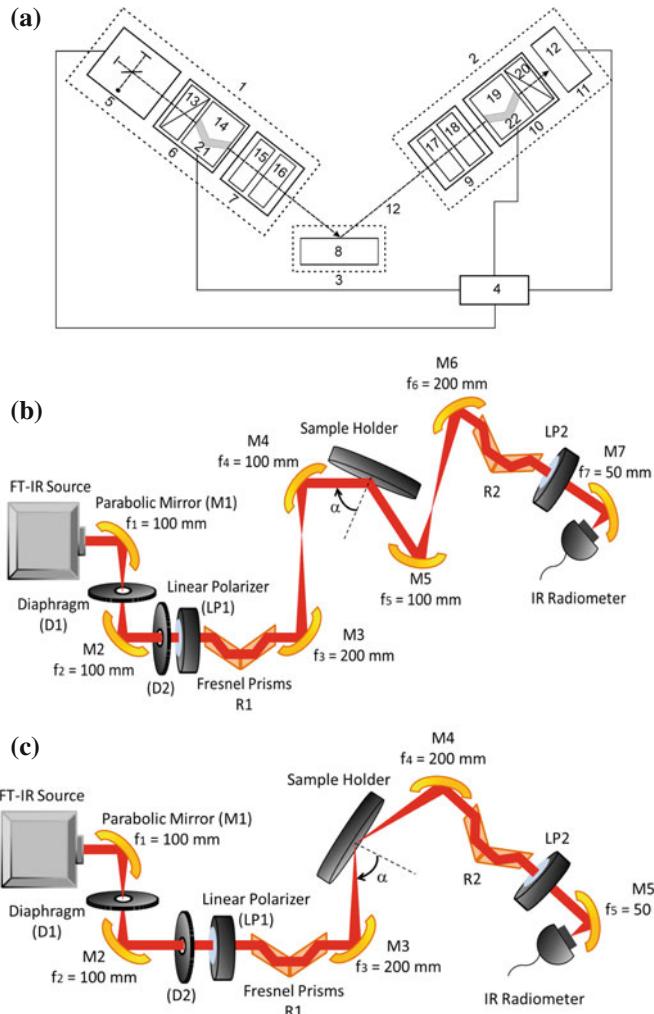
Fig. 2.25 a Functional schema of the Mueller ellipsometer. It includes a FTIR interferometer (5) a PSG (6) and a PSA (10). The PSG includes a linear polarizer (13) and the V-shaped retarder (21). The same elements can be found in the PSA. The detection signal (12) and the motors which control the orientation of the V-shaped retarders are controlled with a computer (4). **b** Schematic representation of the optical chain to create a collimated beam incident on the sample. **c** Idem to create a focused spot size on the sample

theta table as sample-holder to make measurements as a function of the azimuthal angle.

2.6.5.2 Spectroscopic Mueller Ellipsometer with Ferroelectric Liquid Crystals

Liquid crystal modulation Mueller ellipsometers use liquid crystal variable retarders to modulate the polarization without any mechanical rotation. The first prototype was built in 2003 and presented to the 3rd International Conference of Spectroscopic Ellipsometry held in Vienna [68]. After the encouraging preliminary results, the system was patented [69], and the first commercial system appeared in 2005 under the name of MM16. Since then, the product has been developed and new versions are now in the market. For instance, a particular implementation adapted to the measurement of small samples, commercialized under the name of AutoSE has been launched in 2008. The latest version, called SmartSE, combine spectroscopic with imaging capabilities. The spectral range, initially limited to the visible (450–850 nm), has been widened to the near IR (450–1000 nm). The working spectral range of liquid crystal based polarimeters is determined by the transparency of the liquid crystal devices. In the short wavelength range, UV radiation must be avoided because it may induce chemical modifications or even destroy the liquid crystals which are made of delicate organic molecules [70]. In the long wavelength range, the near infrared, the limits are due to the thin conducting oxides that are deposited on the windows of the liquid crystal devices that allow the electrical control the orientation of the liquid crystals. The conducting oxides have a high concentration of free charge carriers which absorbs very efficiently the near infrared ($>$1500 nm) making the devices to become opaque. This range matches quite well that defined in Sect. 2.6.3.4 from the criterion $c(\mathbf{W})$, $c(\mathbf{A}) > 4$, and is also very well suited for the use of standard CCDs.

The optomechanical design of the instrument is shown in Fig. 2.26. The illumination beam comes from a halogen lamp, goes through the PSG, interacts with the sample and traverses the PSA before entering a spectrometer operating in the visible and fitted with a CCD array. For practical purposes, the PSG and the PSA were mounted on an automatic goniometer allowing to vary the angles of incidence and detection from 40° to 90° in steps of 0.01, and to operate the ellipsometer in
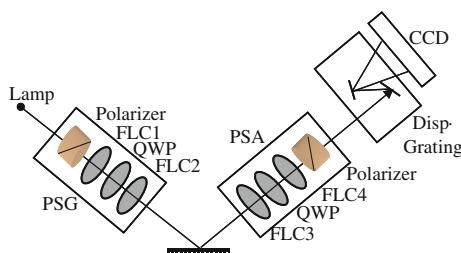

Fig. 2.26 *Left* schematic representation of the general set-up of a Mueller ellipsometer mounted in reflection configuration, showing the PSG the sample and the PSA. *Right* schema of the optical configuration of the PSG. The PSA is identical to the PSG. Pictures taken from [68] with the authorization of the authors and the publisher

reflection and in transmission modes. The sample holder was mounted on an auto-
mated theta table which allows rotation of the sample about a vertical axis. As shown
below, this azimuthal movement is interesting to characterize anisotropic samples
and diffracting structures.

With this setup, full Mueller spectra in the whole spectral range could be taken in
seconds or less, depending on the sample reflectivity. Moreover, most users are essen-
tially interested in classical ellipsometry, for which the accuracy must be comparable
to that of other, very mature, instruments. Actually, normalized Mueller matrix spec-
tra are retrieved with typical accuracies of a few 10^{-3} % a clear demonstration of the
efficiency of the optimization and calibration procedures described earlier.

2.6.5.3 Imaging Mueller Ellipsometer with Nematic Liquid Crystals for Macroscopic Samples
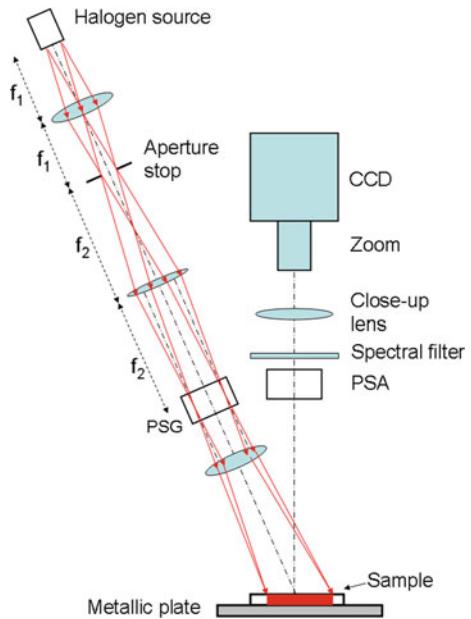
We now describe an imaging ellipsometer for macroscopic samples (up to 4 cm in
size) using PSG and PSA based on nematic liquid crystal variable retarders, previ-
ously described in Sect. 2.6.3.3. This approach seemed indeed the most appropriate
for an instrument used to image static samples, as the acquisition time (of the order of
10 s) was not an issue while the condition numbers of **W** and **A** could be minimized
for each chosen wavelength. Actually, this instrument was installed in 2009 in the
Pathology Department of Institut Mutualiste Montsouris hospital in Paris, and has
been used to image dozens of samples such as those briefly presented in Sect. 2.7.4.

A schematic view of the optical assembly is represented in Fig. 2.27. The illu-
mination part of the set-up comprises a halogen lamp (Olympus CLH-SC 150 W),
a fiber bundle with its output at the focus of an aspherical condenser (Newport
KPA046, $f1 = 37$ mm) followed by an achromatic lens (Edmund Optics NT- 32-
886, $f2 = 150$ mm) whose focal point F coincides with the condenser's one. This
combination is both telecentric in the object and images the fiber output onto the PSG.
As all points of the NLCs are illuminated with the same angular aperture (defined
by the aperture stop), the polarization generated by the liquid crystals is spatially
uniform in spite of the angle-dependency of the liquid crystal retardation. Finally, a
third lens images the PSG on the sample plane, illuminating a 5 cm diameter zone
with very uniform intensity and polarization throughout.

The sample was imaged on a fast CCD camera (Dalsa CAD1, 256×256 pixels,
12 bits) by means of a $12.5 - 75$ mm zoom with an additional 500 mm close-up lens
to "reject" the sample image at infinity, in order to use the zoom. This arrangement
allowed an efficient rejection of any stray light. The wavelengths could be chosen
between 400 and 700 nm in 50 nm steps by means of interference filters with 20 nm
bandwidth placed before the close-up lens.

The PSG and PSA were built according to the optimized design defined in
Sect. 2.6.3.3, [71] with a dichroic linear polarizer (Melles Griot, 03 FPG 007) fol-
lowed by two nematic LCVRs (Meadowlark LVR 300). The four NLCs of the PSG
and PSA were sequentially switched with frequencies in ratios 1, 2, 4 and 8 to gener-
ate the 16 needed polarization states. The dark current, which depended on the room

Fig. 2.27 Scheme of the imaging Mueller ellipsometer in reflection

temperature was measured frequently and subtracted from the signal for each pixel. A satisfactory signal-to-noise ratio was achieved by typically averaging 7 images for each state of polarization. A complete set of 16 averaged raw images was taken in about 11 s.

Typically, the reciprocal condition number of the PSG and the PSA were of the order of 0.45, thus quite close to the theoretical maximum of 0.57. The polarimetric accuracy, evaluated with a polarizer and a retardation plate set a various azimuths in either the illumination or the detection arm, was of the order of 3 % (maximum errors), which turned out to be more than sufficient for polarimetric imaging of surgical samples. However, a "run" of improvement of this instrument is planned for the near future.

Impact of the choice of the reference frame basis. With instruments operating in reflection at, or close to, normal incidence, special attention has to be paid to the definition of the $(\mathbf{s}, \mathbf{p})$ and $(\mathbf{s}', \mathbf{p}')$ bases for incident and emerging beams, whose directions of propagation are defined by the vectors $\mathbf{z}$ and $\mathbf{z}'$ respectively. Far from normal incidence, the most natural (and the most widespread !) choice is that shown, for instance, on Fig. 2.11: $\mathbf{s} = \mathbf{s}'$ and both $\mathbf{p}$ and $\mathbf{p}'$ have positive components on the outgoing surface normal. By doing so, the 3D bases $(\mathbf{p}, \mathbf{s}, \mathbf{z})$ and $(\mathbf{p}', \mathbf{s}', \mathbf{z}')$, are both right handed. If this choice is maintained up to normal incidence, the result is:

$$(\mathbf{p}', \mathbf{s}', \mathbf{z}')^+ = (-\mathbf{p}, \mathbf{s}, -\mathbf{z}) \qquad (2.116)$$

with this (usual) convention the Mueller matrix $\mathbf{M}_m^+$ of a mirror under normal incidence is

$$\mathbf{M}_m^+ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0' & -1 \end{pmatrix} \tag{2.117}$$

The other possible choice for the emerging beam is to take the same basis (s, p) for both directions of propagation:

$$(\mathbf{p'}, \mathbf{s'}, \mathbf{z'})^- = (\mathbf{p}, \mathbf{s}, -\mathbf{z}) \tag{2.118}$$

Of course, in this case the $(\mathbf{p'}, \mathbf{s'}, \mathbf{z'})^-$ is no longer right-handed. But then the Mueller matrix $\mathbf{M}_m^-$ of a mirror under normal incidence is simply the unit matrix. More generally, if $\mathbf{M}^+$ is the Mueller matrix of a system in the usual (+) convention, then the Mueller matrix $\mathbf{M}^-$ of the same system in the other convention is obtained by changing the sign of the lower two lines.

As we will see in the following, many samples of interest for us behave as diagonal depolarizers, with $M_{22} = -M_{33}$ with the usual (+) convention and $M_{22} = M_{33}$ with the other choice. Obviously, it is much easier to visually check this relationship if the elements to be compared are supposed to be equal rather than opposite ! We thus decide to keep the "unusual" choice for the calibration of the imagers operating in reflection.

2.6.5.4 Imaging Mueller Ellipsometer with Nematic LCs Coupled to a Microscope

This instrument can be seen as the ultimate development of the well-known polarized microscopy, as the polarimetric characterization of the sample is complete, in contrast with the usual setups with crossed linear polarizers or left and right circular polarizers. An overall view of the imaging polarimeter [72, 73] is shown in Fig. 2.28.

A microscope objective (Nikon Plan Achromat 100×) with a high numerical aperture (0.90) is illuminated by a halogen source via a fiber bundle followed by an input arm comprising, among other elements:

- An *aperture diaphragm* 1, which is imaged on the objective back focal plane (BFP), and is used to define the *angular distribution* of the light incident on the sample,
- A *field diaphragm* 2, imaged on the sample, which allows to define *the size of the illuminated area* on the sample.
- The PSG 3, to modulate the incident polarization,
- A *nonpolarizing beamsplitter*, with approximately 50 % transmission and reflection coefficients, to steer the beam onto the microscope.
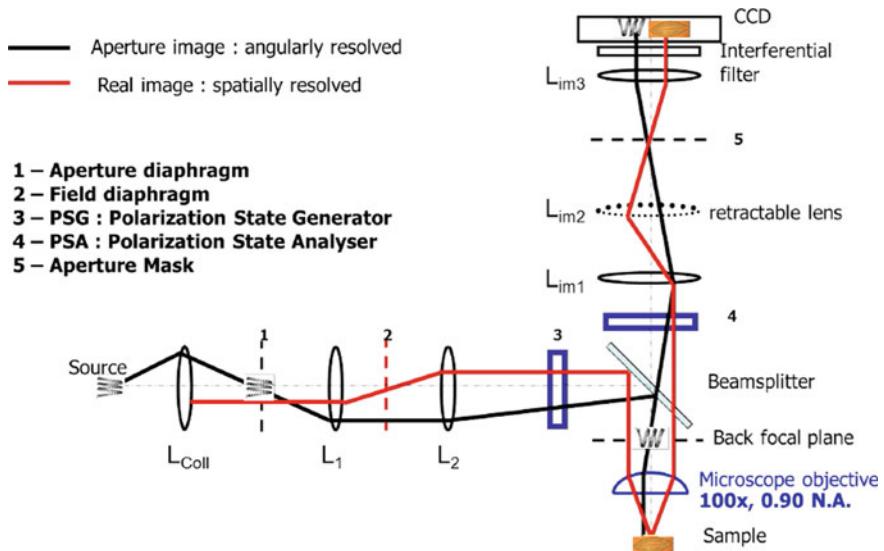
Fig. 2.28 Schematic representation of the imaging/conoscopic Mueller polarimeter

On the detection side, we find

- The *beamsplitter* again, working this time in transmission,
- The PSA 4 to analyse the emerging polarization,
- A set of two lenses which *image the objective back focal plane onto a two dimensional* imaging detector,
- A "retractable" lens which can be inserted in the beam path *to image the sample* instead of the objective back focal plane,
- An *aperture mask* 5 can be set in a plane conjugated with the objective BFP, typically to eliminate some strong contributions in order to see weaker ones, or to select the visualized diffraction orders if the sample is a grating.
- An interferential filter, typically quite narrow for metrological applications.
- The camera, a backthinned and cooled 512×512 pixel CCD from Hamamatsu.

The PSG and PSA operating in this setup are essentially the same as those used in the macroscopic imagers, though of smaller size. This choice was driven by the same considerations as for the macroscopic imager: the samples were static, and we considered it was important to be able to minimize the condition numbers at each wavelength [68, 74].

The angular distribution and the spot size characterizing the beam incident onto the sample can be adjusted independently of each other (at least as long as the illumination beam is far from being diffraction limited, a condition which is always fulfilled in practice with the classical light sources such as the one we use).

The two modes of operation of the microscope are illustrated on Fig. 2.29. The left panel shows the real space image of image of a grating, and the reciprocal space
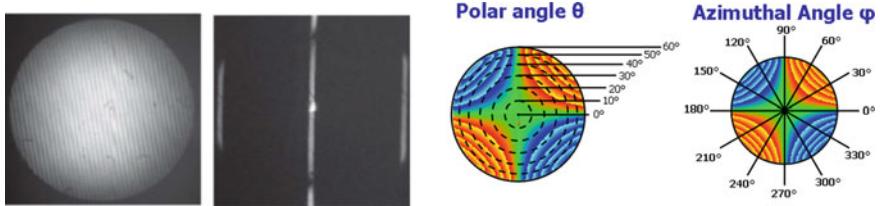
Fig. 2.29 *Left* Real-Fourier space images of a grating. *Right* Angular coordinates in the Fourier space (maximum aperture 62°). Images taken from Ref. [72] with the permission of Willey Science

image obtained with a slit as an aperture diaphragm and the grating as the sample: due to the presence of the slit, the angular distribution of the incident light is almost 1D, which is then diffracted in orders 0 (central line) and $+ - 1$ (lateral images).

The right panel of Fig. 2.29 shows how the angular distribution of the light coming from the sample is actually mapped on the objective back focal plane. Due to Abbe's sine condition [4] a parallel beam emerging from the sample with a polar angle θ and an azimuth ϕ is focussed in the back focal plane on a point with radial coordinates $(f \sin\theta, \phi)$, where f is the objective focal length. Of course, while all the azimuths between 0° and 360° are mapped, the polar angles θ are limited by the numerical aperture of the objective. In our case the nominal values are $\sin \theta_{max} = 0.90$ and $\theta_{max} = 65°$. In practice, it is difficult to achieve the full angular range. The radial coordinate is calibrated on the images by using diffraction patterns obtained with gratings with known pitches, such as the pattern shown in the left panel of Fig. 2.29, and our images are limited to about 60°.

Imaging in reciprocal space may constitute an interesting alternative to the more conventional approach of goniometric ellipsometry/polarimetry if angularly resolved data are to be acquired. Measurements along the polar angle at a fixed azimuthal angle of an image are equivalent to measurements taken at different angles of incidence on a non-imaging system. Accordingly, measurements recorded at a fixed polar angle and along the azimuthal direction on an image are equivalent to measurements taken rotating the sample holder in a non-imaging ellipsometer. With respect to simple conoscopy through crossed polarizers, full polarimetric conoscopy can be very useful to characterize anisotropic materials, as it provides angularly resolved maps of retardation (and diattenuation, if present) which significantly constrains the values of the dielectric tensor from easy and fast measurements [75] Moreover, under a powerful microscope objective, the *spot size can easily be reduced to* 10 μm *or less*, a possibility which can be very useful for some metrological applications, and more particularly in microelectronics, as we will see in Sect. 7.3. Obviously, so small spots sizes are much more difficult to obtain with the usual ellipsometric setups involving narrow beams with low numerical apertures.

As an example of Mueller images in the Fourier space we show in left panel of Fig. 2.30 the data taken on a silica thick plate. At first sight the observed patterns may seem surprising for an isotropic sample. In fact, the isotropy is "broken" by
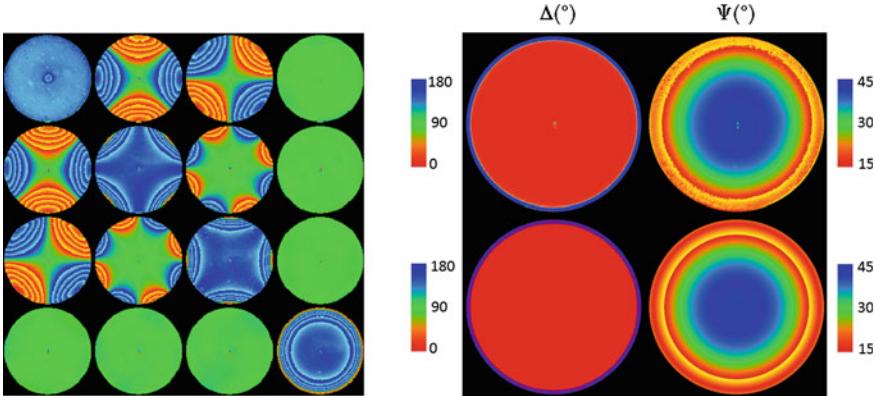
Fig. 2.30 *Left* raw Mueller images in the reciprocal space of a *thick plate* of silica. The basis vectors for the definition of polarisaton are vertical and horizontal all over the image. *Right*, *top* maps of angularly resolved Δ and Ψ derived from the Lu–Chipman decomposition of the data shown in the *left panel*. *Right bottom* corresponding simulations

the choice of the basis used to define the polarization, and which is uniform all over the image with one vector horizontal and the other vertical. Obviously, these are not the usual (*p, s*) vectors defined with respect of the incidence plane, and which would be oriented radially in each point of the image.

If the Lu–Chipman decomposition is applied to the raw image at the left of Fig. 2.30 we obtain essentially zero depolarization, and the angularly resolved values of Δ and Ψ shown in the right panel (top), together with the corresponding simulations (bottom). As expected, once Δ and Ψ are plot according to the conventional (s,p) coordinate system, both parameters display an almost perfect radial symmetry. Moreover Δ remains at zero, and then "jumps" to 180° at the Brewster incidence, while Ψ starts at 45° at the image center, and then decreases in agreement with the theory (the jump from red to light yellow indicates that the plotted value went below the minimum of the scale, here 15°).

These results clearly show that this technique may be very powerful. We give in Sect. 2.7.3 an example of application for the metrology of sub-wavelength gratings. However, it would be extremely difficult to "push" the accuracy of such measurements to the levels reached by usual, non-imaging ellipsometers. The main reason for this is that the objectives used in Fourier configuration may introduce some polarimetric artifacts which cannot be taken into account by the ECM method, as the system must be calibrated with the objective removed [72]. Moreover, even strain free objectives are extremely sensitive to mechanical constraints, and the resulting artifacts would probably evolve in time. In spite of these limitations in accuracy, Mueller microscopes (operating here in reflection, but transmission may be used too) are likely to open new research topics in many areas.

2.7 Examples of Applications of Mueller Ellipsometry

One of the interest of using Mueller ellipsometers is to access the whole Mueller matrix in a single measurement configuration to characterize samples with a complex optical response such as anisotropic layers or diffracting structures. Another interest of having broadband spectral measurements is to access different types of properties of the solids. In the following we provide three examples showing the performances of the above described systems to perform optical metrology of diffracting structures and to characterize the optical (and electrical) properties of thin metal films.

2.7.1 Characterization of Low Emissivity Coatings Using Broadband Mueller Ellipsometry and Standard Ellipsometry in Total Internal Reflection

In the first part of this chapter, it is stated that ellipsometry is an indirect technique, here we provide an example of a common way to work with ellipsometric data from non-depolarizing samples; i.e. the use of parameterized optical models to fit experimental data. This example also illustrates the advantages of using data in a broad spectral range and the interest of combining data from different sources and optical configurations to optimize the amount of information that can be obtained from the sample. In particular we show how to use optical data to study the dependence of the electrical properties (conductivity) of thin metallic layers with their thickness. The link between electrical and optical properties can be established because the infrared optical response of conducting materials is dominated by the dynamics of free charge carriers and can be written in terms of the electrical conductivity in the framework of the Drude model. We have worked with multi-layer stacks of aluminum doped zinc oxide, ZnO:Al, and silver, Ag, with the structure ZnO:Al/Ag/ZnO:Al. Silver layer thickness changed from sample to sample in the range from 8 to 50 nm, while the thickness of the ZnO:Al layers was kept constant ($\approx$20 nm). ZnO:Al and Ag layers were deposited by magnetron sputtering. In the following we present a summary of a detailed study made on the set of samples. More information about measurements, parameterization of the dielectric function of the materials present in the samples, and fitting of data can be found elsewhere [76].

The optical characterization of the samples was performed with two techniques: spectroscopic ellipsometry and infrared reflectometry. Infrared reflectometry measurements were taken with a commercial (Bruker) FTIR Spectrophotometer in the spectral range from 4 to 30 μm. Spectroscopic ellipsometry measurements in the ultraviolet to the near-infrared (270–1600 nm) were performed with a commercial phase-modulated ellipsometer UVISEL (HORIBA Jobin-Yvon). Mueller ellipsometric data in the mid-infrared were measured with the Mueller ellipsometer described in Sect. 2.6.5.1 The angle of incidence was the same, 68.5°, for the UVISEL and
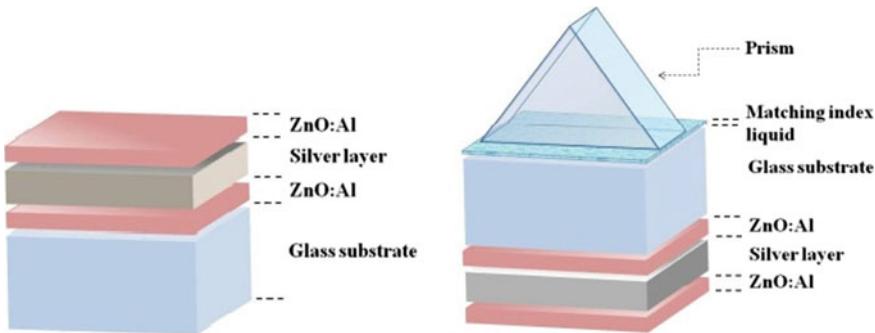
Fig. 2.31 *Left* sketch of the ZnO:Al/Ag/ZnO:Al stack on a glass substrate. *Right* sketch of TIR configuration used and the orientation of the sample. The prism used for TIR measurements was cut at an angle of 57.6°

the broadband Mueller ellipsometer in order to have continuous data. Measurements were performed from the coated side as depicted in Fig. 2.31. Total internal reflection ellipsometric (TIR) measurements were performed using the UVISEL spectroscopic ellipsometer because the most prominent spectral features appear at visible frequencies. The TIR measurement configuration, sketched in Fig. 2.31, consists of a BK7 prism optically matched to the substrate by an oil designed for this purpose. Measurements were performed from the substrate side of samples as depicted in Fig. 2.31.

Experimental data for all samples was interpreted on the basis of a common model of the structure which included as parameters the thickness and dielectric functions of the three layers. The dielectric function of the substrate was considered as well. The dielectric function of the glass at infrared frequencies was described with a combination of four Lorentz oscillators (OS) whereas in the ultraviolet it was represented using the Tauc-Lorentz (TL) model [3, 31, 77]. ZnO:Al showed strong ultraviolet absorption due to interband transitions which were modeled with a TL formula with two oscillators. Silver is a noble metal which shows a noticeable absorption in the whole spectral region from the infrared to the ultraviolet. Above the plasma frequency the absorption is governed by interband transitions, and below it by intraband absorptions due to free carriers. Free carrier contribution was modeled with the Drude expression, whereas interband transitions were modeled with TL expression.

Glass substrate was characterized following the example given in [78]. Ellipsometry and reflectometry measurements were performed on both sides of the uncoated glass substrate. In Fig. 2.32, we provide a comparison between best-fitted data and experimental values obtained for the glass substrate at the air side. In addition, the values of the refractive index and absorption coefficient deduced from the best-fitted parameters are also plotted in the same figure.

For the characterization of optical properties of ZnO:Al we worked with a sample consisting of a glass substrate coated with a 30 nm layer of ZnO:Al. The optical model of the sample was identical to the glass substrate with an additional layer
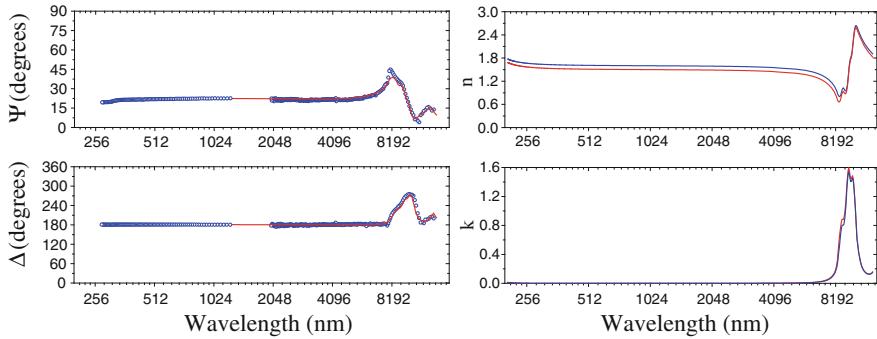
Fig. 2.32 *Left* experimental (*blue dots*) and best-fitted (*red lines*) values for the ellipsometer angles Ψ and Δ corresponding to the air side of the glass substrate. *Right* best-fitted refractive index n, and absorption coefficient k, corresponding to bulk glass (*red line*) and glass with tin (*blue line*). Spectral range spans from the ultraviolet to the mid infrared, the wavelength axis of both figures is represented in a logarithmic scale

representing the ZnO:Al film. In Fig. 2.33, we provide a comparison between the best-fitted and the experimental data obtained for the ZnO:Al sample. The same figure also shows the values of the refractive index and the absorption coefficient deduced from the fitted data.

Unlikely to zinc oxide, it is not possible to characterize a single layer of silver because it oxidizes quickly in contact with the air. In order to work with silver, we had to use the complete ZnO:Al/Ag/ZnO:Al tri-layer samples. We measured the sample having the thickest film, 50 nm. From the set of parameters characterizing the Ag optical response, the most relevant for the purposes of the present example are only two: the plasma frequency, ω_p, and the Drude damping factor, Γ_D. The best-fitted value of ω_p was 9.2 ± 0.2 eV and it was found to be stable from sample to sample [76, 79]. On the contrary, the value of the Drude damping, which is the inverse of the mean time among two collisions of electrons, was expected to strongly depend on the thickness of the silver layers. Considering that the mean free path (MFP) of electrons in bulk silver is about 40 nm, it is not surprising that once dimensions of the layer become of the same order or even smaller than MFP, the interfaces of the film as well as the microcrystalline grain boundaries, have a non-negligible impact on the kinetics of free electrons. The fitted value for Γ_D for the thicker layer was about 0.038 eV, a value which is compatible with the literature.

Once the dielectric functions of the materials were known, it came the problem of fitting the three layer system. It was found that it was not possible to fit properly the thickness of the three layers and the Drude damping using either ellipsometric data in external configuration alone or reflectometry data alone. The problem was the strong numerical correlations between parameters. This limitation is due to the loss of information contained in optical data because of the absorption of the Ag layer. To circumvent this problem we added to our fitting procedure the measurements performed in TIR configuration. This approach allowed to enhance the sensitivity
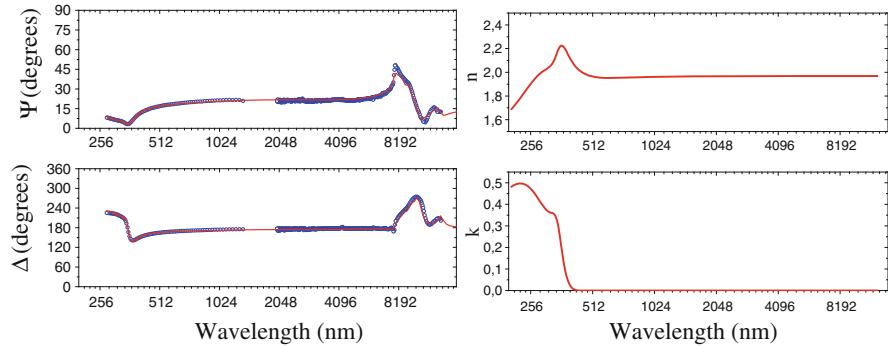
Fig. 2.33 *Left* experimental (*blue dots*) and theoretical (*red lines*) values for the ellipsometric angles Ψ and Δ corresponding to the ZnO:Al single layer sample. *Right* best fitted refractive index, n, and absorption coefficient, k. Spectral features visible in Ψ and Δ at infrared wavelengths larger than 8 μm are created from the glass substrate. Spectral range spans from the ultraviolet to the mid infrared, the wavelength axis of both figures is represented in a logarithmic scale
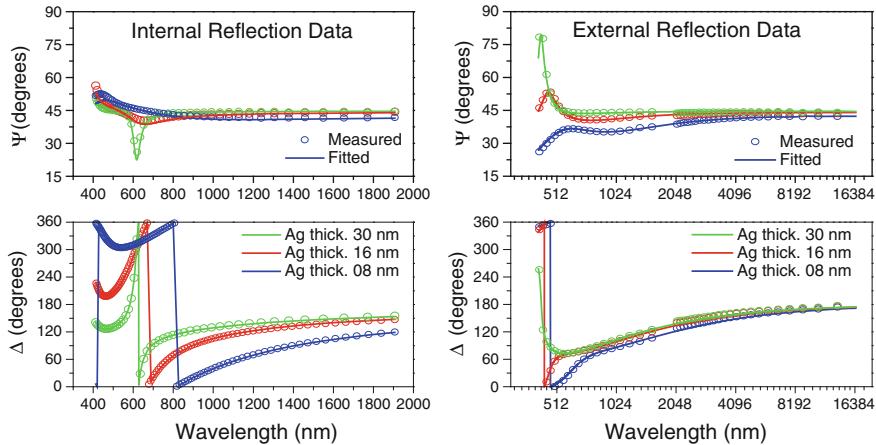
Fig. 2.34 Experimental (*circles*) and best-fitted (*lines*) Ψ and Δ spectra corresponding to samples with silver thickness of 8 (*blue*), 16 (*red*) and 30 (*green*) nm. *Left* ellipsometric data in TIR configuration. *Right* ellipsometric data in external configuration

of measurements and to remove most of the numerical correlations. In Fig. 2.34 we have represented the experimental and the best-fitted ellipsometric data for three representative samples.

The main advantage of using ellipsometer in TIR configuration is the excitation of resonant surface plasmon waves on metal/dielectric interfaces [80–83]. When the metallic surface is covered with thin films, the field distribution and the propagation characteristics of plasmons are strongly influenced by the thickness and refractive index of these layers. Contrarily to standard propagating (plane) waves, the way surface plasmons interact with thin films is substantially non-interferometric, which
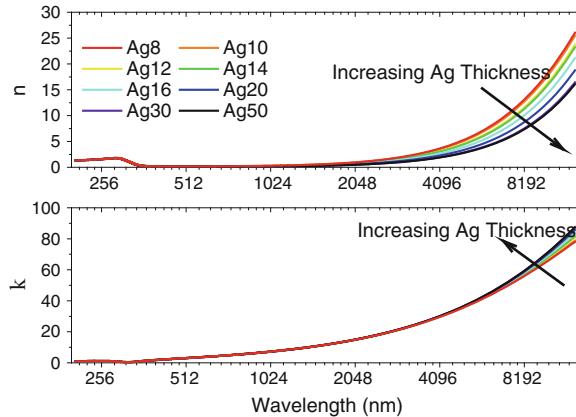
Fig. 2.35 Spectral dependence of the values of the refraction index, n, and the absorption coefficient, k, for each sample studied. The *arrows* indicate spectra corresponding to samples with increasing silver layer thickness. Spectral range spans from the ultraviolet to the mid infrared, the wavelength axis of both figures is represented in a logarithmic scale

explains why surface plasmons can retrieve an information which is not affected by the same limitations as that carried by light beams in external configuration and thus, enhance the sensitivity of ellipsometric measurements.

As expected, we observed that Γ_D increased when the thickness d_{Ag} decreases. Γ_D is proportional to the collision frequency. The observed behavior of Γ_D is explained because when the silver layer thickness decreases, scattering of electrons at the layer interfaces and at grain boundaries becomes non-negligible and it adds-up to the intrinsic bulk scattering by phonons and impurities. The value of Γ_D impacts considerably the refractive index and the absorption coefficient of the Ag layer specially at long wavelengths. As shown in Fig. 2.35, the refractive index decreases and the absorption coefficient increases when Γ_D increases.

In order to study the dependency of Γ_D with the silver layer thickness, the optical conductivity of each sample was calculated from the fitted data and compared with the corresponding values of the measured electrical conductivity. In the framework of the Drude theory the optical conductivity can be calculated as follows: [3]

$$\sigma_{c,opt} = \frac{\varepsilon_0 \omega_p^2}{\Gamma_D} = \frac{1}{\rho_{c,opt}}$$

According to [84, 85] the most convenient is not to represent resistivity, but the resistivity times the thickness of the silver layer, $\rho_c d_{Ag}$, as a function of the silver layer thickness itself. It can be shown that for metals this relation is linear. The slope of the line is related to the intrinsic resistivity of the silver layer, ρ_0, and the cut with the y-axis gives information about the influence of the surface scattering. In Fig. 2.36
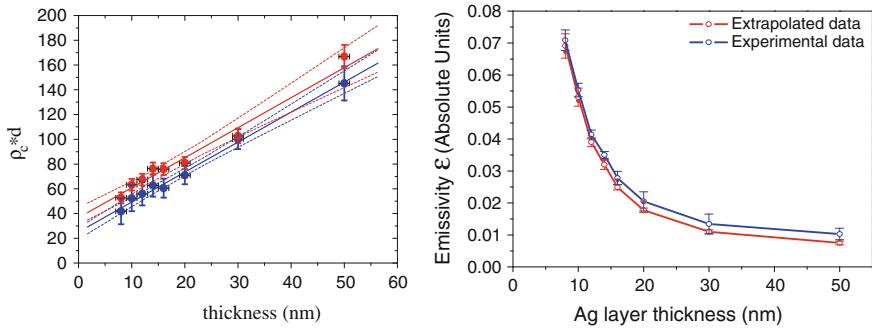
Fig. 2.36 *Left* electrical (*blue spots*) and the optical (*red spots*) results in terms of the $d_{Ag}\rho_c$ product as function silver layer thickness d_{Ag}. Linear regression of the experimental electrical data (*blue line*) and optical data (*red line*) are represented together with the corresponding 95 % confidence intervals (*blue* and *red dotted lines*). *Right* experimental (*blue*) and extrapolated (*red*) emissivity as a function of the Ag layer thickness for the whole set of samples

are represented the values of the product $\rho_c d_{Ag}$ obtained with optical and electrical measurements respectively as a function of d_{Ag}.

The slope of the linear regression (also shown in the figure) corresponding to the electrical and optical data were $(2.1\pm0.9)\mu\,\Omega$ cm and $(2.1\pm0.3)\mu\,\Omega$ cm respectively. These values are slightly higher than the value for bulk monocrystalline silver at room temperature, $1.59\mu\,\Omega$ cm, because the silver in our samples was polycrystalline. To validate the results obtained from optical data fitting, we compared the emissivity extrapolated from the optical model for each sample with experimental data. The results, shown in Fig. 2.38, prove that the high accuracy of the optical model allows to predict emissivity which matches very well with experimental data.

In summary, in this example we wanted to illustrate the way to treat ellipsometric data from non-depolarizing samples and also to show the interest of combining data from different sources in a wide spectral range to optimize the amount of information that can be obtained from the sample. In our particular example we have applied a combined approach to characterize a tri-layer structure in which an Ag layer was stacked between two ZnO:Al layers. We have used Mueller ellipsometric data in the infrared to obtain the optical properties of Ag layers, which are responsible of the thermal emissivity of the coatings. Moreover, the enhanced sensitivity due to the penetration of surface plasmons through the tri-layer stack measured in TIR configuration, allowed us to achieve the better accuracy for the thickness of the layers. The validation of the method using an independent measurement (emissivity), confirmed the accuracy of characterization achieved with ellipsometry.

This example shows how Mueller ellipsometric data in the infrared can be used to study the dependence of the electrical properties (conductivity) of thin metallic layers with their thickness. The link between electrical and optical properties can be done because the infrared optical response of conducting materials is dominated by the dynamics of free charge carriers. We have studied multi-layer stacks of aluminum doped zinc oxide, ZnO:Al, and silver, Ag, forming the structure ZnO:Al/Ag /ZnO:Al

with thin Ag films of following thicknesses: 8, 10, 12, 14, 16, 20, 30, and 50 nm. The thickness of the ZnO:Al layers was kept constant ($\approx$20 nm). ZnO:Al and Ag layers were deposited by magnetron sputtering. In the following we present a summary of a detailed study made on the set of samples, here we want to highlight the benefits of having optical data in a broad spectral range.

2.7.2 Characterization of Spatially Inhomogeneous Samples by Sum Decomposition of the Measured Mueller Matrix

This example illustrates one practical use of the sum decomposition, the retrieval of non-depolarizing components of a depolarizing Mueller matrix. A beam which shines two parts of a sample with different optical properties can lead to depolarizing Mueller matrices because the two responses add-up incoherently in the detector. This situation can easily occur when it comes to measure samples with characteristic size comparable to the size of the probe beam. If the alignment of the sample with the beam is not extremely accurate, a portion of the beam may fall outside of the sample, then the combination in the detector of beam parts having fallen inside and outside the sample causes the measured Mueller matrix to be depolarized. This example shows how to separate the Mueller matrix of the sample from the Mueller matrix of the surrounding substrate provided that both matrices are non-depolarizing and the matrix of the substrate is known. The adequate decomposition to treat this simple example is the Cloude decomposition.

The sample consisted of a one-dimensional diffraction grating etched on a photoresist deposited on the surface of a silicon wafer. More details about the grating profile and the etching procedure can be found elsewhere [86]. The size of the etched area containing the grating was about 3×3 mm, large enough to contain the spot size of the Mueller ellipsometer, $200 \times 400\,\mu$m projected on the sample when the grating was properly aligned. The etching process removed the photoresist coating surrounding the grating, leaving exposed the surface of the bare silicon substrate. Therefore, the measured signal could come from the substrate, from the sample or from a mixture of both depending on the alignment of the beam. The lines of the grating were oriented at 45° respect to the plane of incidence to maximize the difference between the polarimetric optical response of the grating and the surface.

As shown in Fig. 2.37, the measurement process started positioning the beam perfectly aligned to the center of the grating and measuring the corresponding Mueller matrix. Then, the beam was slightly shifted from the initial position and the sample was measured again. The procedure was repeated until the entire beam spot fell well outside the grating. The goal was to measure the depolarization of all the measured matrices, to separate the depolarizing matrices in two components, and to measure the relative weight of each non-depolarizing component as a function of total shift of the beam respect to the initial position. The matrices of the initial and the final steps corresponding to the grating and the substrate respectively, were found to be non-depolarizing. Ellipsometric measurements were performed with a com-
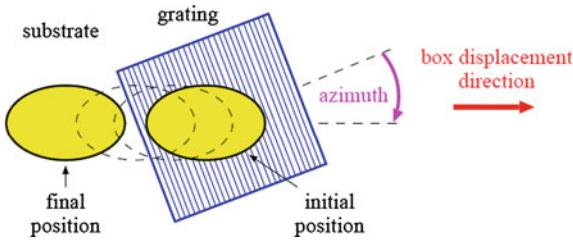
Fig. 2.37 Schematic representation of the sample and the measurement procedure. The *stripped square* represents the grating whereas the *white* area surrounding it represents the silicon substrate. The *yellow ellipses* represent the projection of the spot size on the sample surface in the initial and the final step. A few intermediate measurement steps are represented by the *dashed ellipses*. Picture taken from [86] with the authorization of the authors and the publisher

mercial liquid crystal based spectroscopic Mueller matrix ellipsometer (MM16 from HORIBA Jobin-Yvon), operating in the visible range (450–850 nm with a spectral step of 1.5 nm) previously described in this chapter.

The method to separate the Mueller matrix of the grating from the experimental Mueller matrix is based on the Cloude sum decomposition. This decomposition implies to work with coherence matrices given by expressions (2.43) or (2.18). Because in this particular experience, there are two components that can contribute to the detected signal, the grating and the substrate, the rank of the coherence matrix $\mathbf{N^e}$ associated to the depolarizing measured Mueller matrix, $\mathbf{M^e}$, must be two (2). It is assumed that both the substrate and the grating are non-depolarizing, thus their respective Mueller matrices $\mathbf{M^s}$, and $\mathbf{M^g}$, have associated matrices $\mathbf{N^s}$ and $\mathbf{N^g}$ of rank one. According to the experimental conditions, the normalized depolarizing Mueller matrix $\mathbf{M^e}$ resulting from the sum of two normalized non-depolarizing matrix components $\mathbf{M^s}$ and $\mathbf{M^g}$ can be written in form:

$$\mathbf{M^e} = \frac{1}{1+p} \left(\mathbf{M^g} + p\mathbf{M^s} \right) \qquad (2.119)$$

where the coefficient p accounts for the relative weight of the two non-depolarizing matrices. Therefore, when $\mathbf{M^e}$ and $\mathbf{M^s}$ are known, it is always possible to use find unique real number, α, such that the rank of the matrix $(\mathbf{N^e} - \alpha \, \mathbf{N^s})$ equals one [87]. The solution of this algebraic problem can be performed numerically with an algorithm that searches the value of parameter α by minimizing the values of three eigenvalues associated to the matrix $(\mathbf{N^e} - \alpha \, \mathbf{N^s})$. Ideally, for the matrix $(\mathbf{N^e} - \alpha \, \mathbf{N^s})$ to be of rank one, three of its four eigenvectors must be zero. Because the matrices are experimental, they have some noise, then for practical reasons related to the minimization routine, the noise level is numerically considered as zero. Once the value of the parameter α is known, the normalized Mueller matrix $\mathbf{M^g}$ an be written in terms of $\mathbf{M^e}$ and $\mathbf{M^s}$ as:
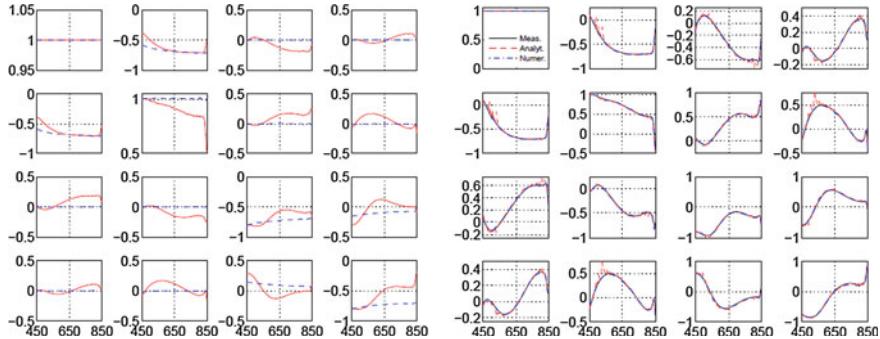
Fig. 2.38 *Left* experimental data corresponding to a normalized depolarizing Mueller matrix $\mathbf{M^e}$ (*solid red line*) and the normalized non-depolarizing Mueller matrix of the substrate $\mathbf{M^s}$ (*dashed blue line*). The different boxes arranged in 4×4 arrays correspond to each one of the sixteen elements of the normalized Mueller matrices. *Right* spectral values of the numerically (*dash-dotted blue line*) retrieved Mueller matrices compared with the directly measured matrix of the grating (*solid black line*). The *dashed blue lines* correspond to an alternative way to calculate the coefficient α, not commented in this chapter. More details about the alternative method can be found in [88] from which the figure has been reproduced with the authorization of the authors and the publisher

$$\mathbf{M^g} = (1 + p)\,\mathbf{M^e} - p\mathbf{M^s} \quad \text{with} \quad p = \left(\frac{\alpha}{1 - \alpha}\right) \tag{2.120}$$
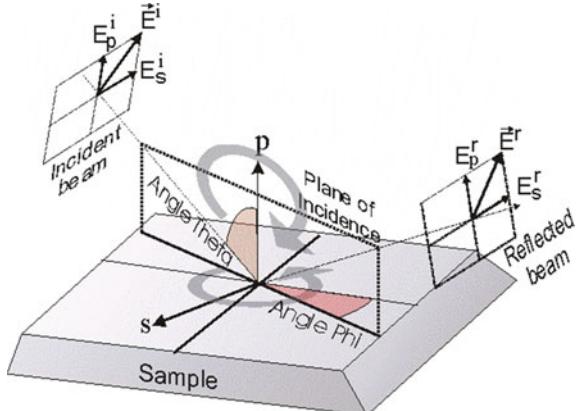
Therefore, the method allows the extraction of the unknown non-depolarizing component $\mathbf{M^g}$ from the original depolarizing matrix by subtracting the right proportion of the known component $\mathbf{M^s}$. The method can be further generalized to extract more than one non-depolarizing component from the depolarizing matrix [88]. Figure 2.38 shows one of the depolarizing Mueller matrices $\mathbf{M^e}$ and the non-depolarizing Mueller matrix $\mathbf{M^s}$. The same figure also shows the matrix $\mathbf{M^g}$ resulting from the decomposition of matrix $\mathbf{M^e}$ together with the experimental matrix of the grating measured alone (first step in the measurement process). The overall correspondence between measured and numerically reconstructed matrices is very good, which shows the validity of the extraction procedure.

In summary, this example illustrates one of the possible uses of the Cloude decomposition, i.e. the extraction of pure non-depolarizing matrices, which can be in turn analyzed using ellipsometric models, from depolarizing matrices. A method to implement the subtraction has been presented. The robustness of the method is illustrated in an example concerning depolarizing Mueller matrices containing two non-depolarizing components, a diffraction grating and a silicon substrate respectively.

2.7.3 Subwavelength Grating Metrology

In this section we focus on the profile reconstruction of diffraction gratings. Optical methods (also called "scatterometry") are fast, non-destructive and may exhibit strong sensitivity to tiny changes in grating profiles [89]. As a result, they are becoming increasingly popular for process control in microelectronics industry [90]. On the other hand, these methods are indirect, and the reconstructed profiles may depend on the model used to fit the data (and the dielectric function of somewhat "ill-defined" materials like resists.) Possible model inadequacies do not necessarily appear in the goodness of fit. Parameter correlations may also constitute a serious issue, as shown in a comprehensive study of the results of scatterometric reconstruction by the usual techniques (normal incidence reflectometry and planar diffraction spectroscopic ellipsometry) of various profiles representing different technological steps [91]. In this context, Mueller ellipsometry may constitute an interesting alternative, provided the data are taken in conical diffraction geometries. In conical diffraction configurations, the symmetry axes of the grating structure are neither parallel or perpendicular respect to the plane of incidence. Indeed, in such geometries the grating Jones matrix is no longer diagonal (and the Mueller matrix no longer block-diagonal). As a result, additional information is available, and may help in constraining the fitting parameters. Moreover, the stability of the optimal values of these parameters when the azimuth is varied may constitute a much better test of the model relevance than goodness of fit at a single azimuth [92]. The principle of ellipsometric measurements is sketched in Fig. 2.39. The relevant parameters characterizing the measurement are the angle of incidence, θ, and the azimuthal angle φ, which is the angle between the plane of incidence and one of the axes of symmetry of the sample. For one-dimensional gratings, the azimuth is defined to be zero when the grooves are aligned perpendicular to the plane of incidence. Only in the case the grooves are ori-

Fig. 2.39 Sketch of a typical Mueller ellipsometric configuration to perform measurements in conical diffraction mode. The relevant angles are the angle of incidence, θ, and the azimuthal angle φ. In the figure only the zero order diffracted beam is represented for simplicity reasons

ented parallel or perpendicular to the plane of incidence the measurement geometry is called planar diffraction.

For non-depolarizing samples, either Jones or Mueller formalisms are valid to represent the properties of the sample and Mueller matrices can be related to Jones matrices with expression (2.14). This particular representation of Mueller matrices is useful to discuss the symmetry properties of the Mueller matrix of a sample measured in conical diffraction. Li demonstrated in [93] that the off-diagonal elements of the Jones matrix are anti-symmetric for symmetric structures as a result of the electromagnetic reciprocity theorem. Translating this condition to expression (2.14) gives a Mueller matrix which satisfies the following symmetry $M_{12} = M_{21}; M_{14} = M_{41}; M_{24} = M_{42}$ and anti-symmetry conditions: $M_{13} = -M_{31}; M_{23} = -M_{32}; M_{34} = -M_{43}$. Figure 2.40 show schematically all these relations:

When the a symmetric grating is rotated azimuthally of $\pm180°$ the resulting geometry is equivalent to the original non-rotated one, except that the s component of the electric field have changed its sign. In other words, "p(φ)" $=$ "p($\varphi \pm 180$)" "s(φ)" $= -$"s($\varphi \pm 180$)". This transformation makes the signs of all the off-diagonal elements of a rotated Mueller matrix to change respect to the elements of the non-rotated matrix. In particular for a symmetric grating the following relations are always satisfied:

$$
\begin{aligned}
M_{13}(\varphi) &= -M_{13}(\varphi \pm 180°) = -M_{31}(\varphi) = M_{31}(\varphi \pm 180°); \\
M_{14}(\varphi) &= -M_{14}(\varphi \pm 180°) = M_{41}(\varphi) = -M_{41}(\varphi \pm 180°); \\
M_{23}(\varphi) &= -M_{23}(\varphi \pm 180°) = -M_{32}(\varphi) = M_{32}(\varphi \pm 180°); \quad (2.121) \\
M_{24}(\varphi) &= -M_{24}(\varphi \pm 180°) = M_{42}(\varphi) = -M_{32}(\varphi \pm 180°); \\
M_{34}(\varphi) &= -M_{34}(\varphi \pm 180°) = -M_{43}(\varphi) = M_{43}(\varphi \pm 180°);
\end{aligned}
$$

These symmetry relations under rotation (2.121), are not fulfilled if the grating is not perfectly symmetric. The lack of symmetry can be due the presence of overlays or just because of imperfections of the grating profile occurred during the fabrication process. In many practical situations, the asymmetric effects due to overlays are small, and may be comparable in magnitude to the measurement errors of standard ellipsometric systems. In those cases, the fact of being able to measure a full Mueller matrix allows to profit the redundancies of all elements of the Mueller matrix to achieve an optimal discrimination between overlay effects and experimental measurement errors. Therefore, when only partial Mueller matrix are available, an extra-care must be taken in order to not mix-up both effects. In the following we show two examples of use of measurements of gratings in conical diffraction to characterize diffraction gratings. The first example is about profile reconstruction using spectroscopic Mueller matrices and the second example illustrates the use of the imaging Mueller ellipsometer to characterize small overlays.

$$
\begin{array}{|c|c|c|c|}
\hline
m_{11} & m_{12} & m_{13} & m_{14} \\
\hline
m_{12} & m_{22} & m_{23} & m_{24} \\
\hline
-m_{13} & -m_{23} & m_{33} & m_{34} \\
\hline
m_{14} & m_{24} & -m_{34} & m_{44} \\
\hline
\end{array}
$$
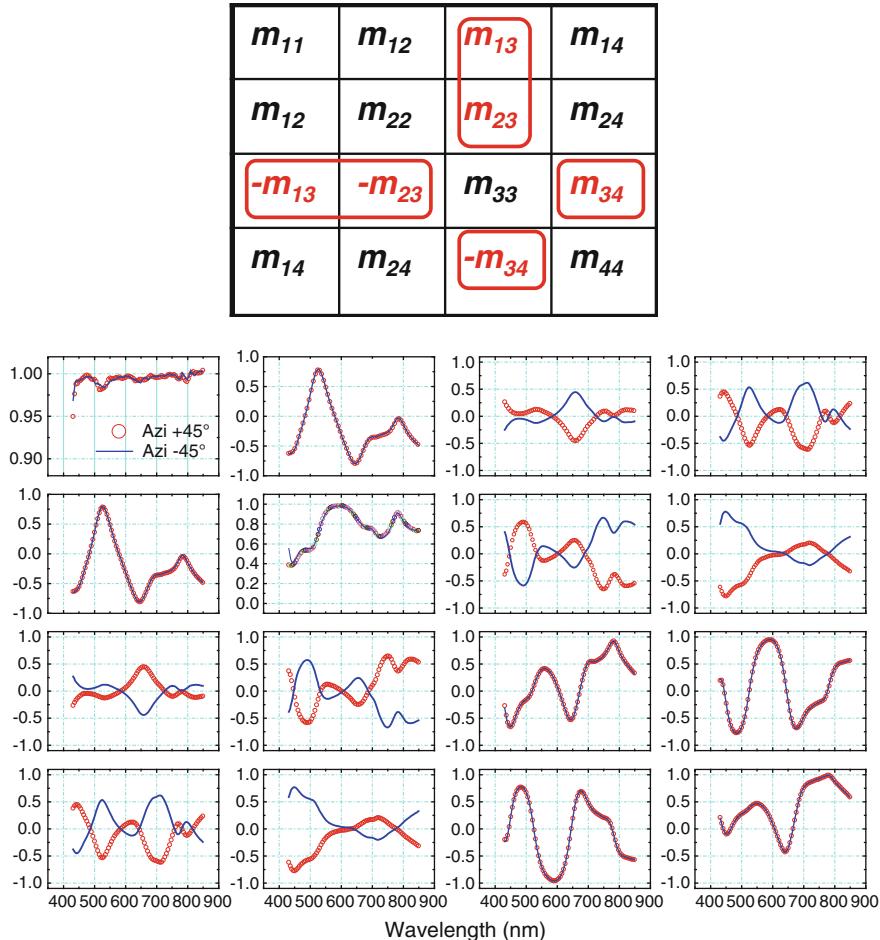
Fig. 2.40 *Top* schematic representation of the Mueller matrix of a symmetric grating, showing in *black* and in *red* the elements showing a symmetric an anti-symmetric relation respectively. Diagonal elements are not subjected to symmetry relations. *Bottom* spectroscopic Mueller matrix of a one-dimensional grating measured at two azimuths, +45° (*red circles*) and −45° (*blue line*) and at angle of incidence of 45°. Two types of symmetries can be seen. The off-diagonal elements of the matrix at azimuth +45° have opposite sign respect to those of the matrix at −45, whereas the rest of elements have equal sign. Moreover for every single Mueller matrix, the symmetries among its respective elements, schematically show in the figure on *top*, are always respected

2.7.3.1 Profile Reconstruction by Spectroscopic Mueller Ellipsometry at Discrete Angles

Here we will summarize the main results of a study which goal was to show the possibility of using Mueller ellipsometry data for reconstruction (optical metrology) of diffraction gratings profile. The sample analyzed consisted of a silicon wafer
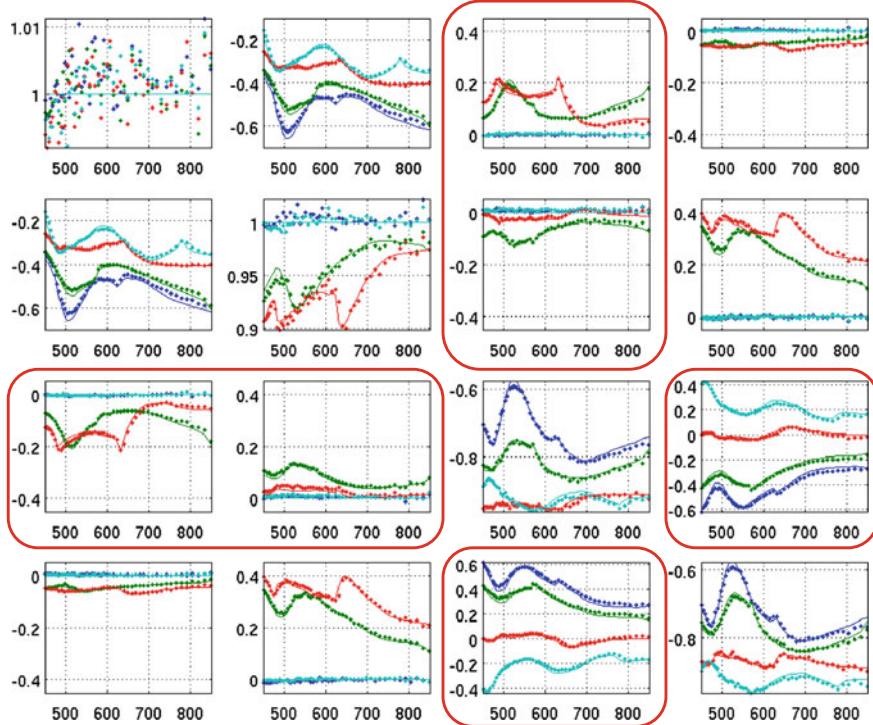
Fig. 2.41 Measured (*dots*) and fitted (*lines*) spectroscopic Mueller matrices for azimuthal angles of 0° (*blue*), +30° (*dark green*), +60° (*red*) and +90° (*light green*) and incidence angle of 45°. Spectral range spanned from 450 to 850 nm. Images taken from Ref. [71] with the authorization of the authors and the publisher

with a series of silicon gratings etched on it using UV beam lithography. Typical dimensions for the gratings were: groove depths around 100 nm, line widths around 130 and 250 nm, and pitches from 500 to 1100 nm. Each individual grating was etched in an area of 3 × 3 mm. Etched silicon gratings were chosen for this study because of their long term dimensional stability, higher refractive index contrast and relevance to semiconductor industry. For reference, the dimensions of the profiles of the gratings were determined by the state of the art 3D AFM microscope. For the sake of simplicity we show here the results corresponding to only one grating. For more details, please refer to [94]. Experimental data was taken by an Horiba Jobin-Yvon Mueller ellipsometer (MM16), operating in the visible (450–850 nm) [68]. A series of measurements were taken varying the azimuth over 360° in steps of 5°. The incidence was kept constant at 45° to make sure the beam diameter at the sample was small enough to safely maintain the spot within the grating. Four of the measured spectroscopic matrices, corresponding to azimuthal angles 0, +30, +60 and 90°, together with the corresponding fits, are shown in Fig. 2.41. The matrix elements are normalized by the element M_{11}, and thus they vary from −1 to 1. The redundant information in the Mueller matrix allows to evaluate the quality of
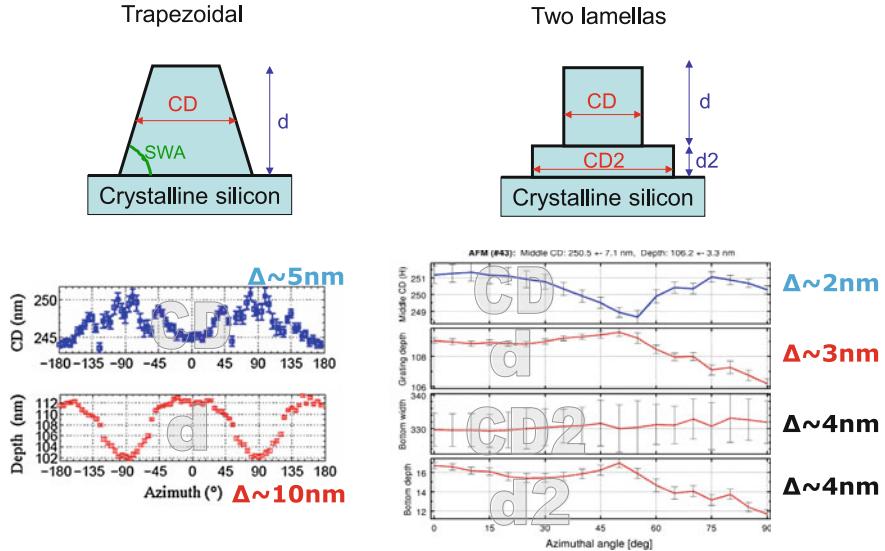
Fig. 2.42 *Left top* profile representing the trapezoidal model with the two characteristic parameters, the CD, the thickness and the SWA. *Left bottom* resulting best-fitted values for the parameters CD and thickness at different azimuths. The corresponding variations are indicated by the symbols Δ. *Right top* profile representing the double lamella model with the four characteristic parameters. *Right bottom* results of the fit of four free parameters of two lamellas model over different azimuthal angles. Error bars in figures denote statistical errors. The maximum variation of each parameter is indicated with the symbol Δ. Images taken from Ref. [94] with the authorization of the authors and the publisher

the measured data by simple criteria such as the degree of polarization, expression (2.19), or the symmetry of the off-diagonal elements (2.121). In the upper left corner panel of Fig. 2.41 there is plotted the degree of polarization, which was found to be very close to 1, indicating that no depolarization phenomenon occurs because of the high quality of both sample and data. The blue and green spectra coincide in the diagonal blocks, while they are opposite in the off-diagonal blocks. These symmetries provide a robust test of the accuracy of both the measurements and the sample azimuthal position.

The measured data were fitted by RCWA simulations [95] formulated in the Mueller–Jones formalism [73]. The profile of the gratings was represented using different models. For the sake of clarity here we discuss only two models. The first model assumed the profile to be trapezoidal. In the second model the grating profile was represented by the superposition of two rectangular lamellas. Both models are sketched in Fig. 2.42. The trapezoidal model depends on three adjustable parameters, the thickness, d, the CD (with) of the lines, and the trapeze angle (SWA). The second model depends on four parameters, the CDs and the thickness of the two lamellas. The resulting best-fitted parameters for both models are presented in Fig. 2.42. In general both models provided fits of same quality, but the most prominent difference among

them was the dependence of the best-fitted parameters with the azimuth angle at which the measurements were taken. Whereas best-fitted parameters corresponding to the model of two lamellas showed a low dependency with the azimuthal angle, CD2 values are dispersed by less than 1.5 nm and the grating depth varies by less than 2 nm around 108 nm, the parameters fitted with the trapezoidal model showed strong fluctuations, 5 nm for the CD and 10 nm for the thickness. The second element that makes the difference between both models is the correlation between fitted parameters. A close look to the values of the CD and the thickness corresponding to the trapezoidal model reveals that them are strongly linearly correlated. This means that the data does not carry the information needed by the model to discriminate the particular influence of each parameter. In contrast, regarding the bi-lamellar model only a small correlation between the overall grating depth and the bottom lamella depth can be observed in this figure. The low amount of correlation and dependency of the fitted parameters with the observation conditions, show that the bi-lamellar model represented better the profile than the trapezoidal model. The adequacy of the model was also confirmed by comparing the obtained profile with AFM measurements. Similar results have been obtained on all the gratings of the sample.

In summary, this example shows that Mueller ellipsometry spectra is a non-destructive technique, with allows accurate metrology of grating profiles. Mueller ellipsometry has the advantage of being faster and cheaper than other tests currently used to in-line quality control in the microelectronics industry.

2.7.3.2 Overlay Characterization by Angle Resolved Mueller Imaging Ellipsometry

The overlay is defined as the misalignment between two layers of a stack. The influence of this error could lead to defective transistors for example if there is no electrical contact between the different constitutive layers. This feature is more and more challenging with the shrinking of the technology node (TN). If this overlay is higher than a set threshold, the whole batch cannot be processed to the new step, this results in a rework, the wafer returns to the previous lithography step and the resist is stripped. Such in the case of grating profile optical metrology, there are several techniques that are considered as a reference for the microelectronic industry. Those techniques include non-optical techniques such AFM or SEM microscopy, and optical techniques based on image analysis (pattern recognition) and on scatterometry. Image analysis also known as Advanced Image Metrology (AIM) is used in this work as reference in order to check the quality of the results obtained by angle resolved Mueller ellipsometry. Angle resolved scatterometry, with a high numerical aperture microscope objective as described above [72, 96], also constitute an interesting scatterometric tool, as it greatly facilitates measurements in extremely tiny targets (less than 5 μm wide), an increasing requirement by semiconductor manufacturers. This would be particularly true for *overlay* (default of positioning of superimposed grids at different layers), a parameter which is becoming increasingly critical and will require in-die dense sampling while current methods involve up to 8 standard (50 μm wide) targets to provide all the relevant information [97]. In the following we provide two

examples, the first one illustrates the use of spectroscopic measurements, and the second one shows the possibilities of the angle-resolved polarimeter.

It has been shown that the choice of proper azimuthal configuration for the measurements with spectroscopic polarimetry is extremely important for the overlay characterization [98, 99]. Given that the angle resolved polarimeter gives an angular signature, it is possible to use the symmetries of the grating to enhance the sensitivity of its angle-resolved signature. The sign of the off-diagonal blocks of the measured Mueller matrix changes when the azimuth φ is changed into $-\varphi$. If the profile is symmetric, the signature is invariant when $\varphi - > \varphi + 180°$ and also for the special case of $\varphi = 90°$ the previous two conditions can only be fulfilled if the off-diagonal blocks are zero. A rupture of symmetry in the structure will violate the above conditions and the off-diagonal blocks will take non-zero values for $\varphi = 90°$. Moreover, given that these blocks change sign upon a mirror symmetry, the information about the sign of the overlay can be unambiguously extracted. In order to highlight the influence of the overlay over the off-diagonal elements of the Mueller matrices the following estimator was defined: $\mathbf{E} = |\mathbf{M}| - |\mathbf{M}|^{T}$ where the superscript t denotes the transposed matrix. The estimator works well either with one-dimensional or two-dimensional gratings and for different types of overlays.

For the sake of clarity we show here a simple example. It consists of the overlay of a one-dimensional grating. As depicted in Fig. 2.43, the overlay is the small shift defined along the direction perpendicular to the lines of the grating. For this particular example, overlay 25 nm, the elements of the estimator matrix $\mathbf{E}$ can reach the value of 0.25 (m14 and m41), i.e. 1/8 of the total scale, which points out the high sensitivity of this estimator.

The fact of being able to measure the full Mueller matrix is advantageous because it provides the "full picture" of the polarization effects. From this full picture, it is possible to optimize the overlay estimation sensitivity for a given type of sample, by selecting a particular matrix element (or a the combination of them) from the whole matrix. However, if a partial Mueller matrix is measured, the evident lack of matrix elements may prevent to evaluate the optimal estimators, needed to properly measure the overlay. For the particular sample discussed here, we choose to work with the estimator element E_{14} for two reasons: The first was that E_{14} showed the strongest values, thus providing the most precise results in terms of signal to noise ratio. The second reason was related to the minimization of the impact of systematic errors in the evaluation of the overlay. For Mueller measurements performed in the Fourier plane at high numerical apertures, the principal source of systematic errors came from the residual birefringence of the microscope objective used to concentrate the light on the sample. Without entering into a detailed description of these errors, it is worth to mention here that it was observed that the Mueller matrix elements which were the less affected by non-idealities of the objective were M_{14} and M_{41}. Therefore the choice of the estimator element E_{14}, which is a combination of M_{14} and M_{41}, was done in order to maximize the accuracy of the measurements. In order to check the linear relation between the values of the estimator $\mathbf{E}$ and the value of the overlay, we compared the maximum value of the element $E_{1,4}$ of the estimator
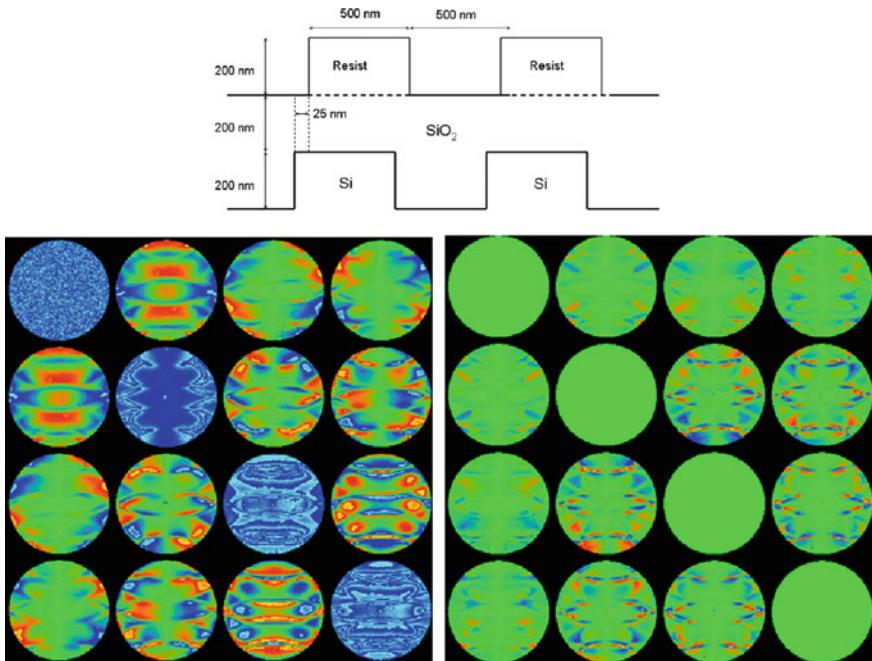
Fig. 2.43 *Top* schematic view of the grating used in the experiences for the overlay characterization. The overlay: 25 nm. *Bottom-Left* experimental angle-resolved Mueller matrix. *Bottom-Right* corresponding estimator matrix **E**. Values taken from [100] with the authorization of the author

matrix with the overlay value obtained by AIM for set of samples. The results are shown in Fig. 2.44.

The fitted linear regression, also shown in Fig. 2.44 is excellent. Moreover, the figure indicates that these results can be extended to negative overlays, the sign of the estimator becoming negative.

In conclusion, it can be said that the overlay errors can be accurately determined provided that we are able to design an apparatus with small enough errors. For this particular case, the magnitude of the systematic errors of the experimental matrices were estimated to be of the order of 1 %, which was associated to an incertitude of about 1nm in the determination of the overlay.

2.7.4 Biomedicine: Cancer Detection and Staging

The field of "optical biopsy", i.e. the diagnosis of the status of tissues by optical means, is currently a very active research area, with many techniques such as optical coherence tomography (OCT), fluorescence imaging, spectral reflectivity, imaging in coherent light (speckles), confocal imaging in vivo and many others. Polarimetric
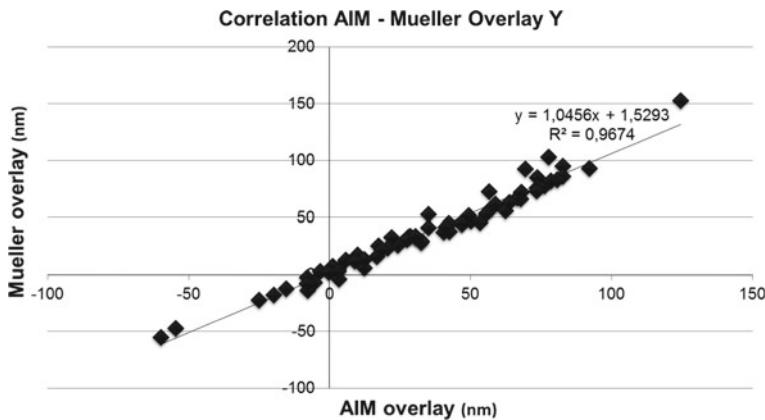
Fig. 2.44 Correlation between measured and AIM Y-overlay for 55 samples with different overlays. Values taken from [100] with the authorization of the author

imaging has received comparatively much less attention, even though it has several potential advantages for a wide use in "real world", among which its intrinsically low cost.

At LPICM we have been exploring this field for several years now, with particular attention to the early detection and staging of uterine cervix cancer. For this disease, the standard practice is as follows:

- the first "alarm" comes with anomalous Pap smear,
- then the patient is examined by a specialist with a long working distance binocular microscope, called colposcope, to give a first diagnosis form the visual aspect of the cervix, with and without staining by diluted acetic acid and iodine. The colposcopist usually takes one to three biopsies to be analyzed in a pathology department.
- If the biopsies are positive for a "precancer" state called dysplasia, the front part of the cervix is removed ("cone biopsy") and examined again by pathologists to pose the final diagnosis.

Now the colposcopic visualization of uterine dysplasias is notoriously very difficult and operator dependent. Any technique able to make this examination easier and more reliable would be welcome, provided it is not too expensive and does not imply too long examination times, which is the case for polarimetric imaging.

During a first clinical trial of a "polarimetric colposcope" which, however, did not provide full Mueller measurements, we realized that the polarimetric response of uterine cervix is much more complex than what we anticipated, and that other samples, where tumors were much easier to see, had to be studied too to get a more sound vision of the origin of the observed polarimetric contrasts. We thus focussed our activity on Mueller ellipsometric imaging of ex vivo samples, mostly colons with cancerous parts, with the instrument presented in Sect. 2.6.5.3, which was installed at the Pathology Department of Institut Mutualiste Montsouris in Paris.
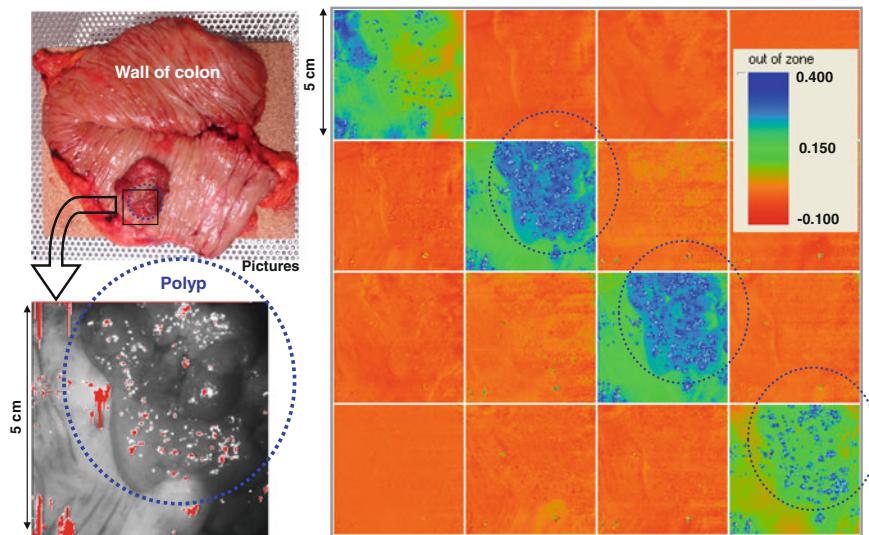
Fig. 2.45 *Left* ordinary (intensity) images of a colon sample with a burgeoning tumor. *Right* full Mueller image of the same sample, with a 5×5 field of view corresponding to the *bottom left* photo

In this contribution we cannot go beyond a short presentation of a few "emblematic" recent images, and a short discussion of the relevance of the technique for "real world" applications in this field.

2.7.4.1 Colon Samples

A first example is shown in Fig. 2.45. A colon sample with a large polyp is shown in ordinary (intensity) imaging and in full polarimetric imaging at 550 nm. The whole tissue appears as a diagonal depolarizer, with

$$M_{22} = M_{33} > M_{44} \qquad (2.122)$$

as expected for a globally isotropic system observed in backscattering. Depolarization is larger for circular than for linear incident polarization, indicating that the contribution of small scatterers is dominant [101]. The key point here is that the tumoral part, at the top right of the images, is clearly *less depolarizing* than the healthy tissue.

Subsequent detailed studies [102] on similar colon samples with tumors at various stages confirmed that this trend is observed essentially at early stages of the disease, when tumoral cells exhibit an exophytic growth over normal tissues. At subsequent stages, the thickness of the tumoral tissue decreases, and the underlying layers are progressively destroyed. Then, the polarimetric responses depend on the structure of the remain tissue. However, by suitable processing of multispectral depolarization images, tumor staging by optical means seems possible.
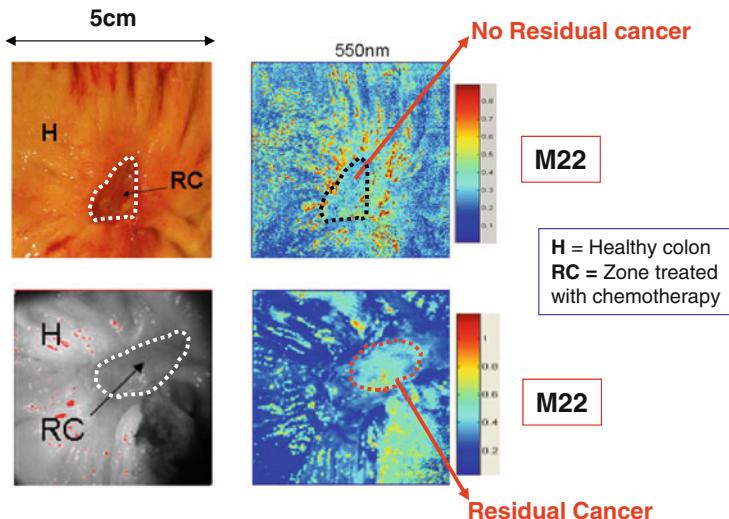
Fig. 2.46 Ordinary (*left*) and linear depolarization (*right*) images of samples of rectum extracted from two different patients after treatment by RC. The "footprint" of the initial tumor is shown on the ordinary images by *white dotted lines*. In the *bottom* images, the tissue visible in the low right corner is the section of the colon wall and should not be considered

Besides the detection and characterization of cancers at early stages, another important issue is the follow-up of treatment by radiochemotherapy (RC). For uterine cervix cancers taken at late evolution stages, RC is very efficient, and in many cases the tumor is totally "burnt" and replaced by fibrotic tissue. If so, no other treatment is actually needed. On the other hand, if some residual tumor is present after RC, it is mandatory to remove it surgically, even though this surgery, on fibrotic tissue, is much more difficult and traumatic than on non-irradiated tissues. In summary it is of paramount importance to know if residual cancer is present or not to take the right decision. Currently residual tumors are searched in CT and PetSCAN images, but none of these techniques is really satisfactory.

We took Mueller images of several rectum samples taken on patients operated after RC. Figure 2.46 shows the results for two such samples, which were subsequently analyzed by pathologists. On the ulcerated regions corresponding to the "footprint" of the initial cancer, M_{22} images reveal a lower depolarization (with respect to surrounding healthy tissue) when a residual tumor is present, while such contrasts are not visible if all the tumor has been "burnt". Of course these results are still preliminary, but they have been considered encouraging enough to start a specific project on this issue for uterine cervix.
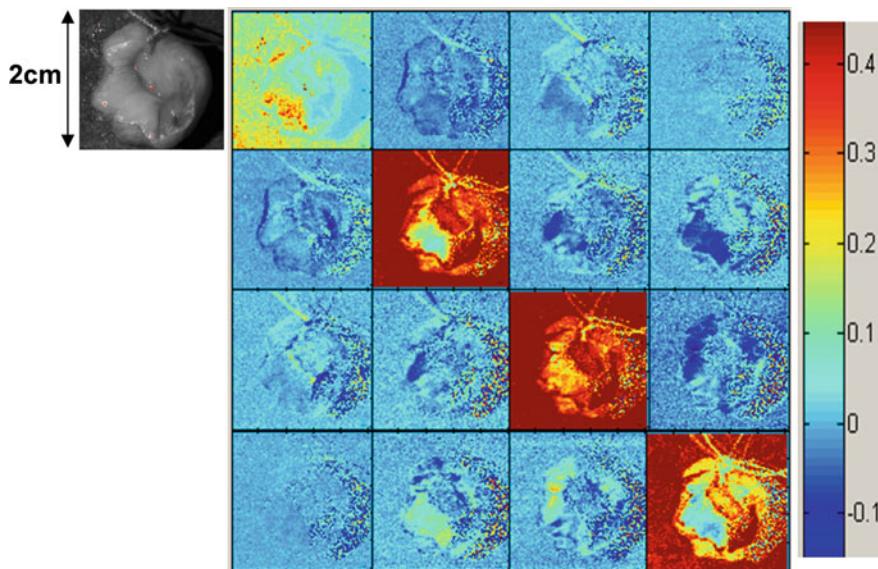
Fig. 2.47 Full Mueller matrix of a uterine cervix cone biopsy

2.7.4.2 Uterine Cervix Cone Biopsies

The colon and rectum samples always show polarimetric responses typical of pure depolarizers. Obviously, to characterize such sample full Mueller polarimetry is not needed and simpler instruments could do the job. However, this is not always so: our clinical trial on incomplete polarimetric imaging clearly showed that in vivo uterine cervix is optically anisotropic, a totally unexpected result for this kind of tissue.

Subsequently, full Mueller matrix imaging of unstained cone biopsies confirmed this very surprising result. An example of such images is shown in Fig. 2.47, where off diagonal terms are clearly quite significant. Again, a Lu–Chipman decomposition has been performed and its result has been compared with the maps of the lesions obtained by a careful (and time consuming !) pathology study of this cone biopsy. The results are shown in Fig. 2.48. The scalar birefringence is a very good marker of healthy regions. In spots where this birefringence disappears, the depolarization power becomes the best indicator of the status of the tissue: as seen previously on colon cancer at early stages, the dysplastic (precancerous) region at the center of the sample is less depolarizing than healthy tissue. Another type of tissue, with an even much lower depolarization, appears at the right bottom corner: this is not really a lesion, and it is very easily distinguished from dysplasia.

These trends have been confirmed with many other cone biopsies. It turns out that the birefringence is a very good marker of healthy regions, while in dysplastic parts the depolarization power seems rather well correlated with the stage of the disease.
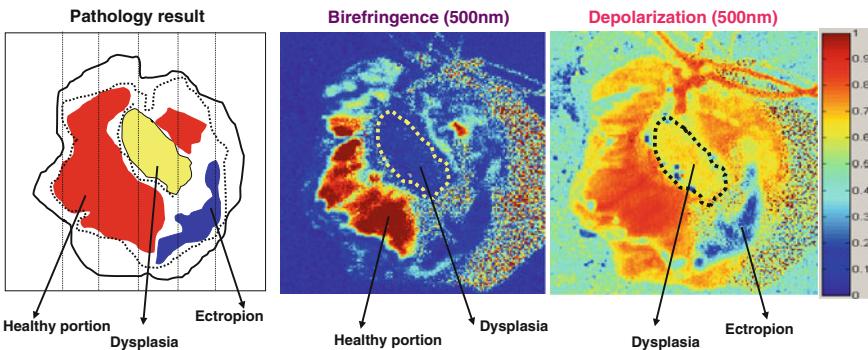
Fig. 2.48 Lu–Chipman decomposition of MM matrix of the polyp shown in Fig. 2.47

2.7.4.3 Summary and Perspectives

The results briefly reported here clearly demonstrate the potential of polarimetric imaging for optical biopsy. Though some tissues, like colon walls, behave as pure depolarizers, in other cases the polarimetric response is too rich to be exploited by incomplete polarimeters. Full Mueller imaging with subsequent image decompositions is mandatory, and can be implemented to be used not only on ex-vivo sample, but also in vivo, for clinical practice. Moreover, to really optimize the reliability of optical biopsy based on Mueller imaging, it will probably be necessary to acquire superimposable images at various wavelengths (for example with a color CCD coupled to FLCs based polarimeter) and then combine data merging and polarimetric analysis. For such an activity to succeed, the motivation of the MDs and the quality of the dialogue between physicists and physicians is crucial.

2.8 Conclusions

In this chapter we reviewed different aspects of Mueller ellipsometric instrumentation and data analysis in relation with standard ellipsometry showing the similarities but also the differences among them. We reviewed the optical formalisms currently used in the framework of standard and Mueller ellipsometry, the Jones and the Stokes formalisms. A special attention has been paid to the description of the fundamental polarization properties. These basic properties are used as "building blocks" to understand the optical response of any sample. We also discussed different ways to decompose general Mueller matrices into simple matrices taking profit of their linear algebraic properties. Mueller matrix decompositions are useful to highlight the physical properties of complex samples, such biological tissues, which cannot be easily modeled. Moreover, matrix decompositions are useful to filter depolarizing Mueller matrices, to remove noise or to isolate non-depolarizing components

which, in turn, can be treated with simple models. Matrix decompositions are powerful tools to characterize samples, but to be applied properly they need to be carried on full Mueller matrices, not on partial matrices. In order to further highlight the advantages of working with full Mueller matrices, we compared the information that can be retrieved by generalized ellipsometry and by Mueller ellipsometry. We showed that whereas for non-depolarizing samples, both, generalized and Mueller ellipsometry provide the same results, for depolarizing samples only Mueller ellipsometry is relevant. In order to disclose the technical possibilities of measuring full Mueller matrices, we reviewed the most frequently used standard ellipsometric configurations showing their characteristics including advantages and limitations. The theory allowing the design, calibration and operation of optimal Mueller polarimeters have been overviewed. We have also shown four examples of complete Mueller ellipsometers developed by us in the past years. The technical description of the Mueller ellipsometers is accompanied by some examples of applications covering topics related to materials science, optical metrology and biomedical imaging. We believe that there is an important potential for the future development of Mueller matrix-based diagnosis tools in medicine and biology. The reason for this is that tissues and living cells show an extremely complex optical behavior, which cannot be sufficiently characterized either by partial polarimetric systems, or by methods neglecting polarization properties of light. We hope to have made clear to the reader the necessity of measuring full Mueller matrices in order to properly study samples showing either depolarization or complex anisotropic response. We also hope to have brought clear and pedagogic information about the general methods and criteria that may be used to build robust and accurate Mueller ellipsometers.

Appendix 1. Mueller Matrices of Some Common Optical Retarders and Diattenuators

In the following we provide specific expressions for Mueller matrices corresponding to different types of general and fundamental optical elements.

Homogeneous Elliptic Diattenuator

The following expression corresponds to the Mueller matrix of a homogeneous elliptic diattenuator oriented with an azimuth angle θ respect to the laboratory axis.

$$\mathbf{M} = \frac{\tau_P}{2} \begin{pmatrix} 1 & C\cos 2\Psi & S\cos 2\Psi \cos(\delta) & S\cos 2\Psi \sin(\delta) \\ C\cos 2\Psi & C^2 + S^2 \sin 2\Psi & CS\cos(\delta)(1 - \sin 2\Psi) & CS\sin(\delta)(1 - \sin 2\Psi) \\ S\cos 2\Psi \cos(\delta) & CS\cos(\delta)(1 - \sin 2\Psi) & (S^2 + C^2 \sin 2\Psi)\cos^2(\delta) + \sin 2\Psi \sin^2(\delta) & S^2 \cos(\delta)\sin(\delta)(1 - \sin 2\Psi) \\ S\cos 2\Psi \sin(\delta) & CS\sin(\delta)(1 - \sin 2\Psi) & S^2 \cos(\delta)\sin(\delta)(1 - \sin 2\Psi) & (S^2 + C^2 \sin 2\Psi)\sin^2(\delta) + \sin 2\Psi \cos^2(\delta) \end{pmatrix}$$

$$(A1.1)$$

In expression (A1.1) the angles Ψ and Δ are the ellipsometric angles with its standard meaning, and the angle δ represents the ellipticity of the eigenvalues of the elliptic retarder.

The following Table A1.1 provides matrices representing some particular cases of the previous general formula corresponding to ideal circular polarizers, dichroic circular polarizers and an ideal elliptical polarizer elliptical oriented at particular angles respect to the laboratory axis.

Table A1.1 provides matrices representing some particular cases of the previous general formula corresponding to ideal linear polarizers and dichroic linear polarizers oriented at particular angles respect to the laboratory axis.

The following expression corresponds to the Mueller matrix of a homogeneous elliptic retarder oriented with an azimuth angle θ respect to the laboratory axis. The retardance provided by the system is Δ. The phase difference between the two linear components needed to build an ellipse is given by the angle φ.

$$
\mathbf{M} = \frac{\tau_P}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & d^2 - e^2 - f^2 + g^2 & 2\,(de + fg) & 2\,(df - eg) \\ 0 & 2\,(de - fg) & -d^2 + e^2 - f^2 + g^2 & 2\,(ef + dg) \\ 0 & 2\,(df + eg) & 2\,(ef - dg) & -d^2 - e^2 + f^2 + g^2 \end{pmatrix}
$$

(A1.2)

$$
d = \cos(2\theta)\sin\left(\frac{\Delta}{2}\right); \quad e = \sin(2\theta)\sin\left(\frac{\Delta}{2}\right)\cos(\varphi);
$$

$$
f = \sin(2\theta)\sin\left(\frac{\Delta}{2}\right)\sin(\varphi); \quad g = \cos\left(\frac{\Delta}{2}\right)
$$

Table A1.2 provides Mueller matrices representing some particular cases of the previous general formula corresponding to circular retarders and linear retarders oriented at particular angles respect to the x–y reference coordinate axis.

Appendix 2. Differential Matrices of Fundamental Polarimetric Properties

In this appendix we provide a detailed expression of the differential matrices corresponding to the eight fundamental polarimetric properties, written according to the 4×4 Stokes formalism and the 2×2 Jones formalism. We also provide the way to deduce them from the original Mueller or Stokes matrices.

Mueller matrix of an homogeneous elliptic retarder

According to [26, 27] the differential matrix **m** of a given Mueller matrix **M** is a 4×4 matrix containing simple expressions of the fundamental polarimetric properties: isotropic refraction, φ isotropic absorption, α, linear birefringence along the coordinate axis x–y, η, linear dichroism along the coordinate x–y, β, linear bire-

Table A1.1 In the table there are summarized the eigenvectors, the Jones matrices, the Mueller matrices and the vector diattenuation for some particular cases of the general expression (A1.1)

Optical element	Eigenvectors	Jones matrix	Mueller matrix	Vector diattenuation
Right circular polarizer ideal	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$ $\frac{1}{\sqrt{2}}\begin{pmatrix}i\\1\end{pmatrix}$	$\sqrt{\tau_p}\begin{pmatrix}1 & -i\\i & 1\end{pmatrix}$	$\frac{\tau_p}{2}\begin{pmatrix}1 & 0 & 0 & 1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\1 & 0 & 0 & 1\end{pmatrix}$	$\begin{pmatrix}0\\0\\1\end{pmatrix}$
Left polarizer ideal	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$ $\frac{1}{\sqrt{2}}\begin{pmatrix}i\\1\end{pmatrix}$	$\sqrt{\tau_p}\begin{pmatrix}1 & i\\-i & 1\end{pmatrix}$	$\frac{\tau_p}{2}\begin{pmatrix}1 & 0 & 0 & -1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\-1 & 0 & 0 & 1\end{pmatrix}$	$\begin{pmatrix}0\\0\\-1\end{pmatrix}$
Dichroic circular polarizer	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$	$z\begin{pmatrix}1 & -i\cos(2\Psi)\\i\cos(2\Psi) & 1\end{pmatrix}$ $z=\frac{\sqrt{\tau_{max}}+\sqrt{\tau_{min}}}{2}$	$\frac{q_1}{2}\begin{pmatrix}1 & 0 & 0 & q_2\\0 & q_3 & 0 & 0\\0 & 0 & q_3 & 0\\q_2 & 0 & 0 & 1\end{pmatrix}$ $q_1=\tau_{max}+\tau_{min}$; $q_2=\cos(2\Psi)$; $q_3=\sin(2\Psi)$	$q_2\begin{pmatrix}0\\0\\-1\end{pmatrix}$

(continued)

Table A1.1 (continued)

Optical element	Eigenvectors	Jones matrix	Mueller matrix	Vector diattenuation
Dichroic elliptical polarizer	$\begin{pmatrix} \cos(\theta) \\ \sin(\theta)\,e^{i\varphi} \end{pmatrix}$	$\dfrac{\sqrt{\tau_{max}}}{2}\begin{pmatrix} c^2 & cse^{-i\varphi} \\ cse^{i\varphi} & s^2 \end{pmatrix}$ $c = \cos(\theta);\ s = \sin(\theta)$	$\dfrac{\tau_{max}}{2}\begin{pmatrix} 1 & C & SX & SY \\ C & C^2 & CSX & CSY \\ SX & CSX & (SX)^2 & SXSY \\ SY & CSY & SXSY & (SY)^2 \end{pmatrix}$ $C = \cos(2\theta)\ ;\ SX = \sin(2\theta)\cos(\varphi)$ $SY = \sin(2\theta)\sin(\varphi)$	$\begin{pmatrix} C \\ SX \\ SY \end{pmatrix}$
Orientation θ	$\begin{pmatrix} -\sin(\theta)\,e^{-i\varphi} \\ \cos(\theta) \end{pmatrix}$			

The cases correspond to elliptical ideal and elliptical dichroic polarizers oriented at particular angles, θ, respect to the x–y reference coordinate frame. The selected cases are representative of elements commonly used in practical optical settings. In this table the values of τ_{max} and τ_{min} correspond to the maximum and minimum transmission of the dichroic polarizer. For an ideal polarizer τ_{min} is null and τ_{max} is equal to the total transmission of the element τ_P. The angle φ is the phase difference between the two linear orthogonal components needed to build an elliptic state. The symbols C, SX, SY and q_2 used in the expressions of the vector diattenuation are the same that these used in the corresponding Mueller matrices

Table A1.2 List of the eigenvectors, the Jones matrices, the Mueller matrices and the vector diattenuation for some particular cases of the general expression (A1.1)

Optical element	Eigenvectors	Jones matrix	Mueller matrix	Vector diattenuation
Linear polarizer Orientation $\theta = 0°$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\sqrt{\tau_p}\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\frac{\tau_p}{2}\begin{pmatrix}1&1&0&0\\1&1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$ $C = \cos(2\theta) = 1\ ;\ S = \sin(2\theta) = 0$	$\begin{pmatrix}1\\0\\0\end{pmatrix}$
Linear polarizer Orientation $\theta = 90°$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\sqrt{\tau_p}\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\frac{\tau_p}{2}\begin{pmatrix}1&-1&0&0\\-1&1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$ $C = \cos(2\theta) = -1\ ;\ S = \sin(2\theta) = 0$	$\begin{pmatrix}-1\\0\\0\end{pmatrix}$
Linear polarizer Orientation $\theta = 45°$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$	$\frac{\sqrt{\tau_p}}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix}$	$\frac{\tau_p}{2}\begin{pmatrix}1&0&1&0\\0&0&0&0\\1&0&1&0\\0&0&0&0\end{pmatrix}$ $C = \cos(2\theta) = 0\ ;\ S = \sin(2\theta) = 1$	$\begin{pmatrix}-1\\0\\0\end{pmatrix}$

(continued)

Table A1.2 (continued)

Optical element	Eigenvectors	Jones matrix	Mueller matrix	Vector diattenuation
Linear polarizer	$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$	$\sqrt{\tau_p}\begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix}$ $c = \cos(\theta)\ ;\ s = \sin(\theta)$	$\dfrac{\tau_p}{2}\begin{pmatrix} 1 & C & S & 0 \\ C & C^2 & CS & 0 \\ S & CS & S^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} C \\ S \\ 0 \end{pmatrix}$
Orientation θ	$\begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$		$C = \cos(2\theta)\ ;\ S = \sin(2\theta)$	
Dichroic linear polarizer	$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$	$\begin{pmatrix} \sqrt{\tau_{max}}\,c^2 & zcs \\ zcs & \sqrt{\tau_{min}}\,s^2 \end{pmatrix}$ $c = \cos(\theta)\ ;\ s = \sin(\theta)\ ;$ $z = \sqrt{\tau_{max}} - \sqrt{\tau_{min}}$	$\dfrac{q_1}{2}\begin{pmatrix} 1 & Cq_2 & Sq_2 & 0 \\ Cq_2 & C^2 + S^2 q_3 & CS\,(1 - q_3) & 0 \\ Sq_2 & CS\,(1 - q_3) & S^2 + C^2 q_3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ $q_1 = \tau_{max} + \tau_{min};\ q_2 = \cos(2\Psi)$ $q_3 = \sin(2\Psi);\ C = \cos(2\theta)\ ;\ S = \sin(2\theta)$	$q_2\begin{pmatrix} C \\ S \\ 0 \end{pmatrix}$
Orientation θ	$\begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$			

These cases correspond to linear ideal and linear dichroic polarizers oriented at particular angles, θ, respect to the x–y reference coordinate frame. They are selected because they represent the optical behavior of optical elements commonly encountered in practical situations. The values of τ_{max} and τ_{min} correspond to the maximum and minimum transmission of the dichroic polarizer. For an ideal polarizer τ_{min} is null and τ_{max} is equal to the total transmission of the element τ_p. The symbols C, S and q_2 used in the expressions of the vector diattenuation are the same that these used in the corresponding Mueller matrices

Table A2.1 This table summarizes the eigenvalues, the Jones matrices, the Mueller matrices and the vector retardances corresponding to some particular cases of the general expression (A1.2)

Optical element	Eigenvectors	Jones matrix	Mueller matrix	Vector diattenuation R
Linear birefringent	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} e^{i\frac{\Delta}{2}} & 0 \\ 0 & e^{-i\frac{\Delta}{2}} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\Delta) & \sin(\Delta) \\ 0 & 0 & -\sin(\Delta) & \cos(\Delta) \end{pmatrix}$	$\Delta \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$
Retarder	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} \cos\left(\frac{\Delta}{2}\right) & i\sin\left(\frac{\Delta}{2}\right) \\ i\sin\left(\frac{\Delta}{2}\right) & \cos\left(\frac{\Delta}{2}\right) \end{pmatrix}$		
Linear birefringent	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$		$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\Delta) & 0 & -\sin(\Delta) \\ 0 & 0 & 1 & 0 \\ 0 & \sin(\Delta) & 0 & \cos(\Delta) \end{pmatrix}$	$\Delta \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$
Retarder orientation $\theta = 45°$	$\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$			
Linear birefringent	$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$	$\begin{pmatrix} c^2 ep + s^2 em & i\sin\frac{\Delta}{2}\sin 2\theta \\ i\sin\frac{\Delta}{2}\sin 2\theta & s^2 ep + c^2 em \end{pmatrix}$ $c = \cos(\theta);\ s = \sin(\theta);$ $ep = e^{i\Delta/2};\ em = e^{-i\Delta/2}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & C^2 + S^2 X & CS\,(1-X) & -SY \\ 0 & CS\,(1-X) & S^2 + C^2 X & CY \\ 0 & SY & -CY & X \end{pmatrix}$ $C = \cos(2\theta)\ ;\ S = \sin(2\theta)$ $X = \cos(\Delta)\ Y = \sin(\Delta)$	$\Delta \begin{pmatrix} C \\ S \\ 0 \end{pmatrix}$
Retarder orientation θ	$\begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$			

(continued)

Table A2.1 (continued)

Optical element	Eigenvectors	Jones matrix	Mueller matrix	Vector diattenuation R
Circular birefringent	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$	$\begin{pmatrix}\cos\left(\frac{\Delta}{2}\right) & \sin\left(\frac{\Delta}{2}\right)\\ -\sin\left(\frac{\Delta}{2}\right) & \cos\left(\frac{\Delta}{2}\right)\end{pmatrix}$	$\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & \cos(\Delta) & \sin(\Delta) & 0\\ 0 & -\sin(\Delta) & \cos(\Delta) & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$	$\Delta\begin{pmatrix}0\\0\\1\end{pmatrix}$
Retarder right	$\frac{1}{\sqrt{2}}\begin{pmatrix}i\\1\end{pmatrix}$			
Circular birefringent	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$	$\begin{pmatrix}\cos\left(\frac{\Delta}{2}\right) & -\sin\left(\frac{\Delta}{2}\right)\\ \sin\left(\frac{\Delta}{2}\right) & \cos\left(\frac{\Delta}{2}\right)\end{pmatrix}$	$\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & \cos(\Delta) & -\sin(\Delta) & 0\\ 0 & \sin(\Delta) & \cos(\Delta) & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$	$\Delta\begin{pmatrix}0\\0\\-1\end{pmatrix}$
Retarder left	$\frac{1}{\sqrt{2}}\begin{pmatrix}i\\1\end{pmatrix}$			

The listed cases correspond to circular retarders and linear retarders oriented at a particular angle, θ, respect to the x–y reference coordinate axis. The symbols C, S and Δ used in the expressions of the vector retardances are the same that these used in the corresponding Mueller matrices

fringence along the bisectors to the coordinate axis x–y, η, linear dichroism along the bisectors to the coordinate axis x–y, γ, circular birefringence, μ, and circular dichroism, δ. The bisectors to the coordinate axis x–y form a coordinate axes rotated 45° respect to the x–y. A particular choice of the x–y axis well adapted for ellipsometric measurements in reflection, or in transmission with tilted samples, is the p-s axis defined respect to the plane of incidence. When light propagates along the z direction in an anisotropic medium, which is considered as homogeneous in the x, y directions, the transformation of the Stokes vector at a given position z, S(z) to a Stokes vector at a given position S(z + Δz) can be described by the Mueller matrix $\mathbf{M_{z,\Delta z}}$. The transformation can be written according to the following expression:

$$\mathbf{S}(z + \Delta z) = \mathbf{M_{z,\Delta z}}\mathbf{S}(z) \tag{A2.1}$$

Subtraction of S(z) from both sides of expression (A2.1) leads to:

$$\mathbf{S}(z + \Delta z) - \mathbf{S}(z) = \mathbf{S}(z)\left(\mathbf{M_{z,\Delta z}} - \mathbf{I}\right) \tag{A2.2}$$

where **I** is the identity matrix. Clearly, if the latter expression is divided by Δz and then the it is extrapolated to the limiting case of Δz → 0, then it is possible to obtain the following expression relating the transformation of the Stokes vector:

$$\frac{d\mathbf{S}}{dz} = \lim_{\Delta z \to 0} \frac{\left(\mathbf{M_{z,\Delta z}} - \mathbf{I}\right)}{\Delta z}\mathbf{S}(z) = \mathbf{m}\mathbf{S}(z) \tag{A2.3}$$

The latter equation is the definition of the matrix **m** in the 4×4 Stokes formalism. The derivation of expression (A2.3) is valid for either a non-depolarizing or a depolarizing medium. The matrix **m** is the expression of the effect of the different optical properties of the medium on the Stokes vector when light travels a differential distance Δz. For this reason the matrix **m** is called the differential propagation matrix, or simply the differential matrix.

The relation of (A2.3) with the expression (2.52) given in the text is easily found. The transformation of an initial Stokes vector **S**(0) of a beam traveling a distance z inside a medium can be written as:

$$\mathbf{S}(z) = \mathbf{M_z}\mathbf{S}(0) \tag{A2.4}$$

After differentiating the previous expression respect to z one gets:

$$\frac{d\mathbf{S}}{dz} = \frac{d\mathbf{M_z}}{dz}\mathbf{S}(0) \tag{A2.5}$$

According to (A2.3) the latter expression can be rewritten as:

$$\frac{d\mathbf{S}}{dz} = \mathbf{m}\mathbf{S}(z) = \mathbf{m}\mathbf{M_z}\mathbf{S}(0) = \frac{d\mathbf{M_z}}{dz}\mathbf{S}(0) \tag{A2.6}$$

Table A2.2 Differential matrices m and n corresponding to the eight fundamental polarimetric properties

Optical property	Differential **m**	Differential **n**
Isotropic refraction	$\begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$	$\dfrac{\varphi}{2}\begin{pmatrix} i&0 \\ 0&i \end{pmatrix}$
Isotropic absorption	$\begin{pmatrix} \alpha&0&0&0 \\ 0&\alpha&0&0 \\ 0&0&\alpha&0 \\ 0&0&0&\alpha \end{pmatrix}$	$\dfrac{\alpha}{2}\begin{pmatrix} 1&0 \\ 0&1 \end{pmatrix}$
Linear birefringence along x–y	$\begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&\eta \\ 0&0&-\eta&0 \end{pmatrix}$	$\dfrac{\eta}{2}\begin{pmatrix} i&0 \\ 0&-i \end{pmatrix}$
Linear dichroism along x–y	$\begin{pmatrix} 0&\beta&0&0 \\ \beta&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$	$\dfrac{\beta}{2}\begin{pmatrix} 1&0 \\ 0&-1 \end{pmatrix}$
Linear birefringence along bisectors of x–y	$\begin{pmatrix} 0&0&0&0 \\ 0&0&0&\nu \\ 0&0&0&0 \\ 0&-\nu&0&0 \end{pmatrix}$	$\dfrac{-\nu}{2}\begin{pmatrix} 0&i \\ i&0 \end{pmatrix}$
Linear dichroism along bisectors of x–y	$\begin{pmatrix} 0&0&\gamma&0 \\ 0&0&0&0 \\ \gamma&0&0&0 \\ 0&0&0&0 \end{pmatrix}$	$\dfrac{\gamma}{2}\begin{pmatrix} 0&1 \\ 1&0 \end{pmatrix}$
Circular birefringence (right)	$\begin{pmatrix} 0&0&0&0 \\ 0&0&\mu&0 \\ 0&-\mu&0&0 \\ 0&0&0&0 \end{pmatrix}$	$-\dfrac{\mu}{2}\begin{pmatrix} 0&-1 \\ 1&0 \end{pmatrix}$
Circular dichroism (right)	$\begin{pmatrix} 0&0&0&\delta \\ 0&0&0&0 \\ 0&0&0&0 \\ \delta&0&0&0 \end{pmatrix}$	$\dfrac{\delta}{2}\begin{pmatrix} 0&-i \\ i&0 \end{pmatrix}$

Thus giving an analogous definition of the matrix **m** which is identical to (2.52):

$$\mathbf{M_z}^{-1}\frac{d\mathbf{M_z}}{dz} = \mathbf{m} \tag{A2.7}$$

An analogous procedure may be performed using Jones vectors instead of Stokes vectors leading to the derivation of a 2×2 matrix called **n**. Obviously the applicability of matrix **n** is restricted to non-depolarizing media.

As an illustration, of the method to obtain the expression of the differential matrices **m** and **n**, let's consider their detailed derivation for the particular case of a homogeneous medium along the propagation direction, of total thickness z, showing linear birefringence characterized by the parameter η. The respective Mueller and Jones matrices associated to this medium are:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\eta z) & \sin(\eta z) \\ 0 & 0 & -\sin(\eta z) & \cos(\eta z) \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} e^{i\frac{\eta}{2}z} & 0 \\ 0 & e^{-i\frac{\eta}{2}z} \end{pmatrix} \quad \text{where} \quad \eta = \frac{4\pi\,birr}{\lambda}$$

(A2.8)

The total retardance created by the medium, commonly expressed by the ellipsometric angle Δ is given by η times the thickness z. The parameter η is the intensive retardance, also called the differential retardance, whereas Δ is the extensive retardance, which is proportional to the path that light has travelled inside the medium. The differential retardance depends of the wavelength of light, λ, and the birefringence of the medium, *birr*.

For a thin section of thickness Δz of the material, the corresponding Jones and Mueller matrices can be calculated expanding the terms of matrices in (A2.8) in a Taylor series expansion respect to z and retaining only the first order terms.

$$\mathbf{M_{\Delta z}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \eta z \\ 0 & 0 & -\eta z & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{J_{\Delta z}} = \begin{pmatrix} 1 + \frac{i\eta z}{2} & 0 \\ 0 & 1 - \frac{i\eta z}{2} \end{pmatrix}$$

(A2.9)

Then according to (A2.3) after subtraction of the respective (2×2) and (4×4) unit matrices and division by z, one gets:

$$\mathbf{m} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta \\ 0 & 0 & -\eta & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{n} = \frac{\eta}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

(A2.10)

The matrices (4×4) and (2×2) **m** and **n** matrices grouped in Table A2.2 the following table are the result of applying the same procedure to each one of the eight fundamental properties:

References

1. R.M.A Azzam, N.M. Bashara, Ellipsometry and Polarized Light (Elsevier, Amsterdam, 1987)
2. H.G. Tompkins, W.A. McGahan, *Spectroscopic Ellipsometry and Reflectometry A User's Guide* (Wiley, New York, 1999)

3. H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (Wiley, West Sussex, 2007)
4. M. Born, E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, 2005)
5. S. Huard, *The Polarization of Light* (Wiley, New York, 1997)
6. R.A. Chipman, Polarimetry, in: Handbook of Optics, vol 2, chap 22, 2nd ed. M. Bass ed. (McGraw Hill, New York, 1995)
7. D. Goldstein, *Polarized Light*, 2nd edn. (Decker, New York, 2003)
8. J.J. Gil, Characteristic properties of Mueller matrices. J. Opt. Soc. Am. **17**, 328–334 (2000). doi:10.1364/JOSAA.17.000328
9. S.R. Cloude, Group theory and polarisation algebra. Optik **75**, 26 (1986). doi:10.1364/JOSAA.18.003130
10. D.G.M. Anderson, R. Barakat, Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix. J. Opt. Soc. Am. A **11**, 2305 (1994). doi:10.1364/JOSAA.11.002305
11. J.J. Gil, E. Bernabeu, A depolarization criterion in Mueller matrices. Opt. Acta **32**, 259 (1985). doi:10.1080/713821732
12. S.Y. Lu, R.A. Chipman, Interpretation of Mueller matrices based on polar decomposition. J. Opt. Soc. Am. A **13**, 1106–1113 (1996). doi:10.1364/JOSAA.13.001106
13. R. Ossikovski, Alternative depolarization criteria for Mueller matrices. J. Opt. Soc. Am. A **27**, 808–814 (2010). doi:10.1364/JOSAA.27.000808
14. R. Espinosa-Luna, G. Atondo-Rubio, E. Bernabeu, S. Hinijosa-Ruiz, Dealing depolarization of light in Mueller matrices with scalar metrics Â'. Optik **121**, 1058–1068 (2010). doi:10.1016/j.ijleo.2008.12.030
15. S.R. Cloude, E. Pottier, IEEE Trans. GRS **34**, 498 (1996)
16. S.R. Cloude, Conditions for the physical realizability of matrix operators in polarimetry. Proc. SPIE **1166**, 177–185 (1989)
17. F. Le Roy-Bréhonnet, B. Le Jeune, Utilization of Mueller matrix formalism to obtain optical targets depolarization and polarization properties. Prog. Quantum Electron. **21**, 109–151 (1997). doi:10.1016/S0079-6727(97)84687-3
18. J. Morio, F. Goudail, Influence of the order of diattenuator, retarder, and polarizer in polar decomposition of Mueller matrices. Optics. Lett. **29**, 2234–2236 (2004). doi:10.1364/OL.29.002234
19. R. Ossikovski, A. De Martino, S. Guyot, Forward and reverse product decompositions of depolarizing Mueller matrices. Opt. Lett. **32**, 689 (2007)
20. R. Ossikovski, Interpretation of nondepolarizing Mueller matrices based on singular-value decomposition. J. Opt. Soc. Am. A **25**, 473–482 (2008). doi:JOSAA.25.000473
21. R. Ossikovski, E. Garcia-Caurel, A. De Martino, Product decompositions of experimentally determined non-depolarizing Mueller matrices. Physica status solidi C (2008). doi:10.1002/pssc.200777794
22. R. Ossikovski, Analysis of depolarizing Mueller matrices through a symmetric decomposition. J. Opt. Soc. Am. A **26**, 1109–1118 (2009)
23. C. Fallet, A. Pierangelo, R. Ossikovski, A. De Martino, Experimental validation of the symmetric decomposition of Mueller matrices. Opt. Express. **18**, 2832 (2009). doi:10.1364/OE.18.000831
24. R. Ossikovski, C. Fallet, A. Pierangelo, A. De Martino, Experimental implementation and properties of Stokes nondiagonalizable depolarizing Mueller matrices. Opt. Lett. **34**, 974 (2009)
25. R. Ossikovski, M. Foldyna, C. Fallet, A. De Martino, Experimental evidence for naturally occurring nondiagonal depolarizers. Opt. Lett. **34**, 2426–2428 (2009). doi:10.1364/OL.34.002426
26. R. Ossikovski, Differential matrix formalism for depolarizing anisotropic media. Opt. Lett. **36**, 2330–2332 (2011). doi:10.1364/OL.36.002330
27. R.M.A. Azzam, Propagation of partially polarized light through anisotropic media with or without depolarization: a differential 4×4 matrix calculus. J. Opt. Soc. Am. **68**, 1756–1767 (1978). doi:10.1364/JOSA.68.001756

28. M. Anastasiadou, S. Ben-Hatit, R. Ossikovski, S. Guyot, A. De Martino, Experimental validation of the reverse polar decomposition of depolarizing Mueller matrices. J. Eur. Opt. Soc. Rapid Publ. **2**, 1–7 (2007). doi:10.2971/jeos.2007.07018
29. M. Gaillet, D. Cattelan, G. Bruno, M. Losurdo, Roadmap on industrial needs in ellipsometry and specifications for the next generation of ellipsometry and polarimetry. NanocharM Report 2009, www.nanocharm.org
30. G.E. Jellison Jr., Data analysis for spectroscopic ellipsometry. Thin Solid Films **234**, 416–422 (1993). doi:10.1016/0040-6090(93)90298-4
31. G.E. Jellison Jr., The calculation of thin film parameters from spectroscopic ellipsometry data. Thin Solid Films **290–291**, 40–45 (1996). doi:0.1016/S0040-6090(96)09009-8
32. P.S. Hauge, Recent developments in instrumentation in ellipsometry. Surf. Sci. **96**, 108–140 (1980). doi:10.1016/0039-6028(80)90297-6
33. D. Thomson, B. Johs, Infrared ellipsometer/polarimeter system, method of calibration and use thereof, US patent No: US5706212, (1998)
34. A. Laskarakis, S. Logothetidis, E. Pavlopoulou, M. Gioti, Mueller matrix spectroscopic ellipsometry: formulation and application. Thin Solid Films **455–456**, 43–49 (2004). doi:10.1016/j.tsf2003.11.197
35. E. Compain, S. Poirier, B. Drévillon, General and self-consistent method for the calibration of polarization modulators, polarimeters, and Mueller-matrix ellipsometers. Appl. Opt. **38**, 3490–3502 (1999). doi:10.1364/AO.38.003490
36. D.S. Sabatke, A.M. Locke, M.R. Descour, W.C. Sweatt, Figures of merit for complete Stokes polarimeters. Proc. SPIE **4133**, 75–81 (2000). doi:10.1117/12.406613
37. J.S. Tyo, Noise equalization in Stokes parameter images obtained by use of variable-retardance polarimeters. Opt. Lett. **25**, 1198–2000 (2000). doi:10.1364/OL.25.001198
38. J.S. Tyo, Design of optimal polarimeters : maximization of signal-to-noise ratio and minimization of systematic error. Appl. Opt. **41**, 619–630 (2002). doi:10.1364/AO.41.000619
39. D.S. Sabatke, M.R. Descour, E.L. Dereniak, W.C. Sweatt, S.A. Kemme, G.S. Phipps, Optimization of retardance for a complete Stokes polarimeter. Opt. Lett. **25**, 802 (2000). doi:10.1364/OL.25.000802
40. J. Zallat, S. Aïnouz, M.P. Stoll, Optimal configurations for imaging polarimeters: impact of image noise and systematic errors. J. Opt. A Pure Appl. Opt. **8**, 807 (2006). doi:10.1088/1464-4258/8/9/015
41. M.H. Smith, Optimisation of a dual-rotating-retarder Mueller matrix polarimeter. Appl. Opt. **41**, 2488 (2002). doi:10.1364/AO.41.002488
42. R.W. Collins, J. Koh, Dual rotating-compensator multichannel ellipsometer: instrument design for real-time Mueller matrix spectroscopy of surfaces and films. JOSA A. **16**, 1997–2006 (1999). doi:10.1364/JOSAA.16.001997
43. E. Compain, B. Drévillon, High-frequency modulation of the four states of polarization of light with a single phase modulator. Rev. Sci. Instrum. **69**, 1574 (1998). doi:/10.1063/1.1148811
44. E. Compain, B. Drévillon, Broadband division-of-amplitude polarimeter based on uncoated prisms. Appl. Opt. **37**, 5938 (1998). doi:10.1364/AO.37.005938
45. E. Compain, B. Drévillon, J. Huc, J.Y. Parey, J.E. Bouree, Complete Mueller matrix measurement with a single high frequency modulation. Thin Solid Films **313–314**, 47–52 (1998). doi:10.1016/S0040-6090(97)00767-0
46. D. Lara, C. Dainty, Double-pass axially resolved confocal Mueller matrix imaging polarimetry. Opt. Lett. **30**, 2879–2881 (2005). doi:10.1364/OL.30.002879
47. See for instance the official website of Woollam Co. www.jawoollam.com
48. G.E. Jellison, F.A. Modine, Two-modulator generalized ellipsometry: experiment and callibration. Appl. Opt. **36**, 8184–8189 (1997). doi:10.1364/AO.36.008184
49. O. Arteaga, J. Freudenthal, B. Wang, B. Kahr, Mueller matrix polarimetry with four photoelastic modulators: theory and calibration, to be published in, Applied Optics (2012)
50. A.E. Oxley, On apparatus for the production of circularly polarized light. Philos. Mag. **21**, 517–532 (1911). doi:10.1080/14786440408637058
51. See for instance the official website of Meadowlark Optics, www.meadowlark.com

52. J. Ladstein, F. Stabo-Eeg, E. Garcia-Caurel, M. Kildemo, Fast near-infra-red spectroscopic Mueller matrix ellipsometer based on ferroelectric liquid crystal retarders. Physica Status Solidi C, Special Issue: 4th International Conference on Spectroscopic Ellipsometry, 5, n/a, doi:10.1002/pssc.200890005

53. P.A. Letnes, I.S. Nerbo, L.M.S. Ass, P.G. Ellingsen, M. Kildemo, Fast and optimal broadband Stokes/Mueller polarimeter design by the use of a genetic algorithm. Opt. Express **18**, 23095–23103 (2010). doi:10.1364/OE.18.023095

54. D. Cattelan, E. Garcia-Caurel, A. De Martino, B. Drevillon, Device and method for taking spectroscopic polarimetric measurements in the visible and near-infrared ranges. Patent application 2937732, Publication number: US 2011/0205539 A1

55. T. Wagner, J.N. Hilfiker, T.E. Tiwald, C.L. Bungay, S. Zollner, Materials characterization in the vacuum ultraviolet with variable angle spectroscopic ellipsometry. Physica Status Solidi A **188**, 1553–1562 (2001). doi:10.1002/1521-396X(200112)188:4<1553:AID-PSSA1553>3.0.CO;2-A

56. A. Zuber, N. Kaiser, J.L. Stehlé, Variable-angle spectroscopic ellipsometry for deep UV characterization of dielectric coatings. Thin Solid Films **261**, 37–43 (1995). doi:10.1016/S0040-6090(94)06492-X

57. E. Garcia-Caurel, J.L. Moncel, F. Bos, B. Drévillon, Ultraviolet phase-modulated ellipsometer. Revi. Sci. Instrum. **73**, 4307–4312 (2002). doi:10.1063/1.1518788

58. D.H. Goldstein, Mueller matrix dual-rotating retarder polarimeter. Appl. Opt. **31**, 6676–6683 (1992). doi:10.1364/AO.31.006676

59. L.L. Deibler, M.H. Smith, Measurement of the complex refractive index of isotropic materials with Mueller matrix polarimetry. Appl. Opt. **40**, 3659–3667 (2001). doi:10.1364/AO.40.003659

60. A. Röseler, *Infrared Spectroscopic Ellipsometry* (Akademie-Verlag, Berlin, 1990)

61. E.H. Korte, A. Röseler, Infrared spectroscopic ellipsometry: a tool for characterizing nanometer layers. Analyst **123**, 647–651 (1998)

62. J.N. Hilfiker, C.L. Bungay, R.A. Synowicki, T.E. Tiwald, C.M. Herzinger, B. Johs, G.K. Pribil, J.A. Woollam, Progress in spectroscopic ellipsometry: applications from vacuum ultraviolet to infrared. J. Vac. Sci. Technol. A **21**, 1103 (2003). doi:10.1116/1.1569928

63. E. Gilli, M. Kornschober, R. Schennach, Optical arrangement and proof of concept prototype for mid infrared variable angle spectroscopic ellipsometry. Infrared Phys. Technol. **55**, 84–92 (2012). doi:10.1016/j.infrared.2011.09.006

64. T.D. Kang, E. Standard, G.L. Carr, T. Zhou, M. Kotelyanskii, A.A. Sirenko, Rotatable broadband retarders for far-infrared spectroscopic ellipsometry. Thin Solid Films **519**, 2698–2702 (2011). doi:10.1016/j.tsf.2010.12.057

65. C. Bernhar, J. Humlıcek, B. Keimer, Far-infrared ellipsometry using a synchrotron light source the dielectric response of the cuprate high Tc superconductors. Thin Solid Films **455–456**, 143–149 (2004). doi:10.1016/j.tsf.2004.01.002

66. T. Hofmann, C.M. Herzinger, A. Boosalis, T.E. Tiwald, J.A. Woollam, M. Schubert, Variable-wavelength frequency-domain terahertz ellipsometry. Rev. Sci. Instrum. **81**, 023101 (2010). doi:10.1063/1.3297902

67. J.M. Bennett, A critical evaluation of rhomb-type quarterwave retarders. Appl. Opt. **9**, 2123–2129 (1970). doi:10.1364/AO.9.002123

68. E. Garcia-Caurel, A. de Martino, B. Drévillon, Spectroscopic Mueller polarimeter based on liquid crystal devices. Thin Solid films **455–456**, 120 (2003). doi:10.1016/j.tsf.2003.12.056

69. B. Drévillon, A. De Martino, Liquid crystal based polarimetric system, a process for the calibration of this polarimetric system, and a polarimetric measurement process. Patent number US 7,196,792 (filing date 2003)

70. T. Scharf, *Polarized Light in Liquid Crystals and Polymers* (Willey, New Jersey, 2007)

71. A. De Martino, Y-K. Kim, E. Garcia-Caurel, B. Laude, B. Drévillon, Optimized Mueller polarimeter with liquid crystals. Opt. Lett. **28**, 619–618 (2003) doi:10.1364/OL.28.000616

72. S. Ben Hatit, M. Foldyna, A. De Martino, B. Drévillon, Angle-resolved Mueller polarimeter using a microscope objective. Phys. Stat. Sol. (a) **205**, 743 (2008). doi:10.1002/pssa.200777806

73. A. De Martino, S. Ben Hatit, M. Foldyna, Mueller polarimetry in the back focal plane. Proc. SPIE **6518**, 65180X (2007). doi:10.1117/12.708627

74. A. De Martino, E. Garcia-Caurel, B. Laude, B. Drévillon, General methods for optimized design and calibration of Mueller polarimeters. Thin Solid Films **455**, 112–119 (2004). doi:10.1016/j.tsf.2003.12.052

75. N.A. Beaudry, Y. Zhao, R.A. Chipman, Dielectric tensor measurement from a single Mueller matrix image. J. Opt. Soc. Am. A **24**, 814 (2007). doi:10.1364/JOSAA.24.000814

76. A. Lizana, M. Foldyna, M. Stchakovsky, B. Georges, D. Nicolas, E. Garcia-Caurel, Enhanced sensitivity to dielectric function and thickness of absorbing thin films by combining Total Internal Reflection Ellipsometry with Standard Ellipsometry and Reflectometry, to appear in Journal of Physics D. Applied Physics

77. G.E. Jellison, F.A. Modine, Parameterization of the optical functions of amorphous materials in the interband region. Appl. Phys. Lett. **69**, 371–373 (1996). doi:10.1063/1.118064

78. R.A. Synowicki, B.D. Johs, A.C. Martin, Optical properties of soda-lime float glass from spectroscopic ellipsometry. Thin Solid Films **519**, 2907–2913 (2011). doi:10.1016/j.tsf.2010.12.110

79. M. Philipp, M. Knupfer, B. Büchner, H. Gerardin, Plasmonic excitations in ZnO/Ag/ZnO multilayer systems: insight into interface and bulk electronic properties. J. Appl. Phys. **109**, 063710–063716 (2011). doi:10.1063/1.3565047

80. H. Raether, *Surface Plasmons on Smoth and Rough Surfaces and on Gratings* (Springer, Berlin, 1988)

81. F. Abelès, Surface electromagnetic waves ellipsometry. Surf. Sci. **56**, 237–251 (1976). doi:10.1016/0039-6028(76)90450-7

82. H. Arwin, M.K. Poksinski, K. Johansen, Total internal reflection ellipsometry: principles and applications. Appl. Opt. **43**, 3028–3036 (2004). doi:10.1364/AO.43.003028

83. T. Lopez-Rios, G. Vuye G, Use of surface plasmon excitation for determination of the thickness and the optical constants of very thin surface layers. Surf. Sci. **81**, 529–538 (1979). doi:10.1016/0039-6028(79)90118-3

84. P. Wissmann, H.-U. Finzel, Electrical Resistivity of Thin Metal Films (Springer Tracts in Modern, Physics, Berlin, 2007)

85. E.H. Sondheimer, The mean free path of electrons in metals. Adv. Phys. **50**, 499–537 (2001). doi:10.1080/00018730110102187

86. M. Foldyna, A. De Martino, R. Ossikovski, E. Garcia-Caurel, C. Licitra, Characterization of grating structures by Mueller polarimetry in presence of strong depolarization due to finite spot size and spectral resolution. Opt. Commun. **282**, 735–741 (2009). doi:10.1016/j.optcom.2008.11.036

87. J.M. Correas, P. Melero, J.J. Gil, Decomposition of Mueller matrices in pure optical media. Mon. Sem. Mat. Garcia de Galdeano **27**, 233–240 (2003). Free PDF can be downloaded www.unizar.es/galdeano/actas_pau/PDF/233.pdf

88. M. Foldyna, E. Garcia-Caurel, R. Ossikovski, A. De Martino, J.J. Gil, Retrieval of a non-depolarizing component of experimentally determined depolarizing Mueller matrices. Opt. Express **17**, 12794–12806 (2009). doi:10.1364/OE.17.012794

89. B.J. Rice, H. Cao, M. Grumski, J. Roberts, The limits of CD metrology. Microelectron. Eng. **83**, 1023 (2006). doi:10.1063/1.2062991

90. See for example the proceedings of the conference "Advanced Lithography" available on line at the site of the SPIE, www.spie.org

91. V. Ukraintsev, A comprehensive test of optical scatterometry readiness for 65 nm technology production. Proc. SPIE **6152**, 61521G (2006). doi:10.1117/12.657649

92. M. Foldyna, A. De Martino, D. Cattelan, F. Bogeat, C. Licitra, J. Foucher, P. Barritault, J. Hazart, Accurate dimensional characterization of periodic structures by spectroscopic Mueller polarimetry. Proc. SPIE **7140**, 71400I (2008). doi:10.1117/12.804682

93. L. Li, Symmetries of cross-polarization diffraction. J. Opt. Soc. Am. A **17**, 881–887 (2000). doi:10.1364/JOSAA.17.000881

94. A. De Martino, M. Foldyna, T. Novikova, D. Cattelan, P. Barritault, C. Licitra, J. Hazart, J. Foucher, F. Bogeat, Comparison of spectroscopic Mueller polarimetry, standard scatterometry and real space imaging techniques (SEM and 3D-AFM) for dimensional characterization of periodic structures. SPIE Proc **6922**, 69221P (2008). doi:10.1117/12.772721
95. M.G. Moharam, T.K. Gaylord, Diffraction analysis of dielectric surface-relief gratings. J. Opt. Soc. Am. **72**, 1385 (1982). doi:10.1364/JOSA.72.001385
96. R.M. Silver, B.M. Barnes, A. Heckert, R. Attota, R. Dixson, J. Jun, Angle resolved optical metrology. Proc. SPIE **6922**, 69221M.1–69221M.12 (2008). doi:10.1117/12.777131
97. P. Leray, S. Cheng, D. Kandel, M. Adel, A. Marchelli, I. Vakshtein, M. Vasconi, B. Salski, Diffraction based overlay metrology: accuracy and performance on front end stack. Proc. SPIE **6922**, (2008) doi:10.1117/12.772516
98. Y.-n. Kim, J.-s. Paek, S. Rabello, S. Lee, J. Hu, Z. Liu, Y. Hao, W. Mcgahan, Device based in-chip critical dimension and overlay metrology. Opt. Express **17**, 21336–21343 (2009). doi:10.1364/OE.17.021336
99. T. Novikova, A. De Martino, R. Ossikovski, B. Drévillon, Metrological applications of Mueller polarimetry in conical diffraction for overlay characterization in microelectronics. Eur. Phys. J. Appl. Phys. **69**, 63–69 (2005). doi:10.1051/epjap:2005034
100. C. Fallet, Polarimétrie de Mueller résolue angulairement et applications aux structures périodiques. Ph.D. Thesis, Ecole Polytechnique, (2011). The manuscript can be freely downloaded at http://pastel.archives-ouvertes.fr/pastel-00651738/
101. M.R. Antonelli, A. Pierangelo, T. Novikova, P. Validire, A. Benali, B. Gayet, A. De Martino, Impact of model parameters on Monte Carlo simulations of backscattering Mueller matrix images of colon tissue. Biomed. Opt. Express **2**, 1836–1851 (2011). doi:10.1364/BOE.2.001836
102. A. Pierangelo, A. Benali, M.R. Antonelli, T. Novikova, P. Validire, B. Gayet, A. De Martino, Ex-vivo characterization of human colon cancer by Mueller polarimetric imaging. Opt. Express **19**, 1582–1593 (2011). doi:10.1364/OE.19.001582

Chapter 3
Data Analysis for Nanomaterials: Effective Medium Approximation, Its Limits and Implementations

Josef Humlicek

Abstract We review here basic theoretical approaches to the optical response of nanostructured materials. We use the well established framework of Effective medium approximation (EMA) and discuss key issues of its use. The treatment of this extensive subject is adapted to the needs of ellipsometric/polarimetric measurements on nanostructured materials. In Sects. 3.1 and 3.2 we formulate the problems and establish notation. Then, we recall and discuss, in Sects. 3.3 and 3.4, several well-known formulae for the effective dielectric function. Sections 3.5 and 3.6 are devoted to a fairly detailed comparison of selected measured data with results of the EMA models. We also assess the uncertainties involved in the EMA approach by visualizing the differences between results of its different versions (Sect. 3.7) and by calculating the differences from exact solutions (Sect. 3.8). Finally, Sect. 3.9 is devoted to the discussion of possible resonant behaviour of EMA mixtures.

3.1 Introduction

Contemporary materials science provides a wealth of unique materials obtained by a fine (nanometer-sized) mixing of different components. An example of a fairly complex artificial nanostructure is shown in Fig. 3.1. Six layers of the nanoscale mixture of molybdenum and SiO_2, separated by very thin SiO_2 spacers (bright horizontal lines), are placed on the oxidized silicon substrate (the bottom part showing the segment of 100 nm of length) and capped with the topmost SiO_2 layer. Differ-

J. Humlicek (✉)
Central European Institute of Technology, Masaryk University Kamenice 735,
62500 Brno, Czech Republic
e-mail: humlicek@physics.muni.cz

J. Humlicek
Department of Condensed Matter Physics, Faculty of Science, Masaryk University Kotlárská 2,
61137 Brno, Czech Republic
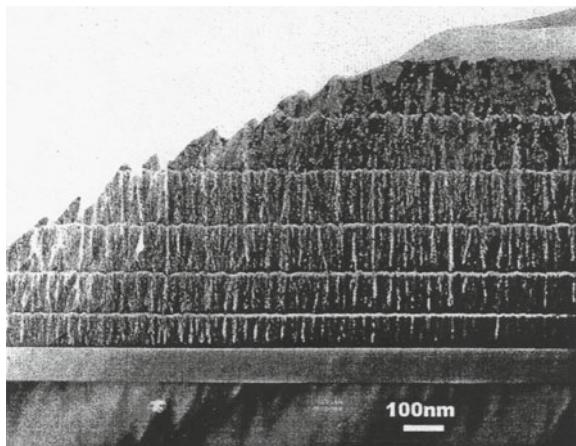
Fig. 3.1 TEM cross section of molybdenum/SiO$_2$ multilayer nanocomposite deposited on SiO$_2$/Si substrate. Courtesy of Ivo Vávra, Bratislava

ent deposition conditions of the individual Mo/SiO$_2$ layers were chosen in order to obtain different microstructures, seen as the changing contrast pattern of the transmission electron microscopy (TEM) picture. The targeted functionality of this kind of structures might be, for example, low-frequency electric conduction and/or optical behaviour in infrared/visible/ultraviolet range.

In general, many fundamental and functional properties of nanocomposites can be conveniently probed using several variants of optical spectroscopy. In particular, using polarized light in ellipsometric measurements proves to be highly efficient. The scheme of reflection and transmission of polarized optical wave interacting with slabs of nanostructured materials is shown in Figs. 3.2 and 3.3. The structure of Fig. 3.2 consists of thin layers extended in the (x,y) plane and stacked along the z-direction, which is typical of contemporary epitaxial heterostructures; that of Fig. 3.3 indicates small particles dispersed in a matrix.

Optical fields in the (meta)materials and the corresponding far-field solutions can be, in principle, obtained from rigorous full-wave analysis. This is usually very tedious, and the results are accompanied by a considerable volume of unwanted information. The fine structure of the optical fields is mostly irrelevant, since only smooth macroscopic averages are detected in actual measurements. Thus, the obvious replacement of mixed materials with "effective medium" can provide a plausible solution. This approach is usually termed EMA (effective medium approximation) or EMT (effective medium theory). The concept of EMA is very old; in fact, Maxwell included a paragraph on the electric conduction in a mixture in his famous *Treatise* [1].
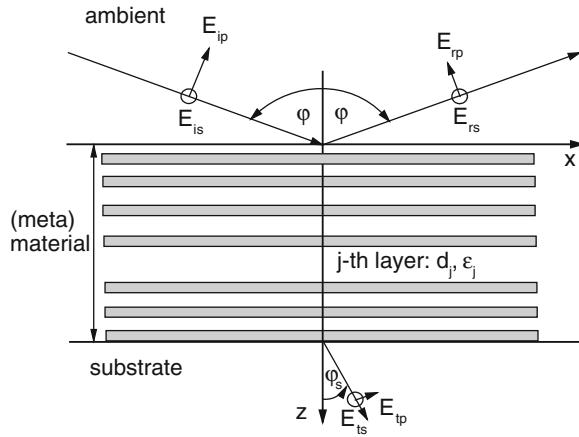
Fig. 3.2 Cross section of a planar ambient-film-substrate system. The film consists of a stack of different materials with interfaces parallel to the surface
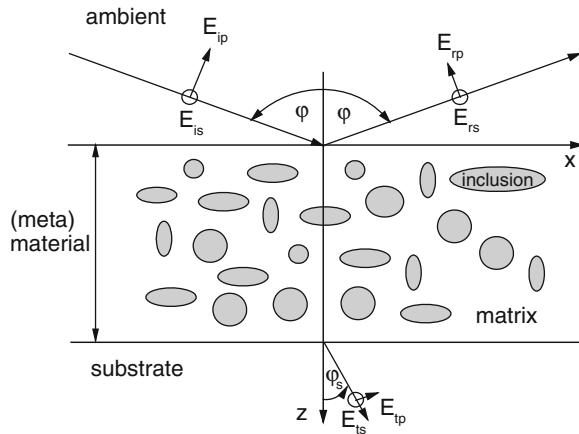

Fig. 3.3 Cross section of a planar ambient-film-substrate system. The film consists of nanoparticles embedded in a matrix material

3.2 Linear Optical Response of Nanostructured Materials

Optical frequencies, $\sim 3 \times 10^{11} - 3 \times 10^{16}$ Hz, cover the range from far-infrared (FIR) to vacuum-ultraviolet (VUV) spectral regions; the longest and shortest vacuum wavelength is about 1 mm and 10 nm, respectively. Since the atomic dimensions are of the order of 0.1 nm, matter behaves as a continuum at the optical frequencies and below (microwaves and radio waves). Namely, the wavelength is large enough to prevent substantial diffraction on the atomic structure, in contrast to the shorter wavelengths of the X-ray range. Although the discrete atomic structure of matter pro-

duces strong spatial variations of the quantities describing optical fields, the smooth averages at macroscopic length scales are usually detected using a light probe from the optical range.

The nanostructured materials are *mixtures* of individual components, possessing their own optical (continuum-like) behaviour. In other words, the atomic structure of individual components enters merely a limited number of macroscopic quantities, characteristic of the corresponding continuum. A second (sometimes called *mesoscopic*) level of averaging might be useful in representing the mixture as another continuum, with negligible diffraction on the mesoscopic structure. The possibility to use this continuum representation is not as universal as that of bulk materials, due to a wide range of possible sizes of the components.

The essential step in treating a mixture as a continuum consists in finding the macroscopic (averaged) field quantities. The actual behaviour of the mixture differs from this approximative treatment (effective medium approximation, EMA), whenever the structuring is not fine enough on the length scale of a single wavelength. Obviously, the fulfilment of this condition depends on the spectral range: a mesoscopic structure with characteristic dimensions of $\sim 10\,\text{nm}$ is likely to behave as a continuum in infrared and visible, but not in ultraviolet.

Maxwell equations govern the spatial and temporal dependencies of the electromagnetic field:

$$\nabla \times \boldsymbol{E} = -\partial \boldsymbol{B}/\partial t, \quad \nabla \times \boldsymbol{H} = \partial \boldsymbol{D}/\partial t + \boldsymbol{j}, \quad \nabla \cdot \boldsymbol{D} = \rho, \quad \nabla \cdot \boldsymbol{B} = 0. \quad (3.1)$$

The electric and magnetic field quantities and their units in the SI system are the following vectors: $\boldsymbol{E}$ [V/m], the electric intensity, $\boldsymbol{H}$ [A/m], the magnetic intensity, $\boldsymbol{D}$ [As/m^2], the electric displacement (flux density), $\boldsymbol{B}$ [Vs/m^2], the magnetic displacement (flux density), $\boldsymbol{j}$ [A/m^2], the current density. The scalar ρ [As/m^3] is the charge density.

The linear response of matter to the monochromatic electromagnetic field, where all of the quantities follow the harmonic time dependence of $\exp(-i\omega t)$ with the angular frequency ω, is described by the *constitutive relations* for the displacements and intensities,

$$\boldsymbol{D} = \varepsilon(\omega)\varepsilon_0 \boldsymbol{E}, \quad \boldsymbol{B} = \mu_0 \boldsymbol{H}, \quad (3.2)$$

where $\varepsilon_0 = 8.85 \times 10^{-12}$ As/Vm is the vacuum permittivity, $\mu_0 = 4\,\pi \times 10^{-7}$ Vs/Am the vacuum permeability, and $\varepsilon(\omega)$ the (dimensionless) relative permittivity. Alternatively, the induced current density $\boldsymbol{j}$ is a linear function of the electric intensity and the proportionality factor, the conductivity σ, is simply related to the permittivity:

$$\boldsymbol{j} = \sigma(\omega)\boldsymbol{E}, \quad \sigma(\omega) = -i\omega[\varepsilon(\omega) - 1]\varepsilon_0. \quad (3.3)$$

The SI unit of conductivity is A/Vm = 1/Ωm. The linear response of Eqs. (3.2) and (3.3) contains complex functions of frequency, $\varepsilon = \varepsilon_1 + i\varepsilon_2$ and $\sigma = \sigma_1 + i\sigma_2$,

reflecting possible phase shifts between the electric field and the induced polar-ization/current at finite frequencies. The second of equations (3.3) is a simple consequence of the indistinguishability of $\partial \boldsymbol{D}/\partial t$ and $\boldsymbol{j}$ in the second of Maxwell equation (3.1) at optical frequencies.

As usual, we use the complex permittivity as the preferred response function. However, the complex conductivity might be more appropriate for conducting struc-tures in IR.

3.3 Average Fields and Effective Permittivity for a Small Contrast

A simple approach to the dielectric response of a mixture is due to Landau-Lifshitz [2]. The mixture is assumed to be finely dispersed, representing a homogeneous and isotropic material with respect to the macroscopic field. Its effective permittivity relates the volume-averaged displacement and intensity,

$$\langle \boldsymbol{D} \rangle = \varepsilon_{eff} \langle \boldsymbol{E} \rangle, \tag{3.4}$$

where the averaging volume V has to be large enough to be representative of the mixture, and, in optical case, small in comparison with the wavelength. Using the averages of the intensity and permittivity,

$$\langle \boldsymbol{E} \rangle = (1/V) \int \boldsymbol{E}(\boldsymbol{r})dxdydz, \quad \langle \varepsilon \rangle = (1/V) \int \varepsilon(\boldsymbol{r})dxdydz, \tag{3.5}$$

the local field intensity and permittivity at the position $\boldsymbol{r} = (x, y, z)$ can be written as

$$\boldsymbol{E}(\boldsymbol{r}) = \langle \boldsymbol{E} \rangle + \delta \boldsymbol{E}(\boldsymbol{r}), \quad \varepsilon(\boldsymbol{r}) = \langle \varepsilon \rangle + \delta \varepsilon(\boldsymbol{r}). \tag{3.6}$$

The essential step in calculating the mean displacement,

$$\langle \boldsymbol{D} \rangle = \langle (\langle \varepsilon \rangle + \delta \varepsilon(\boldsymbol{r}))(\langle \boldsymbol{E} \rangle + \delta \boldsymbol{E}(\boldsymbol{r})) \rangle = \langle \varepsilon \rangle \langle \boldsymbol{E} \rangle + \langle \delta \varepsilon(\boldsymbol{r})\delta \boldsymbol{E}(\boldsymbol{r}) \rangle, \tag{3.7}$$

consists in finding the average of the product $\delta \varepsilon(\boldsymbol{r})\delta \boldsymbol{E}(\boldsymbol{r})$. An approximate treat-ment of Landau-Lifshitz uses the third of Maxwell equation (3.1) with the vanishing charge density ρ, which relates the divergence of the displacement to positional dependencies of the field intensity and permittivity:

$$\nabla \cdot \boldsymbol{D} = \nabla \cdot [(\langle \varepsilon \rangle + \delta \varepsilon(\boldsymbol{r}))(\langle \boldsymbol{E} \rangle + \delta \boldsymbol{E}(\boldsymbol{r}))]$$
$$\approx \langle \varepsilon \rangle \nabla \cdot [\delta \boldsymbol{E}(\boldsymbol{r})] + \langle \boldsymbol{E} \rangle \nabla \cdot [\delta \varepsilon(\boldsymbol{r})] = 0. \tag{3.8}$$

The divergence of the product $\delta\varepsilon(\mathbf{r})\delta\mathbf{E}(\mathbf{r})$ is supposed to be negligible. The second term of the right-hand side of Eq. (3.7) is calculated in two steps. First, the spatial average is performed over the regions of constant $\delta\varepsilon$, i.e., for a given component of the mixture. The corresponding average of $\delta\mathbf{E}$ can be obtained from its divergence

$$\nabla \cdot \langle \delta\mathbf{E}(\mathbf{r})\rangle = \frac{\delta}{\delta x}\langle \delta E_x(\mathbf{r})\rangle + \frac{\delta}{\delta y}\langle \delta E_y(\mathbf{r})\rangle + \frac{\delta}{\delta z}\langle \delta E_z(\mathbf{r})\rangle = 3\frac{\delta}{\delta x}\langle \delta E_x(\mathbf{r})\rangle, \quad (3.9)$$

using the assumption of the isotropy of the 3-dimensional mixture. Choosing the direction of $\langle \mathbf{E}\rangle$ along x, we obtain from Eq. (3.8) the following relations,

$$3\langle\varepsilon\rangle\frac{\delta}{\delta x}\langle \delta E_x(\mathbf{r})\rangle = -\langle E_x(\mathbf{r})\rangle\frac{\delta}{\delta x}\langle\varepsilon(\mathbf{r})\rangle, \quad \langle \delta\mathbf{E}(\mathbf{r})\rangle = -\frac{1}{3\langle\varepsilon(\mathbf{r})\rangle}\langle \mathbf{E}(\mathbf{r})\rangle\delta\varepsilon(\mathbf{r}).$$
$$(3.10)$$

The second stage of finding the average of Eq. (3.7) is performed by multiplying the second of Eq. (3.10) by $\delta\varepsilon$, and averaging over all species in the mixture. The result reads

$$\langle \delta\varepsilon(\mathbf{r})\delta\mathbf{E}(\mathbf{r})\rangle = -\frac{1}{3\langle\varepsilon(\mathbf{r})\rangle}\langle \mathbf{E}(\mathbf{r})\rangle\langle(\delta\varepsilon(\mathbf{r}))^2\rangle. \quad (3.11)$$

Using this in Eqs. (3.7) and (3.4), we find the resulting effective permittivity,

$$\varepsilon_{eff} = \langle\varepsilon\rangle - \frac{\langle(\delta\varepsilon)^2\rangle}{3\langle\varepsilon\rangle}. \quad (3.12)$$

We call the result of Eq. (3.12) the Landau-Lifshitz (LL) formula. It can be expressed approximately in the following appealing way. Neglecting higher-order terms in the third power of the Taylor expansion of the averaged cube root,

$$\langle\varepsilon^{1/3}\rangle^3 = \langle(\langle\varepsilon\rangle + \delta\varepsilon)^{1/3}\rangle^3 \approx \langle\varepsilon\rangle - 3\langle\varepsilon\rangle\frac{1\cdot 2}{3\cdot 6}\frac{\langle(\delta\varepsilon)^2\rangle}{\langle\varepsilon\rangle^2}, \quad (3.13)$$

the right-hand side of Eq. (3.12) is recovered, i.e.,

$$\varepsilon_{eff} = \langle\varepsilon^{1/3}\rangle^3. \quad (3.14)$$

This form of the approximate effective permittivity has been obtained in a different way by Looyenga [3]; the formula of Eq. (3.14) is sometimes labeled as LLL (Landau-Lifshitz-Looyenga). In practice, it does not bring any essential advantage compared with the LL formula of Eq. (3.12).

The above development can be easily modified for structures of lower dimensions. Let us assume translational invariance of the mixture along one spatial dimension (z), with the structuring restricted to the (x,y) plane of the 2-dimensional (2D) system. The material is uniaxial, with the optical axis along z, and the z (extraordinary) component

of its dielectric tensor equals the volume average $\langle \varepsilon \rangle$. In order to calculate the ordinary component, we notice that the divergence of Eq. (3.9) becomes

$$\nabla \cdot \langle \delta \boldsymbol{E}(\boldsymbol{r}) \rangle = \frac{\delta}{\delta x} \langle \delta E_x(\boldsymbol{r}) \rangle + \frac{\delta}{\delta y} \langle \delta E_y(\boldsymbol{r}) \rangle = 2 \frac{\delta}{\delta x} \langle \delta E_x(\boldsymbol{r}) \rangle. \tag{3.15}$$

Similarly, with the structuring vanishing along z and y in a one-dimensional (1D) mixture, the material is uniaxial with the optical axis along x, and the y and z (ordinary) components of the dielectric tensor are equal to the volume average $\langle \varepsilon \rangle$. The divergence of Eq. (3.9) is reduced to

$$\nabla \cdot \langle \delta \boldsymbol{E}(\boldsymbol{r}) \rangle = \frac{\delta}{\delta x} \langle \delta E_x(\boldsymbol{r}) \rangle. \tag{3.16}$$

The corresponding modification of the corresponding tensor components of the effective permittivity of Eq. (3.12) is

$$\varepsilon_{eff} = \langle \varepsilon \rangle - \frac{\langle (\delta \varepsilon)^2 \rangle}{D \langle \varepsilon \rangle}, \tag{3.17}$$

where D =3, 2, 1 for the 3D, 2D, and 1D mixtures, respectively.

Let us note that looking for an approximate representation of Eq. (3.17) using powers, analogous to Eq. (3.14), leads to the following simple result in the 1-dimensional ($D = 1$) case: the right-hand side of Eq. (3.17) is recovered when expanding

$$\varepsilon_{eff}|_{1D} = \langle \varepsilon^{-1} \rangle^{-1}. \tag{3.18}$$

This is actually a precise result of the long-wavelength averaging for general lamellar structures, not restricted to the low contrast of constituents. We will use this fact later in order to test the level of errors introduced by the assumption of the small contrast.

For the simplest mixture consisting of just two components, a and b, the averages can be expressed explicitly in terms of their permittivities, ε_a and ε_b, and the volume fractions, f_a and f_b,

$$f_a = V_a/(V_a + V_b), \quad f_b = V_b/(V_a + V_b) = 1 - f_a \equiv f. \tag{3.19}$$

Thus, the composition of binary mixtures is specified by a single parameter f, the volume fraction of the component occupying volume V_b, with the permittivity of ε_b. The average permittivity of the binary mixture is

$$\langle \varepsilon \rangle = (1 - f)\varepsilon_a + f \varepsilon_b = \varepsilon_a + f(\varepsilon_b - \varepsilon_a), \tag{3.20}$$

the deviations from the mean are

$$\delta \varepsilon_a = \varepsilon_a - \langle \varepsilon \rangle = f(\varepsilon_a - \varepsilon_b), \quad \delta \varepsilon_b = \varepsilon_b - \langle \varepsilon \rangle = (1 - f)(\varepsilon_b - \varepsilon_a), \tag{3.21}$$

the mean of the squared deviation is

$$\langle(\delta\varepsilon)^2\rangle = (1-f)(\delta\varepsilon_a)^2 + f(\delta\varepsilon_b)^2 = f(1-f)(\varepsilon_b - \varepsilon_a)^2, \qquad (3.22)$$

and the LL formula of Eq. (3.17) reads

$$\varepsilon_{\text{eff}} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a) - \frac{f(1-f)(\varepsilon_b - \varepsilon_a)^2}{D[\varepsilon_a + f(\varepsilon_b - \varepsilon_a)]}. \qquad (3.23)$$

The approximation based on neglecting higher-order terms in the derivation of Eqs. (3.8)–(3.17) introduces differences that are, in general, difficult to specify. We will discuss these errors later, using partly precise solutions for selected 1D cases, partly comparisons with predictions of the effective dielectric functions that do not assume the low contrast.

3.4 A Collection of Mixing Rules

A mixture is isotropic in the absence of preferred directions of polarization on macroscopic scale. This condition is evidently not fulfilled in cases of preferential orientations of the interfaces between the components. Two extreme cases of the orientation of aligned *planar* interfaces with respect to the direction of electric field can easily be identified; they are shown in Fig. 3.4.

Applying the boundary conditions for tangential and normal components of E and H [4], the two orientations differ in the average values of the field quantities. Namely, $E_a = E_b = \langle E \rangle$ in the parallel case, leading to the following average value of the displacement,

$$\langle D \rangle = (1-f)\varepsilon_a E_a + f\varepsilon_b E_b = [(1-f)\varepsilon_a + f\varepsilon_b]\langle E \rangle, \qquad (3.24)$$
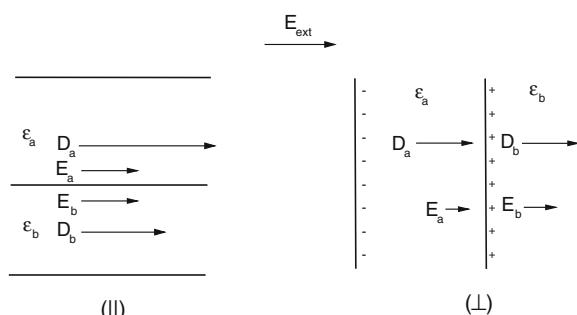

Fig. 3.4 Cross sections of a layered structure with the parallel (*left part*, no screening charges at the interfaces) and perpendicular (*right part*, maximum screening charge density) orientation with respect to the electric field

where f is the volume fraction of the component b. Thus, the effective permittivity of Eq. (3.4) is the simple volume average of the two components,

$$\varepsilon_{eff} = \langle\varepsilon\rangle = (1 - f)\varepsilon_a + f\varepsilon_b. \tag{3.25}$$

On the other hand, normal components of the displacement $\boldsymbol{D}$ are continuous ($\boldsymbol{D}_a = \boldsymbol{D}_b = \langle\boldsymbol{D}\rangle$) across the interfaces in the perpendicular case, leading to the following average value of the electric intensity perpendicular to the planes,

$$\langle\boldsymbol{E}\rangle = (1 - f)\boldsymbol{D}_a/\varepsilon_a + f\boldsymbol{D}_b/\varepsilon_b = [(1 - f)/\varepsilon_a + f/\varepsilon_b]\langle\boldsymbol{D}\rangle. \tag{3.26}$$

Consequently, the inverse of the effective permittivity along the horizontal direction is the volume average of the inverses of the permittivities of the individual components (i.e., the inverse permittivity is additive),

$$1/\varepsilon_{eff} = (1 - f)/\varepsilon_a + f/\varepsilon_b. \tag{3.27}$$

The two orientations of layers in Fig. 3.4. differ in the absence/presence of screening charges at the interfaces. Namely, a surface charge density develops at the interfaces in the perpendicular case, which is related to the discontinuity of the electric field intensity across the interface. On the other hand, no screening charges appear in the parallel case. Curved interfaces lead evidently to more complex field patterns, with positional dependence of the surface charge density.

The case of a *spherical* (more generally, ellipsoidal) inclusion is particularly simple, as it allows an analytical solution [5]. Shown in Fig. 3.5 are the field lines describing the field in and around a dielectric sphere. The inner field, produced by the applied intensity and screening charges unevenly distributed on the surface, is constant [4]:

$$E_b = \frac{3\varepsilon_a}{\varepsilon_b + 2\varepsilon_a}E_a. \tag{3.28}$$

Note that the inner field is weaker than the outer one for $\varepsilon_b > \varepsilon_a$, i.e., for the inclusion more polarizable than the host material (if both permittivities are real). Static fields are assumed in calculations of average quantities; however, the optical case is essentially the same for the diameter of the sphere negligible with respect to the wavelength and the penetration depth of light.

The more general case of an ellipsoid leads to similar results. Let us assume an ellipsoid (permittivity ε_b, semi axes u, v, and w) with u oriented along the electric field intensity. The field inside the ellipsoid is constant [5],

$$E_b = \frac{\varepsilon_a}{(1 - L_u)\varepsilon_a + L_u\varepsilon_b}E_a, \tag{3.29}$$

where

$$L_u = \frac{uvw}{2} \int_0^\infty \frac{dt}{(t^2 + u^2)\sqrt{(t + u^2)(t + v^2)(t + w^2)}} \in \langle 0, 1 \rangle \qquad (3.30)$$

is the *depolarization factor*. The three depolarization factors of any ellipsoid satisfy the condition $L_u + L_v + L_w = 1$. Special cases are

$$L_u = L_v = L_w = \frac{1}{3} \quad \text{for } u = v = w \quad \text{(sphere)}, \qquad (3.30a)$$

$$L_u = L_v = \frac{1}{2}, L_w = 0 \quad \text{for } u = v, w \to \infty \quad \text{(cylinder)}, \qquad (3.30b)$$

$$L_u = 1, L_v = L_w = 0 \quad \text{for } v = w \to \infty \quad \text{(slab)}. \qquad (3.30c)$$

Note that the zero depolarization inserted into Eq. (3.29) reproduces the condition $E_a = E_b$ for the interface parallel to the field; this occurs for the orientation of a cylinder or a slab with the interfaces parallel to the field. On the other hand, the maximum (unit) value of the depolarization factor describes the continuity of the normal component of electric displacement, $\varepsilon_a E_a = \varepsilon_b E_b$, for a slab with its normal oriented along the field.

The simplest version of averaging the microscopic fields (such as those in Fig. 3.5) neglects the dipole-like field pattern in the close neighborhood of the spheres, sparsely dispersed in the host material. Using the volume fraction f of the spheres, the average intensity and displacement is approximately

$$\langle E \rangle \approx (1 - f)E_a + f E_b = \left[1 - f + \frac{f 3\varepsilon_a}{\varepsilon_b + 2\varepsilon_a} \right] E_a, \qquad (3.31a)$$

$$\langle D \rangle \approx (1 - f)\varepsilon_a E_a + f \varepsilon_b E_b = \left[(1 - f)\varepsilon_a + \frac{f 3\varepsilon_a \varepsilon_b}{\varepsilon_b + 2\varepsilon_a} \right] E_a. \qquad (3.31b)$$
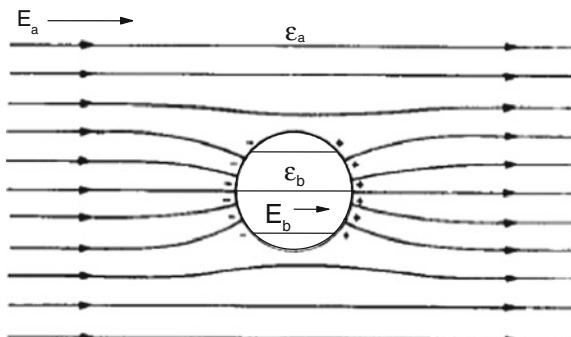

Fig. 3.5 A dielectric sphere embedded in an infinite dielectric medium under the electric field E_a

The effective dielectric function of Eq. (3.4) is then given by the well known Maxwell Garnett [6] formula

$$\varepsilon_{eff} = \frac{\langle D \rangle}{\langle E \rangle} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{3\varepsilon_a}{\varepsilon_b + 2\varepsilon_a - f(\varepsilon_b - \varepsilon_a)}. \tag{3.32}$$

This result explained colors in glasses with nanometer-sized spherical metallic inclusions as early as in 1904.

The Maxwell Garnett (MG) formula is easily generalized to a dilute mixture of aligned ellipsoids in a host matrix as

$$\varepsilon_{eff,u} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{\varepsilon_a}{\varepsilon_a + L_u(1 - f)(\varepsilon_b - \varepsilon_a)}. \tag{3.32a}$$

A mixture with randomly oriented ellipsoids is isotropic, with the scalar permittivity

$$\varepsilon_{eff} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\varepsilon_a\frac{\displaystyle\sum_{j=u,v,w} 1/[\varepsilon_a + L_j(\varepsilon_b - \varepsilon_a)]}{3 - f(\varepsilon_b - \varepsilon_a)\displaystyle\sum_{j=u,v,w} L_j/[\varepsilon_a + L_j(\varepsilon_b - \varepsilon_a)]}. \tag{3.32b}$$

Two limiting cases of very prolate and oblate ellipsoids are useful; the cylinders of Eq. (3.30b) are likely to approximate the behavior of needle-like inclusions, with the effective permittivity of the random orientation

$$\varepsilon_{eff} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{\varepsilon_b + 5\varepsilon_a}{3(\varepsilon_b + \varepsilon_a) + 2f(\varepsilon_a - \varepsilon_b)}. \tag{3.32c}$$

Using the depolarization factors of Eq. (3.30c), we obtain the effective permittivity of disk- or platelet-shaped inclusions with random orientation,

$$\varepsilon_{eff} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{2\varepsilon_b + 2\varepsilon_a}{3\varepsilon_b + f(\varepsilon_a - \varepsilon_b)}. \tag{3.32d}$$

An instructive treatment of mixtures can be based on a simplified *microscopic* model, with the polarization of components represented by point dipoles [7]. For simple geometries, the average values can be found analytically, and the distinction between local field causing the polarization and experimentally accessible macroscopic (averaged) field is straightforward. On the other hand, the actual induced density of dipole moment of either electrons or atomic nuclei in condensed matter is significantly different from the diverging pattern of point dipoles. A discrete set of dipoles of different polarizabilities, representing two different materials, leads to the effective dielectric function obeying the equation

$$f\frac{\varepsilon_b - 1}{\varepsilon_b + 2} + (1 - f)\frac{\varepsilon_a - 1}{\varepsilon_a + 2} = \frac{\varepsilon_{eff} - 1}{\varepsilon_{eff} + 2}. \tag{3.33}$$

Evidently, it can be viewed as a variant of Clausius-Mosotti or Lorentz–Lorenz relation [7]. It is actually of the same form as the Maxwell Garnett formula (3.32); the latter can easily be put in the form

$$f \frac{\varepsilon_b - \varepsilon_a}{\varepsilon_b + 2\varepsilon_a} = \frac{\varepsilon_{eff} - \varepsilon_a}{\varepsilon_{eff} + 2\varepsilon_a}. \tag{3.34}$$

In fact, introducing the permittivity ε_h of the *host material*, the following general form

$$f \frac{\varepsilon_b - \varepsilon_h}{\varepsilon_b + 2\varepsilon_h} + (1 - f)\frac{\varepsilon_a - \varepsilon_h}{\varepsilon_a + 2\varepsilon_h} = \frac{\varepsilon_{eff} - \varepsilon_h}{\varepsilon_{eff} + 2\varepsilon_h} \tag{3.35}$$

covers the Lorentz–Lorenz formula of Eq. (3.33) when taking $\varepsilon_h = 1$ (vacuum), and the Maxwell Garnett formula of Eq. (3.34) with $\varepsilon_h = \varepsilon_a$ (precisely in the spirit of the approximate calculations of the averages of Eq. (3.31a,b). An appealing variant of the mixing is to assume the host material to be the effective medium itself, $\varepsilon_h = \varepsilon_{eff}$. This was suggested by Bruggeman [8], resulting in the mixing formula

$$f \frac{\varepsilon_b - \varepsilon_{eff}}{\varepsilon_b + 2\varepsilon_{eff}} + (1 - f)\frac{\varepsilon_a - \varepsilon_{eff}}{\varepsilon_a + 2\varepsilon_{eff}} = 0. \tag{3.36}$$

The effective permittivity solves the following quadratic equation, with one of its two roots being physical,

$$2\varepsilon_{eff}^2 + \varepsilon_{eff}\left[(3f - 2)\varepsilon_a + (1 - 3f)\varepsilon_b\right] - \varepsilon_a\varepsilon_b = 0. \tag{3.37}$$

The Bruggeman formula is symmetric with respect to interchanging the components, which is attractive for dealing with materials of comparable volume fractions in the mixture. Further, the way of its derivation leads to the expectation of a better performance outside the dilute limit of Maxwell Garnett formula.

Another attractive mixing rule is known as the coherent potential formula; we recall the form for spherical inclusions [9],

$$\varepsilon_{eff} = \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{3\varepsilon_{eff}}{3\varepsilon_{eff} + (1 - f)(\varepsilon_b - \varepsilon_a)}. \tag{3.38}$$

As in the case of Bruggeman formula, the effective permittivity solves a quadratic equation, with only one of its two roots being physical:

$$3\varepsilon_{eff}^2 + \varepsilon_{eff}\left[4(f - 1)\varepsilon_a + (1 - 4f)\varepsilon_b\right] - (1 - f)\varepsilon_a(\varepsilon_b - \varepsilon_a) = 0. \tag{3.39}$$

We use here the label CPA (coherent potential approximation) for the mixing formula of Eq. (3.26). It is based, like the Bruggeman formula, on the assumption of the spherical form of inclusions. However, the averaging procedure of field quantities is different. For dilute mixtures of spheres ($f \ll 1$), both Bruggeman and CPA rules

give the same expansion to the first order in f as the Maxwell Garnett formula, namely

$$\varepsilon_{eff} \approx \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{3\varepsilon_a}{\varepsilon_b + 2\varepsilon_a}. \tag{3.40}$$

For real permittivities, the slope of this linear approximation is smaller than that of the linear interpolation between ε_a and ε_b for $\varepsilon_b > \varepsilon_a$ and vice versa. This is easily understandable with the help of the approximate averaging used in deriving the Maxwell Garnett rule above. It should be noted that the corresponding expansion found from Eq. (3.12),

$$\varepsilon_{eff} \approx \varepsilon_a + f(\varepsilon_b - \varepsilon_a)\frac{4\varepsilon_a - \varepsilon_b}{3\varepsilon_a}, \tag{3.41}$$

and (3.14),

$$\varepsilon_{eff} \approx \varepsilon_a + 3f(\varepsilon_a^{2/3}\varepsilon_b^{1/3} - \varepsilon_a), \tag{3.42}$$

is different from that of Eq. (3.40). This is not surprising, because of different assumptions used in deriving the latter two mixing rules. In fact, the most important requirement was a small contrast of the permittivities. No specific geometry of the mixture has been used; it should not be expected that the result valid for spherical inclusions would be obtained.

The EMA calculations done for the spherical inclusions can be easily modified for aligned ellipsoids, similar to the Maxwell Garnett type of averaging of Eq. (3.32a). The resulting effective medium is anisotropic, since the screening effects depend on the relative orientation of the electric field and the ellipsoids. An instructive case is that of the aligned cylinders, i.e., infinitely elongated ellipsoids with a circular cross-section. There is no screening for the field parallel to the cylinders, and the corresponding tensor component of the effective permittivity is just the volume average of Eq. (3.25). For the field perpendicular to a dilute system of cylinders, the two-dimensional distribution of screening charges leads to the following effective permittivity, which is a modification of Eq. (3.34),

$$f\frac{\varepsilon_b - \varepsilon_a}{\varepsilon_b + \varepsilon_a} = \frac{\varepsilon_{eff} - \varepsilon_a}{\varepsilon_{eff} + \varepsilon_a}, \tag{3.43}$$

sometimes called Rayleigh mixing formula. Obviously, Maxwell Garnett and Rayleigh mixing rules are two intermediate stages between the absence and maximum of screening in Eqs. (3.25) and (3.27), respectively.

  Most of the rules for binary mixtures discussed above can be rewritten in the form of explicit relations for the volume fraction f. This is convenient, since the typical use of the effective medium approach is to estimate the composition from the known permittivities of the components, and the measured value for the mixture. We list in the table below several mixing formulas using the expression for the effective permittivity and the volume fraction (Table 3.1).

Table 3.1 Effective medium rules for binary mixtures

Effective medium	ε_{eff}	f
Layer stack, planes parallel to E	$(1-f)\varepsilon_a + f\varepsilon_b$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}$
Layer stack, planes perp. to E	$\dfrac{\varepsilon_a\varepsilon_b}{(1-f)\varepsilon_b + f\varepsilon_a}$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}\,\dfrac{\varepsilon_b}{\varepsilon_{eff}}$
LLL (small contrast)	$[(1-f)\varepsilon_a^{1/3} + f\varepsilon_b^{1/3}]^3$	$\dfrac{\varepsilon_{eff}^{1/3}-\varepsilon_a^{1/3}}{\varepsilon_b^{1/3}-\varepsilon_a^{1/3}}$
Maxwell Garnett (dilute spheres)	$\varepsilon_a + f(\varepsilon_b-\varepsilon_a)\dfrac{3\varepsilon_a}{\varepsilon_b+2\varepsilon_a - f(\varepsilon_b-\varepsilon_a)}$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}\,\dfrac{\varepsilon_b+2\varepsilon_a}{\varepsilon_{eff}+2\varepsilon_a}$
Rayleigh (dilute cylinders)	$\varepsilon_a + f(\varepsilon_b-\varepsilon_a)\dfrac{2\varepsilon_a}{\varepsilon_b+\varepsilon_a - f(\varepsilon_b-\varepsilon_a)}$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}\,\dfrac{\varepsilon_b+\varepsilon_a}{\varepsilon_{eff}+\varepsilon_a}$
Maxwell Garnett (dilute aligned ellipsoids)	$\varepsilon_a + f(\varepsilon_b-\varepsilon_a)\dfrac{\varepsilon_a}{\varepsilon_a+L(1-f)(\varepsilon_b-\varepsilon_a)}$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}\,\dfrac{\varepsilon_a+L(\varepsilon_b-\varepsilon_a)}{\varepsilon_a+L(\varepsilon_{eff}-\varepsilon_a)}$
Bruggeman (spheres)	$2\varepsilon_{eff}^2 + \varepsilon_{eff}[(3f-2)\varepsilon_a+(1-3f)\varepsilon_b] - \varepsilon_a\varepsilon_b = 0$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}\,\dfrac{\varepsilon_b+2\varepsilon_a}{3\varepsilon_{eff}}$
CPA (spheres)	$3\varepsilon_{eff}^2 + \varepsilon_{eff}[4(f-1)\varepsilon_a+(1-4f)\varepsilon_b] - (1-f)\varepsilon_a(\varepsilon_b-\varepsilon_a) = 0$	$\dfrac{\varepsilon_{eff}-\varepsilon_a}{\varepsilon_b-\varepsilon_a}\,\dfrac{3\varepsilon_{eff}+\varepsilon_b-\varepsilon_a}{4\varepsilon_{eff}-\varepsilon_a}$

The volume fraction of the component a is $1-f$, that of the component b is f. L is the depolarization factor of Eq. (3.30). For the Bruggeman and CPA models the effective permittivity results from the solution of the quadratic equation given in the table

3.5 Tests of EMA: Glass Spheres in Liquids

An illustrative example of using different mixing formulas is the system of glass spheres dispersed in liquids. The effective response has been studied experimentally in the low ($\leq$GHz) frequency range, using the mixture as a dielectric in a condenser or a resonator. Even fairly large glass particles (mostly spheres) warrant the applicability of the continuum approach, since the vacuum wavelength λ_{vac} at the frequency of 1 GHz is 0.3 m and the effective refractive index of the mixture, $n_{eff} = \sqrt{(\varepsilon_{eff})}$, does not exceed 10. Thus, the diameter of the glass spheres, smaller than about 1 mm, is much smaller than the wavelength in the effective medium, λ_{vac}/n_{eff}.

First, we analyze the mixture of glass spheres in the non-polar liquid of carbon tetrachloride; the experimental results of J. A. Reynolds quoted in [3] are shown in Fig. 3.6, together with the long-wavelength (real) permittivity predicted by four selected mixing formulas as a function of the volume fraction f of glass. A slightly bowed dependence is almost coinciding for Bruggeman, Eq. (3.36), Landau-Lifshitz-Looyenga, Eq. (3.14), and CPA, Eq. (3.38), formulas. The prediction of Maxwell Garnett, Eq. (3.32), lies slightly below the three.

Since the mutual differences of the mixing formulas are rather small, we show their differences from the Bruggeman model in Fig. 3.7 on an expanded scale. Except for the Maxwell Garnett model, the effective permittivities are within the ±0.02 margins (less than $\pm1\%$ of the effective values) in the whole composition range. Also the deviations from experimental points are fairly small, as shown in Fig. 3.8. We have calculated the mean square deviations between the model and measured values and listed the results in Fig. 3.8. The least mean deviation occurs for the CPA rule, followed by the Landau-Lifshitz-Looyenga, Bruggeman and Maxwell Garnet
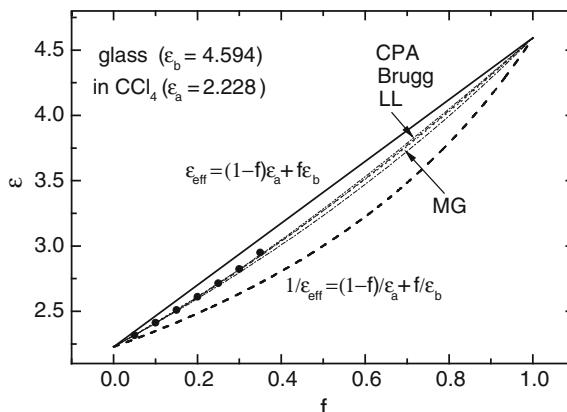

Fig. 3.6 Permittivity of glass spheres dispersed in carbon tetrachloride according to several mixing formulas on the whole range of composition (*lines*); measured data for the volume fraction of glass up to 0.35 (*symbols*)
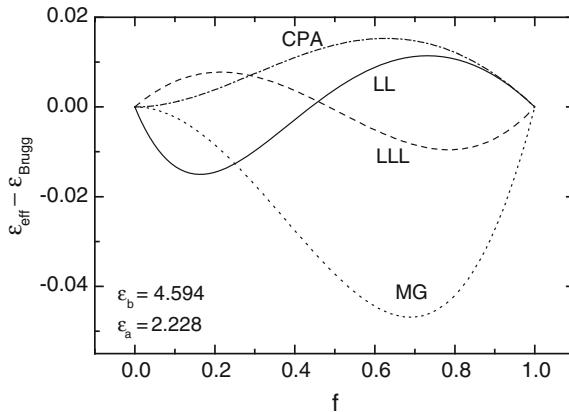
Fig. 3.7 Differences between several mixing formulas for the permittivity of glass spheres dispersed in carbon tetrachloride, data from Fig. 3.6, and the Bruggeman formula

rules. Evidently, the experimental point for the largest value of f might be an outlier; removing it from the data set reduces the mean square deviation of the CPA formula to 0.0015, about a half of the value for the LLL rule.

Another representation of the data is shown in Fig. 3.9. The volume fractions were calculated from the measured value of permittivity and those of the constituents using the same mixing rules as above. Considering the possibility of the data point with the largest value of f being an outlier, the agreement of Bruggeman and CPA with the measured data is excellent. The Maxwell Garnett rule is very good for
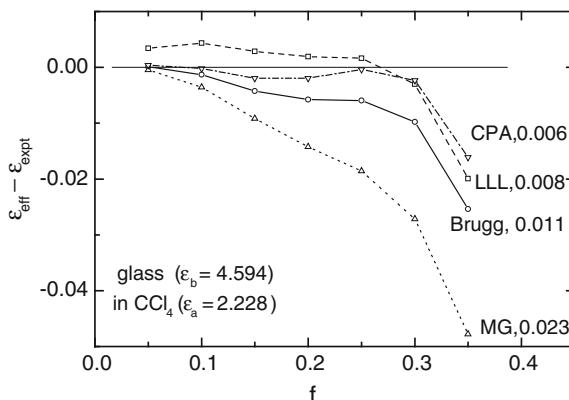

Fig. 3.8 Deviation of the predictions of different mixing formulas from experimental data for glass spheres dispersed in carbon tetrachloride. The mean square deviation for the 7 measured points is given next to the acronym of the mixing formula. The *lines* are guides to the eye
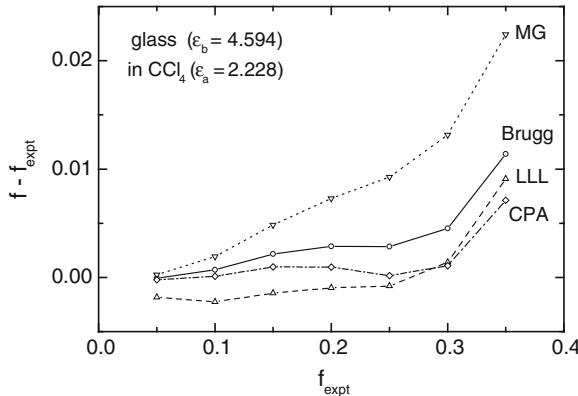
Fig. 3.9 Volume fraction of glass calculated from the measured values of permittivity. The *lines* are guides to the eye

small f, which is expected from the proper account of the screening charges on the glass spheres; however, it overestimates the value of the volume fraction due to the deficiencies of the simplified averaging of Eq. (3.31a,b). The LL rule underestimates the lower values of f; however, it was derived for the condition of small differences of permittivities in the mixture, not satisfied very well in the present case.

The LLL rule of Eq. (3.14) has been derived for a mixture of unspecified geometry, with the only assumption of the isotropy and small variations of the permittivity. The existence of screening charges on inner boundaries is hidden in the manipulation with the divergences of Eqs. (3.8) and (3.9). The corresponding "average screening" lies between the minimum and maximum of the planar interfaces oriented parallel and perpendicular to the electric field, respectively. These two cases would occur for the corresponding orientations of aligned glass platelets dispersed in CCl_4, with the tensor components of the permittivity of Eqs. (3.25) and (3.27). We compare in Fig. 3.10 the compositional dependence resulting from Eq. (3.12), and its approximation of Eq. (3.14), with these two limiting cases. Note that the results of Eqs. (3.12) and (3.14) differ rather markedly; this is obviously due to the large contrast of permittivities of glass and carbon tetrachloride. Further, this way of showing the measured data supports the suspicion of the presence of an outlier (for the largest volume fraction of glass, 0.35).

Another set of experimental data, suitable for testing the EMA mixing rules, has been collected for a dense packing of glass spheres immersed in various liquids [10]. The volume fraction f of soda lime silicate glass (diameter $500\,\mu$m) was fixed at the value of 0.605, the pores between the spheres were filled with different liquid immersions. In our notation, the dielectric function of glass inclusions, ε_b, had the value of 7.6. The permittivity of the immersing liquids, ε_a, span the range from 1 (air) to 78.5 (water, at radio frequencies of the order of $1\,$GHz or smaller used in the measurements). Unfortunately, the uncertainties of the effective permittivity measured using
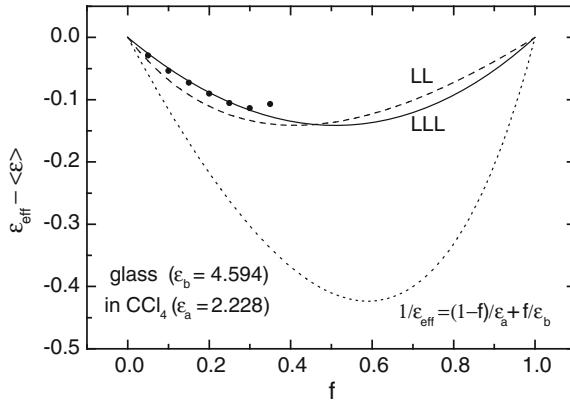
Fig. 3.10 Difference between the effective permittivity of several mixing rules and the volume average of the constituents; LL (*dashed line*) denotes the Landau-Lifshitz formula of Eq. (3.12), LLL (*solid line*) its approximation of Eq. (3.14). Symbols: the difference of measured data from the volume average

the time-domain reflectometry are rather large, the estimate of accuracy quoted in Ref. [10] is ± 0.1. We have digitized the experimental points from Figs. 3.3 and 3.4 of Ref. [10]; they are shown together with the predictions of several mixing rules in Fig. 3.11. Interestingly, the measured points are fairly close to the Maxwell Garnett model for large values of ε_a. A plausible explanation of this fact relies on the averaging procedure of Eq. (3.31), which neglects the fields generated by the surface charges of the spheres. In fact, the large polarizability of the liquids should tend to reduce the contribution of these fields to the average intensity and displacement.
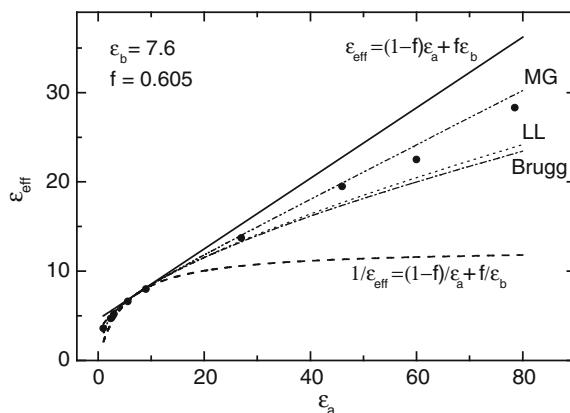

Fig. 3.11 Effective permittivity of glass spheres in different liquids. Experimental data (*symbols*), and several mixing rules (*lines*)
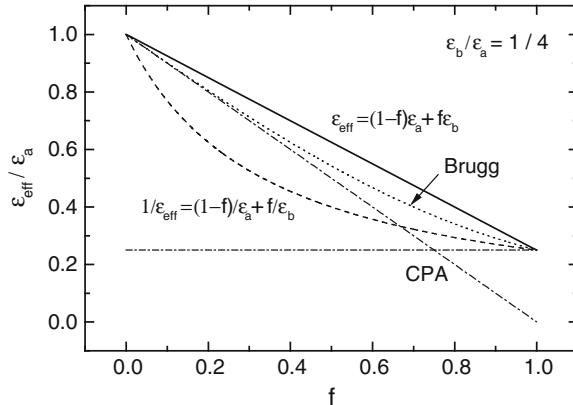
Fig. 3.12 Effective permittivity of several EMA mixing rules for $\varepsilon_b = \varepsilon_a/4$. The *dash-dotted straight lines* are the predictions of the CPA formula of Eq. (3.38)

We have omitted the CPA mixing rule (which provided the best representation of the glass-CCl$_4$ mixture discussed above) from the comparison of Fig. 3.11. The reason is its failure for large values of the ratio $\varepsilon_a/\varepsilon_b$. An indication of peculiar behavior is linked to the possibility of vanishing denominator in the fraction of Eq. (3.38) for $\varepsilon_b > \varepsilon_a$, or the indefinite expression for the volume fraction f resulting for $\varepsilon_{eff} = \varepsilon_a/4$ (see the last row of the table in paragraph 4). Shown in Fig. 3.12 are the two values of the CPA rule of Eq. (3.38) obtained for $\varepsilon_b = \varepsilon_a/4$. One of the roots of the quadratic Eq. (3.39) is $\varepsilon_{eff} = \varepsilon_a/4$; the remaining one is $\varepsilon_{eff} = \varepsilon_a(1-f)$, coinciding with the linear expansion of Eq. (3.40). Thus, the latter root is usable for small volume fractions f, while it even runs out of the Wiener bounds for larger values of f.

The range of the smaller polarizabilities of the immersion liquids is shown on expanded scales in Fig. 3.13. The CPA, Bruggeman and Landau-Lifshitz predictions are fairly close to one another, while the Maxwell Garnett rule deviates from the three. The measured points lie slightly below the lower Wiener bound for the two largest values of ε_a; the increase with increasing ε_a seems to be closer to that of the CPA, Bruggeman and Landau-Lifshitz models than to the Maxwell Garnett model.

3.6 Testing EMA: Water Solutions of Sucrose

An interesting mixture suitable for the investigation of its optical properties is the water solution of sucrose, as suggested by Feynman in his Lectures [11]. The sucrose molecules, $C_{12}H_{22}O_{11}$, remain stable in aqueous solutions at ordinary temperatures. Further, due to their importance in food industry, the solutions have been measured
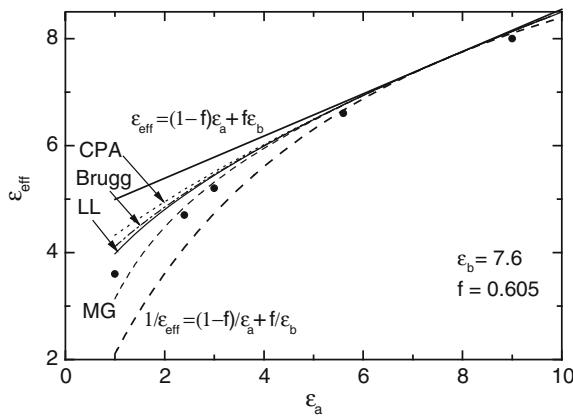
Fig. 3.13 Effective permittivity of glass spheres in different liquids for smaller polarizability of the liquids. Experimental data (*symbols*), and several mixing rules (*lines*)

extensively; reliable, comprehensive datasets in saccharimetry are readily available [12].

Before discussing the optical properties of the solutions, we estimate the volume fractions of their constituents based on the density (mass per unit volume) measurements. The number of molecules in 1 mol of matter is the Avogadro number $N_A = 6.02214 \times 10^{23}$ mol^{-1}. The mass of one solvent (water) molecule and one solute (sucrose) molecule is $m_w = 18.02/N_A$ g and $m_s = 342.3/N_A$ g, respectively. The solution is formed by mixing the masses $M_w = N_w m_w$ and $M_s = N_s m_s$ of the two molecular species, where N_w and N_s are the corresponding numbers of the molecules. The density of the solution is

$$\rho = \frac{N_w m_w + N_s m_s}{V} = c_w m_w + c_s m_s, \tag{3.44}$$

where c_w and c_s are the concentrations of the constituents (i.e., the numbers of molecules per unit volume). Let us assume the total volume V is shared by the corresponding apparent volumes, V_w and V_s, of the constituents, $V = V_w + V_s$. We assume further the apparent volume of the solvent molecule to be that of the pure solvent having the density ρ_w (this assumption is the better the more dilute the solution is). Then, we can rewrite Eq. (3.44) in the following form

$$V_s = \frac{N_w m_w + N_s m_s}{\rho} - \frac{N_w m_w}{\rho_w}. \tag{3.45}$$

Introducing the mass fraction of the solute (a convenient measure of the composition, independent of temperature and pressure),
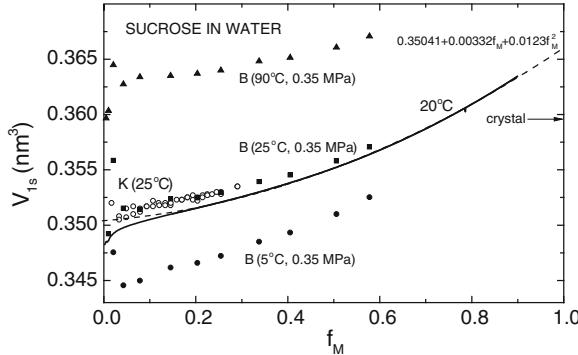
Fig. 3.14 Apparent volume of the sucrose molecule as a function of the mass fraction in water solution, obtained from the density. Tabular data for 20° C of Ref. [12], pp. 343–344 (*thick solid line*); its quadratic approximation (best fit in the range from 0.15 to 0.9, *dashed line*). *Closed symbols* data of Ref. [13] for different temperatures and the pressure of 0.35 MPa. *Open symbols* data of Ref. [14] for 25° C. The *arrow* shows the volume per sucrose molecule in a single crystal at 20° C

$$f_M = \frac{N_s m_s}{N_w m_w + N_s m_s}, \tag{3.46}$$

we use Eq. (3.45) to obtain the apparent volume of one solute molecule:

$$V_{1s} \equiv \frac{V_s}{N_s} = \left(\frac{1}{\rho} - \frac{1 - f_M}{\rho_w} \right) \frac{m_s}{f_M}. \tag{3.47}$$

Shown in Fig. 3.14 is the apparent volume of the sucrose molecule as a function of the mass fraction, obtained from the density data from several sources. We have used the extensive tabulation of Ref. [12] to calculate the dependence shown by the thick solid line in Fig. 3.14. The results are in a fair agreement with the individual data points from two other sources, taking into account the higher temperature of 25° C. As expected, the determination of the apparent volume is rather poor at small mass fractions, when the density of the solution approaches that of the solvent; two close values are subtracted in the parentheses of Eq. (3.45) which magnifies both random and systematic errors.

The increase of the apparent volume with increasing mass fraction indicates the formation of voids adjacent to the sucrose molecules, small enough to prevent filling with water, see the space filling model in Fig. 3.15. The presence of voids should be easily detectable in the optical response of the solutions, provided the dielectric response of water and sucrose molecules was only weakly dependent on the composition of the solution. This assumption has been tested by Feynman [11] using the representation of the water and sucrose molecules by point dipoles, contributing (via their polarizabilities) to the refractive index of the mixture. We will reformulate the problem in terms of the standard EMA picture: the solution is described by the
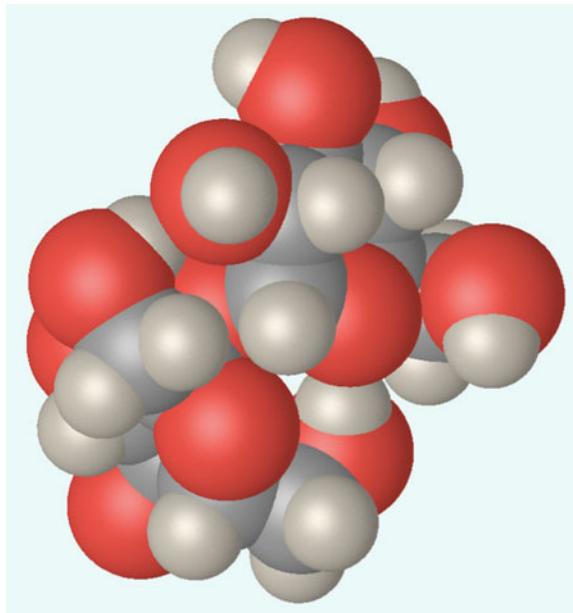
Fig. 3.15 Space filling model of the sucrose $C_{12}H_{22}O_{11}$ molecule, from [15]

three-component mixture with the volume fractions f_w, f_s, and f_v, and permittivities $\varepsilon_w, \varepsilon_s$, and $\varepsilon_v = 1$ of water, sucrose and voids, respectively. We will use the tabulated dependence of the refractive index (i.e., the square root of the permittivity) at the sodium line $\lambda_D = 589.3$ nm at 20 C, covering the compositions from pure water ($f_M = 0$) to $f_M = 0.85$ (85 weight percent of sucrose) [16]. The permittivity of solution increases monotonically from 1.77686 (pure water) to 2.26196 at $f_M = 0.85$, implying the value for the sucrose component in EMA to be below about 2.5; thus, the contrast of the dielectric constants (that of vacuum, water and sucrose) is fairly small, which suggests using the LL or LLL mixing rules of Eqs. (3.12) or (3.14). This choice is further substantiated by the complex geometry of the elongated sucrose molecules and the free volume within their clusters (the averages of the optical fields in the LL approach are independent of the geometry).

Using the LLL mixing formula of Eq. (3.14) for the resulting permittivity ε, we arrive at the following formula for the volume fraction of voids,

$$f_v = \frac{\varepsilon_s^{1/3} - \varepsilon^{1/3}}{\varepsilon_s^{1/3} - 1} - f_w \frac{\varepsilon_s^{1/3} - \varepsilon_w^{1/3}}{\varepsilon_s^{1/3} - 1}. \qquad (3.48)$$

The volume fraction of water is provided by the density ρ, using the same reasoning as that in deriving Eq. (3.45):
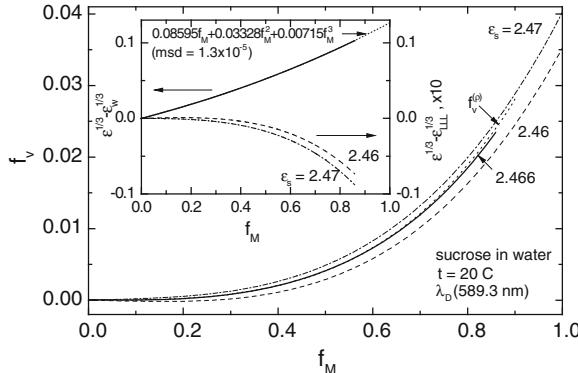
Fig. 3.16 Volume fraction of voids in water solution of sucrose obtained using Eq. (3.48), using ε_s of 2.466 (*solid line*), 2.46 (*dashed line*) and 2.47 (*dash-dotted line*). Dotted line: volume fraction of voids from density, Eq. (3.50). Inset: the difference of the cube root of the dielectric constant from that of pure water (*solid line*), and its polynomial approximant (*dotted line*); magnified (10 times) differences of measured data from the LLL rule for two values of the dielectric constant of sucrose in a two-component mixture (no voids), *dashed* and *dash-dotted lines*

$$f_w = \frac{\rho}{\rho_w}(1 - f_M). \tag{3.49}$$

The only unknown quantity in Eq. (3.48) is the permittivity ε_s, describing the dielectric response of a continuum of sucrose molecules in a hypothetical mixture with no water and no voids. Its value results readily from the plausible requirement of the zero slope of the $f_v(f_M)$ dependence for $f_M \to 0$, which occurs for $\varepsilon_s = 2.466$. This value is larger than the average of the three principal components of the dielectric tensor of the biaxial sucrose crystal (2.427; note that Feynman [11] uses the average of refractive indices to estimate the response of the sucrose component of the solution). Shown in Fig. 3.16 is the volume fraction of voids resulting from Eq. (3.48), using the refractive index (i.e., the square root of the dielectric constant) at 20 C and the wavelength of 589.3 nm given in Ref. [16]. The measured data are represented in the inset of Fig. 3.16 together with its cubic polynomial approximation; the mean square deviation of the two is 1.3×10^{-5} in the whole experimental range of the mass fractions f_M from zero to 0.85. Consequently, we use the polynomial also for the extrapolation of experimental data in order to cover the whole compositional range by the LLL model with ε_s equal to 2.46 (dashed line in Fig. 3.16, negative slope of f_v at $f_M = 0$) and 2.47 (dash-dotted line in Fig. 3.16, positive slope of f_v at $f_M = 0$).

The void fraction can also be estimated from the density ρ of the solution, using the fixed volume of one sucrose molecule from Eq. (3.47) for $f_M \to 0$ throughout the composition range, and attributing the increase of the apparent volume to the presence of the voids. Assuming once more the density of water independent of the composition, we arrive at

$$f_v^{(\rho)} = 1 - \rho \left[\frac{1}{\rho_w} - f_M \left(\frac{1}{\rho_w} - \frac{V_{1s}}{m_s} \right) \right]. \tag{3.50}$$

The resulting void fraction shown in Fig. 3.16 is in a very good agreement with that obtained from the refractive index. Thus, the simple treatment of the packing of the water and sucrose molecules, neglecting the changes of bond lengths and optical polarizabilities with the composition, seems to provide consistent results.

Let us note that the assumption of averaging the microscopic fields of point dipoles using the Clausius-Mossotti relation [11] is very good for dilute solutions. In fact, the conceptually similar treatment of EMA described above results in the volume fraction of voids smaller than 0.001 for the mass fraction of sucrose less than 0.32. On the other hand, the presence of voids is clearly indicated in denser solutions, and discrepancies of a few percent are observed in the EMA description of the dielectric function when using the simplest two-component picture.

The concentrated sucrose solutions were studied by molecular dynamics simulations in [17]. These calculations suggested a non-monotonic dependence of the free volume fraction on the compositions, which is absent in the data of Fig. 3.16 resulting from both density and refractive index. A further notice concerns the temperature dependence of Fig. 3.14: the increase of the free volume with increasing sucrose fraction is steeper at higher temperatures. This indicates the role of dynamical effects in filling the voids between clustered sucrose molecules with water.

Of course, EMA has to be used with caution. Substantial deviations from the simple mixing are expected in the spectral range of strong absorption (of both sucrose and water); the spectral shifts of molecular absorption bands due to the changes of environment of each molecule when changing concentration would be absent in the model spectra. The simplest EMA approach would also fail if it was used in treating the light scattering by the solutions [18]. However, the predictions of EMA described above are very good in predicting the refractive index in the transparent range, in spite of the very small size of the sucrose molecule ($V_{1s} \sim 0.35\,\text{nm}^3$ at room temperature).

3.7 Differences Between Mixing Rules for Binary Dielectric Mixtures

The choice of a mixing rule in a particular situation need not be obvious. It is therefore desirable to estimate the differences between various plausible possibilities. This is easily accomplished in the simplest case of binary mixtures possessing real components of permittivity. Different rules provide differing compositional dependences of the effective permittivity, interpolating between the two constituents. The simplest linear interpolation is exact (in the long-wavelength limit) for layers parallel to the electric intensity, the perpendicular orientation is properly described by addi-
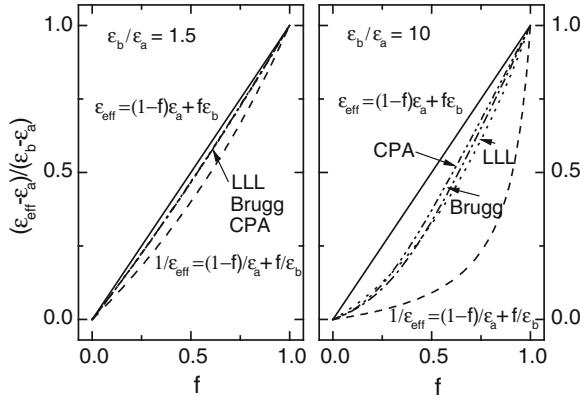
Fig. 3.17 Normalized difference of the effective permittivity of binary mixtures from that of one of the components, ε_a, for $\varepsilon_b = 1.5\,\varepsilon_a$ (*left panel*) and $\varepsilon_b = 10\,\varepsilon_a$ (*right panel*). The mixing rules: Eq. (3.25), *solid line*; Eq. (3.27), *dashed line*; Eq. (3.14), *dotted line*; Eq. (3.36), *dash-dotted line*; Eq. (3.38), *dash-dot-dotted line*

tive inverse permittivities of Eq. (3.27). These two limiting cases are usually called *Wiener bounds* for two-phase mixtures of any microstructure.

The Wiener bounds and the effective permittivities of three selected mixing rules are shown in Fig. 3.17 for two values of $\varepsilon_b/\varepsilon_a$. We have chosen the CPA and Bruggeman rules for spherical inclusions, and the LLL formula for a small difference of the permittivities. For the smaller contrast, the bounds are fairly close and the three rules are indistinguishable on the scale of the figure. However, the large relative value of ε_b in the right panel of Fig. 3.17 shifts the lower bound rather significantly, since the smaller value ε_a becomes more important. In addition, the remaining mixing rules deviate markedly from each other. It should be reemphasized that the LLL rule has been derived for small values of $\varepsilon_b/\varepsilon_a$; we have included it here in order to draw the attention to a possible "accidental" success of any mixing rule. For example, the predictions of CPA and LLL rules coincide for $f \approx 0.29$ for the ratio of permittivities as high as 10.

The differences between the selected mixing rules are better seen on expanded (logarithmic) scales of Fig. 3.18. The difference between the LLL and CPA rule (left panel of Fig. 3.18) crosses zero at $f \approx 0.29$ in a broad range of the dielectric contrasts. At the same time, the CPA and Bruggeman rules agree within less than one percent of the difference between the constituents.

All of the tested rules agree within a few percent of the difference between the permittivities of the components when the latter does not exceed 10. Moreover, the mutual agreement is much better for dilute mixtures (small values of f); this is expected since the averaged fields become less sensitive to the properties of inclusions and their interactions.
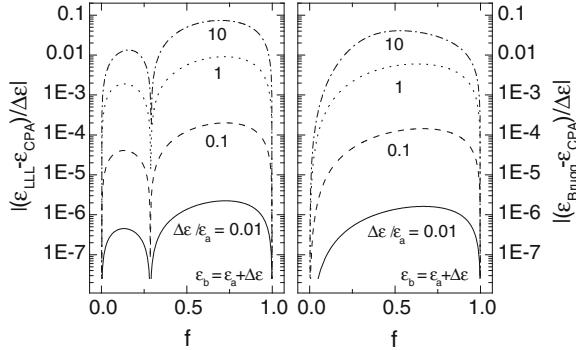
Fig. 3.18 Normalized difference of the effective permittivities of binary mixtures calculated from the rules used in Fig. 3.17: LLL and CPA (*left panel*), Bruggeman and CPA (*right panel*)

3.8 Exact Solutions and EMA for Layered Structures

Layered structures of the type shown in Fig. 3.2 represent a convenient system for evaluating quantitatively the approach of effective medium. Namely, the optical fields can be computed exactly for incident plane waves using the scheme of transfer matrices [19]; moreover, analytical results are available for derivatives of the field amplitudes and ellipsometric quantities, enabling a very efficient fitting of measured data [20]. We show here examples of exact results and their approximation using effective medium approach. Let us start with the positional dependence of the field intensity
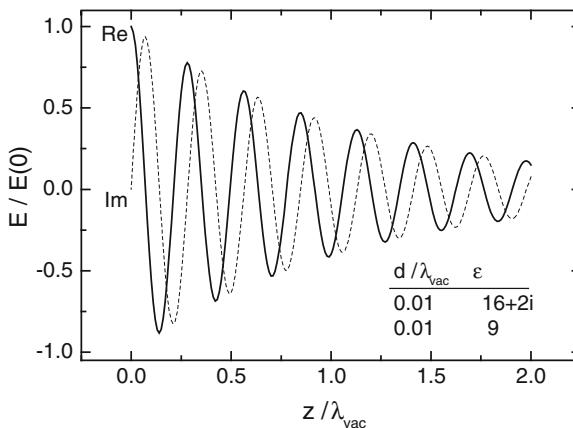

Fig. 3.19 Complex amplitude of the electric intensity of a plane wave propagating along the z axis in the layered system consisting of repeated pairs of films specified by the thickness d and the dielectric function ε

shown in Fig. 3.19. The wave, travelling in the z direction, is damped due to the absorption in one of the layers in the repeated pairs. The electric intensity is parallel to the interfaces, i.e., the angle of incidence in Fig. 3.2 is zero. The exact solution is composed from linear combinations of plane waves in each of the layers, with the electric intensity continuous across the interfaces and the electric displacement undergoing step-like changes. The appropriate effective permittivity is the simple volume average with no screening, Eq. (3.25). The damped plane wave solving the propagation in the effective medium is indiscernible from the exact solution at the scale of Fig. 3.19. The reason is the smallness of the thickness of individual layers with respect to the wavelength of the optical field. The latter is about λ_{vac}/n_{eff}, with the effective refractive index of ~ 3.54 resulting from the volume average of the two permittivities of the mixture. The detailed comparison of the exact and approximate fields shown in Fig. 3.20 reveals a subtle relative difference of the order of 10^{-3}. Even this small difference could be observed via differences in interference patters observed in light reflected from or transmitted by a film of suitable thickness made from this model metamaterial. On the other hand, the differences diminish with decreasing d/λ ratio.

This kind of differences of the inner fields in the metamaterial becomes fairly easily observable in ellipsometric measurements. We show in Fig. 3.21 the spectra of the standard ellipsometric angles computed for different values of d/λ_{vac}, where d is the total thickness of the layered metamaterial of Fig. 3.19. We have used the repeated pairs of the $\lambda_{vac}/1000$ films of the two different dielectric functions, kept constant throughout the calculation of the spectra, calling the metamaterial a "superlattice" (SL). The substrate has been chosen to be the material with the complex permittivity $16 + 2i$, the more polarizable component of the mixture. Because of the large
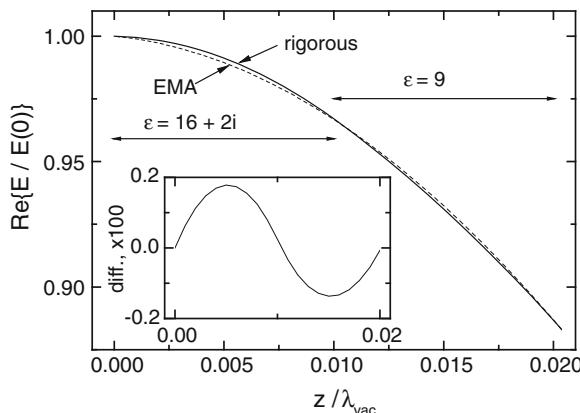

Fig. 3.20 A detail of the electric intensity of Fig. 3.19 obtained by solving rigorously the wave equation in the film stack (*solid line*) and in the appropriate effective medium (*dashed line*). The extent of neighboring individual layers is indicated by the *horizontal arrows*. The *inset* shows the difference between the two intensities on an expanded vertical scale
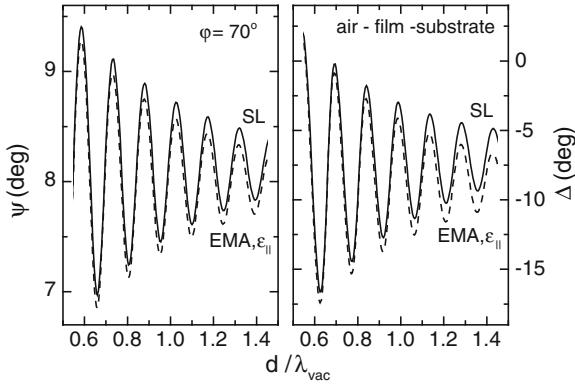
Fig. 3.21 Ellipsometric angles, (ψ, *left panel*) and (Δ, *right panel*), computed for a layer of thickness d, made of the metamaterial of Fig. 3.19 with $d_1 = d_2 = \lambda_{\text{vac}}/1000$, at the angle of incidence of $70°$. *Solid line* exact solution; *dashed line* effective medium approximation, using the isotropic dielectric function along the interfaces

refractive index of the metamaterial, the optical wave is refracted very close to the z direction, and the usual approach is to neglect the anisotropy and use the parallel component of the dielectric function for the appropriate effective medium. The exact and EMA calculations produce similar interference patterns, displaying a pronounced decrease of the amplitude with increasing ratio d/λ_{vac}, due to the absorption in the metamaterial. The EMA calculation deviates from the exact result for the following two reasons: the neglected anisotropy, and the finite values of d_1 and d_2. A closer look at the increase of the differences for decreasing wavelength identifies the second reason as the decisive one.

In order to identify potential problems caused by a too coarse structure of the mixture compared to the wavelength, we have performed the exact calculations of the ellipsometric spectra for different thicknesses of the constituent bilayers, keeping the total thickness d of the metamaterial fixed. In other words, we have used the appropriate number of repetitions of the basic bilayer motif, $d/(d_1 + d_2)$. The results shown in Fig. 3.22 demonstrate a fairly high sensitivity of ellipsometry to the fineness of the mixture; as usual, the phase shift Δ is more sensitive. Note that considerable changes from the behavior of a "true mixture" occur for the individual layer thicknesses well below one percent of the wavelength in the effective medium.

The dramatic changes seen for thicker constituent layers are due to the interferences in the stack of layers with the high and low index of refraction. In fact, the thickness of individual layers close to $\lambda/4$ leads to the appearance of Bragg-like bands of increased reflectivities and corresponding changes in the relative phase shifts.

We can readily transform these model calculations into practical guidelines for the applicability of the effective medium (i.e., continuum) approximation of the layered metamaterial. Namely, at the representative vacuum wavelength of the visible range,
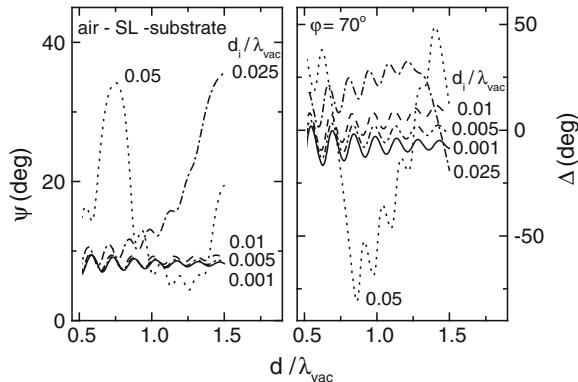
Fig. 3.22 Ellipsometric angles, (ψ, *left panel*) and (Δ, *right panel*), computed for a layer of fixed thickness d, made of the metamaterial of Fig. 3.19 with different values of $d_1 = d_2$, at the angle of incidence of 70°

500 nm, we observe easily detectable (at least one tenths of a degree in ψ, and one degree in Δ) changes for the individual layer thicknesses as low as 2 nm in this particular metamaterial. Of course, this value scales with the wavelength in the mixture.

This level of sensitivity is related to a possible detection of surface or interface roughness in ellipsometric measurements. In fact, the representation of the surface roughness as the presence of an overlayer composed of the topmost material and voids is usually a plausible approximation explaining the measured data. Similarly, the interface roughness can be modeled by inserting a thin layer of mixed composition. These transition layers are graded; due to the usually small extent of the grading, the approximation by a single homogeneous layer of intermediate composition is suitable. However, without independent information concerning the geometry of these mixtures, values of film thicknesses and compositions derived from measured data should be used with care.

3.9 Resonant Behavior of EMA Mixtures

When selecting properly the individual contributions of the constituents of an EMA mixture, we can arrive at a spectacular behavior of the effective optical response; the latter can be traced down to the spectacular behavior of local fields. We outline here the simple cases having analytical solutions, based on the constant field inside an isolated ellipsoid in an infinite host medium as described by Eq. (3.29). Namely, the field intensity inside the ellipsoid becomes infinite whenever

$$\frac{\varepsilon_b}{\varepsilon_a} = 1 - \frac{1}{L_u} \in (-\infty, 0\rangle. \tag{3.51}$$

According to the value of the depolarization factor L_u of Eq. (3.30), this divergence occurs in particular in a

- sphere: $u = v = w$, $L_u = 1/3$, for $\varepsilon_b = -2\varepsilon_a$;
- cylinder, field perpendicular to its axis: $u = v$, $w \to \infty$, $L_u =1/2$, for $\varepsilon_b = -\varepsilon_a$;
- slab, field perpendicular to the interface: $v = w \to \infty$, $L_u = 1$, for $\varepsilon_b = 0$.

To fulfill this condition, the permittivity of the inclusion has either to vanish or to be of the opposite sign as that of the host. Since the response functions are frequency dependent, the diverging (in practice, very large) field intensities can only occur in narrow spectral ranges, i.e., they exhibit resonant behavior. Let us note that the above condition for the slab is related to the occurrence of the surface plasmon resonance, achieved by using a thin metallic film and the obliquely incident, p-polarized wave; the enhanced fields occur for $\varepsilon_b \approx 0$ in the metal.

Assuming a dilute mixture (the volume fraction $f \ll 1$) of aligned ellipsoids, the Maxwell Garnett formula of Eq. (3.32a) leads to a diverging dielectric function for

$$\frac{\varepsilon_b}{\varepsilon_a} = 1 - \frac{1}{L_u(1 - f)}, \tag{3.52}$$

which is close to the condition of diverging field of Eq. (3.51). In the case of flat interfaces perpendicular to the field, the effective dielectric function can be expanded in the form

$$\varepsilon_{eff} \approx \varepsilon_a(1 + f) - \frac{\varepsilon_a^2}{\varepsilon_b} f, \tag{3.53}$$

leading to a divergence whenever ε_b crosses zero (in practice, its imaginary part being small and the real part crossing zero). The spectral lineshape of the resonance, proportional to negative inverse of ε_b, is almost independent of f; it is multiplied by the squared permittivity of the host material, which is large for highly polarizable materials.

The resonance condition of Eq. (3.52) is actually exact for the field perpendicular to flat interfaces (i.e., $L_u = 1$). In fact, the effective dielectric function of Eq. (3.27) diverges for

$$\frac{\varepsilon_b}{\varepsilon_a} = -\frac{f}{1 - f} \tag{3.54}$$

at any volume fraction f. With equal volume fractions of one half (i.e., equal film thicknesses), the resonance occurs for $\varepsilon_b = -\varepsilon_a$. The opposite signs of the per-mittivitites have been achieved by a proper doping of one of the components of a semiconductor superlattice [21], as shown in Fig. 3.23. The undoped AlInAs alloy possesses an (almost constant) positive permittivity ε_a throughout the mid-infrared

50x(80nm i-AllnAs, 80nm n-InGaAs)
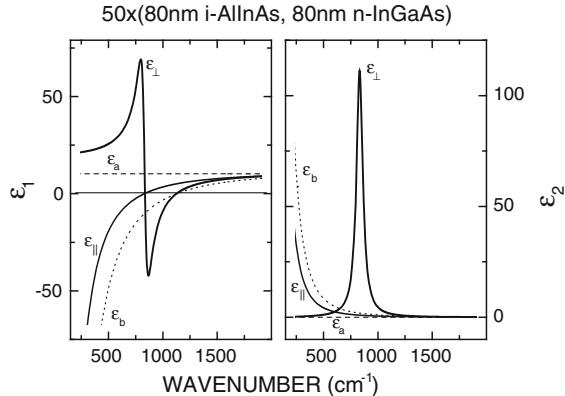
Fig. 3.23 Real (*left panel*) and imaginary (*right panel*) parts of the model dielectric functions of the layered heterostructure formed by alternating components (insulating, ε_a) and (a single Drude term on a constant real background, ε_b) of equal thicknesses. The effective dielectric functions for the electric field parallel and perpendicular (*solid lines*) to the film interfaces

(MIR) range, while the n-type InGaAs alloy has the negative-valued real part of its permittivity, ε_b, up to the wavenumber of about 1130 cm^{-1}, due to the Drude-like contribution of free electrons. The model lineshapes of Fig. 3.23 are based on ellipsometric measurements [22] performed on a superlattice layer (lattice-matched to its InP substrate), exhibiting negative refraction in MIR [21]. With equal film thicknesses, the effective medium response for the field perpendicular to the interfaces displays a strong resonance at 832 cm^{-1}, and a band of negative real part of permittivity from 836 to 1125 cm^{-1}.

Another example of the resonance governed by Eq. (3.54) is the "transverse plasmon" in the c-axis response of superconducting cuprates [23]. In this case, the resonant behavior observed in the far-infrared range is more damped due to the energy losses of normal-state carriers.

The averaging procedure pertinent to the EMA treatment of finely structured metamaterial is instructive also for the strongly anisotropic layer stack of Fig. 3.23. In fact, it explains in simple terms the effect of negative refraction occurring in a fairly broad range of wavenumbers [22]. The averaging procedure is shown in detail for a selected pair of undoped (positive permittivity) and doped (negative permittivity) layers. With an arbitrary selection of the electric field intensity in the former layer, we find all of the intensities and displacements as follows:

$$D_{ax} = \varepsilon_a E_{ax}, \quad D_{az} = \varepsilon_a E_{az} \quad \text{(isotropic material a)},$$
$$E_{ax} = E_{bx}, \qquad D_{az} = D_{bz} \qquad \text{(crossing the interface)},$$
$$D_{bx} = \varepsilon_b E_{bx}, \quad E_{bz} = D_{bz}/\varepsilon_b \quad \text{(isotropic material b)}.$$

Further, the averaging assuming the thicknesses d_a, d_b, $d = d_a + d_b$, and the volume fractions $f_a = d_a/d$, $f_b = d_b/d = 1 - f_a$, leads to

$$\langle E_x \rangle = f_a E_{ax} + f_b E_{bx} = E_{ax},$$
$$\langle E_z \rangle = f_a E_{az} + f_b E_{bz} = (f_a + f_b \varepsilon_a/\varepsilon_b) E_{az},$$
$$\langle D_x \rangle = f_a D_{ax} + f_b D_{bx} = (f_a \varepsilon_a + f_b \varepsilon_b) E_{ax},$$
$$\langle D_z \rangle = f_a D_{az} + f_b D_{bz} = D_{az} = \varepsilon_a E_{az},$$

i.e., to the components of the dielectric tensor of the resulting effective medium:

$$\langle D_x \rangle = \varepsilon_{||} \langle E_x \rangle = (f_a \varepsilon_a + f_b \varepsilon_b) \langle E_x \rangle,$$
$$\langle D_z \rangle = \varepsilon_\perp \langle E_z \rangle = \frac{1}{f_a/\varepsilon_a + f_b/\varepsilon_b} \langle E_z \rangle.$$

This averaging procedure assumes constant fields within each layer, in other words, film thicknesses negligible compared with the wavelength. The average flow of energy, shown also in Fig. 3.24, occurs in the "negative" direction.

The description of the doped multilayer within the EMA framework has been tested by performing MIR ellipsometric measurements [22]. The good agreement of the measured and model spectra (no fitting, the nominal layer thicknesses and a guess of
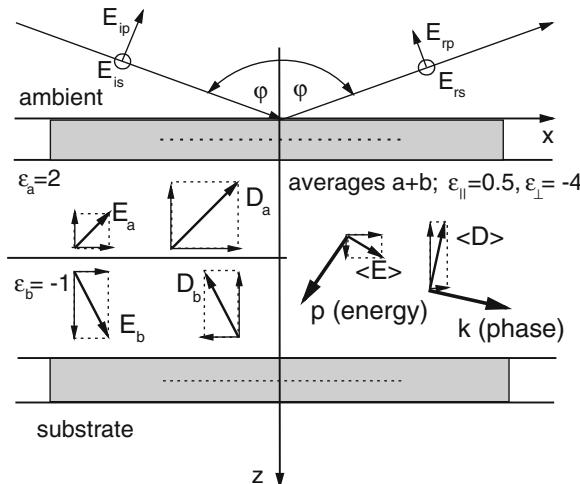

Fig. 3.24 Electric field (E) and displacement (D) amplitudes of a plane wave inside a pair of selected neighboring layers (of equal thicknesses) of the anisotropic layered material, differing by the sign of their real permittivities (*left part*). The *right part* volume averages of the fields, the directions of the propagation of the constant-phase planes (k) and the flow of energy (Poynting vector, p), perpendicular to $\langle D \rangle$ and $\langle E \rangle$, respectively, are shown by the *thick arrows*
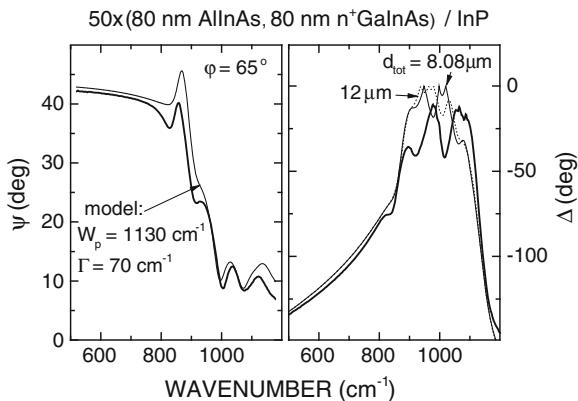
Fig. 3.25 Measured (*thick solid lines*) and model (*thin solid lines*) spectra of the ellipsometric angle ψ (*left panel*) and Δ (*right panel*) at the angle of incidence of $65°$. The model calculation used a single Drude term with the indicated parameters. An additional model spectrum of Δ was calculated for the total thickness of the metamaterial layer of $12\,\mu m$ (*dotted line*)

the Drude parameters of the doped material have been used in the model) confirms the usefulness of the EMA picture.

The individual layer thicknesses of 80 nm are small compared with the wavelength if MIR light in the metamaterial. On the other hand, the interference pattern with the period of about $100\,cm^{-1}$ results from the coherent reflections within the total thickness of 8080 nm of the superlattice. We have tested this interpretation by calculating the model spectra of Δ for two values of the total thickness of the metamaterial layer. Note that the spectra of Δ are folded to the range from -180 to $0°$ (due to the rotating-analyzer measurement setup, which is unable to determine the sign of Δ). The denser interference pattern obtained for the larger total thickness confirms our assumption (Fig. 3.25).

3.10 Conclusions

The approximate treatment of mixtures using effective medium approach is simple and attractive, as it can capture important properties of nanostructured materials. We would encourage using it, whenever the underlying assumptions are fulfilled. The applicability of any specific EMA formula should be assessed carefully, and the level of uncertainties estimated.

It might be surprising to find a consistent EMA picture of the water solutions of sucrose in the visible spectral range, remembering the apparent volume of one sucrose molecule of about $0.35\,nm^3$. Further, the individual layers in a semiconductor superlattice might be as thick as $\sim 0.1\,\mu m$ and still form a proper component in the EMA

continuum in the mid-infrared range, and even provide a plausible explanation of the negative refraction. We could easily find a failure of the EMA models for these systems in other circumstances; however, this would not be a defect of the approach, but that of its improper use.

References

1. J.C. Maxwell, *A Treatise on Electricity and Magnetism*, vol. I, 3rd edn. (Clarendon Press, Oxford, 1904), Sect. 314
2. L.D. Landau, E.M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd edn. (Pergamon Press, Oxford, 1984), Sect. 9
3. H. Looyenga, Physica **31**, 401 (1965)
4. J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962)
5. J.A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941)
6. J.C.M. Garnett, Philos. Trans. R. Soc. Lond. **203**, 385 (1904)
7. D.E. Aspnes, Am. J. Phys. **50**, 704 (1982)
8. D.A.G. Bruggeman, Ann. Phys. (Leipzig) **24**, 636 (1935)
9. A. Sihvola, *Electromagnetic Mixing Formulas and Applications* (IEE, Stevenage, 1999)
10. D.A. Robinson, S.P. Friedman, J. Non-Cryst. Solid **305**, 261 (2002)
11. R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures on Physics*, vol. II (Addison-Wesley, Boston, 1964), Sect. 32–5
12. *International Critical Tables of Numerical Data, Physics, Chemistry and Technology*, vol. 2, First Electronic Edition (Knovel, Norwich, New York, 2003), pp. 334–355
13. B.R. Brown et al., J. Chem. Thermodyn. **37**, 843 (2005)
14. K. Kiyosawa, Bull. Chem. Soc. Jpn. **61**, 633 (1988)
15. http://www.crystmol.com/structures/Sucrose_SpaceFilling.jpg
16. www.icumsa.org, ICUMSA Method SPS-3 (2000)
17. V. Molinero et al., Chem. Phys. Lett. **377**, 469 (2003)
18. M. Kazsuba et al., J. Nanopart. Res. **10**, 823 (2008)
19. F. Abeles, Annls Phys. **5**, 596 (1950)
20. J. Humlíček, Optica Acta **30**, 97 (1983)
21. A.J. Hoffman et al., Nat. Mater. **6**, 946 (2007)
22. J. Humlíček, Thin Solid Films **519**, 2655 (2011)
23. J. Chaloupka, D. Munzar, J. Humlíček, *Advanced mueller ellipsometry instrumentation and data analysis*, ed. by M. Losurdo, K. Hingerl, Ellipsometry at the nanoscale

Chapter 4
Relationship Between Surface Morphology and Effective Medium Roughness

Angel Yanguas-Gil and Herbert Wormeester

Abstract The modeling of surface and interface roughness is a key issue in the interpretation of ellipsometric measurements. Materials properties are often extracted from ellipsometry measurements in an indirect way by modeling the optical response of the material. Since roughness is known to affect the scattering of light on an interface, how roughness is incorporated into these models can affect the outcome of the fitting procedure.

4.1 Introduction

Many practical approaches used to model the optical response of a material rely on Fresnel's reflection and refraction relations. These describe the effect experienced by an incident electric field at a mathematically sharp planar interface between two media with different refractive index. However, roughness poses a clear violation of the concept of a mathematical sharp planar interface, see also Fig. 4.1. The statistical properties of the surface topography can be used to quantify the deviation of the actual surface from the planar interface approximation. Two commonly used parameters are the standard deviation of the height distribution, also known as the root mean square (rms) roughness, and the lateral correlation length scale ξ, the distance above which the heights of two points are statistically uncorrelated. In many practical applications, both rms and lateral correlation length are much smaller than the wavelength of light. This is known as the microscopic roughness regime.

A. Yanguas-Gil (✉)
Argonne National Laboratory, Argonne, IL 60439, USA
e-mail: ayg@anl.gov

H. Wormeester
Physics of Interfaces and Nanomaterials, MESA + Institute for Nanotechnology,
University of Twente, Enschede, The Netherlands
e-mail: h.wormeester@utwente.nl
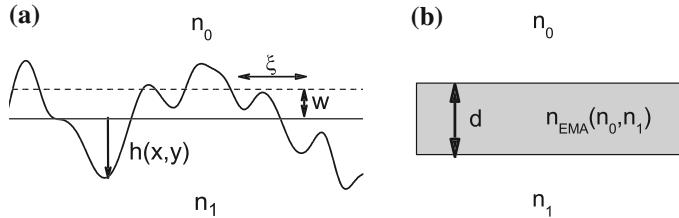
Fig. 4.1 **a** Schematic of interface roughness profile between two media with refractive index n_0 and n_1. The average interface position is indicated by the *solid line*. The local deviation from this is the height h(x,y). Also indicated are the statistical quantities w and ξ. **b** Modeling of roughness by the introduction of a layer of thickness d and an effective refractive index n_{EMA}

In a seminal work by Aspnes et al. [3], the use of an intermediate layer to model the contribution of interface roughness to the optical response of the surface was evaluated and compared with experimental results. This concept is pictorially depicted in Fig. 4.1. The introduced intermediate layer has a thickness d_s that is some measure for the average height variation, while the dielectric properties of this layer are a mix of the two media on either side of the rough interface. The introduction of this intermediate layer creates again two sharp interfaces which allows to use the full Fresnel formalism to evaluate the optical response. An empirical evaluation of several mixing rules for dielectric media, so called effective medium approximations (EMA) was made [3]. These included the Lorentz-Lorenz model, the Maxwell Garnett model [23] and Bruggeman's EMA for spherical entities [6]. They concluded that Bruggeman's EMA consistently provided the best agreement for the materials considered, which include both semiconductor and metal interfaces. Since then, BEMA has been routinely used to take into account the influence of roughness on the optical properties of a film.

BEMA is Bruggeman's solution to a long standing problem of a self-consistent effective medium model. In the most widely used variety of Bruggeman's analysis, two spherical entities are mixed. The size of these spherical entities is much smaller than the wavelength of light. The self-consistent nature of this model is obtained by embedding the spherical particles of different refractive index n_0 and n_1 in a material with an effective refractive index n_{EMA}:

$$f_v \frac{n_0^2 - n_{EMA}^2}{n_0^2 + n_{EMA}^2} + (1 - f_v) \frac{n_1^2 - n_{EMA}^2}{n_1^2 + n_{EMA}^2} = 0 \qquad (4.1)$$

where f_v fraction of species 0 in the mixed layer. The effective refractive index of this intermediate layer is used to represent the dielectric properties of the surface roughness layer.

The EMA model for surface roughness provides two parameters to characterize roughness: the layer thickness d_s and the fraction f_v. However, because the roughness layer is usually quite thin, these two parameters are strongly correlated. Therefore, this fraction is usually fixed to a value of 0.5 [15]. The value of d_s obtained in

many analyses is generally assumed to reflect the rms roughness of the real surface. However, the concept proposed by Aspnes et al. [3] does not provide a straightforward connection between these two parameters. With the advent of surface characterization techniques such as Atomic Force Microscopy (AFM), it was possible to investigate whether a direct relation between surface roughness and the value of d_s indeed exists. However, only a partial correlation was found [15]. Therefore, after more than 30 years since the publication of Aspnes work, two long-standing questions prevail regarding the use of EMA in ellipsometry to model the effect of surface morphology: (1) how is it possible that, with only two parameters, the EMA is able to describe the effect of surface roughness in spectroscopic ellipsometry in such a wide range of experimental systems? and (2) what is the correlation between the thickness of the EMA layer d_s and the surface morphology?

This chapter introduces the current state of roughness analysis and describes alternative approaches for modelling microscopic roughness based on a perturbation of the Fresnel reflection coefficient. This description starts with an evaluation of the statistically averaged quantities that are used to characterize surface roughness. This is followed by a short overview of the correlation between surface roughness and EMA reported in literature. The Rayleigh Rice perturbation theory will be introduced that provides a suitable formalism for describing the effect of microscopic roughness in ellipsometric analysis. We will finish with a short discussion on normal incidence ellipsometry, also known as Reflectance Anisotropy Spectroscopy (RAS) as this provides an experimental configuration for which the Rayleigh Rice approximation condenses to a very transparent formula.

4.2 Quantification of Microscopically Rough Surfaces

4.2.1 Statistical Properties of a Surface Morphology

The simplest mathematical description of a surface assumes that every point of the surface is characterized by a unique height $h(x, y)$. While this definition disregards the influence of overhangs and microporosity, it is a good approximation for most of the experimental systems where spectroscopic ellipsometry is applied. Moreover, this definition allows a one-to-one correlation with experimental measurements. Discrete maps $h(x, y)$ of surface topography at the nanoscale can be obtained by Scanning Probe Microscopy techniques such as AFM or STM. The output of those techniques is a discretized version of the $h(x, y)$ function convoluted with an instrument function due to the finite radius of the tip and the precision in the lateral and vertical displacement of the experimental system. The task is to represent the fluctuation of $h(x, y)$ by numbers that represent statistically averaged quantities [39].

Sampling of the values of $h(x, y)$ over a certain sampling area of characteristic size L allows the determination of a height distribution function $f(h; L)$, so that $f(h; L)dh$ indicates the probability that a point of the surface has a height between

h and $h + dh$. From $f(h; L)$ it is possible to determine a number of parameters frequently used to characterize surface topography, of which two are the most common:

- The average height of the surface is simply the first moment of this height distribution: $\bar{h} = \int h f(h; L) dh$.
- The surface roughness $\sigma(L)$ is defined as the second central moment of the distribution:

$$\sigma^2 = \int (h - \bar{h})^2 f(h; L) dh \qquad (4.2)$$

In many cases, this distribution function resembles a Gaussian. In this case, the second moment provides the width of this height distribution and $\bar{h}$ and σ completely determine the statistical properties of the rough surface. However, if the distribution departs from a Gaussian shape, higher moments of the height distribution become relevant and surface roughness fails to completely capture the statistical properties of the surface morphology.

The surface roughness $\sigma(L)$ generally depends on the length scale L. On real surfaces, microscopic surface roughness generally increases with the length scale, until reaching a constant value above a certain value of L. The interpretation of this dependence is simple: closer points tend to share similar heights; as we increase the sampling area, a more rich morphology is sampled, leading to higher height variations. Eventually, for a wide enough sampling area, we capture all the main surface features, and increasing it even further does not provide any extra information. In this case, the surface roughness reaches an asymptotic value w, so that $\sigma(L) \to w$ for high enough L. Hereafter, we will refer as surface roughness to the asymptotic length scale value w, also commonly known as the mean square roughness (rms).

The fact that roughness changes with length scale is a consequence of surface correlations. It is the fact that closer points tend to exhibit similar heights (their heights are correlated) what leads to a length scale-dependent roughness. In contrast, in a purely uncorrelated surface, where heights at each points are extracted randomly from a given distribution, surface roughness is independent of L.

The height distribution function $f(h; L)$ defined above provides information on the vertical range of the heights of points in a surface. However, it does not contain any information on how those points are spatially distributed. This information is obtained from lateral correlation functions.

The so called height-height correlation function is defined as

$$C(\mathbf{r}) = \left[\langle (h(\mathbf{r_0} - \mathbf{r}) - h(\mathbf{r_0}))^2 \rangle_{SA} \right]^{1/2} \qquad (4.3)$$

where the average extends to a certain surface SA. $C(\mathbf{r})$ measures the statistical correlation of two points separated by the vector $\mathbf{r}$. If two points are completely uncorrelated, then their heights are independent from each other and $C(\mathbf{r}) \to 2w$. For extremely short distances, the correlation function is zero. From the behavior of $C(r)$ a correlation length ξ can be defined. The points separated a distance greater than the correlation length are considered to be statistically uncorrelated. The influence
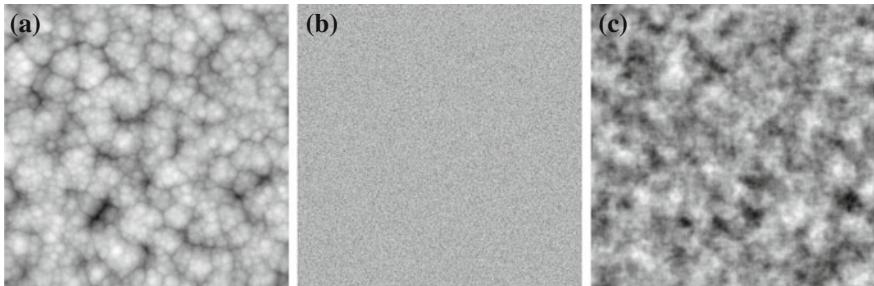
Fig. 4.2 **a** Experimental AFM image of a HfB$_2$ surface. **b** Surface with identical height distribution but spatially uncorrelated. **c** Surface with identical power spectral density constructed using a random phase in its Fourier transform

of correlation length is shown in Fig. 4.2: the surfaces represented in Fig. 4.2a, b are characterized by the same height distribution function, but in Fig. 4.2b the points are completely uncorrelated. The presence of a *cauliflower-like* texture is a consequence of spatial correlations in the surface.

A final important function used to describe the morphology of surfaces is the power spectral density. It is defined as:

$$S(\mathbf{q}) = \langle \hat{h}(\mathbf{q})\hat{h}(-\mathbf{q}) \rangle \tag{4.4}$$

where $\hat{h}(\mathbf{q})$ is the Fourier transform of the surface. The power spectral density is the Fourier transform of the surface autocorrelation function $G(\mathbf{r})$, which is defined as:

$$G(\mathbf{r}) = \langle h(\mathbf{r_0} - \mathbf{r})h(\mathbf{r_0}) \rangle = w^2 - \left[C(\mathbf{r})^2/2 \right] \tag{4.5}$$

Generally, for isotropic surfaces, it is the radial average what matters, and we can define a radially averaged power spectral density $S(k)$ from Eq. 4.4. Surfaces having the same averaged power spectral density exhibit identical roughness, correlation length and height distribution function. However they still can appear significantly different. Thus, Fig. 4.2a, c have the same power spectral density and therefore have the same roughness, correlation length and height distribution function despite their different textures.

As shown in this chapter, the power spectral density is the fundamental variable controlling the scattering of light from a rough surface. Thus, surfaces whose textures are represented by Fig. 4.2a, c have identical response and are equivalent from the point of view of spectroscopic ellipsometry.

4.2.2 Classification of Surface Morphologies

The main criterion to characterize the surface morphology is whether the surface is periodic or random. Real surfaces are rarely periodic, and even periodically textured

surfaces have microscopic roughness. The second criterion is whether the surface is isotropic or anisotropic.

In the case of isotropic randomly rough surfaces, their power spectral density allows to classify them as one of the three following types [39]:

Gaussian surfaces   Gaussian surfaces are characterized by a Gaussian power spectral density. Their power spectral density is given by two parameters, the surface roughness w and the correlation length ξ, so that:

$$S(\mathbf{q}) = \frac{w^2 \xi^2}{4\pi} e^{-q^2 \xi^2/4} \tag{4.6}$$

Gaussian power spectral densities are seldom found in experimental systems, but they are commonly used as model surfaces in scattering problems.

Self-affine surfaces   As mentioned above, roughness changes with the scale of measurements before a certain length scale. In some cases, the dependence of roughness with length scale follows a power law, so that $\sigma(L) \sim L^\alpha$, where α is the so-called roughness exponent. Objects having this property are known as self-affine. The case of $\alpha = 1$ corresponds to a self-similar fractal surface [7].

For these kind of objects, the power spectral density $S(q)$ has two asymptotic behaviors:

$$S(q) = \begin{cases} \text{const.} & q \ll 2\pi/\xi \\ q^{-(2\alpha+2)} & q \gg 2\pi/\xi \end{cases} \tag{4.7}$$

Mound surfaces   Mound surfaces are characterized by a characteristic periodicity Λ_m that reflects the average separation between similar features [28]. The power spectral density is characterized by a maximum at surface wavenumber $q_m \sim \Lambda_m^{-1}$. The width of the peak is related to the correlation length of the surface ξ. At higher wavenumber, the asymptotic value of the power spectra density can be characterized by a power law as in the self-affine case.

4.2.3 Origin and Development of Surface Morphology

Roughness is ubiquitous in real surfaces. Even in the case of thermodynamic equilibrium, a roughening transition takes place at high-index surfaces. In nanotechnology, roughness normally appears in the context of growth or etching processes. The surface morphology develops as a consequence of two competing processes: the incorporation or removal of material from a surface in a random fashion, and surface relaxation processes that generally tend to drive the surface towards a more stable configuration. The main reason why the evolution of morphology is so relevant for our purposes is that, as we will see in the next section, correlations between the EMA and surface morphology have been determined in the context of thin film growth. Therefore, in this section we will provide the required background.

In understanding how surface morphology and surface roughness develops in the context of thin film growth, it is important to distinguish between the early stages of growth and the steady-state growth conditions for thicker films. If materials growth starts on a foreign surface, nucleation can have a strong impact on the surface morphology. Sparse nucleation tend to favor big grains that form rough surfaces after coalescence, and the films exhibit nucleation induced roughness [12, 21]. The equilibrium between deposition and relaxation control the subsequent evolution of the surface morphology at later stages of growth.

The evolution of surface morphology with growth has been studied in detail for the case of self-affine surfaces. In many experimental systems, the growth dynamics leads to the formation of self-affine surfaces [7]. In the simplest case, roughness increases with thickness t as $\sigma \sim t^{\beta}$, where β is the so called growth exponent, until reaching a certain saturation value. That saturation value depends on L so that $\sigma(L) \sim L^{\alpha}$. Likewise, for distances greater than the correlation length ξ, the surface roughness becomes independent of the length scale. The consequence of these two trends is that the surface morphology evolves in such a way that the roughness exponent α is preserved during growth, and the roughness w increases as $w \sim t^{\beta}$. This behavior is summarized by the so called Family Vicsek scaling law, stating that $\sigma(L, t) \sim L^{\alpha} f(t/L^{z})$, with $f(u) \to 1$ for $u \gg 1$ and $f(u) \to u^{\beta}$ for $u \ll 1$ [7].

The power spectral density of systems satisfying the scaling behavior described above evolves with time in such a way that the high spatial frequency components remain constant, and changes take place only at low wavenumbers. Likewise, in these systems the correlation length increases with time following the scaling law: $\xi \sim t^{1/z}$, where $z = \alpha/\beta$. As we will show in later sections, the evolution of the power spectral density with time is crucial in order to establish a good correlation between the EMA roughness and the surface roughness w in self-affine surfaces.

The Family Vicsek scaling law is only one of the many different dynamics that have been experimentally observed. Anomalous scaling behaviors have been reported in many different systems, including situations in which the roughness exponent α changes as a function of time [16, 17, 24, 25, 35, 36]. However, the Family Vicsek is the benchmark against all the experimental results are compared.

4.3 Correlation Between EMA and Surface Morphology: Comparison with Experiments

The EMA has been proved to be extremely effective in modeling the optical response of rough surfaces. One of the most surprising aspects is that, given the number of different parameters required to characterize surface morphology, such good agreement with optical measurements is obtained with at most two parameters. In practice, a strong correlation between the two parameters is found and the void fraction is usually set to 0.5. This leaves the thickness as the sole fitting parameter. The main question is: how does it correlate with the surface morphology?

One of the earliest attempts to extract information on the surface morphology was carried out by Aspnes et al. [3]. In their work on Si surfaces, they prepared samples with different morphologies and fitted the ellipsometric data using different effective medium theories. By leaving the void fraction as a fitting parameter, they were able to extract an effective density of the roughness layer. They then correlated this density value with that calculated using different models for the surface morphology, including a surface layer composed by hemispheres, truncating cones and cylindrical ridges. Their results showed that hemispheres provided density values that were in agreement with the experimental data. Moreover, they determined that Bruggeman's EMA theory provided a better agreement with their experimental data. This work on p-Si was followed up by work on Au surfaces. Both models helped consolidating the EMA as a valid approach to the modeling of microscopic roughness.

Since then, a number of different authors have studied the correlation of EMA roughness with surface morphology using characterization techniques such as profilometry, atomic force microscopy or transmission electron microscopy. One of the earliest insights on this subject was the work of Fang et al., which highlighted the influence of power spectral density on the ellipsometric response of the surface. In their 1996 work, Fang et al. [10] characterized more than forty samples using both single-wavelength ellipsometry and atomic force microscopy. Surface roughness was generated using wet etching and wet etching followed by thermal processing, and measurements were carried out immediately after HF dipping. Their results showed that the optical response of wet-etched samples could be fitted assuming a 1:1 relationship between the EMA model roughness and the surface roughness as measured by AFM and a 40–50 % void fraction. However, the 1:1 relationship in the samples that were also thermally treated could be maintained only if void fractions of 70–80 % were assumed. A direct comparison of the surface morphology of these two films showed a similar surface roughness, while their power spectral density significantly differed. Wet-etched films had a smaller high-wavenumber contribution to the power spectral density, and that directly correlated with a different magnitude change on the ellipsometric angle Δ, therefore providing a compelling evidence on the influence of high-wavenumber contributions of the surface on the polarization state of the reflected light.

The lack of universal correlation between d_s and w was also confirmed by Petrik et al. [27, 29]. In their characterization of polysilicon thin films by spectroscopic ellipsometry and atomic force microscopy, they found d_s/w ratios ranging from 0.8 to more than 1.6. The authors attributed such variation to differences in surface morphology that, as shown in the previous section, cannot be described by a single roughness parameter. The author suggested that, while quantitative determination of surface roughness is possible by spectroscopic ellipsometry, a calibration curve is first needed in order to determine the d_s/w ratio for the desired experimental conditions.

Excellent correlation between surface roughness as measured by Atomic Force Microscopy and the EMA thickness determined by spectroscopic ellipsometry was also found by Fujiwara et al. for a-Si:H films [11, 12]. The linear relationship was estimated to be $d_s = 1.5w + 4$ Å. Best-fit void fractions were higher (40–50 %) for

the early stages of growth characterized by nucleation-induced roughness than for the thicker films (30–40 %). In a later study, Fujiwara et al. determined a correlation between EMA and AFM roughness given by $d_s = 0.88w + 4.9$ Å for a different set of samples of a-Si:H and μc-Si:H, thus confirming that even for the same material, different growth conditions could lead to different d_s/w ratios.

While the d_s/w is a-priori unknown, a linear dependence between d_s and w implies that spectroscopic ellipsometry can be used as an in-situ technique to monitor the change of roughness as a function of film thickness or processing time. Thus, in 2003 Smets et al. [33] studied the morphology evolution of a-Si:H growth as a function of temperature. Both in situ ellipsometry and ex-situ AFM showed that the kinetic roughening of a-Si was characterized by a power law dependence with time: $d_s \sim t^\beta$, which was consistent with the ideal case described by the dynamic scaling theory referred in the previous section and with the results found by other authors. The value of β determined using single-wavelength rotating compensator ellipsometry and AFM were within the experimental error ($\beta_{RCE} = 0.29 \pm 0.03$ vs $\beta(AFM) = 0.27 \pm 0.03$), showing that even if the ratio is unknown, surface ellipsometry can provide a useful indication of the roughness evolution if a linear dependence between the EMA thickness layer and the surface roughness can be demonstrated.

However, even though there is an abundant experimental evidence suggesting a linear dependence between d_s and w, in 2004 Sperling and Abelson [31] provided an interesting counterexample that showed that such linear dependence is far from universal. Starting from rough substrates, they showed that while in-situ spectroscopy ellipsometry indicated a smoothening of the surface, the surface roughness w as determined by AFM increased with time. From their analysis of the power spectral density, the authors determined that, while the low wavenumber component of the power spectral density increased with film thickness, its high wavenumber component was actually decreasing. They concluded that spectroscopic ellipsometry was more sensitive to the high wavenumber contribution of the power spectral density.

The conclusion that we can extract from the literature is that, while in most of cases there is a linear relationship between thickness of the EMA layer and the surface roughness w, that relationship is by no means universal, and in some extreme cases the evolution with thickness of both parameters even follow opposite trends. Also, d_s/w ratios depend on the experimental conditions. First-principle calculations of the interaction of light with microscopically rough surfaces hence become a crucial tool to understand the correlation between EMA and surface morphology, and to determine under which conditions it is reasonable to expect a linear dependence between EMA and surface roughness.

4.4 Application of Scattering Models to Rough Surfaces

4.4.1 Rayleigh-Rice Scattering Theory

A complementary approach to the experiments described in the previous section is the use of electromagnetic scattering theory to describe the optical response of a rough surface. Many different approaches have been described in literature to describe the scattering of electromagnetic radiation from rough surfaces. Since both surface roughness w and the characteristic size of the surface features are much smaller than the wavelength of the incident light λ, the Rayleigh-Rice formalism is a suitable approach to model the interaction of light with microscopically rough surfaces.

Franta and Ohlidal [13] developed such formalism up to second order of the w/λ parameter, and obtained a closed expression for the change in the Fresnel reflection coefficient of a randomly rough surface:

$$\Delta r_{p,s} = w^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{p,s}(q_x, q_y) S_n(q_x - n_0 k_0 \sin \theta_0, q_y) dq_x dq_y \qquad (4.8)$$

Here, $f_{p,s}$ are kernel functions that relate the surface morphology to the optical response that depend on the optical properties of the material and the angle of incidence θ_0, the wavenumber of the radiation $k_0 = 2\pi/\lambda$ and the spatial wavevector q. Expressions for the kernel functions $f_{p,s}$ were derived by Franta and Ohlidal [13] and are reproduced in the appendix. $S_n(\mathbf{q}) = S(\mathbf{q})/w^2$ is the normalized version of the power spectral density defined in the previous section, i.e.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_n(q_x, q_y) dq_x dq_y = 1. \qquad (4.9)$$

Let us consider the ideal case of a surface with a Gaussian power spectral density. Then, the power spectral density can be parametrized using Eq. 4.6, and the optical response depends on two parameters: the surface roughness w and the correlation length ξ. The effect of surface morphology on the pseudodielectric constant is shown in Fig. 4.3 for the case of p-Si. It is clear from Fig. 4.3 that both the surface roughness and the correlation length have an influence on the optical response.

Moreover, by fitting the simulated optical response to an EMA model, it is possible to establish a correlation between the power spectral density of the surface and the EMA roughness. This process was carried out by Franta and Ohlidal [14] for Gaussian microscopically rough surfaces and by Yanguas-Gil et al. [37] for self-affine microscopically rough surfaces.
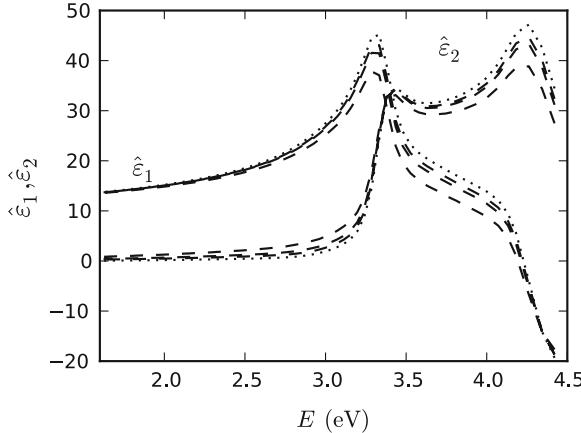
Fig. 4.3 Influence of correlation length ξ and roughness w on the pseudodielectric constant of a p-Si surface with a Gaussian power spectral density

4.4.2 Comparison with Effective Medium Models

Before establishing the connection between the power spectral density and EMA model roughness, it is instructive to compare Eq. 4.8 with the corresponding expression obtained considering an effective medium approximation. The change in the reflection coefficients due to the presence of an EMA roughness d_s can be expressed as:

$$\Delta r_{p,s} = \frac{r_{s1}\left(1 - r_{0s}^2\right)\left(e^{-2i\chi} - 1\right)}{(1 + r_{0s}r_{s1})\left(1 + r_{0s}r_{s1}e^{-2i\chi}\right)} \tag{4.10}$$

where r_{0s} and r_{s1} are the Fresnel reflection coefficients of the air/roughness and roughness/surface interface respectively, and χ is given by:

$$\chi = 2\pi\left(\frac{d_s}{\lambda}\right)n_{\text{EMA}}\cos\theta_s \tag{4.11}$$

Here, n_{EMA} is the effective refraction index of the roughness layer and θ_s is the angle of incidence in the effective medium, which can be obtained from the incidence angle θ_0 using Snell's law.

Restricting Eq. 4.10 to the region of applicability of the Rayleigh-Rice theory, that is, assuming that $d_s \ll \lambda$, one obtains at first order in d_s/λ that:

$$\Delta r_{p,s} = -4i\pi\frac{n_{\text{EMA}}\cos\theta_s}{\lambda}\frac{r_{s1}\left(1 - r_{0s}^2\right)}{(1 + r_{0s}r_{s1})^2}d_s \tag{4.12}$$

From, Eq. 4.12, it is clear that for microscopically rough surfaces, the change in the Fresnel reflection coefficients scales linearly with the EMA roughness d_s.

Equation 4.8, on the other hand, would predict a quadratic dependence of $\Delta r_{p,s}$ on the surface roughness w, due to the normalization condition of the power spectral density. However, as shown in the previous section, in many cases a linear dependence of d_s on w is experimentally found.

This apparent discrepancy was studied by Yanguas-Gil et al. [37], who applied the second-order Rayleigh-Rice scattering theory described above to the case of randomly rough surfaces with a self-affine power spectral density. In this case, as mentioned in the previous section, the surface morphology is described by three parameters: surface roughness w, correlation length ξ and the roughness exponent α, which determines the high surface wavenumber dependence of the power spectral density. To these parameters the authors added a cut-off wavenumber q_c, to account for the breakdown in the scaling behavior at an atomic scale.

Using a parametrization of the power spectral density such that:

$$S(\mathbf{q}) = \frac{\alpha w^2 \xi^2}{\pi} \frac{1}{\left(1 + \xi^2 q^2\right)^{\alpha+1}} \qquad (4.13)$$

$\Delta r_{s,p}$ can be determined for different materials and surface morphologies. As in the Gaussian case, changes in w, ξ and α affect the optical response of the surface, and indeed a quadratic dependence on w is obtained when the rest of surface parameters remain constant. However, when the changes in the reflection coefficients are plotted against the single parameter w^2/ξ^α, the points coalesce onto a single curve, that is:

$$\Delta r_{p,s} \sim w^2/\xi^\alpha \qquad (4.14)$$

Using the standard definition of ρ:

$$\rho = \frac{r_p}{r_s} = \tan \Psi e^{i\Delta} \qquad (4.15)$$

it is easy to see how this dependence translates to Ψ and Δ. This is shown in Fig. 4.4. The change in Ψ and Δ as a function of w^2/ξ^α with respect to the values of a flat surface Ψ_0, Δ_0 is shown for CdS, GaAs, pSi and Cu surfaces for an incident angle of 70° and a wavelength of 500 nm. From the results in Fig. 4.4, it is clear that the parameter w^2/ξ^α captures the influence of the surface morphology on the optical response of a surface.

These results can be extended to a multiple wavelength analysis. As shown by Yanguas-Gil et al. [37] for a-Si:H, by fitting the simulated optical response to the BEMA model a linear relationship between the thickness of the roughness layer in the EMA model d_s and the parameter w^2/ξ^α for a-Si:H (Fig. 4.5). As in the case of the reflection coefficient, d_s changes with the three surface parameters controlling the morphology of a self-affine surface, but again these points coalesce in a single curve if d_s is plotted against w^2/ξ^α.

The physical interpretation of such dependence is simple: the average surface slope in a self-affine surface scales with w, ξ and α as $\delta \sim w/\xi^\alpha$ [26].
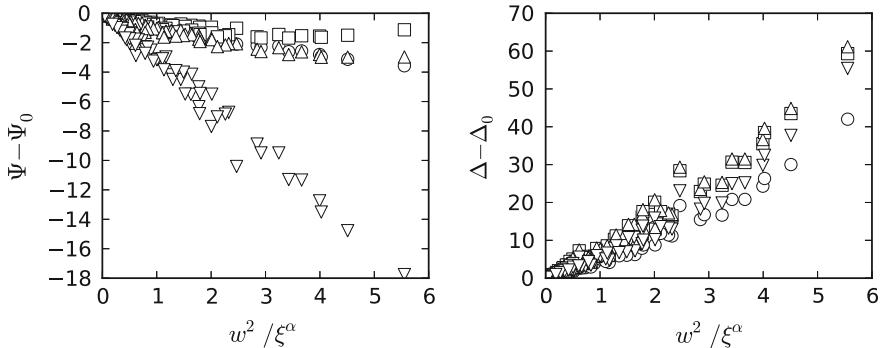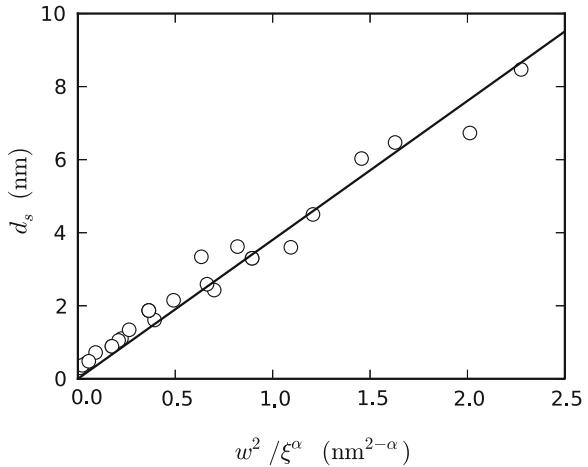
Fig. 4.4 Scaling of the change of the ellipsometric angles Ψ and Δ with the parameter w^2/ξ^α for different materials (CdS, GaAs, pSi, Cu) for an incident angle of 70° and a wavelength of 500 nm

Fig. 4.5 Correlation between the thickness of the roughness layer d_s in the BEMA model and the scaling parameter w^2/ξ^α for a-Si:H

Therefore, both the optical response of a self-affine surface and the corresponding EMA roughness d_s depend on the product of the surface roughness and the average surface slope:

$$d_s \sim w\delta \tag{4.16}$$

Surfaces exhibiting power spectral densities that can be parametrized by Eq. 4.13 are commonplace in the scientific literature, including some of the examples described in the previous sections. Therefore, in all those experimental systems, a linear dependence between the roughness d_s and w is indeed expected, and different proportionality factors experimentally observed between these two variables can be accounted for by considering different surface slopes.

Interestingly, the results presented above for self-affine power spectral densities are also reproduced in the case of an ideal Gaussian power spectral density. Franta and

Ohlidal determined a similar $d_s \sim w\delta$ dependence from fitting the optical response of modeled p-Si surfaces using Eq. 4.8 to the EMA model.

One consequence of the results discussed above is that the ability of EMA models to accurately follow the evolution of roughness with time relies on a small variation of the average surface slope δ with thickness. How common is this behavior? The dynamic scaling theory introduced in the previous section predicts that the correlation length changes with thickness as $\xi \sim t^{\beta/\alpha}$ and surface roughness as $w \sim t^\beta$. Thus, the average surface slope δ is expected to remain constant in systems where growth can be described by these two power laws.

When the growth departs from this behavior, the average surface slope will change with thickness. In experimental systems where this so-called anomalous scaling is present, the surface slope is expected to change as $\delta \sim t^\kappa$, and this would lead to a breakdown of the linear dependence between w and d_s, as experimentally shown by Sperling and Abelson [32]. The general conclusion that can be then extracted from scattering theory is that in general the linear dependence between d_s and w cannot be ensured without a proper evaluation of the roughening dynamics.

4.4.3 Small Correlation Length Approximation

The Rayleigh-Rice formalism allows the rationalization of apparent discrepancies found in the literature. As shown above, the application of this formalism to different materials suggests that the w^2/ξ^α, $w\delta$ dependence should be an intrinsic property of the scattering process. This can be explicitly shown in the limit where the surface correlation length is much smaller than the wavelength of the light used to probe the material investigated: $\lambda \gg \xi$. We will follow the approach presented in Ref. [37].

Assuming that the correlation length is much smaller than the wavelength, so that $\xi k_0 \ll 1$, it is possible to expand Eq. 4.13 in 4.8 in a power series of ξk_0, so that:

$$S_n(q_x - n_0 k_0 \sin\theta_0, q_y) = S_n(q_x, q_y)\left[1 - 2(\alpha + 1)n_0 \sin\theta_0 \frac{\xi q_x}{1 + \xi^2 q^2}(\xi k_0) + \mathcal{O}(\xi k_0)^2\right]$$

(4.17)

This essentially splits the contribution of the power spectral density in an even and odd contribution.

Likewise, the kernel functions $f_p(q_x, q_y)$ and $f_s(q_x, q_y)$ can be divided into even and odd components:

$$f_{p,s} = f_{p,s}^{(e)}(q_x, q_y) + q_x f_{p,s}^{(o)}(q_x, q_y)$$

(4.18)

Combining the two results, the change in the fresnel coefficients can be expressed as:

$$\Delta r_{p,s} = w^2 \int \int f_{p,s}^{(e)} S_n(q) dq_x dq_y \tag{4.19}$$

$$- 2w^2(\alpha + 1)\xi k_0 n_0 \sin \theta_0 \int \int f_{p,s}(q) \frac{\xi q_x^2 S_n(q)}{1 + \xi^2 q^2} dq_x dq_y$$

The even and odd components of $f_p(q_x, q_y)$ are given by:

$$f_p^{(e)}(q_x, q_y) = A_p + \frac{(b-c)}{q^2 + bc} \left[B_{31p} q^2 + B_{62p}(q_y^2 + bc) \right] \tag{4.20}$$

$$f_p^{(o)}(q_x, q_y) = B_{61p} \left[q^2 \frac{(b-c)^2}{(q^2 + bc)} + (q^2 + bc) \right] - \frac{B_{32p}}{q^2 + bc} \tag{4.21}$$

while those for s polarized light are given by:

$$f_s^{(e)}(q_x, q_y) = A_s + B_{5s} \frac{(b-c)(q_x^2 - bc)}{k^2 + bc} \tag{4.22}$$

$$f_s^{(o)}(q_x, q_y) = 0 \tag{4.23}$$

In this expression, $b = \sqrt{n_0^2 k_0^2 - q^2}$ and $c = \sqrt{n_1^2 k_0^2 - q^2}$. A_p, B_{31p}, B_{62p}, B_{61p}, B_{32p}, A_s and B_{5s} are coefficients that depend on the optical properties of the interface defined in Ref. [14], whose values are given in the appendix.

$q \gg k_0$ in most of the integration range of Eq. 4.19. Thus, we can further simplify the expression by determining the asymptotic limit of the optical response functions above. Considering:

$$b = iq \left[1 - \frac{1}{2} \frac{n_0^2 k_0^2}{q^2} + \mathcal{O}\left(\frac{k_0}{q}\right)^4 \right] \tag{4.24}$$

and

$$c = iq \left[1 - \frac{1}{2} \frac{n_1^2 k_0^2}{q^2} + \mathcal{O}\left(\frac{k_0}{q}\right)^4 \right] \tag{4.25}$$

then

$$f_p^{(e)}(q_x, q_y) = A_p + i \frac{n_1^2 - n_0^2}{n_1^2 + n_0^2} q \left[B_{31p} + B_{62p} \left(\frac{q_y^2}{q^2} - 1 \right) \right] \tag{4.26}$$

$$f_p^{(o)}(q_x, q_y) = \frac{2}{n_1^2 + n_0^2} \left(2n_0^2 n_1^2 k_0^2 B_{61p} - \frac{B_{32p}}{k_0^2} \right) \tag{4.27}$$

$$f_s^{(e)}(q_x, q_y) = A_s + i \frac{n_1^2 - n_0^2}{n_1^2 + n_0^2} q B_{5s} \left(\frac{q_x^2}{q^2} - 1 \right) \tag{4.28}$$

$$f_s^{(o)}(q_x, q_y) = 0 \tag{4.29}$$

And the change in the Fresnel coefficients is given by:

$$\Delta r_p = A_p w^2 + i\pi \frac{n_1^2 - n_0^2}{n_1^2 + n_0^2} \left(2B_{31p} - B_{62p}\right) w^2 \int q^2 S_n(q) dq \tag{4.30}$$

$$- \frac{4\pi k_0^2 n_0 \sin\theta_0}{n_1^2 + n_0^2} \left(2n_0^2 n_1^2 B_{61p} - \frac{B_{32p}}{k_0^4}\right)(\alpha + 1) w^2 k_0 \xi^2 \int \frac{q^3 S_n(q)}{1 + \xi^2 q^2} dq$$

$$\Delta r_s = A_s w^2 - i\pi \frac{n_1^2 - n_0^2}{n_1^2 + n_0^2} B_{5s} w^2 \int q^2 S_n(q) dq \tag{4.31}$$

As was shown in Ref. [37], of the three terms in Δr_p, it is the second one the responsible for the scaling behavior determined above. The first and the third show contributions that only depend on the square of the surface roughness w, with the third term being first order in ξk_0, which is normally much smaller than unity. Likewise, in Δr_s the scaling behavior is due to the second term. This means that, in both cases, the influence of the optical properties of the interface and the surface microstructure become separable:

$$\Delta r_{p,s} = C_{p,s}(n_0, n_1, \theta_0, k_0) T(w, \xi, \alpha, q_c) \tag{4.32}$$

where we have defined:

$$C_p(n_0, n_1, \theta_0, k_0) = i\pi \frac{n_1^2 - n_0^2}{n_1^2 + n_0^2} \left(2B_{31p} - B_{62p}\right) \tag{4.33}$$

$$C_s(n_0, n_1, \theta_0, k_0) = -i\pi \frac{n_1^2 - n_0^2}{n_1^2 + n_0^2} B_{5s} \tag{4.34}$$

$$T = w^2 \int q^2 S_n(q) dq \tag{4.35}$$

Here T has been chosen so that it has units of length. As shown in Fig. 4.6, T is responsible for the scaling with w^2/ξ^α predicted by the Rayleigh Rice model.

To conclude this analysis, it is interesting to compare Eq. 4.32 with the linearized expression for the BEMA model, Eq. 4.12. In both cases, since

$$\Delta\rho \approx \rho \left(\frac{\Delta r_p}{r_p} - \frac{\Delta r_s}{r_s}\right) \tag{4.36}$$

the ratios $\Delta\rho^{RR}/T$ and $\Delta\rho^{EMA}/d_s$ depend solely on the optical properties of the materials, angle of incidence and light wavelength. In Fig. 4.7, we show the correlation of the absolute values and phase of these magnitudes for a wavelength range of 1.5–4.5 eV, refractive index ranging from 1.2 to 3.1, attenuation constant from 0.2 to 2 and an angle of incidence of 70°. Despite their different origin and mathematical structure, the agreement between the two models is fairly good, showing a good
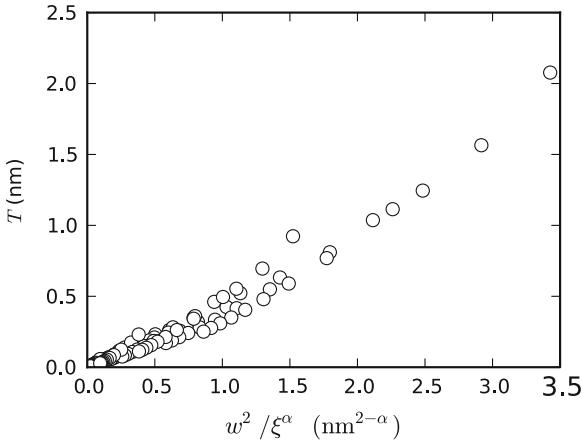
Fig. 4.6 Dependence of the surface term T of the asymptotic RR model with w^2/ξ^α
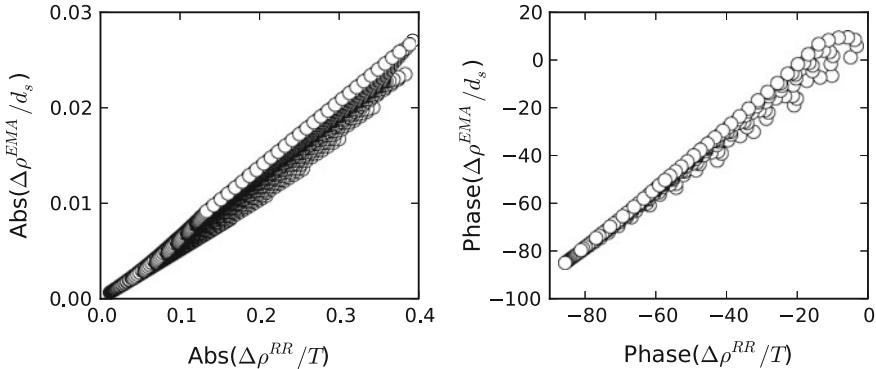

Fig. 4.7 Correlation of the normalized optical response modeled using the BEMA model and using the asymptotic RR theory

linear behavior. Such agreement is the reason why the simulated spectroscopic ellipsometry data using the RR theory can be accurately fitted using the BEMA model, as depicted in Fig. 4.5 [37].

4.5 Optical Characterisation at Normal Incidence

The influence of roughness is particularly well observed on metal surfaces like silver. Several studies showed the emergence of a strong plasmon feature around 335 nm during the growth of a silver film [5, 19, 22, 30]. A thorough treatment

of the influence of roughness on the normal incidence optical response of a silver surface using the Rayleigh Rice approximation was provided by Kretschmann and Kröger [18]. This approach explains the substantial red shift observed for roughness with a large correlation length. The BEMA approximation is not able to reproduce the position of the plasmon peak. This inability of BEMA to represent a plasmon resonance is very similar to that found for the optical characterisation of noble metal colloidal particles [20]. Modelling the roughness of a silver surface with BEMA provides a resonance feature at a much larger wavelength than the 335 nm observed for roughness with a small correlation length. In fact, the BEMA approach for a silver surface reflects the response of roughness with a correlation length of 4 μm. However, this is not the only problem associated with the BEMA approach. Also anisotropic structures like ripples are conceptually difficult to model with this. The optical response of such a nanostructured surface was characterized in detail with RAS measurements [8, 9, 34]. The use of RAS has several advantages compared to the usually employed oblique incidence ellipsometry. The main advantage is that the technique is only sensitive for anisotropic features. Any isotropic roughness present on the initial surface is disregarded. Furthermore, the standard experimental system used in RAS works as an optical bridge, which allows high sensitivity [1, 38]. The analysis of the recorded optical response is very suited for analysis within the Rayleigh Rice formalism introduced in the previous section.

At normal incidence, the rather complex perturbation of the reflectivity for oblique incidence introduced in Eq. 4.8 can be highly simplified. In this situation, the kernel reduces to the form found for s-polarized light, see Eq. 4.22. However, because we are now interested in the response as a result of roughness on a surface ($n_0 = 1$) that is different two orthogonal directions (x, y) on the surface, two kernels associated with these directions have two be considered:

$$\hat{f}_{y,x} = -2k_0 r^{(0)} q \left(\frac{k_0 n_1}{q} + \frac{(\hat{b} - \hat{c})\left(\left(\frac{q_{x,y}}{q}\right)^2 + \hat{b}\hat{c}\right)}{1 + \hat{b}\hat{c}} \right) \tag{4.37}$$

In this the normalized version of the previously introduced parameters was used:

$$\hat{b} = \frac{b}{q} = \sqrt{\left(\frac{k_0}{q}\right)^2 - 1} \tag{4.38}$$

$$\hat{c} = \frac{c}{q} = \sqrt{\left(\frac{k_0 n_1}{q}\right)^2 - 1} \tag{4.39}$$

are used.

With these definitions the reflectivity difference of an anisotropically rough surface on top of an optically isotropic substrate can be evaluated:

$$\frac{\Delta r}{r} = 2 \cdot \frac{(r_x - r_y)}{(r_x + r_y)}$$

$$= \frac{\sigma^2}{\hat{r}^{(0)}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\hat{f}_x - \hat{f}_y) w(\vec{q}) d\vec{(q)} z \qquad (4.40)$$

where $\hat{r}^{(0)} = \frac{1}{2} \cdot (\hat{r}_x^{(0)} + \hat{r}_y^{(0)})$. The difference in the kernel $\hat{f}$ reduces to:

$$\hat{f}_x - \hat{f}_y = -2k_0\hat{r}^{(0)}q \cdot \frac{(\hat{b} - \hat{c}) \cdot \left(\left(\frac{q_y}{q}\right)^2 - \left(\frac{q_x}{q}\right)^2 \right)}{(1 + \hat{b}\hat{c})}$$

$$= -2k_0\hat{r}^{(0)}q \frac{(\hat{b} - \hat{c})}{(1 + \hat{b}\hat{c})} \cdot \cos 2\phi \qquad (4.41)$$

The angle ϕ represents the direction of $\vec{q}$ with respect to the parallel direction. The selection of only the anisotropic roughness contribution to the optical response is made by the $cos2\phi$ term.

For anisotropic roughness in the y direction with a small lateral length scale, i.e. $q \gg k_0$, the kernel can be simplified to:

$$\hat{f}_y = -2ik_0\hat{r}^{(0)}q_y \frac{n_1^2 - 1}{n_1^2 + 1} \qquad (4.42)$$

The spectral contribution of this kernel is reduced to a resonance function $\frac{n_1^2-1}{n_1^2+1}$. For silver this provides the mentioned plasmon resonance at 335 nm. Both rms roughness and spatial periodicity contribute in a similar way to the amplitude of this kernel. The anisotropic optical response is found by integrating the product of the kernel and the specific PSD:

$$\frac{\Delta r}{r} = -2ik_0\sigma^2 \frac{n_1^2 - 1}{n_1^2 + 1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\vec{q}| \cos 2\phi w(\vec{K}q) d\vec{(q)} \qquad (4.43)$$

For anisotropic roughness, this integration over the PSD provides the average spatial wavevector q_0. This leads to an optical response that shows the same dependence on roughness and correlation length derived above, i.e. a change in reflectivity proportional to w^2/ξ

$$\frac{\Delta r}{r} = -2ik_0w^2q_0 \frac{n_1^2 - 1}{n_1^2 + 1} = -\frac{8\pi^2 w^2}{\lambda \xi} \frac{n_1^2 - 1}{n_1^2 n + 1} \qquad (4.44)$$

Aspnes [2] evaluated the optical response of various nanostructures taking into account local field effects and found for a sinusoidal modulation the result of Eq. 4.44. However, in his derivation he used $Re(n_1^2) \gg 1$. The resonance observed for Ag does not comply with this, as the position of the resonance is given by the Frölich
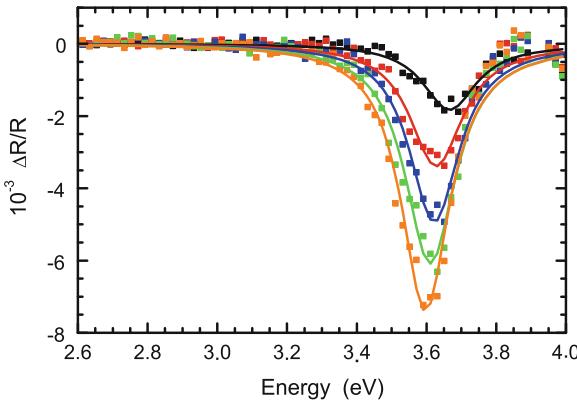
Fig. 4.8 RAS spectrum of a ripple pattern on Ag(001) created by grazing incidence ion sputtering for various ion fluence. A fit of a Lorentz lineshape to the individual spectra is also shown
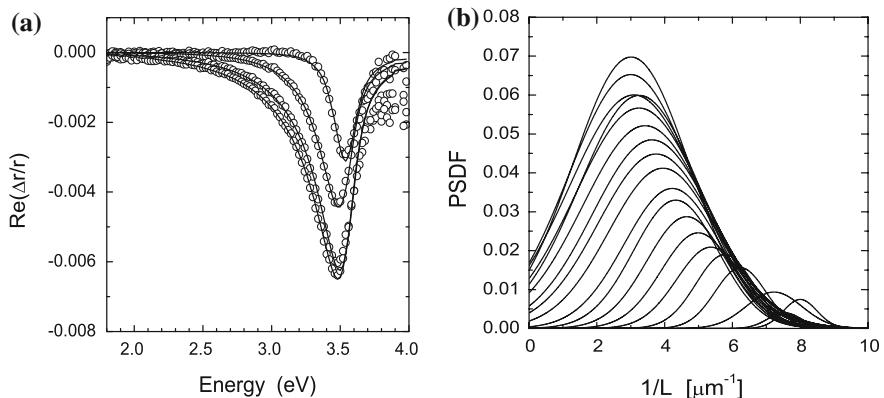

Fig. 4.9 **a** RAS spectrum of a ripple pattern on Ag(001) created by oblique incidence ion sputtering for various ion fluence. Note that only 4 spectra and the and the model fit are shown for clarity. **b** Evaluated PSDF from the fit to the measured spectra

condition $Re(n_1{}^2) = -1$. Local field effects as considered by for instance Aspnes and Barrera et al. thus seem not to have any influence on the optical response for a sinusoidal profile [2, 4].

An anisotropic ripple pattern can be obtained by grazing incidence ion sputtering [8]. The anisotropic optical response of such a pattern created at 320 K is shown in Fig. 4.8. Clearly visible is the plasmon resonance whose shape can be fitted with Eq. 4.44. This lineshape is essentially a Lorentz profile [34]. However, a red shift with increasing ion fluence is observed, albeit that the Lorentz lineshape still fits the optical spectrum. This indicates that the length scale at the prolonged stage is no longer much smaller than the resonance wavelength of 334 nm and in principle Eq. 4.44 is no longer valid. The periodicity of the ripple pattern can be reduced by

sputtering at a lower sample temperature. At 250 K, a pattern with a periodicity of around 50 nm can be made providing the resonance peak at the expected 334 nm wavelength. The red shift observed in Fig. 4.8 allows to retrieve the length scale of the ripple pattern. The amplitude of the RAS spectrum is again proportional to w^2/ξ.

Ripple patterns with a larger periodicity can no longer be modelled with the short correlation length approach used in the derivation of Eq. 4.42. This breach can be experimentally observed as a much larger red shift, see Fig. 4.9a. This occurs for a less grazing incidence sputtering, i.e. a polar angle of 70° and a temperature of 390 K. A fit to the experimental data with a Gaussian distributed periodicity of the ripples provides an evolution of the PSDF that provides both the ripple periodicity distribution and ripple amplitude as depicted in Fig. 4.9b. The observed red shift is related to a substantial increase in the ripple periodicity, reflected by the decrease of the depicted spatial wavevector.

A. Terms of the Rayleigh-Rice Theory

Following the notation used in Ref. [14], the kernel functions in the Rayleigh-Rice theory are given by:

$$f_p(q_x, q_y) = A_p + (b - c)\frac{B_{31p}q^2 + B_{62p}(q_y^2 + bc)}{q^2 + bc}$$

$$+ q_x \frac{B_{61p}\left[q^2(b - c)^2 + (q^2 + bc)^2\right] - B_{32p}}{q^2 + bc} \tag{4.45}$$

$$f_s(q_x, q_y) = A_s + B_{5s}\frac{(b - c)(q_x^2 - bc)}{q^2 + bc} \tag{4.46}$$

where:

$$b = \sqrt{n_0^2 k_0^2 - q^2} \tag{4.47}$$

$$c = \sqrt{n_1^2 k_0^2 - q^2} \tag{4.48}$$

and

$$A_p = k_0^2 n_0 n_1^2 X/D \tag{4.49}$$

$$B_{31p} = k_0 n_0^2 n_1^2 \sin \theta_0 W / D \tag{4.50}$$

$$B_{62p} = k_0 n_0 n_1 \cos \theta_1 X / D \tag{4.51}$$

$$B_{61p} = n_0 n_1 \cos \theta_1 W / (D k_0) \tag{4.52}$$

$$B_{32p} = k_0^3 n_0^2 n_1^2 \sin \theta_0 X / D \tag{4.53}$$

$$A_s = -2k_0^2 n_0 n_1 \cos \theta_0 \cos \theta_1 r_s \tag{4.54}$$

$$B_{5s} = -2k_0 n_0 \cos \theta_0 r_s \tag{4.55}$$

with:

$$D = (n_0 \cos \theta_0 + n_1 \cos \theta_1)(n_0^2 \sin^2 \theta_0 + n_0 n_1 \cos \theta_0 \cos \theta_1) \tag{4.56}$$

$$X = (n_0^2 - n_1^2) \cos \theta_1 t_p \tag{4.57}$$

$$W = (n_1 / n_0^2 - 1/n_1) n_0 \sin \theta_0 t_p \tag{4.58}$$

Here, θ_0 is the incidence angle, k_0 is the light wavenumber, n_0 and n_1 are the air and the material complex refractive index, and r_s, r_p, t_p, t_s are the Fresnel reflection and transmission coefficients for the flat interface.

References

1. D.E. Aspnes, J.P. Harbison, A.A. Studna, L.T. Florez, Application of reflectance difference spectroscopy to molecular-beam epitaxy growth of gaas and alas. J. Vac. Sci. Technol. A **6**, 1327 (1988)
2. D.E. Aspnes, Optical response of microscopically rough surfaces. Phys. Rev. B **41**, 10334 (1990)
3. D.E. Aspnes, J.B. Theeten, F. Hottier, Investigation of effective-medium models of microscopic roughness by spectroscopic ellipsometry. Phys. Rev. B **20**, 3292 (1979)
4. R.G. Barrera, M. Delcastillomussot, G. Monsivais, P. Villasenor, W.L. Mochan, Phys. Rev. B **43**, 13819 (1991)
5. D. Beaglehole, O. Hunderi, Study of the interaction of light with rough metal surfaces i experiment. Phys. Rev. B **2**, 309 (1970)
6. D.A.G. Bruggeman, Ann. Physik **24**, 636 (1935)
7. A.L Barabási, H.E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995)
8. F. Everts, H. Wormeester, B. Poelsema, Optical anisotropy induced by ion bombardment of ag(001). Phys. Rev. B **78**, 155419 (2008)
9. F. Everts, H. Wormeester, B. Poelsema, Anomalous anisotropy in athermal bradley-harper roughening of cu(001). Phys. Rev. B **82**, 081415(R) (2010)
10. S.J. Fang, W. Chen, T. Yamanaka, C.R. Helms, Comparison of si surface roughness measured by atomic force microscopy and ellipsometry. Appl. Phys. Lett. **68**(20), 2837–2839 (1996)
11. H. Fujiwara, M. Kondo, A. Matsuda, Real-time spectroscopic ellipsometry studies of the nucleation and grain growth processes in microcrystalline silicon thin films. Phys. Rev. B **63**, 115306 (2001)

12. H. Fujiwara, J. Koh, P.I. Rovira, R.W. Collins, Assessment of effective-medium theories in the analysis of nucleation and microscopic surface roughness evolution for semiconductor thin films. Phys. Rev. B **61**, 10832–10844 (2000)
13. D. Franta, I. Ohlidal, Ellipsometric parameters and reflectances of thin films with slightly rough boundaries. J. Mod. Opt. **45**(5), 903–934 (1998)
14. D. Franta, I. Ohlidal, Comparison of effective medium approximation and rayleighrice theory concerning ellipsometric characterization of rough surfaces. Opt. Commun. **248**(4–6), 459–467 (2005)
15. H. Fujiwara, *Spectroscopic Ellipsometry* (Wiley, Chichester, 2007)
16. J.H. Jeffries, J.K. Zuo, M.M. Craig, Instability of kinetic roughening in sputter-deposition growth of pt on glass. Phys. Rev. Lett. **76**(26), 4931–4934 (1996)
17. R.P.U. Karunasiri, R. Bruinsma, J. Rudnick, Thin-film growth and the shadow instability. Phys. Rev. Lett. **62**(7), 788–791 (1989)
18. E. Kretschmann, E. Kröger, Reflection and transmission of light by a rough surface, including results for surface-plasmon effects. J. Opt. Soc. Am. **65**, 150–154 (1975)
19. W. Kaspar, U. Kreibig, Surface structure influences on the absorptance of thick silver films. Surf. Sci. **69**, 619 (1977)
20. E.S. Kooij, H. Wormeester, E.A.M. Brouwer, E. van Vroonhoven, A. van Silfhout, B. Poelsema, Langmuir **18**, 4401 (2002)
21. N. Kumar, A. Yanguas-Gil, S.R. Daly, G.S. Girolami, J.R. Abelson, Remote plasma treatment of si surfaces: enhanced nucleation in low-temperature chemical vapor deposition. Appl. Phys. Lett. **95**(14), 144107 (2009)
22. T. Lopez-Rios, Y. Borensztein, G. Vuye, Roughening of ag surfaces by ag deposits studied by differential reflectivity. Phys. Rev. B **30**, 659 (1984)
23. J.C. Maxwell Garnett, Colours in metal glasses and in metallic films. Philos. Trans. R. Soc. Lond. **203A**, 385 (1904)
24. S.G. Mayr, M. Moske, K. Samwer, Identification of key parameters by comparing experimental and simulated growth of vapor-deposited amorphous Zr(65)Al(7.5)Cu(27.5) films. Phys. Rev. B **60**(24), 16950–16955 (1999)
25. F. Ojeda, R. Cuerno, R. Salvarezza, L. Vazquez, Dynamics of rough interfaces in chemical vapor deposition: experiments and a model for silica films. Phys. Rev. Lett. **84**(14), 3125–3128 (2000)
26. George Palasantzas, Static and dynamic aspects of the rms local slope of growing random surfaces. Phys. Rev. E **56**, 1254–1257 (1997)
27. P. Petrik, L.P. Bir, M Fried, T. Lohner, R. Berger, C. Schneider, J. Gyulai, H. Ryssel, Comparative study of surface roughness measured on polysilicon using spectroscopic ellipsometry and atomic force microscopy. Thin Solid Films **315**(1–2), 186–191 (1998)
28. M. Pelliccione, T.-M. Lu, Mounded surfaces. In: *Evolution of Thin Film Morphology, Springer Series in Materials Science*, vol. 108 (Springer, New York, 2008), pp. 47–58
29. P. Petrik, T. Lohner, M. Fried, L.P. Biro, N.Q. Khanh, J. Gyulai, W. Lehnert, C. Schneider, H. Ryssel, Ellipsometric study of polycrystalline silicon films prepared by low-pressure chemical vapor deposition. J. Appl. Phys. **87**(4), 1734–1742 (2000)
30. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, vol. 111 (Springer, Berlin, 1988)
31. Brent A. Sperling, John R. Abelson, Simultaneous short-range smoothening and global roughening during growth of hydrogenated amorphous silicon films. Appl. Phys. Lett. **85**(16), 3456–3458 (2004)
32. B.A. Sperling, J.R. Abelson, Kinetic roughening of amorphous silicon during hot-wire chemical vapor deposition at low temperature. J. Appl. Phys. **101**, 024915 (2007)
33. A.H.M. Smets, W.M.M. Kessels, M.C.M. van de Sanden, Temperature dependence of the surface roughness evolution during hydrogenated amorphous silicon film growth. Appl. Phys. Lett. **82**(6), 865–867 (2003)
34. H. Wormeester, F. Everts, B. Poelsema, Plasmon resonance shift during grazing incidence ion sputtering on Ag(001). Thin Solid Films **519**, 2664 (2011)

35. A. Yanguas-Gil, J. Cotrino, A. Barranco, A.R. Gonzalez-Elipe, Influence of the angular distribution function of incident particles on the microstructure and anomalous scaling behavior of thin films. Phys. Rev. Lett. **96**(23), 236101 (2006)
36. A. Yanguas-Gil, J. Cotrino, A. Walkiewicz-Pietrzykowska, A.R. Gonzalez-Elipe, Scaling behavior and mechanism of formation of SiO_2 thin films grown by plasma-enhanced chemical vapor deposition. Phys. Rev. B **76**(7), 075314 (2007)
37. A. Yanguas-Gil, B.A. Sperling, J.R. Abelson, Theory of light scattering from self-affine surfaces: relationship between surface morphology and effective medium roughness. Phys. Rev. B **84**, 085402 (2011)
38. J.-T. Zettler, K. Haberland, M. Zorn, M. Pristovsek, W. Richter, P. Kurpas, M. Weyers, Real-time monitoring of movpe device growth by reflectance anisotropy spectroscopy and related optical techniques. J. Cryst. Growth **195**, 151–162 (1998)
39. Y. Zhao, G.-C. Wang, T.-M. Lu, *Characterization of Amorphous and Crystalline Rough Surface: Principles and Applications, Experimental Methods in Physical Sciences*, vol. 37 (Academic Press, New York, 2001)

Chapter 5
Plasmonics and Effective-Medium Theory

David E. Aspnes

Abstract The field of plasmonics has recently undergone a renaissance with the development of methods that can systematically produce structures on submicron scales. Plasmons, plasmon polaritons, and dielectric responses at the effective-medium level all involve screening charge that develops at the boundaries between different media, and hence are all related. The purposes of this work are to develop, from a fundamental perspective, these connections; to fill gaps that exist in standard treatments of electrodynamics; and to make contact with the relevant literature. The emphasis is on solutions of the homogeneous Maxwell Equations for composite materials and laminar samples. As a restatement of objectives, we outline what effective-medium theory can teach us about plasmonics; cast plasmonics in a form that can readily be taught to students; and provide introductory references to more extended treatments, for example recent reviews, the light-scattering literature describing grating anomalies, and approaches to metrology.

5.1 Introduction and Overview

When an external field is applied to a homogeneous material, forces on the atomic scale are generated on the charges of the material. In the classical picture, these forces displace the point charges q from their equilibrium positions, generating dipoles $\Delta \vec{p} = q \Delta \vec{r}$, where $\Delta \vec{r}$ is the change of position of a given q resulting from the local field at q [1]. We have neither the temporal nor spatial resolution to follow these

D. E. Aspnes(✉)
Department of Physics, North Carolina State University,
Raleigh, NC 27695-8202, USA

D. E. Aspnes
Department of Physics, Kyung Hee University,
Seoul 130-701, Republic of Korea
e-mail: aspnes@unity.ncsu.edu

displacements on the scale of the individual q, nor would this necessarily lead to useful information even if we could. However, by performing spatial averages on length scales that are large compared to atomic dimensions but small compared to those in the laboratory, we convert these atomic-scale displacements into a macroscopic average [2]. This averaging process yields a *dielectric function ε* that provides a macroscopic-scale description of the atomic-scale response of the material to the applied field. This is genuine progress, because ε can be used to describe macroscopic phenomena such as wave propagation and reflection through the Fresnel equations. Taking advantage of its dependence $\varepsilon(\omega)$ on frequency ω, one can also identify materials by spectral measurements of their transmittances and reflectances.

If the material is a composite consisting of dissimilar regions that are small compared to the wavelength λ of light but still large enough to possess their own dielectric identity, the local fields and the subsequent averaging process are now complicated by the presence of screening charge that develops at the boundaries between regions. This screening charge modifies the local fields and hence the polarizations within the different regions of the composite, and therefore its macroscopic average polarization. However, under certain conditions we can perform a second average, which again yields a macroscopic dielectric function $\langle\varepsilon(\omega)\rangle$, but one that now describes the macroscopic response of the composite to the applied field [3]. The result of this *effective-medium calculation* is an expression that gives $\langle\varepsilon(\omega)\rangle$ not only in terms of the dielectric functions $\varepsilon_a(\omega)$, $\varepsilon_b(\omega)$, ... of its constituent materials, but also of its structure. If the geometry is sufficiently simple so that a closed-form expression results, we can solve the inhomogeneous Maxwell Equations analytically to obtain expressions relating $\langle\varepsilon(\omega)\rangle$ to these material and structural parameters. A relation of this type is termed an *effective-medium theory* (EMT). The prototype EMT is that developed by Garnett [4, 5] in 1904 to explain the well-known red color of dispersions of Au nanoparticles in glass.[1] While the description of the response is now more complicated, it also invests optical measurements with substantially more diagnostic power. This capability is crucial in integrated-circuits technology, where optical measurements have become a nondestructive replacement for the analysis of materials, thin films, and structures used in these ubiquitous electronic assemblies [6].

Under conditions that also depend on the nature and structure of the constituents, the screening charge can oscillate without being externally driven, although an external drive is necessary to initiate the excitation. These excitations are solutions of the *homogeneous* Maxwell Equations, and are termed *plasmons*. A critical examination of these solutions under simplified but instructive conditions is one of the goals of this work. These form a much richer set than those of the corresponding inhomogeneous equations used to obtain the standard effective-medium expressions. These solutions correspond to poles in the macroscopic dielectric response of the composite, or even more generally to poles in expressions for reflectance or transmittance, and can be

---

[1] A famous example of the use of dispersed metal colloids to color glass is the Lycurgus Cup, crafted probably in Rome about the fourth century AD and currently on exhibit at the British Museum (www.britishmuseum.org.)

found by approaching them from this perspective. In the former case the poles are associated with the composite material itself, as for example the aforementioned dispersions of Au nanoparticles in glass. An example of the latter is the "surface" or more accurately *interface plasmon*, which requires only a planar interface between two materials with the appropriate dielectric responses [7]. The interface plasmon to which the term usually refers is an evanescent excitation that decays exponentially on both sides of the interface. However, a second class also exists, consisting of plane waves that radiate from both sides of the interface.

Why a pole in a response function should correspond to a self-sustaining excitation may be easier to accept if we note that a response, for example the reflectance, is the ratio of an outgoing to an incoming field. However, a stand-alone excitation generates an outgoing field with no incoming field. From this perspective it should not be surprising that these excitations influence the spectral dependences of reflectances and transmittances either on or off resonance, and that plasmons are already an intrinsic part of the Fresnel equations. When driven by an external plane wave, these excitations are termed *plasmon polaritons*.

Although the present work is concerned mainly with solutions of the homogeneous Maxwell Equations, to gain a better perspective, scattering needs to be mentioned as well. Scattering is the unavoidable result of incoming or outgoing waves interacting with spatially inhomogeneous systems. With or without effective-medium involvement, standard laminar-sample calculations using the Fresnel equations ignore scattering by assuming that the length scales of all inhomogeneities except layer thicknesses are small compared to λ. In this case reflectance and transmittance can be described by interactions involving the four basic types of plane waves at planar interfaces: incoming, reflected, back-reflected, and transmitted. Resonances are particularly effective at scattering, but for mathematically smooth interfaces the only resonances that can be accessed directly by plane-wave excitation are the effective-medium plasmons in the bulk. However, rough interfaces possess Fourier coefficients that can be obtained by Fourier-transforming their autocorrelation functions. These Fourier coefficients can couple incoming and outgoing waves to interface plasmons in second order, thereby diverting energy to their creation [7]. This process results in plasmonic features in reflection spectra. The diffraction grating represents an extreme case, where the objective is to divert as much of the incoming radiation as possible into scattered radiation. Here, the inhomogeneities are not only of the scale of λ but also periodic. The Hessell and Oliner model [8], which is generally considered to be the landmark for quantitative analysis of grating structures, invokes plasmons directly by viewing the grating as a waveguide interacting with a resonant cavity. The incident and reflected waves correspond to the propagating modes of the waveguide, and the role of the resonant cavity is taken by the plasmon.

In the bulk, effective-medium theory crosses over to scattering in the analysis of critical dimensions (CD) in integrated-circuits technology. CD analysis has reached the stage where optical CD measurements have taken over functions that previously required scanning electron microscopy (SEM) or transmission electron microscopy

(TEM) [6]. The landmark analytic connection is generally considered to be the work of Moharam et al. [9], who model patterned layers as a laminar stack, with the (spatially dependent) dielectric response of a given lamination expressed as a Fourier series. By matching data obtained over a wide spectral range in the visible-near uv using rotating-compensator polarimetry, the coefficients of the different laminations, and hence their dielectric properties and structural dimensions, can be determined. The issue reduces to one of computational speed, although this remains a continuing challenge because scattering calculations are formidable.

The current resurgence of interest in plasmonics (at least the third such) is driven by several factors. First, plasmons increase local fields at boundaries. Attaining the highest possible fields is important in a number of areas, for example surface enhanced Raman scattering (SERS) [10]. Second, the length scales of plasmons are determined by the length scales of the inhomogeneities themselves, thus nominally providing a basis for more compact information-processing technologies [11]. Third, spatial inhomogeneities provide another means of engineering refractive indices to create layers that can index-match other materials and thus for example enhance efficiencies of solar cells [12]. Fourth, and most important, with current lithographic technology we can actually fabricate structures of dimensions considerably less than the wavelength of visible light.

This recent resurgence of interest tends to obscure the fact that plasmonics is a field that is well over a century old. The first papers on *radiative* plasmonic effects are those of Garnett [4, 5]. However, *nonradiative* (interface-plasmon) solutions to Maxwell's Equations have been known since the work of Wood [13] and the follow-up calculations by Wood [14] and Sommerfeld [15]. These *surface plasma waves* (SPWs) were more extensively investigated in the 1960s and early 1970s. Noteworthy treatments include those of Kretschmann [16] and Kretschmann and Kröger [17], Abeles [18], and Otto [19]. Another major resurgence of interest occurred with the discovery of surface-enhanced Raman scattering (SERS) by Fleischman et al. in 1974 [10]. This work was followed by hundreds of papers discussing not only the theory but also methods of systematically roughening surfaces, for example with lithography, to generate the spatial inhomogeneities necessary to control local fields (see for example the theoretical treatments by Kerker et al. [20] and references therein, Gersten and Nitzan [21] and references therein, Moscovitz [22], and Hulteen et al. [23]). Numerous general reviews of plasmonics are also available. Raether's monograph [7] is an excellent summary of knowledge up to its date of publication, with a major focus on interface roughness. Murray and Barnes [24] give a summary of the field from the perspective of materials, Atwater and Polman [11] from applications, and Zhang and Zhang from nanoengineered structures [25].

Given this extensive background, it is unfortunate that the only contact that most students have with spatial inhomogeneity in traditional electricity and magnetism courses is the unphysical cavity derivation of the Clausius-Mossotti relation [26]. Yet a minor extension of the traditional problem of a dielectric sphere in a dielectric medium in a uniform electric field is all that is necessary to introduce plasmonics.

One need only to look for solutions with *no* applied field, which are easy to find as shown below. The derivation of the Maxwell Garnett EMT requires the extra step of spatially averaging the textbook solution, which is surprisingly more difficult than it sounds. My point is not to do detailed calculations or a provide comprehensive review, but to develop the topics at a relatively low level and provide some examples.

5.2 Dielectric Functions and Effective-Medium Theories

A. Basic Equations

We suppose that the material of interest consists of point charges q that respond in some way to applied fields. The simplest classical equation of motion for a given q involves the Lorentz and restoring forces, specifically [27]

$$\vec{F} = m\vec{a} = m\frac{d^2\vec{r}}{dt^2} = q\vec{E} + \frac{q}{c}\vec{v} \times \vec{B} - \kappa \cdot \Delta\vec{r}, \tag{5.1}$$

where $\vec{r}$ is the position of q, $\vec{v} = d\vec{r}/dt$ is the velocity of q, $\vec{E}$ and $\vec{B}$ are the local electric and magnetic fields at q, and κ is a restoring-force tensor. $\vec{E}$ and $\vec{B}$ satisfy the microscopic Maxwell Equations

$$\nabla \cdot \vec{E} = 4\pi\rho; \tag{5.2a}$$

$$\nabla \times \vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0; \tag{5.2b}$$

$$\nabla \cdot \vec{B} = 0; \tag{5.2c}$$

$$\nabla \times \vec{B} = 4\pi\vec{J} + \frac{1}{c}\frac{\partial\vec{E}}{\partial t}, \tag{5.2d}$$

where ρ and $\vec{J} = \rho\vec{v}$ are the microscopic charge and current densities, respectively. We recast these into a form more convenient for solution by defining vector and scalar potentials $\vec{A}$ and ϕ such that $\vec{B} = \nabla \times \vec{A}$ and $\vec{E} = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t} - \nabla\phi$. In the Lorentz gauge $\nabla \cdot \vec{A} + \frac{1}{c}\frac{\partial\phi}{\partial t} = 0$, and the Maxwell Equations reduce to the Helmholtz Equation

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)(\phi, \vec{A}) = -\frac{4\pi}{c}(c\rho, \vec{J}), \tag{5.3}$$

where $(\phi, \vec{A})$ and $(c\rho, \vec{J})$ are the 4-potential and 4-current, respectively. We return to Eq. (5.3) when discussing limits to the validity of EMTs.

B. Atomic-Scale Averaging

We assume that we are deep inside a macroscopically homogeneous medium, far from any boundaries. Our objective is to connect atomic-scale quantities to macroscopic observables by spatial averaging. Following the electrostatics approach used by Ashcroft and Mermin [2], we define for example the macroscopic equivalent $\langle \vec{E}(\vec{r}) \rangle$ of the microscopic field $\vec{E}(\vec{r}')$ according to

$$\langle \vec{E}(\vec{r}) \rangle = \int_{-\infty}^{\infty} d^3 r' W(\vec{r} - \vec{r}') \vec{E}(\vec{r}'), \qquad (5.4)$$

where $W(\vec{r} - \vec{r}')$ is an averaging function that is positive definite, varies slowly with a characteristic length L on the atomic scale but possibly rapidly on the laboratory scale, integrates to 1, and depends only on the spatial difference between the source point $\vec{r}'$ and the observation point $\vec{r}$. Although vague, this definition is sufficient. From a different perspective, the above convolutions of two functions in direct space can be replaced by the product of their Fourier transforms in reciprocal space. The slow spatial variation of $W(\vec{r} - \vec{r}')$ ensures that its Fourier transform is highly localized about zero, thereby largely eliminating a substantial fraction of the wide range of Fourier coefficients of $\vec{E}(\vec{r}')$, which varies rapidly on the atomic scale. Thus $W(\vec{r} - \vec{r}')$ effectively selects only a small range of the Fourier coefficients of $\vec{E}(\vec{r}')$ about zero.

By defining a new integration variable $\vec{r}'' = \vec{r} - \vec{r}'$, taking the divergence of both sides of Eq. (5.3), and performing some mathematics that depend on the integration limits being infinite, we obtain the macroscopic version of Coulomb's Law: $\nabla \cdot \langle \vec{E} \rangle = 4\pi \langle \rho \rangle$. While correct, this is not particularly informative because $\langle \rho \rangle$ includes the change of the original charge density $\langle \rho_o \rangle$ that is the polarization induced by $\langle \vec{E} \rangle$. Since this induced polarization is presumably linear in $\langle \vec{E} \rangle$, we can isolate its contribution. We do this by writing $\langle \rho \rangle = \langle \rho_o \rangle + (\langle \rho \rangle - \langle \rho_o \rangle)$, then considering the difference, noting that for neutral materials $\langle \rho_o \rangle = 0$. If q is a point charge located at $\vec{R}_j$, then $\rho_o(\vec{r}') = \sum_j q \delta(\vec{r}' - \vec{R}_j)$. When $\langle \vec{E} \rangle$ is applied, $\vec{R}_j \rightarrow \vec{R}_j + \Delta \vec{R}_j$, where $\Delta \vec{R}_j$ can be obtained from a suitable model. The induced dipoles themselves are $q \Delta \vec{R}_j$. The δ-function makes the integral trivial, and we convert the resulting sum to an integral, introducing in the process the volume density n of q. Finally, we perform a Taylor-series expansion of the result. After some additional mathematics we find

$$\nabla \cdot \langle \vec{E} \rangle = 4\pi \langle \rho_o \rangle - 4\pi \nabla \cdot \langle \vec{P} \rangle, \qquad (5.5)$$

where

$$\langle \vec{P} \rangle = \int_{-\infty}^{\infty} d^3 r' W(\vec{r} - \vec{r}') n(\vec{r}') q \Delta \vec{R}_j \qquad (5.6)$$

is the macroscopic *dipole density*. Since by assumption $\Delta \vec{R}_j \sim \langle \vec{E} \rangle$, we can bring the polarizability term to the left side of Eq. (5.5), defining the macroscopic *displacement field* $\langle \vec{D} \rangle$ and therefore ε according to $\langle \vec{D} \rangle = \varepsilon \langle \vec{E} \rangle$. We thus not only obtain ε but also develop a procedure for calculating it in terms of atomic-scale parameters. At the same time we have laid the groundwork for the description of anisotropic and nonlinear effects, which are obtained by suitably extending $\Delta \vec{r}$, in the latter case by including higher-order terms.

C. Effective-Medium Theories

EMTs result from a second average. Rather than make use of the weighting function explicitly, the standard procedure is to assume that the material has been placed in an external field $\vec{E}$, then to solve Laplace's Equation exactly for the microscopic electrostatic field $\vec{E}(\vec{r}')$ in terms of the geometry of the configuration and the dielectric functions of its constituents. The external field is assumed to be a constant, $\vec{E} = \vec{E}_o$, being generated for example by well-separated, large-area, parallel capacitor plates. The resulting $\vec{E}(\vec{r}')$ is then volume-averaged along with the microscopic displacement-field $\vec{D}(\vec{r}')$, which is obtained by evaluating $\vec{D}(\vec{r}') = \varepsilon(\vec{r}')\vec{E}(\vec{r}')$, where $\varepsilon(\vec{r}')$ is the value of the dielectric function at $\vec{r}'$. Although we write $\varepsilon = \varepsilon(\vec{r}')$ as if it were a continuous function of $\vec{r}'$, the calculations that follow require ε to be constant within a given region and to change discontinuously at the boundaries between one region and another. The results of the averaging processes are $\langle \vec{D} \rangle$ and $\langle \vec{E} \rangle$, and ε is obtained from $\langle \vec{D} \rangle = \varepsilon \langle \vec{E} \rangle$.

For a real composite an analytic solution is generally impossible. Therefore, we give it a configuration that we can solve, and use limit theorems to estimate possible ranges of error. One easily solvable configuration is a laminar stack of alternating layers of materials of dielectric functions ε_a and ε_b. For $\vec{E}$ parallel to the laminations the continuity condition on tangential $\vec{E}$ yields $\vec{E}(\vec{r}) = \vec{E}_o$, and therefore $\langle \vec{E} \rangle = \vec{E}_o$. The corresponding calculation for $\vec{D}(\vec{r})$ yields $\varepsilon_a \vec{E}_o$ or $\varepsilon_b \vec{E}_o$ according to whether $\vec{r}$ is in material a or b. Performing the volume average we find that $\langle \vec{D} \rangle = f_a \varepsilon_a \vec{E}_o + f_b \varepsilon_b \vec{E}_o$, where f_a and f_b are the volume fractions of materials a and b, where $f_a + f_b = 1$. The EMT describing this configuration is therefore

$$\varepsilon = f_a \varepsilon_a + f_b \varepsilon_b. \tag{5.7}$$

For $\vec{E}$ perpendicular to the laminations, the boundary condition on normal $\vec{D}$ yields a constant value $\vec{D}(\vec{r}) = \vec{D}_o$ throughout, and an analogous calculation yields

$$\frac{1}{\varepsilon} = \frac{f_a}{\varepsilon_a} + \frac{f_b}{\varepsilon_b}. \tag{5.8}$$

In neither case do the layer thicknesses appear, so in this approximation layers may be of any thickness. For the perpendicular-field case $1/\varepsilon$ becomes singular if either ε_a or ε_b goes to zero. We shall say more about this below.

Except for field orientation the laminar configurations are identical, yet the EMTs describing them are different. This is a consequence of screening charge. From the Maxwell Equations and the assumption that ε is constant within a given region, the screening charge σ can be calculated from

$$\sigma = \frac{1}{4\pi}\hat{n} \cdot (\vec{E}_b - \vec{E}_a), \tag{5.9}$$

where $\hat{n}$ is the unit normal vector at the interface and $\vec{E}_a$ and $\vec{E}_b$ are the local fields on either side. For fields parallel to the boundaries it is evident that no screening charge exists, so the EMT expression for ε reduces in this case simply to a weighted average of the dielectric responses of the two constituents. The circuit equivalent is capacitors in parallel. For fields perpendicular to the boundaries the screening charge is clearly maximized, and the weighted average leading to ε now involves reciprocals (capacitors in series). Since there can never be less screening charge than no screening charge, nor more than that which occurs when all boundaries are perpendicular to the field, these two cases are extremes. When plotted as a function of composition for any given ε_a and ε_b, the resulting loci form the Wiener limits [28]. For any two-component composite, ε must lie in the region bounded by these two loci.

The above also represent limiting cases of Maxwell Garnett (MG) EMTs, which follow from intermediate solvable configurations. The standard example follows from the standard introductory electrostatics problem of a dielectric sphere of radius a and dielectric function ε_b embedded in a "host" material of dielectric function $\varepsilon_h = \varepsilon_a$. The relevant expressions for the scalar potential ϕ inside and outside the sphere are

$$\phi_{in} = \sum_{l=0}^{\infty} A_l \left(\frac{r}{a}\right)^l P_l(\cos\theta); \tag{5.10a}$$

$$\phi_{out} = \sum_{l=0}^{\infty} B_l \left(\frac{a}{r}\right)^{l+1} P_l(\cos\theta) - Ez, \tag{5.10b}$$

where the $P_l(x)$ are Legendre polynomials. Applying the boundary conditions on tangential $\vec{E}$ and normal $\vec{D}$ and performing volume averages of the resulting electric and displacement fields we find

$$\frac{\varepsilon - \varepsilon_h}{\varepsilon + 2\varepsilon_h} = f_a \frac{\varepsilon_a - \varepsilon_h}{\varepsilon_a + 2\varepsilon_h} + f_b \frac{\varepsilon_b - \varepsilon_h}{\varepsilon_b + 2\varepsilon_h}. \tag{5.11}$$

In our example the host dielectric function $\varepsilon_h = \varepsilon_a$, so the first term on the right vanishes. In the Bruggeman mean-field approximation [29] we set $\varepsilon_h = \varepsilon$, and the left side of Eq. (5.11) vanishes.

The quantity 2 in Eq. (5.11) is a particular value of the *depolarization factor p*. Rewriting Eq. (5.11) more generally gives

$$\frac{\varepsilon - \varepsilon_h}{\varepsilon + p\varepsilon_h} = f_a \frac{\varepsilon_a - \varepsilon_h}{\varepsilon_a + p\varepsilon_h} + f_b \frac{\varepsilon_b - \varepsilon_h}{\varepsilon_b + p\varepsilon_h}. \tag{5.12}$$

Other values of p for simple configurations include $p = 0$ and ∞ for laminar samples where the applied field is perpendicular and parallel, respectively, to the laminations, and $p = 1$ for inclusions of cylindrical symmetry. In particular, the values $p = 0$ and ∞ reduce Eq. (5.12) to Eqs. (5.8) and (5.7), respectively, so the laminar-sample solutions are limiting cases of the MG EMT. In particular, Eq. (5.12) becomes singular when $\varepsilon_a + p\varepsilon_h$ or $\varepsilon_b + p\varepsilon_h$ vanishes. Again, we will say more about this below.

5.3 Plasmonics

As noted in the Introduction, plasmons are resonances that are solutions of the homogeneous Maxwell Equations. We now look for EMT solutions in the electrostatic (long-wavelength) limit.

A. *Parallel Plates*

The simplest case is the configuration where a material of dielectric function ε is placed between two ideal conducting plates. For the field direction parallel to the plates the boundary condition on normal $\vec{D}$ is satisfied trivially. Because there can be no electric field parallel to the surface of a conductor, the boundary condition on tangential $\vec{E}$ requires $\vec{E} = 0$ throughout. Hence there is no nontrivial solution for this field direction.

We now contrast this to the situation where $\vec{E}$ is perpendicular to the plates. Here, the boundary condition on tangential $\vec{E}$ is satisfied trivially. However, the boundary condition $\vec{D} = 0$ on normal $\vec{D}$ requires that $\vec{D} = \varepsilon\vec{E} = 0$ between the plates. This equation can be satisfied either by setting $\vec{E} = 0$, which is again a trivial solution of no interest, or requiring that $\varepsilon = 0$. Thus if some *plasma frequency* $\omega = \omega_p$ exists such that $\varepsilon(\omega) = \varepsilon(\omega_p) = 0$, then an internal field can exist even though there is no external applied field. The resulting solution is the *bulk plasmon*. The bulk plasmon also follows from Eq. (5.8), where $1/\varepsilon = E/D$ becomes singular if either $\varepsilon = \varepsilon_a$ or ε_b is equal to zero.

This simple example also shows that these solutions cannot be static, but are resonances that occur at specific nonzero frequencies. By Eq. (5.9), they are also characterized by screening charges σ that develop at interfaces. In this laminar case, the screening charges obviously have opposite signs on the two sides. The physics behind the absence of a solution for fields parallel to the boundary is now obvious: no screening charge, no resonance.

The plasmon also obviously has energy associated it. This is stored alternately in the electric field at maximum $\vec{E}$ and in the magnetic field at maximum current density $\vec{J}$. This simple picture remains valid as long as retardation effects are negligible. If the spacing d between plates is too large, then the bulk plasmon is no longer supported. For appropriate values of dielectric functions and sufficiently large d, waveguide modes appear for both orientations of the electric field. However, these are not plasmon solutions and will not be discussed further.

B. Sphere

We consider next the embedded-sphere configuration used to derive the Maxwell Garnett EMA. With no external driving field to impose a symmetry constraint on allowed solutions, Eqs. (5.10) take the more general form

$$\phi_{in}(\vec{r}) = \sum_{l=0}^{\infty} A_l \left(\frac{r}{a}\right)^l Y_l^m(\theta, \varphi), \tag{5.13a}$$

$$\phi_{out}(\vec{r}) = \sum_{l=0}^{\infty} A_l \left(\frac{a}{r}\right)^{l+1} Y_l^m(\theta, \varphi). \tag{5.13b}$$

where the $Y_l^m(\theta, \varphi)$ are spherical harmonics. In Eqs. (5.13a, b) the coefficients are chosen so that ϕ is continuous at $r' = a$, which is a more efficient version of the usual continuity condition on tangential $\vec{E}$. Invoking the boundary condition on normal $\vec{D}$ we obtain

$$A_l(l\varepsilon_b + (l+1)\varepsilon_a) = 0. \tag{5.14}$$

No condition is imposed on m.

Equation (5.14) again presents opportunities for nontrivial solutions. For a given l either $A_l = 0$, or there exists an ω such that $\varepsilon_b(\omega) = -(1 + 1/l)\varepsilon_a(\omega)$ [7]. If this is the case, then for any l we have a set of solutions for integer m ranging from $m = -l$ to $m = +l$. We thus find an entire host of solutions, each of which corresponds to a different plasmon. Note that this set includes, but goes well beyond, the $l = 1$ plasmon singularity that appears in the MG EMT at $\varepsilon_b(\omega_p) = -2\varepsilon_a$ for the sphere. The absence of any other solutions in the MG EMT case follows because the MG EMT was derived assuming the configuration was driven by a uniform external

field, so we obtain only the solution $l = 1, m = 0$. The resonance is evanescent, because the exterior field converges as $1/(r^{l+1})$ for all l.

Spheroidal configurations can also be solved analytically [7]. For rotationally symmetric shapes the $l = 1$ plasmon appears as two separate modes, one of which is doubly degenerate.

C. Cylinder

In this geometry the generic solution for the potential is

$$\phi_{in} = \sum_{m=0}^{\infty} \left(\frac{r}{a}\right)^m [A_m \cos m\varphi + B_m \sin m\varphi], \quad (5.15a)$$

$$\phi_{out} = \sum_{m=0}^{\infty} \left(\frac{a}{r}\right)^m [C_m \cos m\varphi + D_m \sin m\varphi] \quad (5.15b)$$

Following the same procedure as above, the condition for a nontrivial solution with no externally applied field is

$$\varepsilon_b(\omega) = -\varepsilon_a(\omega), \quad (5.16)$$

valid for any m and either harmonic function cos or sin. The solution is again evanescent: for a given m the exterior field converges as $1/r^m$.

D. Shape and Size Effects: Limits to the Validity of the Long-Wavelength Approximation

The above calculations show that particle shapes determine the value that ε must take at the singularities. To make the discussion more specific, we make the usual approximation that the effect of free carriers on the polarizability is adequately summarized by the Drude expression

$$\varepsilon_s(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2}, \quad (5.17)$$

where ω_p is the plasma frequency and $\varepsilon_\infty > 0$ is the *infrared dielectric constant*. Equation (5.17) allows us to bring frequency explicitly into the calculation. While useful for describing qualitative behavior, Eq. (5.17) neglects lifetime effects on the free carriers, which are dictated by scattering. Scattering typically arises from two basic phenomena: finite mean-free paths of carriers in the metal, and inelastic interactions between carriers and the surface. Both act to broaden resonance lineshapes, as discussed in more detail by Kreibig [30]. Viewing Eqs. (5.7), (5.8), (5.14), and

(5.16) in the context of Eq. (5.17), we see that a decrease in screening (increase in p) leads to a red shift of the resonance for a given material.

Regarding particle sizes, the above effective-medium and plasmonics calculations were done in the approximation $c \to \infty$, where retardation is neglected. For finite frequencies the Laplace Equation of electrostatics must be replaced by the Helmholtz Equation, with Cartesian solutions consisting of plane waves $\phi \sim \vec{E} \sim e^{ikz}$, where $k^2 = \omega^2 \varepsilon / c^2$ is the square of the wave vector $\vec{k}$. From Eqs. (5.2a–d) and (5.3), we can also view the use of $\nabla \cdot \vec{E} = 4\pi\rho \approx -\nabla^2\phi$ as being equivalent to using an incorrect value of ρ.

We can estimate where particle size becomes important by calculating the macroscopic average of e^{ikz} from Eq. (5.4). Taking a simple normalized decreasing exponential $e^{-|z|/L}/2L$ for W, the average of e^{ikz} about $z = 0$ is not 1 but $(1 + k^2 L^2)^{-1}$. Thus to lowest order the errors go as $k^2 L^2$. This is negligible for atomic-scale averages but becomes important when particle dimensions become comparable to L. For example the lowest nontrivial solution of the Helmholtz Equation for the interior potential of the sphere is

$$\phi_{in}(z) \sim \ j_1(kz) = \frac{\sin(kz)}{(kz)^2} - \frac{\cos(kz)}{kz}, \tag{5.18}$$

where $j_1(kr)$ is the spherical Bessel function of order 1. Consequently, the MG EMT fails where the approximation

$$j_1(ka) = \frac{ka}{3}(1 - \frac{(ka)^2}{10} + \cdots) \approx \frac{ka}{3} \tag{5.19}$$

is no longer valid. Since $ka = 2\pi na / \lambda$, the factor $(2\pi n)$ causes this to occur at surprisingly small values of a.

The correct treatment of finite-wavelength effects for isolated inclusions follows the above procedure, but requires the use of solutions of the Helmholtz equation for both interior and exterior waves. For the spherical case, the exterior waves are given by Hankel functions. Numerical calculations for simple metals show that resonances red-shift with increasing particle size, along with changes in lineshapes that are a consequence of the dependences of the different scattering mechanisms on particle size. Representative results for some simple metals are given by Derkachova and Kolwas [31].

Several of these points are illustrated in Fig. 5.1, which shows the resonant emissions of Ag particles of different sizes and shapes [24]. The top row shows the TEM micrographs of the particles below. The spectrum shifts from the blue with large screening to the yellow with reduced screening. The fourth (red) emission derives from the same shape as that of the third particle, but the fourth particle is larger.

Interference causes additional difficulties, even for well-separated particles. Although local polarizations depend only on local fields, interference can affect these
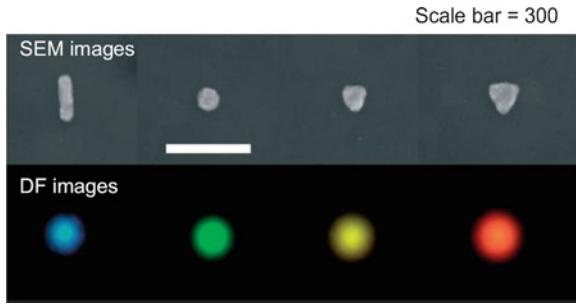
Fig. 5.1 Resonant emission of Ag particles of different shapes and sizes, illustrating the effect of these parameters on resonance wavelengths. The *top* row shows the transmission electron micrographs of the particles (after Ref. [24])

values and hence the weightings and the overall averages. We take interference explicitly into account when evaluating reflectances and transmittances of laminar stacks. But even in the MG case, where the inclusions are assumed to be well separated, scattered fields may add coherently in certain directions. Thus other requirements for the validity of EMTs are that contributions to fields resulting from scattering outside the averaging region are small and phase-incoherent. The diffraction grating is the obvious example where these conditions are not satisfied.

5.4 Interfaces and Laminar Samples

Planar configurations present interesting challenges of a somewhat different nature. In what follows we identify plasmons by an unconventional approach: examining singularities in reflectance expressions.

A. Interfaces

The nonradiative interface plasmon at a planar interface between two semi-infinite media of dielectric functions ε_a (ambient) and ε_s (substrate) is the classic case. It is particularly interesting not only from a physical but also from a mathematical perspective. For TE (s-polarized) light the reflectance r_{TE} is given by

$$r_{TE} = \frac{n_{a\perp} - n_{s\perp}}{n_{a\perp} + n_{s\perp}}, \tag{5.20a}$$

where

$$n_{s\perp} = \sqrt{\varepsilon_s - \varepsilon_a \sin^2\theta}, \tag{5.20b}$$

etc., where θ is the angle of incidence. The denominator can never be zero in this case simply because $n_{a\perp}$ and $n_{s\perp}$ are both restricted to the first quadrant of the complex plane. Thus no pole, and no plasmon, exists for TE polarization.

However, the situation is different for TM (p-) polarization. Here

$$r_{TM} = \frac{\varepsilon_s n_{a\perp} - \varepsilon_a n_{s\perp}}{\varepsilon_s n_{a\perp} + \varepsilon_a n_{s\perp}}, \tag{5.21}$$

We set the denominator equal to zero,

$$\varepsilon_s n_{a\perp} = -\varepsilon_a n_{s\perp}, \tag{5.22a}$$

then square the result to eliminate square-root operations. We find

$$\varepsilon_s^2 (\varepsilon_s - \varepsilon_a) = (\varepsilon_s^2 - \varepsilon_a^2) \sin^2 \theta. \tag{5.22b}$$

Cancelling a common factor we have

$$\sin^2 \theta = \frac{\varepsilon_s}{\varepsilon_s + \varepsilon_a}, \tag{5.23a}$$

$$\cos^2 \theta = \frac{\varepsilon_a}{\varepsilon_s + \varepsilon_a}, \tag{5.23b}$$

so

$$\tan \theta_B = n_s / n_a, \tag{5.23c}$$

where θ_B is the Brewster angle.

But this is not what we expected, because the Brewster angle results from setting the numerator, not the denominator, equal to zero. What went wrong? The difficulty clearly arises when we square Eq. (5.22a) to obtain Eq. (5.22b), so we need to examine Eq. (5.22a) more carefully. This equation can be satisfied in principle if ε_s is negative. But by Eq. (5.20b), $n_{s\perp}$ must then be pure imaginary. We can accept this if $n_{a\perp}$ is also pure imaginary. However, at first sight Eq. (5.20b) indicates that this is impossible, because it would require $\sin^2 \theta > 1$. We therefore need to examine the origin of Eq. (5.21).

Equation (5.21) is the result of combining the dispersion equations, which specify the squares of the propagation vectors in the ambient and substrate, and the continuity condition at the $s - a$ interface, which requires that their parallel components are equal. Defining $n_{a||} = ck_{a||}/\omega$ etc., the dispersion equations can be written

$$\varepsilon_a = n_{a||}^2 + n_{a\perp}^2, \tag{5.24a}$$

$$\varepsilon_s = n_{s||}^2 + n_{s\perp}^2 \tag{5.24b}$$

where the continuity condition requires $n_{a||} = n_{s||} = n_{||}$. By formulating the problem in this way we eliminate any reference to θ. Applying Eq. (5.22a) to Eqs. (5.24a, b) leads to

$$n_{a||}^2 = n_{s||}^2 = n_{||}^2 = \frac{\varepsilon_s \varepsilon_a}{\varepsilon_s + \varepsilon_a}; \tag{5.25a}$$
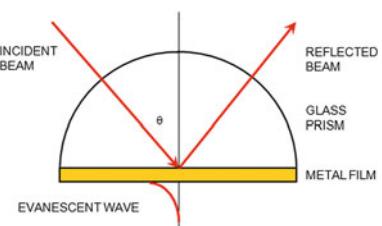
$$n_{s\perp}^2 = \frac{\varepsilon_s^2}{\varepsilon_s + \varepsilon_a}; \tag{5.25b}$$

$$n_{a\perp}^2 = \frac{\varepsilon_a^2}{\varepsilon_s + \varepsilon_a}. \tag{5.25c}$$

We see immediately that a solution does in fact exist for which ε_s is negative and where $k_{a||} = k_{s||} = k_{||}$ is real. It is only necessary to require that $\varepsilon_s < -\varepsilon_a$. From this it follows that $k_{a||}^2 > k_a^2$, so $k_{a\perp}$ is imaginary. Since ε_s is already negative, $k_{s\perp}$ is also imaginary. Therefore, the excitation is an evanescent wave that decays exponentially on both sides of the interface. In the context of Eq. (5.20b) we see that this also requires $\sin^2\theta > 1$. However, since we expect $\sin^2\theta \leq 1$, the reason for not initially recognizing the plasmonic solution becomes clear. We have simply used the wrong notation for the relevant variable.

In practice $k_{a||}^2 > k_a^2$ cannot be achieved by illuminating the interface in an isotropic ambient, and special measures must be taken. One approach is the Kretschman configuration [16], which is illustrated in Fig. 5.2. Here, measurements are made on a thin metal film deposited on a prism of a suitably large refractive index by illuminated the film through the prism. By changing the angle of incidence in the prism, we change the projection angle and therefore $k_{||}$. For the typical case of an air ambient, a sufficiently large angle of incidence in the prism allows us to generate a $k_{||}$ that exceeds anything that can be achieved by illuminating the film in air, so the conditions for the existence of the interface plasmon at the air-film interface are realized. As long as the metal film is not too thick, sufficient radiation can penetrate the film to excite the plasmon. In this three-phase ambient/overlayer/substrate configuration the prism is the ambient. However, in the mathematics that follow, we shall consider the air as the ambient, the film as the substrate, and the projection $k_{||}$ as an independent variable that can exceed k_a.

Fig. 5.2 Kretschmann configuration for exciting an interface plasmon at the air-metal interface by illumination through a prism

We note that the denominator of Eq. (5.21) can also equal zero for $k_{a\perp}$ and $k_{s\perp}$ both real. This solution of the homogeneous equation corresponds to plane waves that propagate *outward* from both sides of the interface. However, Eqs. (5.22a, b) cannot be satisfied with one propagating and one evanescent wave.

B. Dispersion of Interface Plasmons

With $n_{a||}$ as the independent variable, Eqs. (5.25a–c) show that with a given $n_{a||}$ and $\varepsilon_s < -\varepsilon_a$ the homogeneous equation has a solution only for a particular value of ε_s, i.e., for a particular value of ω. To investigate its behavior we assume that $\varepsilon_s(\omega)$ is given by Eq. (5.17), which we now write as

$$\varepsilon_s(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2} = \varepsilon_\infty - x^2, \tag{5.26}$$

where $x = \omega_p/\omega$ is defined for convenience. With this definition Eqs. (5.25a–c) become
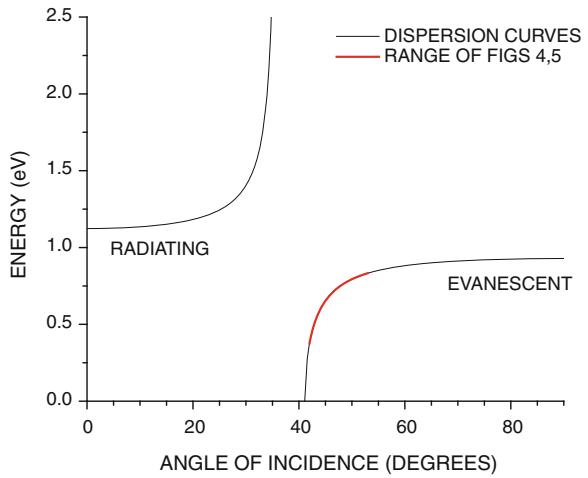
$$n_{a||}^2 = n_{s||}^2 = n_{||}^2 = \frac{c^2 k_{||}^2}{\omega^2} = \frac{\varepsilon_s \varepsilon_a}{\varepsilon_s + \varepsilon_a} = \frac{\varepsilon_a(\varepsilon_\infty - x^2)}{\varepsilon_a + \varepsilon_\infty - x^2}. \tag{5.27}$$

We identify 3 ranges: for sufficiently small ω, $x^2 > \varepsilon_a + \varepsilon_\infty$, Eq. (5.27) is positive and $n_{||}$ is real; for $\varepsilon_a + \varepsilon_\infty > x^2 > \varepsilon_\infty$ Eq. (5.27) is negative and $n_{||}$ is imaginary; and for sufficiently large ω, $\varepsilon_\infty > x^2$ Eq. (5.27) is again positive and $n_{||}$ is real. The first range is the plasmon; the second violates the requirement that $n_{a\perp}$ and $n_{s\perp}$ must have the same phase, and the third corresponds to the solution with two outgoing plane waves.

The solution of Eq. (5.27) for energy as a function of angle of incidence in Fig. 5.3 for $\varepsilon_a = 1$, $E_p = 2.19\,\text{eV}$, $\varepsilon_\infty = 3.80$, and $\varepsilon_{BK-7} = 2.31$, which are parameters describing the example data that we will present in Sect. D. The higher-energy radiative branch begins at $n_{||} = 0$ and terminates at $n_{||}^2 = \varepsilon_s \varepsilon_\infty (\varepsilon_\infty + \varepsilon_s) \cong 0.79$, which occurs at $\theta \cong 35.5°$. The lower-energy plasmon branch begins at the pole at $x^2 = \varepsilon_a + \varepsilon_\infty$, which occurs at $\theta \cong 41.1°$, and ends at an energy $E \cong 0.93\,\text{eV}$. We plot the solution in this form for easier comparison to the data of Fig. 5.4. These solutions are usually plotted as $k_{||}$ versus E.

With a finite real value of $k_{||}$, the plasmon propagates along the interface. If ε_s has an imaginary part, then the interface plasmon is attenuated in the propagation direction with a characteristic length $L = 1/(2\text{Im}(k_{||}))$. The situation for small losses is discussed in Raether [32]. We note that the cylindrical geometry discussed in Sect. C also has propagating solutions, where the radial dependences of the fields are Hankel functions.

Fig. 5.3 Solution of Eq. (5.27) as described in the text. The *red* segment shows the result in the angle-of-incidence range of Figs. 5.4 and 5.5

C. Three-Phase (Substrate/Overlayer/Ambient) Systems

In the above section the existence of three phases was incidental to the generation of the interface plasmon at the air-overlayer interface. However, when all three phases are considered, new possibilities emerge. The reflectance r_{soa} of the three-phase substrate/overlayer/ambient system for either TE or TM polarization is given by the Airy equation [33]

$$r_{soa} = \frac{r_{oa} + r_{so}e^{2ik_{o\perp}d}}{1 + r_{so}r_{oa}e^{2ik_{o\perp}d}}, \tag{5.28}$$

where d is the overlayer thickness and $r_{oa}^{TE} = \frac{n_{a\perp} - n_{o\perp}}{n_{a\perp} + n_{o\perp}}$ and $r_{oa}^{TM} = \frac{\varepsilon_o n_{a\perp} - \varepsilon_a n_{o\perp}}{\varepsilon_o n_{a\perp} + \varepsilon_a n_{o\perp}}$ are the *oa* reflectances for TE- and TM-polarized light, respectively, with corresponding expressions for the *so* interface. The notation is the same as that used in the previous section. Setting the denominator equal to zero yields a variety of solutions ranging from the bulk plasmon for $d \approx 0$ to the interface plasmons that were discussed in the above section, and if d is sufficiently large, to waveguide modes for both TM and TE polarizations.

Possibly the most interesting case is the bulk plasmon that occurs for very thin films. Expanding Eq. (5.28) for $d/\lambda \ll 1$ for TM polarization, we obtain

$$r_{soa}^{TM} \approx r_{sa}^{TM}\left(1 + \frac{4\pi i d n_a \cos\theta}{\lambda}\frac{\varepsilon_s - \varepsilon_o}{\varepsilon_s - \varepsilon_a}\frac{1 - \left(\frac{1}{\varepsilon_o} + \frac{1}{\varepsilon_s}\right)\varepsilon_a \sin^2\theta}{1 - \left(\frac{1}{\varepsilon_a} + \frac{1}{\varepsilon_s}\right)\varepsilon_a \sin^2\theta}\right). \tag{5.29}$$

The interesting part of Eq. (5.29) is the term $(1/\varepsilon_o)$ in the numerator. As noted by Abeles [18], this indicates a singularity (plasmon) at $\varepsilon_o(\omega) = 0$.

However, the above conclusion is not quite accurate owing to the fact that the first-order term becomes infinite for $\varepsilon_o = 0$, and hence the first-order approximation fails drastically. We can obtain a better representation by expanding Eq. (5.28) for small ε_o and d. This yields

$$0 = \varepsilon_o d(\varepsilon_s n_{a\perp} - \varepsilon_a n_{s\perp}) + 4\pi i \varepsilon_s \varepsilon_a^2 \sin^2 \theta. \qquad (5.30)$$

We see from Eq. (5.30) that the singularity is not a pure resonance, as suggested by Eq. (5.29), but has an imaginary part that cannot be set to zero.

D. Example

We illustrate some of these results with the data presented in Fig. 5.4. These data are the relative reflectances R_{TM}/R_{TE} of indium tin oxide (ITO) films of increasing thickness on glass [34]. These films were prepared with a bulk plasma frequency of 2.19 eV by controlling the carrier concentration during deposition and annealing. The purpose of these data is to highlight the points made above. Being driven by an external field, the features in these spectra are technically plasmon polaritons, but the underlying plasmon resonances are evident. For the thinnest films only the bulk plasmon polariton is seen, with a frequency that is essentially independent of angle of incidence, as indicated by Eq. (5.29). For increasing d this feature gradually damps out, as suggested by Eq. (5.30). It is supplanted by an interface plasmon polariton, which exhibits substantial dispersion. A more complete discussion of these results, including a more thorough discussion of plasmons and a comparison of these data to calculations done with the material parameters and the Fresnel equations, is given in Ref. [34].

For intermediate thicknesses the upper and lower interfaces are coupled, and the results are more complicated. However, for thick films the upper and lower interfaces are effectively decoupled, and the dispersion curve of Fig. 5.3 is applicable. Evident in Fig. 5.4 is the movement of the plasmon-polariton structure to higher energies with increasing angle of incidence, as also illustrated in Fig. 5.3.

The fact that the intermediate-thickness situation is complicated suggests that to genuinely compare theory and experiment we must calculate the reflectance ratio using the exact equations. The result is shown in Fig. 5.5. This agreement is achieved with *no* free parameters. Every parameter used in the calculation was determined by independent measurements. The agreement is striking. All features are present, and Figs. 5.4 and 5.5 show the same evolutions of the different features with increasing thicknesses and angles of incidence. This is further confirmation that plasmonic effects are already built into the Fresnel equations and do not need to be added in an *ad hoc* fashion.
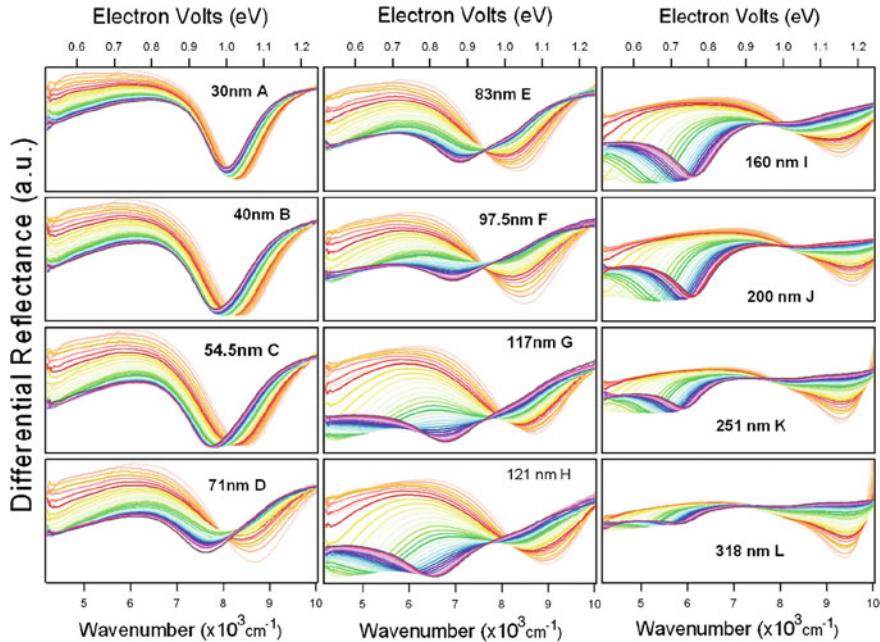
Fig. 5.4 Ratio R_{TM}/R_{TE} of indium tin oxide films of indicated thicknesses deposited on BK-7 glass, measured in the Kretschmann configuration of Fig. 5.2. Angles of incidence range from 42° (*orange*) to 53° (*purple*) in increments of 0.35°. The bulk plasmon polariton structure seen in panel A evolves into an interface plasmon polariton structure at the air-ITO interface at panel I, after which attenuation in the film significantly reduces optical contact to the air-film interface (after Ref. [34])

More generally, the dynamic response of a configuration can be described beyond the range of validity of EMTs with the approach introduced by Ewald and Oseen, which leads to the Ewald-Oseen extinction theorem [35, 36]. Here, an optical response is treated as a four-step process [27]. First, the local field is evaluated at a charge site. Second, the relevant force equation is solved for the motion of q. Third, the radiation from q is calculated. Fourth, the radiation from all q is superposed. For linear optics this is a difficult challenge, because the wavelength of the emerging field is the same as that of the incident field, and the intensities are comparable. Thus the entire calculation must be done self-consistently. Nevertheless, the 4-step process highlights the essential physics. For nonlinear optics this approach works excellently well, since the emerging radiation is much weaker and occurs at a different λ. This is one of the few examples where a nonlinear problem is actually simpler than its linear equivalent [27].
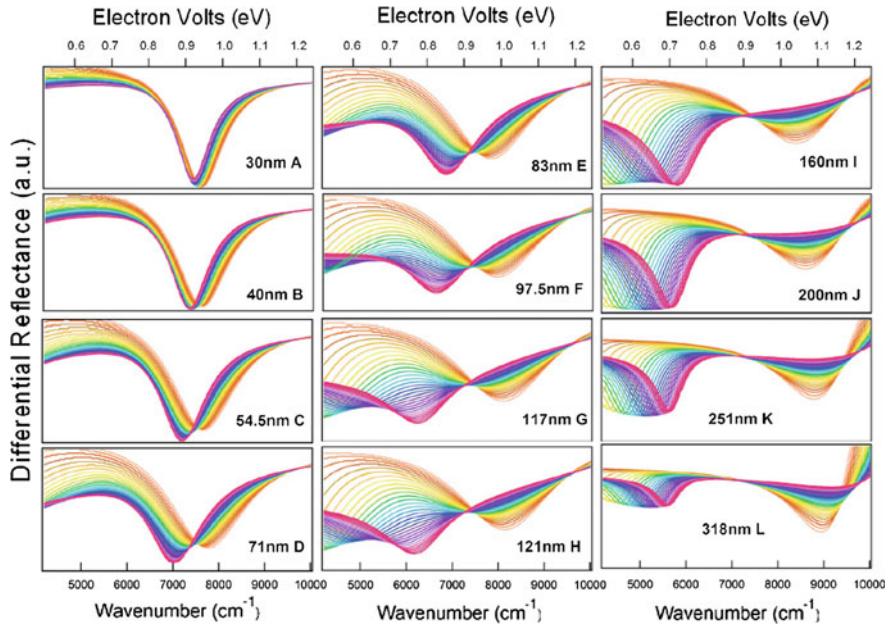
Fig. 5.5 Reproduction of the data of Fig. 5.4 calculated with no adjustable parameters (after Ref. [34])

5.5 Lateral Inhomogeneities and Scattering

Scattering of incoming radiation is described by solutions of the inhomogeneous Maxwell Equations, a topic that is far too large to be discussed here. For example most of Raether's monograph [7] is devoted to the optical properties of systems with rough surfaces, and CD analysis is now an intrinsic part of integrated-circuits technology [6]. However, we gain perspective by briefly summarizing some aspects.

Rough interfaces break the coupling between $n_{||}$ and E with interesting consequences. The Fourier transform of the autocorrelation function of the rough surface supplies momentum so the wave vector of the plasmon can couple to the radiation, leading to a situation where the combination $(k_{||} - k_{AC})^2$ can be less than $(\varepsilon_a + \varepsilon_\infty)$, although it may still be greater than ε_∞. In this case the $n_{||}$ become a continuum [17]. Although the plasmon-radiation coupling is a second-order effect, it allows the incident radiation to generate plasmons and therefore for energy to be extracted. These plasmon-absorption features appear as locally decreased intensities of the reflected beam. The Kretschmann and Kröger work [17] also places more traditional scattering formulations [37] in context, showing that these result from the scalar limit where no plasmon generation is possible.

As also mentioned in the Introduction, the diffraction grating represents an extreme case of a rough surface, where the inhomogeneities are not only of the scale

of λ but also periodic. Hessell and Oliner [8] describe the grating as a waveguide interacting with a resonant cavity, where the incident and reflected waves correspond to the propagating modes of the waveguide, and the role of the resonant cavity is filled by the plasmon. The theory is cast as a Fourier series, with the intensities of the scattered orders expressed as a continued fraction and evaluated analytically for the case of a single sinusoidal component. Although the mathematics is challenging, it provides a concise explanation of Wood's anomalies [38] and identifies the origin of previously unidentified structures in the reflected beams.

5.6 Summary

Plasmonic phenomena and their description have been well known for many years, although the marked increase in general interest has occurred only recently. Solutions of the homogeneous Maxwell equations are not generally considered in standard formulations of electrodynamics, but can be implemented in a straightforward manner. The present work represents an attempt to make plasmonics more accessible to readers who are familiar with classical electrodynamics.

References

1. M. Born, E. Wolf, *Principles of Optics*, 5th edn. (Pergamon, Oxford, 1975), pp. 99 ff, 624 ff
2. N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Saunders College Publishing, New York, 1976)
3. D.E. Aspnes, Thin Solid Films **519**, 2571–2574 (2011)
4. J.C.M. Garnett, Philos. Trans. R. Soc. Lond. Ser. A **203**, 385 (1904)
5. J.C.M. Garnett, Philos. Trans. R. Soc. Lond. Ser. A **205** 237 (1906)
6. J. Opsal, H. Chu, Y.C. Chang, G. Li, Proc. SPIE **4689** 163–172 (2002)
7. H. Raether, Surface plasmons on smooth and rough surfaces and on gratings, in *Springer Tracts in Modern Physics*, vol. 111, ed. by G. Höhler (Springer, Berlin, 1988)
8. A. Hessel, A.A. Oliner, Appl. Opt. **4**, 1275–97 (1965)
9. M.G. Moharam, E.B. Grann, D.W. Pommet, T.K. Gaylord, J. Opt. Soc. Am. **A12**, 1068–1076 (1995)
10. M. Fleischman, P.J. Hendra, A.J. McQuillan, Chem. Phys. Lett. **26**, 163 (1974)
11. H.A. Atwater, A. Polman, Nat. Mater. **9**, 205–213 (2010)
12. M. Bergmaier, M. Huber, K. Hingerl, Appl. Phys. Lett. **89**, 081907 (2006)
13. R.W. Wood, Philos. Mag. **4**, 19–24 (1902)
14. R.W. Wood, Phys. Rev. **48**, 928 (1935)
15. A. Sommerfeld, Ann. Physik **28**, 665 (1909)
16. E. Kretschmann, Z. Phys. **241**, 313 (1971)
17. E. Kretschmann, E. Kröger, J. Opt. Soc. Am. **65**, 150 (1975)
18. F. Abeles, Thin Solid Films **34**, 291 (1976)
19. A. Otto, Zeitschrift für Physik **216**, 398–410 (1968)

20. M. Kerker, D.S. Wang, H. Chew, Appl. Opt. **19**, 3373 (1980)
21. J. Gersten, A. Nitzan, J. Chem. Phys. **73**, 3023 (1980)
22. M. Moskovits, Annu. Rev. Phys. Chem. **42**, 465–499 (1999)
23. J.C. Hulteen, D.A. Treichel, M.T. Smith, M.L. Duval, T.R. Jensen, R.P. Van Duyne, J. Phys. Chem. **B103**, 3854–3863 (1999)
24. W.A. Murray, W.L. Barnes, Adv. Mater. **19**, 3771–3782 (2007)
25. J. Zhang, L. Zhang, Adv. Opt. Photonics **4**, 157 (2012)
26. J.D. Jackson, *Classical Electrodynamics*, 3rd edn. (Wiley, New York, 1998)
27. E.J. Adles, D.E. Aspnes, Phys. Rev. **B77**, 165102 (2008) where a complete radiative version of the theory is developed for nonlinear-optical applications
28. O. Wiener, Abh. Math.-Phys. K1 Konigl. Sachs. Ges. **32**, 509 (1912)
29. D.A.G. Bruggeman, Ann. Phys. (Leipzig) **24**, 636 (1935)
30. U. Kreibig, M. Vollmer, *Optical Properties of Metal Clusters* (Springer, Berlin, 1995)
31. A. Derkachova, K. Kolwas, Eur. Phys. J. Special Top **144**, 93–99 (2007)
32. H. Raether, Surface plasmons on smooth and rough surfaces and on gratings, in *Springer Tracts in Modern Physics*, vol. 111, ed. by G. Höhler (Springer, Berlin, 1988), p. 4 ff
33. M. Born, E. Wolf, *Principles of Optics*, 5th edn. (Pergamon, Oxford, 1975), p. 62
34. C. Rhodes, M. Cerruti, A. Efremenko, M. Losego, D.E. Aspnes, J.-P. Maria, S. Franzen, J. Appl. Phys. **103**, 093108 (2008)
35. P.P. Ewald, Ph.D. Thesis, TU München, Germany, 1912
36. C.W. Oseen, Ann. Phys. **48**, 1 (1915)
37. H.E. Bennett, J. Opt. Soc. Am. **51**, 123 (1961)
38. R.W. Wood, Phil. Mag. **4**, 396 (1902)

Chapter 6
Thin Films of Nanostructured Noble Metals

Herbert Wormeester and Thomas W. H. Oates

Abstract The optical characterisation of nanostructured noble metal films is discussed. A good description of the optical response of a 2D layer of identical metal particles on a support is obtained with the so-called Thin Island Film theory, which is based on the polarisability of an individual nanoparticle. The influence of particle shape, density, and the optical properties of both the support and the surrounding ambient is discussed as well as the ability of the commonly used Maxwell Garnett approximation to model these films. The possibilities and limits of this approach is illustrated with several examples. A short outlook towards the ellipsometric analysis of metamaterials is given.

6.1 Introduction

The plasmonic effects observed in noble metal particles makes them very interesting for a variety of applications, including bio-medical and information technology. The plasmon resonance is related to the confinement of the free electrons in the particles. Although silver and especially gold are widely used, plasmon resonances are also observed for semiconducting particles and thin transparent conducting oxide layers. The dominance of gold and silver in plasmon research is due to the rather simple ways these can be produced. Probably the oldest manufacturing method is in the production of stained glass, a very nice example of which is the so-called Lycurgus

H. Wormeester (✉)
Physics of Interfaces and Nanomaterials MESA+ Institute for Nanotechnology,
University of Twente, Enschede, The Netherlands
e-mail: h.wormeester@utwente.nl

T. W. H. Oates
Forschungsbereich Material- und Grenzflächenanalytik, Leibniz-Institut für Analytische
Wissenschaften - ISAS-e.V, Albert-Einstein-Str. 9, 12489 Berlin, Germany
e-mail: tom.oates@isas.de

cup, demonstrating the high level of craftsmanship in Roman times. Yellow and red colours were produced with the aid of silver and gold nanoparticles, respectively.

In the early 19th century, Faraday produced gold and silver particles via colloidal chemistry, the colours of which depend on the size and the dielectric properties of the nano particles and the solution. A first explanation of the origin of the observed colours was given by Maxwell Garnett [1, 2] in his seminal papers entitled "Colours in Metal Glasses and in Metallic Films". He showed that the coloration is the result of a mixing of the optical properties of the metal particles and the glass. As such, these articles are a landmark in the development of Effective Medium Approximation (EMA) theories that describe the average optical response of a material that is constituted of more than one material.

For many contemporary technological applications, particularly in sensing and photovoltaics, a thin layer of noble metal particles on a suitable substrate is required. Characterisation of this layer consists of describing not only the particle density, but also the optical properties of such a layer, or the change in the plasmon resonance position if molecules are adsorbed on the metal particles. Another area of interest is the growth of a metal film, initially consisting of isolated entities, towards a percolated metal film, with a corresponding change in the electrical resistance. Around the percolation threshold, such a film displays very characteristic and interesting optical behaviour. The real part of the dielectric function at low photon energies has to transit from a positive value characteristic of non-conductive isolated particles to a large negative value characteristic of a conductive continuous metal film. These optical features can be monitored with ellipsometry as it provides a reproducible and fast method for thin film characterisation with great potential for in situ monitoring.

The analysis of ellipsometric spectra requires a model of the measured sample that incorporates the essential features that are optically monitored. Modeling the optical response of thin films which are much thinner that the wavelength of the light generally proceeds by one of two methods [3]. For extremely thin films the dielectric properties of the layer are used to define a dimensionless surface susceptibility which then acts as a correction to the Fresnel reflection coefficients from a single interface (modified two-layer model). For thicker films the thickness and (effective) dielectric properties of the film are used to construct a three-layer model (substrate/film/ambient) and the Airy formulas or Abelès matrix formalism is used to compute the optical response. The effective dielectric properties may be described by an EMA if the material consists of two or more different entities with a characteristic length scale much smaller than the wavelength of the light used to probe the layer. Several EMA's have been developed, among them the Lorentz–Lorenz model, the Maxwell Garnett model and several models developed by Bruggeman [4–6]. Bruggeman's approach for the mix of spherical entities is usually regarded as the most robust. These models were developed for materials that show a similar behaviour in all three orthogonal directions; i.e. isotropic materials.

However, as already noted by Maxwell Garnett, the isotropic situation no longer holds for particles deposited as a thin film. Such a film will show a uniaxial dielectric
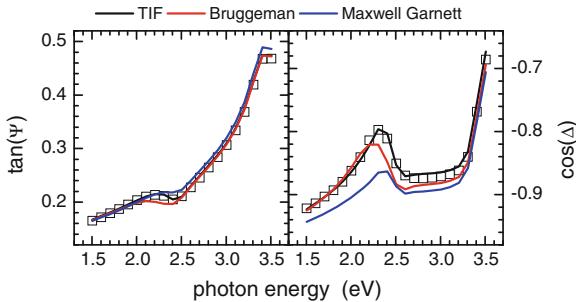
Fig. 6.1 Ellipsometric measurement of a Au nanocolloidal film deposited on Si. From SEM measurements a surface coverage of 12 % (8 % film volume) was determined for this film. The lines show simulations of the ellipsometric response with different EMA's, i.e. the Thin Island Film, the Bruggeman and the Maxwell Garnett approach

response, with the direction perpendicular to the surface different from the in-plane direction. This effect is illustrated by Fig. 6.1, which compares the measured spectrum of a film of 13 nm gold nanocolloids with known polarizability and surface coverage deposited on a silicon substrate with models using the Bruggeman (BEMA) and Maxwell Garnett (MG) EMA. The presence of the gold particles on the surface is directly observed by the characteristic plasmon resonance at 2.3 eV. All model parameters required to calculate the optical response with BEMA and MG were measured prior with different techniques [7]. Since there are no fitted parameters Fig. 6.1 depicts the accuracy of the models used. Both BEMA and MG fail to accurately describe the measured spectrum at the still low surface coverage of 12 %. The BEMA approach is quite reasonable at low and high energies, but it predicts a red shifted plasmon resonance energy which is the important characteristic of this sample. The MG approach gives the correct plasmon resonance energy which explains its popularity in the characterisation of noble metal particle layers. However, the intensity of the plasmon resonance is much smaller than experimentally measured.

Also shown in Fig. 6.1 is the prediction of the Thin Island Film (TIF) model using the pre-measured parameters, which gives a much better representation of measured spectra [7]. The TIF model provides an improved result due to incorporation of the effects of substrate and dipole interactions and also the uniaxial anisotropy of the layer. We shall show in this chapter that for the MG and TIF models the in-plane plasmon resonance can be characterized by a Lorentzian line shape, whose resonance energy provides information about surface coverage and particle shape. Without a detailed modelling of ellipsometric spectra, this fact already allows an interpretation of (in situ) recorded spectra. We will also discuss the limits of the applicability of MG approach and anisotropic particles and metamaterials.

6.2 The Optical Response of a Thin Layer of Metal Particles

6.2.1 The Polarisability of a Drude Metal Nanoparticle

The origin of the various colours of noble metal nanoparticles was a challenging physics problem in the late 19th century. In 1908 Gustav Mie solved Maxwell's equations for spherical geometry that explained the observed red colour of nanometer sized gold particles [8]. This landmark theory provided a complete solution for the scattering and absorption of electromagnetic radiation by a sphere. The full Mie solution for a spherical geometry is not required when the particle size is much smaller than the wavelength of the light used. In this case the so-called *quasistatic approximation* is applied, where the effects of the phase variation of the electric field at different points on the particle (retardation) are ignored [9]. The electrons in the particle thus oscillate in a dipolar mode. The polarizability of a spherical particle in the dipolar limit is given by:

$$\alpha = 4\pi a^3 \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m} \qquad (6.1)$$

where a is the radius of the particle $\varepsilon(\omega) = \varepsilon' + i\varepsilon''$ is the dielectric function of the sphere material and ε_m is the real dielectric constant of the medium surrounding the sphere.

The metallic behaviour of a material is the result of a finite density of states of electrons at the Fermi energy. The optical response as a result of this part of the electronic structure is described by the Drude theory [9, 10]. In this, three parameters are used to characterize the dielectric behaviour: the density of free carriers N_e, the effective electron mass m^* and the characteristic damping time τ. This provides a dielectric function according to:

$$\varepsilon_D = \varepsilon_\infty - \frac{N_e \frac{\hbar^2 e^2}{m^* \varepsilon_0}}{E^2 + i\frac{E}{\tau}} = \varepsilon_\infty - \frac{E_p^2}{E^2 + i\Gamma E}. \qquad (6.2)$$

The combination of the electron density and effective mass are often expressed as the plasma energy E_p. Their individual numbers can not be determined from the optical response of a metal but require a more elaborate characterisation, for instance with the (optical) Hall effect. The damping time is often expressed by the broadening Γ. The Drude formula is valid in the spectral region in which no interband transitions occur. However, the influence of these interband transitions is present since the real and imaginary part of dielectric function have to be Kramers-Kronig consistent. This consistency is obtained with the quantity ε_∞. For a pure Drude metal the value of this quantity is 1. Deviations from this number reflect the influence of the interband transitions [10].

The combination of the dipole approximation for a spherical particle and the Drude response provides a Lorentzian line shape for the polarisability of a single

Table 6.1 Lorentzian parameters for the dipolar spherical particle, Maxwell Garnett and Thin Island Film model. The derivation of these parameters is provided in Appendix A.1, A.2 and A.3

	Spherical particles	MG model	TIF model
$E_{p,L}^2$	$\dfrac{3\varepsilon_a}{(\varepsilon_\infty + 2\varepsilon_a)^2}E_p^2$	$3f\dfrac{\varepsilon_a(1+k_s)}{\left[\varepsilon_\infty(1-f)+\varepsilon_a(k_{s+f})\right]^2}E_p^2$	$f\dfrac{\varepsilon_a(1+k_P)^2}{(\varepsilon_\infty+k_P\varepsilon_a)^2}E_p^2$
$E_{0,L}^2$	$\dfrac{E_p^2}{\varepsilon_\infty+2\varepsilon_a}$	$\dfrac{(1-f)}{\varepsilon_\infty(1-f)+\varepsilon_{a(k_{s+f})}}E_p^2$	$\dfrac{E_p^2}{\varepsilon_\infty+k_P\varepsilon_a}$
$\varepsilon_{\infty,L}$	$\dfrac{\varepsilon_\infty-\varepsilon_a}{\varepsilon_\infty+2\varepsilon_a}$	$1+\dfrac{3f(\varepsilon_\infty-\varepsilon_a)}{\varepsilon_\infty(1-f)+\varepsilon_a(k_{s+f})}$	$1+f(1+k_P)\dfrac{\varepsilon_\infty-\varepsilon_a}{\varepsilon_\infty+k_P\varepsilon_a}$

metal particle:

$$\alpha = 4\pi a^3\left(\varepsilon_{\infty,L}+\frac{E_{p,L}^2}{E_{0,L}^2-E^2-i\Gamma E}\right) \tag{6.3}$$

The limited size of a spherical metal particle changes the optical response. A plasmon-like oscillation is observed and the polarisability has a maximum at the resonance energy:

$$E_{0,L}=\frac{E_p}{\sqrt{\varepsilon_\infty+2\varepsilon_a}} \tag{6.4}$$

The equations for the strength and the offset are given in Table 6.1 and the derivation is provided in Appendix A.1. The expression for the resonance energy includes a number of factors that influence the actual position of the resonance energy [9, 10]:

1. The value of ε_∞: Interband transitions define this parameter. Its influence on the resonance position is illustrated with the comparison of gold and silver. Both materials have a similar electron density and electron mean free path resulting in a similar value for both the plasma energy and the broadening. If the dielectric response is described by Eq. 6.2, the difference between these two metals is the value of ε_∞; 10 and 4 respectively. As a result, the resonance energies for gold and silver nanoparticles are very different, explaining their red and yellow appearances, respectively.
2. The value of ε_a: A change in the nature of the ambient can induce a distinctive shift in the resonance energy, with a larger dielectric constant resulting in a stronger redshift.
3. The value of E_p: The plasma energy is determined by the density of the free electrons. Its magnitude can be influenced by factors such as the chemical adsorption of molecules on a small particle.
4. The value 2: This value is valid for spherical particles. For non-spherical particles, this value increases in the extended direction, while it decreases for the compressed direction. This results in two plasmon resonances for prolate and oblate particles, one at energy below and one above that of a spherical particle. This value is also changed if the particles start to interact with their surrounding,

i.e. other particles or a surface. The latter effect can be very important for the ellipsometric characterisation of a particulate film.

In this analysis, the assumption is still made that the dielectric function of the particle is the same as that for a bulk crystal. However, a few more effects have an influence on the actual dielectric function [9]. Most important is the limited size of a particle compared to the electron mean free path of the bulk material. If this is the case, the value of the broadening, Γ, can be considerably enlarged. Also polydispersity of a particle ensemble can manifest itself as a broadening of the optical response.

6.2.2 Maxwell Garnett EMA for a Drude Metal

The effective dielectric function formulated by Maxwell Garnett for a layer consisting of particulates with dielectric function ε and occupying a volume fraction f in a host medium with dielectric function ε_a is:

$$\frac{\varepsilon_{eff} - \varepsilon_a}{\varepsilon_{eff} + 2\,\varepsilon_a} = f\,\frac{\varepsilon - \varepsilon_a}{\varepsilon + k_s \varepsilon_a}. \tag{6.5}$$

The screening factor k_s takes into account the shape of the inclusions. Rearrangement of the parameters yields an expression for the effective dielectric function:

$$\frac{\varepsilon_{eff}}{\varepsilon_a} = \frac{\varepsilon(1 + 2f) + \varepsilon_a(k_s - 2f)}{\varepsilon(1 - f) + \varepsilon_a(k_s + f)}. \tag{6.6}$$

The effect of metal particles with a Drude metal characteristic leads to an effective dielectric function with a Lorentzian line shape [11]. The Lorentzian parameters, i.e. the plasma energy, the resonance energy and the interband transitions for this effective dielectric function are given in Table 6.1. The derivation of these parameters is provided in Appendix A.2. These Lorentzian parameters are a generalization of the results of Yamaguchi [12], who considered spherical particles ($k_s = 2$) and an ideal metal ($\varepsilon_\infty = 1$) in vacuum ($\varepsilon_a = 1$). In this special case, the Lorentz plasma energy is $E_{p,MG}^2 = fE_p^2$ and the resonance energy $E_{0,MG} = \sqrt{(1 - f)/3})E_p$. However, as discussed above, not even silver shows such ideal Drude behaviour. The influence of these interband transitions via the ε_∞ on the Lorentz line profile parameters is considerable, see Fig. 6.2. Both the location and the strength of the plasmon resonance are strongly affected by the interband transitions, although the plasmon resonance and the interband transitions do not overlap. The dependence on the volume fraction of the particles is also quite different. The plasmon resonance energy is hardly affected by the density of particles **f** for $\varepsilon_\infty = 10$, i.e. for Au like particles. This was also experimentally observed in the range up to 40 % surface coverage [7].

Maxwell Garnett already discussed that for deposited particles two resonances should be observed. A nice illustration of this can be found in the work of Hilger
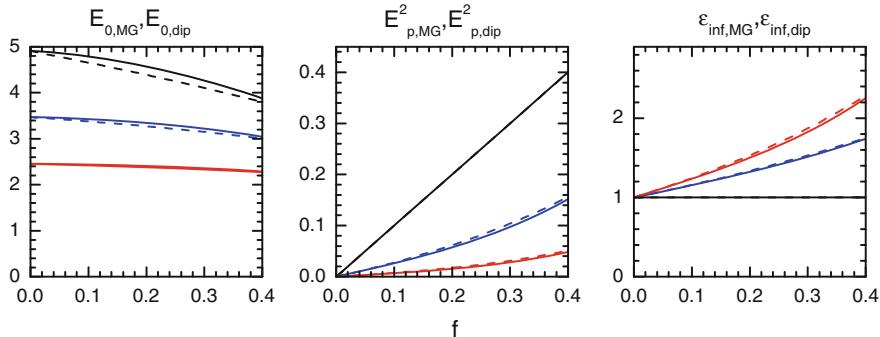
Fig. 6.2 Lorentz line profile parameters as a function of surface coverage for the TIF theory (*solid lines*) and the MG approach (*dashed lines*). The parameters are evaluated for $\varepsilon_\infty = 1$ (*black*), 4 (*blue*) and 10 (*red*). These represent the values found for a real Drude metal, silver and gold respectively

and Kreibig who studied the optical response of deposited silver particles [13]. They observed two separate resonances, one that can be excited with light polarized parallel to the interface and one that can be excited with light polarized perpendicular to the interface. The split in resonance energy increases with coverage and is virtually zero for very low coverage. These in-plane and out-of-plane resonances reflect the interaction between particles; the in-plane resonance is red shifted, and the out-of-plane resonance is blue shifted. This effect was already discussed by Maxwell Garnett himself, who limited the discussion to the in-plane component [1, 2]. His approach shows the red shift in resonance energy, as also illustrated in Fig. 6.2. However, the out of plane resonance for deposited particles is not accounted for by the standard formulation of the MG approach.

The split in resonance energy can also be found using the MG EMA if one takes into account that a monolayer of nanoparticles form a very thin layer. In this case the Drude thin film approach should be valid, which is a special case of the Airy equations [3]. For a film with a thickness d much smaller than the wavelength of light, the Drude thin film expression (first formulated by Rijn van Alkemade [14]) makes a first order approximation and includes only the dipolar terms. The change in the ratio of the p- and s-polarized light measured in ellipsometry is found to be proportional to a value J, called an invariant [15]. The proportionality constant is determined by the incidence angle, the wavelength and the substrate and ambient (ε_a) dielectric function. The value of J is given by the variation in the dielectric properties of a film between the substrate at height zero and the film thickness d.

$$J = \int_0^d \varepsilon_P(z) - \varepsilon_a + \varepsilon_a \left(\frac{1}{\varepsilon_\perp(z)} - \frac{1}{\varepsilon_a} \right) dz. \qquad (6.7)$$

A difference between the dielectric function of the thin film parallel ε_p and perpendicular $\varepsilon_\perp$ to the interface is used. Although these are the same within the MG approach, they effect the optical response in a different manner as in general $\varepsilon \neq 1/\varepsilon$. For a Lorentzian profile of the dielectric function, the reciprocal value is also a Lorentzian:

$$\frac{\varepsilon_a}{\varepsilon_{eff}} = \cfrac{1}{\varepsilon_\infty + \cfrac{E_p^2}{E_0^2 - E^2 - i\Gamma E}}$$

$$= \frac{1}{\varepsilon_\infty} + \frac{E_p^{*2}}{\left(E_0^{*2} - E^2 - i\Gamma E\right)} \tag{6.8}$$

with

$$E_0^{*2} = E_0^2 + \frac{E_p^2}{\varepsilon_\infty} \quad \text{and} \quad (E_p^*)^2 = -\left(\frac{E_p}{\varepsilon_\infty}\right)^2. \tag{6.9}$$

The Lorentz line shape has an attenuated strength and a blue shifted resonance energy. The MG approach thus indeed gives rise to a blue shift of the resonance energy of the perpendicular component of the dielectric function. The difference in resonance energy between the in plane and out of plane component depends again on the interband transition parameter ε_∞. This explains the stronger features and larger resonance splitting observed for Ag compared to Au.

Figure 6.3 shows the measured and evaluated ellipsometric spectra for silver nanocolloids with a surface coverage of 44 %. A difference in resonance energy between the parallel and perpendicular component of about 0.8 eV is measured. The measured spectra could be adequately modelled with the MG theory using a modified Ag dielectric function. The broadening in the Drude part of the original silver dielectric function was changed to incorporate the finite size of the 5 nm diameter silver particles. The MG combined with the Drude thin film approximation gives a 1.0 eV peak splitting, slightly larger than the 0.8 eV from experimental results. The effect of this differences on the ellipsometric angles is relatively small. The applicability of
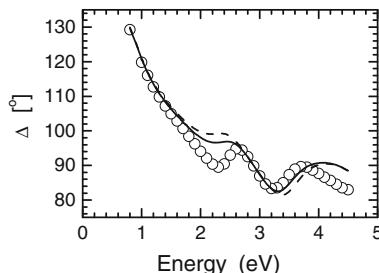
Fig. 6.3 The measured value of the ellipsometric parameter Δ for a nanocolloidal Ag film on HOPG (o). The *solid line* is from the TIF model, the *dashed line* represents the result of the MG model

the MG approach is in this case limited by the slight difference in resonance energy. However, it does not explain the strong difference in simulated and measured spectra observed for gold, see Fig. 6.1. This can only be explained by a more appropriate treatment of the electro-magnetic interactions of particles with each other and the substrate on which they are deposited.

6.2.3 The Thin Island Film Theory

A particle deposited on a surface experiences two interactions that are very different from that in a diluted solution or bulk crystal. First of all, the presence of the substrate interface breaks the symmetry. As a result the optical response will change, which may be mathematically described as an interaction with a mirror image. A second effect is that the surface coverage can easily exceed a few percent and the interactions between particles have to be taken into account. These neighbouring particles also have images with which the original particle will interact. These interactions for the in-plane component are schematically depicted in Fig. 6.4. The incorporation of the effect of the image dipole of a particle on the depolarization factor was explicitly incorporated into a simple dipolar model by Yamaguchi [16] and at virtually the same time by Bedeaux and Vlieger [17]. The image effect leads to an effective polarizability of each individual particle that is different for directions parallel and perpendicular to the interface. The incorporation of the image effect was the start of the development of the Thin Island Film (TIF) theory by Bedeaux and co-workers [18] which also incorporates the neighbor interaction effect and the description of non-spherical particles. This TIF theory is essentially a multipole expansion of the electro-magnetic problem of the response of identical particles by an incident electric field. As shown by Haarmans et al. [19] a noble metal particle film with a coverage
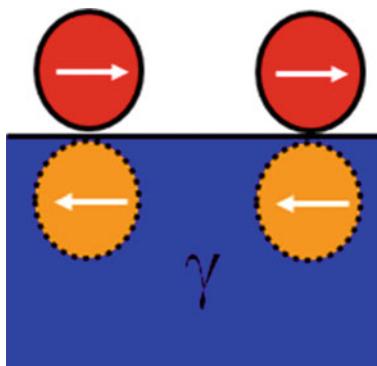
Fig. 6.4 Optical interactions between nanoparticles on a surface giving different contributions to the in-plane excess polarizability γ. These interactions are with the image of the particle, with neighbouring particles and with the images of neighbouring particles

up to 50 % is well described by dipole and quadrupole terms. Wormeester et al. [20] showed that if image effects are negligible, dipolar terms are sufficient to describe the optical response in this coverage range. A computer program that calculates the optical response with the TIF theory up to arbitrary order, called Granfilm, has been published by Lazarri and Simonsen [21].

The TIF theory treats the influence of the shape and the interaction between identical particles as a modification of the polarizability α of the individual particles. The presence of the particles in a 2D layer results in an effective dimensionless polarizability α' that is different parallel and perpendicular to the substrate interface. Similar to the MG approach a dielectric function parallel and perpendicular to the interface can be defined as [11]:

$$\frac{\varepsilon_{\text{eff}}^{\text{P}}}{\varepsilon_a} = 1 + f\alpha'_P \text{ and } \frac{\varepsilon_{\text{eff}}^{\perp}}{\varepsilon_a} = \frac{1}{1 - f\alpha'_{\perp}}. \tag{6.10}$$

For an ellipsoidal particle, the polarizability is modified via the depolarization factor L. For a spherical particle without any interaction this is 1/3. This depolarization is directly related to the screening factor $k_s = (1 - L)/L$. With only dipolar interaction the effective polarizability and the parallel component for the dielectric function are:

$$\alpha'_{P,\perp} = \frac{1}{L_{P,\perp}} \frac{\varepsilon - \varepsilon_a}{\varepsilon + k_{P,\perp}\varepsilon_a} = (1 + k_{P,\perp}) \frac{\varepsilon - \varepsilon_a}{\varepsilon + k_{P,\perp}\varepsilon_a}$$

$$\frac{\varepsilon_{\text{eff}}^{\text{P}}}{\varepsilon_a} = 1 + f(1 + k_P) \frac{\varepsilon - \varepsilon_a}{\varepsilon + k_P\varepsilon_a} \tag{6.11}$$

For Drude metal particles, the effective dielectric function is again a Lorentz profile with the parameters listed in Table 6.1. The derivation is provided in Appendix A.3. The effective dielectric functions are not the common quantities used within the TIF formulation. Instead surface excess polarizabilities are used, reducing the problem to a two-layer model described by modified Fresnel expressions. These quantities are related to the dielectric function via the invariant J from the Drude thin film approach [18] (Eq. 6.7):

$$J = \gamma + \varepsilon_a\varepsilon_s\beta \tag{6.12}$$

with γ and β the surface excess polarizability parallel and perpendicular to the interface, respectively. For particles represented by dipoles on the surface with surface coverage ϕ this excess quantity is written as:

$$\gamma = \phi \, 4a\varepsilon_a \frac{\varepsilon - \varepsilon_a}{\varepsilon + k_{s,p}\varepsilon_a} = \phi \, \frac{4a\varepsilon_a}{3} \frac{\delta\varepsilon}{1 + L_p\delta\varepsilon} \tag{6.13}$$

$$\beta = \frac{\phi \, 4a}{\varepsilon_a} \frac{\varepsilon - \varepsilon_a}{\varepsilon + k_{s,z}\varepsilon_a} = \phi \, \frac{4a}{3\varepsilon_a} \frac{\delta\varepsilon}{1 + L_z\delta\varepsilon}$$
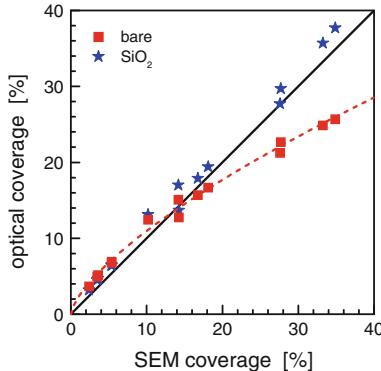
Fig. 6.5 The surface coverage determined with SEM and ellipsometry of a series of samples containing 13 nm diameter Au nanocolloids deposited on a silicon substrate with a thin native oxide. The modelling of the ellipsometric spectra was done with the TIF theory employing a fixed depolarization factor of 1/3 (*red square symbols*) and incorporating image and neighbour interaction effects (*blue stars*). The *dashed line* is a guide to the eye, while the *black solid line* signifies a one-to-one relation between the two methods

In this the term $\delta\varepsilon = (\varepsilon - \varepsilon_a)/\varepsilon_a$ is referred to as the dielectric contrast. These excess polarizabilities show again a Lorentzian lineshape.

The simplest approach to using TIF is to neglect at first any image and neighbour effects, i.e. $k_s = 2$. This route was employed initially by Kooij et al. to analyze their ellipsometric measurements of deposited gold colloid layer as depicted in Fig. 6.1 [7]. They found that the measured spectra could be well represented with only the surface coverage as fit parameter. However, a comparison of the fitted coverage and the coverage determined with SEM showed two differences, see Fig. 6.5. Below 20 % surface coverage, the ellipsometric data overestimated the actual coverage, while for a surface coverage above 20 % an underestimation was found. These were attributed to the effect of image dipoles and the mutual interaction between particles, see Fig. 6.4. If these effects are taken into account a one-to-one relation between the surface coverage determined with ellipsometry and with SEM was found. Note that this approach works well for gold particles as the resonance position is only slightly effected by a change in k_s. This is the result of the dominance of the high value of ε_∞ compared to $k_s\varepsilon_a$. For silver particles the effect of a change in k_s is much larger.

The optical response of an ensemble of interacting dipoles was analyzed by Barrera et al. [22]. They considered particles with radius a placed on a regular lattice with r_0 the nearest neighbor distance. The interaction between the dipoles changed the depolarization factor for the directions parallel and perpendicular to the interface according to:

$$L_p = \tfrac{1}{3}\left(1 - \tfrac{1}{2}\left(\tfrac{a}{r_0}\right)^3 S\right)$$
$$L_z = \tfrac{1}{3}\left(1 + \left(\tfrac{a}{r_0}\right)^3 S\right)$$

$$(6.14)$$
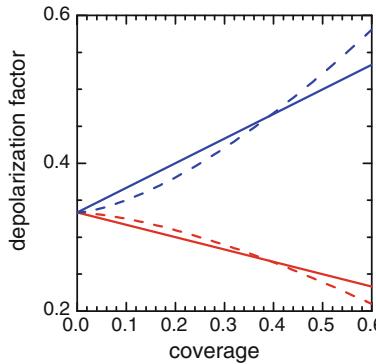
Fig. 6.6 The influence of the surface coverage on the depolarization factor. The lower branch (*red*) and upper branch (*blue*) represent the in plane and out of plane depolarization factor, respectively. The *solid line* represents a uniform distribution while the *dashed line* represents the hexagonal distribution. From Ref. [10], reproduced with permission from Elsevier

With S the weighted sum over all lattice sites. Note that the summation of the depolarization factors $2L_p + L_z = 1$. For a lattice normalized to the nearest neighbor distance, the weighted summation is given by $S = \Sigma 1/r^3$. For a hexagonally close-packed lattice $S = 11.034$ while $S = 9.03$ for a cubic lattice. The factor a/r_0 reflects the surface coverage. The precise relation depends on the actual distribution of the particles on the surface, i.e. hexagonal, square, uniform etc. Figure 6.6 shows the effect of surface coverage for a hexagonal lattice (dashed) and uniform distribution (solid line) on the depolarization factor.

The excess polarizability can be used to evaluate the dielectric response parallel and perpendicular to the interface [11]. Both the parallel and the perpendicular component have their own Lorentzian lineshape. The resonance energy depends via the depolarization factor on the surface coverage. Stronger collective interactions (i.e. a larger surface coverage) results in a stronger deviation of the depolarization factor from 1/3 and as such a larger separation between the two resonances. This interaction thus creates an effective polarizability that mimics a set of non interacting oblate particles. For the already mentioned system of Ag particles on HOPG [20], the TIF model provides a description of the measured data that explains the split in the in-plane and out-of-plane plasmon resonances (see Fig. 6.3). The differences between measurement and simulation were attributed to faults in the description of the dielectric properties of the silver colloids while they are still in solution. Note that no fit parameters were used to generate the simulated spectra.

The coverage dependence of the split in resonance energy explains qualitatively the observations of Hilger and Kreibig [13] for a varying coverage of silver nanoparticles. The measured resonance energies as a function of coverage are shown in Fig. 6.7. The solid lines are a calculation of the variation in resonance energy according to $E_0 = \sqrt{E_p^2 / \varepsilon_\infty + k_p \varepsilon_a}$. The surface coverage effects the depolarization and thus
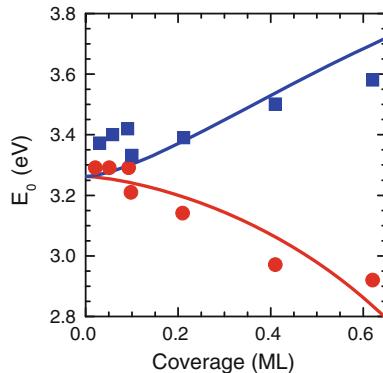
Fig. 6.7 The resonance energy of the in-plane (*red dots*) and out-of-plane (*blue squares*) plasmon resonance observed for silver nanoparticles deposited by Hilger et al. [13]. The *solid lines* indicate the change in resonance energy according to the surface coverage dependence of the depolarization factor. From Ref. [10], reproduced with permission from Elsevier

the screening factor k. The value of E_p was used as a scaling parameter and chosen such that for zero coverage the resonance energy is 3.26 eV. This figure also implies that the reciprocal path can be used: the shift in resonance energy can be used to evaluate the depolarization factor. The value of this factor can be explained in terms of coverage and shape of the particles.

The value of this resonance energy in Fig. 6.7 is smaller than expected for a spherical silver particle. This can be the result of an image effect that would red shift the resonance energy for both the parallel and perpendicular component. The influence of image effects is discussed in detail elsewhere [10]. For spherical particles this effect is relatively small. This is even further reduced in the case of the discussed silver colloids on HOPG. This is because the silver is in a polymer matrix which effectively increases the distance between the center of the metal nanoparticle and the HOPG interface and diminishes the image contribution. The neglect of the image effect also has the consequence that a more precise quadrupole evaluation within TIF theory reduces again to the dipolar form (see also the Appendix). As a result, the simple dipole approach can be used for spherical particles to a coverage of at least 50 %. The quadrupole expansion is given in Appendix A.4. For oblate particles, the image effect is much larger and has to be incorporated at lower coverages. However, also in this case the depolarization factor is strongly affected. As a result, the Drude free electron part will still contribute as a Lorentzian lineshape.

The dipolar approximation of the surface excess strengthens the importance of the depolarization factor and the possibility to analyze experimental spectra from the location of the surface resonance. Equation 6.13 shows that the in plane and out of plane surface resonance will occur at the so-called Fröhlich condition:
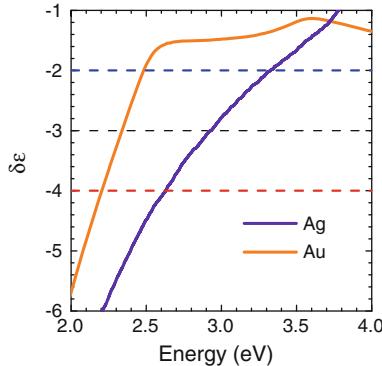
Fig. 6.8 The dielectric contrast of Au (*orange*) and Ag (*purple*) in air as a function of photon energy. The *black dashed line* indicates the Fröhlich resonance condition for spherical particles. The *red* and *blue dashed lines* reflect the in-plane and out-of-plane Fröhlich condition for a surface covered by 44 % with spherical particles. The intersection of the dashed lines and the solid material lines provide the resonance energies

$$\mathrm{Re}(1 + \mathrm{L_p}\delta\varepsilon) = 0$$
$$\mathrm{Re}(1 + \mathrm{L_z}\delta\varepsilon) = 0 \tag{6.15}$$

Next to the value of the depolarization factor, which is mainly determined by shape and surface coverage, it is the dielectric functions of both the nanoparticle and the ambient that determine the resonance energy. For a non-absorbing ambient, this leaves only the real part of the dielectric function as a valid parameter. At the resonance this leads to a value of the dielectric contrast of:

$$\mathrm{Re}(\delta\varepsilon) = -1/\mathrm{L_p}$$
$$\mathrm{Re}(\delta\varepsilon) = -1/\mathrm{L_z} \tag{6.16}$$

Figure 6.8 shows the dielectric contrast for gold and silver particles surrounded by a thiol shell, the situation encountered for the Ag particles on HOPG. For 44 % surface coverage, the values of $\mathrm{L_p}$ and $\mathrm{L_z}$ are respectively 1/4 and 1/2. The dashed blue and red lines indicate the values at which the Fröhlich condition is fulfilled. The plasmon resonance energy for silver particles can be evaluated to occur at 2.6 and 3.3 eV. For gold particles this would be at 2.2 and 2.4 eV. This is a reduction by almost a factor 4 in the split between the in-plane and out-of-plane resonance energy. Taking into account the life time broadening of the resonance peak, these two peaks are difficult to discriminate at this already rather high coverage. Note that the thiol shell enhances the split as it increases the dielectric contrast, see also Eq. 6.16 [20].

A comparison of the Lorentzian parameters in Table 6.1 for both the MG and TIF model is depicted in Fig. 6.2. Only small differences between these two approaches are noted. This explains why the silver particles on HOPG can be described by both MG and TIF. A detailed analysis [11] shows that the fill fraction used in the MG

model nicely mimics the effect of the depolarization factor. However, Fig. 6.1 shows that in case of Au colloids, the MG approach fails to describe the ellipsometric spectra. The reason for this lies in the optical coupling between the nano-entities and the substrate.

6.2.4 The Incorporation of TIF in Ellipsometry

The computation of the optical response of a layer of nanoparticles with TIF does not stop at the evaluation of the surface excess polarizability. This excess polarizability also has to couple to the underlying layer and the ambient above. Effectively, the Fresnel coefficients have to be determined. The formulation of this problem in terms of the Abelès matrix formalism was described by Bohmer et al. [23]. The reflection and transmission coefficients for s-polarized light are:

$$r_\ell^s = \frac{X^s}{1 - X^s} \text{ and } t_\ell^s = \frac{1}{1 - X^s} \tag{6.17}$$

with

$$X^s = \frac{\pi i \gamma}{\lambda n_1 \cos(\theta_1)}$$

and for p-polarized light

$$r_\ell^p = \frac{X^p}{1 - X^p} - \frac{Y^p}{1 - Y^p} \text{ and } t_\ell^p = 1 + \frac{X^p + Y^p - 2X^p Y^p}{(1 - X^p)(1 - Y^p)} \tag{6.18}$$

with

$$X^p = \frac{\pi i \beta (n_1)^3 \sin^2(\theta_1)}{\lambda \cos(\theta_1)} \text{ and } Y^p = \frac{\pi i \gamma \cos(\theta_1)}{\lambda n_1}.$$

The Abelès film matrix for the interface reflection of a particulate film in terms of these reflection and transmission coefficients is given by

$$F = \frac{1}{t_\ell} \begin{pmatrix} 1 & -r_\ell \\ r_\ell & t_\ell^2 - r_\ell^2 \end{pmatrix} \tag{6.19}$$

Note that this matrix describes a non-invariant beam, i.e. the response from the opposite side of the interface is different. The overall matrix becomes

$$A' = F L_F I_{1,3} L_3 I_{3,4}. \tag{6.20}$$

In this the matrices $I_{n,m}$ and L_m are the normal Abelès matrices. The layer matrix L_F, identical to the layer matrices L_m used in the Abelès formalism, has a phase factor
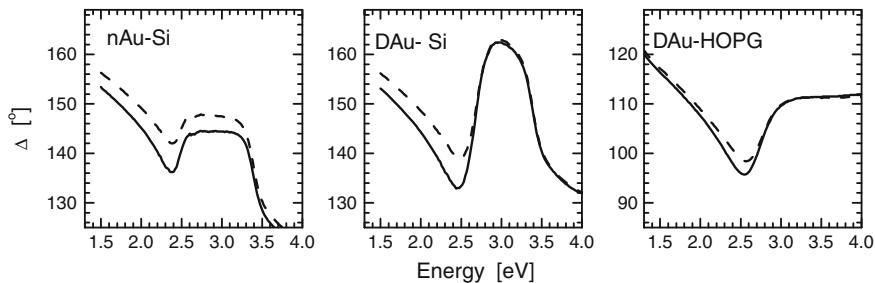
Fig. 6.9 Simulated values of the ellipsometric parameter Δ with the TIF (*solid line*) and MG (*dashed line*) approach for (*left*) use of the actual dielectric function of nanocolloidal Au particles on silicon. (*middle*) use of the Drude function for the Au particles on silicon and (*right*) use of the Drude function for the Au particles on HOPG. An Au volume fraction of 15 % and a particle diameter of 13.6 nm was used

$\Delta = \frac{2\pi}{\lambda} h n_1 \cos\theta_1$, where h is the height of the excess dipoles above the Fresnel interface. If the particles are directly on the substrate surface, h = r.

The influence of the particular coupling between a nanoparticle layer and its underlying layer via the above Fresnel coefficients is the reason for the failure of the MG approach in specific situations. The nature of these specific situations was evaluated for deposited Au particles [11]. Figure 6.9 shows a comparison of ellipsometric spectra evaluated with MG and TIF. The first situation deals with nanoparticles that are made of a Drude metal, i.e. a metal in which the interband transition are far outside the experimental window. If these so-called Drude gold particles would be deposited on HOPG, the two approaches give a similar answer. If instead of HOPG a dielectric substrate like silicon is used, differences are observed. These differences are in the region in which the silicon behaves as a dielectric, far below the first critical point of silicon at 3.4 eV. In the energy range in which silicon has a considerable absorption, the discrepancy vanishes. In the third situation, the dielectric function of real gold is used for the metal particles. In this case a discrepancy between TIF and MG is observed over the whole spectral range. These simulations show that MG works pretty well in the Drude metal region on non-dielectric substrates, i.e. for example silver particles on HOPG. However for Au particles on glass MG predicts the energy resonance quite well, but the intensity is incorrect.

6.3 Examples

6.3.1 The Growth of Silver Films Deposited by Magnetron Sputtering

The initial deposition of silver on a SiO_2 substrate follows the Volmer-Weber (island) growth mode. Initially a large density of silver nanoparticles nucleates on the substrate. These nanoparticles evolve into a conducting ultrathin film with deposition
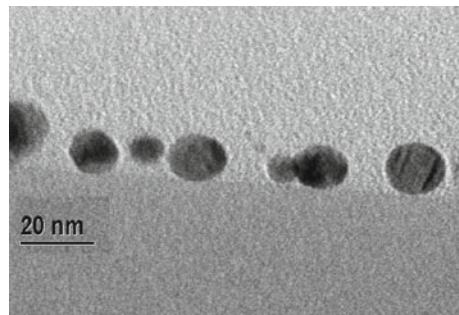
Fig. 6.10 Cross sectional TEM image of silver nanoparticles deposited by magnetron sputtering on SiO$_2$ at 150° C. From Ref. [24], reproduced with permission from Institute of Physics

time. Figure 6.10 shows a TEM image of the silver nanoparticles in the initial stage [24]. A considerable size distribution is observed. The growth and evolution of these nanoparticles into a continuous conducting thin film was characterized with in situ spectroscopic ellipsometry. Figure 6.11 shows the change in dielectric function as obtained from an Arwin-Aspnes (AA) analysis of the measured optical spectra [25]. This is an elegant method that employs the optical interference of the 500 nm thick SiO$_2$ film to establish a value for both the thickness and the real and imaginary part of the dielectric function. The imaginary part of the dielectric function shows a plasmon resonance peak which red-shifts with increasing film thickness. For films beyond 16 nm thickness, the peak maximum lies outside the spectral range. The percolation threshold, marking the transition from a film consisting of isolated particles to a closed film that shows good macroscopic conductivity, can be inferred from the behaviour of the real part. A positive real part signifies a system of isolated nanoparticles while a negative part is related to a continuous film. This rule of thumb puts the percolation threshold around 17 nm. The change in dielectric function with thickness shows that the dielectric properties can be described by a Lorentzian lineshape. Beyond percolation this lineshape transform to a Drude lineshape. This is essentially a Lorentzian with a resonance energy at zero. The evolution of the resonance energy with film thickness is depicted in Fig. 6.12. A sudden drop in resonance energy signifies the percolation threshold. Around this point also the width of the Lorentzian strongly decreases signifying a large increase in the electron mean-free-path. This effect is also directly correlated to percolation. The TIF model couples the decrease in resonance energy directly to a decrease in the depolarization factor of the in-plane plasmon resonance. This is a direct result of the increase in surface coverage of the nanoparticles. At percolation, the depolarization factor will assume its thin film limit value of 0. This results in a resonance energy at 0 eV, i.e. a transition from Lorentzian to Drude lineshape.

It was found that the optical response before percolation could be more accurately described by using two Lorentzians. The second oscillator shows about 0.5 eV lower energy. The data can not resolve whether in the very initial stage this second oscil-
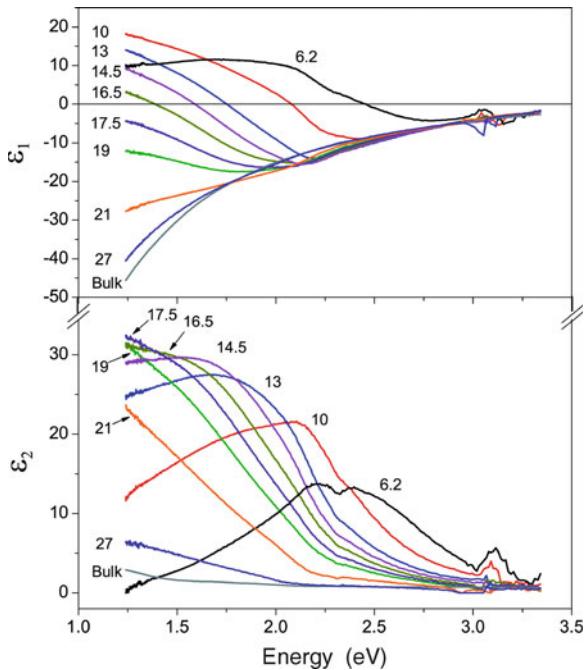
Fig. 6.11 Evolution of the effective dielectric function during silver deposition. Indicated are the evaluated film thicknesses. From Ref. [24], reproduced with permission from Institute of Physics
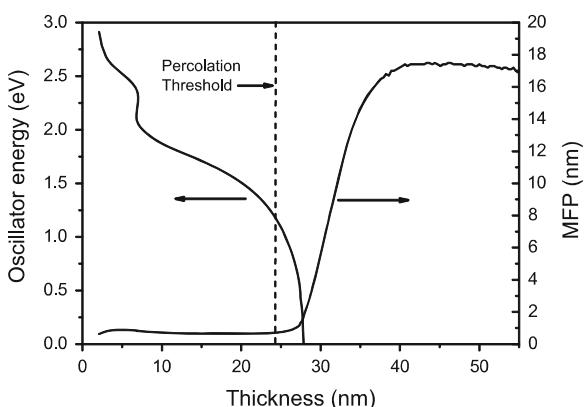

Fig. 6.12 Change with thickness of the oscillator energy and the Mean Free Path used to represent the measured spectra with a Lorentzian lineshape. From Ref. [24], reproduced with permission from Institute of Physics

lator splits of from the original one or that it is a single entity by itself. These two oscillators do not signify an in plane and out of plane plasmon resonance. For such an assignment, they should show an opposite change in resonance energy with surface coverage. The limited size of the particles also excludes a breakdown of the dipole approximation and the observation of a quadrupole contribution. A more likely explanation is in terms of two sets of particles with a different depolarization factor. Initially, the first set characterized by a higher resonance energy is observed. The contribution of this set diminishes with increasing coverage while the second set with the lower resonance energy becomes more prominent up to percolation. The difference in depolarization factor can be attributed to a local variation in the surrounding of the particles or a variation in the particles themselves. TEM images show the presence of particles with a very different size. The size itself should not effect the depolarization factor. However, the larger particles seem to have a slightly oblate shape. Oblate particles have a smaller in plane depolarization factor. A detailed TEM analysis shows that the smaller particle is single crystalline while the larger particle is composed of several crystallites. The two Lorentzians could thus be explained with the initial growth of small spherical particles followed by a coalescence of two or more small particles and the formation of slightly oblate particles with increasing coverage.

6.3.2 A Growing Silver Film

The difference in free energy between silver and an insulating substrate prohibits the growth of an ultra thin film using physical vapour deposition at room temperature and above. Also for electroless deposition of a silver film an island growth is observed initially. In this case the nucleation of small silver islands in the initial phase can be quite difficult leading to very rough films. By using gold colloids as seed particles, the problematic nucleation phase can be circumvented and only a local outgrowth of silver on the already present nuclei can be obtained [26]. Ellipsometric spectra at 3 instants during outgrowth are depicted in Fig. 6.13. Also shown are SEM images at the same point in time as the spectra. After 300s, a spectrum very similar to that of bulk silver is observed. The low energy region of these spectra can be modelled by a growing layer with the dielectric properties represented by a single Lorentzian line shape. The Lorentzian parameters provide the electron density, resonance energy and width of the plasmon resonance. The latter two parameters are shown in Fig. 6.14, with the width transferred to the electron mean free path (MFP). The MFP shows a discontinuous jump around 130s, indicating percolation. The SEM image taken after 150s also shows that percolation has occurred although an additional 150s is required to obtain a continuous layer. At this point, also the fit of the resonance energy reaches 0 eV, signifying the transition towards a Drude behaviour. Note that just before percolation the plasmon resonance energy has considerable red shifted and is outside the measured range. This requires caution in the determination of its position and therefore the determination of the percolation moment. However, as with
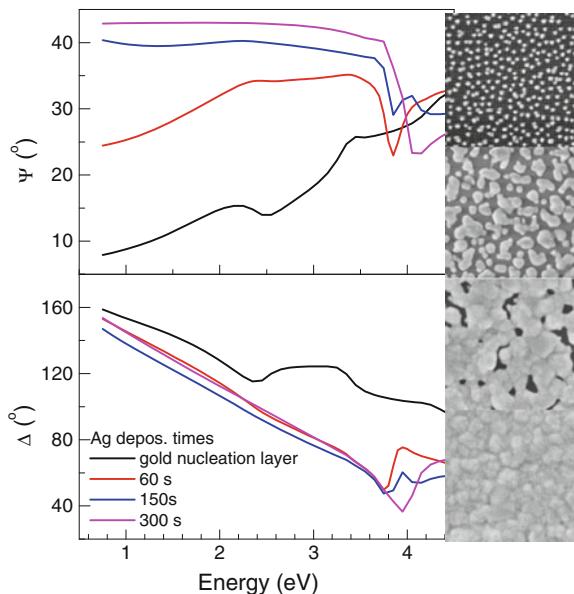
Fig. 6.13 Ellipsometric spectra taken during successive stages of the silver outgrow of gold colloids. The SEM images show the corresponding morphological situation. From Ref. [26], reproduced with permission from Institute of Physics

the previous example, the MFP shows a clear signature of this. Also the sign of the real part of the effective dielectric function of the thin layer helps to limit the error in the resonance energy position. As observed in the previous example, before percolation the low energy tail of the real part is positive, whereas it becomes negative after percolation. However, it also clear that sensitivity in the infrared region is required for an accurate determination of percolation. The value of the MFP is considerably smaller than that of the magnetron deposited film. A comparison with ellipsometric measurements on a silver single crystal as well as resistance measurements showed a factor of 3 smaller MFP.

Within the dipole approximation, the value of the resonance energy can be directly used to calculate the in-plane depolarization factor. It was found that this factor decreases rapidly towards 0, the value of a thin metal sheet [11, 26]. This rapid decrease and thus the strong red shift of the SPP was explained by a non-spherical outgrow of the original spherical gold particles. The formation of slightly oblate particles results in a shape related red shift. However, this does not explain the very strong red shift observed. This can be understood from the fact that for oblate particles the in-plane electromagnetic coupling and image effects are much stronger [18]. An analysis of these effects indicate that these can reduce the depolarization factor by an order of magnitude, leading to a strongly enhanced red shift [11].
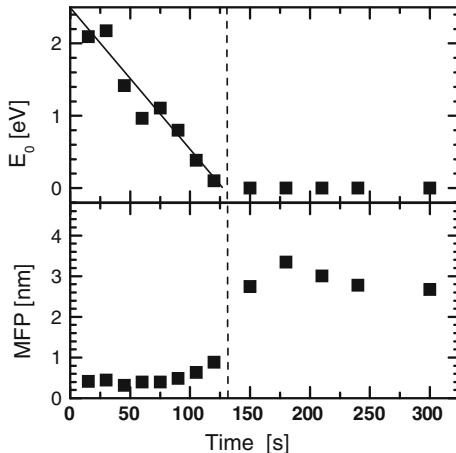
Fig. 6.14 The resonance energy and MFP obtained from a fit of Lorentzian shape dielectric function to the measured ellipsometric spectra. The *dashed line* indicates percolation

6.3.3 The Characterization of Anisotropic Silver Particles

As discussed above, the dielectric function of nanoparticle films is in general uni-axially anisotropic due to the influence of the substrate and the anisotropic particle distribution. This requires us to define an ordinary and extra-ordinary dielectric func-tion for the in-plane and out-of-plane directions, respectively. Experimentally this can be done by making multiple angle-of-incidence ellipsometry measurements and fitting the parameters of a uniaxial model to the measurements simultaneously. Such a study was performed on a magnetron deposited silver island film using incident angles from 15° to 75° [27]. A uniaxial model was defined with a Lorentzian-Tauc Lorentz model for the in-plane dielectric function and a Drude-Tauc Lorentz model (describing the free and bound electrons, respectively) for the out-of-plane dielectric function. As seen in Fig. 6.15a, b an isotropic model cannot accurately describe the out-of-plane resonance near 3.5 eV, while the uniaxial model gives a good fit to the data. Fitting simultaneously over all angles (Fig. 6.15c, d) gives the in-plane (ε_x, ε_y) and out-of-plane (ε_z) dielectric functions, as shown in Fig. 6.16.

While there is little surprise as to the nature of the in-plane DF, showing a strong plasmon resonance at around 2.6 eV, we make note of the nature of the out-of-plane DF. Rather than showing an intrinsic resonance, as would be predicted by the Yamaguchi model ($0 < L_z < 1$), the resonance stems from the entire film polarization (longitudinal mode) when the real part of the effective dielectric function is zero and the imaginary part is small. The out-of-plane DF is thus better described by a simple linear mix of the dielectric and metal, which is equivalent to $L_z = 0$ in the Yamaguchi model. The model therefore describes the particles as cylinders or rods, which is partially correct when one considers the truncation of the spheres on the surface. This also has the advantage that if the islands grow into rods or wires
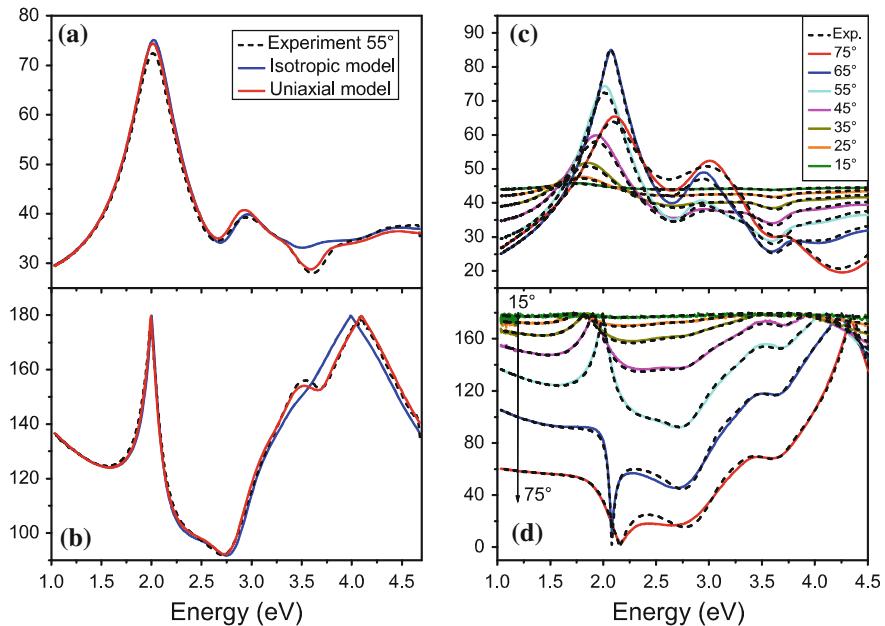
Fig. 6.15 Ellipsometric spectra of a silver nanoparticle film. Depicted are the measured values of **a** Ψ and **b** Δ for an incidence angle of 55° and the results of an isotropic and an uniaxial model. **c** Ψ and **d** Δ as a function of the angle of incidence and the fit of a model with an uniaxial dielectric function for the silver particles. From Ref. [27], reproduced with permission from the Optical Society of America

in the z direction (e.g. via oblique angle deposition) that the effective DF remains constant and only the thickness of the layer changes.

6.4 The Puzzle of Metamaterials

Ongoing advances in miniaturization have reached the point where complex structures can now be accurately fabricated on scales that are far smaller than visible light wavelengths. Composite materials containing arrays (periodic or random) of such tiny structures appear optically homogenous, with the optical properties of the structures (or "meta-atoms") incorporated into the "meta-material". If the meta-atoms are fabricated from metallic materials, sub-wavelength arrays of plasmonic resonators (analogous to electronic circuit elements) may be designed to guide the wave-fronts inside the material—a concept described as transformation optics. The effective macroscopic optical properties of the metamaterials are described by their effective permittivity ε_{eff}, permeability μ_{eff} and refractive index n_{eff}, with particularly interesting cases occurring for ε_{eff} near zero (ENZ), zero refractive index materials (ZIM)
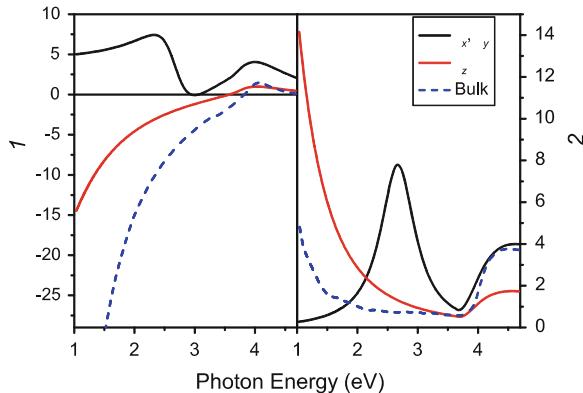
Fig. 6.16 Uniaxial dielectric functions of the silver particles. The confinement of the silver in the *x-y* direction results in a plasmon resonance simulated with a Lorentzian, while the *z* direction is well represented with a Drude lineshape. Also shown is the dielectric function of silver. From Ref. [27], reproduced with permission from the Optical Society of America

and negative refractive index materials (NIM) where both ε_{eff} and μ_{eff} are negative (since n $= (\varepsilon\mu)^{1/2}$).

Experimentally characterizing the optical properties of metamaterials is particularly important for device applications. While spectroscopic ellipsometry is the method of choice for determining the isotropic dielectric function of bulk and thin film materials, extension to anisotropic materials requires the use of generalized ellipsometry. If the material additionally has an inherent magnetic resonance that defines a permeability, or lower symmetry that results in chirality, further measurements and analysis are required.

6.4.1 Sub-Wavelength Wire Arrays

Conceptually stretching the island films of the previous section in the z-direction results in an array of needles or wires standing vertically on the substrate. Such a material has been fabricated by filling the holes of self-organised anodized alumina arrays with silver via electroplating [28], and similarly by oblique angle deposition of silver (albeit with a tilted z-axis)[29]. As pointed out by Pendry [30], if the wires are extremely thin one may assume that the polarization across the wires is negligible and that the medium therefore acts as a diluted metal with a red-shifted plasma frequency (since the electron density is now reduced by an amount proportional to the fill factor). A further surprising result comes when one considers the refraction of light in wire arrays. Due to the extreme anisotropy of the effective dielectric function, a transverse EM wave will be negatively refracted at the interface between the material and an isotropic medium. This does not arise due to a negative refractive index but
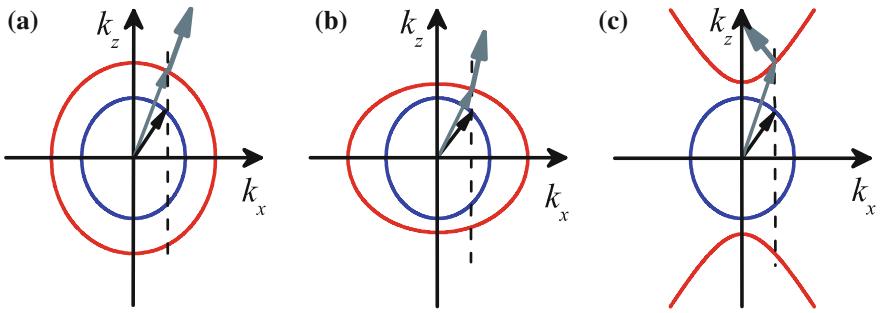
Fig. 6.17 Isofrequency curves of isotropic, anisotropic and indefinite materials. The wavevectors (*thin arrows*) at a given frequency in an isotropic medium (*blue curve*) and the material (*red curve*) form an acute angle. The group velocity (*thick grey arrow*) is normal to the isofrequency surface, which for a hyperbolic dispersion makes an obtuse angle with the wavevector and results in negative refraction. From Ref. [27], reproduced with permission from the Optical Society of America

due to the difference in sign of $\varepsilon_{x,y}$ and ε_z. Such materials are termed indefinite and the dispersion relations are hyperbolic (compared to circular for isotropic materials and elliptical for materials with comparatively small anisotropy). Figure 6.17 illustrates the observation of negative refraction at the interface between a material with an isotropic dielectric function and three different materials with isotropic, weakly anisotropic and hyperbolic anisotropic dispersions of the dielectric functions, respectively. A light vector incident from the isotropic medium (blue) is refracted at the interface. The angle of incidence is indicated by the angle of the wavevector. Upon refraction, the parallel component is conserved (dashed line) and the angle of the refracted wave in the red medium is easily obtained. The group velocity in the red medium is normal to the isofrequency contour. In anisotropic media, the wavevector and the group velocity are no longer parallel. In case of a hyperbolic dispersion an obtuse angle is found between the two. This results in negative refraction effects.

Strong anisotropy can also be introduced in the plane of the substrate by depositing metal island films at an oblique angle onto periodic templates. Numerous variations have been utilized for these templates, ranging from lithographically produced gratings to vicinal surfaces on crystals. Recently the use of low-energy ion beam sputtered rippled patterns as templates has been demonstrated [31, 32]. Aligned silver spheres, rods and wires have been prepared in this way. The plasmon resonances are anisotropic in the plane due to different effective coverage of the particles in the orthogonal directions. By defining a biaxial model and fitting the parameters to measurements made at multiple rotation angles, the orthogonal dielectric function elements have been determined. In this case, generalized ellipsometry must be used since the off-diagonal elements of the Jones matrix will no-longer be zero. The real and imaginary parts of the orthogonal dielectric functions for aligned spheres and rods are shown in Fig. 6.18. The in-plane plasmon resonance is now split into two plasmon peaks, and the splitting increases as the particles go from spheres to rods, i.e. in the direction of the long axis of the particles (y-direction) the plasmon peak is
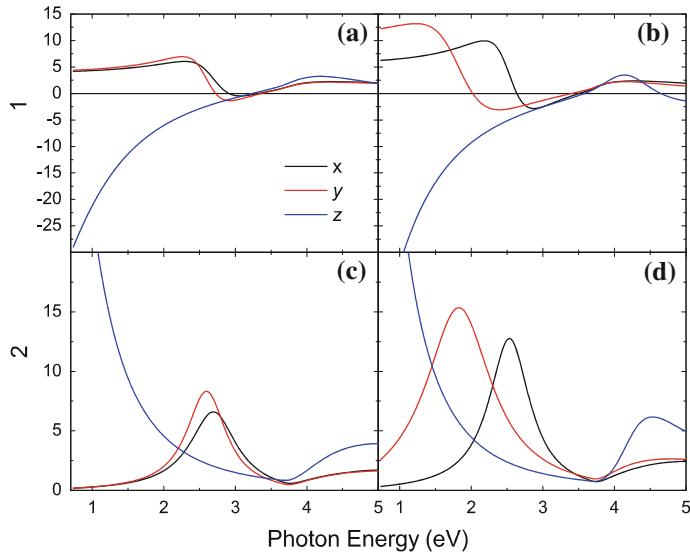
Fig. 6.18 Real (ε_1) and imaginary (ε_2) parts dielectric functions of the aligned spheres **a**, **c** and aligned rods **b**, **d** determined using an anisotropic effective medium model

red-shifted. As more material is deposited the rods eventually connect to form wires and ε_y evolves to resemble the DF of silver with a reduced electron density. Such a DF can be effectively modelled by a linear effective medium approximation.

Wire array metamaterials have also been discussed in the context of ENZ materials. The idea is to move the plasma frequency into the visible and infrared by changing the effective free electron density. While the idea of playing with the plasma frequency has been around for some time, the discussion of ENZ materials has led to some interesting optical concepts. The main interest is the fact that at $\varepsilon = 0$ the phase of the wave in the material is infinite. This implies that EM waves can be tunnelled and squeezed through very narrow channels filled with ENZ materials, and that the incoming planar wavefront is replicated at the output surface, independent of the geometry of the channel [33]. Similar proposals have been made for metamaterials where the permeability, μ, is near zero (MNZ).

6.4.2 Artificial Permeability and Negative Refractive Index

Since there are natural materials where ε is less than zero, manipulating the permeability by fabricating composites of positive and negative ε materials is relatively straightforward. In contrast, at infrared and optical frequencies there are no naturally occurring materials that exhibit permeability other than unity. Therefore manipulating the permeability is not as straightforward. A breakthrough came at the turn of the
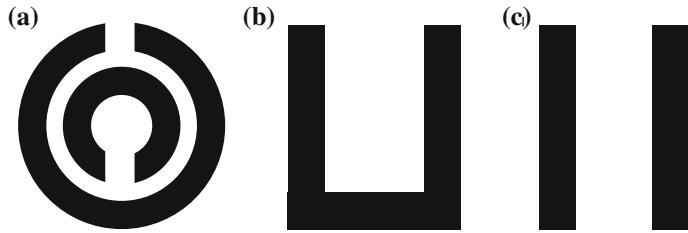
Fig. 6.19 Meta-atoms designs used for artificial magnetism. Shown are **a** concentric split ring resonator, **b** U shaped resonator and **c** coupled rod pair

century when Pendry et al. suggested a material with artificial magnetism composed of dual concentric split ring resonators (SRRs) [34], see Fig. 6.19. A magnetic field directed through the rings will induce a current which will in turn create a magnetic moment, an effect that is enhanced by creating a resonance with a capacitive element—the split in the ring. This concept was demonstrated at microwave frequencies, showing not just artificial magnetism but negative effective permeability and, when combined with a negative permittivity in the form of a wire array, also negative refractive index (NIM) [35].

By shrinking down and simplifying the SRRs to U-shaped resonators a magnetic resonance was also reported in the near infra-red [36]. It was shown that these resonances were in fact due to anti-symmetrically coupled dipolar localised plasmon resonances. In a single SRR, reduced to a U-shaped metallic element, the anti-symmetric resonance was driven by coupling to the base of the "U", (see Fig. 6.19.) Alternatively, coupling can be achieved in parallel rods by using the phase of the incoming light to excite anti-parallel resonances. In such structures one observes at least two resonances; a low frequency symmetric mode and a high frequency anti-symmetric mode. The anti-symmetric mode is associated with a magnetic dipole moment and may be described by an effective permeability. Combining a metamaterial of coupled rods (negative μ) with a wire array medium (negative ε) resulted in the "fishnet" design that has been used to demonstrate negative refractive index in the infrared and visible regions [37].

Characterizing the optical properties of SRRs and fishnets by ellipsometry is challenging. The measurements are relatively straightforward however the modelling required is complicated by spatial dispersion. The original idea of fabricating a material with effective homogenous parameters, ε and μ, being dependent only on the temporal frequency of the electric field, and not on the spatial frequency, is thus invalid. Spatial dispersion must also be considered. This is easily seen when one considers that the difference in phase of the light is used to excite the anti-symmetric resonances in the parallel rods. The dimensions of the unit cell must therefore be appreciable with respect to the wavelength of light. In fact, it is well documented that weak spatial dispersion may in some cases be equivalently described using

ε and μ [38]. Thus the description of an effective permeability is more generally described by spatial dispersion.

6.4.3 Chiral Metamaterials

Chirality has also been proposed as a method to NIMs [39] since the gyrotropic or chiral parameter ξ (which is zero in most materials) is also present in the definition of the complex refractive index

$$N_{L,R} = \sqrt{\varepsilon\mu} \pm \xi$$

If a material could be produced with $\xi > (\varepsilon\mu)^{1/2}$ then in theory a negative refractive index should be possible. In practice ξ is usually very small. It is interesting to note however that chirality (or optical activity) is also a direct result of spatial dispersion in materials without inversion symmetry (see e.g. Landau and Lifschitz [40]). Therefore diffractive arrays should show evidence of chirality. This was recently demonstrated using Mueller matrix (MM) ellipsometry on sub-wavelength hole arrays in gold films [41]. The off-diagonal elements of the MM showed significant resonances which would usually be associated with optical activity in isotropic materials. The authors were able to reconstruct the MM using a model that accounted for spatial dispersion. When interpreting MMs one should always be aware that the off-diagonal elements may also show strong signals due to structural anisotropy (form birefringence) and that decomposing MMs into optical dichroism, birefringence and form dichroism and birefringence is a complicated task [42].

Appendix

A.1 Lorentzian Line Shape for the Polarizability in the Dipole Approximation

The polarizability of a particle with radius a and screening parameter k to incorporate shape effects is:

$$\alpha = 4\pi a^3 \frac{\varepsilon - \varepsilon_a}{\varepsilon + k\varepsilon_a} \tag{A1}$$

Inserting the Drude dielectric function:

$$\alpha = 4\pi a^3 \frac{\varepsilon_\infty - \varepsilon_a - \frac{E_p^2}{E^2 + i\Gamma E}}{\varepsilon_\infty + k\varepsilon_a - \frac{E_p^2}{E^2 + i\Gamma E}}$$

$$= 4\pi a^3 \frac{\left(\frac{\varepsilon_\infty - \varepsilon_a}{\varepsilon_\infty + k\varepsilon_a}\right)(E^2 + i\Gamma E) - \frac{E_p^2}{\varepsilon_\infty + k\varepsilon_a}}{E^2 + i\Gamma E - \frac{E_p^2}{\varepsilon_\infty + k\varepsilon_a}}$$

$$= 4\pi a^3 \frac{\chi\left(E^2 + i\Gamma E\right) - E_{0,\mathrm{dip}}^2}{E^2 - E_{0,\mathrm{dip}}^2 + i\Gamma E}$$

$$= 4\pi a^3 \frac{\chi\left(E^2 - E_{0,\mathrm{dip}}^2 + i\Gamma E\right) + (\chi - 1)E_{0,\mathrm{dip}}^2}{E^2 - E_{0,\mathrm{dip}}^2 + i\Gamma E}$$

$$= 4\pi a^3 \left(\chi + (1 - \chi)\frac{E_{0,\mathrm{dip}}^2}{E_{0,\mathrm{dip}}^2 - E^2 - i\Gamma E}\right). \tag{A2}$$

This is a Lorentzian line shape with the same broadening Γ as the Drude lineshape and with

$$\chi = \frac{\varepsilon_\infty - \varepsilon_a}{\varepsilon_\infty + k\varepsilon_a} \quad \text{and} \quad E_{0,dip}^2 = \frac{E_p^2}{\varepsilon_\infty + k\varepsilon_a}. \tag{A3}$$

A.2 Lorentzian Line Shape in the Maxwell Garnett Approach

The Maxwell Garnett approach provides an effective dielectric function according to:

$$\frac{\varepsilon_{\mathrm{eff}}}{\varepsilon_a} = \frac{\varepsilon(1 + 2f) + \varepsilon_a(k_s - 2f)}{\varepsilon(1 - f) + \varepsilon_a(k_s + f)}. \tag{A4}$$

For a Drude metal particle this leads to:

$$\frac{\varepsilon_{\mathrm{eff}}}{\varepsilon_a} = \frac{\left(\varepsilon_\infty - \frac{E_p^2}{E^2 + i\Gamma E}\right)(1 + 2f) + \varepsilon_a(k_s - 2f)}{\left(\varepsilon_\infty - \frac{E_p^2}{E^2 + i\Gamma E}\right)(1 - f) + \varepsilon_a(k_s + f)}$$

$$= \frac{\varepsilon_{\infty,\mathrm{MG}}\left(E^2 + i\Gamma E\right) - E_{0,\mathrm{MG}}^2 \frac{1 + 2f}{1 - f}}{E^2 - E_{0,\mathrm{MG}}^2 + i\Gamma E}$$

$$= \varepsilon_{\infty,\mathrm{MG}} + \frac{E_{p,\mathrm{MG}}^2}{E_{0,\mathrm{MG}}^2 - E^2 - i\Gamma E} \tag{A5}$$

which is a Lorentzian line shape with

$$\varepsilon_{\infty,\mathrm{MG}} = 1 + \frac{3f(\varepsilon_\infty - \varepsilon_a)}{\varepsilon_\infty(1 - f) + \varepsilon_a(k_s + f)} \tag{A6}$$

$$E_{0,MG}^2 = \frac{(1-f)}{\varepsilon_\infty(1-f) + \varepsilon_a(k_s + f)} E_p^2.$$

The width Γ of the Lorentzian is the same as the original Drude broadening. The enumerator $E_{p,MG}^2$ in the last term of Eq. 6.14 can be rewritten

$$E_{p,MG}^2 = E_{0,MG}^2 \left(\frac{1+2f}{1-f} - \varepsilon_{\infty,MG} \right)$$

$$= 3f \frac{\varepsilon_a(1+k_s)}{[\varepsilon_\infty(1-f) + \varepsilon_a(k_s + f)]^2} E_p^2. \tag{A7}$$

A.3 Lorentzian Line Shape in the TIF Approach

The derivation of the dimensionless polarizability lineshape for a Drude metal is very similar to A1:

$$\alpha' = (1+k)\frac{\varepsilon_\infty - \varepsilon_a - \dfrac{E_p^2}{E^2 + i\Gamma E}}{\varepsilon_\infty + k\varepsilon_a - \dfrac{E_p^2}{E^2 + i\Gamma E}}$$

$$= (1+k)\left(\chi + (1-\chi)\frac{E_{0,dip}^2}{E_{0,dip}^2 - E^2 - i\Gamma E} \right). \tag{A8}$$

This is again a Lorentzian line shape with the same broadening Γ as the Drude lineshape and with as for the polarizability

$$\chi = \frac{\varepsilon_\infty - \varepsilon_a}{\varepsilon_\infty + k\varepsilon_a} \quad \text{and} \quad E_{0,dip}^2 = \frac{E_p^2}{\varepsilon_\infty + k\varepsilon_a}. \tag{A9}$$

The parallel component of the dielectric function is thus (Eq. 6.6):

$$\frac{\varepsilon_{effP}}{\varepsilon_a} = 1 + f(1+k_P)\chi + \frac{f(1+k_P)(1-\chi)E_{0,dip}^2}{E_{0,dip}^2 - E^2 - i\Gamma E} \tag{A10}$$

$$= \varepsilon_{\infty,dip} + f\frac{\varepsilon_a(1+k_P)^2}{\varepsilon_\infty + k_P\varepsilon_a} \frac{E_{0,dip}^2}{E_{0,dip}^2 - E^2 - i\Gamma E}$$

$$= \varepsilon_{\infty,dip} + \frac{E_{p,dip}^2}{E_{0,dip}^2 - E^2 - i\Gamma E} \tag{A11}$$

with

$$\varepsilon_{\infty,dip} = 1 + f(1 + k_P)\frac{\varepsilon_\infty - \varepsilon_a}{\varepsilon_\infty + k_P\varepsilon_a} \tag{A12}$$

$$E^2_{p,dip} = f\frac{\varepsilon_a(1 + k_P)^2}{(\varepsilon_\infty + k_P\varepsilon_a)^2}E^2_p.$$

The resonance energy $E_{0,dip}$ in the dipolar approximation of the TIF model looks at first the same as that of the isolated particles. However, do note that the screening factor k not only includes shape effects, but also image and neighbour interaction effects.

A.4 Quadrupole Expression for the Excess Polarizability

Haarmans and Bedeaux [18] derived an explicit form for the expression up to quadrupole order for excess surface polarizability γ and β:

$$\gamma_{qu} = \phi\frac{4a\varepsilon_a}{3}\frac{\delta\varepsilon\left(1 + L1_p\delta\varepsilon\right)}{\left(1 + L_p\delta\varepsilon\right)\left(1 + L1_p\delta\varepsilon\right) + \Lambda_p\delta\varepsilon^2} \tag{A13}$$

$$\beta_{qu} = \phi\frac{4a}{3\varepsilon_a}\frac{\delta\varepsilon\left(1 + L1_z\delta\varepsilon\right)}{\left(1 + L_z\delta\varepsilon\right)\left(1 + L1_z\delta\varepsilon\right) + \Lambda_z\delta\varepsilon^2} \tag{A14}$$

Here, L_p and L_z represent the dipolar correction terms and $L1_p$ and $L1_z$, Λ_p and Λ_z are the quadrupole depolarization factors.
The image effect is modulated by the factor B_{sa} that describes the contrast between ambient and substrate:

$$B_{sa} = (\varepsilon_a - \varepsilon_s) / (\varepsilon_a + \varepsilon_s) \tag{A15}$$

This contrast is quite considerable for semiconductors and metals ($B_{sa} \approx -1$), while it is quite reduced for dielectrics. For instance for a glass substrate in a water ambient, B_{sa} and thus the image effect, is reduced by an order of magnitude.
The dipolar and quadrupole depolarization factors as a function of surface coverage are:

$$L_p = \frac{1}{3}\left[1 - \frac{\phi}{2} + B_{sa}\left(\frac{1}{8} - \frac{\phi}{4\sqrt{2}}\right)\right] \tag{A16}$$

$$L_z = \frac{1}{3}\left[1 + \phi + B_{sa}\left(\frac{1}{4} - \frac{\phi}{2\sqrt{2}}\right)\right]$$

$$L1_p = \frac{2}{5}\left[1 - \frac{\phi}{8} + \frac{B_{sa}}{8}\left(1 - \phi\left(\frac{1}{\sqrt{2}}\frac{3}{8}\right)\right)\right]$$

$$L1_z = \frac{2}{5}\left[1 + \frac{3}{16}\phi + \frac{3B_{sa}}{16}\left(1 - \phi\left(\frac{1}{\sqrt{2}}\frac{3}{8}\right)\right)\right]$$

$$\Lambda_p = -\frac{B_{sa}^2}{640}\left(-1 + \frac{\phi}{\sqrt{2}}\right)^2$$

$$\Lambda_z = -\frac{3B_{sa}^2}{640}\left(-1 + \frac{\phi}{\sqrt{2}}\right)^2$$

The factors Λ_p and Λ_z are quite small and are only present if image effects play a role. Very often, Λ_p and $\Lambda_z \ll 1$ in a coverage range up to 50 %. In this case the quadrupole contribution vanishes, and only the dipole contribution remains. Note that as a results of the image effect the sum of the depolarization factors

$$2L_p + L_z = 1 + \frac{B_{sa}}{3}\left(\frac{1}{2} - \frac{\phi}{\sqrt{2}}\right) \tag{A17}$$

no longer equals 1. This rule is broken as a result of the image effect.

References

1. J.C.M. Garnett, Phil. Trans. Roy. Soc. Lond. **203A**, 385 (1904)
2. J.C.M. Garnett, Phil. Trans. Roy. Soc. Lond. ibid **205A**, 237 (1906)
3. O.S. Heavens, Optical properties of thin films. Rep. Prog. Phys. **23**, 2–65 (1960)
4. D.A.G. Bruggeman, Ann. Phys. (Leipzig) **24**, 636 (1935)
5. D.E. Aspnes, Am. J. Phys. **50**, 704 (1982)
6. H. Fujiwara, *Spectroscopic Ellipsometry* (Wiley, Hoboken, 2007)
7. E.S. Kooij, H. Wormeester, E.A.M. Brouwer, E. van Vroonhoven, A. van Silfhout, B. Poelsema, Langmuir **18**, 4401 (2002)
8. G. Mie, On the optics of turbid media, especially colloidal metal solutions. Ann. Phys. -Berlin **25**, 377–445 (1908)
9. U. Kreibig, M. Vollmer, *Optical Properties of Metal Clusters* (Springer, New York, 1995)
10. T.W.H. Oates, H. Wormeester, H. Arwin, Prog. Surf. Sci. **86**, 328 (2011)
11. H. Wormeester, E.S. Kooij, B. Poelsema, Phys. Stat. Sol. A **205**, 756 (2008)
12. S. Yamaguchi, J. Phys. Soc. Jpn. **15**, 1577 (1960)
13. A. Hilger, M. Tenfelde, U. Kreibig, Appl. Phys. B **73**, 361 (2001)
14. A. Rijn van Alkemade, Ph.D. thesis, University of Leiden, 1881.
15. J. Lekner, *Theory of Reflection* (Martinus Nijhof Publishers, Dordrecht, 1987)
16. T. Yamaguchi, S. Yoshida, A. Kinbara, Thin Solid Films **21**, 173 (1973)
17. D. Bedeaux, J. Vlieger, Phys. A **67**, 55 (1973)
18. D. Bedeaux, J. Vlieger, *Optical Properties of surfaces* (Imperial College Press, London, 2002)
19. M.T. Haarmans, D. Bedeaux, Thin Solid Films **224**, 117 (1993)
20. H. Wormeester, A.I. Henry, E.S. Kooij, B. Poelsema, M.P. Pileni, J. Chem. Phys. **124**, 204713 (2006)
21. R. Lazzari, I. Simonsen, Thin Solid Films **419**, 124 (2002)
22. R.G. Barrera, M. Delcastillomussot, G. Monsivais, P. Villasenor, W.L. Mochan, Phys. Rev. B **43**, 13819 (1991)
23. M.R. Bohmer, E.A. van der Zeeuw, G.J.M. Koper, J. Colloid Interface Sci. **197**, 242 (1998)
24. T.W.H. Oates, A Mucklich. Nanotechnology **16**, 2606–2611 (2005)

25. H. Arwin, D.E. Aspnes, Thin Solid Films **113**, 101 (1984)
26. A.J. de Vries, E.S. Kooij, H. Wormeester, A.A. Mewe, B. Poelsema, J. Appl. Phys. **101**, 053703 (2007)
27. T.W.H. Oates, M. Ranjan et al., Highly anisotropic effective dielectric functions of silver nanoparticle arrays. Opt. Express **19**(3), 2014–2028 (2011)
28. J. Yao, Z.W. Liu et al., Optical negative refraction in bulk metamaterials of nanowires. Science **321**(5891), 930–930 (2008)
29. Y.J. Jen, A. Lakhtakia et al., Vapor-deposited thin films with negative real refractive index in the visible regime. Opt. Express **17**(10), 7784–7789 (2009)
30. J.B. Pendry, A.J. Holden et al., Low frequency plasmons in thin-wire structures. J. Phys. Condens. Matter **10**(22), 4785–4809 (1998)
31. T.W.H. Oates, A. Keller et al., Aligned silver nanoparticles on rippled silicon templates exhibiting anisotropic plasmon absorption. Plasmonics **2**(2), 47–50 (2007)
32. T.W.H. Oates, A. Keller et al., Self-organized metallic nanoparticle and nanowire arrays from ion-sputtered silicon templates. Appl. Phys. Lett. **93**(6), 3 (2008)
33. M. Silveirinha, N. Engheta, Tunneling of electromagnetic energy through subwavelength channels and bends using epsilon-near-zero materials. Phys. Rev. Lett. 97(15) (2006).
34. J.B. Pendry, A.J. Holden et al., Magnetism from conductors and enhanced nonlinear phenomena. IEEE Trans. Microw. Theory Tech. **47**(11), 2075–2084 (1999)
35. D.R. Smith, W.J. Padilla et al., Composite medium with simultaneously negative permeability and permittivity. Phys. Rev. Lett. **84**(18), 4184–4187 (2000)
36. T.J. Yen, W.J. Padilla et al., Terahertz magnetic response from artificial materials. Science **303**(5663), 1494–1496 (2004)
37. G. Dolling, M. Wegener et al., Negative-index metamaterial at 780 nm wavelength. Opt. Lett. **32**(1), 53–55 (2007)
38. V.M. Agranovich, Y.N. Gartstein, Spatial dispersion and negative refraction of light. Phys. Usp. **49**(10), 1029–1044 (2006)
39. S. Tretyakov, I. Nefedov et al., Waves and energy in chiral nihility. J. Electromagn. Waves Appl. **17**(5), 695–706 (2003)
40. L.D. Landau, L.P. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1984)
41. B. Gompf, J. Braun et al., Periodic nanostructures: spatial dispersion mimics chirality. Phys. Rev. Lett. **106**, 185501 (2011)
42. R. Ossikovski, M. Anastasiadou et al., Depolarizing Mueller matrices: how to decompose them? Phys. Stat. Sol. A **205**, 720 (2008)

Chapter 7
Spectroscopic Ellipsometry on Metallic Gratings

Michael Bergmair, Kurt Hingerl and Peter Zeppenfeld

Abstract During the last decades light was successfully modelled by micro and nano structures such as photonic crystals, gratings, wave couplers, and, in past years, also negative index materials. For the visible range a further miniaturization of the feature sizes of these metamaterials is necessary. Advanced fabrication facilities as well as investigation techniques are required by science and industry. This chapter deals with spectroscopic ellipsometry (SE) applied to the investigation of metallic gratings and the comparison with theoretical simulations. The comparison of theoretical and experimental spectra of metallic gratings provides a useful basis for more complicated metamaterials, such as negative index materials. This analysis is shown in the beginning of this chapter. Deviations in the structure are identified and it is shown that RCWA and SE constitute a powerful combination to analyze nanostructures. It is followed by the study of surface plasmons on metallic grating which are excited due to the periodic structure. These excitation frequencies are clearly visible in the ellipsometric spectra due to a phase shift and the ratio of reflected intensities between s and p polarization.

M. Bergmair(✉) · K. Hingerl
ZONA—Center for Surface- and Nanoanalytics,
Johannes Kepler University, Altenbergerstr. 69,
4040 Linz, Austria
e-mail: michael.bergmair@jku.at

K. Hingerl
e-mail: kurt.hingerl@jku.at

P. Zeppenfeld
Atomic Physics and Surface Science Division at the Institute of Experimental Physics,
Johannes Kepler University, Altenbergerstr. 69,
4040 Linz, Austria
e-mail: peter.zeppenfeld@jku.at

7.1 Introduction

During the last years the fabrication of structured surfaces has made remarkable progress towards the nano scale. Various techniques now open the possibility to create structures that are highly reproducible, have sharp contours and even offer three dimensional structures by precisely orientated stacking of layers [5].

One of these techniques is nanoimprint lithography (NIL) which allows scientists and industry the realization of ordered structures over large areas. A stamp which is fabricated by electronic beam writing is used as a positive or negative template. Therefore the desired structure can be reproduced in short fabrication time as only the writing of a large-scale master structure with e-beam is time consuming. This provides a method to fabricate nanostructures with very small feature sizes (around 100 nm) over several centimeters.

The characterization of nanostructures is an important issue as many breakdowns may occur in the fabrication process. In general, micro- and nano-structures are investigated by techniques that spatially resolve the structure (scanning electron microscopy (SEM), transmission electron microscopy (TEM), atomic force microscopy (AFM), ...). These investigation tools provide images of the sample and its composition but the investigation time is large and the scanning techniques are cumbersome, as e.g., vacuum is needed. Micro- and nano-structured samples that are extending over centimeters can be characterized by other techniques. One of them is spectroscopic ellipsometry (SE). The reflectivity spectrum of a sample yields next to the absolute values also phase information of the structure and is therefore sensitive to structural variations on the nanometer scale. Furthermore, this method is non-invasive and non-destructive and can also be used in-situ to monitor fabrication processes.

In this part we study gratings that are fabricated by NIL. Metallic gratings exhibit special optical features which are seen as resonances in the spectrum. These resonances are sensitive to the periodicity of the structure. We show that one dimensional metallic gratings exhibit these surface plasmonic resonances because their lattice vector and its multiples close the wavevector gap in reciprocal space between the incident light and the surface plasmon dispersion relation. Contrary to the Kretschmann configuration [28], where a prism is used to provide the missing wavevector, metallic gratings allow more than one resonance energy.

This part is organized as following: First we give a short introduction to the fabrication procedure that allows us to make different gratings. The gratings differ in their geometrical parameters and material composition. It is necessary to study the gratings with cross correlation techniques, that are explained in Sect. 7.3.1 to ensure perfect periodicity and regularity of the structure. As the main goal is the characterization of gratings and their geometry by SE, we give a short introduction of this technique in the following section. Before introducing the exact calculation procedure in Sect. 7.4.1 we summarize the results that are obtained by effective medium theories. They provide a first approximation as the gratings can be roughly approximated by homogeneous layers of mixed materials. We compare the results

with the exact one of the real grating structures; the latter is obtained by the rigorous coupled wave analysis (RCWA). We introduce this powerful tool to calculate reflected and transmitted fields in Sect. 7.4.1.

Before comparing the measured and calculated optical spectra for silicon, gold and silver gratings (Sects. 7.5.1–7.5.3), we introduce a model to explain certain peaks in the spectrum. We show that surface plasmonic resonances are clearly visible in ellipsometric spectra and match with the predictions of the exact calculations. We analyze the line shape of the resonances and identify them as Fano type. Finally, we calculate the complex reflection and transmission spectra for normal incidence in Sect. 7.6. These spectra are used to test a homogenization algorithms discussed in recent literature and the resulting effective parameters are presented and analyzed.

7.2 Sample Structure and Fabrication

The gratings that are subject of our investigation in this work are fabricated by nanoimprint lithography (NIL) and a subsequent two layer resist lift-off process. First the transfer layer LORA1 and the UV-curable resist mr-UVCur21 are spin coated on top of a Si wafer [56]. The NIL stamp with a structured area of a few cm^2 contains line and space structures of different periods, e.g., ranging from $P = 2.4\,\mu$m to $P = 800$ nm which are located on different fields of the stamp. Using this NIL stamp the mr-UVCur21 is structured by bringing stamp and substrate into contact and curing the resist with UV light. After separation a negative copy of the stamp remains on the substrate (see Fig. 7.1a). A subsequent etching step through the transfer layer down to the substrate and controlled development of the transfer layer ends up in a line and space resist structure (b) with recessed sidewalls (c) which are necessary for a successful final lift-off process (e). After creating recessed sidewalls thin layers of metal are deposited. The deposited layer thicknesses of the metals determine the height of the grating (d). A Ti layer is needed to increase the adhesion of Au or Ag to the substrate. The resist and, hence, also the metal deposited on top of the resist are removed by putting the sample in the developer again (e), such that only the metal which was deposited on the substrate remains. The results are gold or silver gratings on an area of 2×2 cm^2 with periods ranging from 2.4 μm to 800 nm for our stamp designs.[1]

In Fig. 7.2a we see a schematic of the fields produced by the stamp. There are 16 fields arranged in four different sectors. Each field is 3×3 mm^2 and four different periods are realized [16]. Therefore, we have four different gratings which are four times repeated on a single sample. This allows us to investigate reproducibility of the structures by SE and cross correlation measurements. The lines and spaces of the grating are always oriented in north (N)–south (S) direction. We identify the fields according to their periodicity, i.e., *Field* (a) has a period of 2400 nm, at *Field* (b) $P = 1600$ nm, *Field* (c) $P = 1200$ nm and at *Field* (d) the period is 800 nm

---

[1] The fabrication of the grating was performed at PROFACTOR GmbH by *Iris Bergmair*.
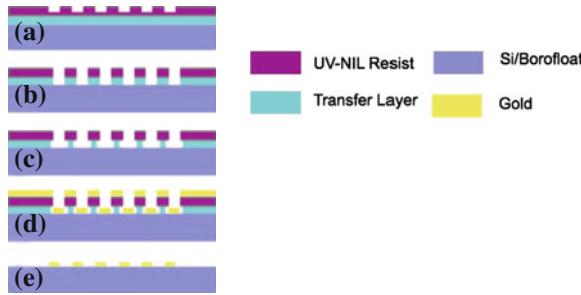
Fig. 7.1 Fabrication steps **a–e** to produce a metallic grating: **a** imprinting in two layer resist system; **b** etching of residual layer of UV-NIL resist and through transfer layer; **c** development of transfer layer to obtain recessed sidewalls; **d** deposition of metal (gold); **e** lift-off
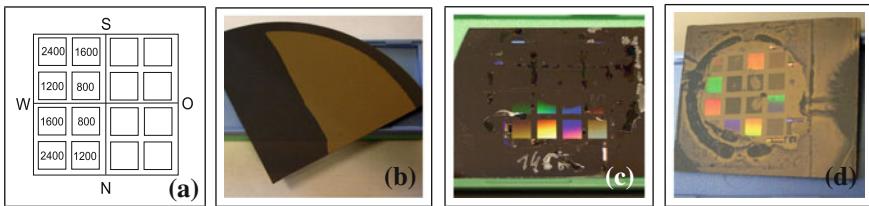

Fig. 7.2 Schematic drawing and pictures of gratings: **a** the schematic of the sample with four times four different Fields. In **b** is a piece of a reference wafer with an area of gold and titanium (*left side*). At the edges of the wafer no metal was evaporated and, therefore, the thickness of natural silica on top of silicon can be determined. A silicon grating is shown in **c** with the 4 × 4 different fields. In **d** we show the various fields of a gold grating

(see Fig. 7.2c, d for the fields on the sample). The line and space width have a ratio of 1:3 on the original stamp. This means for the period of 2400 nm the linewidth is $L = 600$ nm whereas the space has a width of $P - L = 1800$ nm. For the fabrication of the metallic grating working stamps are used [38], this means a negative copy of the stamp. By replicating the working stamp twice one gets a positive copy of the stamp. These two different copies give the possibility to fabricate gold gratings with two different linewidths but the same period. One time with a small protruding metal linewidth and one time with wider metal linewidth.

To characterize the produced grating it is necessary to know the thickness of all materials evaporated on the silicon wafer. The determination of these layer thicknesses is demanding without destroying the sample. The easiest way is to measure the oxide thickness by SE before evaporating the various metals on it. The thickness of the metallic layers are obtained by placing an unstructured wafer in the evaporation chamber and afterwards measuring the SE spectrum (see Fig. 7.2b—homogeneous control wafer). In the visible range this works quite well except for optically thick metals (e.g., for a gold thickness >40 nm). In the case of optically thick metal layers one can prepare a small lamella of the control wafer by focused ion beam and investigate the cross section by TEM. Therefore, we have a good cross-check on

the sample geometry and can compare it with the results from SE (we describe the procedure in detail in the next section) and from theoretical simulations. By fitting the thickness of the control wafer we find out whether our deposited metal has voids as the dielectric response of the metal is changed due to the evaporation procedure. This has to be considered in the simulations.

The gratings have several parameters that are still unknown. Besides the different materials and their composition, the geometrical parameters like the period P of the structure, the linewidth L, the spacewidth $P - L$ and height h of the (metallic) grating have to be checked. To this end, we use several cross-checking techniques which are introduced in the next section.

7.3 Measurement of Gratings

7.3.1 Cross Correlation Techniques

The measurement of nano structures is a necessity to ensure a correct fabrication process. There are several techniques which we introduce below that are more or less time intensive. Furthermore some methods need special facilities like high vacuum.

The nano structured gratings that we investigate in this work are produced on an area of several square millimeters which provides the possibility to investigate them by SE. One main goal of this work is to show that ellipsometry is a good choice in the characterization of nano structures as it is a fast, non-invasive and non-destructive method. We also want to emphasize that for a complete characterization other methods than SE are needed, which are introduced below (for further information, see [15]).

7.3.1.1 Scanning Electron Microscopy

The surface morphology of a sample can be investigated by scanning electron microscopy (SEM). The surface is scanned with an electron beam and the back-scattered electrons are detected. In Fig. 7.3 we see two typical top view pictures[2] of a grating.

In (a) we see the top view of a silicon grating. As the substrate and the lines of the grating consist of the same material the structure has a low contrast in the SEM image. We see that the grating bars are almost rectangular and have a well defined periodicity. The SEM software allows the user to set markers on the sample image to determine the geometrical dimensions of the sample. Therefore, **by SEM we determine the periodicity, the line and the space width.** In (b) we show the SEM picture of a gold grating. As the grating lines are made of gold and the underlying

---

[2] The SEM pictures and the analysis were made by Iris Bergmair at the University Linz on behalf of Profactor GmbH.
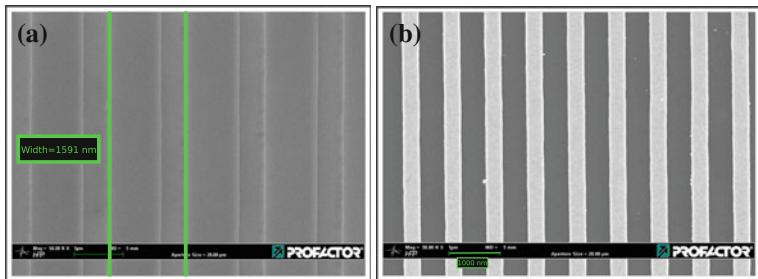
Fig. 7.3 SEM images of a silicon grating with $P = 1600$ nm (**a**) and a gold grating with a periodicity of $P = 800$ nm (**b**)

substrate is silicon, we get a high image contrast. From the image we conclude that the structure is well fabricated. Some bright spots at the edge of the lines and between them indicate left-overs of gold.

7.3.1.2 Atomic Force Microscopy

An atomic force microscope (AFM) is an appropriate investigation tool to measure a height profile of a nano structured sample. In an AFM, which was invented by Binning et al. [7], the vertical force between the sample and a tip is measured. The tip, which has a radius of around 20–50 nm, is fixed at the end of a cantilever, which oscillates with its resonance frequency. The forces (long and short range as, e.g., chemical binding, frictional, Van der Waals, electrostatic) between the tip and the surface are detected by the reflection signal of a laser beam which is pointed at the cantilever. Two main modes can be used: In contact mode, the tip touches the sample surface. The surface scan can be done in constant height or constant force mode. In the other, so called non-contact mode, the tip and the surface are separated and, therefore, mainly long-range interactions are measured as the cantilever oscillation frequency and amplitude are changed due to forces between the tip and surface.

The AFM tip is moving relative to the sample and may show some artificial broadening effects due to the finite size of the tip and/or a too fast scanning velocity. In Fig. 7.4 we see two AFM screenshots of the software of (a) a silicon grating and (b) a gold grating. The surface profile is recorded in a very clear way; we get information about the period, linewidth and the height of the structure. The first two parameters are slightly enhanced by possible broadening effects but for gratings with a sufficiently large spacewidth the height of the grating is exactly determined by AFM.[3] In this example the height of the silicon grating is determined to $h = 21$ nm and the gold bars have a height of $h = 55$ nm. In addition, we obtain information on

---

[3] The AFM measurements were done by Ahmed Saeed. The measurement was done in the tapping mode which is partially contact mode.

**(a)**   **(b)**
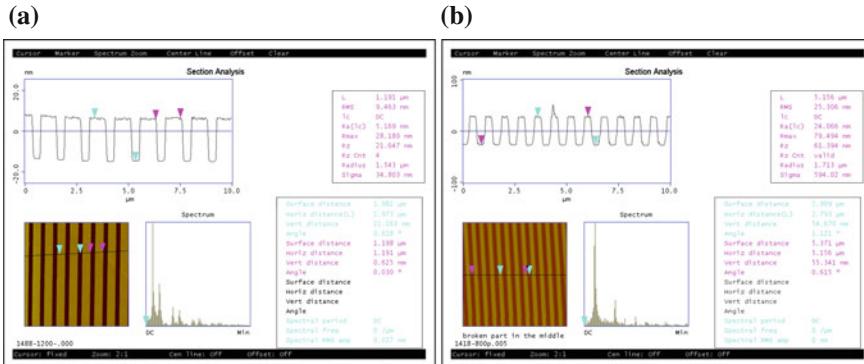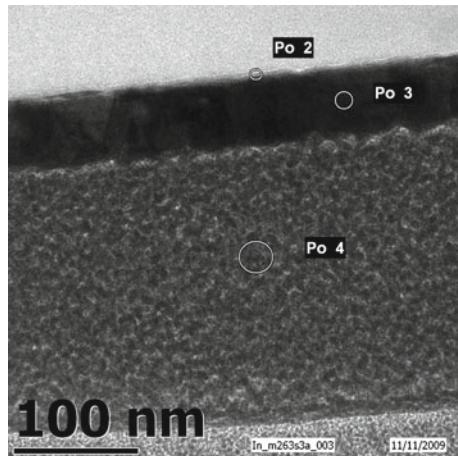
Fig. 7.4 Two AFM images of samples: **a** a silicon grating with a period $P = 1200$ nm and a height of $h = 21$ nm; **b** a gold grating with $P = 800$ nm and a height of $h = 55$ nm

Fig. 7.5 A TEM image of a thin lamella of a gold layer on a silicon wafer (control wafer)

the periodicity of the sample and the overall shape of the grating (e.g., in (b) we see a spike on one gold bar indicating a left-over of the resist stripping).

7.3.1.3 Transmission Electron Microscopy

In a transmission electron microscope (TEM) a beam of electrons is transmitted through a very thin sample. In our case the grating is placed on a silicon substrate in vertical direction and in horizontal direction we have dimensions on a macroscopic scale.

Therefore we have to provide a thin lamella which is cut with a focussed ion beam (FIB) which takes several hours. After the small specimen is prepared we have to transfer the sample to the TEM chamber which is at ultra-high vacuum.

Table 7.1 A list of various characterization techniques, their advantages and the average time for a measurement

Technique	Information	Time
SE	thickness, $\varepsilon(\omega)$	minute(s)
AFM	Geometrical parameters of a grating: P, L, h	hour(s)
SEM	Geometrical parameters P and L, lineshape	hour(s)
Fib-Rem/TEM	Layer thickness, layer composition (diffusion)	hour(s)

Therefore an investigation of a sample takes about a full day. We can only visualize cross sections with TEM to obtain information about layer thicknesses. Using EDX (electron-dispersive X-ray spectroscopy) we further obtain information on the material distribution at certain spots in a sample (see three circles in Fig. 7.5). We, therefore, obtain information about possible diffusion or mixing effects which may occur during or after the evaporation process.

In Fig. 7.5 we have investigated a lamella of a homogeneous gold layer from the control wafer (thickness $h \sim 50$ nm).[4] We estimate the thickness of the gold layer (at Po 3) and can also make a rough guess about the interface with the Si substrate (Po 2). The large area (Po 4) is a copper layer which is the actual sample holder.

Summarizing, in Table 7.1 we show the cross correlation techniques in comparison with SE. We list the information that can be obtained with each technique.

Note that, with SE and numerical simulations we also determine layer thicknesses and/or the dielectric function of the materials. It is by far the fastest technique to characterize structures. SE does not give any direct information about feature sizes or surface compositions. Therefore, we first have to check with unstructured, homogeneous test samples the structure by other techniques to be sure to have the desired structure and periodicity. Then we measure the ellipsometric spectrum and compare the result with simulations of the appropriate sample geometry to get confidence about the model.

The real power of SE is always based on a combination with theoretical simulations. The model we introduce in Sect. 7.4.1 is exact and allows us to derive reflectance and polarimetric spectra for given geometries. Thus we determine in which way deviations change the spectra and, therefore, SE can be (an in industry is) used for the control of fabrication processes.

7.4 Theoretical Description

As we have seen in the previous section, a metallic grating with structure sizes comparable to the incident wavelength cannot be modelled by effective medium approximations. Therefore, we follow an analytical formalism that allows us to cal-

---

[4] This work was performed by *Sajjad Tollabimazraehno*.

culate the reflection and transmission properties of structured materials exactly. We further discuss features in grating spectra that do not appear in EMA results and introduce a theoretical model to predict the excitation frequencies of these surface plasmonic resonances.

7.4.1 Simulation Procedure

As mentioned before, the standard ellipsometry software only allows the user to calculate SE spectra for homogeneous layers, which might also be an effective medium of randomly mixed constituents.[5] We are interested in structured, periodic surfaces and therefore have to simulate the SE spectra (Ψ as well as Δ) by a different method than the stratified layer approach implemented in standard ellipsometry software. This model, based on Fourier expansion of the structure, is called rigorous coupled wave analysis (RCWA).

For a system without free carriers and currents, the curl equations of Maxwell read

$$\nabla \times E = -\mu \frac{\partial H}{\partial t} \tag{7.1a}$$

$$\nabla \times H = \varepsilon \frac{\partial E}{\partial t} \tag{7.1b}$$

We assume a harmonic time dependence of the field and use the convention $E(t) = E_0 e^{-\iota \omega t}$. Due to our geometrical setup, we further reduce the six differential equations for the two times three field components of Eq. (7.1) and by assuming s-polarization, i.e., the electric field has only one component perpendicular to the plane of incident or p-polarization, where the magnetic field has one component perpendicular to the plane of incidence we end with the two Helmholtz equations in one variable.

As shown in Fig. 7.6, the sample is decomposed into several layers and the Helmholtz equations are solved by expanding the fields and dielectric functions in a Fourier series. From the continuity conditions the reflected and transmitted complex fields are derived.

We use a commercial program called `reticolo` [23] which is a numerical implementation and improvement of the RCWA method [17, 22, 29, 36, 37, 43].

As we have periodic boundary conditions along the grating, `reticolo` derives the electric and magnetic fields diffracted by the structure according to the Rayleigh expansion. Some of the diffracted orders, which are determined by the number of Fourier harmonics, are evanescent waves. The exact number depends on the energy, angle of incidence and period of the structure. `Reticolo` returns propagating orders in the far field.

---

[5] Besides this limitation, the ellipsometry software is still at a very high standard, including models for anisotropy, depolarization, roughness, . . .
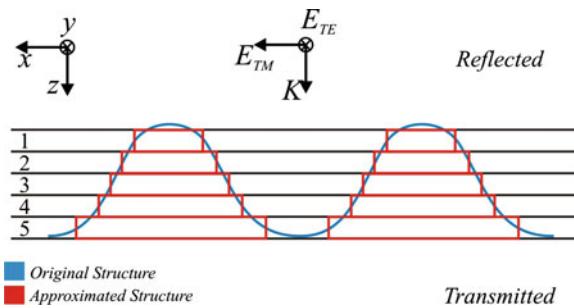
Fig. 7.6 A sinusoidal structure is approximated by layers (staircasing method) in the RCWA calculation. The reflected and transmitted fields, here for normal incidence, are calculated exactly for all diffraction orders

Technically, the RCWA method is a scattering matrix approach comparable to the transmission matrix method. We do not describe the algorithm in detail, here, but only mention that textures like those in Fig. 7.6 have to be defined. For a metallic grating one needs to define five different textures:

- Si substrate (wafer) with infinite thickness (The top and bottom layer extend to ∞ and $-\infty$, respectively, independent on the chosen thickness value in the software. The thickness of the bottom layer defines the origin of the coordinate system perpendicular to the grating),
- SiO$_2$ layer (homogeneous),
- Ti (line) and air (space) layer—the Ti underneath the gold grating with the same linewidth L as the grating,
- Au (line) and air (space) layer—the grating linewidth L,
- air as top layer.

For each layer the user has to provide the optical data of the constituting materials. We use the same material data as given in the Woollam software WVASE32 [55] and material parameters for Si and SiO$_2$ as listed in [20].

For a given polarization (s or p) and angle of incidence the program `reticolo` returns the reflected, scattered and transmitted fields and efficiencies for all propagating diffraction orders. The ellipsometric angles Ψ and Δ are derived, as **no cross polarization occurs** if the lines of the grating are orientated perpendicular to the plane of incidence.

When characterizing a nanostructured sample by SE, the MSE is of limited usefulness. The mean square error represents the difference from measured to fitted data. In the case of nanopatterned structures it is more important to correctly fit main features, such as peak positions, of the experimental spectra. A good reliability factor, or r-factor (as discussed in [41]) should be sensitive to peak positions rather than small variations in the amplitude of the signal. At a later part in this work we exclusively use second derivatives of the Ψ spectrum with respect to the energy to characterize the surface plasmon resonances caused by the metallic gratings.

Finally, we remark that the simulations presented here and below were carefully checked for convergence. This means that the spectrum (and the MSE) does not change by further increasing the number of Fourier coefficients. In the next sections the MSE is used to compare the measured spectra which those computed for different geometrical parameters.

7.4.2 Surface Plasmons at Gratings

At the interface of metals and dielectrics surface bound electromagnetic excitations can be excited due to the sudden change of the permittivity across the interface. This excitation can be localized, e.g., on a metallic particle or propagating along the interface. Localized surface plamsons arise from the electromagnetic boundary conditions at the surface of metallic particles and due to the effective restoring forces on the electrons, which are driven by an external electromagnetic field, resonances can occur. These resonances may rise large field amplifications on the surface and can be excited directly by illumination with light at specific frequencies.

On planar surfaces these states are called surface plasmon polaritons, but are propagating electromagnetic waves confined in only one dimension (z-direction). The dispersion relation describing their propagation along the surface (x, y) is derived from the continuity conditions from Maxwell's equations. The dispersion relation provides a resonance condition for a certain frequency $\omega(k_\parallel \rightarrow \infty)$ which is called the surface plasmon frequency. This resonance is the more pronounced the smaller the intrinsic damping of metal, as an increasing imaginary part reduces the divergence of the dispersion relation. In the long-wavelength regime ($k \rightarrow 0$) the dispersion relation is linear. A special property of the plasmon dispersion curve is that it always lies below the light line, i.e., $c' < c$, where c is the speed of light within the dielectric medium (see e.g. [31, 44, 59]). This implies that the surface excitation is decaying exponentially in the direction normal to the surface as $k_\perp$ must be imaginary. On the other hand surface plasmons cannot be coupled in or out from a metallic surface in the form of propagating electromagnetic waves, due to the mismatch of the in-plane wavevector $k_\parallel$. This mismatch can be overcome in several ways: The simplest one is roughness of the surface but this coupling cannot be controlled and is random. Surface plasmons can be excited in a more controllable way by attenuated total reflection (ATR) [28, 31, 39, 44] which means that the additional wavevector is provided by the material of the prism.

In this work we use metallic gratings as periodic structure with period P provide a fundamental reciprocal (in-plane) lattice vector

$$k_\parallel = k_P = \frac{2\pi m}{P} \tag{7.2}$$

and multiples $m k_P$ where $m \in \mathbb{Z}$ is a positive or negative integer number. The reciprocal lattice vectors match the difference between surface plasmon dispersion
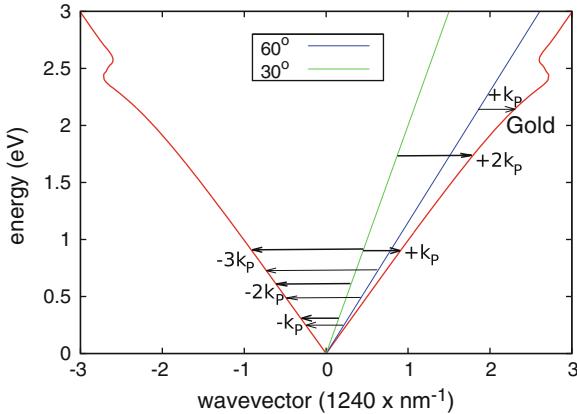
Fig. 7.7 Dispersion relation of surface plasmons at the gold/air interface (*red*). For the gold dielectric function, experimental data from [40] were taken. The *blue* and *green straight lines* indicate light lines for an incidence under $\phi = 60°$ and incidence under $\phi = 30°$. The plasmonic dispersion curve is also drawn for negative wavevectors indicating plasmons travelling in the opposite direction as the incident light

relation and the light line at specific energies $E_m = \hbar \omega_m$. These frequencies depend on the surface projection of the light line, i.e., the angle of incidence ϕ:

$$\omega(\phi) = \frac{ck_\phi}{\sin \phi} \tag{7.3}$$

In Fig. 7.7 we plot the dispersion of surface plasmons and the light line (in green for an angle of incidence $\phi = 30°$ and in blue for $\phi = 60°$). The resonance condition for ω_m thus becomes[6]

$$k_{sp} = k_\phi + mk_P = \frac{\omega_m}{c} \sin \phi + \frac{2\pi m}{P} \tag{7.4}$$

To solve this equation for ω_m we have to consider the surface plasmon dispersion relation, i.e., the parallel wavevector component k_{sp} as a function of frequency ω. This relation is given by the dispersion of surface plasmons on a flat metallic interface [44]

$$k_{sp} = \pm \frac{\omega}{c} \sqrt{\frac{\varepsilon_1(\omega)\varepsilon_2(\omega)}{\varepsilon_1(\omega) + \varepsilon_2(\omega)}} \tag{7.5}$$

where $\varepsilon_1(\omega)$ and $\varepsilon_2(\omega)$ are the dielectric functions at the interface. In the present case (in Fig. 7.7) we have used the dielectric function for gold $\varepsilon_1(\omega) = \varepsilon_{Au}(\omega)$ from literature [40] and a constant $\varepsilon_2(\omega) = 1$ for air.

---
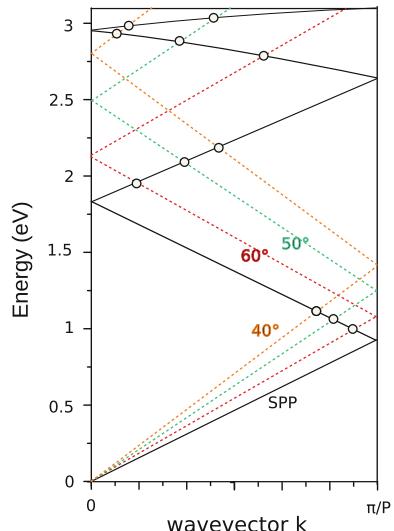
[6] The ideas of exciting surface plasmons by negative lattice vectors were developed by *Peter Zeppenfeld*.

The arrows in Fig. 7.7 represent the reciprocal lattice vectors mk_P according to Eq. (7.2). Both positive and negative values of m can bridge the gap between surface plasmon and light line dispersion, indicating that the incident light may excite surface plasmons with $\pm k_{sp}$, i.e., propagation in opposite directions. Likewise multiples of the fundamental lattice vector give rise to more than just one resonance at each value of ϕ. One notices that shorter fundamental wavevectors k_P (or larger periodicities P) provide a larger set of resonances. The bold arrows in Fig. 7.7 connect the surface plasmon dispersion curve and the light line for incidence under an angle of $\phi = 30°$ to the normal of the grating surface and the thin arrows with the light line for an angle of incidence $\phi = 60°$. Surface plasmonic excitation including negative m values show a red shift with increasing ϕ, whereas excitations with a positive wavevector mk_P yield a blueshift. For normal incidence ($\phi = 0°$) the solution for the positive and negative reciprocal lattice vectors become degenerate ($k_{sp} = \pm 2\pi/m$ according to Eq. (7.4)).

The coupling of surface plasmons and incident light can be regarded in a slightly different view as well. For a highly periodic system it is sufficient to consider the dispersion in the first Brillouin zone, only. Due to the lateral symmetry, the dispersion curve is backfolded at the zone edges and forms a band structure consisting of multiple branches. This causes crossings of the back folded branches with the straight dispersion curve of the incident light beam and an excitation of surface plasmons becomes possible: wherever the wavevector of the light line matches one of the branches of the backfolded surface plasmon dispersion. In Fig. 7.8 this alternative view is depicted.

From Eq. (7.4) we see that for a fixed period of the grating and a fixed angle of incidence we can excite surface plasmons at several energies. For special values of the grating period and angle of incidence ϕ, two different, counter propagating

Fig. 7.8 The dispersion relation of surface plasmons is backfolded in the first Brillouin-zone due to the periodicity of the grating in the in-plane direction. The backfolded dispersion crosses the light line, indicating a possible excitation of surface plasmons at certain resonance energies for various angles of incidence

branches of the surface plasmon dispersion can have the same excitation energy (e.g., at $E \sim 1\,\mathrm{eV}$ the $-3k_P$ and $+k_P$ wavevector in Fig. 7.7). These two counter propagating and degenerate surface plamsons will create a standing wave, i.e., a localized, non-dispersive state. In other words, the two plasmon dispersion branches do not cross and may hybridize for an existing perturbation and an energy gap may be formed [21, 44]. We will investigate this anticrossing behavior in experiments and simulations in Sect. 7.5.3.3.

As we investigate metallic gratings by SE we now want to address the question **how these resonance transitions show up in the ellipsometric spectra**. Surface plasmons can only be excited with light where the magnetic field vector component is parallel to the surface of the metal which corresponds to p-(or TM)polarization. We expect that r_p will decrease if we excite a surface plasmon and therefore we should observe a dip in the SE spectrum for Ψ according to the definition of $\tan \Psi = |r_p|/|r_s|$.

Surface plasmons are collective oscillations of the electrons in a metal. By the incident light we obtain a resonant scattering phenomena with two scattering channels (amplitudes). One is due to scattering within a continuum of states, namely the simple reflection of light from the sample surface, which shows a smooth variation of the scattering amplitude with the photon frequency ω. The second one is due to an excitation of a discrete state (the resonant excitation of a surface plasmon) at a well defined resonance frequency which lies within the continuum of the so-called "background states". In this constellation, the interference between these two scattering amplitudes gives rise to an effective cross section (reflectance) with an asymmetric lineshape, the so-called Fano resonance lineshape. This phenomena was also found for the optical response of photonic crystals [18] and was originally derived by Fano [13].

In the next sections we present measurements and simulations of various gratings. We show that besides the agreement between simulation and experiment, we can clearly identify surface plasmonic excitations at specific frequencies in the spectrum. The transition matrix element and hence the spectral lineshape of this excitation is, indeed, well described by Fano's formula which has the form of
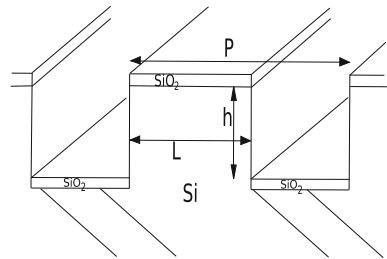
$$R_P \sim \frac{(q\gamma/2 + \omega - \omega_r)^2}{(\omega - \omega_r)^2 + \gamma^2/4} \tag{7.6}$$

where ω_r is the frequency of the resonance, γ is a damping parameter and q is the Fano parameter which measures the ratio between the resonant and the direct (or background) scattering amplitudes.

7.5 Analysis of Gratings

In the previous sections we have introduced the fabrication procedure, cross correlation measurements, SE and numerical simulation techniques for gratings. Additionally, we have shown that metallic gratings support the excitation of surface plasmons

Fig. 7.9 A schematic of the silicon grating fabricated by NIL

at more than one energy. In this part we analyze silicon, gold and silver gratings by spectroscopic ellipsometry, simulate their spectra by rigorous coupled wave analysis (RCWA) and focus on details of the obtained results.

7.5.1 Silicon Gratings

The simplest gratings are produced by structuring the plane wafer without evaporation of any additional material. To this end, we apply the fabrication procedure as illustrated in Fig. 7.1 directly to a silicon wafer. After structuring the residual resist with the stamp we perform a reactive ion etching to obtain the silicon grating. Finally we remove the resist from the sample. A schematic drawing of the resulting pattern is presented in Fig. 7.9.

Before studying the SE spectra of the gratings we take a closer look on the dielectric function of silicon, which we take from [20] and the according Ψ and Δ spectra of a natural wafer with a thin overlayer of SiO_2 in Fig. 7.10. The dielectric function of silicon is governed by two resonances in the visible at 3.4 and 4.25 eV, respectively. These resonances are excitations between different bands at the so-called critical points where an occupied and an unoccupied band are almost parallel in reciprocal
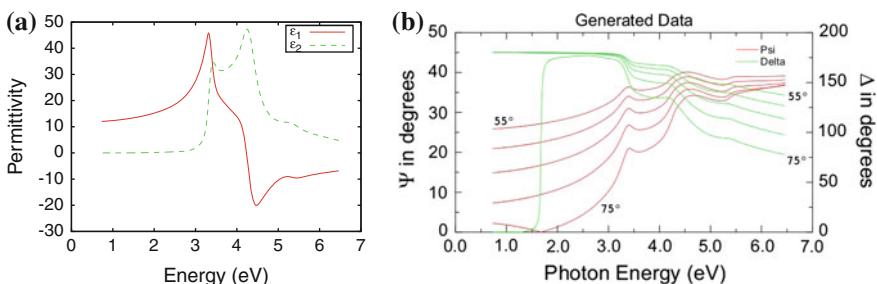
Fig. 7.10 a Real (*red*) and imaginary (*green*) part of the dielectric function of silicon. **b** Calculated SE spectra (Ψ and Δ) of a silicon wafer without oxide for several angles of incidence ϕ between 55° and 75°
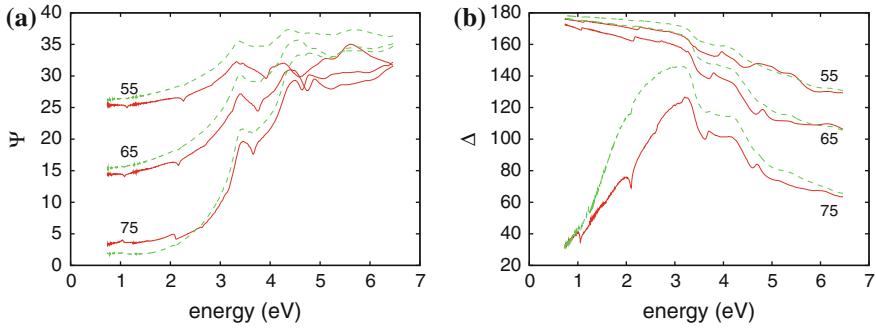
Fig. 7.11 SE spectra (Ψ (**a**) and Δ (**b**)) for Field (c) ($P = 1200$ nm) for two different sample sets with a linewidth $L \sim 720$ nm: in *red*, the height of the grating is $h = 35$ nm and in *green* $h = 20$ nm
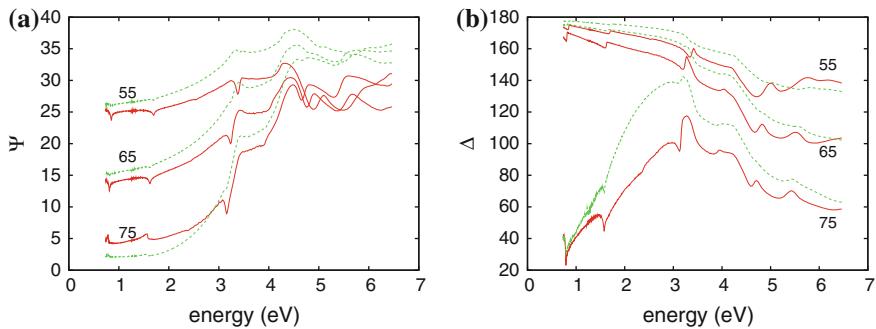

Fig. 7.12 SE spectra (Ψ (**a**) and Δ (**b**)) for Field (d) ($P = 800$ nm) for two different sample sets with a linewidth $L \sim 460$ nm: in *red*, the height of the grating is $h = 35$ nm and in *green* $h = 20$ nm

space [58]. The critical points of Si are also recognized in the SE spectrum (b) but not as sharp peaks. To characterize critical points from SE spectra several numerical manipulations of the SE data, such as Fourier transformation and analysis of the coefficients, have to be performed. This is known as critical point analysis ([57] and references therein). The dielectric function of silicon is positive below $E = 4.15$ eV and negative above.

Due to the small fields of 3×3 mm^2 we have to use the focussing probes of the SE [55]. In Fig. 7.11 we present the ellipsometric spectra for *Field* (*c*) with a period of $P = 1200$ nm and in Fig. 7.12 for *Field* (*d*) for $P = 800$ nm (see Fig. 7.2a). We clearly distinguish peaks due to the grating at certain energies. The two different curves in each plot represent measurements of two sample sets which only differ in height and linewidth. The difference in height is simply realized by a longer etching time which also causes a reduced linewidth.

We have measured the ellipsometric data for three different angles of incidence, namely for $\phi = 55°$, $65°$ and $75°$ as for silicon Ψ and Δ are more sensitive for large angles of incidence. The deeper grating shows enhanced features in the spectrum.
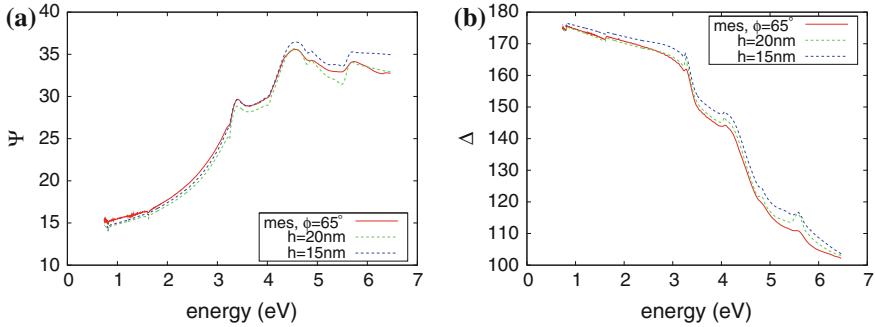
Fig. 7.13 Measured (*red curve*) and simulated (*blue and green*) curve for a silicon grating of Field (d) for Ψ (**a**) and Δ (**b**). We compare two different simulations with the experiment to deduce the real height of the grating

Comparing Figs. 7.11 and 7.12 with Fig. 7.10b, we see that the overall shape of Ψ and Δ is characteristic for silicon. Two critical points of silicon are also visible in the spectrum of the gratings but are slightly reduced in amplitude due to the natural oxide on the surface.

The peaks/dips, especially for the measurements of the deeper grating, are explained by surface plasmon excitations. Although Si is not metallic in this energy range we are allowed to insert the dielectric function of silicon for $\varepsilon_2(\omega)$ in Eq. (7.5). The measured resonance frequencies indeed match with the ones calculated from Eq. (7.4) using the periodicities of $P = 1200$ and $P = 800$ nm of the real samples.

In Fig. 7.13 we compare the simulation results with the measured spectra. In Sect. 7.3.1 we already saw from AFM and SEM images that the linewidth and height may slightly vary around a mean value. Only the period is constant as it is given by the stamp design and repeated all over the sample. In the simulations, geometrical parameters of the grating have been varied. In this way we are able to adjust the simulated spectrum to the measured one and derive the best fit parameters for the geometry from the best agreement between simulation and experiment. In the present example we have verified the period $P = 800$ nm (*Field (d)*) and concluded that the lineshape is almost rectangular with a mean value of $L = 480$ nm which is in agreement with the SEM images.[7]

The only unknown geometrical parameter is the height of the grating h. The green line in Fig. 7.13 shows the SE spectra of a grating with $h = 20$ nm and the blue one the spectrum for $h = 15$ nm. From the Ψ spectra we see that both simulations do not exactly match with the measured spectrum (red curve) but the simulation for $h = 20$ nm shows the same qualitative features as the measured one and only has a vertical offset of around $2°$ in Ψ.

The quantitative analysis yields an MSE $= 35$ for the experiment and simulation with $h = 20$ nm and an MSE $= 54$ for the experiment compared with calculations

---

[7] Similar to the one presented in Fig. 7.3a but for a period of $P = 800$ nm.

assuming $h = 15$ nm. We thus conclude that the grating height must be around $h = 20$ nm which was confirmed by AFM measurements (see Fig. 7.4).

We have performed various simulations concerning the influence of the silicon surface oxide layer. First, we investigate whether a thin SiO_2 layer on the side walls of the grating influences the spectra. Second, we changed the thickness of the top oxide layer (1, 2 or 3 nm). All these variations had no significant effect on the SE spectra and we conclude that we can neglect the oxide on the sidewalls and a rough estimate of the oxide thickness is enough.

7.5.2 Gold Gratings

In science there is a growing interest in metallic and especially gold structures. The search for negative index materials whose resonant frequency band is in the near infrared or even in the visible spectral range is a driving force for this interest. The aim is to create "magnetic resonances", which are needed to tune the effective permeability to negative values. For this reason split rings, rods or fishnet structures containing metallic elements are investigated (Ref. [50] and references therein).
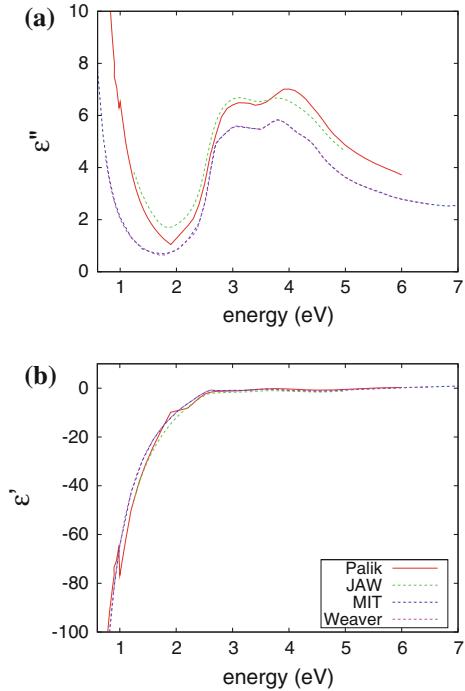
These metallic nanostructures have a complicated geometry and we rather describe and analyze the ellipsometric response of one dimensional gold gratings in this section as a simple model case. Before doing this, we present and discuss the various dielectric functions for gold found in literature and measured by ourselves. Next, we show the modelling of the grating and compare the results to actual SE measurements. Finally, we again study surface plasmonic excitations which are more pronounced then in Si gratings as we have metal-dielectric interfaces enhancing surface plasmonic effects.

7.5.2.1 Bulk Gold

In literature several approaches have been applied to measure the dielectric function of gold. As the data sets differ considerably, a summary and qualitative analysis of the optical properties of gold was done in [4]. It was shown that the spectrum of gold consists of two different parts around a band gap at $E = 2.5$ eV. Below this value the optical response is Drude like due to the free electron contribution and the imaginary part $\varepsilon''(\omega)$ of the dielectric function is influenced by grain size which affects the damping constant (relaxation time) as well as by roughness. The lowest values for ε'' were achieved for non-annealed samples. In the interband region for energies $E > 2.5$ eV a different effect influences the dielectric function, namely the amount of voids which are described by an effective medium approximation (Bruggeman). This is why the lowest values of ε'' in this spectral range were achieved for annealed samples as the amount of voids was reduced.

In Fig. 7.14 we present data sets which are included in the ellipsometry software [55] and found in literature. The first two are the permittivity data sets of

Fig. 7.14 Different dielectric functions of gold from literature: $\varepsilon' = \mathrm{Re}(\varepsilon)$ and $\varepsilon'' = \mathrm{Im}(\varepsilon)$ (from [40]) (*red curves*), ε' and ε'' from Woollam software (*green curve*), dielectric functions from [24] (*blue*) and from [53] (*magenta*)

Palik [40] (red curve) and of the Woollam company (green curve). The latter is only available in an energy range from $E \in [1.2, 5]\,\mathrm{eV}$ and has to be extrapolated for smaller energies. The next one from literature is found in [53] and plotted in magenta. These three data sets are obtained from actual measurements of gold.

The other data set which we present in blue are obtained by fitting gold data to a model dielectric function [24] (we label this as MIT-data set). All metals can be approximated by a Drude-Lorentz model which consists, as the name indicates, of a Drude and a Lorentz part

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega(\omega + \iota\gamma_p)} + \sum_i \frac{A_i}{\omega_{r,i}^2 - \omega^2 - \iota\gamma_i\omega} \tag{7.7}$$

The Drude part describes the free electron contribution with a plasma frequency ω_p and damping coefficient γ_p and of several Lorentzian like parts with resonance frequencies $\omega_{r,i}$, oscillator strengths A_i and damping factors γ_i. The Lorentzian oscillator (or possible Tauc-Lorentz oscillators for semiconductor or amorphous materials) represent interband transitions [1, 53]. Therefore, the total dielectric function consists of a low energy part where interband transitions or free electron motion is mainly and a high energy part with dominating intraband contributions.
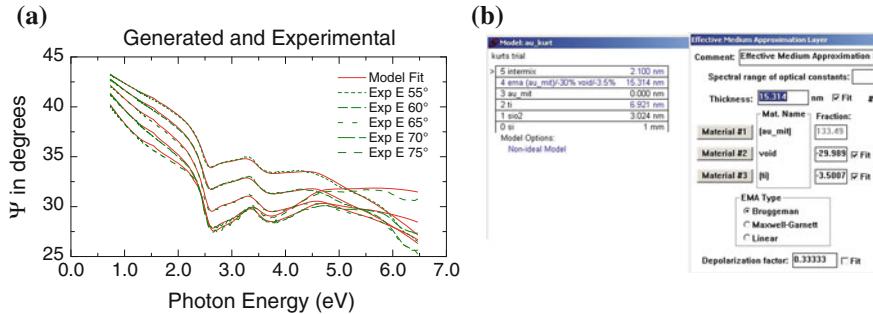
Fig. 7.15 SE data (*green*) for a control sample consisting of a Si wafer with natural oxide. Above a thin Ti layer was deposited for the adhesion of the gold layer on top. In *red* we show the model fit according to the parameters plotted in the *right panel*. The negative fraction indicates a more dense material than the original data set

Within this model we can try to fit measured data of gold: As described before, intraband transitions are dominating below energies of $E \sim 2.5\,eV$ which can be fitted by the classical free electron model (Drude). Grain sizes influence the mean free path of the electrons and thereby γ_p. Above this energy, interband transitions occur, which are modelled by (Tauc)-Lorentz models [1, 4, 58]. In Fig. 7.14 we see that the theoretical values in blue and the compiled data from literature of [53] (in magenta) match almost perfectly. Unfortunately no information about grain sizes and surface roughness was given in the corresponding publications.

As described previously, we have used a control wafer during the evaporation process to determine the height of the gold layer. The measured SE data (green spectra) are presented in Fig. 7.15a for five different angles of incidence from $\phi = 55°$ to 75°.

We have tried to fit the measured data by constructing a multilayer model consisting of homogeneous Si/SiO$_2$, Ti (we use the dielectric function of polycrystalline titanium for the evaporated film) and Au layers using the known thickness of the oxide (a measured before evaporation). This ansatz yields an MSE of around 100. Also the use of different gold models in the fitting process does not reduce the error significantly.

The best fitting result for the control wafer was achieved by creating a complicated model which is presented in Fig. 7.15b. We replace the gold layer by an effective medium. This effective medium is derived by Bruggemans formula and consists of three different materials (MIT gold, voids and titanium). Several parameters as filling fractions, thickness of this EMA and the thickness of the titanium were assumed as free fit parameters (indicated in blue in Fig. 7.15). To reduce the MSE we have added a "intermix" surface which consists of a 50 % mixture between air and the underlying Au layer. This "intermix" is mainly used to mimic the surface roughness. With this model, the MSE was reduced to ~16 which seems perfect at first glance.

Although we have reduced the MSE significantly with this model, we know that the result is unphysical. During the evaporation process we made a rough guess that
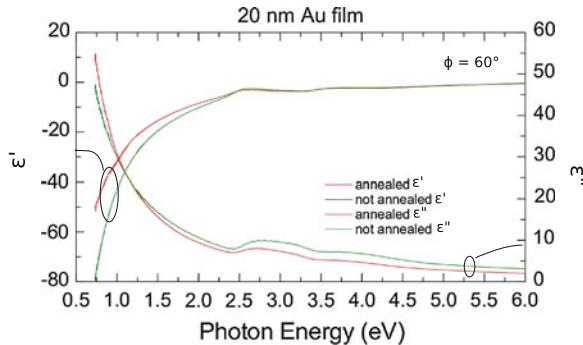
Fig. 7.16 Pseudo dielectric function $\langle\varepsilon\rangle$ calculated from ellipsometric spectra before and after annealing (in UHV at $350\,°C$ for $4\,h$) of a thin gold film of $20\,nm$ on a natural Si/SiO$_2$ wafer. The angle of incidence was $65°$

not more than $5\,nm$ of titanium were evaporated on the sample (the fit yielded $7\,nm$ plus some volume fraction in the gold layer). Also the thickness of the gold layer was overestimated. Therefore, we conclude that the real result must be somewhere in between: the metals segregate and form intermixed layers whose contribution could only be determined exactly by a FIB-REM cut (see Fig. 7.5) and X-ray photoelectron spectroscopy (XPS). Assuming that the Ti layer is diffused completely and thus neglecting Ti in the fitting procedure, we could reach a reasonable MSE with a thickness of the Au layer of $\sim$50 nm which is more consistent with the expectation based on the evaporation settings. This shows that the lowest values of the MSE does not guarantee correct results as the precise sample composition and other constraints have to be considered. Without an incomplete model structure, the MSE is a rather useless quantity.

To further investigate the gold layer we have annealed the sample after evaporation. The annealing was done under ultra high vacuum at $350\,°C$ for $4\,h$. In Fig. 7.16 we present the pseudo dielectric function for the annealed and non annealed sample, calculated from the ellipsometric spectra using the formula for the pseudo dielectric function (which depends on ρ and angle of incidence). The measurements were done on a thin film ($20\,nm$) of gold on a natural Si wafer. We see that the real part of the dielectric function has increased after annealing. The imaginary part is enhanced for energies $E < 1\,eV$ and reduced above. According to [4] this would suggest a reduction of the voids and an increase of the grain size, which was indeed confirmed by AFM measurements.[8]

Although annealing improves the material properties of gold [8], we focus in the next sections on non-annealed samples. First of all, the fit to the SE data of thin gold films is better for non annealed samples, also yielding a lower MSE as well. Furthermore, annealing of grating structures changes its profile and the lines become

---

[8] The annealing procedure and AFM investigations were performed by Maria Losurdo at the IMIP, Bari, Italy.

rounded. This effect cannot be avoided as annealing before the lift off destroys the resist.

Therefore, in all RCWA simulations of the gratings below we will model the gold lines by an effective medium of MIT bulk gold [24] with $\sim -11\,\%$ voids indicating a rather dense gold film (this is the gold model which yielded SE spectra that were closest to the measured ones without assuming any surface roughness and mixtures with other materials than air). As stated above, an EMA of gold and voids is appropriate for $E > 2.5\,eV$ but due to the decent agreement also for lower energies and for reasons of simplicity we use the same EMA mixture throughout the whole energy range.

7.5.2.2 Modelling of Gratings

In the previous sections we have described the fabrication of gold gratings and their structure and have discussed the overall shape and certain features appearing in the ellipsometric spectra. Now, we focus on the comparison of measurements of gold gratings with RCWA simulations and highlight the sensitivity of the SE data with respect to changes in the geometrical parameters. We study in detail a variation of the linewidth L, period P and the height h of the grating as well as the thickness of the homogeneous layers underneath. As mentioned before, the first two/three parameters are rather well known from the cross correlation techniques SEM/AFM. Thicknesses or possible diffusion effects could only be investigated by TEM images of cross-sections (and would therefore destroy the sample).

In Fig. 7.17 we present several simulations and compare them to SE measurements of *Field* (*d*) which has a period $P = 800\,nm$ and a height $h \sim 50\,nm$. This comparison is done for the whole spectrum but for a single angle of incidence
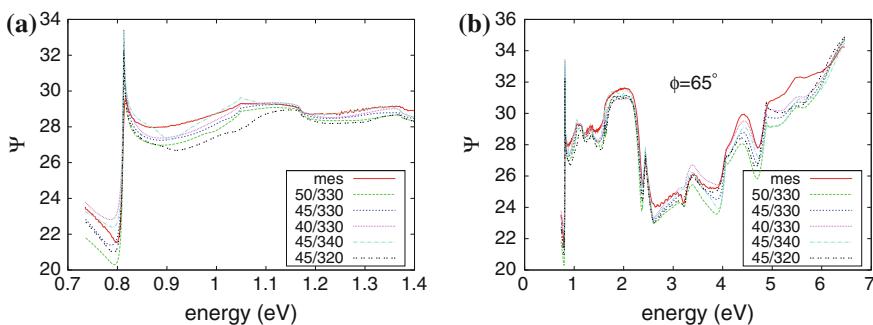

Fig. 7.17 SE spectra of a grating of field (d) ($P = 800\,nm$) with $h \sim 50\,nm$ (*red curve*) in comparison with different simulations (h/L): 50/330 in *green*, 45/330 in *blue*, 40/330 in *magenta*, 45/340 in *turquoise*, 45/320 in *black*. In **a** the low energy region of the spectrum is shown to highlight the sensitivity of the simulation with respect to the variation of geometrical parameters. In **b** the whole spectrum is displayed

Table 7.2 MSE for different simulations of linewidth and period

Geometry (h/L)	50/330	45/330	40/330	45/340	45/320	50/bulk
MSE	56	65	103	47	86	44

($\phi = 65°$). We have determined the thickness of the SiO_2 to be 2.8 nm and assumed a homogeneous layer of 2 nm of titanium consistent with the SE data recorded on the homogeneous reference sample. The other geometrical parameters as h, P and L were measured by SEM and AFM. In the simulations we vary them to see the effect in Ψ and Δ. The numbers shown in the legend of Fig. 7.17a indicate the height h and the linewidth L of the grating (numbers in nm) used in the simulations. Generally speaking, the line shape is in good agreement with the measured spectrum for the first three parameter sets in green (50/330), blue (45/330) and magenta (40/300) but deviates, especially in the range of $E \in [0.75-1.4]$ eV (see Fig. 7.17a), for the last two simulations (turquoise (45/340) and black (45/320) dotted curves). Obviously, changing the linewidth L results in different shapes of the spectra, even for a variation of 10 nm in lateral direction, only. This highlights the sensitivity of the simulations with respect to the linewidth.

The green, blue and pink dotted curves illustrate the variation of the height of the gold grating. According to the simulation, h should be closer to 40 nm than to 50 nm—a result not completely consistent with the AFM measurements ($h = 50$ nm). The origin of this discrepancy is manifold: The gold lines have a rough surface and residual resist was present after stripping (see Fig. 7.4b). Furthermore, geometrical details like the somewhat trapezoidal shape of the gold lines were neglected in the simulations.

The calculation of the MSE for gold gratings is not very useful as it yields values of $MSE \sim 2000$ for the whole spectrum. The deviation between the simulations and the experiment is particularly large for high energies. If we restrict the MSE to an energy range $E < 3.1$ eV we obtain the results presented in Table 7.2. For comparison, in the last column the MSE for a homogeneous, reference sample with $h = 50$ nm is also shown, which is the lower limit for the MSE. Concluding, in this energy range, the simulations of the gold gratings match the experimental spectrum rather well as the best MSE is just slightly larger than the error for an unstructured gold layer.

We see that the MSE is best for a simulation with a wrong linewidth. Apart from that, we get a better MSE for a height of $h = 50$ nm than for $h = 45$ nm. Therefore, we conclude that in this case the MSE is not a reliable quantity for interpreting the simulations. Much more important is the fact that the simulation correctly reproduces the sharp plasmon related features in the spectra, rather than the smooth regions in between.

We have shown the sensitivity of the simulations to variations in h and L. We now focus on the variation of the period for a gold grating of *Field* (a) on the same sample set ($h = 50$ nm). To this end, we have fixed the linewidth to $L = 760$ nm in the simulations. In Fig. 7.18 the results are presented. The periodicity was assumed
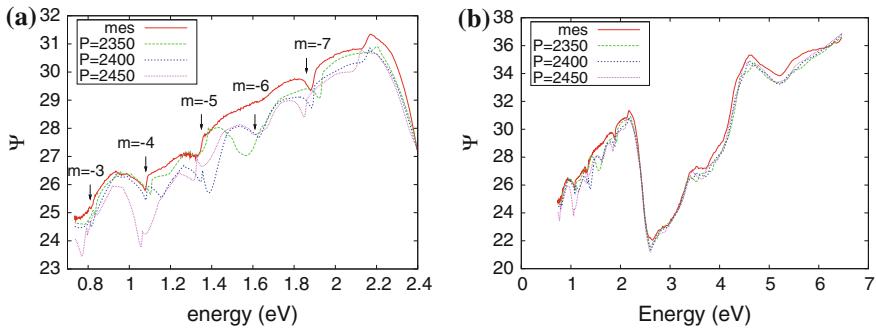
Fig. 7.18 Simulation of field **a** of a gold grating with $h = 50$ nm. In *green*, *blue* and *magenta*, different simulation results are displayed for $P = 2350$, 2400 and 2450 nm. The *left plot* (**a**) shows the low energy part of the spectrum; in **b** the whole spectrum is presented

to be $P = 2350$ nm (green), $P = 2400$ nm (blue) and $P = 2450$ nm (magenta). In the left figure (a) where we show the low energy part of the spectrum, the three simulated spectra are compared to the measurement (red curve). We can clearly see that the dips at ∼1.1, 1.35 and 1.9 eV strongly shift with the period as predicted by Eq. (7.4). Therefore, the periodicity of the grating can be precisely determined from SE and the corresponding simulations with a very high accuracy.

In Fig. 7.18b the whole spectrum is displayed. Apart from a shift of the dip position the simulated spectra can reproduce the overall shape of the measured curves. This indicates that the height h and the linewidth L closely match the values of the real sample (in the simulation we have used $h = 45$ nm and $L = 760$ nm which was in good agreement with the SEM images).

7.5.2.3 SE and Surface Plasmonic Response

As a first example we have investigated in Sect. 7.5.1 a silicon grating which only consists of silicon and its oxide. In this case, the surface plasmonic effects are not well pronounced due to the semiconducting properties the material.

Due to its metallic behavior we expect much stronger surface plasmonic resonances on a gold grating. As the imaginary part of the dielectric function of gold is much smaller than that of silicon for energies above $E > 3.2$ eV, the resonance peaks are expected to be sharper.

In Table 7.3 we have listed the resonance energies as calculated from Eq. (7.4) for two different periods. For grating *Field* (*a*) with $P = 2.4 \, \mu$m we have used an angle of incidence of $\phi = 65°$ and for grating *Field* (*d*) with $P = 800$ nm, the energies were derived for $\phi = 55°$, 65° and 75°. The subscript of E corresponds to the integer number (resonance order) m in Eq. (7.4).

In Fig. 7.19 we present the measured and simulated spectra of a gold grating with $P = 2400$ nm. The resonance frequencies from Table 7.3 (fourth column) are

Table 7.3 Resonance energies for an angle of incidence of $\phi = 55°, 65°$ and $75°$ for *Field* (d) ($P = 800$ nm) and *Field* (a) (with $P = 2400$ nm) and for $\phi = 65°$

Field (d) 55° (eV)	Field (d) 65° (eV)	Field (d) 75° (eV)	Field (a) 65° (eV)
$E_{-1} = 0.85$	$E_{-1} = 0.81$	$E_{-1} = 0.78$	$E_{-3} = 0.81$
$E_{-2} = 1.68$	$E_{-2} = 1.61$	$E_{-2} = 1.56$	$E_{-4} = 1.08$
$E_{-3} = 2.47$	$E_{-3} = 2.34$	$E_{-3} = 2.28$	$E_{-5} = 1.35$
$E_{-4} = 3.40$	$E_{-4} = 3.24$	$E_{-4} = 3.14$	$E_{-6} = 1.61$
$E_{-5} = 4.24$	$E_{-5} = 4.05$	$E_{-5} = 3.93$	$E_{-7} = 1.86$

The subscript of E denotes the order of the resonance m
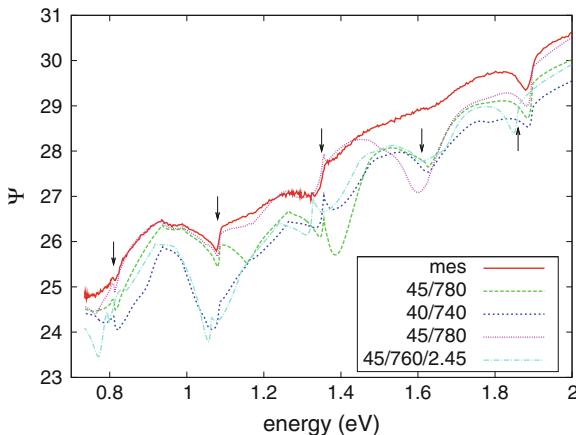
Fig. 7.19 Measurement and simulations of Field (a): The surface plasmonic excitation energies only shift for a change in the periodicity as we have fixed $\phi = 65°$. The *arrows* indicate the excitation energies from Table 7.3. The height and width of the gold lines were assumed 45/780 (*green*), 40/740 (*blue*), 45/780 (*magenta*) and 45/760 with $P = 2450$ nm (*turquoise*)

marked by arrows. Similar to the previous section the numbers at the top right in the plot indicate the height and width of the gold lines, respectively. As the periodicity is now three times larger than for the grating shown in Fig. 7.17, the sensitivity of the simulations with respect to a change in line width decreases. We see that all simulations show kinks in the spectrum at these energies except for the turquoise curve. The latter was calculated for a periodicity of $P = 2.45\,\mu$m which shifts the excitation energy of surface plasmons to lower energies $\sim$20 meV. Therefore, the plasmon resonances are clear indicators for the angle of incidence and the periodicity.

In the measured spectrum the low energy resonances are identified but certain plasmonic excitations, e.g., E_6 are hardly visible in the measured spectrum. Furthermore, we see that the simulations sometimes deviate from the experiment in certain energy regions. These differences have their origin, as we mentioned before, in

- the modification of the dielectric functions due to the evaporation process (we use the experimental determined dielectric function of gold for simulations)
- possible statistical variations of the fabricated gratings
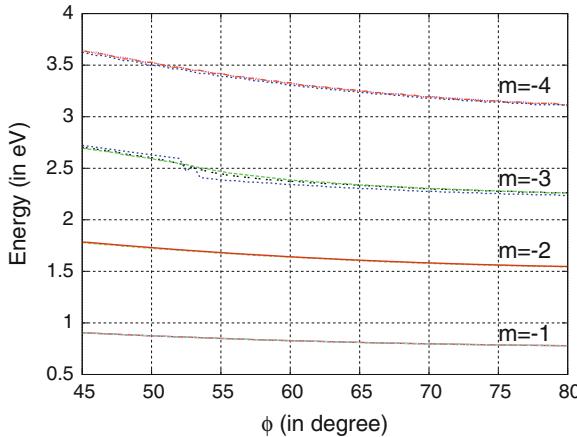
Fig. 7.20 Resonance energies for a gold grating with a period of 800 nm as a function of the angle of incidence for three different gold dielectric functions (Palik, Woollam, MIT)

For a better illustration of the first point, we have plotted the surface plasmonic resonance energies over the angle of incidence for a grating with $P = 800$ nm in Fig. 7.20. For the simulations have used the three different gold dielectric functions which were introduced in Sect. 7.5.2.1. We see that the choice of the dielectric function has only a minor influence on the resonance energies except for the $m = -3$ excitation around $E \sim 2.5$ eV, where the data of Palik exhibit a pronounced kink (see Fig. 7.14).

Concerning the second point, we note that the simulations are based on a structure that consists of perfectly periodic, sharp lines of gold, where the height of the gold lines does not vary and the grating has no surface roughness. This cannot be realized when fabricating the real grating samples. However, we can conclude from these results that the better the agreement between experiment and simulation is, the better the sample quality.

In Fig. 7.21 we present the phase spectrum $\Delta(E)$ for *Field* (*d*) for three different angles of incidence. We plot the spectrum for the lower energy range, as in the Δ spectrum the peaks of the surface plasmons can be easier identified. The arrows again indicate expected resonance energies as listed in Table 7.3 which all correspond quite well to the experimental observations.

To demonstrate the sensitivity of the surface plasmonic resonances we focus on the first peak ($m = -1$) in Fig. 7.21 which is strongly enhanced compared to the resonances at high energies. The surface plasmons, that are excited for $m < 0$ show a blueshift with decreasing angle of incidence which is also seen from Fig. 7.20. This shift becomes stronger the smaller ϕ gets. However, as we have to use focussing probes, measurements below $\phi = 45°$ are prohibited. Between $\phi = 45°$ and $\phi = 55°$ the peaks, which are best seen in the Δ spectrum, shift by about 6 meV per degree of ϕ. As the detector is sensitive to 2 meV in this energy region, we reach a sensitivity of one third to one fourth of a degree.
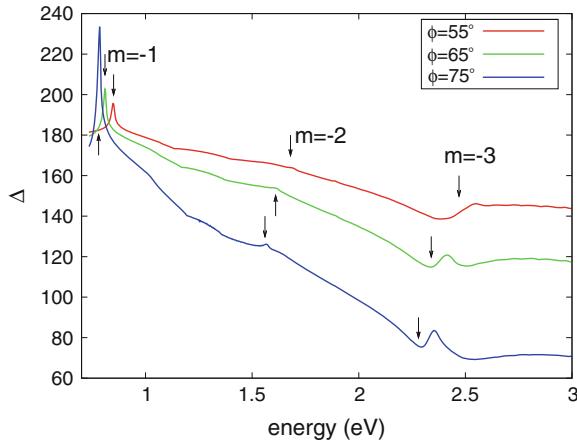
Fig. 7.21 Δ-Spectrum of a grating with $P = 800$ nm for $\phi = 55°$ and $75°$. The *arrows* indicate the resonance energies of the surface plasmonic excitation
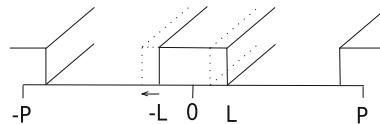

Fig. 7.22 Supercell calculations where the central gold bar is shifted to the left in order to create a new period doubling the one of the original grating

7.5.2.4 Supercell Calculations and Effects of Linewidth

We have shown that the ellipsometric spectrum is affected by changing the geometrical parameters of the grating. Now, we focus on large-scale variations of the period P and line width L. As the stamp has a fixed design we exclusively perform simulations for these two parameters.

The variations are interesting because we demonstrate that dips and peaks in the SE spectra are really caused by surface plasmonic excitations. As a first hint, the effect is polarization dependent. We have highlighted the sharp features in the spectra of Ψ and Δ in the previous section. The latter one represents the phase shift between the two different polarizations (p and s, respectively) (Fig. 7.22).

To emphasize the relation between periodicity and resonance frequencies we set up a supercell with a periodicity of $P = 1600$ nm. Within this supercell we place an additional gold bar. By putting this exactly in the middle, we create a periodicity of $P = 800$ nm—the red curve in Fig. 7.23 shows the corresponding result. During the other simulations the central bar is shifted in steps of 40 nm out of its central position (blue curve indicates a shift of $\delta = 40$ nm and the pink curve of $\delta = 80$ nm). All gold bars were assumed to have a height of 45 nm and a linewidth $L = 100$ nm. The angle of incidence was $\phi = 65°$.
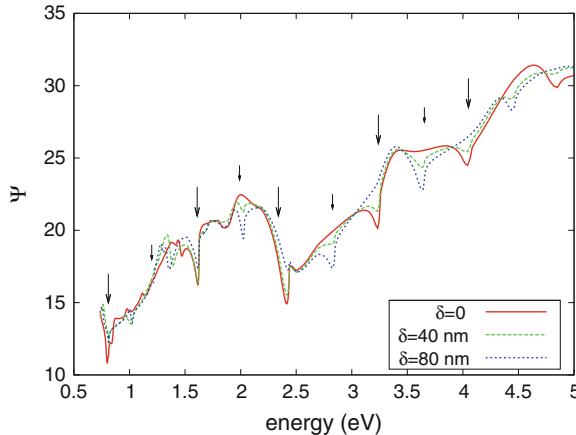
Fig. 7.23 Supercell calculations: The spectrum in *red* is for a periodicity of $P = 800\,\text{nm}$. The other spectra are supercell simulations with an overall periodicity of $P = 1600\,\text{nm}$ and with a gold bar that is displaced from the central position to destroy the periodicity. Therefore, only the large period of $P = 1600\,\text{nm}$ (*blue* and *pink curves*) remains. This doubles the number of surface plasmonic resonances which are indicated by the *larger* ($P = 800\,\text{nm}$) and *shorter arrows* ($P = 1600\,\text{nm}$)

Table 7.4 Additional resonance frequencies (odd values of m) for a periodicity of $P = 1600\,\text{nm}$

Order of resonance	-3rd	-5th	-7th	-9th
Energy	$E = 1.2\,\text{eV}$	$E = 1.99\,\text{eV}$	$E = 2.827\,\text{eV}$	$E = 3.653\,\text{eV}$

The angle of incidence was $\phi = 65°$

The larger arrows in Fig. 7.23 are the plasmon peaks for a period of $P = 800\,\text{nm}$ whereas the shorter arrows indicate the additional resonances occurring for a periodicity of $P = 1600$ nm. We see that the simulations for the offset bar (green and blue) exhibit the additional resonances of the larger periodicity, whereas the simulation for the central bar and, hence, reduced periodicity of $P = 800\,\text{nm}$ shows only resonances at the larger arrows. The additional resonance energies for $P = 1600\,\text{nm}$ are displayed in Table 7.4 for $\phi = 65°$ (and the excitation energies for $P = 800\,\text{nm}$ are listed in Table 7.3).

From these supercell simulations we conclude that the resonances in the SE spectra are very sensitive to the periodicity. Resonances are strictly related to the appropriate reciprocal wavevector $2\pi/P$ and even a small deviation of the gold bars affects the resonance frequencies of the surface plasmons.

Next, we investigate a change in linewidth and its effect on the SE spectrum. We perform simulations for gold gratings with a height of $h = 45\,\text{nm}$ on a homogeneous layer of $SiO_2 = 2\,\text{nm}$ for a period of $P = 800\,\text{nm}$ and an angle of incidence of $\phi = 65°$. For this structure we know the surface plasmon resonance frequencies from Table 7.3 (second column). In the following we have changed the linewidth of the gold bars between 1 and 640 nm and the simulation results are plotted in Fig. 7.24. The vertical lines indicate the expected SP resonance positions.
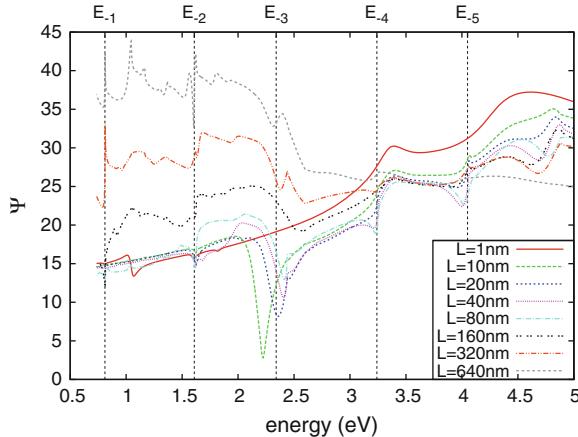
Fig. 7.24 Simulated spectra for gratings with a periodicity of $P = 800\,$nm and a height of $h = 45\,$nm. The linewidth was changed for the different spectra: $L = 1\,$nm (*red*), $L = 10\,$nm (*green*), $L = 20\,$nm (*blue*), $L = 40\,$nm (*magenta*), $L = 80\,$nm (*turquoise*), $L = 160\,$nm (*black*), $L = 320$ (*orange*), $L = 640\,$nm (*gray*)

The simulation for a grating with very small linewidth of $L = 1\,$nm (red spectrum) is similar to the bulk behavior of silicon and the grating has only an effect in the low energy region where a peak at around 1 eV occurs. Already an increase of the linewidth to 10 nm (green curve) yields several surface plasmonic resonances due to the periodicity.

We focus now on the lowest three excitations, namely $E_{-1} = 0.81\,$eV, $E_{-2} = 1.61\,$eV and $E_{-3} = 2.34\,$eV. The resonance for the first lattice vector ($m = -1$) becomes more pronounced with an increasing linewidth. Below $L = 40\,$nm this resonance is hardly visible. Further, with increasing width the resonance becomes sharper, which is caused by vertical surface plasmon resonances as explained in detail in Sect. 7.5.3.3. The resonances E_{-2} and E_{-3} have different strengths and are not visible for certain values of L. This behavior may be a consequence of the Rayleigh-Rice theory (RRT) which states that the reflectance of a structure is determined by the power spectrum of the surface geometry [14, 45]. RRT is only valid for small heights of a structure but as we are changing the linewidth, we are pronouncing different Fourier coefficients of the power spectrum of the grating.

Apart from this, the E_{-3} excitation shows a strong blueshift with increasing linewidth between $L = 10$ to about 40 nm. We see that in the low energy region an increase of L leads to a positive offset of Ψ and negative one in the high energy region. This cross over is caused by the fact that for small line width the signal is dominated by the underlying silicon substrate, whereas for large values of L a gold like spectrum is obtained.

7.5.3 Silver Gratings

In the previous section we have characterized gold gratings and studied the surface plasmon energies for different angles of incidence. Surface plasmons are excitations on metallic surfaces and their excitation/propagation becomes more enhanced the smaller the damping of the metal is. The absorption of silver is in some regions smaller than the one of gold and, therefore, we expect that surface plasmonic excitations are more pronounced.

The main disadvantage of silver regarding technology aspects is its aging effect as (thin) silver films oxidize relatively fast. Therefore, the material properties change after some time of exposing of the silver grating to air. Thus, in literature and applications mainly gold is used for the fabrication of nanostructures [50].

As a consequence, the technology for the fabrication of gold gratings is better developed as for silver structures. However, we have processed a silver sample with the same stamp (and geometric dimensions) as the gold and silicon gratings. This silver sample was not as perfect as the gold samples but some areas on the sample were well processed and, therefore, a proper measurement of the ellipsometric response using the focussing probes was possible.
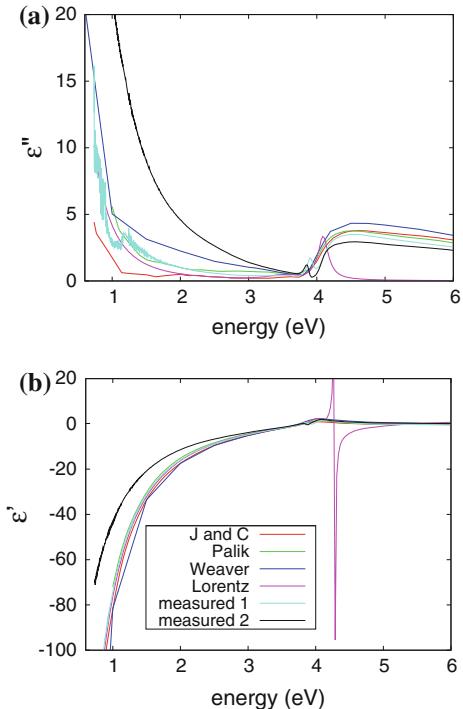
7.5.3.1 Bulk Silver

Similar to gold, many measurements and data for bulk silver exist. In Fig. 7.25 we present several different dielectric functions for the metal. The first one is the well known Johnson Christy data set [26] and the according real and imaginary part of the permittivity are presented in red. The next data set (plotted in green) are Palik's data [40], which have a slightly larger imaginary part in the low energy range. The physics data booklet of Weaver [53] also contains silver data, which are presented by the blue curve. The discretization of the permittivity in the visible spectral range is not sufficient and, therefore, we reject this data set.

The data set plotted in magenta is included in the Woollam software and is a sum of Lorentz oscillators as described by Eq. (7.7). It is only valid for energies $E < 4.1 \, \mathrm{eV}$ as seen by the artificial oscillator for $E > 4.1 \, \mathrm{eV}$. It contains a Drude term with the parameters $\hbar \omega_p = 8.53 \, \mathrm{eV}$ and $\gamma_P = 0$ and four Lorentzians at resonance energies $E_0 = \{0.33, 3.51, 4.09, 4, 28\} \, \mathrm{eV}$ with the amplitudes $A_i = \{2.46, 8.8, 5.4, 0.586\} \, \mathrm{eV}^2$ and damping parameters $\gamma_i = \{0.18, 0.47, 0, 0.5\} \, \mathrm{eV}^{-1}$. This dielectric function yields the best description of the silver films we have fabricated, among all data sets in the energy range $E < 4 \, \mathrm{eV}$. As the surface plasmonic resonances are best visible below this range this limited spectral range is sufficient.

The pseudo dielectric function $\langle \varepsilon \rangle$ obtained from the SE measurement of a silver film of 50 nm thickness is presented in turquoise. We have performed this measurement directly after processing the film. In addition, the black curve shows the pseudo dielectric function for the same sample but measured after six months of exposure to

Fig. 7.25 Different dielectric functions of silver from literature: ε' and ε'' from [26] (*red curve*), from [40] (*green*), from [53] (*blue*), dielectric model function described by four Lorentz oscillators and a Drude term as included in Woollam software (*magenta*). The pseudo dielectric function $\langle\varepsilon\rangle$ derived from measurements of a 50 nm silver film are shown in *turquoise* (directly after evaporation) and *black* (several month after sample processing)

air. The imaginary part as well as the real part of $\varepsilon(\omega)$ have increased considerably, indicating oxidation effects of silver.

We see that the measured sample shows a kink in the dielectric function slightly below 4 eV. This peak is not visible in published data sets and is clearly not an effect of silver. In general, the reason for the difference in the measured values for silver and the tabulated ones from literature is due to the evaporation process, similar as for gold. Grain size, surface roughness and the density of silver, i.e., the amount of voids, influences the quality of the silver film.

In Fig. 7.26 we compare the Drude-Lorentz silver (taken from the Woollam software) data to the gold dielectric function of [53]. We have chosen these two data sets as they describe our samples best. We see that ε' of silver (red curve) is always lower than the real part of gold (green). The imaginary part of gold (magenta line) is smaller than for silver (blue) below $E \sim 1.5$ eV but becomes larger above. Therefore, in our samples the conductivity of gold is better than for silver for low energies, which should not be the case for well prepared silver films. In this case, ε'' for silver is always smaller than gold, which can be seen, e.g., by comparing the gold data of Weaver with Johnson/Christy data for silver.
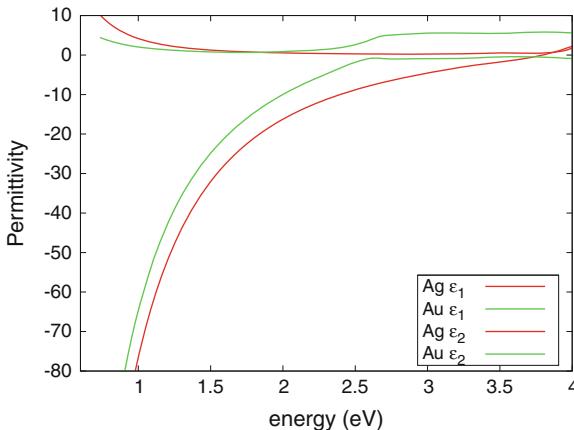
Fig. 7.26 Comparison between the silver Drude-Lorentz model (real part—*red curve* and imaginary part of the permittivity—*green curve*) and the gold data [53] (real part—*blue* and imaginary part—*magenta*)

7.5.3.2 Ellipsometric Response and Surface Plasmons

Although the technology for the fabrication of gold nanostructures is advanced, silver gratings are highly interesting due to the smaller damping constant of silver for well prepared films. The small imaginary part of the dielectric function of silver requires a higher number of Fourier components in the RCWA simulation to avoid artefacts and ensure convergence.

For silver gratings, we have fabricated samples with the same geometrical parameters as discussed before, e.g., the grating in *Field* (d) with a period of $P = 800$ nm, a linewidth of $L = 330$ nm and a height of $h = 50$ nm, which can thus be compared directly with the corresponding gold grating.

In Fig. 7.27a, b we present the measured Ψ and Δ spectra, respectively, for a silver (red) and a gold grating (blue) for the given geometry. The spectral range is reduced as we discard values larger than 4 eV due to the artificial peak of the used silver dielectric function at $E = 4.1$ eV. For low energies $E < 2.5$ eV, the gold and silver gratings show qualitatively very similar spectra although the absolute Ψ values are slightly lower for the silver grating. Above $E > 2.5$ eV the spectra are clearly different which is caused by the different behavior of the dielectric function (see Fig. 7.26).

The lower plasmonic resonance energies E_{-1} and E_{-2} are basically equal as the slope of the surface plasmon dispersion is close to the light line for small frequencies. We see that the E_{-3} peak for the silver sample is more pronounced than for a gold grating as $\varepsilon''(\omega)$ of silver is smaller than that of gold in this energy range.

In Fig. 7.28 we compare the SE measurement of a silver grating (Ψ in (a) and Δ in (b)) with the RCWA simulations with parameters $P = 800$ nm, $L = 330$ nm and
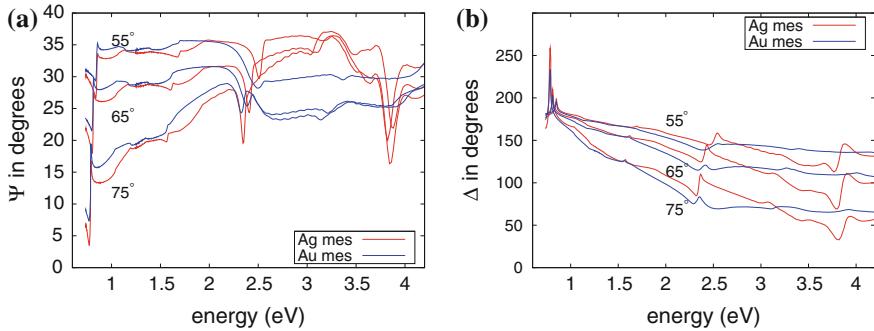
Fig. 7.27 Measured Ψ spectra (**a**) for a silver grating (*red*) and a gold grating (*blue*) of type (d) with a period $P = 800$ nm, linewidth $L = 330$ nm and height $h = 50$ nm. **b** Δ-spectra of the same samples

$h = 50$ nm. The agreement between simulation and experiment is almost perfect up to 3 eV although the silver grating is far from ideal. The linewidth of $L = 330$ nm was determined by SEM.

For Ψ as well as Δ the simulated results differ from the measured ones in the energy range between $E \in [3, 3.5]$ eV. This deviation arises from material issues as a variation of both, L and h, did not yield any better results and periodicity changes are clearly visible as a shift in the surface plasmonic resonances.

The surface plasmon resonance energies for silver are again calculated by Eq. (7.4) ($\varepsilon(\omega)$ is the permittivity of silver). The resonances for a fixed periodicity of a grating depend on the energy as well as the angle of incidence ϕ. In Fig. 7.28a, b the black dotted lines are the solutions of Eq. (7.4) and show the dispersion of surface plasmons excited with different integer multiples m of the fundamental reciprocal lattice vector. As mentioned in Sect. 7.4.2, the dispersion of the positive and negative solution with the same $|m|$ become degenerate at normal incidence $\phi = 0°$. By increasing the angle of incidence the surface plasmons propagating in opposite direction to the incident light (solution with negative m) show a red shift whereas the resonances with positive values of m shift to higher energies.

At certain energies and angles of incidence two different dispersion lines for different values of m can cross. The color map in Fig. 7.29b shows a 2D representation of the amplitude of the second derivative of the Ψ spectrum with respect to energy. In this way, peaks in the SE spectra (surface plasmonic resonances) are better identified. We see, that for all experimentally accessible angles of incidence $\phi \geq 44°$ and below $E < 2.5$ eV, the dispersions of the surface plasmon resonances coincide well with the ones predicted by Eq. (7.4).

For smaller values of ϕ the agreement between the resonance position as calculated with RCWA and those predicted by Eq. (7.4) is only good up to an energy of $E \sim 2$ eV. At $\phi = 20°$ and $E \sim 2.2$ eV the E_1 and E_{-2} solutions cross. In simulations this happens at slightly higher energies than predicted by the simple model of Eq. (7.4). From the color map it cannot be seen whether the crossing is avoided by
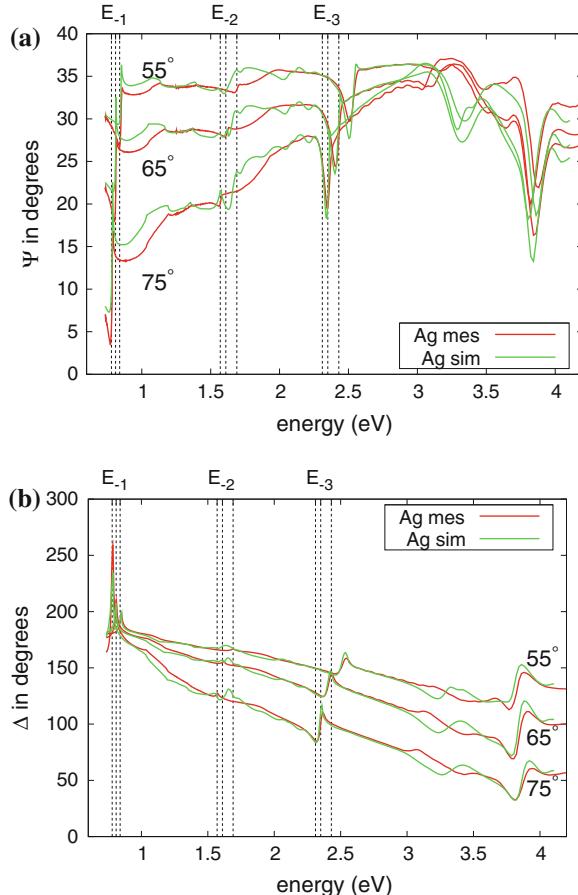
Fig. 7.28 a Measured (*red*) and simulated (*green*) Ψ spectra for a silver grating of type (d) with period $P = 800$ nm, linewidth $L = 330$ nm and height $h = 50$ nm. **b** The Δ spectrum of the same sample. The *grid* shows the predicted resonance energies and serves as a guide for the eye

forming a hybrid solution and introducing a band gap. Concerning the E_1 solution it is astonishing that the surface plasmonic resonance curve does not follow the prediction of the simple model, as for larger energies and angles of incidence than this crossing, the resonance positions in the RCWA calculations lie well below the black line for $m = 1$. This indicates that another effect is influencing the dispersion which is not identified, yet.

In (a) we show a measurement (left) and a simulation (right) of the Ψ spectrum for $\phi \in [44, 75]°$ in steps of $1°$. The simulated spectra exhibit more features, which are also visible in the $\frac{d^2\Psi}{dE^2}$ color map, than in the experimental ones. This is due to the imperfection of the grating as peaks are broadened and some features may be washed out. The additional streaks in the double differentiated RCWA spectrum,
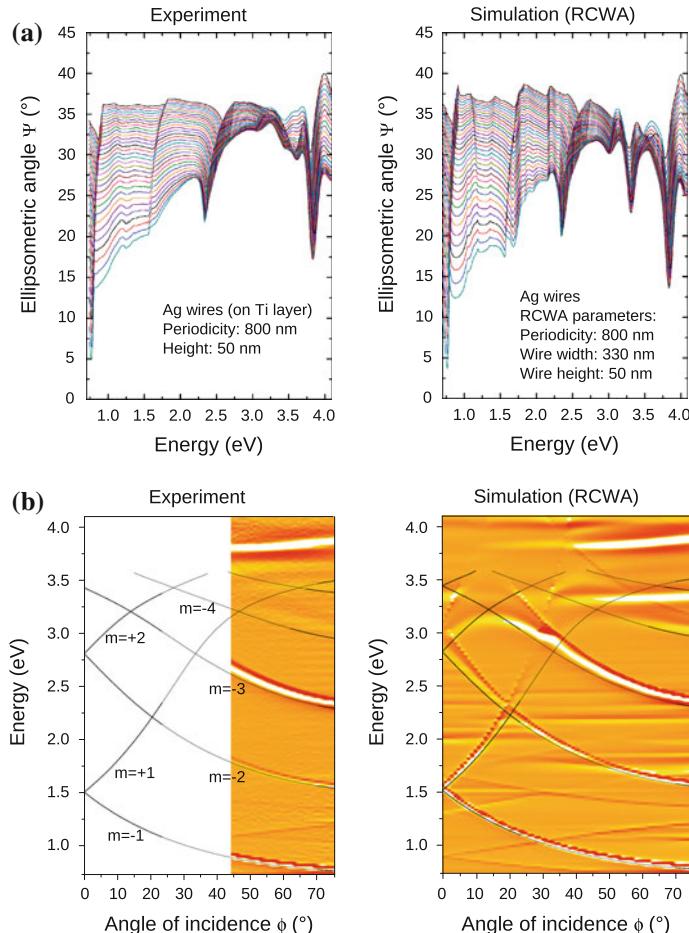
Fig. 7.29 **a** Measured Ψ spectra (*left*) compared to the RCWA simulations (*right*) for various angles of incidence ($\phi \in [44, 74]°$ in steps of 1°), for a field of type (d) with $P = 800$ nm, $L = 330$ nm and $h = 50$ nm. **b** Double differential spectra $d^2\Psi/dE^2$ plotted over E and ϕ for the same structure. The *black lines* indicate the resonance conditions for various values of m calculated from Eq. (7.4)

which do not follow the black lines, can be separated into two different classes: In the low energy region, lines appear that change with the angle of incidence and energy but cannot be described by Eq. (7.4), whereas in the medium energy region $E \sim 1.75$–2.5 eV horizontal streaks appear at constant energies over large ranges of ϕ. The latter effects are caused by the material and not by the grating structure itself.

In Fig. 7.30a we show the measurement (left panel) and corresponding RCWA simulations (right) for a silver grating with a period $P = 2400$ nm, $L = 800$ nm and a height $h = 55$ nm. The Ψ spectra were again measured over the accessible range $\phi = 44°$ and $\phi = 75°$ with $\Delta\phi = 1°$. The measured spectra are compared to
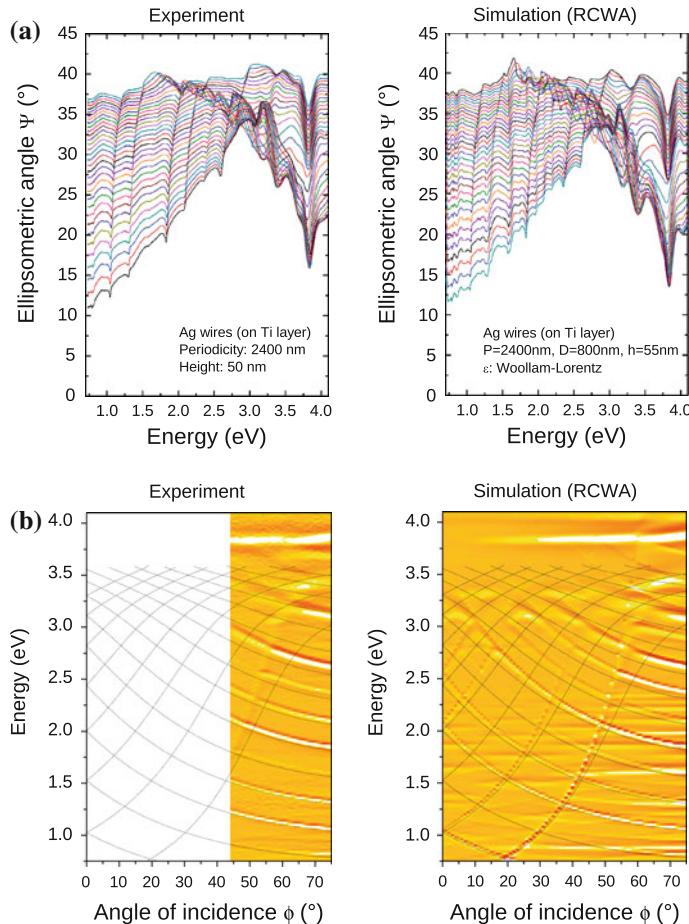
Fig. 7.30 **a** Measured Ψ spectra (*left*) compared to RCWA simulations (*right*) for various angles of incidence ($\phi \in [44, 74]°$,) for a field (a) with $P = 2400\,\text{nm}$, $L = 800\,\text{nm}$ and $h = 55\,\text{nm}$. **b** Double differential spectra $d^2\Psi/dE^2$ plotted over E and ϕ for the same structure. The *black lines* indicate resonance conditions for various values of m calculated by Eq. (7.4)

the simulated ones in the top panel. As can be clearly seen, we obtain more surface plasmonic solutions due to the larger periodicity. As before, we see more features in the simulated spectra but all surface plasmonic resonance peaks are also found in the experimental spectra. In (b) we plot the double differentiated spectrum $\frac{d^2\Psi}{dE^2}$ over the angle of incidence and energy and compare the resulting color map with the dispersion of the resonance energies calculated from Eq. (7.4) (black lines). The number and dispersion of the resonances differs from Fig. 7.29 as the period of the structure is changed to $P = 2400\,\text{nm}$. The dispersion curves for negative values of m are well reproduced by the simulated and measured results of the sample. However,

the resonance energies for positive orders of m are only correct for small angles of incidence and the deviation between calculated energies and experimental/RCWA spectra increases above.

In conclusion, although silver suffers from the well known aging effects, the investigation of the surface plasmonic excitations is very interesting as the peaks are sharp and meet theoretical predictions over a wide range of energies. Especially for low energies, experiment and RCWA simulations match quite well. At larger angles of incidence or higher energies, we observe discrepancies of both with respect to the simple analytical predictions of the resonance positions which indicates additional effects that are not included in the simple theory leading to Eq. (7.4).

7.5.3.3 Mode Pattern of Surface Plasmons

The excitation of surface plasmons at metallic gratings is clearly visible in SE spectra and their derivatives as this effect is strongly dependent on the polarization of the incident light. The fields are enhanced and confined at the metal-dielectric interface and decay exponentially perpendicular to the metallic surface. Therefore, it is very interesting to investigate the field distributions within one period of the grating, monitor the energy flow across the boundaries and check the reflection and transmission coefficients.

We only perform numerical simulations in this section and use a model grating with $P = 800$ nm, $L = 300$ nm and $h = 50$ nm. The dielectric function of silver in this part is taken from Johnson and Christy [26] as this data set is widely used in literature.

In Fig. 7.31a we present the ellipsometric spectrum (red) of the grating together with the number of non-evanescent diffraction orders (green—right scale). The number of diffracted orders increases with increasing energy and as soon as a new order shows up, the parallel wavevector component $k_{x+G} = k_x + \frac{2\pi m}{P}$ becomes real for an increasing integer number $|m|$, which denotes the diffraction order. At the threshold energy, each newly arising order $m < 0$ has an angle of $\phi_m = -90°$.

In Fig. 7.31b, we investigate the field distribution $|H_y|^2$ within the $x-z$ plane, which is the total sum of all diffraction orders of the grating described above. The angle of incidence is chosen to be $\phi = 65°$. The dashed lines show the structure, i.e., the silicon wafer, the 2 nm thick SiO_2 layer and the silver grating on top of it. For the chosen p-polarization the magnetic field component is parallel to the silver bars of the grating.

The energies for which the field distributions are plotted, were chosen such that the frequency of the incident light is slightly below or above a change in diffraction orders ($m = -1$: 0.81 eV, $m = -2$: 1.62 eV, $m = -3$: 2.4 eV) and at energies away from plasmonic resonances ($E = [0.74, 1.21, 2.00, 2.80]$ eV). We see that the amplitude of the magnetic field is largest at the silver/air interface for energies around the calculated plasmonic excitation frequencies. If a new propagating diffraction order is allowed, the field is lowered as energy can propagate in a new diffraction channel.
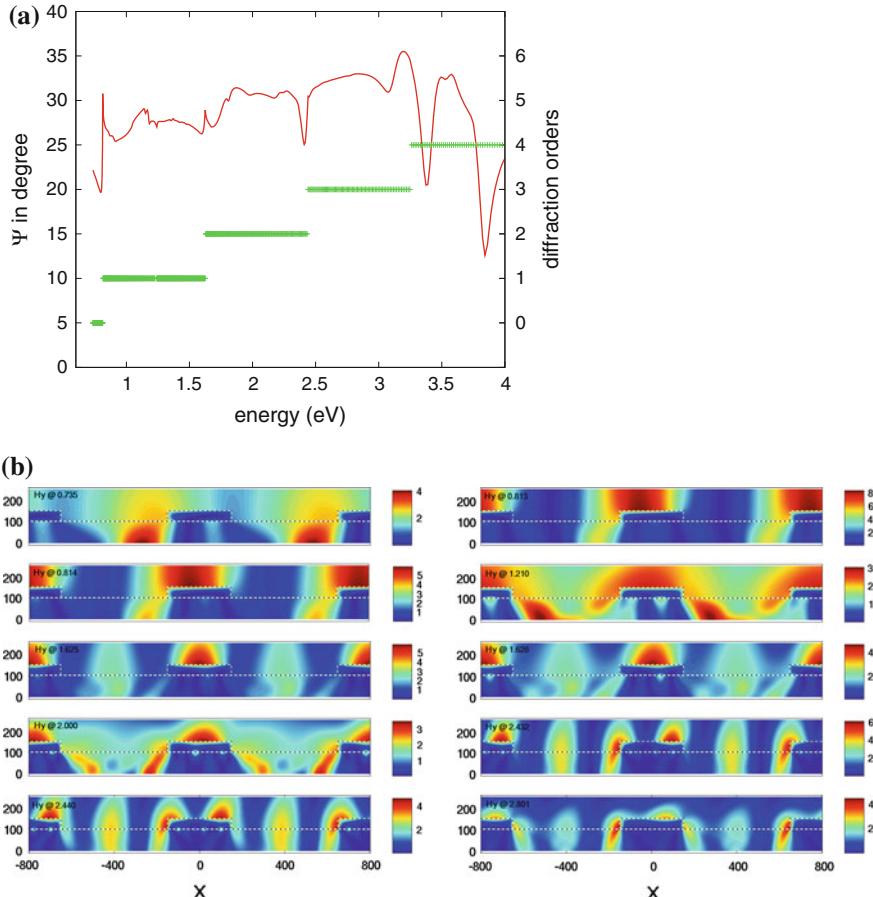
Fig. 7.31 **a** Ellipsometric Ψ spectrum (*red*) and the number of non evanescent diffraction orders (*green*) for a silver grating with $P = 800$ nm, $L = 300$ nm and $h = 50$ nm. **b** Magnetic field distributions for *p*-polarization (where H_y is perpendicular to the plane of incidence) at different energies of the incident light ($\phi = 65°$). The energies are denoted at the *top right*: $E = [0.735, 0.813, 0.815, 1.210, 1.625, 1.628, 2.000, 2.432, 2.440, 2.801]$ eV

For intermediate values of the energy, the field distribution has its maxima between two silver lines and the amplitude of $|H_y|^2$ is smaller.

The time evolution of the fields can also be monitored by RCWA (the origin of the coordinate system has to be varied to obtain all phases) or by using finite elements method (FEM)[9] by multiplying the solutions with $e^{\iota \omega t}$. Another possibility are real space methods like the finite difference time domain method (FDTD) [51].

---

[9] Martin Huber has visualized the time evolution using an FEM code. The movies are available on request.

With RCWA and its implementation in `reticolo`, we are able to investigate the energy flow inside a grating by evaluating the energy balance equation. This relation is the main subject of the second part of this thesis ("Energy density in absorbing media") and it is discussed in detail there. The relation reads

$$- \int_V \nabla (E \times H) d^3 r = - \oint_S (E \times H) da = \int_V \left(\frac{d}{dt} w + \omega \varepsilon_2(\omega) E^2 \right) d^3 r \quad (7.8)$$

The volume integral of the negative divergence of the Poynting vector $S = E \times H$ is transformed into a surface integral which yields a term that is proportional to a time derivative of the energy density w and a lossy term $v = \omega \varepsilon_2(\omega) E^2$. In the next part we show that the averaged energy density w vanishes for monochromatic excitations (which are used in RCWA) and we are able to derive the loss of an arbitrary structure from Eq. (7.8).

This is done by either evaluating the loss in a direct way by the volume integral $\int_V v \, d^3 r$ or indirectly by $\oint_S E \times H \, da$ which is the integration of the normal component of the Poynting vector over the surface of a unit cell. As we have a two dimensional problem and our system has a continuous symmetry normal to the plane of incidence, the dimensions of the integrals are reduced by one. For the indirect expression $\oint_S E \times H \, da$ (which becomes $\oint E \times H \, dx$), the normal component of Poynting's vector is integrated along the boundaries of a unit cell. As we have periodic boundary conditions, the energy flow through the right boundary is equal to the flow at the left boundary. Therefore, the loss is determined by the flow through the top boundary, which is equal to the total reflection, minus the flow through the bottom boundary, which is the transmission. The direct expression $\int_V v(r) \, d^3 r$ reduces in the case of metallic gratings to an integral of the loss over the area of a unit cell $\int_A v(r) \, da$.

In Fig. 7.32a we show the loss of one period of the silver grating with the same dimensions as before $P = 800 \, nm$, $L = 300 \, nm$, $h = 50 \, nm$. The loss was calculated by the flow of the Poynting vector (green) and the direct expression via the surface integral over v (blue). The results are identical except for a slight deviation above 4 eV. We compare this result for a silver grating with a homogeneous silver layer, which has the same thickness as the grating (50 nm). In order to excite a surface plasmon, we use in our simulations the Kretschmann geometry, i.e., the ambient has a refractive index $n = 1.4$ and the substrate is air. The resulting resonance frequency is at $E \sim 3.4 \, eV$ for an angle of incidence of $\phi = 65°$.

From this comparison we see that the absorption of a homogeneous silver layer and a silver grating have the same qualitative shape. In the energy region where the wavelength is comparable with typical structure sizes of the grating ($E < 3 \, eV$), we obtain a more complicated loss spectrum with additional features arising from the grating structure.

In Fig. 7.32b, c we analyze these features by plotting the line integral of the energy flux along the left/right boundary S_x (left and right boundary of a unit cell have to be equal due to the periodic boundary conditions) and along the bottom surface (S_z) of the silver layer/grating, respectively. The in-plane component of the Poynting
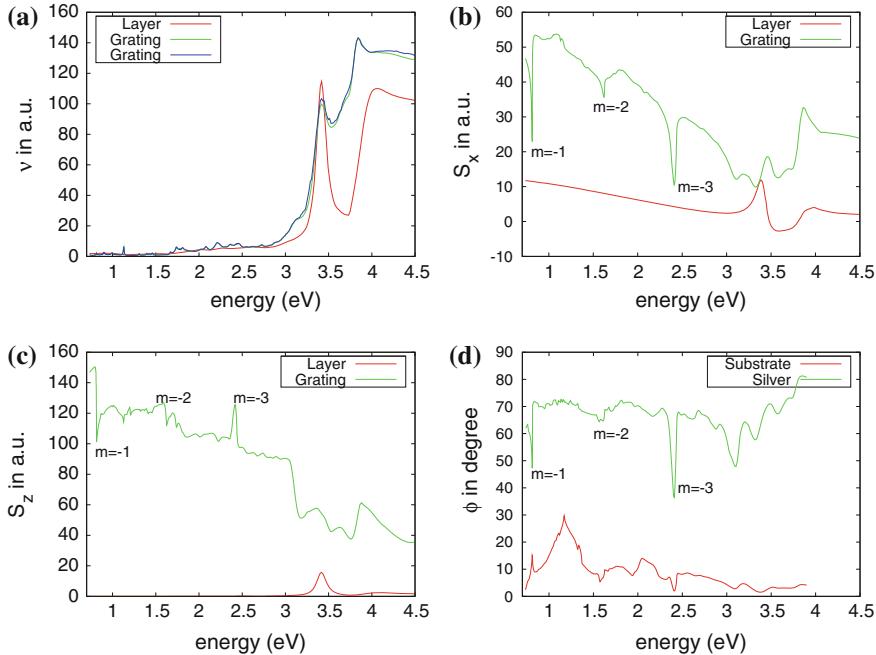
Fig. 7.32 a Absorption (loss ν) in a unit cell of a grating derived by line integrals of the Poynting vector along the boundaries (*green*) and by volume integration (*blue*) compared with the absorption of a Ag layer ($h = 50$ nm) with a fictive periodicity ($P = 800$ nm) (*red*). **b** The line integral along the left boundary of the unit cell compared for a layer of silver (*red*) and a grating (*green*). **c** The line integral of the Poynting vector along the bottom interface of the silver layer (*red*) and the grating (*green*). **d** Averaged Poynting vector at the bottom of the silver structure (*green*) and inside the substrate (*red*). The angle of incidence is $\phi = 65°$ for all plots

vector S_x for a homogeneous layer is enhanced for the surface plasmonic resonances, whereas it shows pronounced dips at the resonance energies of the silver grating ($E_{-1} = 0.8$ eV, $E_{-2} = 1.6$ eV, $E_{-3} = 2.4$ eV). This is caused by the different propagation characteristics as the surface plasmons in the grating are excited by negative in plane wavevectors ($m < 0$).

Investigating the energy flow through the bottom interface of the silver (Fig. 7.32c) we see, that the energy flux is enhanced at the plasmonic resonances, especially for the first energy resonance $E_{m=-1}$. Therefore, we plot in (d) the angle of the averaged Poynting vector $\tan \phi = \sum_{x_i} S_x(x_i, z_0) / \sum_{x_i} S_z(x_i, z_0)$ in the substrate (i.e., z_0 is in the substrate—red curve) and at the interface of the silver lines to SiO_2 (i.e., z_0 is chosen to be at the interface—green curve). We average along the x-direction over one period (at an according height (z_0) in the structure) and see that the energy flux changes at the resonance condition. This method can also be used to investigate negative index materials as the angle ϕ representing the energy flow has

to be negative. However, as the result changes strongly in the vicinity of the grating one has to average at a suitable vertical position z_0.

The effect of surface plasmon resonances is also seen in the specular reflection and transmission spectra, which are shown in Fig. 7.33a. The amplitude of the reflectance $R = |r|^2$ and transmittance $T = |t|^2$ for a silver grating are plotted. For p-polarization and angle of incidence $\phi = 65°$ we have indicated the calculated surface plasmon resonance energies by arrows. The first, very sharp, resonance differs from the other two resonances, as we obtain a very high transmission. Additionally, we see discontinuities in the spectrum at the energies indicated by the dotted vertical lines. As a new propagating diffraction order appears, intensity is removed from the specular reflection and transmission and transferred into this new channel (Woods anomaly [44]). As a consequence, a sharp feature appears in the R and T spectra.

Figure 7.33b shows R and T for s-polarization. One can distinguish two broader peaks at 1.8 and 2.75 eV, respectively, and several sharp peaks between 1 and 1.8 eV. The broad peaks are caused by the finite thickness of the silicon (Fabry-Perot resonances) as they change their position upon variation of the thickness (increasing the thickness of the silicon substrate layer leads to a redshift of these peaks). The sharp peaks in the transmission spectra occur due to the possibility of coupling out one diffraction order.

Contrary to the previous results, for the simulation of R and T we have used a silver grating structure on top of a 100 nm thick silicon layer instead of a semi infinite substrate in order to obtain a finite transmission through the absorbing silicon layer. We fixed the thickness of the silicon layer to 100 nm to avoid too many Fabry-Perot resonances. The sample was again illuminated under an angle of incidence of $\phi = 65°$.

From the field plots in Fig. 7.31b and from Fig. 7.33a we see that the transmission is enhanced at the first resonance at $E_{-1} = 0.8$ eV. This phenomenon was previously observed and discussed for sub-wavelength hole arrays [12]. In our case, the enhanced transmission is also a sub-wavelength phenomenon as we have a periodicity of $P = 800$ nm but a plasmon wavelength of $\lambda_{m=-1} = \hbar c / E_{m=-1} = 1530$ nm which is about twice as large as the grating period P (a very similar geometry situation compared to the 2D hole array presented in the literature). In Fig. 7.34 the transmission spectrum of the grating can be compared with the results from literature [12]. At higher wavelengths, the spectrum has the same shape as found in [12] supporting the findings that the extraordinary transmission for $\lambda > P$ does not depend on the shape of the holes [33]. For slightly lower wavelengths than the zero order transmission peak at $\lambda \sim 1520$ nm and $\lambda \sim 1280$ nm, respectively, both spectra exhibit a kink, which is caused by the appearance of the first diffraction order (Wood's anomaly). In the relevant figure in [12] the thickness of the silver film was 200 nm, the cylindrical holes had a diameter of 150 nm and were orientated in a 2D array with a periodicity of $P = 900$ nm. In our case the transmission peak is much more intense as we have a thin silver grating and present simulation results from a perfect grating. From the field distribution in Fig. 7.31b we observe an enhancement of the field between the grating lines indicating this large transmission effect (which is also apparent in the movies). In a recent study [33] a transmission model was derived to describe the extraordinary
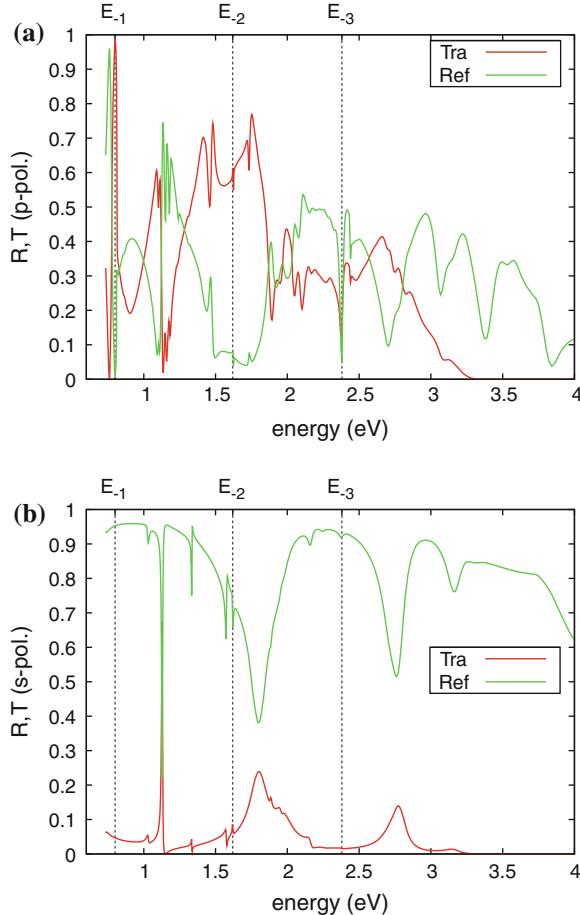
Fig. 7.33 Reflectance R and transmittance T of a silver grating of type (d) with a period $P = 800\,\text{nm}$, linewidth $L = 330\,\text{nm}$ and $h = 50\,\text{nm}$ placed on a layer of $100\,\text{nm}$ of silicon. The angle of incidence is $\phi = 65°$. The incident light is p-polarized in **a** and s-polarized in **b**

transmission of the 2D hole arrays. Using RCWA the experimental spectra could be retrieved exactly as it uses the transmission matrix formalism explicitly and yields exact results.

Finally, we investigate the lineshape of the resonance which we assume to be Fano type as a direct scattering process is coupled with an indirect one. Therefore, we use Eq. (7.6) to investigate three surface plasmon resonances ($m = -1$, $m = -2$ and $m = -3$) in the reflectivity spectra for p-polarization. In Fig. 7.35 we present these resonances for 4 different linewidths. The silver grating again has a period $P = 800\,\text{nm}$, a height $h = 50\,\text{nm}$ and the linewidth takes the values of $L = 200\,\text{nm}$ (magenta), $L = 300\,\text{nm}$ (blue), $L = 330\,\text{nm}$ (green) and $L = 350\,\text{nm}$ (red). In (a)-(c), where we present the $m = -1$, $m = -2$ and $m = -3$ resonance, respectively, it
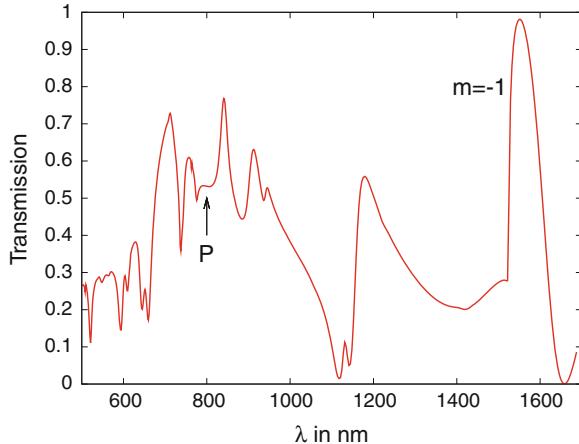
Fig. 7.34 Transmission of a silver grating with $P = 800$ nm, $L = 300$ nm and $h = 50$ nm on a 100 nm thick silicon layer. The angle of incidence is $\phi = 65°$
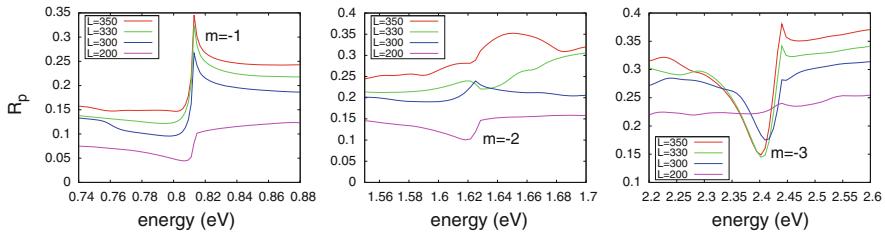

Fig. 7.35 The reflectivity for p-polarization at the surface plasmon resonances $m = -1$ (**a**), $m = -2$ (**b**) and $m = -3$ (**c**) for four different linewidths ($L = 350$ nm (*red*), $L = 330$ nm (*green*), $L = 300$ nm (*blue*) and $L = 200$ nm (*magenta*)). The angle of incidence is $\phi = 65°$, the period is $P = 800$ nm and $h = 50$ nm

is clearly visible that the resonances change significantly by varying the linewidth. We have analyzed this quantitatively by fitting the Fano formula to the resonance. From Table 7.5 we see that the values for q are significantly different for $L = 200$ nm compared to the other linewidths. We further see, that the Fano parameters are not very large (e.g., in [18] Fano parameters of $q \sim 1.4-3$ were presented for a photonic crystal) which indicates a weak re-radiation of the surface plasmons and, therefore, a weak coupling of surface plasmons to the scattered light. However, by changing the linewidth of the structure we influence this coupling and may enhance it.

The influence of the linewidth is not considered in the simple model to describe the resonance position which we have presented in Sect. 7.4.2. Here, we clearly see that for a more sophisticated description we have to go beyond the model presented in this work.

In conclusion, we have shown that different kinds of surface plasmons are excited on a metallic structure. Following the notation from [10] we can divide them into

Table 7.5 Fano parameter q for different geometries and resonance types ($m = -1, -2, -3$)

L (nm)	$q_{m=-1}$	$q_{m=-2}$	$q_{m=-3}$
200	0.772 ± 0.05 [0.75;0.85]	0.30 ± 0.02 [1.5;1.8]	2.04 ± 0.18 [2.3;2.74]
300	1.46 ± 0.11 [0.75;0.85]	1.82 ± 0.11 [1.58;1.7]	0.27 ± 0.02 [2.3;2.55]
330	1.29 ± 0.08 [0.75;0.85]	–	0.33 ± 0.04 [2.3;2.55]
350	1.14 ± 0.08 [0.75;0.85]	0.79 ± 0.10 [1.59;1.68]	0.43 ± 0.03 [2.3;2.6]

The numbers in the square brackets indicate the fit range

"horizontal" surface plasmons which are propagating in the plane at the metallic grating surface. These are absorbed at the metal/air (and substrate) interface. In contrast, in the long wavelength region where $\lambda > P$ one "vertical" excitation exists that may enhance the transmission significantly.

7.6 Homogenization Methods

In the past years there was a rising interest in artificial materials which exhibit specific optical characteristics. This was achieved by creating structures with dimensions comparable with the operated wavelength, as for instance the gratings described in the previous sections or photonic crystals [25].

During the last decade, special attention has been paid to negative index materials (NIMs) which consist of structure sizes that are well below the operation wavelength. These materials are "artificial" concerning their optical response, as materials with a negative n cannot be found in nature.

In case such a material would be available, very exciting properties such as perfect lensing [42] and cloaking [46] could be realized. In 2006, the first negative refraction was experimentally demonstrated in the microwave region [46] and since then big efforts were set on the goal, to find smaller structures that allow to have negative refraction in the visible spectral range. A good overview and summary up to the beginning of 2007 is given in [50].

The key ideas about materials with a negative effective refractive index [32, 52], pointed out, that one has to use the correct sign of the square root in the equation

$$n = \sqrt{-\varepsilon}\sqrt{-\mu} = -\sqrt{\varepsilon\mu} \tag{7.9}$$

as the square root has to be taken from ε and μ separately. This relation points out that if the permeability and the permittivity are both negative, the negative sign of the square root has to be used and the resulting refractive index will be negative. In the visible, a negative dielectric function is realized by using metallic structures as $\varepsilon(\omega)$ is negative in this frequency range (see, e.g., the dielectric function of gold in Fig. 7.14). A negative permeability however, cannot be found for optical frequencies as spin

resonances occur in the Gigahertz regime but not for higher energies. Therefore, all materials have a permeability $\mu \sim 1$ in the visible range.

To overcome this limitation, structures were fabricated that should cause magnetic resonances. The first theoretical and experimental approaches were done with split ring resonators [42], later on with open split rings in a "U" shaped form, followed by small rods and fishnet-structures. In the last two years mainly fishnet and swiss cross designs were investigated, mostly fabricated of gold.

Although many efforts were undertaken to find better structures and to scale them down, the concept of resonant micro- and nanostructures for creating magnetic resonances is doubtable. By comparing the magnitude of electric and Lorentz-forces we see, that magnetic effects are a factor $1/c$ smaller than electric features. Therefore, electric resonances are dominating and magnetic resonances caused by the electric field, even by induction, are very weak.

In general, for all artificial structures the same theoretical problem arises than discussed before in ellipsometry, namely how one defines and calculates effective parameters? For nanostructures, the homogeneity is lost and due to highly oriented and periodic structures a spatial dependence of the response is expected.

7.6.1 Effective Parameters

For bulk materials the electric fields vary on atomic scales a_0 and a homogenization procedure over larger length scales $l \gg a_0$ is done [2]. The electric field varies significantly on the atomic scale, but if the optical wavelength is large compared to atomic dimensions $\lambda \gg a_0$ the response field shows no features of these atomic variations.

In linear response theory the response to an applied electric field is given by

$$D(r, t) = \iint dt'dr'\varepsilon(r', r; t', t)E(r', t') \tag{7.10}$$

with a similar expression concerning the magnetic fields B and H. For homogeneous media, the response function $\varepsilon(\omega)$ does not explicitly depend on the coordinates r and r' but only on the distance of $|r - r'|$. Then Eq. (7.10) becomes

$$D(r, t) = \iint dt'dr'\varepsilon(|r' - r|; |t' - t|)E(r', t') \tag{7.11}$$

(time is always homogeneous) which now has the form of a convolution integral. A convolution integral becomes a simple multiplication in Fourier space and we obtain

$$D(k, \omega) = \varepsilon(k, \omega)E(k, \omega) \tag{7.12}$$

When talking about permittivities and permeabilities we think of quantities in Fourier space (k, ω) and keep in mind the assumption of homogeneity. A possible k dependence can be dropped if the distance $|r - r'|$ on which ε is changing, is small compared to the wavelength of the incident light.

At this point a serious problem arises: Gratings, NIMs, or generally speaking all metamaterials, are not homogeneous and often even not periodic media in the direction of the incident propagating light. This destroys the basic assumption of homogeneous media as we have a layered medium and further, the internal structure size is on the same lengthscale of the incident wavelength.

In literature, several ways have been presented to extract effective parameters. One possible way is to use the pseudo-dielectric formula

$$\langle \varepsilon \rangle = \sin^2 \phi \left[1 + \left(\frac{1 - \rho}{1 + \rho} \right)^2 \tan^2 \phi \right] \tag{7.13}$$

which is the inversion of ellipsometric data assuming a semi-infinite effective medium. For a thin overlayer a three phase model is derived [34]. Both inversions are analytically correct and can be applied to homogeneous layers. The limits of applicability are well known in ellipsometry, and violations, e.g., too thick overlayers yield an effective permittivity with a negative imaginary part due to interference effects.

Another ansatz is to use effective medium approximations which are useful in finding an effective index for random mixtures of materials. In the EMTs, the models are restricted to random mixtures of mainly spherical inclusions as the Clausius-Mossotti formula is used as a basis to describe the relation between external and local field (see e.g. [47]). Previously, we have shown that the Bruggeman EMA fails to describe the ellipsometric response of a metallic grating due to the highly orientated periodic structure.

The most recent idea in assigning effective parameters to a metamaterial whose internal structure is smaller than the incident wavelength was done by inverting the Airy formula for a three phase system: The refractive index of the ambient n_1 and the substrate n_S are well known and the layer in between (homogeneous or structured) with a thickness d is described by effective parameters

$$\varepsilon_{\text{eff}} = \frac{n_{\text{eff}}}{z_{\text{eff}}} \qquad \mu_{\text{eff}} = n_{\text{eff}} z_{\text{eff}} \tag{7.14}$$

with the effective refractive index n_{eff} and effective impedance z_{eff}. The two complex reflection and transmission parameters for each polarization are now measured or simulated and by solving the inversion problem (for normal incidence $\phi = 0°$) we obtain from

$$z_{\text{eff}} = \pm \sqrt{\frac{(1 + r)^2 - t^2}{(1 - r)^2 - t^2}} \qquad n_{\text{eff}} = \pm \frac{1}{kd} \left[\arccos \left(\frac{1 - r^2 + t^2}{2t} \right) + 2\pi m \right] \tag{7.15}$$

all effective parameters. In Eq. (7.15) it was assumed that the slab we want to "homogenize" was placed in air or vacuum [48], i.e., $n_1 = n_S = 1$.

The term $2\pi m$ represents the different branches arising from Fabry-Perot resonances of the slab. To obtain the correct branch, the right values of m as well as the correct signs have to be found in order to guarantee that the physical constraints of the effective parameters are fulfilled. These requirements are that "for passive materials, $\mathrm{Re}(z)$ and $\mathrm{Im}(n)$ must be greater than zero" [48]. This is **not** equivalent to positive imaginary parts of $\varepsilon'' = (-n'z'' + n''z')/|z|^2$ and $\mu'' = n'z'' + n''z'$. This issue is discussed in [9] yielding the additional condition of

$$|n'z''| \leq n''z' \tag{7.16}$$

In various publications this inversion algorithm was extended to more complicated systems as the slab was composed of a NIM with a glass substrate, a NIM on ITO and a glass substrate [27] or a fishnet structure on a dielectric spacer, metal film and glass substrate [30]. This kind of retrieval algorithm allows to calculate effective parameters for slabs which may be arbitrarily complicated, as long as one can measure or calculate the complex transmission and reflection data (Fig. 7.36).

The theory of homogenization was generalized by a group from Jena [35] to arbitrary angles of incidence. They derived the complex transmission and refraction coefficients by introducing an in-plane wavevector component (k_x in the publication—the angle of incidence is therefore given by $\tan \phi = k_x/k_z = k_{||}/k_{\perp}$). For this case, the degeneracy between TE (or s-polarization) and TM (p-polarization) is lifted and the transmission and reflection parameters become polarization dependent. For a system that consists of substrate, film and a cladding these coefficients can be inverted and effective parameters are derived. The authors show that for normal incidence the equations of Ref. [48] (Eq. (7.15)) are retrieved.
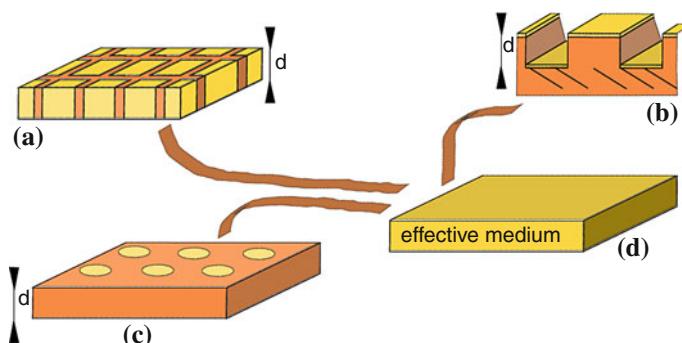

Fig. 7.36 Metamaterials like a negative index material (fishnet) (**a**), a grating (**b**) or a photonic crystal (**c**) with thickness d are replaced by an effective layer (**d**) with the same thickness. The cladding/substrate above and below are assumed to be the same

This ansatz of inverting Airy's formula without approximations looks promising at a first glance: The complex reflection and transmission coefficients for a slab are derived exactly and two complex parameters are obtained. These parameters are inverted to a complex refractive index n_{eff} and a complex impedance z_{eff} and finally a complex effective permittivity and effective permeability is derived.

The homogenization algorithm in the form of Eq. (7.15) as well as in the general formulation [35] fail under certain conditions. First, the structure is not homogeneous in propagation direction and, therefore, we expect a wavevector dependent (or ϕ dependent) effective index. Second, for structure sizes comparable to the wavelength, diffraction occurs and the higher diffraction orders are neglected in the derivation.

In general, effective permittivities and permeabilities have to fulfill certain boundary conditions which were derived in static an quasi static approximations. The dielectric properties of any mixture of materials are constrained by the dielectric function of the components in the sense that

$$\min\{\varepsilon_1, \varepsilon_2\} \leq \varepsilon_{\text{eff}} \leq \max\{\varepsilon_1, \varepsilon_2\} \tag{7.17}$$

and similar for the permeability. For mixtures of two materials with ε_1 and ε_2 and a volume fraction of f and $1 - f$, respectively, these so called Wiener bounds were derived in the beginning of the last century and read

$$\varepsilon_{\text{eff,max}} = f\varepsilon_1 + (1 - f)\varepsilon_2 \tag{7.18a}$$

$$\varepsilon_{\text{eff,min}} = \frac{\varepsilon_1 \varepsilon_2}{f\varepsilon_2 + (1 - f)\varepsilon_1} \tag{7.18b}$$

depending on whether the inclusions are orientated in the direction along the flux of the electric field or perpendicular to it [3, 6, 47, 54].

Stricter bounds than formulated by Wiener were found by using variation theorems [19] and the resulting Hashin-Shtrikman bounds read [47]

$$\varepsilon_{\text{eff,max}} = \varepsilon_2 + 3f\varepsilon_2 \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + 2\varepsilon_2 - f(\varepsilon_1 - \varepsilon_2)} \tag{7.19a}$$

$$\varepsilon_{\text{eff,min}} = \varepsilon_1 + 3(1 - f)\varepsilon_1 \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_2 - 2\varepsilon_1 - (1 - f)(\varepsilon_2 - \varepsilon_1)} \tag{7.19b}$$

The Hashin-Shtrikman bounds are equal to the Maxwell-Garnett mixing rules for spherical inclusions. Both, the Wiener and the Hashin-Shtrikman bounds, are fully symmetrical to the exchange of materials, i.e., $\varepsilon_1 \rightarrow \varepsilon_2$, $\varepsilon_2 \rightarrow \varepsilon_1$ and $f \rightarrow 1 - f$. Due to the duality between electrostatic and magnetostatic problems, the Hashin-Shtrikman bounds are also valid for the permeability. Therefore, these results from electrostatics suggests, that no effective magnetic permeability different from $\mu_{\text{eff}} = 1$ is possible as all components of metamaterials posses a permeability very close to unity.

In the next section, we retrieve the effective parameters for a metallic grating and compare them with effective medium theories and the Hashin-Shtrikman bounds.

7.6.2 Testing the Retrieval Algorithm

As an example to test retrieval algorithms, we simulate reflection and transmission coefficients for a silver grating with a periodicity $P = 300$ nm, linewidth $L = 100$ nm and a height of $h = 50$ nm for normal incidence ($\phi = 0°$). This grating is completely embedded in air and for simplicity no layer or substrate underneath is present. We have chosen a rather small periodicity as in cases where the wavelength is comparable to the structure sizes, diffraction occurs (in this case for $\lambda \sim P \Rightarrow E \sim 4.1$ eV). For shorter wavelength than the period, energy is lost in higher diffraction orders and is not considered in the homogenization algorithm as the effective parameters are only retrieved from specular orders. Therefore, we restrict ourselves to the limitation of $kP \ll 1$ [49].

Although we have normal incidence we have to distinguish between s- and p-polarization as the grating breaks the symmetry. The configuration corresponds to a filling factor $f = L/P = 1/3$ for material 1 with $\varepsilon_1 = \varepsilon_{Ag}(\omega)$. In the RCWA simulation for $r_{s,p}$ and $t_{s,p}$ as well as for the calculation of the Hashin-Shtrikman bounds we use the Johnson-Christy data for silver [26].

In Fig. 7.37 we show the bounds for an effective material that consists of $2/3$ air and $1/3$ of silver. The red curves in (a) and (b) show the real and the green curves the imaginary part of the maximum/minimum effective dielectric index derived by Hashin-Shtrikman bounds Eq. (7.19). By comparison, we see that the minimum bound corresponds to effective dielectric function for s-polarization, which we have obtained from the retrieval algorithm (7.15). The real part, which is shown in blue is always larger than the minimum bound in red up to an energy of around $E \sim 3.6$ eV. The imaginary part obtained by the retrieval formula (magenta) is always smaller than the Hashin-Shtrikman bound. Further we see, that the minimum effective index has a kink at around $E = 3.8$ eV which is caused by the vanishing real part of the denominator in the formula for the minimum bound Eq. (7.19b). At $E = 4.1$ eV the retrieved index shows a dis-continuous behavior coming from the arising first order reflection and transmission order.

The result of the retrieval procedure for p-polarization is shown in Fig. 7.37b. For energies below $E \sim 3$ eV the retrieved real index is lower than the maximum bound Eq. (7.19a). The Hashin-Shtrikman bound shows a strong resonance 3.24 eV which is caused by the vanishing real part of the denominator in Eq. (7.19a). This resonance is also present in the retrieved dielectric function but the resonance frequencies is 0.1 eV lower than predicted by the maximum bound.

These results show, that we have violated the quasi static limit which is approximated by $\lambda/2\pi > P$. Wiener bounds and Hashin-Shtrikman bounds are, similar to all effective medium theories, derived for the electrostatic problem. This implies that wave propagation features, i.e., time and space variation of the incident electric and magnetic field, were neglected.

In Fig. 7.38 we show the result for the effective magnetic permeability as retrieved from Eq. (7.15) for the two different polarizations (red/blue: real part of μ_{eff} for s-/p-polarization; green/magenta: imaginary part). According to the homogenization
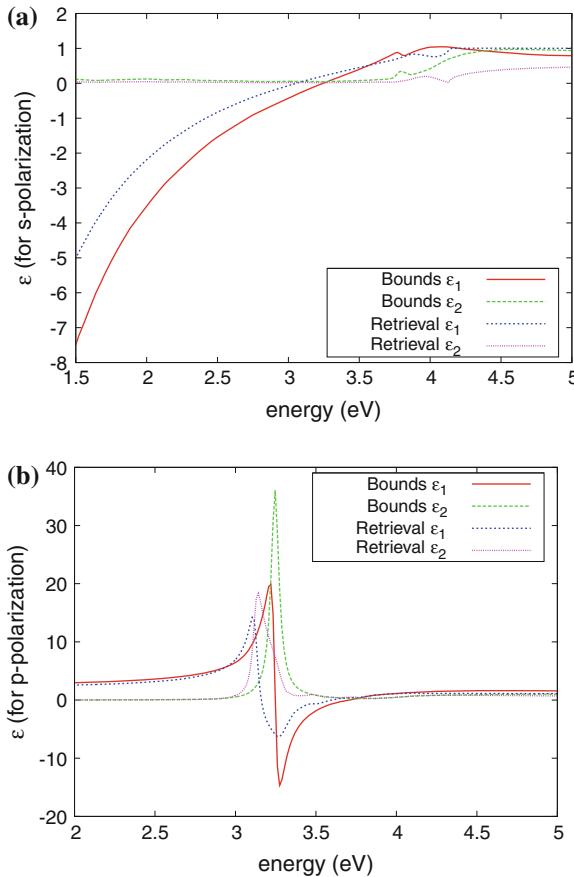
Fig. 7.37 Comparison of Hashin-Shtrikman bounds (*red* real part of the permittivity, *green* imaginary part) and results of the retrieval algorithm (*blue* real part, *magenta* imaginary part of the effective dielectric function) for a silver grating with $P = 300\,$nm, $L = 100\,$nm and $h = 50\,$nm embedded in air. The angle of incidence was $\phi = 0°$. In **a**, the incident light was *s*-polarized, in **b** we used *p*-polarization

algorithm, we obtain an effective permeability different from one for both polarizations (we did not take care about the correct sign as only the magnitude is relevant here). This shows that the retrieval algorithm yields wrong results and artificially creates magnetic effects even for wavelengths smaller than the structure sizes. In the used grating structure no materials were used with $\mu \neq 1$ and no magnetic resonance structure was provided.

We therefore conclude, that the concept of the underlying homogenization algorithms is misleading. Periodically structured materials have a dielectric function that is changing over distances comparable to the wavelength and the characteristic length scale is not vanishingly small compared to the incident wavelength. We
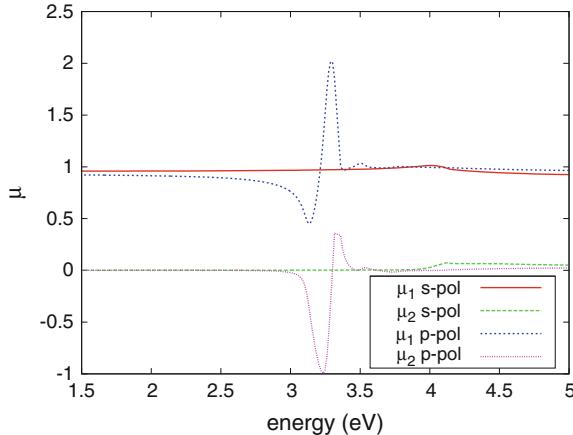
Fig. 7.38 Effective permeability for a silver grating with $P = 300\,\mathrm{nm}$, $L = 100\,\mathrm{nm}$ and $h = 50\,\mathrm{nm}$ surrounded by air for normal incidence ($\phi = 0°$). Due to the structured surface we obtain two "effective" indices

therefore expect that a suitable description of metamaterials should be based on a spatial dispersion with a resulting effective dielectric function $\varepsilon(k, \omega)$. For NIMs, it was already shown, that the effective refractive index obtained by homogenization algorithms is a function on the angle of incidence [11]. The effective magnetic permeability however, should not be different from unity since the Hashin-Shtrikman bounds for materials with $\mu = 1$ only allow an effective permeability of unity and any possible "magnetic resonances" are a factor $1/c$ smaller than electric effects.

7.7 Conclusion and Outlook

Micro- and nanostructures gained more and more interest in science and technology during the last two decades. Since the development of nanoimprint lithography a production over macroscopic scales can be realized which allows the use of spectroscopic ellipsometry for their characterization. Ellipsometry is an easy, fast, non-invasive and non-destructive technique. As one dimensional, metallic gratings provide a good basis for typical nanostructures we have investigated the optical properties by variable angle spectroscopic ellipsometry and numerical simulations.

After giving a short introduction about the fabrication process and necessary cross characterization techniques we have demonstrated the principles of ellipsometry and have introduced the simulation procedure based on the rigorous coupled wave analysis (RCWA). With this technique, we have simulated ellipsometric spectra. Specifically, silicon and gold gratings were investigated with respect to their material and geometric parameters, the matching of experimental and simulated spectra was analyzed and possible deviations were identified. By comparing the exact results with

those obtained with effective medium theories, we have shown that the response of metallic gratings is beyond the reach of standard effective medium approximations and a proper modelling of the structure is necessary.

In the ellipsometric spectra of gold and silver gratings resonances have shown up which are sensitive to the angle of incidence, the energy and the periodicity of the structure. As these resonances were sensitive to one polarization of the incident light and by their characteristic field distribution we have identified them as surface plamsons. We have introduced a model which includes the coupling of surface plasmons and the incident light via the reciprocal lattice vector of the grating. The resulting solutions of this model match the resonances in the spectra exactly. In detail, we have investigated the dispersion of surface plasmons on silver gratings and analyzed the energy flow inside a unit cell.

In this chapter we have considered the question of effective indices and the optical response. We have introduced the pseudo-dielectric function, the three phase model for thin overlayers, various effective medium approximations and, finally, recent homogenization algorithms and have discussed their limits of applicability. We have proposed that nanostructures yield a wavevector dependent response function as homogeneity is lost.

The techniques and results presented in this part are already used to analyze two dimensional structures, e.g., fishnet and swiss cross designs but the computational effort becomes quite large in two dimensions. For these geometries it is important to investigate sidewall effects as trapezoidal layers change reflection and transmission properties. For complicated structures the energy flow can be analyzed as presented in this work. Further it would be interesting to study non specular reflection and transmission orders experimentally, as they are accessible with numerical simulations.

In future works, we investigate the Fano resonances in detail. As the coupling parameter serves as a measure for the coupling between direct scattering and second order processes (excitation of a surface plasmon and re-radiation), the Fano formula can be exploited to find appropriate grating geometries to foster surface plasmonic resonances.

Another interesting point is to reduce the height of the gratings and apply Rayleigh-Rice theory to describe the specular reflection of one and two dimensional structures. Although effects of surface plasmonic resonances become weaker, it is interesting to compare the results of the Rayleigh-Rice theory to those of the exact RCWA theory and to experimental data.

References

1. F. Abelès, *Optical Properties of Solids* (North-Holland, Amsterdam, 1972)
2. N. Ashcroft, N. Mermin, *Festkörperphysik* (Oldenbourg, München, 2005)
3. D. Aspnes, Local-field effects and effective-medium theory: a microscopic perspective. Am. J. Phys. **50**, 704 (1981)

4. D. Aspnes, E. Kinsbron, D. Bacon, Optical properties of AU: sample effects. Phys. Rev. B **21**(8), 3290 (1980)

5. I. Bergmair, M. Mühlberger, K. Hingerl, E. Pshenay-Severin, T. Pertsch, E. Kley, H. Schmidt, R. Schöftner, 3d materials made of gold using nanoimprint lithography. Microelectron. Eng. **87**, 1008 (2010)

6. M. Bergmair, M. Huber, K. Hingerl, Band structure, Wiener bounds, and coupled surface plasmons in one dimensional photonic crystals. Appl. Phys. Lett. **89**, 081907 (2006)

7. G. Binnig, C.F. Quate, Ch. Gerber, Atomic force microscope. Phys. Rev. Lett. **56**(9), 930–933 (1986)

8. K.-P. Chen, V. Drchev, J. Borneman, A. Kildishev, V. Shalaev, Drude relaxation rate in grained gold nanoantennas. Nano Lett. **10**, 916–922 (2010)

9. X. Chen, T. Grzegorczyk, B. Wu, J. Pacheco, J. Kong, Robust method to retrieve constitutive parameters of metamaterials. Phys. Rev. E **70**, 016608 (2004)

10. S. Collin, F. Pardo, R. Teissier, J. Pelouard, Horizontal and vertical surface resonances in transmission metallic gratings. J. Opt. A Pure Appl. Opt. **4**, S154–S160 (2002)

11. B. Dastmalchi, Berreman retrieval. ICSE V—Albany, 2010

12. T. Ebbesen, H. Lezec, H. Ghaemi, T. Thio, P. Wolff, Extraordinary optical transmission through sub-wavelength hole arrays. Nature **391**, 667 (1998)

13. U. Fano, Effects of configuration interaction on intensities and phase shifts. Phys. Rev. **124**(6), 1866 (1961)

14. D. Franta, I. Ohlìdal, Comparison of effective medium approximation and rayleigh-rice theory concerning ellipsometric characterization of rough surfaces. Opt. Commun. **248**, 459–467 (2005)

15. R. Gajic, Z. Dohcevic-Mitrovic, Z. Popovic, Spectroscopic ellipsometry and correlation measurements. Technical Report, NanoCharM, 2009, Deliverable D5.1a, http://www.nanocharm.org

16. T. Glinsner, P. Lindner, M. Mühlberger, I. Bergmair, R. Schöftner, K. Hingerl, H. Schmidt, E.B. Kley, Fabrication of 3d-photonic crystals via uv-nanoimprint lithography. J. Vac. Sci. Technol. B **25**, 2337 (2007)

17. G. Granet, B. Guizal, Efficient implementation of the coupled-wave method for metallic lamellar gratings in TM polarization. J. Opt. Soc. Am. A **13**, 1019 (1996)

18. R. Harbers, S. Jochim, N. Moll, R. Mahrt, D. Erni, J. Hoffnagle, W. Hinsberg, Control of fano line shapes by means of photonic crystal structures in a dye-doped polymer. Appl. Phys. Lett. **90**, 201105 (2007)

19. Z. Hashin, S. Shtrikman, Conductivity of polycrystals. Phys. Rev. **130**(1), 129 (1963)

20. C. Herzinger, B. Johs, W.A. McGahan, J.A. Woollam, W. Paulson, Ellipsometric determination of optical constants for silicon and thermally grown silicon dioxide via a multi-sample, multi-wavelength, multi-angle investigation. Appl. Phys. **83**, 3323–3336 (1998)

21. I. Hooper, J. Sambles, Dispersion of surface plasmon polaritons on short-pitch metal gratings. Phys. Rev. B **65**(16), 165432 (2002)

22. M. Huber, Radiation efficiency enhancement and diffraction effects of photonic crystals. Master's Thesis, JKU, 2006

23. J. Hugonin, P. Lalanne, Reticolo software for grating analysis. Technical Report, Institut d'Optique, Orsay, France, 2005

24. A. Jaffer, Freesnell: refractive index spectra, http://people.csail.mit.edu/jaffer/FreeSnell/nk.html

25. J.D. Joannopoulos, R.D. Meade, J.N. Winn, *Photonic Crystals* (Princeton University Press, Princeton, 1995)

26. P. Johnson, R. Christy, Optical constants of the noble metals. Phys. Rev. B **6**(12), 4370–4379 (1972)

27. A. Kildishev, W. Cai, U. Chettiar, H. Yuan, A. Sarychev, V. Drachev, V. Shalaev, Negative refractive index in optics of metal-dielectric composites. J. Opt. Soc. Am. B **23**(3), 423 (2006)

28. E. Kretschmann, H. Raether, Radiative decay of non-radiative surface plasmons excited by light. Z. Naturf. A **23**, 2135 (1968)

29. P. Lalanne, G. Morris, Highly improved convergence of the coupled-wave method for TM polarization. JOSAA **17**, 779–784 (1996)
30. L. Lin, R. Blaikie, Negative permeability using planar-patterned metallic multilayer structures. J. Opt. A Pure Appl. Opt. **9**, S385 (2007)
31. S. Maier, *Plasmonics: Fundamentals and Applications* (Springer, New York, 2007)
32. L. Mandelstam, Group velocity in crystal lattice. Zh. Eksp. Teor. Fiz. **15**, 475 (1945)
33. L. Martin-Moreno, F. Garcia-Vidal, H. Lezec, K. Pellerin, T. Thio, J. Pendry, T. Ebbesen, Theory of extraordinary optical transmission through subwavelength hole arrays. Phys. Rev. Lett. **86**, 1114 (2001)
34. J.D.E. McIntyre, D.E. Aspnes, Differential reflection spectroscopy of very thin surface films. Surf. Sci. **24**(2), 417–434 (1971)
35. C. Menzel, C. Rockstuhl, T. Paul, F. Lederer, T. Pertsch, Retrieving effective parameters for metamaterials at oblique incidence. Phys. Rev. B **77**, 195328 (2008)
36. M. Moharam, E. Grann, D. Pommet, T. Gaylord, Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. J. Opt. Soc. Am. A **12**, 1068 (1995)
37. M. Moharam, D. Pommet, E. Grann, T. Gaylord, Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach. J. Opt. Soc. Am. A **12**, 1077 (1995)
38. M. Mühlberger, I. Bergmair, A. Klukowska, A. Kolander, H. Leichtfried, E. Platzgummer, H. Loeschner, Ch. Ebm, G. Grützner, R. Schöftner, UV-NIL with working stamps made from ormostamp. Microelectron. Eng. **86**(4–6), 691–693 (2009). MNE '08—The 34th International Conference on Micro- and Nano-Engineering (MNE)
39. A. Otto, Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. Z. Phys. **216**, 398 (1968)
40. E. Palik, *Handbook of Optical Constants of Solids* (Academic Press, New York, 1997)
41. J. Pendry, Reliability factors for leed calculations. J. Phys. C **13**, 937 (1980)
42. J. Pendry, Negative refraction makes a perfect lens. Phys. Rev. Lett. **85**(18), 3966–3969 (2000)
43. E. Popov, M. NeviFre, Grating theory: new equations in Fourier space leading to fast converging results for TM polarization. J. Opt. Soc. Am. A **17**, 1773 (2000)
44. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer, Berlin, 1988)
45. S. Rice, Reflection of electromagnetic waves from slightly rough surfaces. Commun. Pure Appl. Math. **4**, 351–378 (1951)
46. D. Schurig, J. Mock, B. Justice, S. Cummer, J. Pendry, A. Starr, D. Smith, Metamaterial electromagnetic cloak at microwave frequencies. Sci. Express **314**, 977–980 (2006)
47. A. Sihvola, *Electromagnetic Mixing Formulas and Applications*. IEE Electromagnetic Wave Series, vol. 47 (The Institution of Electrical Engineers, London, 1999)
48. D. Smith, S. Schultz, P Markos, C. Soukoulis, Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients. Phys. Rev. B **65**, 195104 (2002)
49. D. Smith, D. Vier, T. Koschny, C. Soukoulis, Electromagnetic parameter retrieval from inhomogeneous metamaterials. Phys. Rev. E **71**, 036617 (2005)
50. C. Soukoulis, S. Linden, M. Wegener, Negative refractive index at optical wavelengths. Science **315**, 45 (2007)
51. A. Taflove, S. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House, Norwood, 2005)
52. V. Veselago, The electrodynamics of substances with simultaneously negative values of ε and ν. Sov. Phys. Usp. **10**(4), 509 (1968)
53. J. Weaver, C. Krafka, D. Lynch, E. Koch, *Optical Properties of Metals*. Number 18-2 in Physics Data (Fachinformationszentrum, Karlsruhe, 1981)
54. O. Wiener, Without title. Abh. Math-Phys. Kl.Königl.Saechs. Ges. **32**, 509 (1912)
55. J. Woollam, Software for: Spectroscopic ellipsometry, data acquisition and analysis. Technical Report, J.A. Woollam, 2009

56. http://www.microresist.de
57. S. Yoo, D. Aspnes, Elimination of endpoint-discontinuity artifacts in the analysis of spectra in reciprocal space. J. Appl. Phys. **89**, 8183 (2001)
58. P. Yu, M. Cardona, *Fundamentals of Semiconductors*, 3rd edn. (Springer, Berlin, 2001)
59. A. Zayats, I. Smolyaninov, A. Maradudin, Nano-optics of surface plasmon polaritons. Phys. Rep. **408**, 131–314 (2005)

Chapter 8
Ellipsometry at the Nanostructure

Yasuhiro Mizutani and Yukitoshi Otani

Abstract Recently, surface profiles of nanostructures have been reduced in size in order to develop microfabrication techniques. In particular, feature sizes of a few tens of nanometers are common in the semiconductor industry. This chapter mention about the Stokes vector, the Mueller matrix, the Mueller matrix decomposition to evaluate the surface profiles of nanostructures. The profile of the structure is determined from the Mueller matrix, which expresses all the polarization properties of the sample by experimental measurements and calculated values. Furthermore, the Mueller matrix is decomposed for more precise detection. In the final part of this chapter, the experimental results after decomposition are compared to the values obtained by numerical analysis.

8.1 Ellipsometry for Nanoscale Structures

The expansion of the field of semiconductors has resulted in a rapid advance of micromachining technology, which has entailed a demand for increased precision of measurements during processing. Currently, memory modules using 32 nm circuit patterns are available as commercial products manufactured with semiconductor micromachining techniques such as lithography and nanoimprinting. In addition, optical elements making use of polarization characteristics of nanoscale structures, such as wire-grid polarizers and antireflection structures, are becom-

Y. Mizutani (✉)
Institute of technology and science, The university of Tokushima,
2-1 Minami-Josanjima-cho, Tokushima 770-8506, Japan
e-mail: mizutani@tokushima-u.ac.jp

Y. Otani
Center for Optical Research and Education, Utsunomiya University,
7-1-2 Yoto, Utsunomiya, Tochigi 321-8585, Japan
e-mail: otani@cc.utsunomiya-u.ac.jp

ing increasingly widespread. According to Moore's law, the integration density of semiconductors increases exponentially, and there is demand in micromachining technology for in-line shape evaluation methods with precision of the order of nanometers.

At present, nanoscale shape evaluation methods make use of devices such as scanning electron microscopes (SEM), which scan the sample surface with an electron beam, and atomic force microscopes (AFM), which measure surface shapes by scanning the sample surface with a cantilever. Such methods perform direct evaluation of shapes on the sample surface and achieve a level of precision allowing for sub-nanometer shape evaluation. However, in-line shape evaluation based on such methods has problems such as large device size, damage to the sample surface and long measurement times.

Ellipsometry is used in scatterometry as a fast method for evaluating nanoscale shapes, enabling the characteristics and shape of the sample material to be evaluated from certain optical properties of nanostructures, such as their absorption and polarization characteristics [1–3]. Scatterometry allows for contactless non-destructive evaluation, with the advantage of being faster than scanning based on spot measurements since evaluation is performed within the area of the measuring beam.

When a nanoscale structure is irradiated with light, the polarization characteristics of the reflected light change in accordance with the shape of the structure as a result of birefringence [4]. In this way, the nanostructure can be evaluated by measuring changes in the polarization characteristics. Since light irradiated onto a nanostructure is scattered due to the anisotropy and non-uniformity of the structure, ellipsometry analyzing depolarization is required.

Mueller matrix measurements are effective for ellipsometry with depolarization since in addition to properties of fully polarized light such as birefringence, absorption and optical rotation, Mueller matrices can also represent depolarization [5]. Furthermore, Mueller matrix reconstruction as reported in recent years allows for evaluation through decomposition of the sample into depolarization components, phase difference components and absorption components.

8.2 Nanoscale Phenomena in Ellipsometry

Light reflected from nanoscale structures is affected in various ways depending on the structure. Although higher-order diffracted light carries shape information if the scale of the nanostructure is larger than the wavelength, at the nanometer scale, shape information is carried by zero-order diffracted light in the form of polarization. For example, according to the effective medium theory, nanostructures can be regarded as thin films; thus, changes in the polarization of reflected light can be considered in a straightforward manner, which in turn allows for ellipsometry to be used for measurements [4, 6].

Furthermore, Rayleigh scattering occurs when high-precision nanostructures have minuscule dimensional variations [7]. For this reason, while the polarization does

not change in the case of nanostructures repeating with ideal dimensional accuracy, in reality dimensional variations result in depolarization, and such dimensional variations can be evaluated by analyzing depolarization. Therefore, the conventional parameters Δ and Ψ are insufficient when using ellipsometers for depolarizing nanoscale objects, and the characteristics of fully polarized light (including depolarization) have to be measured. In other words, the Mueller matrix has to be evaluated.

8.3 Mathematical Treatment of Polarization Characteristics at the Nanoscale

8.3.1 The Stokes Vector

The Stokes vector S is used to characterize arbitrary polarization states. The Stokes vector is expressed using the intensity of polarized light, and its definition is as follows:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_x + I_y \\ I_x - I_y \\ I_{45°} - I_{-45°} \\ I_R + I_L \end{pmatrix}. \tag{8.1}$$

Here, S_0 is the intensity of irradiated light, and S_1 represents the difference of the intensity I_y in the y direction and the intensity I_x in the x direction for linearly polarized light. Furthermore, S_2 represents the difference of the intensity $I_{-45°}$ in the $-45°$ direction and the intensity $I_{+45°}$ in the $+45°$ direction for linearly polarized light, and S_3 represents the intensity I_L of left-hand circularly polarized light subtracted from the intensity I_R of right-hand circularly polarized light. The Stokes vector can be expressed as follows by using the respective electric field vectors E_x and E_y in the x and y directions and the phase difference δ between E_x and E_y:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} E_x^2 + E_y^2 \\ E_x^2 - E_y^2 \\ 2E_x^2 E_y^2 \cos \delta \\ -2E_x^2 E_y^2 \sin \delta \end{pmatrix}. \tag{8.2}$$

8.3.2 The Mueller Matrix

Taking S_{in} and S_{out} as the Stokes vectors of the polarized light before and after irradiation onto the sample, the relation between S_{in} and S_{out} can be expressed

through a linear transformation as

$$S_{out} = M S_{in} \tag{8.3}$$

$$= \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix}. \tag{8.4}$$

Here, the 4×4 matrix M is referred to as a Mueller matrix, which reflects the polarization characteristics of the object. For more details also check Chap. 2.

8.3.3 Mueller Matrix Decomposition

Changes in the polarization characteristics caused by nanostructures involve a mixture of various components, such as birefringence, depolarization and absorption, in which case the Mueller matrix have to be reconstructed for each polarization characteristic. In this section, we introduce a reconstruction method proposed by Chipman et al. [8], in which the Mueller matrix M can be decomposed into a phase difference component M_R, a depolarization component M_Δ and an absorption component M_D as follows:

$$M = M_\Delta M_R M_D. \tag{8.5}$$

The absorption component M_D can be expressed as

$$M_D = \begin{bmatrix} 1 & \vec{D} \\ \vec{D} & m_D \end{bmatrix}, \tag{8.6}$$

where the element D of M is

$$\vec{D} = \frac{1}{m_{00}} \begin{bmatrix} m_{01} & m_{02} & m_{03} \end{bmatrix}^T. \tag{8.7}$$

Furthermore, from Eq. (8.7), m_D can be written as

$$m_D = \sqrt{1 - D^2} I + (1 - \sqrt{1 - D^2}) \frac{\vec{D}}{|\vec{D}|} \left(\frac{\vec{D}}{|\vec{D}|} \right)^T, \tag{8.8}$$

where I is the identity matrix.

Eliminating the absorption coefficient from M, we obtain

$$M' = M M_D^{-1} = M_\Delta M_R. \tag{8.9}$$

With the absorption component eliminated, the matrix M', the depolarization component M_Δ and the phase difference component M_R can be written as

$$M' = \begin{pmatrix} 1 & \vec{0} \\ \vec{P}_\Delta & m' \end{pmatrix}, \tag{8.10}$$

$$M_R = \begin{pmatrix} 1 & \vec{0} \\ \vec{0} & m_R \end{pmatrix}, \tag{8.11}$$

$$M_\Delta = \begin{pmatrix} 1 & \vec{0} \\ \vec{P}_\Delta & m_D \end{pmatrix}. \tag{8.12}$$

Here, m' is a submatrix of M', and $\vec{P}_\Delta$ can be expressed as

$$\vec{P}_\Delta = \frac{\vec{P} - m\vec{D}}{1 - D^2}, \tag{8.13}$$

$$\vec{P} = \frac{1}{m_{00}} \begin{bmatrix} m_{10} & m_{20} & m_{30} \end{bmatrix}^T. \tag{8.14}$$

From Eqs. (8.8) and (8.9), the relation between submatrices can be derived as

$$m' = m_\Delta m_R. \tag{8.15}$$

From this, the submatrix m is written as

$$m_\Delta = \pm \left[m'(m')^T + (\sqrt{\lambda_1\lambda_2} + \sqrt{\lambda_2\lambda_3} + \sqrt{\lambda_3\lambda_1}) \right]$$
$$\times \left[(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3})m'(m')^T + \sqrt{\lambda_1\lambda_2\lambda_3} \right], \tag{8.16}$$
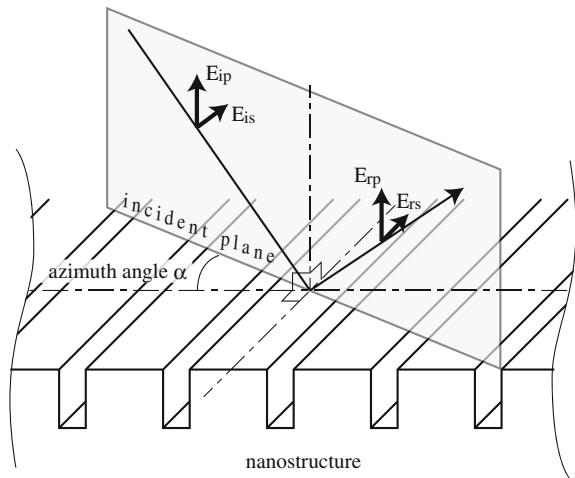
where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of m'. The phase difference component M_R can be calculated from Eqs. (8.9) and (8.12).

Using the above relations, the Mueller matrix M can be decomposed into a phase difference component M_R, a depolarization component M_Δ and an absorption component M_D.

8.3.4 Rigorous Analysis

Rigorous solutions to the Maxwell equations have to be obtained in order to calculate the fluctuations in electromagnetic waves caused by nanostructures. Existing calculation methods include the finite differential time domain method and rigorous coupled-wave analysis (see also Chap. 7) [9]; however, in this chapter we discuss a method where numerically calculated [[fluctuations in]] electromagnetic waves are substituted for the elements of the Mueller matrix. Denoting the Jones vectors of

Fig. 8.1 Light reflected at the
nanostructure

incident and reflected light as E_i and E_r, the electromagnetic fields incident to and
reflected from the sample are shown in Fig. 8.1.

First, the zero-order electromagnetic waves in the direction of each axis of the
xyz coordinate system in the numerical model are calculated as shown in Fig. 8.1.
Denoting the respective electric fields in each direction as E_x, E_y and E_z, the respec-
tive electric field components E_p^* and E_s^* in p and s direction can be expressed as
follows by combining the electric field vectors in the xyz coordinate system:

$$E_p^* = E_x + E_z, \tag{8.17}$$

$$E_s^* = E_y. \tag{8.18}$$

Here, the p and s components E_p and E_s with respect to the incidence plane can
be written as follows by using Eq. (8.18) and the rotation matrix A of the incidence
azimuth angle α^o:

$$E_p = AE_p^*, \tag{8.19}$$

$$E_s = AE_s^*. \tag{8.20}$$

Next, the electric field components corresponding to the p and s components of
the incident electric field in Fig. 8.1 are denoted as E_{ip} and E_{is}, and their respec-
tive amplitudes and phases are r_{ip}, δ_{ip} and r_{is}, δ_{is}. Furthermore, the electric field
components corresponding to the p and s components of the reflected electric field
are denoted as E_{rp} and E_{rs}, with respective amplitudes and phase differences r_{rp},
δ_{rp} and r_{rs}, δ_{rs}. Consequently, the relation between the incident electric field E_i and
the reflected electric field E_r can be written as follows by using the Jones matrix T
together with the amplitude ratio r_{nm} and phase difference δ_{nm} of the incident and

reflected electric fields.

$$T = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix} \tag{8.21}$$

$$= \begin{bmatrix} \left[\dfrac{|r_{rp}|\exp(i\delta_{rp})}{|r_{ip}|\exp(i\delta_{ip})} \right]_{E_{is}=0} & \left[\dfrac{|r_{rp}|\exp(i\delta_{rp})}{|r_{is}|\exp(i\delta_{is})} \right]_{E_{ip}=0} \\ \left[\dfrac{|r_{rs}|\exp(i\delta_{rs})}{|r_{ip}|\exp(i\delta_{ip})} \right]_{E_{is}=0} & \left[\dfrac{|r_{rs}|\exp(i\delta_{rs})}{|r_{is}|\exp(i\delta_{is})} \right]_{E_{ip}=0} \end{bmatrix}. \tag{8.22}$$

Next, the Mueller matrix can be expressed through the Jones matrix T and the transformation matrix A:

$$M = A < TT^* > A, \tag{8.23}$$

where the transformation matrix A is

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix}. \tag{8.24}$$

Solving Eq. (8.23), the Mueller matrix M is obtained as

$$A = \begin{bmatrix} \frac{1}{2}(E_1 + E_2 + E_3 + E_4) & \frac{1}{2}(E_1 - E_2 - E_3 + E_4) & F_{13} + F_{42} & -G_{13} - G_{42} \\ \frac{1}{2}(E_1 - E_2 + E_3 - E_4) & \frac{1}{2}(E_1 + E_2 - E_3 - E_4) & F_{13} - F_{42} & -G_{13} + G_{42} \\ F_{14} + F_{32} & F_{14} - F_{32} & F_{12} + F_{34} & -G_{12} + G_{34} \\ G_{14} + G_{32} & G_{14} - G_{32} & G_{12} + G_{34} & F_{12} - F_{34} \end{bmatrix}, \tag{8.25}$$

where

$$E_i = T_i T_i^* = |T_i|^2, \tag{8.26}$$

$$F_{ij} = -F_{ij} = Re(T_i T_j^*) = Re(T_j T_i^*), \tag{8.27}$$

$$G_{ij} = -G_{ji} = Im(T_i T_j^*) = -Im(T_j T_i^*), \tag{8.28}$$

$$i, j = 1, 2, 3, 4. \tag{8.29}$$

Furthermore, the elements of the matrix can be expressed through the amplitude ratio and the phase difference as follows:

$$m_{00} = \frac{1}{2}(|r_{pp}|^2 + |r_{ss}|^2 + |r_{sp}|^2 + |r_{ps}|^2), \tag{8.30}$$

$$m_{01} = \frac{1}{2}(|r_{pp}|^2 - |r_{ss}|^2 - |r_{sp}|^2 + |r_{ps}|^2), \tag{8.31}$$

$$m_{02} = |r_{pp}||r_{sp}|\cos(\delta_{pp} - \delta_{sp}) + |r_{ps}||r_{ss}|\cos(\delta_{ps} - \delta_{ss}), \qquad (8.32)$$

$$m_{03} = -|r_{pp}||r_{sp}|\sin(\delta_{pp} - \delta_{sp}) - |r_{ps}||r_{ss}|\sin(\delta_{ss} - \delta_{ps}), \qquad (8.33)$$

$$m_{10} = \frac{1}{2}(|r_{pp}|^2 - |r_{ss}|^2 + |r_{sp}|^2 - |r_{ps}|^2), \qquad (8.34)$$

$$m_{11} = \frac{1}{2}(|r_{pp}|^2 + |r_{ss}|^2 - |r_{sp}|^2 - |r_{ps}|^2), \qquad (8.35)$$

$$m_{12} = |r_{pp}||r_{sp}|\cos(\delta_{pp} - \delta_{sp}) - |r_{ps}||r_{ss}|\cos(\delta_{ps} - \delta_{ss}), \qquad (8.36)$$

$$m_{13} = -|r_{pp}||r_{sp}|\sin(\delta_{sp} - \delta_{pp}) + |r_{ps}||r_{ss}|\sin(\delta_{ss} - \delta_{ps}), \qquad (8.37)$$

$$m_{30} = |r_{pp}||r_{ps}|\sin(\delta_{pp} - \delta_{ps}) + |r_{sp}||r_{ss}|\sin(\delta_{sp} - \delta_{ss}), \qquad (8.38)$$

$$m_{31} = |r_{pp}||r_{ps}|\sin(\delta_{pp} - \delta_{ps}) - |r_{sp}||r_{ss}|\sin(\delta_{sp} - \delta_{ss}), \qquad (8.39)$$

$$m_{32} = |r_{pp}||r_{ss}|\sin(\delta_{ss} - \delta_{pp}) + |r_{sp}||r_{ps}|\sin(\delta_{ps} - \delta_{sp}), \qquad (8.40)$$

$$m_{33} = |r_{pp}||r_{ss}|\cos(\delta_{pp} - \delta_{ss}) - |r_{sp}||r_{ps}|\cos(\delta_{sp} - \delta_{ps}). \qquad (8.41)$$

This method allows us to calculate the Mueller matrix for the incident and reflected electric fields for nanoscale structures.

8.4 Application of the Mueller Matrix Polarimeter to Nanostructures

In this section, some experimental results measured by a double rotating retarder polarimeter after decomposition which agree well with the values obtained by numerical analysis based on RCWA analysis are shown [9–12]. Figure 8.2 shows a surface profile and a typical cross sectional profile of a subwavelength structure measured by

(a) 3D-view   **(b)** cross-sectional view
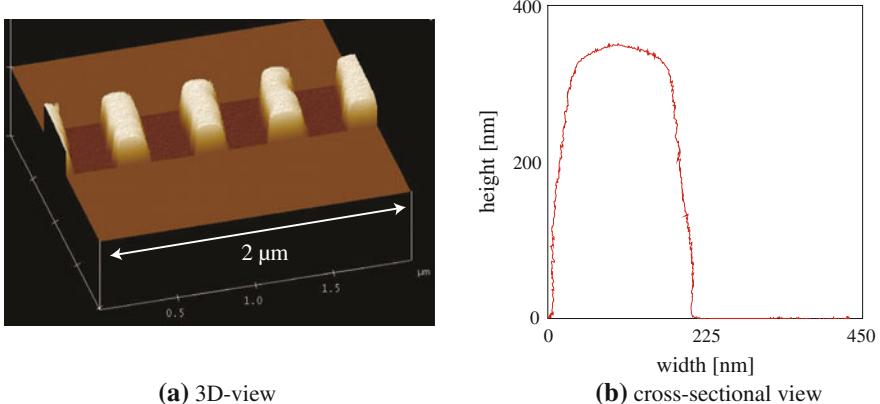
Fig. 8.2 Cross sectional profile of subwavelength structure for anti-refraction obtained by CD-AFM
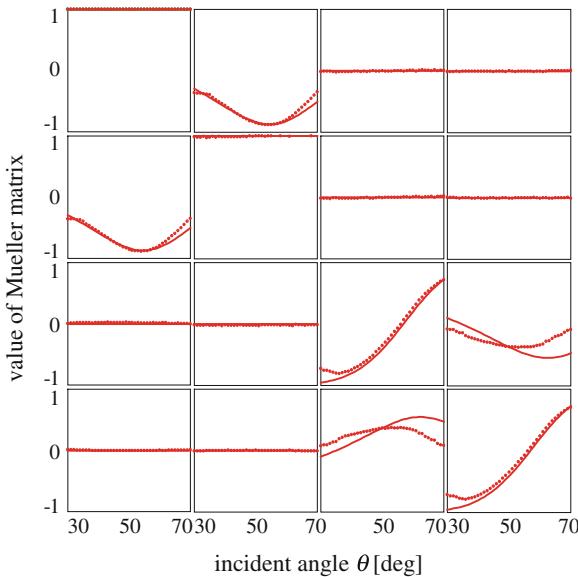
Fig. 8.3 Mueller matrix of subwavelength structure before decomposition. *Dots* are experimental results and *solid lines* are simulation results

a critical dimensional atomic force microscope. The structure was manufactured by nanoimprinting process. A material of the structure is PMMA and it is suitable for anti-reflection of the light of 633 nm. The structure has a width of 200 nm, a height of 355 nm, and a period of 433 nm. Using the obtained profile, a simulation model for RCWA was constructed.

Figure 8.3 shows an undecomposed Mueller matrix of the subwavelength structure detected by the Mueller matrix polarimeter using He-Ne laser as a light source. The experimental results are not in agreement with the simulation results, especially, the elements of m_{01}, m_{10}, m_{22}, m_{23}, m_{32} and m_{33}, because there are some structural fluctuation as shown in Fig. 8.2.

Figures 8.4 and 8.5 show diattenuation and depolarization of the measured Mueller matrix, respectively. In Fig. 8.4, results obtained by experimental and simulation are in good agreement. However, in Fig. 8.5, the experimental results are not agreement with the simulation results. The simulation results represent the unit vector which means there are no depolarization effect because there are no fluctuation in the simulation model. In other words, there are manufacturing errors in the subwavelength structure. There is a possibility of estimation for the frequently errors of the subwavelength structure less than 0.1λ.

The Mueller matrix for the subwavelength structure was measured by double rotating polarimeter and was calculated using RCWA. Comparing the experimental results with the calculated values demonstrated the effectiveness using Mueller matrix to detect the surface of the subwavelength structure. By decomposition of the Mueller matrix, the frequently error of the subwavelength structure has been

Fig. 8.4 Diattenuator Mueller matrix of subwavelength structure after decomposition. *Dots* are experimental results and *solid lines* are simulation results
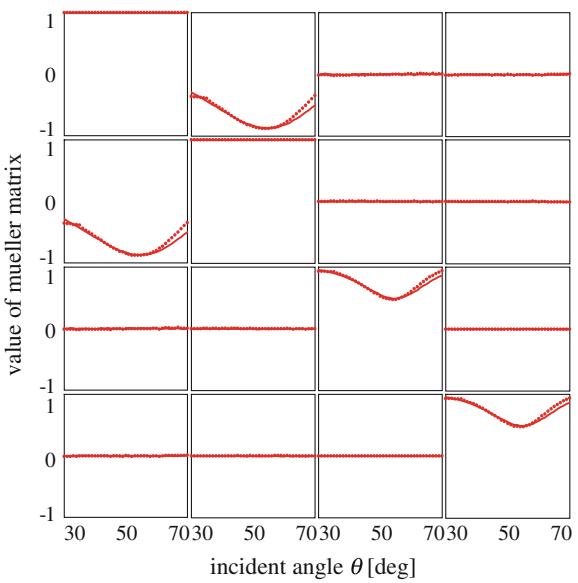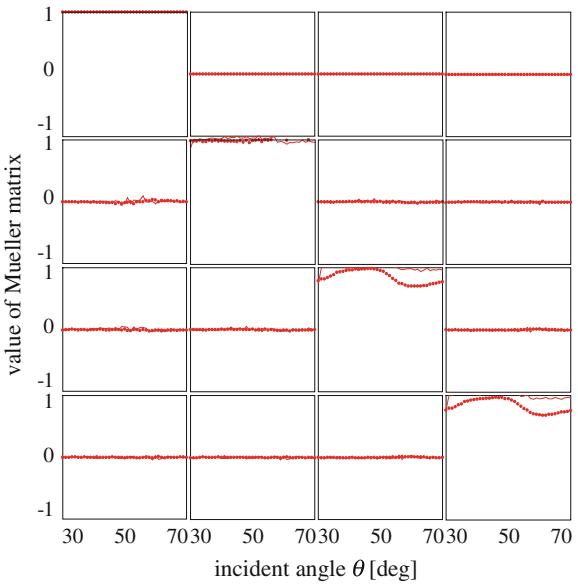

Fig. 8.5 Depolarization Mueller matrix of sub-wavelength structure after decomposition. *Dots* are experimental results and *solid lines* are simulation results

detected and possibility of an estimation by the depolarization Mueller matrix. This will require using an analysis that considers the depolarization due to scattering.

References

1. K.C. Hickman, S.M. Gaspar, K.P. Bishop, S.S.H. Naqvi, J.R. McNeil, G.D. Tipton, B.R. Stallard, B.L. Draper, Use of diffracted light from latent images to improve lithography control. proc. SPIE **1464**, 245–257 (1991)
2. R. Antos, I. Ohlidal, J. Mistrik, K. Murakami, T. Yamaguchi, J. Pistora, M. Horie, S. Vinovsky, Spectroscopic ellipsometry on lamellar gratings. Appl. Surf. Sci. **244**, 225–229 (2005)
3. See for example, T. Novikova, A. De Martino, S.B. Hatit, B. Drevillon, Application of Mueller polarimetry in conical diffraction for critical dimension measurements in microelectronics. Appl. Opt. **45**, 3688–3696 (2006)
4. H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (Wiley, Chichester, 2003)
5. R.M.A. Azzam, N.M. Bashara, Ellipsometry and polarized light (North-Holland Personal Library, Amsterdam, 1977)
6. H. Kikuta, Y. Ohira, H. Kubo, K. Iwata, Effective medium theory of two-dimensional subwavelength gratings in the non-quasi-static limit. J. Opt. Soc. Am. A **15**, 1577–1585 (1998)
7. C.F. Bohren, D.R. Huffman, Absorption and Scattering of Light by Small Particles. (Wiley, Weinheim, 1983)
8. S.-Y. Lu, Russell A. Chipman, Interpretation of Mueller matrices based on polar decomposition. J. Opt. Soc. Am. **13**, 1106–1113 (1996)
9. M.G. Moharam, T.K. Gaylord, Rigorous coupled-wave analysis of planar-grating diffraction. J. Opt. Soc. Am. **71**, 811–818 (1981)
10. R.M.A. Azzam, Photopolarimetric measurement of the Mueller matrix by Fourier analysis of a single detected signal. Opt. Lett. **6**, 146–150 (1978)
11. D. Goldstein, *Polarized Light* (Marcel Dekker Inc., New York, 2003)
12. Y. Mizutani, U. Yoshiyuki, K. Tomohito, O. Yukitoshi, U. Nrihiro, Detection of subwavelength structure profile by decomposition of Mueller matrix, in *XIX IMEKO World Congress Fundamental and Applied Metrology*, pp. 618–619, 2009

Chapter 9
Spectroscopic Ellipsometry and Magneto-Optical Kerr Spectroscopy of Magnetic Garnet Thin Films Incorporating Plasmonic Nanoparticles

Satoshi Tomita

Abstract Ferrimagnetic yttrium iron garnet (YIG) thin films incorporating plasmonic Au nanoparticles were prepared. Magneto-optical (MO) Kerr spectroscopy was carried out. The polar MO Kerr spectra in wavelength between 400 and 800 nm show that, by incorporating the Au nanoparticles, Kerr rotation angles become negative values in the region where the localized surface plasmon (LSP) resonance of the Au nanoparticles is located. Spectroscopic ellipsometry was performed and complex electric permittivity (dielectric function, ϵ) of the films was obtained. An anomalous dispersion of Re$[\epsilon]$ is clearly observed in the visible region originating from LSP of Au nanoparticles. The influence of ϵ on the MO Kerr properties of the nanocomposite films and the physics underlying the anomalous Kerr rotation are discussed.

9.1 Introduction

Localized surface plasmon (LSP) is a coupled mode of electromagnetic waves and collective oscillations of free electrons in nanostructures of noble-metals, for example, gold (Au) or silver (Ag). The LSP is accompanied by an optical near-field [1]. Manipulating the optical near-field at a submicronic scale and enhancing the local field in nanostructured metals is now one of the central problems in the growing field of plasmonics, which is of importance for surface-enhanced Raman scattering (SERS) [2], electromagnetic waveguide [3], sub-wavelength lithography [4], and non-linear optical effect [5]. When the noble-metal nanoparticles are embedded in a magneto-optical (MO) medium, it is expected that the LSP of the nanoparticles is

S. Tomita (✉)
Graduate School of Materials Science,
Nara Institute of Science and Technology (NAIST),
8916-5 Takayama, Ikoma, Nara 630-0192, Japan
e-mail: tomita@ms.naist.jp

coupled to the MO effects of the medium, leading to a modification and enhancement of the MO properties; this might be considered as an MO counterpart of SERS [6].

Coupling between MO effects and bulk plasma of charge carriers in metals was initially investigated by Feil and Haas [7] and by Katayama et al. [8]. Safarov et al. [9] studied an MO Kerr property of Au/Co/Au multilayer structures in total reflection geometry, and demonstrated a strong enhancement of the MO figure of merit of the system. This indicates that the traveling surface plasmon polariton (SPP) generated in the Au layers successfully coupled to the MO effect of the Co layer via an evanescent field. Such surface-enhanced MO effects in multilayer systems have been studied intensively owing to their potential application for MO media [10–13]. On the other hand, little is known about an MO effect coupled with the LSP of noble-metal nanoparticles via an optical near-field, although such kind of contribution has been theoretically predicted [6, 14].

We consider plasmonic metal nanoparticles, which are MO-inactive, to be included in an MO-active medium. Owing to the large specific surface of the nanoparticles, a small volume fraction of the particles effectively influences MO properties of the surrounding medium. In contrast to the SPP on planar metal surfaces, the LSP of the nanoparticles can be excited directly by the traveling light wave. Frequency of the LSP depends on the size, shape, and arrangement of the nanoparticles, as well as the kind of metal. This results in a tailored LSP resonance located in a wide range from visible to near-infrared. Moreover, the penetration length of the near-field by the LSP depends not on the wavelength of the incident light but on the size of the nanoparticles. These enable us to study how the near-field contribution influences the MO properties and to provide an insight into the physics underlying the coupling between the LSP and MO effects.

Ferrimagnetic $Y_3Fe_5O_{12}$ (yttrium iron garnet: YIG) is a well known insulating MO medium, which is almost transparent in the visible region above 500 nm [15]. In this work, we prepared YIG thin films incorporating Au nanoparticles, the LSP resonance of which is located at about 600 nm [16, 17], as described in Sect. 9.2. Volume fraction of Au (f_{Au}) was increased up to 27 %. In Sect. 9.3, the structures of the films were studied using X-ray diffraction and transmission electron microscopy. Polar MO Kerr effects of the films were studied in the visible region in Sect. 9.4. It was found that, by incorporating the Au nanoparticles at least $f_{Au} = 4.8$ %, the Kerr rotation angles at about 600 nm show negative values, indicating a possible coupling between the MO effect of YIG and the LSP of Au nanoparticles [16]. In order to study the influence of the permittivity tensor on the MO Kerr properties, we carried out spectroscopic ellipsometry (SE) for the films [17] in Sect. 9.5. SE is the sole technique to evaluate the permittivity (ϵ) of films [18–20]. The diagonal part of ϵ around 600 nm of YIG thin films with f_{Au} of 4.8 % is almost identical to that of pure YIG. This suggests that the negative Kerr angle observed in the polar MO Kerr spectra originates from the modification of the off-diagonal part due to the excitation of the LSP of Au nanoparticles. The physics underlying the anomalous Kerr rotation is discussed in Sect. 9.6. Section 9.7 concludes this chapter.

MO properties of matter have been addressed mainly from chemistry so far. For example, Bi was doped into YIG to enhance the MO effect [21, 22]. On the other hand,

Table 9.1 Deposition conditions and microstructures of samples

f_{Au} (%)	d_{sp} (min)	t_{sp} (nm)	t_{el} (nm)
0	90	531	359
4.8	90	531	380
27	30	177	136

f_{Au} is the volume fraction of Au in the films. d_{sp} is duration of deposition. t_{sp} is the film thickness estimated using d_{sp}. t_{el} is the film thickness evaluated from ellipsometry

in this study the MO effect is modified by introducing artificial nanostructures of Au. Artificial materials consisting of subwavelength-sized units are called metamaterials [23]. Metamaterials manifest several intriguing properties, for example, negative index of refraction [24, 25], electromagnetic cloaking [26], and narrow-band perfect absorption [27], which are never observed in nature. A key concept of metamaterials is an assignment of distinct functions to different units. When realizing negative index of refraction by metamaterials [28], one may assign magnetic resonance to split-ring resonators [29] and electric response to metallic cut-wires [30]. In this context, the composite consisting of Au nanoparticles and YIG matrices created in this study is a sort of metamaterial for MO effect—MO effect are assigned to YIG matrices, and the modification are assigned to Au nanoparticles instead of Bi. The assignment to different units may help us to tune the wavelength of modification and obtain tailored MO effects by changing the size, shape, and arrangement of the Au nanoparticles, as well as the kind of metal.

9.2 Film Preparation

A mixture of Au and YIG was deposited onto quartz substrates (1 mm in thickness) using rf co-sputtering [31] of a YIG target (4 in. in diameter) with Au chips in an Ar gas atmosphere (20 mTorr). The sputtering power was 50 W. The sputtering rate of YIG (s_{YIG}) measured using a quartz microbalance was 5.9 nm/min. The number of Au chips was varied in order to obtain different volume fractions of Au [f_{Au} (%)] in the films. In this work, we study three films: with f_{Au} values of 0, 4.8, and 27%. The atomic ratio of Au and Y in the films (R_Y/R_{Au}) was measured using electron-probe microanalysis (JEOL JXA-8200). The f_{Au} was evaluated with

$$f_{Au} = \frac{3 \times 100(M_{Au}/\rho_{Au})}{3(M_{Au}/\rho_{Au}) + (R_Y/R_{Au})(M_{YIG}/\rho_{YIG})}, \tag{9.1}$$

where $M_{Au} = 197$ (the atomic weight of Au) and $M_{YIG} = 738$ (the molecular weight of YIG). $\rho_{Au} = 19.28\,\text{g/cm}^3$ and $\rho_{YIG} = 4.45\,\text{g/cm}^3$ are the densities of Au and YIG, respectively.

Table 9.1 shows the f_{Au}, duration of deposition [d_{sp} (min)], and film thickness [t_{sp} (nm)] (roughly estimated from $s_{YIG} \times d_{sp}$). The thickness was decreased in
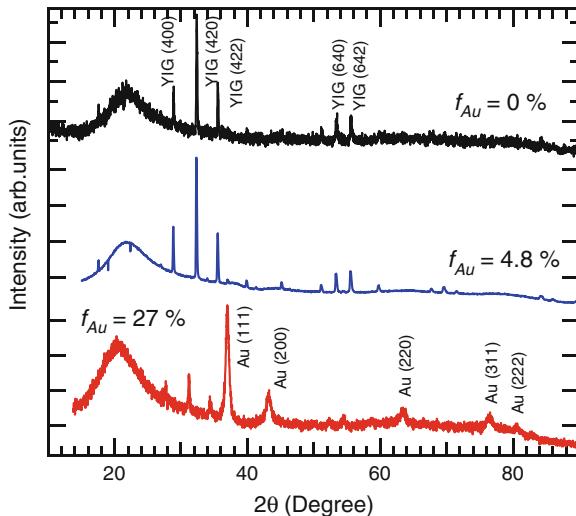
Fig. 9.1 XRD profiles for films with $f_{Au} = 0$ (*black line*), 4.8 (*blue line*), and 27% (*red line*) annealed at 900°C

proportion to the f_{Au} increase because, in order to study the ultraviolet-visible (UV-Vis) transmission spectroscopy on the samples, optical density of the films should be kept almost constant. All the films were annealed in N_2 gas at 900°C for 30 min. During the post-annealing, YIG matrices were crystallized and Au nanoparticles were grown. We obtained crystalline YIG thin films containing metallic Au nanoparticles on quartz substrates.

9.3 Sample Structures

Figure 9.1 shows X-ray diffraction (XRD) profiles of the films annealed at 900°C. MAC Science M21X with a Cu Kα radiation ($\lambda_{CuK\alpha} = 1.54$ Å) was used. A XRD profile of a film without Au, i.e., $f_{Au} = 0\%$, (black line) is consistent with that of YIG having the garnet structure. This suggests that deposited YIG is crystallized and a polycrystalline YIG film is formed by annealing at 900°C. The XRD profile of a film with $f_{Au} = 4.8\%$ (blue line) exhibits diffraction peaks, where positions agree with those of polycrystalline YIG film. However, it shows no traces of Au. This is probably caused by the low volume fraction of Au in the film. In a profile of a film with $f_{Au} = 27\%$ (red line), five diffraction peaks assigned to fcc-Au emerge in addition to weak signals corresponding to YIG, demonstrating that metallic Au is grown and crystalline YIG and Au coexist in the film.

A cross-section of the film was observed with a transmission electron microscope (TEM) operated at 200 kV (JEOL JEM-2000EX). The specimens for the TEM obser-
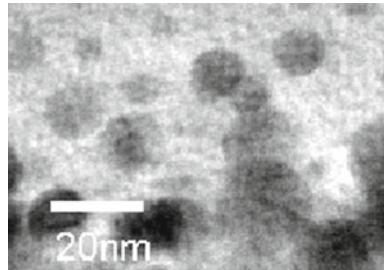
**Fig. 9.2** A cross-sectional TEM image of a film with $f_{Au} = 27\%$
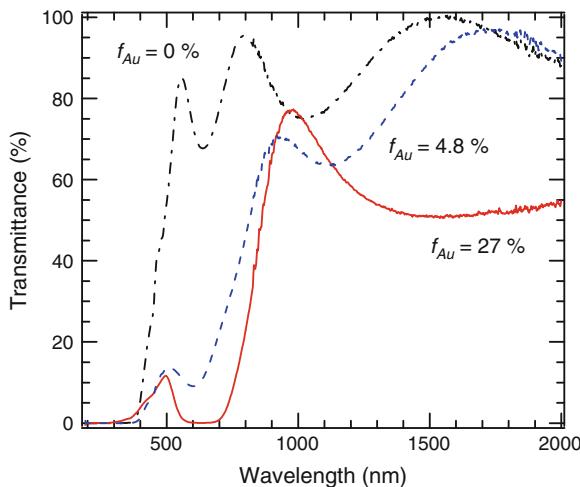
Fig. 9.3 UV-Vis transmission spectra of films with $f_{Au} = 0\%$ (*black dot-dashed curve*), 4.8% (*blue dashed curve*), and 27% (*red solid curve*)

vations were prepared through standard procedures including mechanical and Ar-ion thinning techniques. Figure 9.2 shows a cross-sectional TEM image of a film with $f_{Au} = 27\%$. We see that spherical particles are embedded randomly in the film. These are most likely Au nanoparticles. The average diameter of the Au nanoparticles observed was about 12.3 nm. The XRD and TEM studies allow us to depict that, in a film with $f_{Au} = 27\%$, spherical metallic Au nanoparticles about 12 nm in diameter are embedded in the crystalline YIG matrices.

Figure 9.3 shows UV-Vis transmission spectra for the films in wavelength (λ) ranging from 190 to 2000 nm. The spectra were obtained with a double-beam-type spectrometer (Shimadz UV-3101PC). A YIG film with $f_{Au} = 0\%$ (black dash-dotted curve) shows strong absorption below 500 nm corresponding to a charge transfer (CT) type electron transition in YIG [22]. Contrastingly, the film is almost transparent above 500 nm even though there are weak dips at 620 and 1000 nm caused by interference.

Incorporation of Au with $f_{Au} = 4.8\%$ (blue dashed curve) results in a strong absorption at about $\lambda = 600$ nm. In a spherical small inclusion with permittivity $\epsilon_i(\lambda)$ embedded into a medium with $\epsilon_m(\lambda)$, a surface mode (Fröhlich mode) is excited at λ_F, where $Re[\epsilon_i(\lambda_F)] = -2Re[\epsilon_m(\lambda_F)]$ is satisfied [32]. If we adopt permittivity of Au [33] to $\epsilon_i(\lambda)$ and that of YIG [15] to $\epsilon_m(\lambda)$, λ_F is calculated to be about 600 nm. Therefore, the strong absorption originates from the LSP of Au nanoparticles in YIG matrices. This indicates that a film with $f_{Au} = 4.8\%$ contains metallic Au nanoparticles, which are similar to those in a film with $f_{Au} = 27\%$, in the crystalline YIG matrices although an XRD profile shows no traces of Au. The absorption at about 600 nm becomes stronger with increasing f_{Au} up to 27 % as shown by a red solid curve.

9.4 MO Kerr Spectroscopy

The polar MO Kerr spectra at room temperature were measured in λ ranging from 400 to 800 nm using a polarization plane modulation technique with a Faraday cell. A Xe lamp was used for the light source. The monochromatic light was illuminated on a film under the magnetic field applied up to ± 15 kOe in a perpendicular direction to the film plane. The Kerr angle was determined from a rotation angle at about 3 kOe because magnetization hysteresis curves obtained with a superconducting quantum interference device magnetometer (not shown here) indicated that the film magnetization was saturated at about 3 kOe. The resolution of the Kerr rotation angle was about $0.001°$.

Figure 9.4 shows the Kerr spectra of the samples with $f_{Au} = 0\%$ (black crosses), 4.8 % (blue triangles), and 27 % (red circles). For a YIG film ($f_{Au} = 0\%$), we see a large Kerr rotation and strong oscillation of the spectrum at about 450 nm, which is probably due to a CT transition at $\lambda = 490$ nm in YIG. Above $\lambda = 500$ nm, although the Kerr angle still oscillates, it retains positive values between 0.07 and $0.01°$. This is thought to be caused by a tail of the CT transition. For the $f_{Au} = 4.8\%$ sample, the Kerr spectra below $\lambda = 500$ nm is similar to that for $f_{Au} = 0\%$. Contrastingly, the spectrum above 500 nm is much different from the control spectrum. In particular, it should be noted that the Kerr rotation in the $f_{Au} = 4.8\%$ sample changes the sign at $\lambda = 530$ nm and the rotation angles become negative in λ ranging between 540 and 600 nm. As f_{Au} increases up to 27 %, the negative Kerr angle appears in a wider range of λ from 520 to 710 nm although the oscillation below $\lambda = 500$ nm is smeared out.

Figure 9.5 shows the Kerr loops between ± 15 kOe obtained for the $f_{Au} = 27\%$ sample. In Fig. 9.5a, the loop at $\lambda = 400$ nm exhibits a positive Kerr angle of about $0.04°$ at 3 kOe. On the other hand, the Kerr loop at $\lambda = 650$ nm (Fig. 9.5b) shows a negative Kerr angle of about $-0.01°$ although the value is relatively small. These results verify that the polarization plane of the incident light rotates into the opposite direction between 400 and 650 nm in the $f_{Au} = 27\%$ sample.
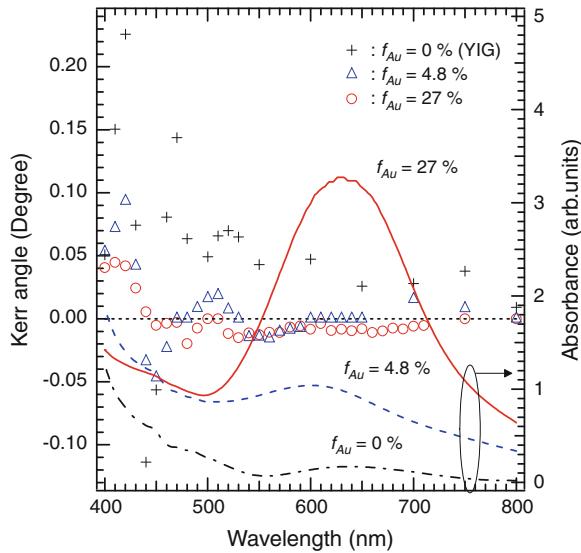
Fig. 9.4 Polar Kerr spectra for films with $f_{Au} = 0\%$ (*black crosses*), 4.8% (*blue triangles*), and 27% (*red circles*). Absorption spectra in wavelength ranging from 400 to 800 nm are also shown for $f_{Au} = 0\%$ (*black dot-dashed curve*), 4.8% (*blue dashed curve*), and 27% (*red solid curve*)
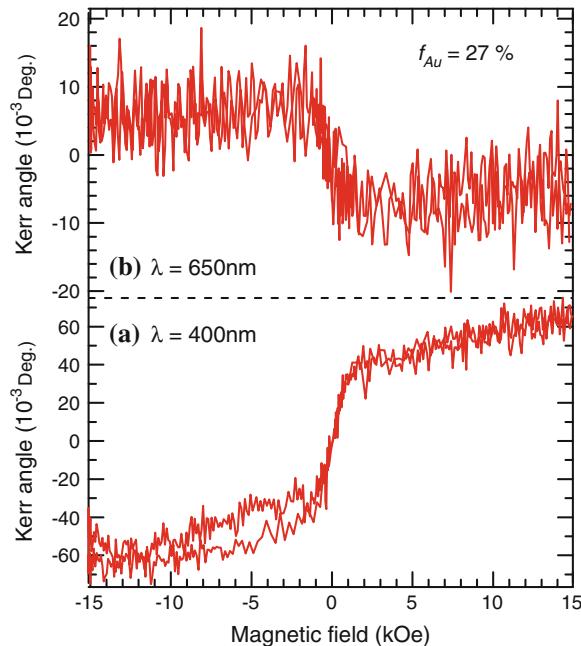

Fig. 9.5 Kerr loops of a film with $f_{Au} = 27\%$ at wavelength of 400 nm (**a**) and 650 nm (**b**)

In Fig. 9.4, the absorption spectra converted from the transmission spectra in Fig. 9.3 are also exhibited. The absorbance at λ [$A(\lambda)$] was calculated through an equation, $A(\lambda) = -log_{10}[T(\lambda)/100]$, where $T(\lambda)$ is the transmittance at λ. The LSP in Au nanoparticles is excited in the vicinity of 602 nm for the $f_{Au} = 4.8\%$ sample and of 634 nm for the $f_{Au} = 27\%$ sample. An increase in f_{Au} leads to a red-shift of the LSP absorption peak [34]. It should be noticed here that the position of the absorption peaks due to the LSP excitation is consistent with the region, where the Kerr rotation angles assume negative values. A red shift of the absorption peak for $f_{Au} = 27\%$ brings about the negative Kerr rotation angle at a longer wavelength up to 710 nm. These point out a possible coupling between the MO Kerr effects of YIG and the LSP of Au nanoparticles.

9.5 Spectroscopic Ellipsometry

The angle of the polar MO Kerr rotation is phenomenologically determined by both the diagonal (ϵ_{xx}) and off-diagonal part (ϵ_{xy}) of electric permittivity tensor. The ϵ_{xx} of the films is modified by incorporating Au into YIG films and forming the composite. In order to study the influence of ϵ_{xx} on the MO Kerr properties, SE analysis was performed with a spectroscopic phase-modulated ellipsometer (HORIBA Jobin Yvon UVISEL NIR). The incident angle was 60°. A pair of ellipsometrical angles Ψ and Δ was measured for each wavelength. Detailed measurement and analysis procedures are reported elsewhere [35].

Figure 9.6 shows the measured Ψ (dotted) and Δ (crosses). Figure 9.6a corresponds to the $f_{Au} = 0\%$ YIG thin film. To analyze this, we constructed an optical model consisting of void (substrate) and three layers: quartz, (composite) film, and surface layer (from bottom to top). By representing the substrate as a void in this model, the reflection at the bottom of the quartz layer was taken into consideration [36]. The quartz layer in the model is assigned to the quartz substrate in the measurement. To represent the surface roughness of the film, a surface layer several nm in thickness, using an effective medium approximation with voids of 50%, was considered. Optical constants of SiO_2 available in the database of the SE apparatus were used to represent the quartz substrate. Ψ and Δ were calculated using our optical model with a dispersion formula consisting of four oscillators for the film. The oscillators were Lorentz oscillators for the pure YIG film. In order to evaluate the thickness of the YIG layer from ellipsometry [t_{el} (nm)], fitting was carried out by varying the values of the parameters in the dispersion formula and the t_{el} to achieve minimization of mean square error χ^2 between the measured and calculated Ψ and Δ spectra. The t_{el} (nm) for the YIG film was 359 nm. The calculated spectra (not shown here) matched those that were measured ($\chi^2 = 1.0523$).

Using the calculated thickness and the measured Ψ and Δ, point-by-point calculation was performed to obtain directly the refractive index n and extinction coefficient κ. Ψ and Δ spectra after point-by-point calculation are shown by solid lines in Fig. 9.6a. The matching between measured and calculated spectra was better
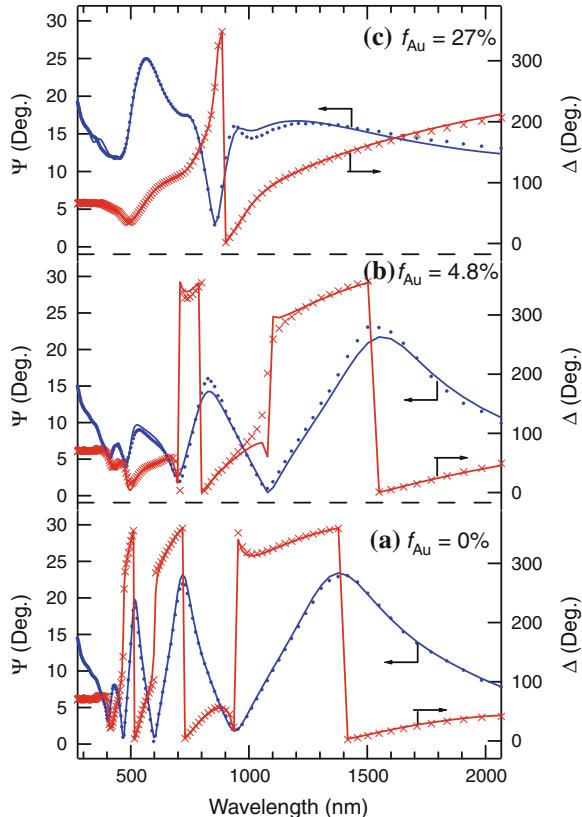
Fig. 9.6 Measured Ψ (*dotted*, *blue*) and Δ (*crosses*, *red*) of (**a**) a pure YIG thin film and (**b–c**) YIG thin films with Au nanoparticles of f_{Au} = 4.8, and 27.0%. *Solid lines* correspond to the point-by-point calculated spectra

after the point-by-point calculation (χ^2 = 0.9367). From the obtained n and κ, we evaluated the complex ϵ using $\epsilon = (n + i\kappa)^2$. Real and imaginary parts of the ϵ are plotted in Fig. 9.7a, b. The values of the Re[ϵ] of YIG in the region of 275–2000 nm are between 4.5 and 6.7. There is a broad peak around 375 nm. That peak can be assigned to charge transfer (CT) type transitions in YIG [15, 22]. Crystal field (CF) type transitions at about 610 and 700 nm of YIG were not observed owing to the parity forbidden character.

Figure 9.6b, c shows the Ψ and Δ of the YIG thin films that have Au nanoparticles with an f_{Au} of 4.8 and 27.0 %, respectively. A fitting was performed to determine the t_{el} using a dispersion formula based on a quantum-mechanical theory of absorption [37, 38], which is better than the formula with Lorentz oscillators in the quality of the fit. The quantum-mechanical approach is based on the existence of band gap energy for which κ has its absolute minimum. In addition, if the Kramers-Kronig dispersion relation is used to derive n from κ, the number of independent parameters can be
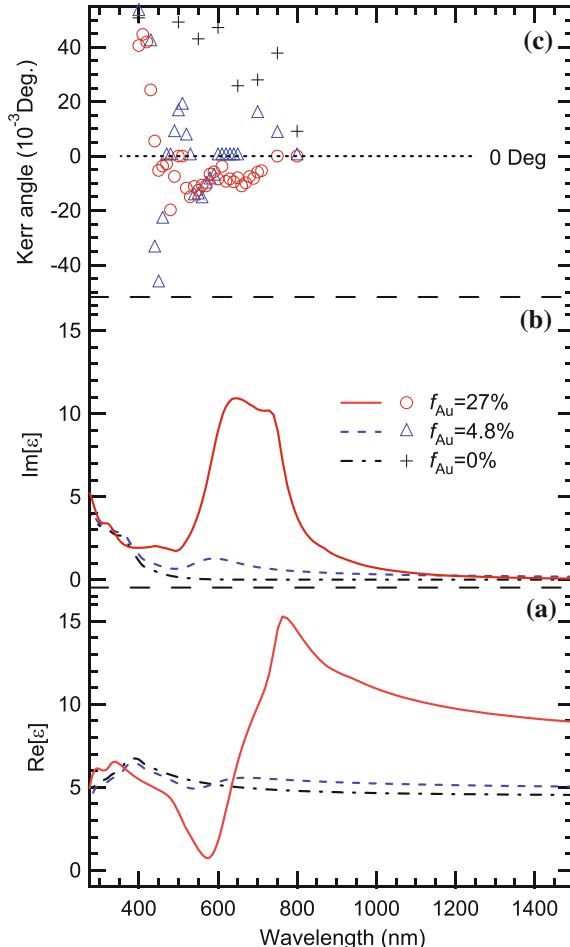
Fig. 9.7 Real (**a**) and imaginary (**b**) parts of ϵ of a YIG thin film (*black*) and composite films with $f_{Au} = 4.8$ (*blue*) and 27.0 % (*red*) as a function of wavelength. **c** MO Kerr rotation angle is also plotted

reduced significantly. This relation has been widely used because of its simplicity and ability to provide a good fit for various materials [39]. The t_{el} (nm) for the composite film with a f_{Au} of 4.8 and 27.0 % was 380 and 136 nm, respectively. The evaluated t_{el} are also summarized in Table 9.1. We see that t_{sp} is overestimated because it is roughly evaluated using s_{YIG} and the density of bulk YIG.

We derived the n and κ from the point-by-point calculation and evaluated the Re[ϵ] and Im[ϵ] as plotted in Fig. 9.7a, b. χ^2 of samples with f_{Au} of 4.8 and 27.0 % are 3.4528 and 2.2499, respectively. Figure 9.7a shows that an anomalous dispersion of the Re[ϵ] emerges around 590 nm because Au nanoparticles were incorporated. The dispersion causes a peak in the Im[ϵ] at about 590 nm, as shown in Fig. 9.7b. As

the f_{Au} increases up to 27 %, the dispersion of the Re[ϵ] and the peak in the Im[ϵ] become significant. These results suggest that the anomalous dispersion originates from the LSP of the Au nanoparticles in the YIG matrices. The spectra of the Im[ϵ] are very similar to the absorption spectra obtained by a UV-Vis spectrometer shown in Fig. 9.4.

Of special note is the region of 500–600 nm in Fig. 9.7a because the negative Kerr rotations are observed in the polar MO Kerr spectra, as shown in Fig. 9.7c. In this region, a YIG thin film shows an almost constant Re[ϵ] of about 5.0, as in Fig. 9.7a. Although Au nanoparticles of $f_{Au} = 4.8$ % were incorporated, the Re[ϵ] did not change significantly. The Re[ϵ] was around 4.5. Contrastingly, when $f_{Au} = 27.0$ %, the incorporation of Au nanoparticles brings about a significant modification of the Re[ϵ].

The complex polar MO Kerr rotation (Φ_K) is described as

$$\Phi_K = -i\epsilon_{xy}\{\sqrt{\epsilon_{xx}}(\epsilon_{xx} - 1)\}, \tag{9.2}$$

where ϵ_{xx} is the diagonal part and ϵ_{xy} is the off-diagonal part of the dielectric tensor. The equation indicates that the Kerr angle is determined by both ϵ_{xx} and ϵ_{xy}. In the present SE study, we obtained the ϵ_{xx} of the nanocomposite films. The ϵ_{xx} in the MO response is less dependent on the magnetization (M) of materials because the $\epsilon_{xx}(M)$ is developed by the even order of the M. Therefore, although somewhat over-simplified, the following is a discussion of the MO response with the use of ϵ_{xx}, which we obtained in the present SE study without an applied magnetic field [40–42].

9.6 Mechanism for Anomalous Kerr Rotation

The present experimental results show that Au nanoparticles only about 4.8 % in volume fraction influence the MO properties of YIG matrices. This is primarily due to a large specific surface of the nanoparticles. When we discuss the physics of the LSP of particles, we should take the particle size compared with the wavelength of light into consideration [32]. TEM studies showed that Au nanoparticles in the $f_{Au} = 27$ % sample were about 12 nm in average diameter. In the film with $f_{Au} = 4.8$ %, the size is much smaller. Since the size of Au particles in both films is sufficiently small compared with the wavelength, in which the Kerr rotation angles change the sign and show negative values (500–700 nm), the retardation effect of light is negligible; the LSP wavelength of the particles is not affected by the particles size. Therefore, we think that a small variation in the Au particle diameter in this study does not influence the sign of the Kerr rotation.

The Re[ϵ_{xx}] of the nanocomposites with Au nanoparticles of $f_{Au} = 4.8$ % embedded in YIG are similar to that of pure YIG thin films. It is unlikely that only such small modification causes the negative Kerr rotation in the region of 600 nm even if the multiple reflection is taken into consideration. The incorporation of Au nanopar-

ticles and excitation of the LSP probably modifies the ϵ_{xy} around 590 nm of the films [14, 43].

From the microscopic point of view, ϵ_{xy} is related to the electron transition energy, oscillator strength of the transition, and spin-orbit (SO) splitting of the excited states [22]. It is known that YIG in the visible region has one CT transition at 490 nm and two crystal field (CF) type transitions at about 610 and 700 nm. The CT transition is responsible for an MO Kerr rotation in the visible region due to the large oscillator strength, while the CF transitions at 610 and 700 nm do not contribute. This is because of intra-$3d$ transition nature having even parity of the CF transitions; owing to the parity forbidden character, the oscillator strength of the CF transitions is very weak. The oscillator strength is, however, enhanced when electric-dipole-allowed excitation, which can be mixed to relieve the parity constraint, lie close by in energy [15]. In the present study, the excited LSP (Fröhlich mode) in Au particles can be considered as an electric dipole. The excitation energy of the LSP lies close by that of the CF transitions of YIG. This may relieve the parity constraint and enhance the oscillator strength of the CF transitions, resulting in an anomalous Kerr rotation in the vicinity of 600 nm.

Furthermore, a modification of SO coupling of the excited states in the CT transition is possible. In a Bi substituted rare-earth iron garnet, the MO Faraday rotation angles in the visible region become negative with increasing Bi concentration [21, 44]. The negative rotation was believed to be caused by the change of the sign of the SO coupling constants in the excited CT states in the iron garnet due to the substitution by Bi [22]. However, in the present study, the formation of an Au substituted YIG can be ruled out from the XRD study. The Hamiltonian of the SO coupling is described as

$$H_{SO} = (\hbar/4mc^2)(\nabla V \times \vec{p}) \cdot \vec{\sigma}, \tag{9.3}$$

where V is the external potential, $\vec{p}$ is the momentum, and $\vec{\sigma}$ is the Pauli spin operator [45, 46]. This equation indicates that the external electric potential V may influence the SO coupling in the magnetized materials. As already mentioned above, the electromagnetic energy concentration in the resonant state of the LSP causes a strong optical near-field around the nanoparticles even with modest input power. It is thus plausible that the LSP accompanied by an enhanced electric field modifies the sign of the effective SO coupling constant in the excited CT states of YIG in the vicinity of the interfaces with Au nanoparticles.

Recently Belotelov et al. [47] reported a theoretical consideration for enhanced MO Faraday and Kerr effects in bilayer heterostructures consisting of an Au film, periodically perforated with hole arrays, deposited on a thin uniform magnetic garnet film. The periodically perforated Au film is similar to the two-dimensional plasmonic crystal. By choosing an appropriate thickness of the magnetic garnet, the edges of the photonic band gaps for both principal modes coincide, leading to the enhancement in Faraday effect with high optical transmittance. Moreover, they studied experimentally transverse MO Kerr effect for magnetic garnet films with Au grating [48]. An enhanced Kerr effect with this artificial material was observed. These works, however, utilize traveling SPP.

Very recently, several experimental works concerning enhanced MO effects using LSP were reported [49, 50]. Jain et al. [49] demonstrated a Faraday rotation enhancement at 530 nm using LSP in Au-coated iron oxide nanoparticles, which is chemically synthesized. The excitation of LSP in the core-shell type nanoparticles provides strong electric dipoles spectrally and spatially close to a spin-flip electron pair transition that has a weak oscillator strength. They suggested that the LSP could lead to an increase in the transition strength, allowing this transition to contribute the MO response. The true origin of the LSP-enhanced MO effect is still under debate.

9.7 Conclusion

In conclusion, we have studied the MO Kerr spectroscopy and spectroscopic ellipsometry of YIG thin films incorporating Au nanoparticles prepared using a co-sputtering technique. TEM and XRD studies suggested that metallic Au nanoparticles are grown in the YIG matrices. UV-Vis transmission spectroscopy showed the LSP resonance of Au nanoparticles at about 600 nm. The polar MO Kerr spectra obtained at wavelength ranging from 400 to 800 nm demonstrated that, by incorporating the Au nanoparticles, the Kerr rotation angles exhibit negative values in the vicinity of 600 nm. These results indicate a plausible coupling between the MO effect of YIG and the SPP through the near-field contribution.

Spectroscopic ellipsometry analysis was used to evaluate the diagonal part of permittivity tensor (ϵ_{xx}) of the composite films. In general, the incorporation of Au affects the ϵ_{xx}. However, the ϵ_{xx} in the region of 500–600 nm of YIG thin films with f_{Au} of 4.8 % is almost identical to that of pure YIG. This suggests that the observed negative Kerr angle in the polar MO Kerr spectra originates from the modification of the off-diagonal part (ϵ_{xy}) due to the excitation of the LSP of Au nanoparticles. As f_{Au} was increased up to 27.0 %, the ϵ_{xx} of the nanocomposite films nearly reaches unity around 560 nm.

An enhancement of the oscillator strength of the CF transitions in YIG is plausible for the anomalous Kerr rotation. A modification of effective SO coupling of YIG via the strong optical near-field is also a possible mechanism. The present study suggests a possibility of controlling magnetic properties of nanocomposite magnetic materials via an optical near-field.

References

1. R. Luppin, in *Electromagnetic Surface Modes*, ed. by A.D. Boardman (Wiley, Chichester, 1982)
2. S. Nie, S.R. Emory, Science **275**, 1102 (1997)
3. S.A. Maier, P.G. Kik, H.A. Atwater, S. Meltzer, E. Harel, B.E. Koel, A.A.G. Requicha, Nat. Mater. **2**, 229 (2003)
4. X. Luoad, T. Ishihara, Appl. Phys. Lett. **84**, 4780 (2004)
5. S. Kim, J. Jin, Y.-J. Kim, I.-Y. Park, Y. Kim, S.-W. Kim, Nature **453**, 757 (2008)
6. V.A. Kosobukin, SPIE Proc. **2535**, 9 (1995)
7. H. Feil, C. Haas, Phys. Rev. Lett. **58**, 65 (1987)
8. T. Katayama, Y. Suzuki, H. Awano, Y. Nishihara, N. Koshizuka, Phys. Rev. Lett. **60**, 1426 (1988)
9. V.I. Safarov, V.A. Kosobukin, C. Hermann, G. Lampel, J. Peretti, C. Marlière, Phys. Rev. Lett. **73**, 3584 (1994)
10. T.J. Silva, S. Schultz, D. Weller, Appl. Phys. Lett. **65**, 658 (1994)
11. C. Hermann, V.A. Kosobukin, G. Lampel, J. Peretti, V.I. Safarov, P. Bertrand, Phys. Rev. B **64**, 235422 (2001)
12. J. Bremer, V. Vaicikauskas, F. Hansteen, O. Hunderi, J. Appl. Phys. **89**, 6177 (2001)
13. N. Bonod, R. Reinisch, E. Popov, M. Nevière, J. Opt. Soc. Am. B **21**, 791 (2004)
14. M. Abe, T. Suwa, Phys. Rev. B **70**, 235103 (2004)
15. F.J. Kahn, P.S. Pershan, J.P. Remeika, Phys. Rev. **186**, 891 (1969)
16. S. Tomita, T. Kato, S. Tsumnashima, S. Iwata, M. Fujii, S. Hayashi, Phys. Rev. Lett. **96**, 167402 (2006)
17. S. Tomita, M. Fujii, S. Hayashi, A. Terai, N. Nabatova-Gabain, Jpn. J. Appl. Phys. **46**, L1032 (2007)
18. D. Dalacu, L. Martinu, J. Appl. Phys. **87**, 228 (2000)
19. D. Dalacu, L. Martinu, J. Opt. Soc. Am. B **18**, 85 (2001)
20. J.C.G. de Sande, R. Serna, J. Gonzalo, C.N. Afonso, D.E. Hole, A. Naudon, J. Appl. Phys. **91**, 1536 (2002)
21. H. Takeuchi, Jpn. J. Appl. Phys. **14**, 1903 (1975)
22. K. Shinagawa, in *Magneto-Optics*, ed. by S. Sugano, N. Kojima (Springer, Berlin, 1999)
23. D.R. Smith, J.B. Pendry, M.C.K. Wiltshire, Science **305**, 788 (2004)
24. V.G. Veselago, Sov. Phys. Usp. **10**, 509 (1968)
25. J.B. Pendry, Phys. Rev. Lett. **85**, 3966 (2000)
26. D. Schurig, J.J. Mock, B.J. Justice, S.A. Cummer, J.B. Pendry, A.F. Starr, D.R. Smith, Science **314**, 977 (2006)
27. X. Liu, T. Tyler, T. Starr, A.F. Starr, N.M. Jokerst, W.J. Padilla, Phys. Rev. Lett. **107**, 045901 (2011)
28. D.R. Smith, W.J. Padilla, D.C. Vier, S.C. Nemat-Nasser, S. Schultz, Phys. Rev. Lett. **84**, 4184 (2000)
29. J.B. Pendry, A.J. Holden, D.R. Robbins, M.J. Stewart, IEEE Trans. Microwave Theory Tech. **47**, 2075 (1999)
30. J.B. Pendry, A.J. Holden, W.J. Stewart, I. Youngs, Phys. Rev. Lett. **76**, 4773 (1996)
31. S. Tomita, M. Hagiwara, T. Kashiwagi, C. Tsuruta, Y. Matsui, M. Fujii, S. Hayashi, J. Appl. Phys. **95**, 8193 (2004)
32. C.F. Bohren, D.R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983)
33. E.D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, New York, 1997)
34. J.C. Maxwell-Garnet, Philos. Trans. R. Soc. London **203**, 385 (1904)
35. T. Tsuboi, Y. Wasai, N. Nabatova-Gabain, Thin Solid Films **496**, 674 (2006)
36. M. Kildemo, R. Ossikovski, M. Stchakovsky, Thin Solid Films **313–314**, 108 (1998)
37. A.R. Forouhi, I. Bloomer, Phys. Rev. B **34**, 7018 (1986)
38. A.R. Forouhi, I. Bloomer, Phys. Rev. B **38**, 1865 (1988)

39. G.E. Jellison Jr, F.A. Modine, Appl. Phys. Lett. **69**, 371 (1996)
40. X. Gao, D.W. Glenn, S. Heckens, D.W. Thompson, J.A. Woollam, J. Appl. Phys. **82**, 4525 (1997)
41. R.J. Lange, S.J. Lee, K.J. Kim, P.C. Canfield, D.W. Lynch, Phys. Rev. B **63**, 035105 (2000)
42. J.L. Menéndez, B. Bescós, G. Armelles, R. Serna, J. Gonzalo, R. Doole, A.K. Petford-Long, M.I. Alonso, Phys. Rev. B **65**, 205413 (2002)
43. G. Armelles, J.B. Gonzalez-Diaz1, A. Garcia-Martin, J.M. Garcia-Martin, A. Cebollada, M.U. Gonzalez, S. Acimovic, J. Cesario, R. Quidant, G. Badenes, Opt. Express **16**, 16104 (2008)
44. Z. Šimša, H. Le Gall, J. Šimšova, J. Koláček, A. Le Paillier-Malécot, IEEE Trans. Magn. **MAG-20**, 1001 (1984)
45. S. LaShell, B.A. McDougall, E. Jensen, Phys. Rev. Lett. **77**, 3419 (1996)
46. E.I. Rashba, Physica E **20**, 189 (2004) (and references there in)
47. V.I. Belotelov, L.L. Doskolovich, A.K. Zvezdin, Phys. Rev. Lett. **98**, 077401 (2007)
48. V.I. Belotelov, I.A. Akimov, M. Pohl, V.A. Kotov, S. Kasture, A.S. Vengurlekar, A.V. Gopal, D.R. Yakovlev, A.K. Zvezdin, M. Bayer, Nat. Nanotechnol. **6**, 370 (2011)
49. P.K. Jain, Y. Xiao, R. Walsworth, A.E. Cohen, Nano Lett. **9**, 1644 (2009)
50. G.X. Du, T. Mori, M. Suzuki, S. Saito, H. Fukuda, M. Takahashi, Appl. Phys. Lett. **96**, 081915 (2010)

Chapter 10
Generalized Ellipsometry Characterization of Sculptured Thin Films Made by Glancing Angle Deposition

Daniel Schmidt, Eva Schubert and Mathias Schubert

Abstract Generalized ellipsometry, a non-destructive optical characterization technique, is employed to determine geometrical structure parameters and anisotropic dielectric properties of highly spatially coherent three-dimensionally nanostructured thin films in the spectral range from 400 to 1700 nm. The analysis of metal slanted columnar thin films fabricated by glancing angle deposition reveals their monoclinic optical properties and their optical response can be modeled with a single homogeneous biaxial layer. This homogeneous biaxial layer approach is universally applicable to sculptured thin films and effective optical properties of the nanostructured thin films are attained. We provide a nomenclature and categorization for sculptured thin films based on their geometry and structure. A piecewise homogeneous biaxial layer approach is described, which allows for the determination of principal optical constants of chiral and achiral multi-fold and helical sculptured thin films. It is confirmed that such sculptured thin films have modular optical properties. This characteristic can be exploited to predict the optical response of sculptured thin films grown with arbitrary sequential substrate rotations. As an alternative model approach, an anisotropic effective medium approximation based on the Bruggeman formula is presented, which provides results comparable to the homogeneous biaxial layer approach and in addition provides the volume fraction parameters for slanted columnar thin films and their depolarization factors.

10.1 Introduction

Nanotechnology in the twenty-first century enabled revolutions in the fields of information technology, cellular, and molecular biology with profound impact on our economy and society. Progress in the interdisciplinary field of nanotechnology

D. Schmidt (✉) · E. Schubert · M. Schubert
University of Nebraska-Lincoln, Lincoln, NE, USA
e-mail: schmidt@engr.unl.edu

allowed for miniaturization of electronic components leading to portability of affordable products with improved functionality. Engineering research in nanotechnology provided and continuous to provide the key component for further technological enhancements.

Today's nanotechnology is mostly a planar (two-dimensional) technology. Another technology leap is foreseen with appropriate utilization of the third dimension employing self-assembled nanostructures as building blocks, for example. Sophisticated techniques and growth processes lead to self-organized three-dimensional nanostructures and novel materials and phenomena are currently incorporated into next generation micro- and nanosystems. The fabrication of metallic nanostructures with tailored geometry and material is a very interesting challenge in nanotechnology because geometrical and material parameters are responsible for the resulting optical, electrical, mechanical, chemical, or magnetic properties of such structures.

Amongst the emerging technologies for fabrication of metallic nanostructures is a physical vapor deposition process called glancing angle deposition. The particular growth geometry combined with dynamic substrate movement allows for in-situ sculpturing of self-organized, highly spatially coherent, three-dimensional achiral and chiral geometries at the nanoscale. This type of nanostructure growth was demonstrated using many elemental and compound materials and, in principle, can be achieved with any material that can be sputtered or evaporated. Such engineered nanostructured materials, which have been termed sculptured thin films (STFs), constitute a new realm of solid state materials, and carry a huge potential for applications in the fields of nano-photonics [1], nano-electromechanics [2], nano-magnetics [3], nano-electromagnetics [4], and nano-sensors [5–8]. For example, tailored effective optical constants by controlling porosity and shape of the nanostructured films are highly desirable for many applications such as broadband antireflection coatings [9, 10], omnidirectional reflectors [11, 12], Bragg reflectors [13, 14], optical resonators [15], light emitting diodes [16], and optical interconnects [17]. The controllable porosity and the large surface area may also be considerably beneficial for existing technologies such as solar cells [18–20] and thin film batteries [21]. Ferromagnetic sculptured thin films, in particular, exhibit interesting magnetic phenomena due to anisotropic structure effects and hence can be exploited for new magnetic or magneto-optical storage media [22–25].

An entire new material class is envisioned when combining the inorganic nanostructures with functionalized polymers or chemical and biological recognition elements thereby creating nanohybrid functional materials. The new nanohybrids are anticipated to offer unmatched tunability in terms of electronic, optical, mechanical, ferroelectric, magnetic, and magneto-optical properties, thereby opening the door to a new family of sensing principles and, ultimately, new classes of ultra sensitive, broad range, portable, inexpensive sensors and detectors [26–28].

Figure 10.1 depicts conceptualized areas of interest for sculptured thin films. Hollow-core nanohelices fabricated with an inner diameter matching dimensions of helical viral protein shells (capsids with typical diameters in the range of 5–30 nm) may become useful for purification and detection. Nanostructure surfaces can be coated with self-assembled monolayers and functional groups to facilitate viral
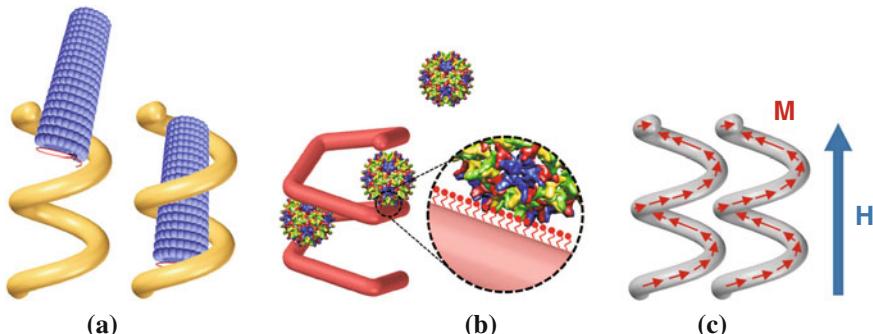
Fig. 10.1 Schematic illustration of novel application areas for sculptured thin films in sensing and detection, and nanomagnetism: **a** selective capsid (protein shell of virus) capturing in hollow-core nanohelices with matched dimensions, **b** viral attachment on bio-functionalized nanoscaffold surfaces, and **c** chiral magnetic domain alignment in ferromagnetic nanohelices (Capsid in (**a**) modified from [29], and Hepatitis B virus in (**b**) adapted from [30])

attachment. The adsorbed biomaterial will change the anisotropic optical response of the functionalized hybrid nanostructures and can be detected by generalized ellipsometry, for example [7]. Nanomagnetism and the arrangement and switching of chiral magnetic fields is in the scope of interest also because ferromagnetic nanostructures might have the ability, upon application of external magnetic fields, to control capture or release of modified adsorbates [24, 25].

In order to systematically utilize sculptured thin films in future applications, however, physical properties of these nanosized objects need to be understood such that targeted geometry engineering with tailored properties from selected materials will be possible. Non-invasive and non-destructive optical techniques are preferred, however, due to the complexity of sculptured thin films, optical characterization is a challenge. Spectroscopic generalized ellipsometry within the Mueller matrix formalism is the most general polarization-dependent linear-optical spectroscopy approach and provides an excellent tool to determine the dielectric functions of anisotropic optical systems. Generalized ellipsometry allows for characterization of sculptured thin films of arbitrary geometry and materials upon analyzing the anisotropic polarizability response.

This section elucidates fundamental linear optical properties of complex sculptured thin films in the visible and near infrared spectral region and is organized as follows:

The physical vapor glancing angle deposition technique used for fabrication of the sculptured thin films is presented in Sect. 10.2. Also, a nomenclature for STFs based on their geometry and structure is introduced.

The concept of spectroscopic generalized ellipsometry is presented in Sect. 10.3. Necessary mathematical formalisms to describe light propagation in stratified media are outlined and the treatment of the external electromagnetic plane wave response of an optical system using the Mueller matrix formalism is given. Model approaches for analysis of ellipsometry data, valid for complex sculptured thin films, are dis-

cussed. Sections 10.4 and 10.5 summarize experimental and structural parameters for samples discussed here. Optical properties of metal sculptured thin films of type F1-STFS are discussed in Sect. 10.6. Metal F1-STFs are found to possess monoclinic optical properties. A model is proposed that explains the origin of the monoclinicity due to the specific spatial arrangement of the nanostructures. It is demonstrated that the homogeneous biaxial layer approach is a universally valid approach for slanted columnar thin films. Subsequently, optical properties of more complex STFs are presented. It is discussed why the optical plane wave response of STFs can be reduced to the determination of the optical constants of the individual building blocks. These building blocks can be assembled in a modular conception mimicking the true geometry, and the optical properties of the film can be predicted from this model arrangement. It is further discussed how optical properties of sculptured thin films are influenced upon ambient changes. The section ends with a comparison of results obtained with the homogeneous biaxial layer approach and anisotropic Bruggeman effective medium approximation calculations.

The present chapter concludes with a summary of the fundamental findings and a brief outlook in Sect. 10.7.

10.2 Fabrication of Sculptured Thin Films: Glancing Angle Deposition

Glancing angle deposition (GLAD) is a bottom-up fabrication technique that employs a physical vapor deposition (PVD) process at oblique angles where the trajectory of the incoming particle flux is not parallel to the substrate normal. The technique allows to engineer the columnar structure of PVD grown films and is a promising self-organized fabrication process for micro- and nanotechnology. The three-dimensionally shaped, highly orientationally coherent but randomly distributed nanostructured thin films are called sculptured thin films (STFs). The GLAD process is schematically shown in Fig. 10.2.

The first report on growth of metallic thin films by PVD at oblique angles with a stationary substrate was published more than 120 years ago. Kundt [31], credited for growing the first slanted columnar thin films (F1-STFs, see Table 10.1), observed birefringence in these metal thin films and concluded that the optical anisotropy was due to the microstructure.

In 1950, König and Hellwig [32] recognized the self-shadowing mechanism responsible for a columnar microstructure developing during deposition at oblique angles. The incident atoms stochastically condense on the substrate and form nucleation clusters. At oblique angles, due to physical shadowing at the atomic scale, a competing three-dimensional growth of these clusters starts since no incoming particles can reach the geometrically shadowed area. Given favorable conditions, such as limited adatom mobility (surface diffusion) and collimated particle flux, the resulting thin film consists of self-organized, highly spatially coherent, slanted nanocolumns. The columns are oriented toward the vapor source, however, the growth direction is

not parallel to the incoming vapor flux but rather tilted toward the substrate normal [33, 34].

Another important step toward the emergence of three-dimensionally shaped thin films was reported by Young and Kowal [35] in 1959. The authors introduced a continuous substrate rotation around the substrate normal during deposition at incident angles of $30° < \theta_i < 60°$ to realize chiral[1] polarization filters. Substrate rotation is equivalent to a constant angular change of the incoming particle flux direction and hence equivalent to an apparent rotation of the vapor source around the substrate normal [36, 37]. This changes the shadowing dynamics and the column growth follows the perceived change in source location.

Robbie et al. [38] demonstrated that at very oblique incident angles ($\theta_i > 60°$) highly porous nanostructured thin films can be fabricated with densities as low as 15 % bulk and coined the term glancing angle deposition. At glancing angles of $\theta_i > 80°$ and in combination with a controlled substrate motion distinct nanostructures can be "sculpted" in-situ [39].

Depending on the azimuthal substrate motion, different STF geometries and combinations thereof can be achieved: simple slanted columnar thin films (F1-STFs) will form at oblique angles with no substrate rotation. If the substrate is rotated stepwise at fixed growth intervals chevron-like (lF2-STFs; 180° steps)[2] or staircase-like with a square footprint (lF4$^\pm$-STFs; 90° steps) can be fabricated, for example. Slow and continuous rotation will result in chiral hollow-core helical sculptured thin films (tH$^\pm$-STF) where the rotation speed determines the inner diameter of the H-STFs (Fig. 10.2). The pitch of the helices is a measure of the vertical periodicity and defined as the vertical distance between two adjacent windings. As the angular velocity of the substrate rotation is increased the inner diameter and the pitch decreases until eventually solid-core screw-like structures will form. Yet faster rotation speed where
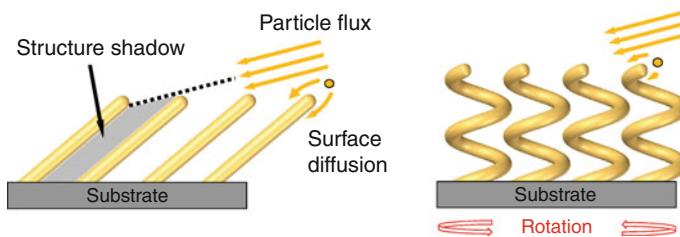

Fig. 10.2 Schematic drawing of two representative GLAD situations in case of lateral ordering (deposited on patterned substrate). (*left*) With a steady substrate and an obliquely incident particle flux highly spatially coherent F1-STFs (slanted columns) will grow. (*right*) Slow and continuous substrate rotation, which is equivalent to a steady change in the direction of the incoming particle flux, results in H-STFs (hollow-core nanohelices) (Reprinted with permission from [40] Copyright 2010, Schmidt)

---

[1] The term chiral is derived from the Greek word for hand and is used to describe an object that is non-superposable on its mirror image.

[2] New nomenclatures are introduced for different sculptured thin film geometries. See Table 10.1.

Table 10.1 Proposed nomenclature for STFs derived from basic building block configurations

Proposed nomenclature	Chiral	Footprint	Example geometry	Description
F1-STF	No			(Slanted) columnar thin film
l F2-STF	No			Chevron or zig-zag; example shown: 3F2
l F4$^{\pm}$-STF	Yes			Four-fold staircase; example shown: 9F4^{+}
t H$^{\pm}$-STF	Yes			Helical thin film; example shown: 2H^{-}
V-STF	No			(Vertical) columnar thin film

The x in Fx stands for the number of equally spaced rotation steps within one full substrate turn and with same rotation sense during fabrication. l is an integer number (>1) and denotes the number of layers. Chiral STFs, starting with 2F3$^{\pm}$-STF (three-fold symmetry, not shown here) and higher-order folds, have to be additionally characterized by their handedness—indicated by a '$+$' for right-handed (clockwise) and a '$-$' for left-handed (counterclockwise). For helical (continuously rotated) STFs (H$^{\pm}$-STF), t indicates the number of turns

the pitch becomes smaller than the column diameter will result macroscopically in the loss of the helical geometry, and the structure degenerates into vertically oriented columns (V-STFs) [33, 34]. An additional degree of freedom can be introduced by also changing the deposition angle during growth, which alters the lateral density of the respective STF [6, 41].

Sculptured thin films by GLAD can be fabricated from a wide variety of materials, including insulators, metals, semiconductors, and organic materials, vaporized by sputtering [42–44], pulsed laser deposition [45], thermal or (most commonly) electron beam evaporation [33, 34, 46, 47]. Electron-beam evaporation is particularly favorable since the impinging atoms have very low energy (<1 eV) and larger amounts of material can be vaporized at constant conditions compared to thermal evaporation. STFs are fabricated under low-adatom-mobility conditions, where the

sticking coefficient (ratio of adsorbed adatoms and total number of adatoms arriving within the same period of time) is essentially unity, and hence substrate temperatures of less than 10 % of the melting point of the evaporant are desired for columnar growth [33]. Therefore STFs can be deposited on virtually any substrate material (e.g., glass and polymers) because the substrate can be kept at room-temperature.

10.2.1 Organized In-plane Growth

GLAD on flat substrates results in random in-plane distribution of nanostructures with a "quasi-periodic" topology because of an average intercolumnar spacing due to the shadowing characteristics [48, 49]. The random in-plane distribution originates from the self-organized growth due to the stochastic condensation process on the substrate surface and the subsequent competing growth mechanism. Lateral coherence can be achieved when depositing on patterned substrates. Patch- or dot pattern, for example, may determine initial shadowing conditions and serve as nucleation and condensation seeds for the incoming particle flux [50]. Electron-beam lithography is a widespread method for prepatterning substrates, however, only small areas can be patterned (in the order of $100 \times 100 \, \mu m^2$), it is costly, and the resulting seeds are >20 nm in diameter [51–54]. Self-assembly large-scale patterning techniques such as nanosphere or diblock-copolymer nanolithography are advantageous over electron beam lithography because smaller seed sizes can be achieved and yet they are more economical [55–61].

Diblock-copolymer nanolithography.   Diblock-copolymer or micelle nanolithography is a self-assembly process to place metal nanodots in a regular pattern onto a flat substrate. Diblock copolymers comprise a polar and non-polar polymer block dissolved in a non-polar solvent. Once a certain concentration is reached these diblock copolymers aggregate into inverse micelles thereby forming a core-shell structure. The micelle nanoreactor permits selective dissolution of metal salt into the polar micelle core. A dip-coating process step with a flat substrate allows for surface coating of a regular monolayer of metal-loaded micelles. After solvent evaporation, the organic part of the film is selectively etched away by an oxygen plasma treatment leaving the inorganic nanodot pattern behind. The micelle diameter and hence interparticle spacing can be controlled by the size of the block copolymers whereas the nanodot size is predominantly controlled by the amount of metal salt added to the micelle solution [56, 57, 60, 62].

10.2.2 Proposed Sculptured Thin Film Nomenclature

Reports on a wide variety of differently shaped STFs have been published in recent years, and each research group has used their own terminology to describe the specific shape of the STF under investigation. Therefore, a universal nomenclature scheme

for STFs based on their building blocks is proposed here to unambiguously identify the STF geometry (Table 10.1). Further simplification for large Fx-STFs with many substrate rotations can be achieved by using the first unit as the building block and the number of repetitions as a subscript in analogy to structure formulas in organic chemistry: for example, a four-fold staircase with five full turns, 20F4$^+$-STF, may be written as (4F4$^+$)$_5$. A combination of different geometries can be described by concatenation of individual building blocks. This nomenclature is used throughout this chapter.

10.3 Generalized Spectroscopic Ellipsometry

Spectroscopic ellipsometry (SE) determines the complex-valued ratio ρ of linearly independent electric field components of polarized electromagnetic plane waves; i.e., the change of the polarization state of an electromagnetic plane wave upon interaction with a sample. Traditionally, this ratio is measured in reflection or transmission for light polarized parallel (p), and perpendicular (s)[3] to the plane of incidence and expressed by the two real-valued ellipsometric parameters Ψ and Δ [63][4]:

$$\rho = \left(\frac{B_p}{B_s}\right) / \left(\frac{A_p}{A_s}\right) = \tan \Psi e^{i\Delta}, \tag{10.1}$$

where amplitudes A stand for incident and B for exiting plane waves with p- and s-polarization components with respect to the plane of incidence as defined in the reflection arrangement in Fig. 10.3.[5] The absolute value of the complex ratio is defined by $\tan \Psi$, and Δ denotes the relative phase change of the p and s components of the electric field vector [64–67].

The complex ratio ρ can be addressed within different presentations of the electromagnetic plane wave response. Also, depending on the sample properties, i.e., for anisotropic samples, which cause mode conversion between p- and s-polarized light upon reflection (or transmission), the ellipsometric parameter set must be further expanded into the so-called generalized ellipsometry parameter set. In such cases the Jones matrix presentation provides a convenient and sufficient frame. However, depolarization of light upon interaction with an optical system cannot be treated with the Jones formalism. In this case the Mueller matrix presentation is the appropriate choice.

---

[3] The abbreviation "s" comes from the German word senkrecht for perpendicular.

[4] Unless used unambiguously as running index, the symbol "i" addresses the imaginary unit $\sqrt{-1}$.

[5] Considerations are given for a reflection set up, but hold for the complex-valued ratio of polarized plane wave components in the transmission arrangement as well.
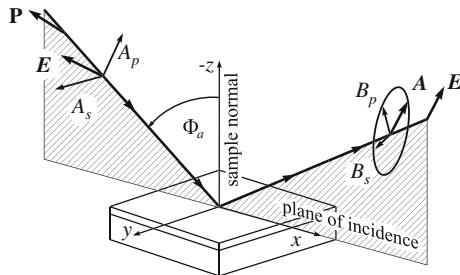
Fig. 10.3 The wavevectors of the incident and emerging plane waves (incident and reflected at an angle Φ_a) and the sample normal define the plane of incidence (x-z plane). A_p, A_s, B_p, and B_s, denote the complex amplitudes of the p and s modes before and after reflection, respectively. P and A are the azimuth angles of the linear polarizers used, for example, in the standard arrangement of rotating-analyzer (polarizer) ellipsometers (Redrawn from Ref. [68])

10.3.1 Definition of the Optical Constants

The ellipsometric quantities Ψ and Δ are related to wave optics through a solution of the wave equation: $\mathbf{E} = E_0 \exp\{i\mathbf{k}\mathbf{r}\}$, with E_0 being the amplitude of the electric-field intensity $\mathbf{E}$ at spatial variable $\mathbf{r}$. The propagation vector $\mathbf{k}$ is a function of the complex-valued refractive index (optical constants) of the medium $N = n + ik$. The refractive index n follows experimentally from Snell's law:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2, \tag{10.2}$$

where θ_j is the angle of incidence counted toward the interface between two materials with n_1 and n_2. The extinction coefficient k is connected to the absorption, measured by intensity ($I = \mathbf{E}\mathbf{E}^*$) loss upon wave propagation over a distance d,

$$I = I_0 \exp\{-\alpha' d\}, \tag{10.3}$$

with the absorption coefficient α' being

$$\alpha' = \frac{4\pi}{\lambda} k. \tag{10.4}$$

Refractive index n and extinction coefficient k, also known as "propagation constants", are defined for propagating waves along direction $\mathbf{k}$ in a material, and for a given direction $\mathbf{E}$, such that both n and k would occur as in (10.2) and (10.4). As will be discussed later, for materials with monoclinic and triclinic optical properties such experiments cannot be designed, instead, coupling of propagating waves with different propagation constants will occur in general. Generalized ellipsometry (GE) is the only appropriate tool to differentiate by experiment between the intrinsic propagation constants, the refractive index and extinction coefficient for major polarizability axes $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, and to determine axes a, b, and c.

10.3.2 Jones and Mueller Matrix Presentation

The Jones Matrix Presentation. For non-depolarizing samples, the so-called Jones matrix provides a complete mathematical description for the electromagnetic plane wave response and allows for ellipsometric data analysis [63, 68–71].

The Jones reflection matrix $\mathbf{J}$ in the $p - s$ reference frame, for a sample with plane parallel boundaries, connects the incident A modes (p, s) with emergent B plane wave modes (p, s):

$$\begin{pmatrix} B_p \\ B_s \end{pmatrix} = \mathbf{J} \begin{pmatrix} A_p \\ A_s \end{pmatrix} = \begin{pmatrix} r_{pp} & r_{ps} \\ r_{sp} & r_{ss} \end{pmatrix} \begin{pmatrix} A_p \\ A_s \end{pmatrix}. \tag{10.5}$$

The Jones matrix $\mathbf{J}$ contains four complex-valued elements, which are also known as the anisotropic Fresnel reflection coefficients.[6] The off-diagonal elements of $\mathbf{J}$ are nonzero for optical systems that convert p into s waves and vice versa.

The Mueller Matrix Presentation. An alternative description of the polarized sample response is the Mueller matrix and the Stokes vector formalism. This approach is equivalent to the Jones matrix formalism for non-depolarizing samples, and can furthermore completely account for depolarization. The four real-valued Stokes parameters[7] (S_j, $j = 0 \ldots 3$) of an electromagnetic plane wave are defined in terms polarized intensities:

$$S_0 = I_p + I_s, \tag{10.6a}$$
$$S_1 = I_p - I_s, \tag{10.6b}$$
$$S_2 = I_{45} - I_{-45}, \tag{10.6c}$$
$$S_3 = I_{\sigma+} - I_{\sigma-}, \tag{10.6d}$$

where I_p, I_s, I_{45}, I_{-45}, $I_{\sigma+}$, and $I_{\sigma-}$ denote the intensities for the $p-$, $s-$, $+45°$, $-45°$, right-, and left-handed circularly polarized light components, respectively [63, 72]. The degree of polarization D_P for any state of polarization can be expressed by the Stokes parameters as [73]

$$D_P = \frac{I_{pol}}{I_{tot}} = \frac{(S_1^2 + S_2^2 + S_3^2)^{1/2}}{S_0}, \qquad 0 \le D_P \le 1, \tag{10.7}$$

where I_{pol} is the intensity of the sum of polarization components and I_{tot} is the total intensity of the beam. A value of $D_P = 1$ corresponds to completely polarized light, $D_P = 0$ corresponds to unpolarized light, and $0 < D_P < 1$ corresponds to

---

[6] In this notation the first index denotes the incident polarization mode, and the second index refers to the outgoing polarization mode.

[7] The Stokes parameters have dimensions of intensities.

partially polarized light. Arranging the Stokes parameters into a column vector, the Mueller matrix then describes the changes of each quantity upon interaction of the electromagnetic plane wave with an optical system[8] as

$$
\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{out} = \begin{pmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{in} . \tag{10.8}
$$

The advantage of the Mueller matrix concept is the ability to handle situations with partial polarization of the electromagnetic plane wave. Further details, the application to ellipsometry, and the relation to the Jones concept have been outlined previously by Azzam and Bashara [63], Röseler [72], and Jellison [74–76].

In a rotating-analyzer-ellipsometer, for example, the Mueller matrix elements of the 4th row and the 4th column cannot be measured. However, this does not impair the accessibility of the normalized Jones matrix elements, (10.5), except for its relative phase, which can only be obtained by including compensator(s) [67, 72].

For a non-depolarizing system, a one-to-one relation exists between matrices **J** and **M** [63]:

$$
M_{11} = \frac{1}{2} \left(r_{pp} r_{pp}^* + r_{ss} r_{ss}^* + r_{sp} r_{sp}^* + r_{ps} r_{ps}^* \right), \tag{10.9a}
$$

$$
M_{12} = \frac{1}{2} \left(r_{pp} r_{pp}^* - r_{ss} r_{ss}^* - r_{sp} r_{sp}^* + r_{ps} r_{ps}^* \right), \tag{10.9b}
$$

$$
M_{13} = \mathrm{Re} \left(r_{pp} r_{sp}^* + r_{ss}^* r_{ps} \right), \tag{10.9c}
$$

$$
M_{14} = \mathrm{Im} \left(r_{pp} r_{sp}^* + r_{ss}^* r_{ps} \right), \tag{10.9d}
$$

$$
M_{21} = \frac{1}{2} \left(r_{pp} r_{pp}^* - r_{ss} r_{ss}^* + r_{sp} r_{sp}^* - r_{ps} r_{ps}^* \right), \tag{10.9e}
$$

$$
M_{22} = \frac{1}{2} \left(r_{pp} r_{pp}^* + r_{ss} r_{ss}^* - r_{sp} r_{sp}^* - r_{ps} r_{ps}^* \right), \tag{10.9f}
$$

$$
M_{23} = \mathrm{Re} \left(r_{pp} r_{sp}^* - r_{ss}^* r_{ps} \right), \tag{10.9g}
$$

$$
M_{24} = \mathrm{Im} \left(r_{pp} r_{sp}^* - r_{ss}^* r_{ps} \right), \tag{10.9h}
$$

$$
M_{31} = \mathrm{Re} \left(r_{pp} r_{ps}^* + r_{ss}^* r_{sp} \right), \tag{10.9i}
$$

$$
M_{32} = \mathrm{Re} \left(r_{pp} r_{ps}^* - r_{ss}^* r_{sp} \right), \tag{10.9j}
$$

$$
M_{33} = \mathrm{Re} \left(r_{pp} r_{ss}^* + r_{ps}^* r_{sp} \right), \tag{10.9k}
$$

$$
M_{34} = \mathrm{Im} \left(r_{pp} r_{ss}^* - r_{ps}^* r_{sp} \right), \tag{10.9l}
$$

---

[8] Sample, mirrors, rotators, optical devices within the light path, and any combinations thereof.

$$M_{41} = -\mathrm{Im}\left(r_{pp}r_{ps}^* + r_{ss}^* r_{sp}\right), \qquad (10.9\mathrm{m})$$

$$M_{42} = -\mathrm{Im}\left(r_{pp}r_{ps}^* - r_{ss}^* r_{sp}\right), \qquad (10.9\mathrm{n})$$

$$M_{43} = -\mathrm{Im}\left(r_{pp}r_{ss}^* + r_{ps}^* r_{sp}\right), \qquad (10.9\mathrm{o})$$

$$M_{44} = \mathrm{Re}\left(r_{pp}r_{ss}^* - r_{ps}^* r_{sp}\right), \qquad (10.9\mathrm{p})$$

where $\{\cdot\}^*$ denotes the complex conjugate. The Mueller matrix for an isotropic sample is given by [76]

$$\mathbf{M} = \begin{pmatrix} 1 & -N_{\mathrm{M}} & 0 & 0 \\ -N_{\mathrm{M}} & 1 & 0 & 0 \\ 0 & 0 & C_{\mathrm{M}} & S_{\mathrm{M}} \\ 0 & 0 & -S_{\mathrm{M}} & C_{\mathrm{M}} \end{pmatrix}. \qquad (10.10)$$

The quantities N_{M}, S_{M}, and C_{M} provide access to the ellipsometric parameters

$$N_{\mathrm{M}} = \cos 2\Psi, \qquad (10.11\mathrm{a})$$

$$S_{\mathrm{M}} = \sin 2\Psi \sin \Delta, \qquad (10.11\mathrm{b})$$

$$C_{\mathrm{M}} = \sin 2\Psi \cos \Delta. \qquad (10.11\mathrm{c})$$

N_{M}, S_{M}, and C_{M} are not independent, and are constrained for non-depolarizing samples by the relation:

$$N_{\mathrm{M}}^2 + S_{\mathrm{M}}^2 + C_{\mathrm{M}}^2 = 1. \qquad (10.12)$$

The complex ratio ρ can be obtained from N_{M}, S_{M} and C_{M}

$$\rho = \frac{C_{\mathrm{M}} + i\,S_{\mathrm{M}}}{1 + N_{\mathrm{M}}}. \qquad (10.13)$$

10.3.3 Generalized Ellipsometry

In contrast to standard ellipsometry, in the generalized ellipsometry situation Ψ and Δ depend on the polarization state of the incident plane wave. This concept is valid within both, the Mueller matrix as well as within the Jones matrix formalism. Within the Jones matrix presentation six real-valued generalized ellipsometry angles Ψ_{ij} and Δ_{ij} are defined by three ratios of the four available complex-valued elements of the Jones reflection matrix $\mathbf{J}$:

$$R_{pp} \equiv \frac{r_{pp}}{r_{ss}} = \tan \Psi_{pp} \exp(i\Delta_{pp}), \tag{10.14a}$$

$$R_{ps} \equiv \frac{r_{ps}}{r_{ss}} = \tan \Psi_{ps} \exp(i\Delta_{ps}), \tag{10.14b}$$

$$R_{sp} \equiv \frac{r_{sp}}{r_{pp}} = \tan \Psi_{sp} \exp(i\Delta_{sp}). \tag{10.14c}$$

Note that the on-diagonal elements from the same column of the Jones matrix are used to normalize the off-diagonal matrix elements. This choice is convenient for rotating-analyzer ellipsometry [69].

The generalized ellipsometry concept is required if the response of the optical system is anisotropic, i.e., p modes are converted in s modes and vice versa. This results in non-zero off-diagonal elements of the Jones (r_{ps} and r_{sp}) and Mueller matrix (M_{kl} and M_{lk} with $k = 1, 2; \ l = 3, 4$).

10.3.4 Anisotropic Dielectric Function Tensor

10.3.4.1 General Description

In condensed matter with non-cubic symmetry, the dielectric function is represented by a complex-valued second-rank tensor $\boldsymbol{\varepsilon}$, which can be expressed in Cartesian coordinates (x, y, z):

$$\mathbf{D} = \varepsilon_0 \left(\mathbf{E} + \mathbf{P}\right) = \varepsilon_0 \boldsymbol{\varepsilon} \mathbf{E} = \varepsilon_0 \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{pmatrix} \mathbf{E}, \tag{10.15}$$

where the field-phasors displacement $\mathbf{D}$, polarization field $\mathbf{P}$, and electric field $\mathbf{E}$ are given along the unit directions $\mathbf{x}, \mathbf{y}, \mathbf{z}$ (ε_0 is the vacuum permittivity):

$$\mathbf{D} = \mathbf{x}D_x + \mathbf{y}D_y + \mathbf{z}D_z, \tag{10.16a}$$

$$\mathbf{E} = \mathbf{x}E_x + \mathbf{y}E_y + \mathbf{z}E_z, \tag{10.16b}$$
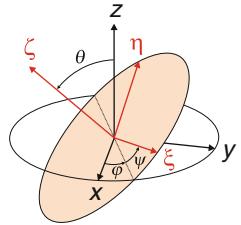
$$\mathbf{P} = \mathbf{x}P_x + \mathbf{y}P_y + \mathbf{z}P_z. \tag{10.16c}$$

In general, the dielectric function tensor $\boldsymbol{\varepsilon}$ is a function of the photon energy $\hbar\omega$ due to non-local response within the time domain (frequency dispersion). Furthermore, the $\boldsymbol{\varepsilon}$ tensor may be non-symmetric due to non-local response (chiral) in space [77].

10.3.4.2 Orthogonal Rotations

Interior and exterior Cartesian coordinate axes of a sample under consideration are related by orthogonal rotations. In order to address the $\boldsymbol{\varepsilon}$ tensor appropriately, a

Fig. 10.4 Definition of the
Euler angles φ, θ, and ψ and
the orthogonal rotations as
provided by **A**. (ξ, η, ζ), and
(x, y, z) refer to the Cartesian
auxiliary and laboratory coor-
dinate systems, respectively
[Redrawn from Ref. [68]]

right-handed Cartesian system (x, y, z) with origin at the sample surface as defined
in Fig. 10.3 is set by the plane of incidence (x, z) and the sample surface (x, y).
The real-valued Euler angles φ, θ, and ψ can be used to rotate between the Cartesian
laboratory (x, y, z) and the Cartesian auxiliary coordinate system (ξ, η, ζ) as defined
in Fig. 10.4:

$$\boldsymbol{\varepsilon}(x, y, z) = \mathbf{A}\boldsymbol{\varepsilon}(\xi, \eta, \zeta)\mathbf{A}^{-1}, \tag{10.17}$$

where the unitary matrix[9] **A** is the orthogonal rotation matrix [73]

$$\mathbf{A} = \begin{pmatrix} \cos\psi\cos\varphi - \cos\theta\sin\varphi\sin\psi & -\sin\psi\cos\varphi - \cos\theta\sin\varphi\cos\psi & \sin\theta\sin\varphi \\ \cos\psi\sin\varphi + \cos\theta\cos\varphi\sin\psi & -\sin\psi\sin\varphi + \cos\theta\cos\varphi\cos\psi & -\sin\theta\cos\varphi \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix}. \tag{10.18}$$

First, rotation φ is performed around the z-axis, then the coordinate system is rotated
by θ around the new x-axis, and a final rotation of ψ around ζ completes the coor-
dinate system rotation.

10.3.4.3 Bond Polarizability Model

Intrinsic bond polarizations (eigenvectors) set up a spatial non-Cartesian (monoclinic,
triclinic), or Cartesian (orthorhombic, tetragonal, hexagonal, trigonal, and cubic)
center-of-gravity system, with axes described by vectors $\mathbf{a} = \mathbf{x}a_x + \mathbf{y}a_y + \mathbf{z}a_z$,
$\mathbf{b} = \mathbf{x}b_x + \mathbf{y}b_y + \mathbf{z}b_z$, and $\mathbf{c} = \mathbf{x}c_x + \mathbf{y}c_y + \mathbf{z}c_z$. The linear polarization response is
additive, and may be split into

$$\mathbf{P} = \mathbf{P}_a + \mathbf{P}_b + \mathbf{P}_c, \tag{10.19}$$

where

$$\mathbf{P}_a = \varrho_a \, (\mathbf{aE}) \, \mathbf{a}, \tag{10.20a}$$

$$\mathbf{P}_b = \varrho_b \, (\mathbf{bE}) \, \mathbf{b}, \tag{10.20b}$$

$$\mathbf{P}_c = \varrho_c \, (\mathbf{cE}) \, \mathbf{c}. \tag{10.20c}$$

---

[9] Note that $\mathbf{A}^{-1} = \mathbf{A}^T$, where $\{\cdot\}^T$ denotes the transpose of a matrix.

The complex-valued scalar major polarizabilities $\varrho_a, \varrho_b, \varrho_c$ must obey Kramers-Kronig consistency, and correspond to the intrinsic center-of-gravity bond polarization system. Under restrictions to linear polarization, the second-rank susceptibility tensor χ is defined by

$$\mathbf{P} = \varepsilon_0 \chi \mathbf{E} = \varepsilon_0 \begin{pmatrix} \chi_{xx} & \chi_{xy} & \chi_{xz} \\ \chi_{yx} & \chi_{yy} & \chi_{yz} \\ \chi_{zx} & \chi_{zy} & \chi_{zz} \end{pmatrix} \mathbf{E}, \tag{10.21}$$

and the electric displacement can be written as

$$\mathbf{D} = \varepsilon_0 (1 + \chi) \mathbf{E} = \varepsilon_0 \begin{pmatrix} 1 + \chi_{xx} & \chi_{xy} & \chi_{xz} \\ \chi_{yx} & 1 + \chi_{yy} & \chi_{yz} \\ \chi_{zx} & \chi_{zy} & 1 + \chi_{zz} \end{pmatrix} \mathbf{E}. \tag{10.22}$$

Accordingly, the corresponding part of the symmetric[10] dielectric function tensor ε is easily deconvoluted by expanding (10.22) and (10.20) [68]:

$$\varepsilon_{xx} = 1 + a_x a_x \varrho_a + b_x b_x \varrho_b + c_x c_x \varrho_c, \tag{10.23a}$$
$$\varepsilon_{xy} = a_x a_y \varrho_a + b_x b_y \varrho_b + c_x c_y \varrho_c, \tag{10.23b}$$
$$\varepsilon_{xz} = a_x a_z \varrho_a + b_x b_z \varrho_b + c_x c_z \varrho_c, \tag{10.23c}$$
$$\varepsilon_{yy} = 1 + a_y a_y \varrho_a + b_y b_y \varrho_b + c_y c_y \varrho_c, \tag{10.23d}$$
$$\varepsilon_{yz} = a_y a_z \varrho_a + b_y b_z \varrho_b + c_y c_z \varrho_c, \tag{10.23e}$$
$$\varepsilon_{zz} = 1 + a_z a_z \varrho_a + b_z b_z \varrho_b + c_z c_z \varrho_c. \tag{10.23f}$$

Note that $\varepsilon_{ij} = \varepsilon_{ji}$ and that (10.23) describe the most general form of an anisotropic dielectric symmetric tensor. For anisotropic materials, besides the coordinates of the unit axes $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, three polarizability functions ϱ_j $(j = a, b, c)$ need to be differentiated, which can be identified by major-axes dielectric function spectra $\varepsilon_j(\omega)$ for certain symmetries only.

For orthorhombic, tetragonal, hexagonal, trigonal, and cubic symmetry, a real-valued rotation matrix $\mathbf{A}$ independent of wavelength can be found such that ε is diagonal in a given orthogonal axes system $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$:

$$\varepsilon = \mathbf{A} \begin{pmatrix} \varepsilon_a & 0 & 0 \\ 0 & \varepsilon_b & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix} \mathbf{A}^{-1}, \tag{10.24}$$

where $\varepsilon_a \equiv 1 + \varrho_a$, $\varepsilon_b \equiv 1 + \varrho_b$, and $\varepsilon_c \equiv 1 + \varrho_c$.

---

[10] For purely dielectric material without internal or external magnetic fields, due to invariance upon time-reversal, there is no directional dependence along one axis (Onsager principle).

For the monoclinic and triclinic crystal systems such a wavelength-independent rotation matrix does not exist, in general.

Isotropic Materials. In the most simple case, with no directional dependence, ε is a scalar and the electric displacement reads

$$\mathbf{D} = \varepsilon_0 \varepsilon \mathbf{E}. \tag{10.25}$$

Uniaxial Materials. For uniaxial materials with trigonal, tetragonal, and hexagonal symmetry, $\boldsymbol{\varepsilon}$ has two identical in-plane components $\varepsilon_a = \varepsilon_b = \varepsilon_\perp$ and one out-of-plane component $\varepsilon_c = \varepsilon_\parallel$, and the electric displacement takes the form

$$\mathbf{D} = \mathbf{x}\varepsilon_0 \varepsilon_\perp E_x + \mathbf{y}\varepsilon_0 \varepsilon_\perp E_y + \mathbf{z}\varepsilon_0 \varepsilon_\parallel E_z. \tag{10.26}$$

Orthorhombic Materials. For biaxial materials with orthorhombic symmetry, with their coordinate system coincident with a given laboratory system ($a_y = a_z = b_x = b_z = c_x = c_y = 0$, $a_x = b_y = c_z = 1$) one obtains from (10.23)

$$\mathbf{D} = \mathbf{x}\varepsilon_0(1 + \varrho_a)E_x + \mathbf{y}\varepsilon_0(1 + \varrho_b)E_y + \mathbf{z}\varepsilon_0(1 + \varrho_c)E_z, \tag{10.27a}$$
$$\mathbf{D} = \mathbf{x}\varepsilon_0 \varepsilon_a E_x + \mathbf{y}\varepsilon_0 \varepsilon_b E_y + \mathbf{z}\varepsilon_0 \varepsilon_c E_z, \tag{10.27b}$$

where the polarization vector in a given direction exclusively depends upon the electric field component in that same direction.

Monoclinic and Triclinic Materials. For biaxial materials with non-Cartesian monoclinic and triclinic systems, (10.23) can be represented by a virtual orthogonal basis for $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and a projection matrix $\mathbf{U}$ [78]:

$$\mathbf{U} = \begin{pmatrix} \sin\alpha & (\cos\gamma - \cos\alpha\cos\beta)(\sin\alpha)^{-1} & 0 \\ 0 & (1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)^{\frac{1}{2}}(\sin\alpha)^{-1} & 0 \\ \cos\alpha & \cos\beta & 1 \end{pmatrix}. \tag{10.28}$$

Parameters α, β, γ are the internal angles between major polarizability axes $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and which differentiate monoclinic ($\beta \neq \alpha = \gamma = 90°$) and triclinic ($\alpha \neq \beta \neq \gamma$) biaxial optical properties. In the definition of $\mathbf{U}$, use was made of the following choice of free coordinates: within the auxiliary Cartesian system, $\mathbf{c}$ is chosen to coincide with the z-axis, thus $c_x = c_y = 0$ in (10.23). $\mathbf{a}$ is chosen to be located within the $\{x, z\}$-plane, thus $a_y = 0$ and $a_x = \sqrt{1 - a_z^2}$, and b_x, b_z follow accordingly where $b_y = \sqrt{1 - b_x^2 - b_z^2}$. Thus, free parameters in (10.28) are a_z, b_x, b_z, or equivalently α, β, γ as depicted in Fig. 10.5. If all angles $\alpha = \beta = \gamma = 90°$ then the so called direct structure matrix $\mathbf{U}$ takes the form $\mathbf{U} = diag\{1, 1, 1\}$, where $diag\{\cdot\}$ indicates the diagonal 3×3 matrix, which represents the orthorhombic symmetry. Explicitly, the dielectric tensor $\boldsymbol{\varepsilon}_m$ for a biaxial material with monoclinic symmetry takes the form
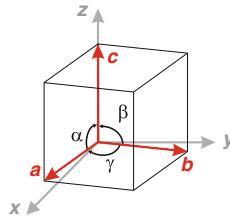
Fig. 10.5 Definition of the angles α, β, and γ as used in WVASE32®(J. A. Woollam Co., Inc.). Orthorhombic ($\alpha = \beta = \gamma = 90°$), monoclinic ($\alpha = \gamma = 90° \neq \beta$), and triclinic ($\alpha \neq \beta \neq \gamma \neq 90°$) systems can be distinguished amongst materials with biaxial ($\varrho_a \neq \varrho_b \neq \varrho_c$) properties when evaluating their external Euler angles as a function of wavelength (Reprinted with permission from [40]. Copyright 2010, Schmidt)

$$\boldsymbol{\varepsilon}_{\mathrm{m}} = \mathbf{U} \begin{pmatrix} \varepsilon_a & 0 & 0 \\ 0 & \varepsilon_b & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix} \mathbf{U}^T = \begin{pmatrix} 1 + \varrho_a & 0 & 0 \\ 0 & 1 + \sin^2\beta\varrho_b & \sin\beta\cos\beta\varrho_b \\ 0 & \sin\beta\cos\beta\varrho_b & 1 + \cos^2\beta\varrho_b + \varrho_c \end{pmatrix},$$
(10.29)

where β is the monoclinic angle between axes **b** and **c**.[11] In the most general form, the dielectric tensor $\boldsymbol{\varepsilon}_{\mathrm{t}}$ for a triclinic system reads

$$\boldsymbol{\varepsilon}_{\mathrm{t}} = \begin{pmatrix} \varrho_a \sin^2\alpha + \varrho_b \frac{1}{\sin^2\alpha}\Gamma^2 & \varrho_b \frac{1}{\sin^2\alpha}\Gamma\Lambda & (\varrho_a + \varrho_b \frac{1}{\sin^2\alpha}\Gamma)\sin\alpha\cos\beta \\ \varrho_b \frac{1}{\sin^2\alpha}\Gamma\Lambda & -\varrho_b \frac{1}{\sin^2\alpha}\Upsilon & \varrho_b \frac{1}{\sin\alpha}\cos\beta\Lambda \\ (\varrho_a + \varrho_b \frac{1}{\sin^2\alpha}\Gamma)\sin\alpha\cos\beta & \varrho_b \frac{1}{\sin\alpha}\cos\beta\Lambda & \varrho_a \cos^2\alpha + \varrho_b \cos^2\beta + \varrho_c \end{pmatrix},$$
(10.30)

with

$$\Gamma = -\cos\alpha\cos\beta\cos\gamma + \cos\gamma, \tag{10.31a}$$

$$\Lambda = (-\cos^2\alpha - \cos^2\beta + 2\cos\alpha\cos\beta\cos\gamma + \sin^2\gamma)^{\frac{1}{2}}, \tag{10.31b}$$

$$\Upsilon = \cos^2\alpha + \cos^2\beta - 2\cos\alpha\cos\beta\cos\gamma + \cos^2\gamma - 1. \tag{10.31c}$$

Experimentally, monoclinic and triclinic properties can only be distinguished by analyzing measured GE data over a wide spectral range. This phenomenon is due to dispersion, i.e., wavelength dependencies of functions $\varrho_a(\omega)$, $\varrho_b(\omega)$, $\varrho_c(\omega)$. The rotations to diagonalize (10.23) depend explicitly on ϱ_a, ϱ_b, ϱ_c and are thus wavelength dependent. Within a narrow spectral region an orthogonal rotation matrix with Euler angles φ, θ, ψ can always be found such that $\boldsymbol{\varepsilon}$ can be diagonalized (Sect. 10.3.4.2). However, choosing a different spectral region, this rotation matrix will be different exhibiting the wavelength-dependent character and revealing mon-

---

[11] Note the uncommon assignment of unit cell angles. This notation is chosen here because it is equivalent to the definition of the ellipsometric analysis software WVASE32® [79] and is illustrated in Fig. 10.5.

Table 10.2 Symmetries and dielectric tensor properties of dielectric materials with symmetric dielectric tensor

Symmetry	Classification	Major dielectric polarizabilities ($\varepsilon_j = 1 + \varrho_j$)	Internal angles
Isotropic	Cubic	$\varepsilon_a = \varepsilon_b = \varepsilon_c$	$\alpha = \beta = \gamma = 90°$
Uniaxial	Trigonal	$\varepsilon_a = \varepsilon_b \neq \varepsilon_c$	$\alpha = \beta = \gamma = 90°$
	Tetragonal		
	Hexagonal		
Biaxial	Orthorhombic	$\varepsilon_a \neq \varepsilon_b \neq \varepsilon_c$	$\alpha = \beta = \gamma = 90°$
	Monoclinic	$\varrho_a \neq \varrho_b \neq \varrho_c$	$\beta \neq \alpha = \gamma = 90°$
	Triclinic	$\varrho_a \neq \varrho_b \neq \varrho_c$	$\alpha \neq \beta \neq \gamma \neq 90°$

For definition and explanation of major dielectric polarizabilities refer to Sect. 10.3.4.3

oclinic or triclinic properties. Considering for intrinsic monoclinic or triclinic properties by using projections **U** and allowing for internal angles α, β, γ, a wavelength-independent set of Euler angles φ, θ, ψ must be found.

10.3.4.4 Connection Between Intrinsic Polarizabilities and Dielectric Tensor

In general, symmetric $\boldsymbol{\varepsilon}$-tensor materials can be organized in three distinct groups with respect to their optical properties, as summarized in Table 10.2:

- Materials with a *cubic* symmetry are optically isotropic (for example, amorphous material). All three axes of major dielectric polarizabilities are equivalent and mutually orthogonal.
- Materials with *trigonal*, *tetragonal*, and *hexagonal* symmetry (three-fold, four-fold, and six-fold rotation axes exist, respectively) have uniaxial optical properties. All axes of major dielectric polarizabilities are mutually orthogonal, however, only two out of the three are equivalent axes. One dielectric principal axis coincides with the rotation axis, while any two remaining and equivalent directions are perpendicular to the principal axis of rotation.
- Materials with *orthorhombic*, *monoclinic*, and *triclinic* symmetry are optically biaxial. All three major dielectric polarizabilities are different.

If (10.15) of a transparent material is transformed to its principal axes (Sect. 10.3.4.2), the principal refractive indices $n_j \equiv \sqrt{\varepsilon_j}$ ($j = a, b, c$) geometrically represent, in general, an ellipsoid, also called the index ellipsoid [73, 80]. The terms uniaxial and biaxial refer to the number of optical axes. An optical axis is defined as the normal to a plane intersecting the ellipsoid, where the circumference of the section is a circle, and its center coincides with the center of the ellipsoids. The optical axis presents a direction along which the speed of propagation is independent of polarization. Two such axes can be found in a general ellipsoid (biaxial), a spheroid has one (uniaxial), and a sphere an infinite number (isotropic). Materials with more than one principle dielectric constant exhibit birefringence.
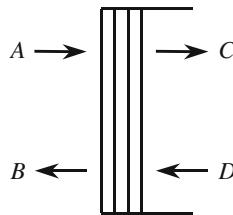
Fig. 10.6 Schematic presentation of incident A, reflected B, and transmitted C plane waves across a sample with plane parallel interfaces, and multiple layer stacks at the front side of the supporting substrate. D modes (if present) are incident from the right. The substrate may totally absorb C and/or D (Reprinted with permission from [68]. Copyright 2004, Springer-Verlag Berlin Heidelberg)

10.3.5 Light Propagation in Layered Anisotropic Media

The electromagnetic response of stratified anisotropic materials with plane parallel boundaries can be conveniently calculated in either Jones or Mueller matrix presentation with a 4×4 matrix approach, often cited as the Berreman-formalism [63, 68, 70, 81, 82]. Whereas the Jones and Mueller matrix formalism describe the measurable electromagnetic field components, the 4×4 matrix algebra treats the electromagnetic field components within the sample, i.e, internal sample polarization-modifying processes responsible for the external polarizing sample properties. The response of the entire multiple-layered structure (Fig. 10.6) is described by a transfer matrix $\mathbf{T}$:

$$\begin{pmatrix} A_s \\ B_s \\ A_p \\ B_p \end{pmatrix} = \mathbf{T} \begin{pmatrix} C_s \\ D_s \\ C_p \\ D_p \end{pmatrix}. \tag{10.32}$$

A so-called characteristic transfer matrix, $\mathbf{T}_\mathrm{p}$, accounts for the optical properties of a single homogeneous layer within a stratified sample.[12] Matrices for incident (ambient, $\mathbf{L}_\mathrm{a}$) and exit mediums ($\mathbf{L}_\mathrm{f}$) embed the layer stack according to the layer stack surrounding. The transfer matrix $\mathbf{T}$ results from the ordered product of all m layers' matrices $\mathbf{T}_\mathrm{p}$ according to their position within the layer stack, starting and ending with the incident and exit matrices [69, 82]:

$$\mathbf{T} = \mathbf{L}_\mathrm{a}^{-1} \mathbf{T}_\mathrm{p1}^{-1} \dots \mathbf{T}_\mathrm{pm}^{-1} \mathbf{L}_\mathrm{f}. \tag{10.33}$$

---

[12] Further details and analytical solutions for $\mathbf{T}_\mathrm{p}$ can be found in the literature [68, 82].

The ambient matrix

$$\mathbf{L}_a^{-1} = \frac{1}{2} \begin{pmatrix} 0 & 1 & -(n_a \cos \Phi_a)^{-1} & 0 \\ 0 & 1 & (n_a \cos \Phi_a)^{-1} & 0 \\ (\cos \Phi_a)^{-1} & 0 & 0 & 1/n_a \\ -(\cos \Phi_a)^{-1} & 0 & 0 & 1/n_a \end{pmatrix} \qquad (10.34)$$

depends on the angle of incidence Φ_a and the index of refraction n_a of the (isotropic) ambient material. The exit matrix depends on the angle of propagation within the exit medium Φ_a and its complex valued index of refraction $N_f = \sqrt{\varepsilon_f} = n_f + ik_f$:

$$\mathbf{L}_f = \begin{pmatrix} 0 & 0 & \cos \Phi_f & -\cos \Phi_f \\ 1 & 1 & 0 & 0 \\ -N_f \cos \Phi_f & N_f \cos \Phi_f & 0 & 0 \\ 0 & 0 & N_f & N_f \end{pmatrix}. \qquad (10.35)$$

The angle Φ_f is calculated from Snell's law:

$$\cos \Phi_f = \sqrt{1 - ([n_a/N_f] \sin \Phi_a)^2}. \qquad (10.36)$$

For a given layer of thickness d with index of refraction N, the matrix $\mathbf{T}_p$ can be obtained from the exponential function

$$\mathbf{T}_p \equiv \exp\left\{i\frac{\omega}{c}\Delta d\right\}, \qquad (10.37)$$

where c is the speed of light, ω is the light wave orbital frequency, and Δ is the characteristic coefficient matrix of the layer. The characteristic coefficient matrix Δ is derived from Maxwell's equations as shown by Schubert [82] and depends on the dielectric tensor $\boldsymbol{\varepsilon}$ and the wavevector component k_x:

$$\Delta = \begin{pmatrix} -k_x \frac{\varepsilon_{zx}}{\varepsilon_{zz}} & -k_x \frac{\varepsilon_{zy}}{\varepsilon_{zz}} & 0 & 1 - \frac{k_x^2}{\varepsilon_{zz}} \\ 0 & 0 & -1 & 0 \\ \varepsilon_{yz}\frac{\varepsilon_{zx}}{\varepsilon_{zz}} - \varepsilon_{yx} & k_x^2 - \varepsilon_{yy} + \varepsilon_{yz}\frac{\varepsilon_{zy}}{\varepsilon_{zz}} & 0 & k_x \frac{\varepsilon_{yz}}{\varepsilon_{zz}} \\ \varepsilon_{xx} - \varepsilon_{xz}\frac{\varepsilon_{zx}}{\varepsilon_{zz}} & \varepsilon_{xy} - \varepsilon_{xz}\frac{\varepsilon_{zy}}{\varepsilon_{zz}} & 0 & -k_x \frac{\varepsilon_{xz}}{\varepsilon_{zz}} \end{pmatrix}, \qquad (10.38)$$

$$k_x = n_a \sin \Phi_a. \qquad (10.39)$$

The matrix Δ should not be confused with the real-valued ellipsometric parameter Δ. $\mathbf{T}_p$ connects the in-plane components of the electric and magnetic fields at layer interfaces separated by d and includes the effects of all multiple reflections if a part of the wave is traveling along a direction with no or weak absorption. $\mathbf{T}_p$ is computed with Δ as input:

$$\mathbf{T}_p \equiv \exp\left\{i\frac{\omega}{c}\Delta d\right\} = \beta_0\mathbf{E} + \beta_1\Delta + \beta_2\Delta\Delta + \beta_3\Delta\Delta\Delta. \tag{10.40}$$

The scalars β_j are obtained from the following linear relations:

$$\exp\left\{i\frac{\omega}{c}q_k d\right\} = \sum_{j=0}^{3}\beta_j q_k^j, \qquad k = 1\ldots 4, \tag{10.41}$$

where q_k denote the eigenvalues of Δ associated with one of the four electromagnetic eigenmodes Ξ_k within the layer ($k = 1\ldots 4$). Two solutions have a positive real part and constitute the forward traveling plane waves with respect to the chosen laboratory coordinate system. The other solutions with negative real parts are due to the backward-traveling wave components.

In order to calculate $\mathbf{T}_p$ parameters ϱ_a, ϱ_b, ϱ_c, φ, θ, ψ, and α, β, γ are needed, and which then represent the current orientation of the polarizability system ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$) relative to the laboratory coordinate system (sample surface and plane of incidence) and relative to the ellipsometry measurement (plane of incidence and angle of incidence).

Piecewise homogeneous layers. Explicitly, the transfer matrix for a 3F2-STF (chevron with three layers) depicted in Fig. 10.9a reads

$$\mathbf{T} = \mathbf{L}_a^{-1}\mathbf{T}_{p1}^{-1}\mathbf{T}_{p2}^{-1}\mathbf{T}_{p3}^{-1}\mathbf{L}_f. \tag{10.42}$$

The STF is virtually separated into three layers with specific thickness d such that each partial transfer matrix accounts for each slanted columnar layer, and together with incident and exit matrix result in the transfer matrix of the layer stack. Note that within each layer the dielectric tensor $\boldsymbol{\varepsilon}$ of the (virtual) orthorhombic basis for $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ is in diagonalized form when $\varepsilon_a = \varepsilon_x$, $\varepsilon_b = \varepsilon_y$, and $\varepsilon_c = \varepsilon_z$, and oriented as depicted in Fig. 10.15. Hence, as discussed also later, if the slanting planes are parallel to the plane of incidence in Fig. 10.9a, the Euler angles within $\mathbf{T}_{p1}$, for example, are $\varphi = +90°$, $\theta = +45°$, and $\psi = 0°$.

Continuously rotated layers. Similar to the piecewise homogenous layer, the transfer matrix for a continuously rotated H-STF (Fig. 10.9b) explicitly takes the form

$$\mathbf{T} = \lim_{\substack{\delta d \to 0 \\ m \to \infty}} \mathbf{L}_a^{-1}\left(\prod_{j=1}^{m}\mathbf{T}_{pj}^{-1}(\delta d)\right)\mathbf{L}_f. \tag{10.43}$$

Here, it is implied in that ϱ_a, ϱ_b, ϱ_c, α, β, γ, θ, and ψ are identical for each layer and layers differ only by $\delta\varphi$. The total number m of partial transfer matrices depends on the level of discretization of the structure, i.e., in how many sublayers (slices) with thickness δd the structure is split up. A higher level of discretization results in more, thinner slices, and therefore finer approximations can be achieved. However, the computational effort increases considerably.

Explicit solutions have been reported for continuously rotated layers at normal incidence as well as a fast approximation scheme for oblique angles of incidence [83, 84].

10.3.6 Ellipsometry Model Description for Sculptured Thin Films

10.3.6.1 Homogeneous Biaxial Layer Approach

It is known that a thin film with columnar microstructure causes optical birefringence [31]. The microstructural asymmetry described as shape anisotropy (columns exhibit a slightly elliptical shape due to non-existence of structure shadowing in the direction perpendicular to the incoming particle flux) or preferentially bunched columns along the direction perpendicular to the deposition plane causes one setup of the so called form birefringence [85]. Also due to these "non-idealities" the optical nature of a slanted columnar thin film is biaxial. Hodgkinson and Wu [86] adapted the Herpin index method[13] to transparent biaxial thin films and concluded that non-absorbing slanted columnar thin films (F1-STFs) can be considered as an effective medium with biaxial properties.

This consideration has been adapted and augmented here to create a generalized optical model, which is valid for slanted columnar thin films of any material (absorbing and non-absorbing). Examples for absorbing STFs prepared from different metals are presented in Sect. 10.6.

The optical equivalent description for a F1-STF can be, in general, a single dielectrically homogeneous (along z) biaxial layer, thereby describing an effective medium (Fig. 10.7). The biaxial layer comprises parameters thickness d, corresponding to the actual thickness of the nanostructured thin film, three complex, wavelength-dependent functions $\varrho(\omega)_a$, $\varrho(\omega)_b$, and $\varrho(\omega)_c$ pertinent to intrinsic axes **a**, **b**, and **c**, their internal angles α, β, and γ, and (external) Euler angles φ, θ, and ψ determining the orientation of the columns and sample during a particular measurement [88, 89].
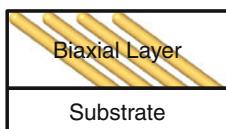
Fig. 10.7 Optical layer model of an absorbing F1-STF (slanted columnar thin film). The anisotropic nanostructure can be described as an effective medium with a single biaxial layer on top of a substrate

---

[13] At one wavelength, a symmetrical thin-film combination (periodically stratified medium) is equivalent to a single film, characterized by an equivalent index and equivalent thickness [87].
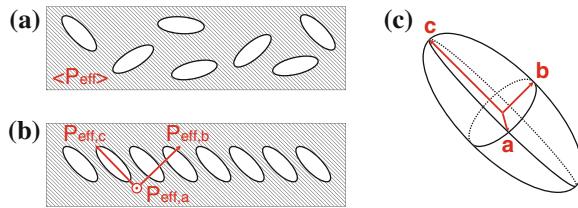
Fig. 10.8 Effective medium scenarios with mixtures of elliptical inclusions [depicted in (**c**)] and a homogeneous host medium. The mixture with randomly oriented inclusions **a** exhibits an average effective polarizability $\langle \mathbf{P}_{eff} \rangle$ whereas the mixture with aligned inclusions **b** shows anisotropic properties with three effective polarizabilities $\mathbf{P}_{eff,j}$ according to the shape of the inclusions (Reprinted with permission from [90]. Copyright 2012, American Institute of Physics)

This homogenous biaxial layer approach has major advantages over other existing effective medium approximations: (i) no initial assumptions such as optical properties of the constituents or packaging fractions are necessary, (ii) it is valid for absorbing and non-absorbing materials, (iii) it does not depend on packaging fractions, and (iv) it does not depend on the structure size. Note that the actual structure size is disregarded in this homogenization approach. This procedure is considered valid since the lateral dimension of the nanostructures (diameter) is much smaller than the probing wavelength. Care must be taken when properties at shorter wavelengths are evaluated, because diffraction and scattering phenomena may be present.

10.3.6.2 Anisotropic Bruggeman Effective Medium Approximation

Effective medium approximations (EMAs) are physical models based on the properties and the relative fraction of its components and describe the macroscopic properties of a medium. The Bruggeman formalism, for example, is a homogenization process with absolute equality between the phases in mixture, and was originally developed for a randomly inhomogeneous medium with, for example, spherical inclusions [91]. This Bruggeman approach has been extensively discussed and generalized to treat materials consisting of crystallites of arbitrary shape by introducing depolarization factors [93, 96]. Depolarization factors L_a^D, L_b^D, L_c^D are proportional to the polarizabilities along major polarizability axes **a**, **b**, **c** and are representative for relative dimensions of elliptical inclusions. In case of a host medium with randomly oriented inclusions the mixture macroscopically exhibits an isotropic effective polarizability $\langle \mathbf{P}_{eff} \rangle$ with an isotropic effective dielectric function ε_{eff} due to an averaging over all major polarizability axes. The generalized Bruggeman EMA can also be applied ad-hoc to highly oriented inclusions and may then be called anisotropic Bruggeman EMA (AB-EMA). For aligned inclusions with ellipsoidal shape embedded in a host matrix the average for the biaxial (orthorhombic) effective dielectric functions $\varepsilon_{eff,a}$, $\varepsilon_{eff,b}$, $\varepsilon_{eff,c}$ is then only taken along **a**, **b**, and **c**, respectively (Fig. 10.8) [94, 95]. The AB-EMA formulae for the three effective major dielectric functions $\varepsilon_{eff,j} \equiv \varepsilon_j$ for a mixture of m constituents in implicit form are:

$$\sum_{n=1}^{m} f_n \frac{\varepsilon_n - \varepsilon_{\text{eff},j}}{\varepsilon_{\text{eff},j} + L_j^{\text{D}}(\varepsilon_n - \varepsilon_{\text{eff},j})} = 0, \qquad j = a, b, c, \qquad (10.44)$$

where permittivities ε_n are the bulk dielectric functions of the respective film constituent with material volume fractions f_n. The sum of all material fractions f_n has to equal unity. L_j^{D} are the depolarization factors of the inclusion ellipsoids along the three orthogonal major polarizability axes $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and the sum of all three depolarization factors must obey unity [92, 96]:

$$L_a^{\text{D}} + L_b^{\text{D}} + L_c^{\text{D}} = 1. \qquad (10.45)$$

Note that in order to preserve the symmetry of the original Bruggeman theory, structural equivalence between constituents is assumed (one set of L_j^{D} renders the effective medium geometry).The upper and lower bounds on $\varepsilon_{\text{eff},j}$, where the depolarization factors are 0 and 1, correspond to minimum and maximum charge screening effects, respectively, and are called Wiener bounds [91, 97, 98]. These two cases can be understood considering capacitors connected either in parallel or in series. The effective permittivities of prolate spheroids aligned along the substrate normal can be approximated from (10.44) with $L_a^{\text{D}} = L_b^{\text{D}} = 0.5$ and $L_c^{\text{D}} = 0$. Consequently, the case of $L_a^{\text{D}} = L_b^{\text{D}} = L_c^{\text{D}} = \frac{1}{3}$ corresponds to spherical inclusions, and for two constituents simplifies to the original quadratic equation derived by Bruggeman [86, 96, 99].

A real-valued rotation matrix $\mathbf{A}$ with wavelength-independent Euler angles φ, θ, ψ can then be found to transform the Cartesian laboratory coordinate frame into the material coordinate frame 10.3.4.2. In order to describe biaxial materials with a non-Cartesian system the obtained virtual orthogonal basis from 10.44 has to be transformed by the projection matrix $\mathbf{U}$ (Sect. 10.3.4.3). Examples of metal slanted columnar thin films with monoclinic optical properties ($\mathbf{P}_{\text{eff},b}$ is not parallel to the semiaxis b of the ellipsoid but rather tilted towards c by a monoclinic angle β) are presented in Sect. 10.6.4.

In general, the AB-EMA is only valid in the long wavelength approximation and hence if the structure size approaches the order of the wavelength of the probing light this model procedure may result parameters which are not representative anymore for the film under investigation [67, 91, 100].

10.3.6.3 Piecewise Homogeneous Biaxial Layer Approach

If substrate rotation is involved during the growth process of STFs, a single biaxial layer accounting for the film is not sufficient anymore to describe the dielectric polarization response.[14] For the piecewise homogeneous biaxial layer approach two types

---

[14] Very fast substrate rotation (<2 nm vertical growth per revolution) results in V-STFs, i.e., a screw degenerates to a straight column because the pitch is too small. Optical properties of V-STFs are not discussed in this chapter, however, the nanostructured film has uniaxial properties with the ordinary
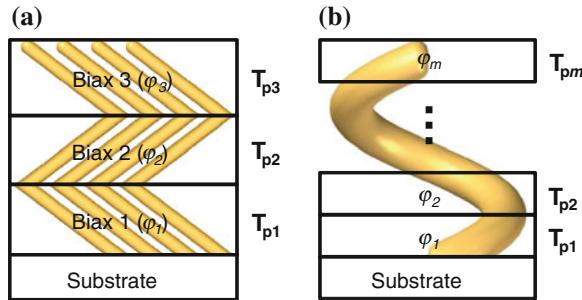
Fig. 10.9 Optical layer model for arbitrarily oriented STFs. **a** A 3F2-STF (chevron with three layers), for example, is optically modeled with three biaxial layers with alternating orientations. **b** A H-STF is approximated by n biaxial layers, which are rotated with respect to each other by $\delta\varphi$. $\mathbf{T}_{pm}$ corresponds to the partial transfer matrix in the 4×4 algebra for layer m, and depends on $\varrho_a, \varrho_b, \varrho_c, \alpha, \beta, \gamma, \varphi, \theta, \psi$, and d

of STFs are distinguished here: (i) F-STFs (except F1; fabricated with sequential substrate rotations) and (ii) H-STFs (fabricated with continuous substrate rotation). It is assumed that the STF is made of m F1-STF slices, where within each slice (layer) the dielectric properties are homogeneous [33, 101].

F-STFs. F-STFs (all but F1) are grown while the substrate is rotated step-wise (abruptly) after a certain pause time. If a sequential substrate rotation of 180° is employed, for example, the resulting F2-STFs, also called chevrons or zig-zags, can be considered as stratified (or a cascade of) F1-STFs with opposite slanting directions in adjacent slabs. Consequently, the optical model for the chevron thin film with three layers (3F2-STF),[15] depicted in Fig. 10.9a, consists of three homogeneous anisotropic (biaxial) layers on top of a layer accounting for the substrate. The Euler angles for each layer $(\varphi_j, \theta_j, \psi_j)$, which transform the Cartesian coordinate system (x, y, z) into the sample coordinates (ξ, η, ζ), represent the orientation of each slanted column (building block) in the nanostructure. In case the angle of the incoming particle flux θ_i was kept constant during deposition, a common dielectric tensor, with three major polarizabilities $\varrho_a, \varrho_b, \varrho_c$ pertinent to the intrinsic axes **a**, **b**, **c**, and internal angles α, β, γ and Euler angles φ, θ, ψ can be assigned to each biaxial layer. Deposition at constant θ_i results in equal packaging fractions in subsequent layers and therefore common major polarizabilities may be assumed. Furthermore, all layers have an individual thickness parameter d_j such that the total thickness is equal to the overall film thickness ($d = d_1 + d_2 + \ldots + d_m$). This approach is valid, in general, for arbitrarily oriented F-STFs and examples are presented in Sect. 10.6.3. As will be shown later, the Euler angle θ is identical to the geometrical slanting

---

(Footnote 14 continued)

dielectric constant in the substrate interface and the extraordinary along the columns and normal to the substrate.

[15] For nomenclature see Table 10.1.

angle, φ is the rotation of the slanting plane with respect to the laboratory coordinate frame, and ψ is found to be zero.

H-STFs. If the substrate is continuously rotating around the normal during deposition, helical STFs (H-STFs) are growing since the sample rotation is equivalent to a constant angular change of the incoming vapor flux direction around the substrate normal and thus the self-shadowed regions change dynamically. H-STFs, schematically shown in Fig. 10.9b, represent rotationally inhomogeneous anisotropic material with a twist along the sample normal. Such chiral nanostructured thin films can be considered as "frozen" cholesteric liquid crystals [82–84]. Here the dielectric tensor $\boldsymbol{\varepsilon}(z)$ depends on the spatial position with respect to the z axis. In order to model the electromagnetic plane wave response of H-STFs the thin film has to be virtually separated into m homogeneous anisotropic layers with subsequently shifted Euler angle parameters $\varphi_1, \varphi_2, \ldots, \varphi_m$ with individual thickness parameters $\delta d = d/m$. These layers represent piecewise rotation with respect to each other by $\delta\varphi$ to resemble the twisted character such that

$$\boldsymbol{\varepsilon}_j = (z) = \mathbf{A}(\varphi_j)\boldsymbol{\varepsilon}\mathbf{A}(\varphi_j)^{-1}, \qquad j = 1, \ldots, m. \tag{10.46}$$

Here, $\boldsymbol{\varepsilon}_j$ is the dielectric function tensor that describes the first virtual layer corresponding to its orientation during measurement with respect to the plane of incidence. Physical quantities such as principal dielectric functions (as a function of photon energy and z), orientation, overall thickness, handedness, and thickness of the helical structure can be thereby retrieved from the ellipsometry model calculations. In contrast to F-STFs, for H-STF the Euler angle ψ is found to be not equal to zero (Sect. 10.6.3.4).

10.3.7 Ellipsometry Data Analysis

The Jones (r_{ij}) or Mueller (M_{kl}) matrix elements are functions of the photon energy $\hbar\omega$, the (major-axes) dielectric functions $\varepsilon_j(\omega)$[16] ($j = a, b, c$) and its Euler orientation angles ψ, φ, and θ, the thickness d, the ambient material's dielectric function, and the angle of incidence Φ_a. For multiple layers, ε_j, the Euler angles, internal polarizability angles (α, β, γ), as well as d may be layer-specific. The standard model for analyzing ellipsometry data is based on a sequence of homogeneous (isotropic or anisotropic) layers with smooth and parallel interfaces. In case of an anisotropic sample, the ellipsometric measurement depends also on the orientation of axes $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ with respect to the plane of incidence, and the polarization state of the incident light beam [63, 71, 82]. Depending on the parameters of interest and the sample properties (layer sequence, anisotropy) different analysis approaches can be employed.

---

[16] For materials with monoclinic and triclinic symmetry $\varepsilon_j(\omega)$ depend on the polarization functions $\varrho_a, \varrho_b, \varrho_c$ and their non-Cartesian axes $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ as described in Sect. 10.3.4.

10.3.7.1 Wavelength-by-Wavelength Analysis

Traditionally, wavelength-by-wavelength best-match model calculations (point-by-point best-match model calculations) are performed when the dielectric function values of interest are extracted from the experimental data for each wavelength, and independent of all other spectral data points. For this procedure, the thickness of the particular layer as well as the dielectric functions and the thicknesses of all other sample constituents have to be known. However, in order to obtain values of physically relevant parameters (such as critical point energies and broadening parameters) and to ensure Kramers-Kronig consistency the dielectric function obtained from the point-by-point best-match model calculation needs to be compared with a line-shape model.

10.3.7.2 Parameterized Model Dielectric Function Analysis

A commonly employed, robust procedure is matching parameterized model dielectric functions (MDFs) to experimental data simultaneously for all spectral data points. This provides a direct connection between measured data and physical parameters of interest. Parametric models further prevent wavelength-by-wavelength measurement noise from becoming part of the extracted dielectric functions and greatly reduce the number of free parameters. With the use of parametric models a certain risk is involved for subtle spectral features to be neglected by the lineshape of the model function. Nevertheless, parameterizations of ε_j based on a physical model is the best choice for ellipsometry data analysis, especially when the wavelength-by-wavelength best-match model calculation method is inapplicable.

In this work, two physical lineshape parameterization models have been used in order to match experimental data in the measured visible to near-infrared spectral region.

Harmonic Lorentzian oscillator model. A simple calculation of the complex dielectric function assumes that the response of the material to electromagnetic radiation can be represented by an ensemble of non-interacting harmonic oscillators. The harmonic Lorentz oscillator model equation is given by

$$\varepsilon(E) = (n + ik)^2 = \varepsilon_{\text{off}} + \sum_j \frac{A_j}{E_{\text{c},j}^2 - E^2 + i\gamma_j E}, \qquad (10.47)$$

for the dielectric function expressed in terms of the photon energy E. Parameters A_j, $E_{\text{c},j}$, γ_j are determined in the best-match model calculation and denote amplitude, center energy, and broadening of the jth oscillator, respectively. ε_{off} is an offset parameter to account for contributions outside the measured spectral range.

Drude model for free-charge carriers. The classical Drude expression for free-charge carrier contributions is given by setting $E_{\text{c}} = 0$ in (10.47):

$$\varepsilon(E) = (n + ik)^2 = -\frac{N_v q^2}{m_{\text{eff}} \varepsilon_0 (E^2 - i\gamma E)}. \tag{10.48}$$

The amplitude parameter $A = N_v q^2 / m_{\text{eff}} \varepsilon_0$, where m_{eff} denotes the effective mass of the free-charge carriers with volume-density N_v. The vacuum dielectric permittivity and the charge of the free-charge carriers is given by ε_0 and q, respectively. The broadening parameter γ is related to the energy-independent relaxation time τ and the free-charge carrier mobility μ as

$$\gamma = \tau^{-1} = \frac{q}{m_{\text{eff}} \mu}. \tag{10.49}$$

Equation (10.48) leads with $\varepsilon = i\sigma/(\varepsilon_0 E)$ in the low frequency limit ($E \to 0$) to the classical DC Drude conductivity expression $\sigma_0 = N_v q^2 \tau / m_{\text{eff}}$ [66, 68].

10.3.7.3 Ellipsometry Test Functions

During the data analysis model parameters are varied until calculated and measured data match as close as possible (best-match model calculation). For fast convergence the Levenberg–Marquardt algorithm can be used, for example, in order to vary the adjustable model parameters until the weighted test function ξ_{SE} (mean square error, MSE) is minimized (maximum likelihood approach) [102]:

$$\xi_{\text{SE}}^2 = \frac{1}{2S - K + 1} \sum_{j=1}^{S} \left[\left(\frac{\Psi_j - \Psi_j^c}{\sigma_j^\Psi} \right)^2 + \left(\frac{\Delta_j - \Delta_j^c}{\sigma_j^\Delta} \right)^2 \right], \tag{10.50}$$

where S denotes the number of measured data pairs (Ψ_j, Δ_j), K is the number of real-valued model parameters, Ψ_j^c and Δ_j^c are the calculated ellipsometric parameters at photon energy $E = \hbar\omega_j$, and ($\sigma_j^\Psi, \sigma_j^\Delta$) are the standard deviations obtained during the measurement [103, 104].

For the generalized ellipsometry situation, the test function is set up accordingly for the Jones ($\xi_{\text{GE-J}}^2$) and Mueller ($\xi_{\text{GE-M}}^2$) matrix presentation:

$$\xi_{\text{GE-J}}^2 = \frac{1}{6S - K + 1} \sum_{j=1}^{S} \left[\left(\frac{\Psi_{pp,j} - \Psi_{pp,j}^c}{\sigma_{pp,j}^\Psi} \right)^2 \right.$$
$$\left. + \left(\frac{\Psi_{ps,j} - \Psi_{ps,j}^c}{\sigma_{ps,j}^\Psi} \right)^2 + \left(\frac{\Psi_{sp,j} - \Psi_{sp,j}^c}{\sigma_{sp,j}^\Psi} \right)^2 \right]$$

$$+ \frac{1}{6S - K + 1} \sum_{j=1}^{S} \left[\left(\frac{\Delta_{pp,j} - \Delta_{pp,j}^c}{\sigma_{pp,j}^{\Delta}} \right)^2 \right.$$

$$\left. + \left(\frac{\Delta_{ps,j} - \Delta_{ps,j}^c}{\sigma_{ps,j}^{\Delta}} \right)^2 + \left(\frac{\Delta_{sp,j} - \Delta_{sp,j}^c}{\sigma_{sp,j}^{\Delta}} \right)^2 \right],$$

$$\text{(10.51)}$$

$$\xi_{\text{GE-M}}^2 = \frac{1}{16S - K} \sum_{j=1}^{S} \sum_{k=0}^{3} \sum_{l=0}^{3} \left(\frac{M_{kl,j} - M_{kl,j}^c}{\sigma_j^{M_{kl}}} \right)^2. \tag{10.52}$$

Similar to standard ellipsometry, in addition to Ψ_{pp}, Δ_{pp}, Ψ_{ps}, Δ_{ps}, Ψ_{sp}, Δ_{sp}, and M_{kl} ($k, l = 0 \ldots 3$), their respective standard deviations σ^{Ψ}, σ^{Δ}, and $\sigma^{M_{kl}}$ are measured and propagated into the test functions.

Note that generalized ellipsometry MSE values for anisotropic samples cannot be directly compared with MSE values obtained with standard ellipsometry best-match model calculations for isotropic samples, where the "rule-of-thumb" indicates best achievement for MSE near or less than unity. The data set included for GE model analysis is generally larger than for SE analysis. In GE data analysis, in addition to closeness of the best-match model data to a given particular measured spectrum, match to the same data set versus sample rotation as well as its angle of incidence dependence is of equal importance.

In the result of the regression analysis, the correlation between different adjusted parameters and the confidence limit of the individual model parameter are of particular importance. Both quantities can be derived from the curvature matrix $\boldsymbol{\alpha}$:

$$\alpha_{kl} = \sum_{j=1}^{N} \left(\frac{1}{\sigma_{\Psi_j}^2} \frac{\delta \Psi_j^C \delta \Psi_j^C}{\delta \alpha_k \delta \alpha_l} + \frac{1}{\sigma_{\Delta_j}^2} \frac{\delta \Delta_j^C \delta \Delta_j^C}{\Delta \alpha_k \Delta \alpha_l} \right), \tag{10.53}$$

which is the inverse of the covariance matrix $\mathbf{C} \equiv \boldsymbol{\alpha}^{-1}$. The standard 90 % confidence limit L for the jth parameter is then given by [105]

$$L_j = \pm 1.65 \sqrt{C_{jj}} \, \xi, \tag{10.54}$$

where 1.65 is a statistically derived constant. Since ξ (MSE) has been also introduced into the expression, the confidence limits become larger when the quality of the best-match model calculation degrades. If not otherwise stated, the uncertainties given in this work are the respective confidence limits. The parameter correlation coefficients η_{jk} can be obtained from $\mathbf{C}$ by:

$$\eta_{jk} = \frac{C_{jk}}{\sqrt{C_{jj}} \sqrt{C_{kk}}}. \tag{10.55}$$
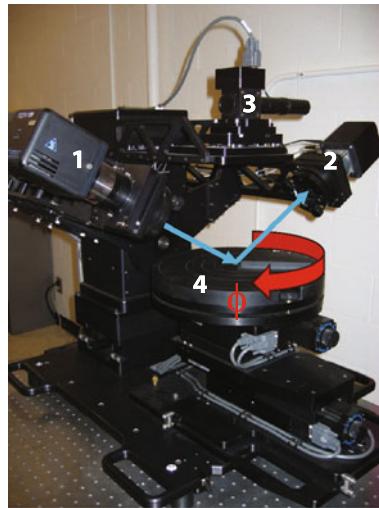
Fig. 10.10 J. A. Woollam M-2000VI$^{®}$ ellipsometer (*1, 2*) with automated rotation stage (*4*) and alignment laser (*3*) for automated *z*-alignment (Reprinted with permission from [40]. Copyright 2010, Schmidt)

A value of $\eta_{jk} \sim 1$ indicates correlation between the *j*th and the *k*th parameter. For correct and unique analysis no or only small correlation may occur between model parameters.

10.3.7.4 Multi-Sample and Multi-Sample-Configuration Analysis

If model parameters correlate or confidence limits are too large, either modifications to the model or other options have to be considered such as including further ellipsometric experimental data from similar samples and employ the multi-sample-analysis technique. In this approach, measured data sets of multiple samples are simultaneously analyzed with different models, which share a common set of parameters. These common parameters are assumed to be identical for each individual sample.

In a similar manner, measurement data obtained from a single sample but with a modified external parameter (applied magnetic field or temperature, for example) can be analyzed simultaneously. The effect of the varied quantity on the MDF has to be known and implemented in the model analysis. This approach is referred here to as multi-sample-configuration analysis.

10.3.7.5 Difference Spectra Analysis

The difference spectra analysis is particularly useful when changes of the sample's optical response due to a modification of an external parameter are small. Difference spectra are obtained by subtracting measured data sets acquired, for example, at

opposite (externally applied) magnetic field directions. This approach reveals most directly the change in the optical response upon variation of the external parameter.

10.3.8 Experimental Setup

Spectroscopic ellipsometry measurements were performed with a commercial J. A. Woollam Co., Inc. M-2000VI® multichannel ellipsometer covering the visible to near infrared spectral region, and which is mounted on an automated stage (Fig. 10.10). The polarization state generator unit (1) comprises a polarizer and rotating compensator and a 50 W halogen lamp serves as the light source. After reflection off the sample surface the light passes a rotating analyzer (part of the polarization state analyzer, 2), is spectrally separated by a prism and directed onto two CCD arrays. One CCD array detects a total of 390 wavelengths within the spectral range of 371–1000 nm (1.24–3.34 eV) whereas the second one detects another 200 wavelengths between 1000 and 1690 nm (0.73–1.24 eV). Hence, data for all 590 wavelengths can be acquired at the same time. The rotating-compensator-type ellipsometer is capable of measuring 11 out of 16 Mueller matrix elements normalized to M_{11} (except for elements in fourth row).[17] The sample tilt adjustment procedure is done with an additional alignment laser and four-quadrant detector (3) and is followed by a computer-controlled z-alignment (sample height adjustment). The motorized goniometer together with the horizontal sample stage (4) enables automated angle resolved (angle of incidence Φ_a and rotation angle ϕ) measurements, where Φ_a can be varied from 45° to 90° and ϕ from 0° to 360°. Furthermore, the M-2000VI® is equipped with a beam shutter to perform automated DC offset calibrations.[18] The WVASE32® software, which controls the ellipsometer hardware, allows for programming customized measurement routines through the add-on program WVASEscript [79].

10.4 Experimental Parameters

STFs discussed in this section were deposited at room-temperature in a customized UHV chamber onto (001) Si substrates with a native oxide of approximately 2.0 nm. All deposition materials were purchased in form of pellets (1/4" diameter × 1/4" length) from the Kurt J. Lesker Company.[19] Alumina (Al_2O_3) crucible liners were used for the deposition of Co, whereas graphite liners were used for Ti and supermal-

---

[17] Mueller matrix elements of the fourth row cannot be resolved because the polarization state analyzer (2) does not comprise a rotating compensator [66, 67].

[18] A DC offset calibration determines the detector noise level without source illumination.

[19] Cobalt has a specified purity of 99.95 % and titanium 99.995 %. Supermalloy is composed of 79.8 % Ni, 15.1 % Fe, and 5.1 % Mo.

Table 10.3 Overview of STF samples and their material discussed in this chapter with sample identification number (ID)

ID	Material	Geometry	Deposition rate (Å/s)	Growth recipe $[p_t, c_t] = $ min, $[r_t] = $ s
1	Co	F1	5.0 ± 0.3	$p_t = 10$
2	Ti	F1	4.0 ± 0.1	$p_t = 10{:}30$
3	NiFeMo	F1	3.6 ± 0.2	$p_t = 11$
4	Ti	2F2	5.1 ± 0.1	$p_{t1} = 5; r_t(-\pi) = 15, p_{t2} = 5$
5	Ti	2F4$^+$	5.0 ± 0.2	$p_{t1} = 5; r_t(-\frac{\pi}{2}) = 10; p_{t2} = 5$
6	Ti	2F4$^-$	5.0 ± 0.1	$p_{t1} = 5; r_t(+\frac{\pi}{2}) = 10; p_{t2} = 5$
7	Co	3F4$^+$	4.2 ± 0.3	$p_{t1} = 16; r_{t1}(-\frac{\pi}{2}) = 10; p_{t2} = 16;$ $r_{t2}(-\frac{\pi}{2}) = 10; p_{t3} = 16$
8	Co	1H$^+$	4.0 ± 0.1	$c_t(-0.1) = 11$
9	Co	1H$^-$	4.0 ± 0.2	$c_t(+0.1) = 11$

A detailed growth recipe is given for each sample with coded substrate dynamics: $r_t(\pm x)$ denotes a stepwise substrate rotation by x rad, $c_t(\pm y)$ denotes a continuous rotation at y rpm, and p_t stands for pause time. Values for p_t and c_t are given in minutes whereas r_t is given in seconds

loy. For all STFs discussed here, the deposition angle θ_i was set to $85 \pm 1°$. The deposition rate was monitored with a quartz crystal microbalance. The electron-beam power was controlled manually by adjusting the emission current to maintain a constant rate typically between 4 and 5 Å/s (with respect to the deposition controller at normal incidence $\theta_i = 0°$). This rate results in a growth time of approximately 10 to 12 min for a 100 nm thick STF deposited at $\theta_i = 85°$. The base pressure was approximately 10^{-9} mbar and did not rise above 10^{-8} mbar during depositions.

Table 10.3 summarizes deposition parameters of all samples, for which optical analysis is discussed in the following sections. A growth recipe is given with respect to substrate action. For example, sample #4 is a Ti 2F2-STF where the first layer was grown for $p_{t1} = 5$ min, then the substrate was rotated 180° counterclockwise in 15 s ($r_t(-\pi) = 15$), and afterwards kept still for $p_{t2} = 5$ min to grow the second layer. During growth of samples #8 and #9, the substrate was continuously rotated at 0.1 rpm (counterclockwise for the 1H$^+$ and clockwise for the 1H$^-$) for 11 min, which results in a total substrate rotation of 396°. The additional 10 % have been added to compensate for the formation of the nucleation layer during the initial growth period, based on investigations on previous H-STFs.

10.5 Structural Properties of Metal Sculptured Thin Films

10.5.1 Scanning Electron Microscopy Micrograph Analysis

Examples for metal F1-STFs resulting from the deposition procedure on flat substrates discussed in Sect. 10.4 are presented in Fig. 10.11. Cross-sectional and top-view high resolution SEM images of Co (sample #1), Ti (sample #2), and supermalloy (sample #3) F1-STFs illustrate the different film morphologies.
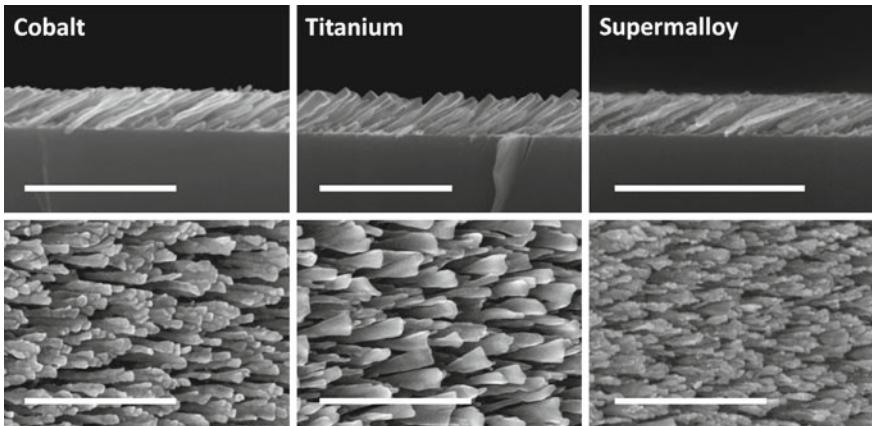
Fig. 10.11 SEM micrographs of F1-STF from different materials. The panels show cross-sectional and *top-view* images of cobalt, titanium, and supermalloy, respectively. For all of the images the incoming particle flux direction is from the right. Scale bars are 500 nm (Reprinted with permission from [40]. Copyright 2010, Schmidt)
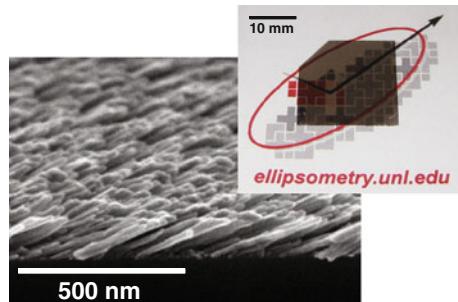

Fig. 10.12 Scanning electron micrograph of a cobalt F1-STF taken at a sample tilt of approximately 15° and photograph of an almost identical film deposited on a glass slide. The approximately 100 nm thick Co F1-STF is highly transparent as can be seen in the photograph, where the glass slide with deposited F1-STFs is placed onto a printed image (Reprinted with permission from [40]. Copyright 2010, Schmidt)

Figure 10.12 shows a side view of a Co F1-STFs on top of a Si substrate tilted by 15° and a photograph of a 100 nm thick F1-STF deposited under similar conditions as sample #1 onto a glass slide. The stack was then placed on a printed image and details of the image can still be seen, which illustrates the high degree of transparency of ferromagnetic Co F1-STFs.

10.5.2 Determination of Structural Parameters by Ellipsometry Data Analysis

As will be shown for different STF samples in Sect. 10.6, Euler angles θ and φ, determining the orientation of major polarizability axes, are identical to the geometrical slanting angle and the rotation of the slanting plane with respect to the laboratory coordinate frame, respectively. Furthermore, the thickness d is obtained as a direct result of the best-match model calculations. These parameters can be individually determined for each layer within complex layered STFs and are in very good agreement with SEM image analysis.

10.6 Optical Properties of Metal Sculptured Thin Films

This section discusses the determination of intrinsic optical properties of metal STFs deposited on flat Si substrates. Monoclinic optical properties of slanted columnar thin films (F1-STFs) are determined by the analysis of spectroscopic generalized ellipsometry measurements with a homogeneous biaxial layer approach. A universality regarding the monoclinic and intrinsic major polarizability functions is found in F1-STFs. Complex layered STFs can be considered as cascaded F1-STFs with different slanting directions. Therefore, optical properties of manifold STFs can be predicted by using the optical model developed here for the F1-STFs. The optical model for complex STFs comprises appropriately stacked model layers accounting for F1-STFs in a modular conception thereby mimicking the cascaded STF geometry. It is shown that the piecewise homogeneous biaxial layer approach is also valid for hollow-core helical STFs. All samples discussed here are analyzed using the multi-sample-configuration analysis method including multiple in-plane orientations ϕ into the best-match model calculations.

10.6.1 Monoclinicity of Metal F1-STFs

Analysis of spectroscopic GE measurements have revealed that metal F1-STFs exhibit monoclinic optical properties. A model based on the specific thin film geometry is proposed, which explains this monoclinic behavior based on dielectric polarization charge coupling effects across neighboring slanted but electrically isolated nanocolumns. A detailed explanation is presented on the example of an achiral Co F1-STF (sample #1).

10.6.1.1 Optical Constants

Experimentally obtained Mueller matrix spectra (as described in Sect. 10.3) for the Co F1-STF were analyzed with an optical model containing a single anisotropic

(biaxial) layer on top of the substrate as discussed in Sect. 10.3.6. Consequently, the computed wavelength-independent Euler angles (φ, θ, ψ) that transform the Cartesian coordinate system (x, y, z) into the sample coordinates $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ represent the orientation of the nanostructure and the internal angles α, β, γ differentiate between orthorhombic, monoclinic, or triclinic properties of the film.

The linear polarizability response of F1-STFs due to an electric field $\mathbf{E}$ is a superposition of contributions along certain directions and may be written as $\mathbf{P} = \varrho_a \mathbf{a} + \varrho_b \mathbf{b} + \varrho_c \mathbf{c}$. Axis $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and major polarizability functions $\varrho_a, \varrho_b, \varrho_c$ in F1-STFs are determined by nanostructure geometry rather than crystallographic unit cells. The experiment shows, that for F1-STFs, $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ span a monoclinic coordinate system with $\mathbf{c}$ along the long axis of the columns, $\mathbf{a}$ perpendicular to $\mathbf{b}$ and $\mathbf{c}$ and parallel to the substrate surface, and a monoclinic angle β between $\mathbf{c}$ and $\mathbf{b}$. Along these axes, major polarizability functions $1 + \varrho_j = \varepsilon_j$ $(j = a, b, c)$ can be determined which may vary with frequency (see also Sect. 10.3.4). ϱ_j were first determined on a wavelength-by-wavelength basis. The point-by-point extracted data have then been parameterized with MDFs (Lorentzian oscillators and Drude terms) and the best-match model calculation procedure was repeated facilitating the MDF approach for final results.

For the Co F1-STF, functions ϱ_a and ϱ_b were parameterized with three Lorentz oscillators, respectively, and four Lorentz oscillators and one Drude term were incorporated for ϱ_c. Individual parameters are listed in Table 10.9.

The MDF parameters are presented exemplarily only without further evaluation. The identification of physically meaningful quantities is not within the scope of this chapter. Hence, for other STF discussed below MDF parameters are not listed individually.

Note that principal refractive indices and extinction coefficients along major polarizability axes $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ do not exist for monoclinic (and triclinic) materials, instead intrinsic polarizability functions $(\varrho_a, \varrho_b, \varrho_c)$ need to be discussed. Intrinsic complex-valued optical constants (N_a, N_b, N_c) imply that directions exist, along which wave propagation with such indices can be obtained (and measured upon refraction). However, such directions do not exist in materials with monoclinic and triclinic symmetry and therefore, N_a, N_b, N_c do not exist. Alternatively, since presentation of optical functions in terms of $\varrho_a, \varrho_b, \varrho_c$ is not common, in this chapter, $N_j \equiv \sqrt{1 + \varrho_j} = n_j + i k_j$ with $j = a, b, c$, are presented and may be considered as effective optical constants.

Experimental data. Figure 10.13 depicts selected GE Mueller matrix data for four different angles of incidence obtained from a ferromagnetic Co F1-STF (sample #1) at an exemplary wavelength of $\lambda = 630$ nm. The film was measured approximately 1 h after deposition, therefore it is assumed that data are representative for pure Co nanostructures and a possible oxide shell can be neglected at this point. The graphs show non-redundant[20] Mueller matrix elements of the measured 4×3 part of the

---

[20] In case the sample under investigation does not exhibit non-reciprocal properties, Mueller matrix elements not shown in Fig. 10.13 can be obtained by symmetry operations: $M_{21}(\varphi) = M_{12}(\varphi + \pi)$ and $M_{3j}(\varphi) = -M_{j3}(\varphi + \pi)$ with $j = 1, 2$. No inversion operation is necessary to convert $M_{12}(\varphi)$
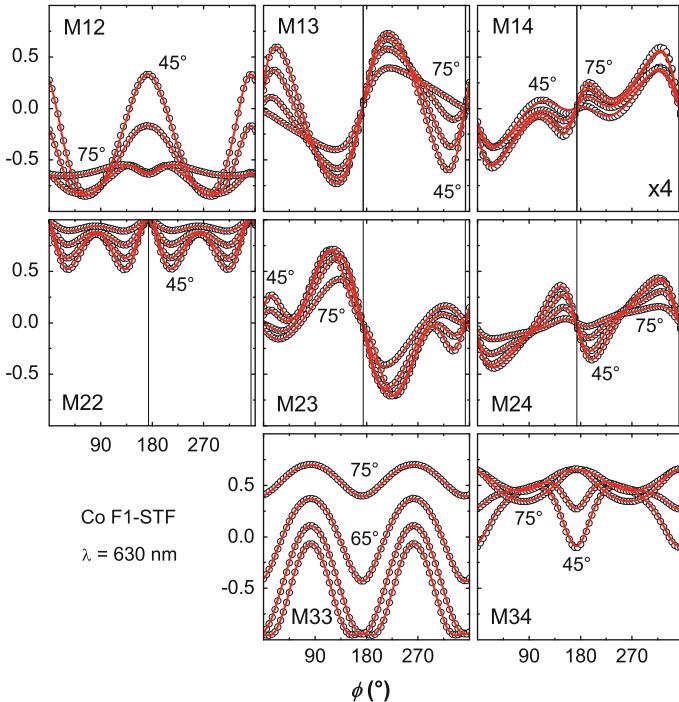
Fig. 10.13 Exemplary experimental (*circles*) and best-match calculated (*solid lines*) GE data of a Co F1-STF versus sample azimuth angle ϕ at $\lambda = 630$ nm. The four graphs for each Mueller matrix element are four different angles of incidence $\Phi_a = 45°, 55°, 65°, 75°$. Note the pseudo-isotropic orientations (indicated by the *vertical bars*); i.e., the sample positions ϕ at which all off-diagonal Mueller matrix data ($M_{13}, M_{14}, M_{23}, M_{24}$) vanish and $M_{22} = 1$, which is the case near $\phi = 0°$ and $\phi = 180°$. Such orientations occur when the slanting plane coincides with the plane of incidence. Element M_{14} is magnified $\times 4$ (Reprinted with permission from [40]. Copyright 2010, Schmidt)

matrix versus sample azimuth. Model and experimental data are in excellent agreement for all wavelengths in the investigated spectral region from 400 to 1700 nm. The off-diagonal Mueller matrix data ($M_{13}, M_{14}, M_{23}, M_{24}$) exhibit the highly anisotropic nature of the F1 nanostructures. These elements are zero for all angles of incidences Φ_a at all wavelengths for isotropic samples. So-called pseudo-isotropic sample orientations can be identified at $\phi \approx 0°$ and $\phi \approx 180°$, which coincide with orientations of the sample when the slanting direction of the nanocolumns is parallel to the plane of incidence (Euler angles $\varphi = 90°$ and $\varphi = 270°$). Hence, no *p*-polarized light is converted into *s*-polarized light and vice versa in this particular setup. It can be seen that there is no repetition of data over one full rotation except for symmetry with respect to pseudo-isotropic ϕ positions, and data over one full in-plane rotation should be measured in order to fully evaluate the optical properties. The element with the

---

into $M_{21}(\varphi + \pi)$ because these elements depend on the symmetric cos function only whereas this is not true for all other elements. See for example (10.10). π denotes a sample rotation by 180°.
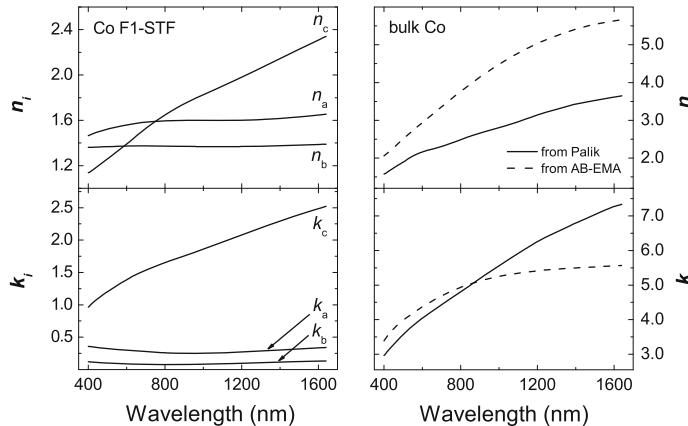
Fig. 10.14 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with monoclinic arrangement of a Co F1-STF, and isotropic bulk optical constants of Co. The bulk properties taken from Palik [106] are also compared to wavelength-by-wavelength determined optical constants from the AB-EMA approach (Sect. 10.6.4). From the anisotropic optical constants of the F1-STF only those along axis **c** have a similar dispersion with the isotropic bulk optical constants

smallest amplitude is M_{14} (multiplied by a factor of 4), which carries information on circularly polarized light. Interestingly, M_{22} exhibits a four-fold symmetry, whereas M_{33} only shows a two-fold symmetry with respect to sample in-plane orientation.

Functions N_a, N_b, N_c obtained with parameterized MDFs (Table 10.9) are depicted in Fig. 10.14. Refractive indices n_j and absorption coefficients k_j are pertinent to axes **a**, **b**, and **c** of the coordinate system shown in Fig. 10.15 and differ drastically from those of bulk material (right panel).[21] Note that the reported optical constants are effective values since the optical model assumes a homogeneous layer. Strong birefringence and dichroism can be observed in the investigated spectral region between all polarizabilities. The index of refraction n_c along the slanted nanocolumns (**c**-axis) is intersecting with both other refractive indices n_b and n_a (Fig. 10.14). Such intersections are not present for the extinction coefficients k_j which follow the same order as the refractive indices above $\lambda = 800$ nm ($k_c > k_a > k_b$). In general, n_c and k_c have a strong wavelength dependence in contrast to the optical constants along the **a**- and **b**-axes. There is almost no absorption along axis **b**. Amongst the three directions, the polarizability along the **c**-axis of the nanostructures is the only one with a similar dispersion compared to bulk Co.

The optically determined structural parameters film thickness ($d = 106.9 \pm 0.1$ nm) and inclination angle (Euler angle $\theta = 63.68 \pm 0.01°$) are in very good agreement with data from SEM micrograph analysis ($d = 115 \pm 5$ nm and $\theta = 65 \pm 3°$, respectively). Note that the GE thickness parameter is an optical thickness, which is generally less than SEM estimates. Even though thickness estimates from

---

[21] Optical constants for bulk Co have been taken from Palik [106].
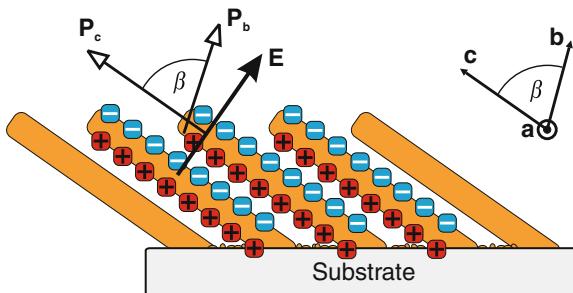
Fig. 10.15 Schematic drawing of the electric charge transfer mechanism causing monoclinic optical properties. The polarization $\mathbf{P}_b$ due to an electric field $\mathbf{E}$ within the plane of incidence is tilted toward the substrate normal (Reprinted with permission from [40]. Copyright 2010, Schmidt)

cross-sectional SEM images are typically obtained from a different region of the thin film than that measured optically, the optical thickness is assumed to originate mainly from differences between the idealized model and the real STF. In particular, the optical model at this present point does not account for a nucleation layer (typically <5 nm) or a surface roughness layer, for example. Besides that, in the homogenization approach, the columns are assumed to be homogeneous along the long axis. These idealizations might affect the optically determined thin film thickness parameter (for SEM micrographs of F1-STFs see Fig. 10.11).

The investigated Co F1-STF possesses monoclinic optical properties with an angle $\beta = 83.69 \pm 0.09°$ ($\alpha = \beta = 90°$). The MSE, which is a measure for the quality of the match between model and experimental data is 7.77. The Euler angle $\psi = 0$ and hence not included in the best-match model calculations. This can be seen by the pseudo-isotropic orientations in the angle-resolved experimental Mueller matrix data (Fig. 10.13). Such orientations with no conversion between p- and s-polarization states would not occur if the sample coordinate system was further rotated by ψ (**b** out of the slanting plane, and **a** out of the substrate surface).

10.6.1.2 Monoclinicity

The monoclinic angle reveals an intrinsic optical thin film property of F1-STFs and results also from the specific arrangement of the coherently tilted nanostructures. As discussed in Sect. 10.3.4, monoclinic properties can only be identified by considering a broad spectral range during data analysis.

The monoclinic angle can be understood as a characteristic due to charge transfer leading to anisotropic charge distribution in slanted columnar thin films (F1-STFs) prepared from electrically conductive materials. At the bottom of the structure, charge exchange is possible due to a conducting nucleation (wetting) layer whereas charge transfer is not possible at the isolated top of the column. The slanting of nanocolumns causes an anisotropic distribution of charges due to the mutual screening of charge dipoles across adjacent columns (Fig. 10.15). Therefore, the effective overall dipole

moment for electric fields perpendicular to the columns and within the slanting plane ($\mathbf{P}_b$) is tilted toward the surface normal. Further examples of monoclinic F1-STF are discussed in Sect. 10.6.2. It can be expected that the monoclinic angle depends on the overall film thickness, the tilting angle of the columns, the intercolumnar spacing, and the properties of the nucleation layer. Systematic investigations have yet to be conducted, however, trends observed in currently available samples hint in this direction.

10.6.2 Universality of Metal F1-STFs

Analysis have shown that there are several characteristics common for metal F1-STF investigated here. Angle-resolved Mueller matrix data for F1-STF prepared from different materials exhibit two pseudo-isotropic orientations, which occur when the slanting plane is parallel to the plane of incidence. The consequence is that the Euler angle $\psi = 0$ and can be excluded from best-match model calculations. For each Mueller matrix element versus ϕ, except element M_{34}, the graph representing data for $\Phi_a = 75°$ exhibits less azimuthal variations than $\Phi_a = 45°$.

The attained refractive indices and extinction coefficients along major polarizability axes for each material show similar dispersion relations and exhibit strong birefringence and dichroism. The order of the refractive indices in the near-infrared spectral region found here is always identical with $n_c > n_a > n_b$, and which is also the order of the extinction coefficients within the entire investigated spectral region. n_c and k_c exhibit bulk-like dispersion, whereas k along axes $\mathbf{a}$ and $\mathbf{b}$ shows almost no absorption. n_c always intersects n_a and then n_b within the visible spectral region. Furthermore, all investigated F1-STFs exhibit monoclinic optical properties.

The observed universality is demonstrated here for two additional F1-STFs made from titanium (sample #2) and supermalloy ($Ni_{80}Fe_{15}Mo_5$, sample #3). Experimentally obtained angle-resolved Mueller matrix data for both films have been analyzed in the same manner as explicitly described for Co F1-STFs in Sect. 10.6.1.1.

10.6.2.1 Titanium F1-STFs

The optical model for Ti F1-STFs contains parameterized MDFs for optical constants along the major polarizability axes $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ of a monoclinic system. ϱ_a and ϱ_b can be parameterized by two and three Lorentzian oscillators, respectively, and ϱ_c by four Lorentzian oscillators and a Drude term.

Refractive indices n_j and absorption coefficients k_j, depicted in Fig. 10.14, differ significantly from those of bulk material (right panel, taken from Palik [106]). In general, the optical constants have similar properties and dispersion relations compared to the ones from ferromagnetic Co F1-STFs (Sect. 10.6.1.1). Again, n_c and k_c exhibit similarities with bulk optical constants from Ti. Especially the local maxi-
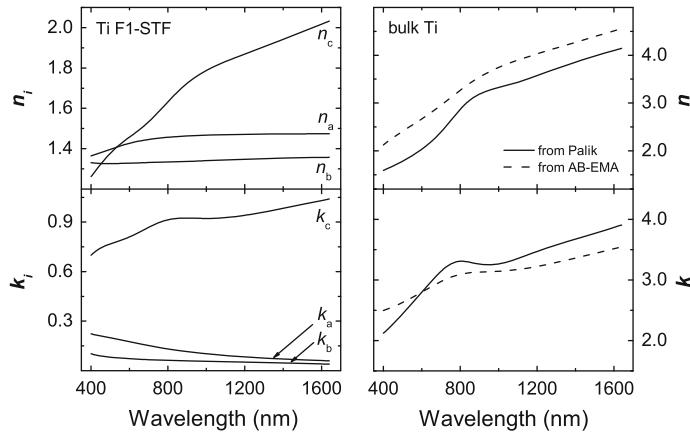
Fig. 10.16 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with monoclinic arrangement of a Ti F1-STF, and isotropic bulk optical constants of Ti. The bulk properties taken from Palik [106] are also compared to wavelength-by-wavelength determined optical constants from the AB-EMA approach (Sect. 10.6.4)

mum in k around 800 nm caused by interband transitions is present, although not as pronounced as in the case of bulk Ti [107].

The investigated Ti F1-STF possesses monoclinic optical properties with an angle $\beta = 80.2 \pm 0.1°$ ($\alpha = \gamma = 90°$). Other structural parameters determined from optical analysis and in comparison to SEM micrograph analysis are summarized in Sect. 10.6.4, Table 10.7.

10.6.2.2 Supermalloy (Ni$_{80}$Fe$_{15}$Mo$_5$) F1-STFs

The optical model for supermalloy F1-STFs contains parameterized MDFs for major polarizabilities ϱ_a, ϱ_b, ϱ_c, where for model analysis two and three Lorentzian oscillators are incorporated into ϱ_a and ϱ_b, respectively; ϱ_c contains two Lorentzian oscillators and a Drude term (Fig. 10.16).

Refractive indices n_j and absorption coefficients k_j, depicted in Fig. 10.17, differ significantly from those of bulk material (right panel).[22] n_c and k_c exhibit similarities with bulk optical constants, however, the broad shoulder in n_c between 800 and 1200 nm (Fig. 10.17) is flattened out for the F1-STF. Supermalloy STF optical constants are almost identical to the ones from ferromagnetic Co F1-STFs (Fig. 10.14).

The obtained monoclinic angle $\beta = 89.52 \pm 0.08°$ indicates almost purely orthorhombic properties. Structural parameters determined optically through best-match model calculations are summarized in Sect. 10.6.4, Table 10.8.

---

[22] Bulk optical constants have been generated with an isotropic Bruggeman EMA ($L_{iso}^D = \frac{1}{3}$ for spherical inclusions) and optical constants for Ni, Fe, and Mo were taken from Palik [106]. For further details on Bruggeman EMA see also Sect. 10.3.6.
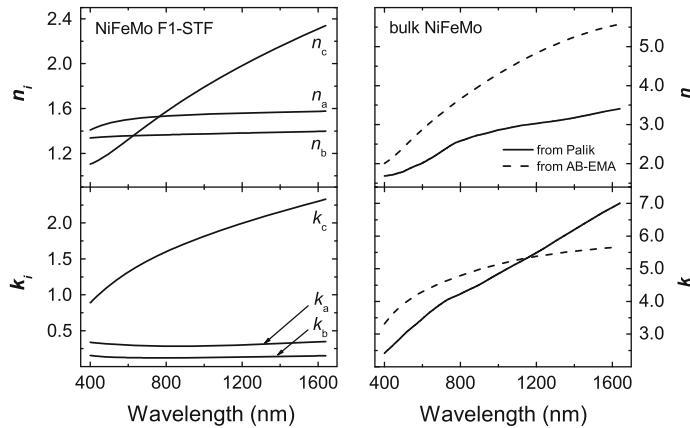
Fig. 10.17 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with monoclinic arrangement of a supermalloy F1-STF, and isotropic bulk optical constants of $Ni_{80}Fe_{15}o_5$. The bulk properties taken from Palik [106] are also compared to wavelength-by-wavelength determined optical constants from the AB-EMA approach (Sect. 10.6.4)

10.6.3 Modularity of Complex Sculptured Thin Films

Optical constants determined for an F1-STF (for a certain material and deposition setup) can be used in principle to predict the optical response for any STF with an arbitrary sequence of rotation steps during growth (lFx-STFs as well as H-STFs). Complex layered STFs can be considered as cascaded F1-STFs with different slanting directions. Therefore, optical properties of manifold STFs can be predicted by using the optical model discussed above for F1-STFs as a basic module. The optical model for complex STFs then comprises appropriately stacked model layers accounting for F1-STFs in a modular conception thereby mimicking the cascaded STF geometry. The modularity is exemplary discussed for a total of four differently grown complex layered STF from cobalt and titanium. Furthermore, it is shown that this piecewise homogeneous biaxial layer approach is also valid for hollow-core helical STFs.

Information on growth conditions for each sample discussed here can be found in Sect. 10.4 and Table 10.3.

10.6.3.1 Chevron-like Sculptured Thin Films (2F2-STFs)

Morphologically one step advanced with respect to F1-STFs are F2-STFs. These nanostructured thin films, also called chevrons or zig-zags, consist of m layers of slanted columns with alternating tilting direction and may therefore be considered as cascaded F1-STFs (Fig. 10.18). The change in geometry required for this growth can be obtained by rotating the substrate by half a turn around its normal axis. Note that all layers of a chevron nanostructure share a common plane containing the slanted
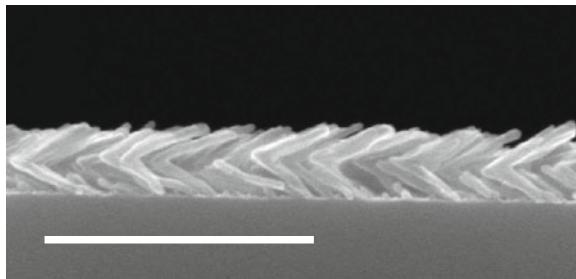
Fig. 10.18 SEM micrograph of the Ti 2F2-STF. The scale bar is 500 nm (Reprinted with permission from [40]. Copyright 2010, Schmidt)

column and the substrate normal and therefore, 2F2-STFs are achiral. Exemplarily, the optical analysis of a 2F2-STF from titanium (sample #4) is discussed here.

In contrast to the previously discussed F1-STFs (Sect. 10.6.2.1), the optical model for the achiral 2F2-STF is composed of two anisotropic (biaxial) layers with opposite azimuthal orientation ($\varphi_1 = -\varphi_2$), which account for both slanting directions. Hence, the thin film is virtually separated into two F1-STFs for optical analysis. The single set of major polarizabilities ϱ_a, ϱ_b, ϱ_c used in both layers has been parameterized with MDFs. Functions ϱ_a and ϱ_b contain two and three Lorentz oscillators, respectively, and four Lorentz oscillators and one Drude term were incorporated for ϱ_c.

Selected experimental and best-match model calculated GE Mueller matrix data for four different angles of incidence obtained from the Ti 2F2-STF are depicted in Fig. 10.19. A single wavelength $\lambda = 630$ nm was chosen for the graphs and similar results with an excellent agreement between model and experimental data have been obtained for all wavelengths in the investigated spectral region from 400 to 1700 nm. Also for 2F2-STFs pseudo-isotropic sample orientations can be identified at $\phi \approx 0°$ and $\phi \approx 180°$, which coincide with sample directions when the set of slanting planes is parallel to the plane of incidence. This reveals that also for a two-layered STF, where both slanting directions share the same plane, similar properties as for F1-STFs can be found.

The obtained set of optical constants common for both biaxial layers is in very good agreement with optical constants determined from Ti F1-STFs deposited in a comparable manner (Fig. 10.20). Even the shoulder in n_c and k_c around 800 nm caused by interband transitions [107] can still be observed. Differences between Ti 2F2- and F1-STF optical constants are attributed to structure non-idealities, differently chosen deposition rates, and the assumed ideal interface between bottom and top layers in the optical model.

10.6.3.2 L-shape Sculptured Thin Films (2F4-STFs)

The most primitive chiral STF is a 2F4-STF, "L-shape" STF, where handedness is introduced based on the rotation direction of the substrate during growth. The required
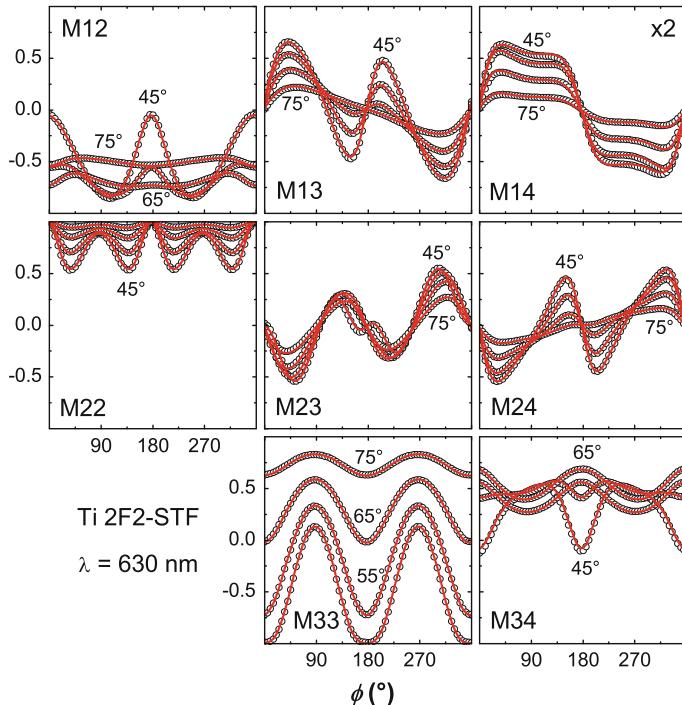
Fig. 10.19 Exemplary experimental (*circles*) and best-match calculated (*solid lines*) GE data of a NiFeMo F1-STF versus sample azimuth angle ϕ at $\lambda = 630$ nm. The four graphs for each Mueller matrix element are four different angles of incidence $\Phi_a = 45°, 55°, 65°, 75°$. Note the pseudo-isotropic orientations; i.e., the sample positions ϕ at which all off-diagonal Mueller matrix data ($M_{13}, M_{14}, M_{23}, M_{24}$) vanish and $M_{22} = 1$, which is the case near $\phi = 0°$ and $\phi = 180°$. Such orientations occur when the slanting planes of both layers coincide with the plane of incidence. Element M_{14} is magnified ×2 (Reprinted with permission from [40]. Copyright 2010, Schmidt)

change in geometry for this growth can be obtained by rotating the substrate by ±90° around its normal axis after depositing the first layer. If the second layer is rotated counterclockwise by 90° with respect to the first one, the nanostructured L-shape thin film is termed here 2F4$^+$-STF (right-handed; sample #5; Fig. 10.21), otherwise 2F4$^-$-STF (left-handed, sample #6). Both, a right- and left-handed 2F4-STFs from titanium are discussed and compared.

Similar to 2F2-STFs, the optical model for the chiral 2F4-STFs is composed of two anisotropic (biaxial) layers, which are azimuthally rotated with respect to each other by Euler angle ±$\delta\varphi$. Hence, the thin film is virtually separated into two model layer F1-STFs for optical analysis.

Here, the single set of major polarizabilities $\varrho_a, \varrho_b, \varrho_c$ used in both biaxial model layers is similar for the right- and lefthanded chiral STF and has been parameterized with MDFs. Functions ϱ_a and ϱ_b contain two Lorentz oscillators each, and four Lorentz oscillators and one Drude term were incorporated for ϱ_c.
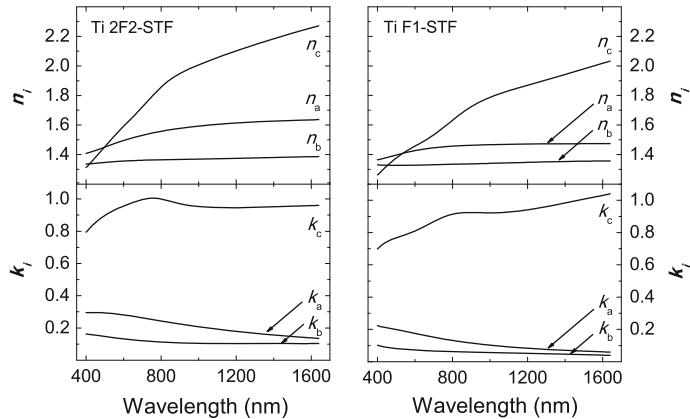
Fig. 10.20 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with monoclinic arrangement of a Ti 2F2-STF (*left panel*) in comparison with Ti F1-STF (*right panel*) (Reprinted with permission from [40]. Copyright 2010, Schmidt)
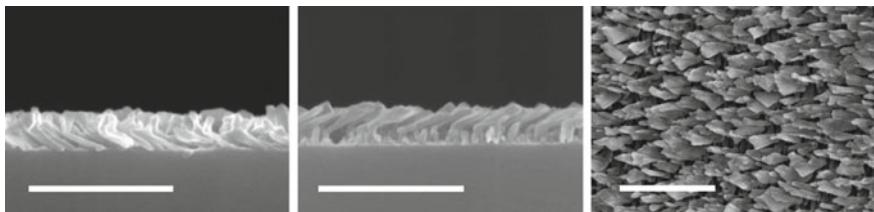
Fig. 10.21 Cross-section and top-view SEM micrographs of the Ti 2F4$^+$-STF. The *left* and *middle* images are orientated such that the first and the second layer are in the plane of the image, respectively. The top-view image (*right*) illustrates the high porosity and reveals that structure fanning (broadening) is comparable to that of a single layer (Fig. 10.11). Scale bars are 500 nm (Reprinted with permission from [40]. Copyright 2010, Schmidt)

Experimental and best-match model calculated GE Mueller matrix data for four different angles of incidence obtained from the Ti 2F4$^+$-STF and 2F4$^-$-STF are shown in Fig. 10.22a and b, respectively. Each graph depicts eight non-redundant Mueller matrix elements versus sample azimuth at an exemplary wavelength $\lambda = 630$ nm. For ideal samples, off-diagonal elements M_{13}, M_{14}, M_{23}, and M_{24} from the Ti 2F4$^+$-STF can be transferred into the same elements of the Ti 2F4$^-$-STF by inversion around (180°,0), whereas a mirror operation at $\phi = 180°$ is necessary for elements M_{12}, M_{22}, M_{33} and M_{34}. Another characteristics for this chiral STF is that there is no symmetry within a single Mueller matrix elements as compared to all other achiral STFs discussed above. Most importantly, in both chiral STFs no pseudo-isotropic orientation can be observed.

Optical constants obtained from both chiral nanostructured films depicted in Fig. 10.23 are very similar and show the same dispersion relation. Differences between left and right handed 2F4-STF can be observed in absolute values of n_c
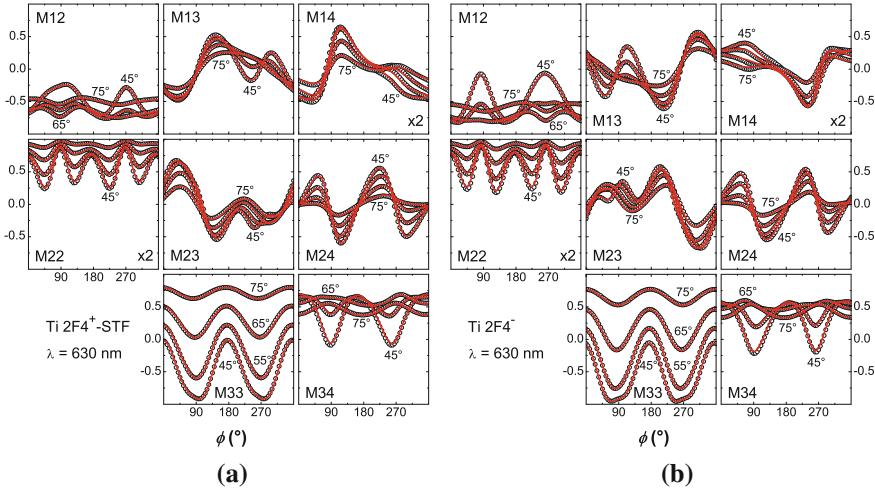
Fig. 10.22 Exemplary experimental (*circles*) and best-match calculated (*solid lines*) GE data of a Ti 2F4$^+$-STF (**a**) and a Ti 2F4$^-$-STF (**b**) versus sample azimuth angle ϕ at $\lambda = 630$ nm. The four graphs for each Mueller matrix element are four different angles of incidence $\Phi_a = 45°, 55°, 65°, 75°$. Note that for these chiral STFs no pseudo-isotropic orientation is present. Elements M_{14} and M_{22} are magnified $\times 2$ (Reprinted with permission from [40]. Copyright 2010, Schmidt)
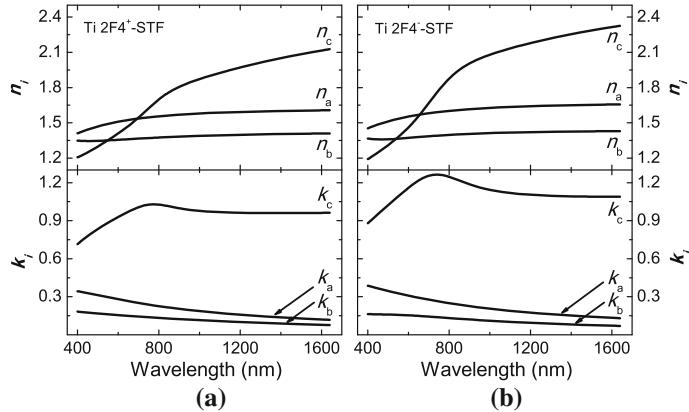
Fig. 10.23 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with monoclinic arrangement of **a** a Ti 2F4$^+$-STF and **b** a Ti 2F4$^-$-STF. Optical constants depicted here are in good agreement with optical constants of Ti F1-STFs and Ti 2F2-STFs shown in Fig. 10.16 and Fig. 10.20, respectively (Reprinted with permission from [40]. Copyright 2010, Schmidt)

and k_c. In general, both sets of optical constants are in good agreement with the optical constants obtained from Ti F1-STFs (Fig. 10.16). Furthermore, n and k along major polarizability axes **a**, **b**, **c** of the 2F4$^+$-STF are in excellent agreement with optical constants obtained from the 2F2-STF (Fig. 10.20). Differences may be well

Table 10.4 Best-match model results for the Ti 2F4$^+$- and 2F4$^-$-STF and results from SEM micrograph analysis of Ti 2F4$^+$

Parameters	2F4$^+$	2F4$^-$	SEM 2F4$^+$
d_1 (nm)	65.3(3)	63.7(2)	63(3)
d_2 (nm)	50.7(1)	50.7(1)	65(5)
θ_1 (°)	60.4(1)	60.5(1)	63(6)
θ_2 (°)	62.0(1)	65.4(1)	67(5)
β_1 (°)	78.0(2)	76.5(3)	–
β_2 (°)	90	90	–
$\delta\varphi$ (°)	73.3(2)	85.7(2)	–
MSE	13.34	15.01	–

The error limits given in parentheses denote the uncertainty of the last digit (90 % reliability)

explained with structure non-idealities due to non-constant evaporation source conditions, differences in deposition rates (Sect. 10.4), and the model assumption of ideal interfaces between bottom and top layers.

Other optically determined best-match parameters for both 2F4$^+$-STF and 2F4$^-$-STF are summarized in Table 10.4. Thickness d_j and slanting angle θ_j for bottom ($j = 1$) and top ($j = 2$) layer, and the monoclinic angle β_1 are in very good agreement between both films. The main difference between both STFs is the angle between the two deposition steps $\delta\varphi$, which is nominally 90°. A deviation from 90° indicates that the two slanting planes are not orthogonal as expected from the growth parameters. It is not clear where the deviation of almost 17° for the 2F4$^+$-STF is coming from but a possible source could be sample manipulator rotation non-idealities (see also Sect. 10.6.3.4). Best-match model calculations revealed that β_2 was not changing and stayed constant around $90.0 \pm 0.1°$. Consequently it was not included into the final best-match calculation. The monoclinic angle of the bottom layer (β_1) is comparable to the monoclinic angle of the Ti F1-STF discussed in Sect. 10.6.2.1. Hence, the second layer with perpendicular slanting direction has almost no influence on the monoclinic properties of the bottom layer in contrast to 2F2-STFs (Sect. 10.6.3.1). On the other hand, similar to 2F2-STFs, the top layer exhibits orthorhombic properties because there is no lateral conductive channel along which the necessary charge transfer could occur.

Best-match model parameters are well within the range determined by SEM micrograph analysis and the only discrepancy is the thickness of the second layer. This might be due to differences in density between both layers and hence, since identical optical constants for both layers are assumed, the denser layer is optically less thick.

10.6.3.3 U-shape Sculptured Thin Films (3F4-STFs)

Adding another layer to a 2F4-STF by further rotating the substrate 90° in the same direction as for the second layer results in a chiral U-shaped nanostructure abbreviated according to the proposed nomenclature 3F4-STF. Hence, slanting planes of top and
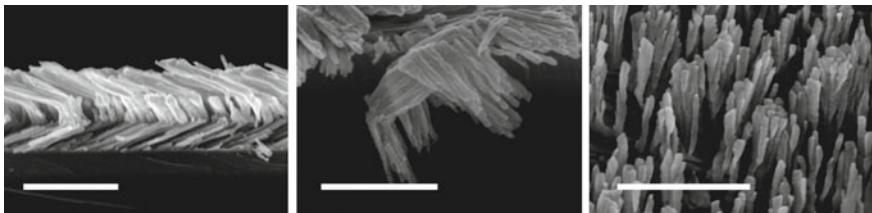
Fig. 10.24 SEM micrographs of the Co 3F4$^+$-STF: (*left*) cross-section view where both top and bottom layer are in the plane of the image, (*middle*) view on a structure ruptured from the substrate, which exhibits the homogeneous columns in form of a U-shape, and (*right*) top-view image revealing that some column bunching occurs but also free standing structures can be observed. Scale bars are 500 nm (Reprinted with permission from [40]. Copyright 2010, Schmidt)

bottom layer are parallel and nanocolumns are tilted in opposite directions, and the slanting plane of the sandwiched middle layer is perpendicular to top and bottom layers. The optical analysis of a 3F4$^+$-STF (sample #7) from cobalt is exemplarily discussed (Fig. 10.24).

For the optical analysis the three-dimensional chiral Co 3F4$^+$ nanostructures were decomposed into three individual F1-STF layers. Consequently, the optical model was similar to the 2F4$^+$-STF but with a third biaxial layer accounting for the third deposition sequence (for model details see Sects. 10.3.6 and 10.6.3.2). The single set of major polarizabilities $\varrho_a, \varrho_b, \varrho_c$ used in all three biaxial model layers has been parameterized with MDFs. Functions ϱ_a and ϱ_b contain three Lorentz oscillators each, and two Lorentz oscillators and one Drude term were incorporated for ϱ_c. For this particular thin film, the large error associated with the monoclinic angle and correlation between β and θ hindered best-match model calculations with a monoclinic arrangement. Therefore, intrinsic biaxial properties were assumed to be orthorhombic ($\alpha = \beta = \gamma \stackrel{!}{=} 90°$). Experimental and best-match model calculated GE Mueller matrix data for four different angles of incidence obtained from the Co 3F4$^+$-STF are shown in Fig. 10.25. The graph depicts selected Mueller matrix elements versus sample azimuth at an exemplary wavelength $\lambda = 630$ nm. Similar to the chiral Ti 2F4$^\pm$-STFs, no pseudo-isotropic orientation can be observed. Note in particular the discrepancy of $\Phi_a = 45°$ in element M_{24}, and which does not intersect with the other Φ_a-graphs at $M_{24} = 0$. However, sample orientations with minimum $p - s$ mode coupling can be identified near $\phi = 180°$ and $\phi = 360°$, which coincide with sample directions when both slanting planes of bottom and top layer are parallel to the plane of incidence. Note that value and position of the minimum are wavelength-dependent. This behavior of orientations with minimum $p - s$ mode coupling is not due the non-perfect alignment $\delta\varphi$ between two deposition steps (Table 10.5) but rather a characteristic of 3F4$^\pm$-STFs, as calculations with $\delta\varphi_{jk} = 90°$ have revealed.

The set of optical constants derived under conditions of the above described model, common for all three biaxial layers, is in good agreement with optical constants determined from Co F1-STFs (Fig. 10.26). Deviations are attributed to the
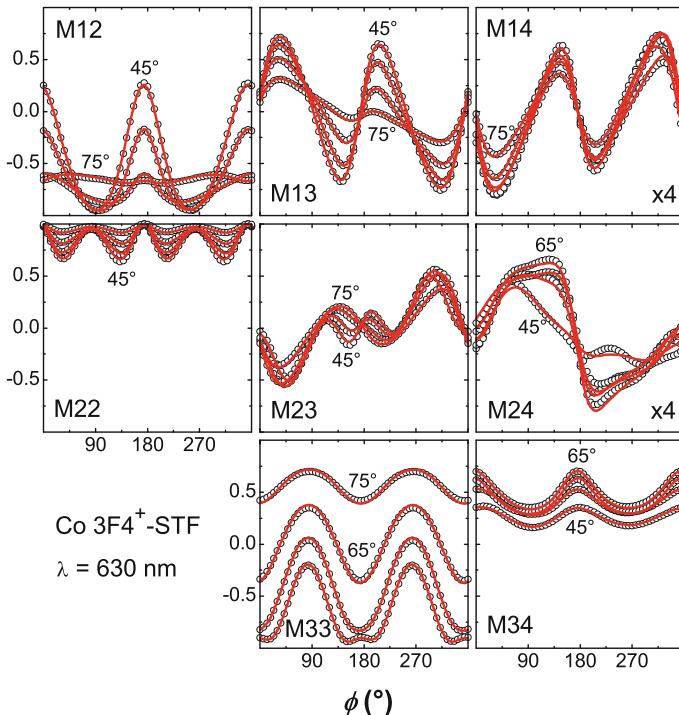
Fig. 10.25 Exemplary experimental (*circles*) and best-match calculated (*solid lines*) GE data of a Co 3F4$^+$-STF versus sample azimuth angle ϕ at $\lambda = 630$ nm. The four graphs for each Mueller matrix element are four different angles of incidence $\Phi_a = 45°, 55°, 65°, 75°$. Note that for this chiral STF no pseudo-isotropic orientation is present. Elements M_{14} and M_{24} are magnified $\times 4$ (Reprinted with permission from [40]. Copyright 2010, Schmidt)

idealized interface between subsequent layers and the assumption of an orthorhombic symmetry for the optical model of the 3F4$^+$-STF.

Further optically determined best-match parameters for the 3F4$^+$-STF are summarized in Table 10.5. The MSE is considerably higher than for F1-STF, which is probably due to structure non-idealities such as column bunching as well as the non-consideration of possible monoclinic properties.

10.6.3.4 Helical Sculptured Thin Films (H-STFs)

In contrast to F-STFs, which are fabricated by sequential substrate rotations, helical chiral STFs are the consequence of a slow continuous substrate rotation. Optical analysis of a right- and left-handed Co H-STF with one turn each, 1H$^+$ (sample #8) and 1H$^-$ (sample #9), are discussed exemplarily here. Cross-sectional and top-view SEM images of both H-STFs are depicted in Fig. 10.27.
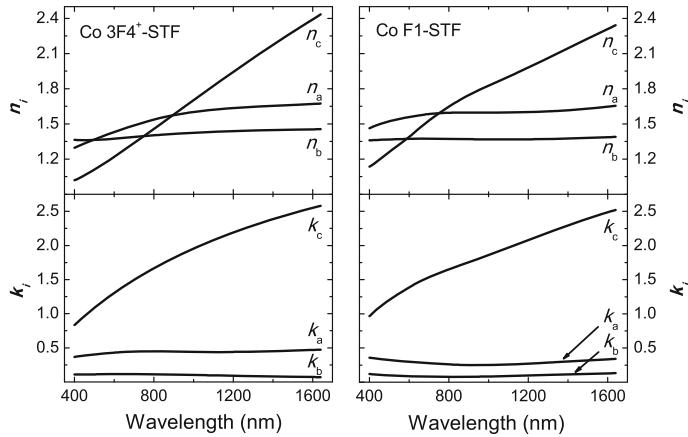
Fig. 10.26 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with orthorhombic arrangement of a Co 3F4$^+$STF (*left panel*) in comparison with Co F1-STF (monoclinic arrangement; *right panel*) (Reprinted with permission from [40]. Copyright 2010, Schmidt)

Table 10.5 Best-match model results for the Ti 3F4$^+$ in comparison with values determined by SEM micrograph analysis

Parameters	GE	SEM
d_1 (nm)	99(1)	135(10)
d_2 (nm)	87.1(8)	130(15)
d_3 (nm)	117.2(4)	130(15)
θ_1 (°)	71.5(4)	66(3)
θ_2 (°)	68.4(2)	65(5)
θ_3 (°)	66.6(1)	65(7)
$\delta\varphi_{12}$ (°)	93.9(6)	–
$\delta\varphi_{23}$ (°)	75.3(2)	–
MSE	21.56	–

Subscripts 1, 2, and 3 denote bottom, middle, and top layer, respectively. $\delta\varphi_{jk}$ is the angle between layer j and k. The error limits given in parentheses denote the uncertainty of the last digits (90 % reliability)

The optical model for H-STFs consists of multiple sublayers (slices) with dielectric function tensor descriptions rotated stepwise with respect to the sample normal between adjacent layers according to the handedness. This in-plane rotation φ (ideally homogeneous) from the substrate interface to the top of the structure accounts for the helical nature of the chiral nanostructures (for further model details see Sect. 10.3.6.3). The major polarizabilities ϱ_a, ϱ_b, ϱ_c, common to all sublayers, have been parameterized with MDFs. For both H-STFs with different handedness, function ϱ_b contains two Lorentz oscillators and ϱ_c one Lorentzian oscillator and a Drude term, respectively, ϱ_a was parameterized with two and three Lorentzian oscillators for 1H$^+$- and 1H$^-$-STF, respectively.
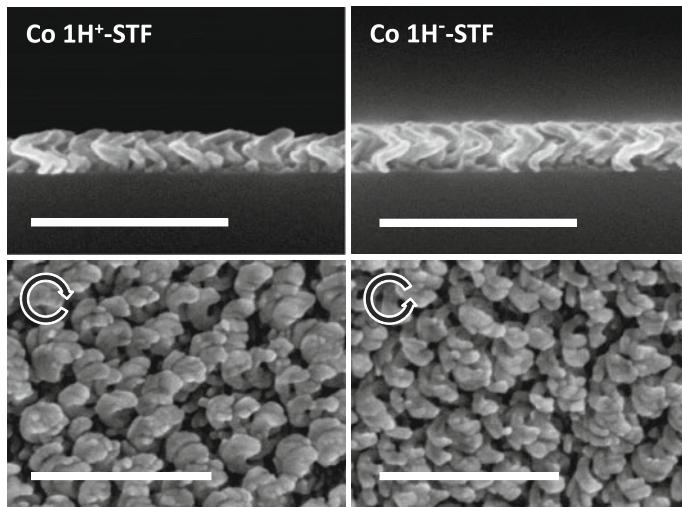
Fig. 10.27 Cross-sectional (*top row*) and *top-view* (*bottom row*) SEM micrographs of the Co 1H$^+$- and Co 1H$^+$-STF. The *arrows* in form of a circle in the *top-view* image denote the handedness of each H-STF. Scale bars are 400 nm (Reprinted with permission from [40]. Copyright 2010, Schmidt)

Each H-STF was subdivided into 21 homogeneous anisotropic layers with a piece-wise $\delta\varphi$ rotated orthorhombic axes system (**a**, **b**, **c**; $\alpha = \beta = \gamma = 90°$; Fig. 10.9). This level of discretization was sufficient to reach close match between model and experiment. However, in order to achieve best-match model results, it was not pos-sible to find a homogeneous φ rotation along z (substrate normal). A stepwise z profile for $\varphi(z)$ was introduced with 8 nodes, and each node was divided into 3 slices. The nodes were fixed and equally spaced along z (spacing depends on the overall film thickness), and the rotation $\delta\varphi$ between two nodes was an additional free parameter during best-match model calculations (Fig. 10.28). Sample 1H$^+$-STF reveals fairly homogeneous rotation with a small disturbance around 225°. Sample 1H$^-$-STF reveals a similar profile with a more pronounced disturbance around the same φ rotation. These deviations from a homogeneous z-profile might be due to a sample wobble during substrate rotation.

Experimental and best-match model calculated GE Mueller matrix data for four different angles of incidence obtained from the Co H$^+$- and H$^-$-STF are shown in Fig. 10.29a and b, respectively. The graph depicts selected Mueller matrix elements versus sample azimuth at an exemplary wavelength $\lambda = 630$ nm. No pseudo-isotropic sample orientations can be identified.

Similar to the Ti chiral 2F2-STFs with opposite handedness, in case of ideal structures, off-diagonal elements M_{13}, M_{14}, M_{23}, and M_{24} from the H$^+$-STF can be transferred into the same elements of the H$^-$-STF by inversion around (180°, 0) whereas a mirror operation at $\phi = 180°$ is necessary for elements M_{12}, M_{22}, M_{33}, and M_{34}. Another characteristics for these chiral STFs is that there is no symmetry
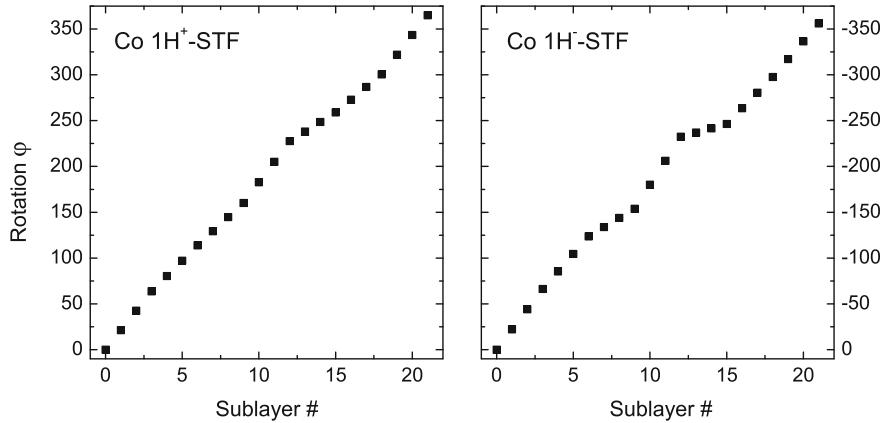
Fig. 10.28 Non-homogeneous piecewise layer model for rotation of the principal dielectric function along z (Reprinted with permission from [40]. Copyright 2010, Schmidt)
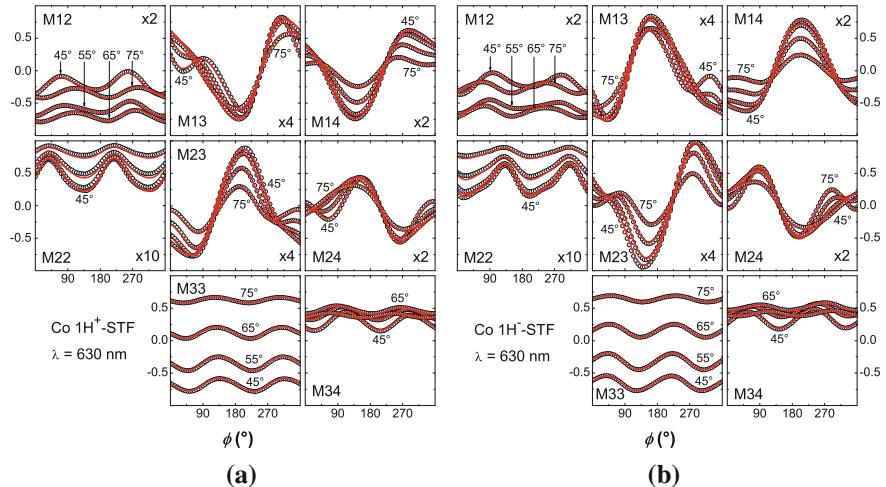
Fig. 10.29 Exemplary experimental (*circles*) and best-match calculated (*solid lines*) GE data of a Co H$^+$-STF (**a**) and a Co H$^-$-STF (**b**) versus sample azimuth angle ϕ at $\lambda = 630$ nm. The four graphs for each Mueller matrix element are four different angles of incidence $\Phi_a = 45°, 55°, 65°, 75°$. Note that for both chiral H-STFs no pseudo-isotropic orientation is present. Numbers in *upper* or *lower right* corner of elements denote the factor with which data are magnified (Reprinted with permission from [40]. Copyright 2010, Schmidt)

within a single Mueller matrix elements as compared to other achiral STFs discussed above.

Optical constants obtained from both nanostructured thin films compare well and are depicted in Fig. 10.30. Note that the order of k_a and k_b is exchanged between both chiral H-STFs.
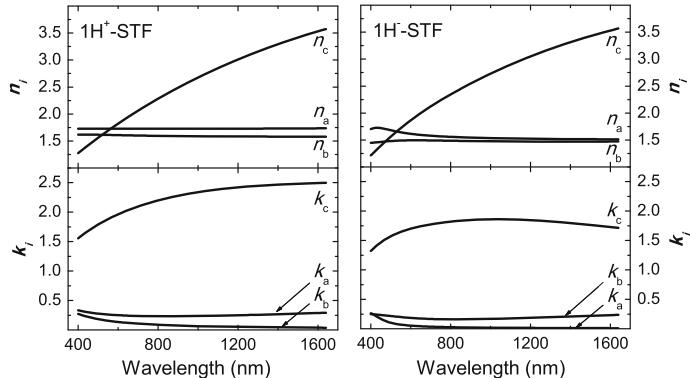
Fig. 10.30 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with orthorhombic arrangement of (*left panel*) a Co H$^+$-STF and (*right panel*) a Co H$^-$-STF. Optical constants depicted here are in good agreement with optical constants of Co F1-STFs and Co 3F4-STFs shown in Fig. 10.14 and Fig. 10.26, respectively (Reprinted with permission from [40]. Copyright 2010, Schmidt)

Table 10.6 Best-match model results for the Co H$^+$- and H$^-$-STF

Parameters	H$^+$-STF	H$^-$-STF
d (nm)	70.7(2)	76.2(2)
δd (nm)	3.37	3.63
φ_{tot} ($^\circ$)	365(3)	356(3)
θ ($^\circ$)	55.1(1)	59.6(1)
ψ ($^\circ$)	25.4(9)	$-15.2(5)$
$\varnothing_{\text{h}}$ (nm)	48.6	45.2
MSE	8.457	6.353

The error limits given in parentheses denote the uncertainty of the last digit (90 % reliability)

Other optically determined best-match parameters for both H$^+$- and H$^-$-STF are summarized in Table 10.6. Note that here the major polarizabilities within an orthorhombic arrangement are rotated with all three Euler angles and hence no principal axis is parallel to the substrate interface anymore. The inclination angle θ is representative of the tilt of the **c**-axis, which determines the slope of the helical "windings". The **a**- and **b**-axes orientations with respect to the substrate surface normal can be understood as effective polarization radii (due to the coordinate system rotation) of the chiral nanostructures. One may interpret these radii as effective coupling distances within which the individual nanostructures couple their dielectric polarization response.

Helix diameter estimation. Together with the total film thickness d, the inclination angle θ, and the overall in-plane rotation φ_{tot} the pitch $P = \frac{2\pi}{\varphi_{\text{tot}}} \cdot d$ and consequently the diameter $\varnothing_{\text{h}}$ of a single helix can be computed as $\varnothing_{\text{h}} = P / \tan\theta$. This calculated diameter is not accounting for any "wire" thickness and has to be compared with an average between inner ($\varnothing_{\text{in}}$) and outer ($\varnothing_{\text{out}}$) diameter of the true helix determined from SEM image analysis [$\varnothing_{\text{h}} \stackrel{\wedge}{=} \varnothing_{\text{avg}} = 0.5 \cdot (\varnothing_{\text{in}} + \varnothing_{\text{out}})$]. The calculated

diameters for both H-STFs are with approximately 45 nm in good agreement with a SEM image estimates of the averaged diameter of $\varnothing_{\mathrm{avg}} = 50 \pm 5\,\mathrm{nm}$.

10.6.4 Effective Medium Approximations for F1-STFs

Bulk material optical constants can be modified using effective medium approximations (EMAs) to attain approximate effective optical constant of a thin film comprising more than one constituent. The advantage of using EMAs for ellipsometric data analysis is the direct access to a material fraction parameter related to the amount of the mixed constituents. The anisotropic Bruggeman EMA (AB-EMA) as described in Sect. 10.3.6.2 has been applied here to estimate the void fraction in F1-STFs by mixing bulk optical constants (determined during analysis) with void. Note that no satisfactory match between experimental and model generated was achieved by using available literature values for bulk dielectric functions. Therefore, bulk-like optical constants for the respective F1-STF material have been determined by wavelength-by-wavelength and parameterized MDF analysis. The effective anisotropic optical constants obtained by MDFs are compared to results from the homogeneous biaxial layer approach.

Model description. The optical model for F1-STFs comprises a single biaxial (monoclinic) layer. Optical constants of the respective bulk material are determined with simple parameterized MDFs (Lorentzian oscillators), mixed with fractions of void f_{v} (optical constants $n_{\mathrm{v}} = 1$, $k_{\mathrm{v}} = 0$), and weighted with depolarization factors (L_j^{D}, $j = a, b, c$) for the three biaxial effective dielectric functions. L_j^{D} is proportional to the polarizability of the jth axis of the "inclusion" and hence ratios define the shape of the aligned particle. Depolarization factors determine the difference between n_j and k_j along axes **a**, **b**, **c**. Euler angles φ, θ, ψ transform the Cartesian laboratory coordinate frame into the material coordinate frame and the internal (monoclinic angle) β converts the virtual orthogonal basis into a monoclinic system. The total film thickness d completes the best-match model parameter list.

10.6.4.1 Cobalt F1-STF

The Co bulk-like optical constants determined with the AB-EMA model approach are compared to values found in Palik [106] and are depicted in Fig. 10.14. The extinction coefficients are comparable within the visible spectral range only and differ significantly in the near-infrared wavelength range. The refractive index as a result of the AB-EMA model is considerably larger than bulk literature values but exhibits a similar dispersion. The difference might be due to the idealized optical model as well as nanostructure confinement effects and the inherent large surface to volume ratio of the F1-STF used here (approximately $200\,\mathrm{m}^{-1}$), which may affect the overall polarizability of the material. Besides that, a readily grown native metal oxide
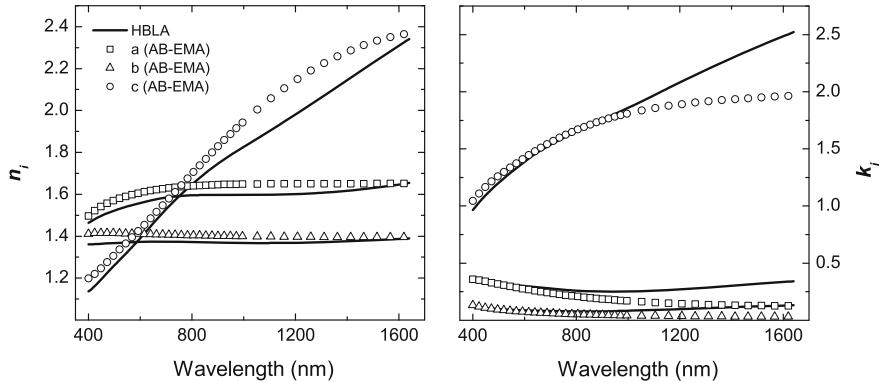
Fig. 10.31 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** of a Co F1-STF determined by the homogeneous biaxial layer approach (HBLA *solid line*) compared to optical constants determined by the AB-EMA (*symbols*)

and airborne contaminations which are not taken into account but may also affect the film birefringence and hence the AB-EMA bulk Co optical constant best-match results. Resulting effective optical constants are compared to data obtained with the homogeneous biaxial layer approach (Sect. 10.6.1.1)[23] and plotted in Fig 10.31. The general trend of optical constants determined with both approaches is in very good agreement, especially within the visible spectral region. The MSE is 11.9 and slightly higher than for the homogeneous biaxial layer approach (MSE = 7.8, see Table 10.9). Also structural parameters such as thickness $d = 101.4 \pm 0.1$ nm and columnar slanting angle $\theta = 62.9 \pm 0.02°$ are in very good agreement with values determined with the homogeneous biaxial layer approach and the best-match value for the monoclinic angle is $\beta = 80.8°$. The void fraction $f_v = 75.38 \pm 0.01\%$ reflects the high porosity of the film and which is characteristic for GLAD at very oblique angles [108]. The depolarization factors ($L_a^D = 0.378$, $L_b^D = 0.504$, and $L_c^D = 0.118$) show that the structural unit is extended in the **c**-direction since L_c^D is considerably smaller than the other two parameters, and the fact that $L_a^D \neq L_b^D$ indicates that the film is rendered with biaxial properties [109].

Based on existing literature [109, 110] and investigations presented in this contribution an empirically found order of depolarization factors for absorbing metal F1-STFs is $L_a^D > L_b^D > L_c^D \neq 0$. Because the columns are not infinitely long along the **c**-axis the depolarization factor L_c^D should not be assumed to be equal to zero as suggested previously [108]. These depolarization factors then are representative for a prolate ellipsoid extended along the **c**-axis, since L_c^D is smaller than the other two parameters, and $L_a^D \neq L_b^D$ shows that the film has biaxial properties. $L_a^D > L_b^D$ is also in agreement with the observed elliptical shape of the columns with a longer axis perpendicular to the incoming vapor flux due to anisotropic shadowing effects

---

[23] Here, data is compared that has been taken immediately after deposition and consequently no oxide layer was included within the AB-EMA.
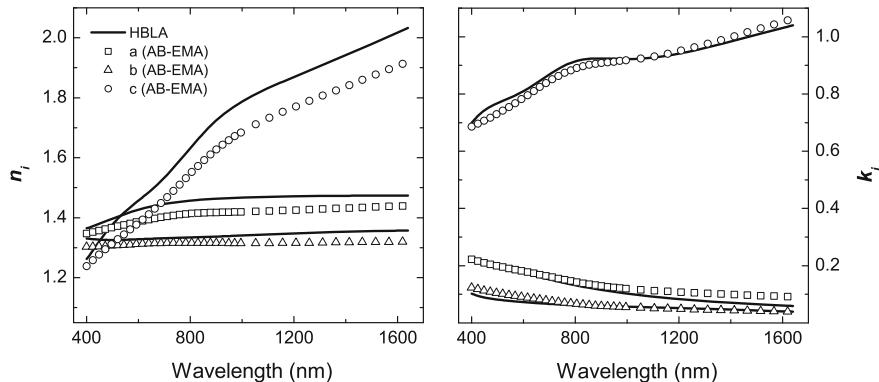
Fig. 10.32 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** of a Ti F1-STF determined by the homogeneous biaxial layer approach (HBLA *solid line*) compared to optical constants determined by the AB-EMA (*symbols*)

during oblique angle deposition [54, 86]. In general, structural parameters and void fraction attained with the AB-EMA are in very good agreement with SEM analysis and may serve as a good estimate. The additionally gained information about the film porosity is very valuable and is key for the importance of this approach.

10.6.4.2 Titanium F1-STF

The Ti bulk-like optical constants determined with the AB-EMA model approach are compared to values published in Palik [106] and found to be in very good agreement (Fig. 10.16). Even the shoulder around 800 nm caused by interband transitions is present. Figure 10.32 shows the effective optical constants determined with the homogeneous biaxial layer approach and with the AB-EMA for Ti F1-STFs. Parameters for the Ti F1-STF obtained with both approaches as well as SEM image analysis are summarized in Table 10.7. Structural values and the monoclinic angle determined with both approaches are in very good agreement and especially the slanting angle matches SEM estimates very well. In general, depolarization factors and the void fraction compare well to values determined for the Co F1-STF above.

10.6.4.3 Supermalloy F1-STF

Supermalloy bulk optical constants for the AB-EMA layer have been generated as described in Sect. 10.6.2.2 and are depicted in Fig. 10.17. The comparison to bulk-like values determined by the AB-EMA model analysis reveals similar discrepancies as found for Co F1-STFs. Especially the refractive index n is larger compared to the literature values. Figure 10.33 shows the effective optical constants determined with

Table 10.7 Best-match model results from homogeneous biaxial layer approach (HBLA), AB-EMA, and SEM analysis for the Ti F1-STF

Parameters	HBLA	AB-EMA	SEM
d (nm)	100.2(2)	103.5(2)	112(4)
θ (°)	57.27(6)	56.95(3)	58(4)
β (°)	80.2(1)	81.3(1)	–
f_v (%)	–	78.30(4)	–
L_a^{D}	–	0.394(1)	–
L_b^{D}	–	0.494(1)	–
L_c^{D}	–	0.112(1)	–
MSE	6.23	10.86	–

The error limits given in parentheses denote the uncertainty of the last digit (90 % reliability)
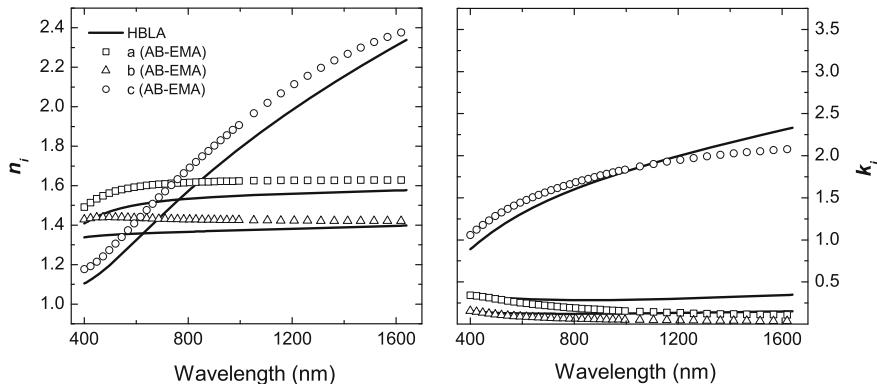
Fig. 10.33 Refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** of a NiFeMo F1-STF determined by the homogeneous biaxial layer approach (HBLA *solid line*) compared to optical constants determined by the AB-EMA (*symbols*)

the homogeneous biaxial layer approach and with the AB-EMA for supermalloy F1-STFs. Obtained parameters with both approaches as well as SEM image analysis are summarized in Table 10.8. The monoclinic angle and the void fraction is comparable to values obtained for the same F1 geometry from Co and Ti, respectively. Interestingly, best-match depolarization factors for supermalloy are almost identical to the ones from a Ti F1-STF, which is not apparent from top view SEM micrographs (Fig. 10.11).

In summary, optically determined structural properties obtained with the AB-EMA model approach are in very good agreement with parameters calculated by the homogeneous biaxial layer approach and hence may serve as excellent estimates. The key advantage of the AB-EMA is a direct access to the material fractions. Furthermore, the number of model-parameters is reduced with respect to the homogeneous biaxial layer approach since the effective optical constants along the major polarizability axes are not determined individually. The empirically found order of depolarization factors for metal F1-STFs is $L_a^{\mathrm{D}} > L_b^{\mathrm{D}} > L_c^{\mathrm{D}} \neq 0$ and L_c^{D} should not

Table 10.8 Best-match model results from homogeneous biaxial layer approach (HBLA), AB-EMA, and SEM analysis for the NiFeMo F1-STF

Parameters	HBLA	AB-EMA	SEM
d (nm)	88.4(1)	80.4(1)	100(4)
θ (°)	63.16(3)	61.13(2)	64(4)
β (°)	89.53(8)	82.40(6)	–
f_v (%)	–	75.0(1)	–
L_a^D	–	0.393(1)	–
L_b^D	–	0.494(1)	–
L_c^D	–	0.112(1)	–
MSE	5.18	10.12	–

The error limits given in parentheses denote the uncertainty of the last digit (90 % reliability)

be zero because the columns have a finite length. Similarly to the modularity principle discussed with regard to the piecewise homogeneous biaxial layer approach, appropriately stacked "F1-STF AB-EMA model layers" accounting for the cascaded STF geometry may be used to forward calculate approximate optical responses of more complex STFs.

10.6.5 Aging Effects and Environmental Influences on Optical Properties

As soon as a metal STF leaves the ultra-high vacuum environment of the deposition chamber and is exposed to ambient conditions the large pristine metal surface area reacts with air. Typically a self-limiting native metal oxide layer starts to grow at room-temperature but the self-termination depends on the surface quality and small changes may still be observed after months [111]. Furthermore, water molecules due to the ambient humidity and airborne contaminations may adsorb on any exposed surface area. These uncontrolled processes add additional constituents to the initially pristine metal film, affect the nanostructure polarizability and thus the effective optical constants. Hence, care must be taken when evaluating the optical properties since the oxide layer may grow over a longer period of time and subtle changes in the ambient cause differences in the film birefringence. If the uncontrolled native oxidation can be prevented, however, such anisotropic metal STFs may be used as very sensitive sensing elements.

10.6.5.1 Effective Optical Constant Changes due to Aging

In Fig. 10.34 optical constants of the Co F1-STF (sample #1) determined by the homogeneous biaxial layer approach from measurements taken 1 h after deposition are plotted (same as Fig. 10.14) and compared to n_j and k_j resulting from

measurements acquired 90 d after deposition (for model and analysis details see Sect. 10.6.1.1). Between both measurements, the sample was stored in a closed yet not airtight container and thus exposed to ambient air. Therefore, differences between both results are attributed mainly to the growth of a metal oxide layer; however, also changes in ambient humidity, and airborne contaminations are possible. In general, directions **a** and **b** show a very similar dispersion. In both cases, $n_{a,b}$ and $k_{a,b}$ from data taken after 90 d exhibit a positive offset, fairly constant over the investigated spectral region with respect to data acquired 1 h after deposition. No common trend is observed along direction **c**. The MDF parameterization delivered excellent results for both measurements and best-match model parameters are summarized in Table 10.9. Structural parameters are in very good agreement with SEM micrograph analysis ($d = 114 \pm 4$ nm and $\theta = 65 \pm 3°$).

10.6.5.2 STF Passivation by Atomic Layer Deposition

In order to prevent uncontrolled native oxidation of metal layers a thin stable oxide capping layer can be deposited. For coating the three-dimensionally shaped STFs a deposition technique is desired which allows for a thin homogeneous passivation shell. It has been shown that atomic layer deposition (ALD) is an excellent technique to conformally coat complex nanostructures and STFs [112–114]. To passivate an Co F1-STF, the thin film was transferred immediately after the growth process to the ALD reactor (Fiji 200, CambridgeNanoTech Inc.). A subsequent thermal Al_2O_3 process [115] of 60 cycles was carried out at a substrate temperature of 60 °C in order to avoid sample property alterations by annealing effects. Both precursors, trimethylaluminum and nanopure water, were pulsed at 60 ms each followed by a 30 s purge time, which warrants a homogeneous and conformal coating of the columns. Note that even though ALD is a well-known technique to achieve highly conformal coatings of high aspect ratios the uniformity is affected by non-ideal sample areas where columns are partially in contact. A schematic drawing of the final core-shell nanocolumnar thin film as well as cross-section scanning electron microscope (SEM) micrographs before and after ALD are depicted in Fig. 10.35. In order to determine the Al_2O_3 layer thickness a 100 nm thick solid Co reference sample has been coated under identical ALD conditions. The resulting growth per cycle for the first 60 cycles determined with ellipsometry was 0.548 Å.

Two measurements of the same sample (Co F1-STF passivated with a thin Al_2O_3 coating) are compared, which have been conducted both immediately after deposition and again 60 d later. The stratified optical model is composed of three layers accounting for Si substrate, native SiO_2, and a AB-EMA layer for the F1-STF. The best-match model parameters within the AB-EMA layer for all for measurement scenarios are summarized in Table 10.10 and respective wavelength-by-wavelength determined optical constants depicted in Fig. 10.36.

The AB-EMA layer of the Al_2O_3 passivated Co SCTF comprises bulk-like optical constants of three constituents: Co (model parameters), Al_2O_3, and void ($n = 1, k = 0$). Notably, except for material fractions, all best-match model parame-

Table 10.9 Best-match model results for Co F1-STF 1 h and 90 d after deposition, respectively

Structural parameters		MDF parameters along direction		
		a	b	c
1 h after deposition (MSE = 7.77)				
$d = 106.9(1)$ nm	A_1 (eV)	1.00(1)	0.348(4)	0.3(3)
$\theta = 63.68(1)°$	E_{c1} (eV)	0.633(4)	0.685(6)	1.43(6)
$\beta = 83.69(9)°$	γ_1 (eV)	0.90(3)	1.18(5)	0.6(3)
	A_2 (eV)	0.08(1)	0.12(2)	0.7(6)
	E_{c2} (eV)	1.96(1)	2.56(9)	1.83(4)
	γ_2 (eV)	1.1(1)	2.4(2)	1.2(5)
	A_3 (eV)	0.88(1)	0.416(8)	0.8(7)
	E_{c3} (eV)	5.2(1)	4.18(9)	2.7(1)
	γ_3 (eV)	8.2(5)	2.2(3)	2(1)
	A_4 (eV)	–	–	1.5(5)
	E_{c4} (eV)	–	–	5(2)
	γ_4 (eV)	–	–	3(5)
	ρ (Ωcm)	–	–	$7.58(4) \times 10^{-4}$
	τ (fs)	–	–	0.325(3)
	ε_{off}	1.41(1)	1.62(1)	1.2(4)
90 d after deposition (MSE = 6.77)				
$d = 106.5(1)$ nm	A_1 (eV)	1.13(1)	0.435(4)	0.5(5)
$\theta = 62.95(3)°$	E_{c1} (eV)	0.672(3)	0.700(5)	1.42(9)
$\beta = 80.92(9)°$	γ_1 (eV)	0.89(2)	1.16(4)	0.9(3)
	A_2 (eV)	0.13(2)	0.25(2)	0.8(8)
	E_{c2} (eV)	2.02(1)	2.89(8)	1.85(5)
	γ_2 (eV)	1.4(1)	2.7(1)	1.3(6)
	A_3 (eV)	1.05(1)	0.47(2)	1.0(7)
	E_{c3} (eV)	5.2(1)	3.90(9)	2.8(1)
	γ_3 (eV)	8.2(5)	1.0(2)	2.3(9)
	A_4 (eV)	–	–	1.5(4)
	E_{c4} (eV)	–	–	5.6(1.5)
	γ_4 (eV)	–	–	3.2(4.6)
	ρ (Ωcm)	–	–	$7.11(4) \times 10^{-4}$
	τ (fs)	–	–	0.459(4)
	ε_{off}	1.50(2)	1.97(1)	1.5(3)

Parameters A_j, E_{cj}, γ_j correspond to amplitude, center energy, and broadening of the jth Lorentzian-type oscillator, respectively, whereas ρ, τ represent the resistivity and scattering time of a Drude term, respectively. The error limits given in parentheses denote the uncertainty of the last digit (90 % reliability)

ters are virtually identical between measurements taken on the day of deposition and 60 d thereafter, respectively. This shows that the film properties of passivated STFs are not affected by storage in ambient air and expected aging effects due to material oxidation not present due to the conformal Al_2O_3 capping layer. The determined layer thickness t and slanting angle θ are in very good agreement with cross-section SEM analysis (Fig. 10.35). The observed imbalance between f_{void} and $f_{Al_2O_3}$ as well
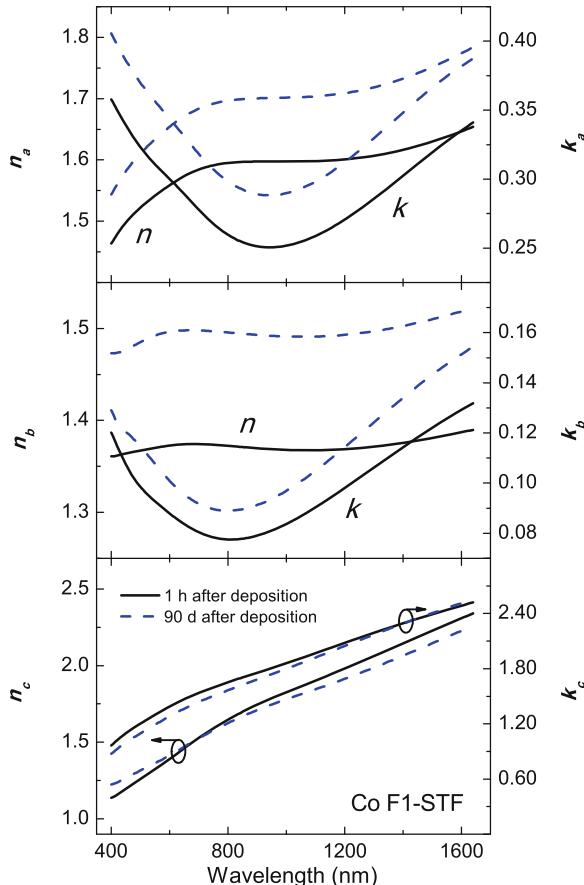
Fig. 10.34 Comparison between refractive indices n_j and extinction coefficients k_j along major polarizability axes **a**, **b**, **c** with monoclinic arrangement of a Co F1-STF determined from measurements taken 1 h (*solid line*; same as Fig. 10.14) and 90 d (*dashed line*) after deposition, respectively (Reprinted with permission from [40]. Copyright 2010, Schmidt)

as the slight changes in best-match Co bulk optical constants are mainly attributed to differences in ambient humidity at the two measurement days. This is supported by the fact that there is virtually no effective optical constant variation along **c**. It is well-known that such nanostructured thin films are very sensitive high-speed capacitive humidity sensors [116, 117]. The wavelength-by-wavelength determined best-match Co bulk optical constants obtained from model analysis for the passivated Co SCTF after deposition and after 60 d are in excellent agreement with each other; however, differ from literature values found in Palik [106] (see also Sect. 10.6.4). The effective biaxial optical constants only differ for n_a and n_b, which suggests no material change due to oxidation (n_c unaffected) but rather changes in the void regions cause by possible humidity changes (Fig. 10.36). Investigations of water
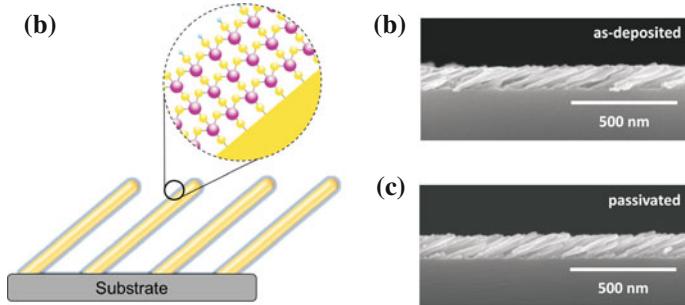
Fig. 10.35 Schematic drawing of a core-shell Co F1-STF coated with an Al_2O_3 passivation layer (3 ALD cycles depicted, **a**). High-resolution cross-section SEM micrographs reveal the structural equivalence before (**b**) and after (**c**) the ALD passivation process (Reprinted with permission from [90]. Copyright 2012, American Institute of Physics)
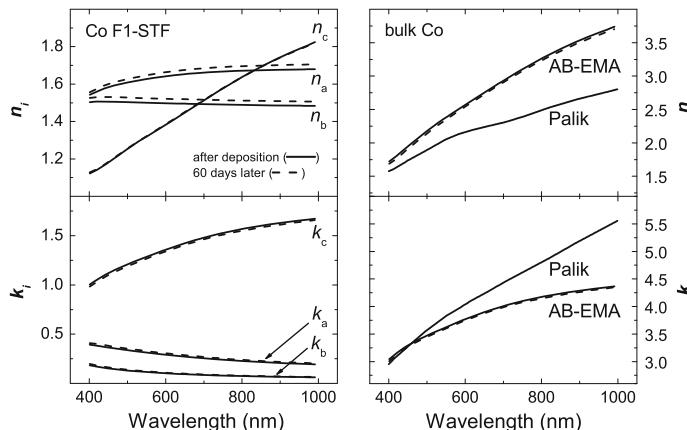
Fig. 10.36 Wavelength-by-wavelength determined optical constants for the Al_2O_3 passivated Co F1-STF measured immediately after deposition (*solid lines*) and 60 days later (*dashed lines*), respectively. The *right panel* shows the bulk-like optical constants n and k, which are input parameters for the AB-EMA approach and also compared to literature values from Palik [106]. The *left panel* depicts the effective optical constants (n_{eff} and k_{eff}) along major axes of polarizability **a**, **b**, **c**, which are due to the different depolarization factors

vapor adsorption in a controlled environment show that even small differences of only 5 % in relative humidity can be easily measured as a change in the film birefringence.

The calculated Al_2O_3 shell layer thickness based on best-match AB-EMA constituent fraction parameters is with 2.4 nm in excellent agreement with the reference sample layer thickness of 3.29 nm. Note that for this calculation ideal and separated nanocolumns with an average diameter of 17.5 nm are assumed. High resolution SEM top-view and cross-section image analysis show that the diameter of individual columns is increased by approximately 4 nm after ALD Al_2O_3 coating [90, 118].

Table 10.10 Overview of the best-match model parameters for both measurement scenarios, after deposition (0 d) and 60 days later (60 d), respectively for the passivated Co SCTF

Parameter	Co+Al$_2$O$_3$ (after deposition)	Co+Al$_2$O$_3$ (60 d later)
t (nm)	89.46(3)	89.86(3)
θ (°)	62.69(1)	62.80(1)
β (°)	82.96(2)	83.39(2)
f_{void} (%)	62.03(7)	59.33(6)
$f_{\text{Al}_2\text{O}_3}$ (%)	13.98(8)	16.74(6)
L_a^{D}	0.4035(2)	0.4030(2)
L_b^{D}	0.5267(2)	0.5283(2)
L_c^{D}	0.0698(7)	0.0688(6)
MSE	8.45	10.82

t = thickness; θ = tilting angle with respect to the substrate normal; β = monoclinic angle between axes c and a; f_x = volume fraction of x; L_j^{D} = depolarization factor along j. The error limits given in parentheses denote the uncertainty of the last digit (90 % reliability)

10.6.6 Discussion

10.6.6.1 Early Optical Investigations on Sculptured Thin Films

Kundt [31] reported on birefringence in metal thin films deposited at oblique angles already in 1886 and concluded that the specific microstructure may be the origin, while electron microscopy or similar techniques were unavailable. Smith, Cohen, and Weiss [119] determined with polarized transmission measurements that the absorption coefficient in obliquely deposited metal films is a periodic function of the sample azimuth and that the differential absorption (parallel versus perpendicular to the slanting plane of the columns) increases with increasing deposition angle θ_i. King and Talim [120] considered the effect of columnar microstructure on the optical anisotropy of thin films deposited onto substrates at normal incidence with the help of ellipsometry and other techniques and postulated an uniaxial model. Inspired by this work, Hodgkinson et al. [121] performed polarization dependent reflection measurements on obliquely deposited transparent ZrO$_2$ and TiO$_2$ F1-STFs. They proposed an orthorhombic biaxial model due to the existence of form birefringence; i.e., the obliquely deposited columns do not exhibit a perfectly round shape but are rather elliptically in shape. The authors realized that there is no mode coupling between p and s polarization for light incident in the plane containing the direction of deposition and the substrate normal. These orientations are equivalent to the pseudo-isotropic orientations discussed for example in Sect. 10.6.1 (Fig. 10.13).

10.6.6.2 Effective Medium Approximations for F1-STFs

In order to quantify birefringence and porosity values of F1-STFs in the visible and near-infrared spectral region from transparent oxides (metal oxides and SiO$_2$),

initially empirical equations [122, 123] have been reported and then existing EMAs were applied (Bragg-Pippard [86, 124], Maxwell-Garnett [101, 125], and AB-EMA [126–130]).

Hodgkinson and Wu [86] reported, based on optical constants determination with EMAs, that in dielectric biaxial F1-STF the optical constants generally follow the empirically found order $n_c > n_a > n_b$. The same order is found here for all investigated STFs in the near infrared spectral region. However, presented data in this chapter over an extended spectral region reveal the order of refractive indices for metal STFs is different in the visible spectral region since n_c is intersecting with n_a and n_b.

Depolarization factors deliver information about the shape of the inclusions. However, it is controversially discussed, which value of L_c^{D} should be used for F1-STF. Often, it is claimed that the depolarization factor along the long axis of ellipsoids (**c**-axis) should be zero because of minimum charge screening effects along this direction and many authors have therefore assumed $L_c^{\mathrm{D}} = 0$ for their best-match model calculations in order to determine optical constants and porosity values of F1-STFs [86, 96, 108, 124–127, 131, 132]. Mbise et al., however, reported on analysis of polarized transmittance measurements using an orthorhombic AB-EMA to quantify optical anisotropy of Cr F1-STFs. The authors determined depolarization factor values $0.14 < L_c^{\mathrm{D}} < 0.45$ and found that optically determined structural properties such as film thickness and column inclination are in fair agreement with SEM investigations [109, 110]. Gospodyn and Sit also determined L_c^{D} based on model calculations using the Bruggeman EMA for a variety of MgF_2 and SiO_2 F1-STFs grown at different deposition angles. They report values of L_c^{D} between 0.01 and 0.1 for their lossless nanostructured films [128]. Recently, Wakefield et al. published an extensive experimental characterizations and a model for metal oxide F1-STFs. The authors provide a recipe to calculate the in-plane birefringence and other film parameters of slanted columnar thin films deposited with different variable azimuthal substrate rotation techniques. Within their approach they fix the depolarization factor $L_c^{\mathrm{D}} = 0.1$.

Hofmann et al. have investigated 500 nm thick Co F1-STFs in the THz spectral range and found that an orthorhombic AB-EMA approach leads to a very good approximation of structural and physical properties. While the structural parameters (thickness, slanting angle, and void fraction) are in good agreement with SEM data and generalized ellipsometry in the visible spectral region, depolarization factors differ significantly from values determined by model analysis of experimental data obtained within the visible spectral range [133]. These results suggest that depolarization factors might be wavelength-dependent and a Bruggeman approach with a different (energy-dependent) depolarization factor definition could possibly be applied here [134, 135]. Furthermore, the authors have shown that also at THz frequencies the optical response strongly depends on the dielectric properties of the surrounding ambient medium. These findings indicate that with sensors comprising STFs minute amounts of dielectrics can be detected also with millimeter long wavelengths [8]. The interested reader is also referred to Chap. 11 for a detailed description of the aforementioned investigations within the THz spectral range.

10.6.6.3 Optical Properties of Complex and Hybrid STFs

Chevron-like nanostructures (*l* F2-STFs). Motohiro and Yaga [136] rotated their sample manually by 180° to grow two successive F1-STF layers with opposite slanting direction and experimentally determined that a non-absorbing metal oxide 2F2-STF can be used as a quarter-wave plate and can compete with conventional types of retardation plates. The authors realized that 2F2-STFs show superior retardation properties with respect to F1-STFs.

Podraza et al. [101] matched Jones matrix data of a non-absorbing MgF$_2$ 2F2-STF with a similar approach than the one used in this chapter and found good agreement between optically determined structural properties and their SEM analysis. However, they assumed uniaxial properties in the transparent region of their MgF$_2$ films and included measured data from only one in-plane orientation into their best-match model calculations. For determination of optical properties and thin film birefringence the authors used the Maxwell-Garnett EMA [96, 137] with depolarization factors $L_a^D = L_b^D = 1$ and $L_c^D = 0$ perpendicular and parallel to the long axis of the nanostructure, respectively, to parameterize both principal dielectric functions.

Four-fold staircase nanostructures (*l* F4-STFs). The geometry of a 3F4-STF can be seen as the three-dimensional equivalent of a two-dimensional split ring resonator [138–140]. Such three-dimensional metamaterials from metal have gained research interest because effective negative index and magnetic resonances have been proposed [141, 142]. Besides that, dielectric (4F4)$_x$-STFs with $x = 4, 5$ are found to act as three-dimensional photonic bandgap crystals with wide bandgaps [143]. However, no reports on intrinsic optical properties of F4-STFs have been found.

Helical nanostructures (*t* H-STFs). In search of new materials to miniaturize existing polarization rotators (Reusch rotator) and potentially create thin film Šolc color filters [144], Young and Kowal [35] were the first ones to report on *in-situ* substrate rotation during oblique angle evaporation thereby creating CaF$_2$ H-STFs. However, even though optical activity and large polarization rotation was experimentally confirmed their paper from 1959 went largely unnoticed. Recently efforts have been made mostly in the theoretical description of light propagation in H-STFs lead by Lakhtakia [33, 145, 146]. Experimental reports about chiral H-STFs are dealing with selective transmission of left- and right- circularly polarized light and optical rotary power. H-STFs, which can be physically considered as "frozen" cholesteric liquid crystals [147] are found to be good circular polarizers since light with the same handedness as the helix is blocked, whereas the other handedness is transmitted (within a certain frequency range) [148–153].

However, except for Zhong et al. [154], who modeled the optical properties of helical ITO thin films with a Cauchy dispersion model, all other reports mentioned above do not report on intrinsic optical properties of the investigated H-STFs.

Hybrid STFs. Photonic characteristics, for example, may be modified and tuned by infiltrating a dielectric material in void [124, 155, 156]. Active control over optical properties of hybrid materials can be achieved by either combining nanoparticles with polymers, which change their properties upon exposure to gases [157] or inorganic porous layers with temperature sensitive liquid crystals [158], for example. There-

fore, optical constants of hybrid nanoporous thin films are effective optical constants, which depend not only on the geometry but also on the dielectric properties of the material infiltrated into void [26].

10.7 Summary and Outlook

Glancing angle deposition was presented as an advanced physical vapor deposition technique capable of bottom-up fabrication of three-dimensionally shaped, highly spatially coherent nanostructures. Paired with a dynamic substrate motion, this technique allows for engineering self-assembled, self-organized highly spatially coherent achiral and chiral *sculptured thin films*. A new nomenclature scheme has been proposed based on basic building blocks to unambiguously identify sculptured thin film geometries by their name.

Spectroscopic generalized ellipsometry in the visible and near-infrared spectral region was demonstrated to be a highly suitable, non-destructive tool for investigation of low-symmetry and highly anisotropic absorbing sculptured thin films. The intrinsic polarizabilities inherent to the biaxial nanostructured materials examined here have been determined for a series of different metal sculptured thin film geometries and effective principal optical constants are reported. Strong optical birefringence and dichroism are quantified for different sculptured thin film samples, and the complex-valued dielectric function tensor differs significantly from the respective bulk material. In particular, achiral metal slanted columnar thin films (F1-STFs) are found *monoclinic* due to dielectric polarization charge coupling effects across neighboring slanted but electrically isolated nanocolumns. The validity of the homogeneous biaxial layer approach applied to model the anisotropic electromagnetic plane wave response of metal F1-STFs was discussed for F1-STFs from three different materials: cobalt, titanium and supermalloy ($Ni_{80}Fe_{15}Mo_5$). Physical properties such as birefringence, dichroism, and monoclinicity, for example, are found to be common amongst all F1-STFs discussed here (*universality*).

For the first time, accurate sets of optical constants for complex manifold and helical sculptured thin films are presented. It is found that complex sculptured thin films may be considered cascaded F1-STFs building blocks and can be optically approximated as a stratified medium comprising F1-STFs with different slanting orientations. Therefore, once the building blocks are characterized the anisotropic polarization response of complex layered sculptured thin films can be predicted by a modular conception. The piecewise homogeneous layer approach enables modular assembly of F1-STF building blocks thereby creating an optical model that mimics the true geometry of the sculptured thin film. The *modularity* was exemplarily discussed for a total of six different sculptured thin film geometries from cobalt and titanium.

It has been shown that structural properties including multiple constituent fractions of F1-STFs determined with the anisotropic Bruggemann effective medium approximation may serve as good estimates. Based on existing literature and investigations presented in this chapter an empirically found order of depolarization factors

is presented ($L_a^D > L_b^D > L_c^D \neq 0$), which is in accordance with the shape of the columns (ellipsoids).

With the findings presented in this contribution it is now possible to predict optical properties of sculptured thin films. This will allow for engineering desired anisotropic physical properties of three-dimensionally nanostructured thin films and pave the way for next generation micro- and nanosystems, especially with respect to the exploitation of nanohybrid functional materials for novel detection principles.

References

1. O. Toader, S. John, Science **292**, 1133 (2001)
2. J.P. Singh, D.-L. Li, D.-X. Ye, R.C. Picu, T.-M. Lu, G.-C. Wang, Appl. Phys. Lett. **84**, 3657 (2004)
3. M.T. Umlor, Appl. Phys. Lett. **87**, 082505 (2005)
4. D.J. Bell, J. Dong, Y. Sun, L. Zhang, B.J. Nelson, D. Grützmacher, in *Proceedings of the 5th IEEE Conference on Nanotechnology* (Nagoya, Japan, IEEE, 2005), p. 15
5. S.V. Kesapragada, P. Victor, O. Nalamasu, D. Gall, Nano Lett. **6**, 854 (2006)
6. J.J. Steele, A.C. van Popta, M.M. Hawkeye, J.C. Sit, M.J. Brett, Sens. Actuat. B-Chem. **120**, 213 (2006)
7. K.B. Rodenhausen, D. Schmidt, T. Kasputis, A.K. Pannier, E. Schubert, M. Schubert, Opt. Express **20**, 5419 (2012)
8. T. Hofmann, D. Schmidt, A. Boosalis, P. Kühne, C. Herzinger, J. Woollam, E. Schubert, M. Schubert, Mat. Res. Soc. Symp. Proc. **1409**, CC13-31 (2012)
9. J.A. Dobrowolski, D. Poitras, P. Ma, H. Vakil, M. Acree, Appl. Opt. **41**, 3075 (2002)
10. D. Poitras, J.A. Dobrowolski, Appl. Opt. **43**, 1286 (2004)
11. J.-Q. Xi, M. Ojha, J.L. Plawsky, W.N. Gill, J.K. Kim, E.F. Schubert, Appl. Phys. Lett. **87**, 031111 (2005)
12. J.-Q. Xi, M. Ojha, W. Cho, J.L. Plawsky, W.N. Gill, T. Gessmann, E.F. Schubert, Opt. Lett. **30**, 1518 (2005)
13. S.-T. Ho, S.L. McCall, R.E. Slusher, L.N. Pfeiffer, K.W. West, A.F.J. Levi, G.E. Blonder, J.L. Jewell, Appl. Phys. Lett. **57**, 1387 (1990)
14. R. Sharma, E.D. Haberer, C. Meier, E.L. Hu, S. Nakamura, Appl. Phys. Lett. **87**, 051107 (2005)
15. Q. Xu, V.R. Almeida, R.R. Panepucci, M. Lipson, Opt. Lett. **29**, 1626 (2004)
16. J.K. Kim, T. Gessmann, E.F. Schubert, J.-Q. Xi, H. Luo, J. Cho, C. Sone, Y. Park, Appl. Phys. Lett. **88**, 013501 (2006)
17. A. Jain, S. Rogojevic, S. Ponoth, N. Agarwal, I. Matthew, W.N. Gill, P. Persans, M. Tomozawa, J.L. Plawsky, E. Simonyi, Thin Solid Films **398–399**, 513 (2001)

18. G.K. Kiema, M.J. Colgan, M.J. Brett, Sol. Energ. Mat. Sol. C. **85**, 321 (2005)
19. J. Bouclé, P. Ravirajan, J. Nelson, J. Mater. Chem. **17**, 3141 (2007)
20. H.-Y. Yang, M.-F. Lee, C.-H. Huang, Y.-S. Lo, Y.-J. Chen, M.-S. Wong, Thin Solid Films **518**, 1590 (2009)
21. M.D. Fleischauer, J. Li, M.J. Brett, J. Electrochem. Soc. **156**, A33 (2009)
22. F. Tang, D.-L. Liu, D.-X. Ye, Y.-P. Zhao, T.-M. Lu, G.-C. Wang, A. Vijayaraghavan, J. Appl. Phys. **93**, 4194 (2003)
23. F. Tang, D.-L. Liu, D.-X. Ye, T.-M. Lu, G.-C. Wang, J. Mag, Mag. Mat. **283**, 65 (2004)
24. D. Schmidt, T. Hofmann, C.M. Herzinger, E. Schubert, M. Schubert, Appl. Phys. Lett. **96**, 091906 (2010)
25. C.B.D. Schmidt, E. Schubert, M. Schubert, Mat. Res. Soc. Symp. Proc. **1408**, BB15 (2012)
26. D. Schmidt, C. Müller, T. Hofmann, O. Inganäs, H. Arwin, E. Schubert, M. Schubert, Thin Solid Films **519**, 2645 (2011)
27. E. Schubert, in *CAREER: Chiral Nanostructure Hybrid Materials for Applications in Terahertz Resonator and Magnetic Storage Devices*, ECCS-0846329, National Science Foundation Award (2009–2014)
28. F.F. Choobineh, RII Project: Nanohybrid Materials & Algal Biology, EPS-1004094, National Science Foundation Award (2010–2015)
29. http://en.wikipedia.org/wiki/File:Helical_capsid.jpg. Accessed 27 Jan 2012
30. T.D. Goddard, C.C. Huang, T.E. Ferrin, Structure **13**, 473 (2005)
31. A. Kundt, Ann. Phys. **263**, 59 (1886)
32. H. König, G. Helwig, Optik **6**, 111 (1950)
33. A. Lakhtakia, R. Messier, *Sculptured Thin Films: Nanoengineered Morphology and Optics* (SPIE Press, Bellingham, 2005)
34. M.M. Hawkeye, M.J. Brett, J. Vac. Sci. Technol. A **25**, 1317 (2007)
35. N.O. Young, J. Kowal, Nature **183**, 104 (1959)
36. B. Goldstein, L. Pensak, J. Appl. Phys. **30**, 155 (1959)
37. J.M. Nieuwenhuizen, H.B. Haanstra, Philips Tech. Rev. **27**, 87 (1966)
38. K. Robbie, L.J. Friedrich, S.K. Dew, T. Smy, M.J. Brett, J. Vac. Sci. Technol. A **13**, 1032 (1995)
39. K. Robbie, M.J. Brett, A. Lakhtakia, J. Vac. Sci. Technol. A **13**, 2991 (1995)
40. D. Schmidt, Generalized ellipsometry on sculptured thin films made by glancing angle deposition. Ph.D. thesis, University of Nebraska-Lincoln (2010)
41. K. Kaminska, T. Brown, G. Beydaghyan, K. Robbie, Appl. Opt. **42**, 4212 (2003)
42. B. Dick, M.J. Brett, T. Smy, J. Vac. Sci. Technol. B **21**, 23 (2003)
43. E. Schubert, F. Frost, H. Neumann, B. Rauschenbach, B. Fuhrmann, F. Heyroth, J. Rivory, E. Charron, B. Gallas, M. Schubert, Adv. Solid State Phys. **46**, 309 (2007)
44. F. Tang, T. Karabacak, L. Li, M. Pelliccione, G.-C. Wang, T.-M. Lu, J. Vac. Sci. Technol. A **25**, 160 (2007)
45. D. Vick, Y.Y. Tsui, M.J. Brett, R. Fedosejevs, Thin Solid Films **350**, 49 (1999)
46. K. Robbie, J.C. Sit, M.J. Brett, J. Vac. Sci. Technol. B **16**, 1115 (1998)
47. K. Robbie, G. Beydaghyan, T. Brown, C. Dean, J. Adams, C. Buzea, Rev. Sci. Instrum. **75**, 1089 (2004)
48. T. Karabacak, G.-C. Wang, T.-M. Lu, J. Appl. Phys. **94**, 7723 (2003)
49. T. Karabacak, G.-C. Wang, T.-M. Lu, J. Vac. Sci. Technol. A **22**, 1778 (2004)
50. M.O. Jensen, M.J. Brett, IEEE Trans. Nanotechnol. **4**, 269 (2005)
51. M. Malac, R.F. Egerton, M.J. Brett, B. Dick, J. Vac. Sci. Technol. B **17**, 2671 (1999)
52. B. Dick, M.J. Brett, T.J. Smy, M.R. Freeman, M. Malac, R.F. Egerton, J. Vac. Sci. Technol. A **18**, 1838 (2000)
53. E. Schubert, T. Höche, F. Frost, B. Rauschenbach, Appl. Phys. A-Mater. **81**, 481 (2005)
54. M.A. Summers, M.J. Brett, Nanotechnology **19**, 415203 (2008)
55. Y.P. Zhao, D.X. Ye, G.C. Wang, T.M. Lu, Nano Lett. **2**, 351 (2002)
56. R. Glass, M. Möller, J.P. Spatz, Nanotechnology **14**, 1153 (2003)

57. R. Glass, M. Arnold, E.A. Cavalcanti-Adam, J. Blümmel, C. Haferkemper, C. Dodd, J.P. Spatz, New J. Phys. **6**, 101 (2004)
58. C. Patzig, B. Rauschenbach, B. Fuhrmann, H.S. Leipner, J. Appl. Phys. **103**, 024313 (2008)
59. C. Patzig, B. Rauschenbach, B. Fuhrmann, H.S. Leipner, J. Nanosci. Nanotechnol. **9**, 1985 (2009)
60. L. Chi (ed.), Nanostructured Surfaces. Nanotechnology, vol. 8 (Wiley-VCH, Weinheim, 2010)
61. A. Dolatshahi-Pirouz, T. Jensen, T. Vorup-Jensen, R. Bech, J. Chevallier, F. Besenbacher, M. Foss, D.S. Sutherland, Adv. Eng. Mater. **12**, 899 (2010)
62. B.N. Mbenkum, On the anomalies in gold nanoparticles prepared by micelle nanolithography and their impact on one-dimensional material synthesis role of substrate, size effects and impurity. Ph.D. thesis, Ruperto-Carola University of Heidelberg (2007)
63. R.M.A. Azzam, N.M. Bashara, *Ellipsometry and Polarized Light* (North-Holland Publ. Co., Amsterdam, 1984)
64. D. Kliger, J. Lewis, C. Randall, *Polarized Light in Optics and Spectroscopy* (Academic Press, New York, 1990)
65. H.G. Tompkins, *A User's Guide to Ellipsometry* (Academic Press, New York, 1993)
66. H.G. Tompkins, E.A. Irene (eds.), *Handbook of Ellipsometry* (William Andrew Publishing, Norwich, 2005)
67. H. Fujiwara, *Spectroscopic Elliposmetry: Principles and Applications* (Wiley, Chichester, 2007)
68. M. Schubert, in *Infrared Ellipsometry on Semiconductor Layer Structures: Phonons, Plasmons, and Polaritons*. Springer Tracts in Modern Physics, vol. 209 (Springer, Berlin, 2004)
69. M. Schubert, *Handbook of Ellipsometry*, ed. by E.A. Irene, H.G. Tompkins. (William Andrew Publishing, Norwich, 2005)
70. P. Yeh, *Optical Waves in Layered Media* (Wiley, New York, 1988)
71. M. Schubert, Thin Solid Films **313–314**, 323 (1998)
72. A. Röseler, *Infrared Spectroscopic Ellipsometry* (Akademie, Berlin, 1990)
73. D. Goldstein, *Polarized Light*, 2nd edn. (Marcel Dekker, New York, 2003)
74. G.E. Jellison, F.A. Modine, Appl. Opt. **36**, 8184 (1998)
75. G.E. Jellison, F.A. Modine, Appl. Opt. **36**, 8190 (1998)
76. G.E. Jellison, *Handbook of Ellipsometry*, ed. by E.A. Irene, H.G. Tompkins. (William Andrew Publishing, Norwich, 2005)
77. W.S. Weiglhofer, in *Introduction to Complex Mediums for Optics and Electromagnetics*, ed. by W.S. Weiglhofer, A. Lakhtakia. (SPIE Press, Bellingham, 2003)
78. M. de Graef, M.E. McHenry, *Structure of Materials: An Introduction to Crystallogrpahy, Diffraction, and Symmetry* (Cambridge University Press, Cambridge, 2007)
79. J.A. Woollam Co., Inc., Guide to Using WVASE32®: Spectroscopic Ellipsometry Data Acquisition and Analysis Software (2008)
80. M. Born, E. Wolf, *Principles of Optics*, 7th edn. (Cambridge University Press, Cambridge, 2003)
81. D.W. Berreman, J. Opt. Soc. Am. **62**, 502 (1972)
82. M. Schubert, Phys. Rev. B **53**, 4265 (1996)
83. M. Schubert, B. Rheinländer, C. Cramer, H. Schmiedel, J.A. Woollam, C.M. Herzinger, B. Johs, J. Opt. Soc. Am. A **13**, 1930 (1996)
84. M. Schubert, C.M. Herzinger, Phys. Status Solidi (a) **188**, 1563 (2001)
85. M.S. Cohen, J. Appl. Phys. **32**, S87 (1961)
86. I.J. Hodgkinson, Q.H. Wu, *Birefringent Thin Films and Polarizing Elements* (World Scientific, Singapore, 1997)
87. L.I. Epstein, J. Opt. Soc. Am. **42**, 806 (1952)
88. D. Schmidt, B. Booso, T. Hofmann, E. Schubert, A. Sarangan, M. Schubert, Appl. Phys. Lett. **94**, 011914 (2009)
89. D. Schmidt, B. Booso, T. Hofmann, E. Schubert, A. Sarangan, M. Schubert, Opt. Lett. **34**, 992 (2009)
90. D. Schmidt, E. Schubert, M. Schubert, Appl. Phys. Lett. **100**, 011912 (2012)

91. D.A.G. Bruggeman, Ann. Physik (Leipzig) **24**, 636 (1953)
92. D. Polder, J.H. van Santen, Physica **12**, 257 (1946)
93. C.G. Granquist, O. Hunderi, Phys. Rev. B **16**, 3513 (1977)
94. G.B. Smith, Opt. Commun. **71**, 279 (1989)
95. C.G. Granqvist, D.L. Bellac, G.A. Niklasson, Renew. Energy **8**, 530 (1996)
96. A. Shivola, in *Electromagnetic Mixing Formulas and Applications*. Electromagnetic Waves, vol. 47. (The Institution of Electrical Engineers, London, 1999)
97. O. Wiener, Abh. Sächs. Ges. Akad. Wiss., Math.-Phys. Kl. **32**, 509 (1912)
98. W.L. Bragg, A.B. Pippard, Act. Cryst. **6**, 865 (1953)
99. H.K. Pulker, E. Jung, Thin Solid Films **7**, 57 (1971)
100. D.E. Aspnes, Thin Solid Films **89**, 249 (1982)
101. N.J. Podraza, C. Chen, I. An, G.M. Ferreira, P.I. Rovira, R. Messier, R.W. Collins, Thin Solid Films **455–456**, 571 (2004)
102. W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1988)
103. C.M. Herzinger, P.G. Snyder, B. Johs, J.A. Woollam, J. Appl. Phys. **77**, 1715 (1995)
104. C.M. Herzinger, H. Yao, P.G. Snyder, F.G. Celii, Y.-C. Kao, B. Johs, J.A. Woollam, J. Appl. Phys. **77**, 4677 (1995)
105. B. Johs, Thin Solid Films **234**, 395 (1993)
106. E.D. Palik (ed.), *Handbook of Optical Constants of Solids* (Academic Press, Boston, 1991)
107. D.W. Lynch, C.G. Olson, J.H. Weaver, Phys. Rev. B **11**, 3617 (1975)
108. G. Beydaghyan, C. Buzea, Y. Chi, C. Elliott, K. Robbie, Appl. Phys. Lett. **87**, 153103 (2005)
109. G.W. Mbise, G.A. Niklasson, C.G. Granqvist, S. Palmer, J. Appl. Phys. **80**, 5361 (1996)
110. G.W. Mbise, D.L. Bellac, G.A. Niklasson, C.G. Granqvist, J. Phys. D: Appl. Phys. **30**, 2103 (1997)
111. G. Hass, A.P. Bradford, J. Opt. Soc. Am. **47**, 125 (1957)
112. M. Knez, K. Nielsch, L. Niinistö, Adv. Mater. **19**, 3425 (2007)
113. O. Albrecht, R. Zierold, C. Patzig, J. Bachmann, C. Sturm, B. Rheinländer, M. Grundmann, D. Görlitz, B. Rauschenbach, K. Nielsch, Phys. Status Solidi B **247**, 1365 (2010)
114. O. Albrecht, R. Zierold, S. Allende, J. Escrig, C. Patzig, B. Rauschenbach, K. Nielsch, D. Görlitz, J. Appl. Phys. **109**, 093910 (2011)
115. C.-C. Wang, C.-C. Kei, Y.-W. Yu, T.-P. Perng, Nano Lett. **7**, 1566 (2007)
116. J.J. Steele, M. Taschuk, M.J. Brett, IEEE Sens. J. **8**, 1422 (2008)
117. M.M. Hawkeye, K.M. Krause, M.J. Brett, Proc. SPIE **7356**, 73560G (2009)
118. D. Schmidt, E. Schubert, M. Schubert, Mat. Res. Soc. Symp. Proc. **1409**, CC13 (2012)
119. D.O. Smith, M.S. Cohen, G.P. Weiss, J. Appl. Phys. **31**, 1755 (1960)
120. R.J. King, S.P. Talim, Opt. Acta **28**, 1107 (1981)
121. I. Hodgkinson, F. Horowitz, H.A. Macleod, M. Sikkens, J.J. Wharton, J. Opt. Soc. Am. A **2**, 1693 (1985)
122. I. Hodgkinson, Q.H. Wu, J. Hazel, Appl. Opt. **37**, 2653 (1998)
123. I. Hodgkinson, Q.H. Wu, Appl. Opt. **38**, 3621 (1999)
124. R.A. May, D.W. Flaherty, C.B. Mullins, K.J. Stevenson, J. Phys. Chem. Lett. **1**, 1264 (2010)
125. K.Y. Bang, E.H. Choi, J.S. Kyung, I. An, S.-H. Woo, Y.J. Park, C.K. Hwangbo, J. Korean Phys. Soc. **46**, 137 (2005)
126. G. Beydaghyan, K. Kaminska, T. Brown, K. Robbie, Appl. Opt. **43**, 5343 (2004)
127. K. Kaminska, A. Amassian, L. Martinu, K. Robbie, J. Appl. Phys. **97**, 013511 (2005)
128. J. Gospodyn, J.C. Sit, Opt. Mater. **29**, 318 (2006)
129. S.-H. Hsu, E.-S. Liu, Y.-C. Chang, J.N. Hilfiker, Y.D. Kim, T.J. Kim, C.-J. Lin, G.-R. Lin, Phys. Status Solidi (a) **205**, 876 (2008)
130. N.G. Wakefield, J.B. Sorge, M.T. Taschuk, L.W. Bezuidenhout, M.J. Brett, J.C. Sit, J. Opt. Soc. Am. A **28**, 1830 (2011)
131. I.S. Nerbø, S.L. Roy, M. Foldyna, M. Kildemo, E. Søndergård, J. Appl. Phys. **108**, 014307 (2010)

132. M. Mansour, A.-S. Keita, B. Gallas, J. Rivory, A. Besnard, N. Martin, Opt. Mater. **32**, 1146 (2010)
133. T. Hofmann, D. Schmidt, A. Boosalis, P. Kühne, R. Skomski, C.M. Herzinger, J.A. Woollam, M. Schubert, E. Schubert, Appl. Phys. Lett. **99**, 081903 (2011)
134. T. Mackay and A. Lakhtakia, J. Nanophoton. **6**, 069501 (2012)
135. D. Stroud, Phys. Rev. B **12**, 3368 (1975)
136. T. Motohiro, Y. Yaga, Appl. Opt. **28**, 2466 (1989)
137. J.C.M. Garnett, Phil. Trans. Soc. Lond. A **203**, 385 (1904)
138. S. Linden, C. Enkrich, M. Wegener, J. Zhou, T. Koschny, C.M. Soukoulis, Science **306**, 1351 (2004)
139. C. Rockstuhl, F. Lederer, C. Etrich, T. Zentgraf, J. Kuhl, H. Giessen, Opt. Express **14**, 8827 (2006)
140. C. Rockstuhl, T. Zentgraf, E. Pshenay-Severin, J. Petschulat, A. Chipouline, J. Kuhl, T. Pertsch, H. Giessen, F. Lederer, Opt. Express **15**, 8871 (2007)
141. R. Abdeddaïm, G. Guida, A. Priou, B. Gallas, J. Rivory, Appl. Phys. Lett. **94**, 081907 (2009)
142. B. Gallas, N. Guth, J. Rivory, H. Arwin, R. Magnusson, G. Guida, J. Yang, K. Robbie, Thin Solid Films **519**, 2650 (2011)
143. M.O. Jensen, M.J. Brett, Opt. Express **13**, 3348 (2005)
144. I. Šolc, J. Opt. Soc. Am. **55**, 621 (1965)
145. A. Lakhtakia, W. Weiglhofer, Proc. R. Soc. Lond. A **448**, 419 (1995)
146. F. Chiadini, A. Lakhtakia, Microw. Opt. Technol. Lett. **42**, 135 (2004)
147. M. Schubert, T. Hofmann, C.M. Herzinger, J. Opt. Soc. Am. A **20**, 347 (2003)
148. P.I. Rovira, R.A. Yarussi, R.W. Collins, R. Messier, V.C. Venugopal, A. Lakhtakia, K. Robbie, M.J. Brett, Appl. Phys. Lett. **71**, 1180 (1997)
149. K. Robbie, D.J. Broer, M.J. Brett, Nature **399**, 764 (1999)
150. A.C. van Popta, J.C. Sit, M.J. Brett, Appl. Opt. **43**, 3632 (2004)
151. J.B. Sorge, A.C. van Popta, J.C. Sit, M.J. Brett, Opt. Express **14**, 10550 (2006)
152. K.M. Krause, M.J. Brett, Adv. Funct. Mater. **18**, 3111 (2008)
153. J.K. Gansel, M. Thiel, M.S. Rill, M. Decker, K. Bade, V. Saile, G. von Freymann, S. Linden, M. Wegener, Science **325**, 1513 (2009)
154. Y. Zong, Y.C. Shin, C.M. Kim, B.G. Lee, E.H. Kim, Y.J. Park, K.M.A. Sobahan, C.K. Hwangbo, Y.P. Lee, T.G. Kim, J. Mater. Res. **23**, 2500 (2008)
155. X. Chen, K. Jiang, Nanotechnology **19**, 215305 (2008)
156. D.W. Flaherty, R.A. May, S.P. Berglund, K.J. Stevenson, C.B. Mullins, Chem. Mater. **22**, 319 (2010)
157. N.A. Luechinger, S. Loher, E.K. Athanassiou, R.N. Grass, W.J. Stark, Langmuir **23**, 3473 (2005)
158. G. Mertens, R.B. Wehrspohn, H.-S. Kitzerowa, S. Matthias, C. Jamois, U. Gösele, Appl. Phys. Lett. **87**, 241108 (2005)

Chapter 11
THz Generalized Ellipsometry Characterization of Highly-Ordered Three-Dimensional Nanostructures

Tino Hofmann, Daniel Schmidt and Mathias Schubert

Abstract Routine generalized ellipsometry measurements at terahertz (THz) frequencies have become available only very recently, and we present and discuss the application to highly-ordered three-dimensional nanostructure thin films. Such nanostructure thin films are obtained from glancing angle deposition, and consist of slanted columnar structures with high spatial coherence. The slanted columnar thin films reveal strong birefringence due to electronic coupling and screening phenomena. Despite their extreme smallness compared with the THz wavelength equivalent, slanted columnar nanostructure thin films can be used as sensors for dielectric fluids in transmission or reflection geometries, where measurements can be made through the back side of THz-transparent substrates. We describe an anisotropic biaxial effective medium dielectric function approach which comprises structural, geometrical and constituent fraction information, and which enables quantitative analysis of THz generalized ellipsometry measurements. We further describe a frequency-domain generalized ellipsometer setup which incorporates backward wave oscillator sources.

11.1 Introduction

Ellipsometric investigations with electromagnetic radiation at THz frequencies offer unique access to high-frequency electrical properties that are of great interest for

---

T. Hofmann (✉) · D. Schmidt · M. Schubert
Department of Electrical Engineering and Center for Nanohybrid Functional Materials, University of Nebraska-Lincoln, Lincoln, NE68588-0511, USA
e-mail: thofmann@engr.unl.edu

D. Schmidt
e-mail: schmidt@huskers.unl.edu

M. Schubert
e-mail: schubert@engr.unl.edu

materials research, for design of high-speed communication systems, and for novel security screening applications [1–3], for example. Recent success in development of THz source and detector technologies has spurred growing interest in spectroscopic THz instrumentation and in novel sensing and communication systems [4–9]. Availability and access to spectroscopic ellipsometry instrumentation is a prerequisite for successful characterization of THz optical properties of materials, and optical system components for novel sensing and communication devices [10, 11]. In the same way as for infrared-visible-ultraviolet (IR-VIS-UV) optics, accurate THz material optical constants are needed for designing components, verifying those designs, and for developing and controlling manufacturing processes. Accurate measurements of optical properties of materials at THz frequencies are of fundamental importance also for basic scientific research [9]. There exists a wealth of fascinating excitation mechanisms with eigenfrequencies in the THz domain in condensed and soft matter such as spin transitions, collective modes of biological molecules, local free charge carrier oscillations, dynamic motion of magnetic domains, ferroelectric domains, or collective charge phenomena, as examples. Measurement of the complex optical properties at THz frequencies may further allow exploration of novel physical phenomena as observed in artificially structured meta-materials [12].

Ellipsometry is established as the preeminent method used to precisely determine the optical properties for wavelengths from the far infrared (FIR) through the vacuum ultra violet (VUV) [13]. Ellipsometers are employed in a wide variety of applied and basic research fields because they can produce highly accurate, non-destructive measurements suitable for quantitative analysis. Ellipsometry in traditional FIR-VIS-VUV spectral ranges utilizes essentially steady-state illumination and modulation frequencies much lower than the illuminating light optical frequency. The key element of ellipsometry, as compared to other optical measurement techniques, is that measurements are made synchronously with the modulation of the optical polarization state. At the most basic level, one polarization state is measured relative to another polarization state. Ellipsometry is therefore much less sensitive to non-idealities of the measurement system than intensity-normalized measurement techniques. Ellipsometry in the THz frequency-domain has only recently become available as a research tool, and experimental reports are not exhaustive [14–17].

THz Ellipsometry   THz ellipsometry instrumentations were reported in both the time-domain as well as in the frequency-domain configurations. The first time-domain THz ellipsometry setup was demonstrated by Nagashima and Hangyo [18]. Nagashima and Hangyo augmented a THz time-domain spectrometer by fixed polarizers in order to measure the p- and s-polarized reflectivities. Thereby the complex optical constants of a moderately phosphorous-doped *n*-type silicon substrate were determined. A similar approach was used later by Ino et al. [19] in magneto-optical Kerr effect measurements on InAs. The first frequency-domain THz generalized spectroscopic ellipsometry (GSE) experiments and magneto-optic GSE experiments using highly brilliant THz Synchrotron radiation were reported by Hofmann, Schubert and others for the determination of free charge carrier properties in ZnMnSe/GaAs and highly oriented pyrolytic graphite [20]. Recently, a

frequency-domain THz ellipsometer setup using a high-power, spectrally tunable, single-wavelength backward wave oscillator was described and applied for the determination of free charge carrier diffusion profiles in silicon [21, 22]. A similar setup is used here for analysis of optically anisotropic highly-ordered three-dimensional nanostructure thin films.

Sculptured Thin Films   Sculptured thin films (STFs) present an interesting class of self-organized, artificially made materials with three-dimensional, highly spatially coherent arrangements of nanostructures. The intrinsic building block is formed of columns with a controllable inclination angle towards the substrate surface (Fig. 11.1). Geometry, dimension, and fabrication material are accessible design parameters, and STFs can be tailored to obtain desired physical properties. STF properties such as reflectance, transmittance, and conductance may differ significantly from the constituent material's bulk form [23–25]. Contemporary interest in materials for THz electronic, optoelectronic, and optical applications is redrawing attention to STFs that may enable designed optical properties for the THz frequency region. In this frequency range, existing materials are either purely dielectric and transparent, or electrically conductive and highly absorbing. STFs prepared from electrically conductive materials offer the interesting opportunity to design transparent materials with very large dielectric polarizabilities. The strong polarizability originates from the highly coherent electrically insulated arrays of sub-wavelength dipole antennas. Coupling between adjacent dipoles leads to a directionally dependent electromagnetic response, observable as optical anisotropy. Few information is available on the optical properties and anisotropic refraction and extinction coefficients of metal STFs. Recently, we have demonstrated that GSE in the visible spectral range serves as an ideal tool for the determination of the anisotropic dielectric properties of STFs
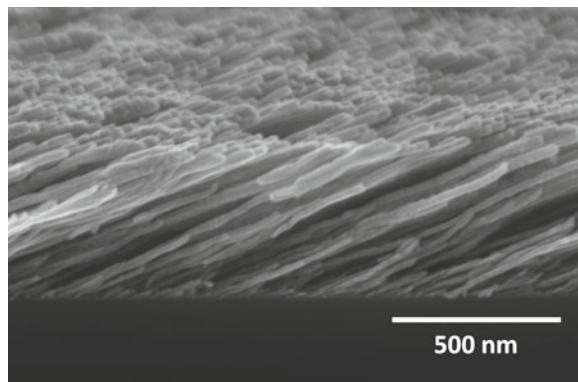
Fig. 11.1   SEM micrograph (sample tilted 15°) of a STF composed of *slanted* Co nanocolumns (F1-STF) which was deposited by glancing angle deposition onto a silicon substrate. The slanting angle is 65° (Reprinted with permission from Ref. [12]. Copyright 2011, American Institute of Physics.)

as described in detail in Chap. 10 of this editorial. No information seems to exist on
metal STF optical properties in the THz region.

11.2 THz Ellipsometry Instrumentation

The THz ellipsometer system used in this work is described in Ref. [14]. The
instrument operates in both the polarizer-sample-rotating analyzer as well as in the
polarizer-sample-rotating compensator analyzer arrangements [13] and permits ellip-
sometric measurements in the spectral range from 0.2 to 1.5 THz. The experimental
setup is based on a $\theta - 2\theta$ high precision goniometer which allows for measurements
at angles of incidence Φ_a from 30° to 90°. A sketch of the experimental setup is given
in Fig. 11.2. A backward wave oscillator (BWO) is employed as the radiation source.
The nearly linearly polarized light emitted from the BWO is collimated by a lens
(L) before passing an optical copper (C). A polarization rotation system (PR) [26] is
used in combination with a wire-grid polarizer (P) to control the incident polarization
state. The mirrors M1-M3 steer the beam and focus the THz radiation to the sample
stage (S) on the high-resolution goniometer (HG). A continuously rotating polarizer
is used as the polarization state analyzer (A). The THz radiation reflected from the
sample is focused (M4) on a Golay cell (GC) which serves as detector. The BWO
tube used here emits quasi linearly polarized light with a very narrow bandwidth of
approximately 2 MHz and a very high output power of 0.1–0.01 W in the frequency
range from 107 to 177 GHz. The high output power in this base band is converted
to higher frequency bands using a set of frequency multipliers. Further instrumen-
tation details, data acquisition and analysis procedures, and demonstration for the
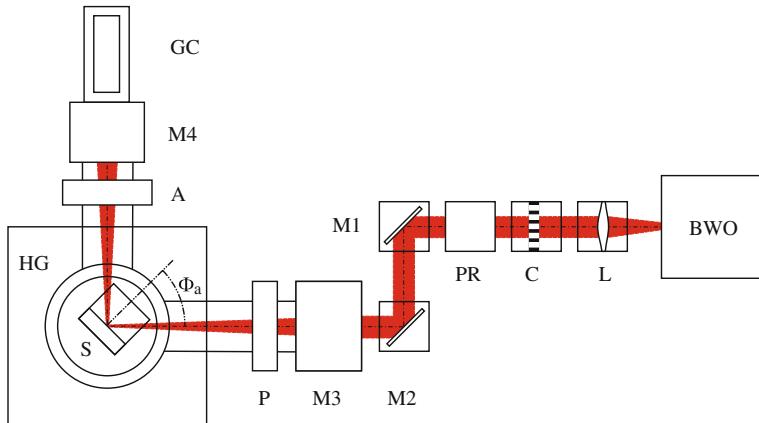
Fig. 11.2 Beam path of the THz ellipsometer employed for the spectral range from 0.2 to 1.5 THz.
The system is operating in a polarizer (P)—sample (S)—rotating analyzer (A) configuration [13].
See text for descriptions, and Ref. [14] for details (Reprinted with permission from Ref. [14].
Copyright 2010, American Institute of Physics.)
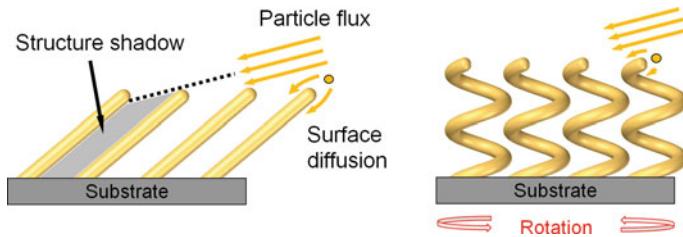
Fig. 11.3 Schematic drawing of deposition process for highly-ordered three-dimensional nanostructure thin film by glancing angle deposition. *Left* with steady substrate orientation columnar thin films with common slanting direction are obtained. *Right* slow and continuous substrate rotation, which is equivalent to a steady change in the direction of the incoming particle flux results in helical structures, for example. See Refs. [27, 28, 34] for further details, or Chap. 10 in this editorial [35]

determination of complex-valued dielectric constants in semiconductor materials are reported in Ref. [14].

11.3 Sculptured Thin Films

Highly-ordered three-dimensional nanostructure thin films such as sculptured thin films can be fabricated by glancing angle deposition (Fig. 11.3) from a wide variety of materials, including insulators, metals, semiconductors, and organic materials, vaporized by sputtering [27–29], pulsed laser deposition [30], thermal or (most commonly) electron beam evaporation [24, 31–33]. STFs are fabricated under low-adatom-mobility conditions, where the sticking coefficient (ratio of adsorbed adatoms and total number of adatoms arriving within the same period of time) is essentially unity, and hence substrate temperatures of less than 10 % of the melting point of the evaporant are desired for columnar growth [31]. STFs can be deposited on many substrate materials (e.g., glass, polymers, etc.) since the substrate may be kept at room-temperature (Fig. 11.3).

11.4 THz Ellipsometry Data Analysis

The experimental THz Jones or Mueller matrix elements depend on the thickness d and the dielectric function ε of all constituents of the sample. However, in order to obtain ε and d of materials in layered samples, stratified model calculations have to be performed [36]. The standard model for analyzing spectroscopic ellipsometry data is based on a sequence of homogeneous (isotropic or anisotropic) layers with smooth and parallel interfaces. The light propagation within the sample is calculated using standard matrix formalism for (isotropic or anisotropic) multilayered systems with plane parallel interfaces.

Two approaches are commonly used to determine the ε values of interest. In the wavelength-by-wavelength analysis approach the ε values are extracted from the experimental data for each wavelength and independent of all other spectral data points. For this procedure d and ε spectra of all other sample constituents have to be known. Physically relevant parameters, such as phonon mode frequencies and broadening parameters, for instance, are determined by comparing the dielectric function obtained from the wavelength-by-wavelength analysis with a line shape model. Alternatively, parameterized model dielectric functions can be fitted directly to the experimental data to extract relevant physical parameters. The major advantage of this approach is that it prevents experimental noise in the ellipsometric data from becoming part of the extracted dielectric function. However, in some cases, e.g., for bulk samples, the wavelength-by-wavelength analysis is preferable because the experimental data are directly inverted numerically without additional model assumptions to obtain the ε values of interest.

11.4.1 THz Model Dielectric Function Approaches for Slanted Columnar Thin Films

The anisotropic dielectric functions of Co F1-STFs can be described by an anisotropic Bruggeman effective medium approximation (AB-EMA) [12]. The Bruggeman formalism describes the homogenization for randomly oriented elliptical inclusions in a homogeneous host medium. This formalism can be generalized for the case of highly oriented elliptical inclusions (Fig. 11.4) [37–40]. For highly spatially coherent, oriented elliptical inclusions the effective dielectric function tensor is described by the three major components ε_a, ε_b, and ε_c along the major axes a, b, and c of an orthorhombic system. ε_j with $j = a, b, c$ are given in implicit form by [41]:

$$\varepsilon_j = \varepsilon_{\mathrm{m}} + \frac{f(\varepsilon_{\mathrm{i}} - \varepsilon_{\mathrm{m}})\varepsilon_j}{\varepsilon_j + L_j(\varepsilon_{\mathrm{i}} - \varepsilon_j)}, \tag{11.1}$$

where the dielectric permittivity and volume fraction of the nanocolumnar inclusions are denoted by ε_{i} and f, respectively. ε_{m} is the permittivity of the host medium. L_j are the depolarization factors of the elliptical inclusions along the major polarizability axes. Regardless of the ellipsoidal shape, the sum of the three depolarization factors satisfies $L_a + L_b + L_c = 1$ [40].

For metal nanocolumnar structures the THz dielectric permittivity of the metal inclusions, ε_{i}, can be described here by the classical Drude formalism [42, 43]:

$$\varepsilon_{\mathrm{i}}(\omega) = \varepsilon_{\infty} - \varepsilon_{\infty}\frac{\omega_{\mathrm{p}}^2}{\omega(\omega + i/\tau)}, \tag{11.2}$$

where ω_p is the screened plasma frequency, τ is the average energy-independent scattering time. The DC conductivity σ_{dc} is related to ω_p by $\sigma_{dc} = \varepsilon_{\infty}\tilde{\varepsilon}_0\omega_p^2\tau$, where
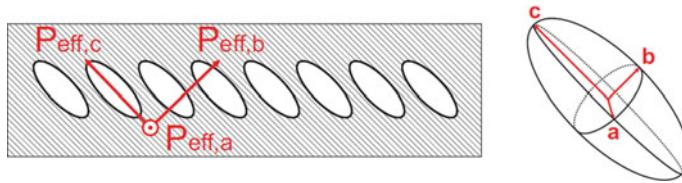
Fig. 11.4 Effective medium scenario with mixtures of ellipsoidal inclusions and a homogeneous host medium. The aligned inclusions produce anisotropic effective properties with three effective polarizabilities $P_{\mathrm{eff},j}$ according to the shape of the inclusions [35]

$\omega_p^2 = Nq^2/(\varepsilon_\infty \tilde{\varepsilon}_0 m)$ and μ is the mobility given by $\mu = q\tau/m$, $\tilde{\varepsilon}_0$ is the vacuum permittivity, N is the free charge carrier concentration, and q denotes the charge. The high frequency dielectric constant and the effective mass of the free charge carriers are denoted by ε_∞ and m, respectively.

The dielectric permittivity of the host medium, ε_{m}, is equal to 1 if the host medium is air as for the measurements obtained from as-grown Co nanocolumnar structures described in Sect. 11.6. However, if the nanostructures are immersed in a polar liquid, water for instance, the dielectric permittivity of the host medium has to be changed accordingly as described in Sect. 11.7. A Debye model with two relaxation terms was used to describe the dielectric permittivity of water here [44, 45]:

$$\varepsilon_{\mathrm{m}}(\omega) = \varepsilon_\infty + \sum_{j=1}^{2} \frac{\varepsilon_j - \varepsilon_{j+1}}{1 + i\omega\tau_j}, \qquad (11.3)$$

where ε_1 is the static dielectric constant, ε_2 is an intermediate step in the dielectric function and $\varepsilon_3 = \varepsilon_\infty$ is the high frequency limit. τ_1 and τ_2 are the relaxation time constants of the two relaxation processes.

11.5 THz Dielectric Function of Isotropic Thin Film Materials

The THz ellipsometric data of two different phosphorous-doped n-type Si substrates are examined to demonstrate the analysis of thin film materials with an isotropic THz optical response. One sample has a high dopant concentration and a nominal resistivity of $0.02\,\Omega$ cm, the other sample a low dopant concentration and a nominal resistivity of $5\,\Omega$ cm. The low-doped sample is transparent in the THz range and was double side polished in order to provide two well-defined plane parallel interfaces.

The experimental and best-model calculated $-M_{12,21}$-specta and M_{33}-spectra of the highly phosphorous-doped n-type Si substrate obtained at an angle of incidence $\Phi_a = 75°$ are depicted in Fig. 11.5a, b, respectively. The data obtained from the THz ellipsometer setup and the commercial mid infrared (MIR) ellipsometer are in excellent agreement with the calculations derived from a common model. A five
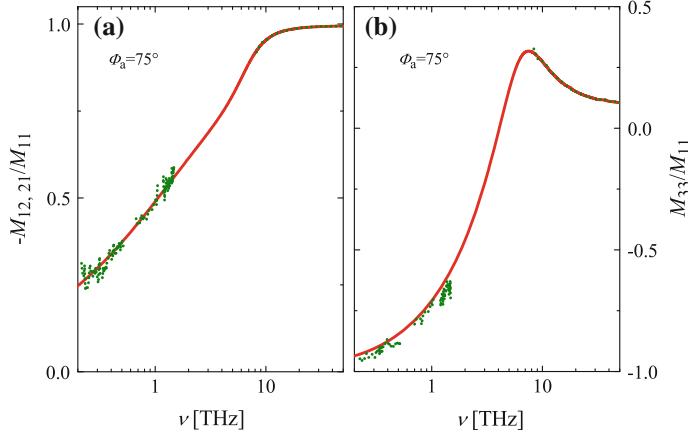
Fig. 11.5 Experimental (*dotted lines*) and best-model calculation (*solid lines*) $-M_{12,21}$-spectra (**a**) and M_{33}-spectra (**b**) of the highly doped *n*-type Si substrate at $\Phi_a = 75°$ (Reprinted with permission from Ref. [14]. Copyright 2010, American Institute of Physics.)

phase model (ambient/ native oxide/depletion region/highly doped Si/ambient) was employed for the simultaneous analysis of the THz and MIR data sets.

The dielectric function $\varepsilon(\omega)$ for doped silicon, governed by Eq. (11.2), dominates the features in the THz and infrared spectral domains. However, in order to fully describe the experimental data throughout the entire THz-MIR region, the contributions of the oxide layer and the carrier depletion region at the substrate surface have to be included in the model. The dielectric function of the oxide consists of three Lorentzian harmonic oscillators modeling the optically active Si–O rocking and stretching vibrations [46]. The oxide produces very subtle effects in the M_{33}-spectra between 13 and 37 THz. The presence of the depletion layer is expressed as small offsets primarily in the MIR spectral region. The model analysis yields the thickness of the SiO$_2$ layer and the carrier depletion region of $d_{SiO2} = 2.2 \pm 0.2$ nm and $d_{depl} = 25 \pm 4$ nm, respectively. For the substrate resistivity and scattering time we obtain $\rho = 0.023 \pm 0.001 \, \Omega$ cm and $\tau = 26.9 \pm 0.5$ fs. Assuming an effective electron mass of $0.26 \, m_0$ (Ref. [47]) this corresponds to a free charge carrier concentration $N = (1.51 \pm 0.01) \times 10^{18} \, \text{cm}^{-3}$ and a mobility $\mu = 182 \pm 3 \, \text{cm}^2/\text{Vs}$.

Figure 11.6 shows the experimental and best-model calculated $-M_{12,21}$-spectra and M_{33}-spectra obtained from the low phosphorous-doped silicon substrate for two angles of incidence $\Phi_a = 55°$ and $65°$. The spectra are dominated by oscillations which originate from multiple internal reflections between the top and bottom silicon-air interface. It can be clearly seen that the amplitude of the Fabry-Pérot oscillations pattern is decreasing for lower frequencies. This can be attributed to free charge carrier absorption in the sample. The effects of the surface depletion and the SiO$_2$ layer to the $-M_{12,21}$-spectra and M_{33}-spectra are negligible and a simple three phase model (ambient/Si/ambient) was used for the data analysis consisting only of a doped Si layer. The best-model parameters obtained for the resistivity and scattering time are $\rho = 5.01 \pm 0.16 \, \Omega$ cm and $\tau = 239 \pm 9$ fs, respectively. This corresponds to
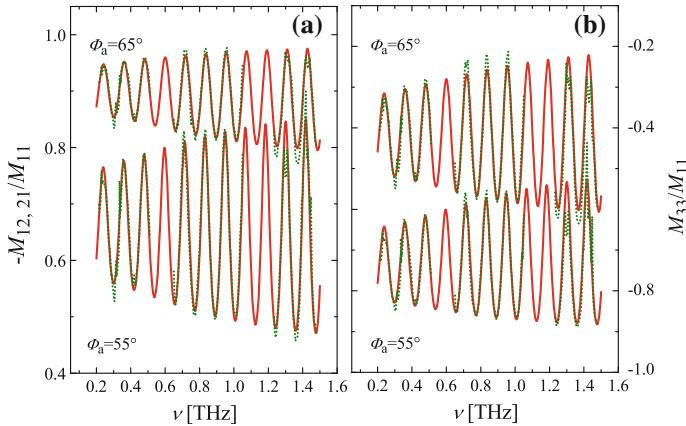
Fig. 11.6 Experimental (*dotted lines*) and best-model calculation (*solid lines*) $-M_{12,21}$-spectra (**a**) and M_{33}-spectra (**b**) of the low doped *n*-type silicon substrate at $\Phi_a = 55°$ and $65°$ (Reprinted with permission from Ref. [14]. Copyright 2010, American Institute of Physics.)
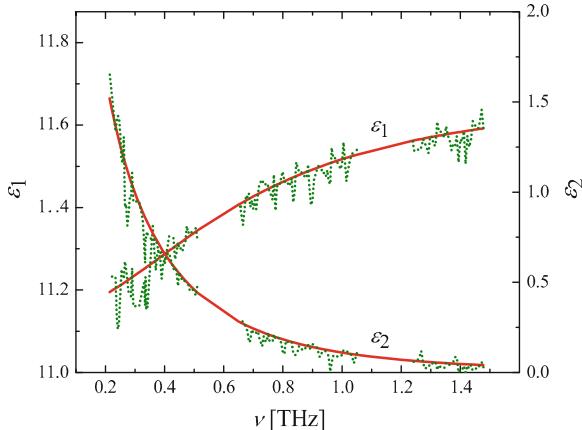
Fig. 11.7 Wavelength-by-wavelength extracted (*dotted lines*) and Drude model calculated real (ε_1) and imaginary (ε_2) part of the dielectric function of the low phosphorous-doped *n*-type silicon substrate (Reprinted with permission from Ref. [14]. Copyright 2010, American Institute of Physics.)

a free charge carrier concentration of $N = 7.7 \pm 0.4 \times 10^{14}\,\mathrm{cm}^{-3}$ and mobility of $\mu = 1617 \pm 80\,\mathrm{cm}^2/\mathrm{(Vs)}$ if an averaged electron effective mass of $0.26\,m_0$ (Ref. [47]) is assumed. The plasma frequency is approximately $0.14\,\mathrm{THz}$. For the thickness of the wafer we obtain $384.2 \pm 0.1\,\mu\mathrm{m}$ which is highly consistent with the results obtained from the infrared spectral range.

The real (ε_1) and imaginary (ε_2) part of the spectroscopic dielectric values for the doped silicon were also extracted on a model free wavelength-by-wavelength basis (dotted lines) and are compared with the previous determined Drude functional model (solid lines) in Fig. 11.7. The wavelength-by-wavelength extracted data are

in excellent agreement with the Drude model calculated data and show the typical increase in ε_2 for lower frequencies.

11.6 THz Dielectric Anisotropy of Metal Slanted Columnar Thin Films

Highly-ordered three-dimensional nanostructured metal thin films exhibit unique anisotropic optical properties in the THz spectral range. A 450 nm thick nanostructured cobalt film composed of slanted nanocolumns was grown by electron-beam glancing angle deposition is investigated. A high-resolution field-emission SEM edge view image of the sample is shown in Fig. 11.1. The film was deposited in a customized ultrahigh vacuum chamber onto a low phosphorous-doped n-type (001) silicon substrate. The deposition was carried out at an angle of 85° between the incident particle flux direction and the substrate normal. Further details about the growth process are omitted here and the interested reader is referred to Ref. [48]. The slanting angle of the Co nanocolumns is 65°.

The THz-GSE measurements were carried out in the spectral range from 0.65 to 1.00 THz with a resolution of 1 GHz using the custom-built frequency-domain THz ellipsometer described in Sect. 11.2. Experimental THz-GSE data, which are represented here using the Mueller matrix formalism, were obtained at two different angles of incidence $\Phi_a = 55°$ and 65° and for three different in-plane sample rotation angles $\varphi = 90°$, 135°, and 180°, and for which the columnar slanting plane is oriented parallel, oblique (135°), and perpendicular to the plane of incidence, respectively.

The experimental THz-GSE data sets were analyzed using a stratified layer model calculation where all model calculated data were matched simultaneously as closely as possible to the experimental THz-GSE data by varying relevant physical model parameters.

For best-match model calculations, a three phase model consisting of a low doped n-type silicon substrate/Co F1-STF layer/ambient was implemented. Figure 11.8 depicts the normalized block-diagonal THz-GSE data M_{12}, M_{21}, and M_{33} of the Co F1-STF sample at an angle of incidence $\Phi_a = 55°$ for three different in-pane rotations $\varphi = 90°$, 135°, and 180°. For comparison, the Mueller matrix spectra of the silicon substrate before Co F1-STF deposition are plotted. The experimental (dotted lines) and best-model calculated (solid lines) data are found to be in very good agreement. The Mueller matrix spectra are dominated by Fabry-Pérot interference pattern originating from the entire sample. The frequency of the Fabry-Pérot interference pattern is determined by the layer thicknesses of all sample constituents, but it is dominated by the thickness of the substrate which is $378 \pm 1 \, \mu m$. At the first glance, one can observe that the interference amplitude is reduced for the Co F1-STF sample in comparison to the bare silicon substrate due to the presence of the nanocolumnar film. It can be seen in Fig. 11.8 that the Fabry-Pérot pattern observed for the Co F1-STF sample change as a function of the in-plane rotation angle. With

Fig. 11.8 Experimental (*dotted lines*) and best-model calculated (*solid lines*) block-diagonal Mueller matrix spectra (*upper panel* M_{12} and M_{21}, *lower panel* M_{33}) for the Co F1-STF sample for three different in-plane rotation angles $\varphi = 90°$, $135°$, and $180°$. The spectra for the silicon substrate before Co STF deposition are shown for comparison (Reprinted with permission from Ref. [12]. Copyright 2011, American Institute of Physics.)
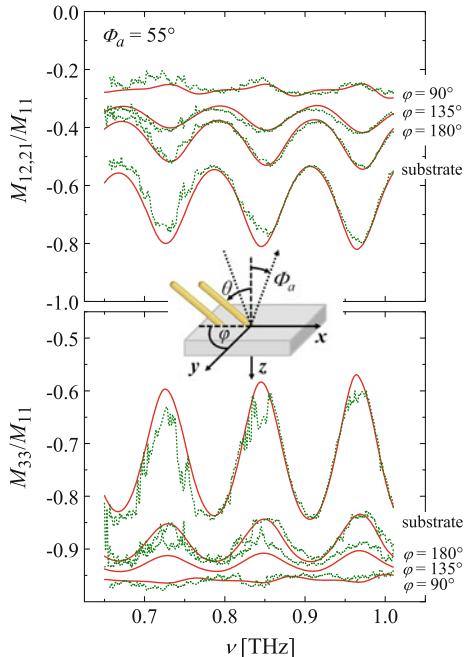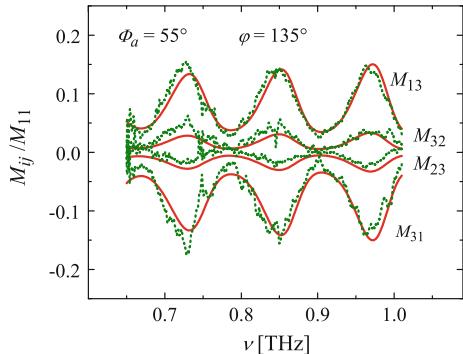
Fig. 11.9 Experimental (*dotted lines*) and best-model calculated (*solid lines*) off-diagonal Mueller matrix spectra for the Co F1-STF sample for an in-plane rotation angle $\varphi = 135°$ (Reprinted with permission from Ref. [12]. Copyright 2011, American Institute of Physics.)

increasing φ the orientation of the slanted nanocolumns change from parallel to perpendicular to the plane of incidence and the amplitudes of the interference oscillations decrease.
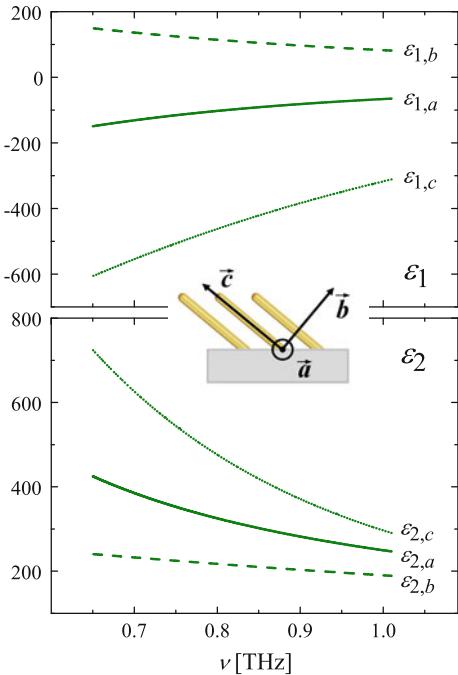
Figure 11.9 depicts off-diagonal block elements of the Mueller matrix of the Co F1-STF sample obtained at an angle of incidence of $\Phi_a = 55°$ and $\varphi = 135°$.[1] The

---

[1] The instrument which was employed for the THz GSE measurements operates in a polarizer-sample-rotating analyzer scheme which gives access to the upper 3×3 block of the Mueller matrix only [13, 14].

anisotropy of the optical response of the nanocolumnar thin film is evident from the non-zero off-diagonal block elements M_{13}, M_{31}, M_{23}, M_{32} which would vanish for an isotropic sample. The best-model parameters for the AB-EMA depolarization factors (Eq. 11.1) are: $L_a = 0.3333(1)$, $L_b = 0.3355(1)$, and $L_c = 0.3312(1)$. The depolarization factors render the dipole shape at the location of the inclusion, and are wavelength dependent. In the THz frequency range, the polarization fields are produced by thousands of neighboring nanocolumns over which one wavelength equivalent of the electric field averages the sample response. Probed at large wavelengths, factors L_j render the shape of the induced polarization almost as a sphere, however, with subtle differences along and perpendicular to the columnar axis. These subtle differences are responsible for the anisotropic dielectric functions, which differ all drastically from the corresponding bulk (cobalt) dielectric properties. The best-model parameters for the STF thickness and volume fraction of the nanocolumns, $d = 441(5)$ nm, and $f = 33(1)$, respectively, are in good agreement with SEM data obtained from Fig. 11.1 and GSE measurements in the NIR-VIS spectral range.[2] As additional structural parameter, we also determined the slanting angle of the intrinsic Cartesian coordinate system of the polarizabilities ε_j, as described previously [48–50], and which is identical with the physical slanting angle of the nanocolumns. For the Drude model parameters (Eq. 11.2) of the Co nanocolumns $\rho = 1.2(3) \times 10^{-5}\,\Omega\,\text{cm}$ and $\tau = 239(7)$ fs were obtained. These parameters were found to be in very good agreement to those obtained for a bulk-like 100 nm thick Co film on a low doped silicon substrate which was analyzed in the same spectral range for comparison. The resistivity model parameter of the Co nanocolumns is approximately twice as large as reported for bulk Co samples [51]. Overall, the agreement with bulk and thin film conductivity is exciting given the limitations of the AB-EMA model approach. The best-model parameters for the silicon substrate are $N = 1.08(4) \times 10^{15}\,\text{cm}^{-3}$ and $\mu = 1375(111)\,\text{cm}^2/(\text{Vs})$ if an electron effective mass of $0.26\,m_0$ (Ref. [47]) is assumed. The anisotropic dielectric functions ε_j and their coordinate system are depicted in Fig. 11.10. The anisotropy induced by the columnar film structure is very large. Note that while the dielectric polarizability (real part of ε_j) is larger than for most dielectric materials, the absorption (imaginary part of ε_j) is reduced by approximately two orders of magnitude compared with bulk or thin film Co. The AB-EMA model approach predicts upon slight modifications of Drude, fraction and/or depolarization parameters that targeted optical properties of STF in the THz range can be achieved by variation of slanting angle, lateral column density, and material.

---

[2] A commercial instrument (M2000VI, J. A. Woollam Co., Inc., Lincoln, NE) was used for the angle-resolved GSE measurements with the spectral range from 400 to 1680 nm. On the fully automated system Φ_a was varied from 45° to 75° in steps of 10°, while φ was varied from 0° to 360° in steps of 6°. Further details on the analysis of the NIR-VIS GSE measurements are omitted here.

Fig. 11.10 Real (*upper panel* $\varepsilon_{1,a}$, $\varepsilon_{1,b}$, and $\varepsilon_{1,c}$) and imaginary (*lower panel* $\varepsilon_{2,a}$, $\varepsilon_{2,b}$, and $\varepsilon_{2,c}$) part of the THz orthorhombic dielectric functions obtained from the best-model calculation for the Co F1-STF sample. The inset depicts the axes locations of the intrinsic Cartesian polarizability system (Reprinted with permission from Ref. [12]. Copyright 2011, American Institute of Physics.)

11.7 THz In-situ Ellipsometry Sensing of Dielectric Fluids Using Metal Slanted Columnar Thin Films

The anisotropic optical response at THz frequencies strongly depends on the dielectric properties of the ambient surrounding the slanted columnar thin films. Exemplarily the difference in the THz range response between STFs in air and STFs immersed in water is discussed. A 450 nm thick STF composed of Co nanocolumns was grown by glancing angle deposition on a THz-transparent silicon substrate. THz measurements were carried out with the sample mounted onto a custom-built liquid cell with the Co STF facing the inside of the cell (Fig. 11.11).

Measurements were recorded in the spectral range from 0.65 to 1.00 THz with a resolution of 1 GHz: (i) with the empty cell (air ambient) and (ii) with the cell filled with nanopure water. Further details can be found in Ref. [52]. The angle of incidence was $\Phi_a = 55°$. The in-plane orientation of the sample was $\varphi = 225°$ during these experiments, i.e., the columnar slanting plane was oriented oblique to the plane of incidence in order to maximize the anisotropic optical response [12]. Figure 11.12 shows the non-zero experimental (dotted lines) and best-model calculated (solid lines) Mueller matrix spectra M_{13}, M_{31}, M_{23}, and M_{32} which are evidence for the anisotropy of the optical response of the STF. The experimental and model calculated
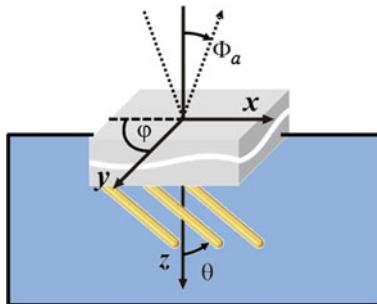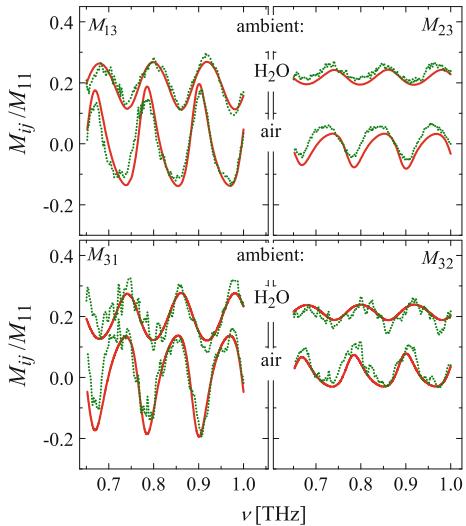
Fig. 11.11 Measurement configuration with THz beam reflected off the backside from the THz-transparent silicon substrate with the STF in contact with fluid ambient. For the in-situ experiments described here, the angle φ is 225° and accordingly the slanting plane is at an oblique angle with respect to the plane of incidence (Reprinted with permission from Ref. [52]. Copyright 2012, Cambridge University Press.)

data are in very good agreement for both the STF in air and water ambient. The Mueller matrix spectra are dominated by a Fabry-Pérot interference pattern, which originates from the double-side polished silicon substrate. One can observe a decrease of the amplitude of the Fabry-Pérot oscillations if the ambient of the STF is changed from air to water. The THz response of the metal STF strongly depends on the dielectric properties of the ambient surrounding the slanted columnar thin films, as well as the dielectric properties of particles attached within the nanostructures. Model calculations using the AB-EMA approach described above deliver quantitative information about amount and THz dielectric constants of the material attached or incorporated within the nanostructures.

For the analysis of the experimental data obtained for the Co F1-STF on a silicon substrate a two layer model composed of air ambient/Co F1-STF/Si substrate/air ambient was used. The measurements for the Co F1-STF on a silicon substrate, where the Co F1-STF side was immersed in a liquid medium, were analyzed using a stratified four layer model calculation composed of air ambient/Si substrate/Co F1-STF/liquid cell medium. The following best-model parameters were obtained for the STF thickness and volume fraction of the nanocolumns: $d = 441(5)$ nm and $f = 33(1)$, respectively. These parameters are in good agreement with SEM data obtained from Fig. 11.1. The best-model parameters for the free charge carrier concentration and mobility of the silicon substrate are $N = 1.08(4) \times 10^{15}$ cm^{-3} and $\mu = 1375(111)$ cm^2/(Vs), respectively, assuming an electron effective mass of 0.26 m_0 [47]. The slanting angle of the intrinsic Cartesian coordinate system of the polarizabilities ε_a, ε_b, and ε_c was found to be identical to the physical slanting angle of the nanocolumns. The following best-model parameters have been obtained for the AB-EMA depolarization factors: $L_a = 0.3333(1)$, $L_b = 0.3355(1)$, and $L_c = 0.3312(1)$ [12]. The best-model Drude parameters resistivity and scattering time for the Co nanostructures were $\rho = 1.2(3) \times 10^{-5}$ Ω cm and $\tau = 239(7)$ fs,

Fig. 11.12 Experimental (*dotted lines*) and best-model calculated (*solid lines*) THz ellipsometry (Mueller matrix) spectra of a Co F1-STF in air and immersed in water (Reprinted with permission from Ref. [52]. Copyright 2012, Cambridge University Press.)

respectively. If immersed in water the dielectric response of the Co nanostructures is screened. This can be clearly observed in a change of the AB-EMA depolarization factors: $L_a = 0.3375(1)$, $L_b = 0.3373(1)$, and $L_c = 0.3252(1)$. Furthermore changes in the best-model Drude parameters resistivity and scattering time for the Co nanostructures are observed. The resistivity increases to $\rho = 1.1(5) \times 10^{-4}\,\Omega\,cm$ and the scattering time decreases to $\tau = 144(37)$ fs. This might be attributed to continued oxidation of the Co nanocolumns. The best-model parameters obtained for the dielectric contribution of water are $\varepsilon_\infty = 3.3(5)$, $\varepsilon_1 = 71(4)$, $\varepsilon_2 = 5.0(1)$, $\tau_1 = 8.3(2)$ ps, and $\tau_2 = 0.18$ ps (not varied during analysis) which are in very good agreement with literature values [44].

The data shown in Fig. 11.12 suggest a new avenue for detection of dielectric properties of liquid materials through THz-transparent windows in fluidic channels or containers. Figure 11.13 depicts calculated off-diagonal Mueller matrix elements M_{13} based on the best-match model obtained from the experimental data shown in Fig. 11.12 for ultra-pure water. In Fig. 11.13 the best-match optical model was recalculated by varying the dielectric optical constant parameter of the fluid constituent, $\varepsilon_{H2O} + \delta\varepsilon_{off}$, and the differences for $M_{13}(\varepsilon_{H2O} + \delta\varepsilon_{off}) - M_{13}(\varepsilon_{H2O})$ are plotted as a 3D surface versus THz frequency. Strong variations of the anisotropic mode coupling properties of the structure substrate + STF can be seen at spectral positions where the Fabry-Pérot interference pattern indicate reflection maxima. The sensitivity to small changes in the fluid dielectric constant suggests this sample system as a sensing device for the actual dielectric constant of the fluid immersing the metal slanted columnar thin film. Note that the sensitivity increases substantially if spectroscopic measurements (data over multiple Fabry-Pérot interferences) are included for data analysis.
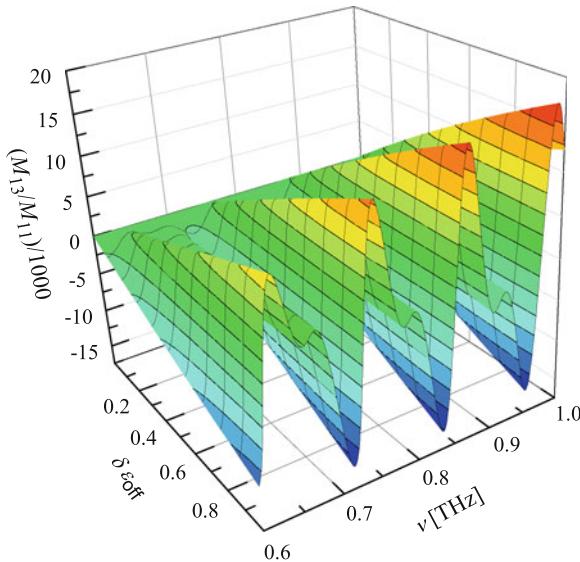
Fig. 11.13 Simulated THz off-diagonal Mueller matrix element difference data for the slanted Co nanocolumns in Fig. 11.12, if immersed in fluids of varying dielectric polarizability, relative to data shown in Fig. 11.12 for ultra-pure water ($\delta\varepsilon_{off}$)

11.8 Summary and Outlook

In summary, we have presented and discussed the application of generalized ellipsometry measurements at THz frequencies for the investigation of highly-ordered three-dimensional nanostructure thin films. We demonstrate that frequency-domain generalized ellipsometric measurements of the anisotropic optical properties of metal nanostructure thin films in the THz spectral range are readily available by using backward wave oscillator light sources emitting intense and highly coherent THz radiation in standard rotating analyzer (or compensator) ellipsometry configurations. The sculptured thin films consisting of three-dimensional, highly spatially coherent arrangements of nanostructures can be routinely obtained from glancing angle deposition. THz GSE measurements of thin films composed of slanted metal nanocolumns reveal strong birefringence due to electronic coupling and screening phenomena. Despite the extreme smallness of the nanostructures compared with the THz wavelength equivalent, screening effects caused by changes in the ambient permittivity lead to measurable changes in the THz birefringence. This effect can be used for sensors of dielectric fluids in transmission or reflection geometries, where measurements can be made through the backside of THz-transparent substrates. We describe an anisotropic biaxial effective medium dielectric function approach which comprise structural, geometrical and constituent fraction information, and which enables quantitative analysis of THz generalized ellipsometry measurements.

Our investigations suggest controlled variability of dielectric polarizability and anisotropy in the THz range by choice of geometry, material, and structure. This represents a new avenue for false-mode-suppressing nanostructured sensors with a wide application range. As a read-out technique THz generalized ellipsometry is envisioned and which will provide non-contact, non-destructive access to the nanostructured sensor properties and allows accurate determination of structural, geometrical and constituent fraction information which are valuable for basic research, and industry.

References

1. P. Siegel, *Terahertz Sensing Technology: Electronic Devices and Advanced Systems Technology*. THz Technology: An Overview, (World Scientific, Singapore, 2003), pp. 1–44
2. D. Mittelman, *Sensing with Terahertz Radiation* (Springer, New York, 2004)
3. D. Woolard, E. Brown, M. Kemp, M. Pepper, Proc. IEEE **93**, 1722 (2005)
4. M. Abo-Bakr, J. Feikes, K. Holldack, P. Kuske, W.B. Peatman, U. Schade, G. Wüstefeld, H.-W. Hübers, Phys. Rev. Lett. **90**, 94801 (2003)
5. F. Floreani, H.W. Koops, W. Elser, Nucl. Instrum. Methods **483**, 488 (2002)
6. R. Kohler, A. Tredicucci, F. Beltram, H.E. Beere, E.H. Linfield, A.G. Davies, D.A. Ritchie, R.C. Iotti, F. Rossi, Nature **417**, 156 (2002)
7. B. Ferguson, X. Zhang, Nat. Mater. **1**, 26 (2002)
8. S.E. Korbly, A.S. Kesar, J.R. Sirigiri, R.J. Temkin, Phys. Rev. Lett. **94**, 054803 (2005)
9. M. Tonouchi, Nat. Photonics **1**, 97 (2007)
10. G.P. Williams, Rep. Prog. Phys. **69**, 301 (2006)
11. J.S. Melinger, S.S. Harsha, N. Laman, D. Grischkowsky, J. Opt. Soc. Am. B **26**, A79 (2009)
12. T. Hofmann, D. Schmidt, A. Boosalis, P. Kühne, R. Skomski, C.M. Herzinger, J.A. Woollam, M. Schubert, E. Schubert, Appl. Phys. Lett. **99**, 081903 (2011)
13. H. Fujiwara, *Spectroscopic Ellipsometry* (Wiley, New York, 2007)
14. T. Hofmann, C.M. Herzinger, A. Boosalis, T.E. Tiwald, J.A. Woollam, M. Schubert, Rev. Sci. Instrum. **81**, 023101 (2010)
15. T. Hofmann, A. Boosalis, P. Kühne, C.M. Herzinger, J.A. Woollam, D.K. Gaskill, J.L. Tedesco, M. Schubert, Appl. Phys. Lett. **98**, 041906 (2011)
16. T. Hofmann, C.M. Herzinger, J.L. Tedesco, D.K. Gaskill, J.A. Woollam, M. Schubert, Thin Solid Films **519**, 2593 (2011)
17. S. Schöche, J. Shi, A. Boosalis, P. Kühne, C.M. Herzinger, J.A. Woollam, W.J. Schaff, L.F. Eastman, M. Schubert, T. Hofmann, Appl. Phys. Lett. **98**, 092103 (2011)
18. T. Nagashima, M. Hangyo, Appl. Phys. Lett. **79**, 3917 (2001)
19. Y. Ino, R. Shimano, Y. Svirko, M. Kuwata-Gonokami, Phys. Rev. B **70**, 155101 (2004)
20. T. Hofmann, U. Schade, W. Eberhardt, C.M. Herzinger, P. Esquinazi, M. Schubert, Rev. Sci. Instrum. **77**, 63902 (2006)
21. T. Hofmann, C.M. Herzinger, T.E. Tiwald, J.A. Woollam, M. Schubert, Appl. Phys. Lett. **95**, 032102 (2009)

22. A. Boosalis, T. Hofmann, J. Sik, M. Schubert, Thin Solid Films **519**, 2604 (2011)
23. I.J. Hodgkinson, Q.H. Wu, *Birefringent Thin Films and Polarizing Elements* (World Scientific, Singapore, 1998)
24. M.M. Hawkeye, M.J. Brett, J. Vac. Sci. Technol. A **25**, 1317 (2007)
25. C. Buzea, K. Kaminska, G. Beydaghyan, T. Brown, C. Elliott, C. Dean, K. Robbie, J. Vac. Sci. Technol. B **23**, 2545 (2005)
26. C. Herzinger, S. Green, B. Johs, Odd bounce image rotation system in ellipsometer systems, U.S. Patent No. 6,795,184 B1 (2004)
27. B. Dick, M.J. Brett, T. Smy, J. Vac. Sci. Technol. B **21**, 23 (2003)
28. E. Schubert, F. Frost, H. Neumann, B. Rauschenbach, B. Fuhrmann, F. Heyroth, J. Rivory, E. Charron, B. Gallas, M. Schubert, Adv. Solid State Phys. **46**, 309 (2007)
29. F. Tang, T. Karabacak, L. Li, M. Pelliccione, G.-C. Wang, T.-M. Lu, J. Vac. Sci. Technol. A **25**, 160 (2007)
30. D. Vick, Y. Tsui, M. Brett, R. Fedosejevs, Thin Solid Films **350**, 49 (1999)
31. A. Lakhtakia, R. Messier, *Sculptured Thin Films: Nanoengineered Morphology and Optics* (SPIE Press, Bellingham, 2005)
32. K. Robbie, J.C. Sit, M.J. Brett, J. Vac. Sci. Technol. B **16**, 1115 (1998)
33. K. Robbie, G. Beydaghyan, T. Brown, C. Dean, J. Adams, C. Buzea, Rev. Sci. Instrum. **75**, 1089 (2004)
34. K. Robbie, M.J. Brett, A. Lakhtakia, Nature **384**, 616 (1996)
35. D. Schmidt, *Generalized Ellipsometry on Sculptured Thin Films made by Glancing Angle Deposition*, Dissertation, University of Nebraska-Lincoln (2012)
36. G.E. Jellison, Thin Solid Films **313–314**, 33 (1998)
37. D. Bruggeman, Ann. Phys. **24**, 636 (1935)
38. D. Aspnes, Thin Solid Films **89**, 249 (1982)
39. G. Smith, Opt. Commun. **71**, 279 (1989)
40. A. Sihvola, *Electromagnetic Mixing Formulas and Applications* (The Institution of Electrical Engineers, London, 1999)
41. B.C. Bergner, T.A. Germer, T.J. Suleski, J. Opt. Soc. Am. A **27**, 1083 (2010)
42. C. Pidgeon, in *Handbook on Semiconductors*, ed. by M. Balkanski (North-Holland, Amsterdam, 1980)
43. M. Schubert, *Infrared Ellipsometry on Semiconductor Layer Structures: Phonons, Plasmons and Polaritons*. Springer Tracts in Modern Physics, vol. 209 (Springer, Berlin, 2004)
44. J.T. Kindt, C.A. Schmuttenmaer, J. Phys. Chem. **100**, 10373 (1996)
45. J. Barthel, R. Buchner, Pure Appl. Chem. **63**, 1472 (1991)
46. C.T. Kirk, Phys. Rev. B **38**, 1255 (1988)
47. M. van Exter, D. Grischkowsky, Appl. Phys. Lett. **56**, 1694 (1990)
48. D. Schmidt, A.C. Kjerstad, T. Hofmann, R. Skomski, E. Schubert, M. Schubert, J. Appl. Phys. **105**, 113508 (2009)
49. D. Schmidt, B. Booso, T. Hofmann, E. Schubert, A. Sarangan, M. Schubert, Appl. Phys. Lett. **94**, 011914 (2009)
50. D. Schmidt, B. Booso, T. Hofmann, E. Schubert, A. Sarangan, M. Schubert, Opt. Lett. **34**, 992 (2009)
51. P. Pureur, J.V. Kunzler, W.H. Schreiner, D.E. Brando, Phys. Stat. Sol. (a) **70**, 11 (1982)
52. T. Hofmann, D. Schmidt, A. Boosalis, P. Kühne, C.M. Herzinger, J.A. Woollam, E. Schubert, M. Schubert, Mat. Res. Soc. Symp. Proc. **1409**, cc13 (2012)

Chapter 12
Infrared Ellipsometric Investigations of Free Carriers and Lattice Vibrations in Superconducting Cuprates

Jiří Chaloupka, Dominik Munzar and Josef Humlíček

Abstract Since the discovery of superconductivity in copper oxides by Bednorz and Müller [1] in 1986, these compounds attracted unprecedented interest of solid state researchers, both experimentalists and theoreticians. The understanding of cuprates at the microscopic level remains one of central problems of the present solid state physics [2, 3]. Among the key issues, the origin of the superconductivity and the unusual pseudogap phase are most intensely debated. In this chapter we discuss the spectral ellipsometry as a very powerful experimental technique used to gain insight into the peculiar nature of the cuprates. Since the energy scales of the superconducting phenomena belong to the far-infrared range, the corresponding experimental setup is rather different to that commonly used in the visible and ultraviolet spectral range.

12.1 Superconducting Cuprates

The phenomenon of superconductivity was discovered by Heike Kamerlingh-Onnes in 1911 by measuring the dc resistivity of mercury at low temperatures [4]. Using liquid helium to cool down the sample, he observed a sudden drop of the resistivity at 4.2 K. In the following years superconductivity was observed in many metals and alloys. The first phenomenological theory of superconductivity was established by London brothers in 1935 [5, 6]. The next important step represents the work by Ginzburg and Landau [7] providing its description in terms of a complex order parameter ψ. A microscopic understanding of the phenomenon was missing for

J. Chaloupka (✉) · D. Munzar · J. Humlíček
Central European Institute of Technology, Masaryk University,
Kamenice 735, 62500 Brno, Czech Republic
e-mail: chaloupka@physics.muni.cz

J. Chaloupka · D. Munzar · J. Humlíček
Department of Condensed Matter Physics, Faculty of Science,
Masaryk University, Kotlářská 2, 61137 Brno, Czech Republic
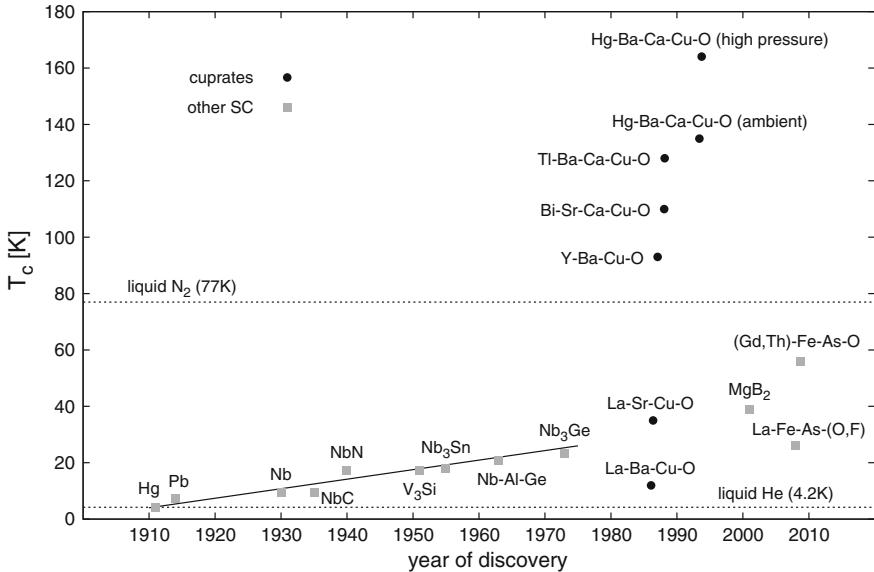
Fig. 12.1 Time evolution of the highest superconducting transition temperature achieved since the discovery of the superconductivity in 1911. The *dashed lines* indicate two temperatures of practical interest—the boiling temperature of liquid helium and that of liquid nitrogen

almost fifty years after the discovery until the famous theory by Bardeen, Cooper and Schrieffer was developed in 1957 [8].

One of the most important parameters of a superconductor is its transition temperature T_c. Below this temperature, the superconductor is characterized by a dissipationless dc transport due to the absence of excitations below a certain energy gap, an ideal diamagnetism and other unusual phenomena. The history of the transition temperature records is captured by Fig. 12.1. Until eighties, the transition temperature has grown roughly linearly reaching about 30 K. The materials include metallic elements and intermetallic compounds like Nb_3Sn, widely used nowadays in high magnetic field applications. The linear trend seemed robust for a large part of the last century, giving only a little hope to reach high temperatures in a reasonable time. A great surprise came in 1986, when Bednorz and Müller found superconductivity with T_c of about 30 K in an oxide system $La_{5-x}Ba_xCu_5O_{5(3-y)}$ [1]. Shortly thereafter the low temperature properties of this family of copper oxides were intensively studied and superconductivity was found in many of them, sometimes with T_c well above the liquid nitrogen boiling point. The highest T_c under ambient pressure observed so far is 138 K in $HgBa_2Ca_2Cu_3O_{8+\delta}$ [9]. Unfortunately, the initially steep increase in T_c saturated quite quickly so that achieving room temperature superconductivity still remains one of the greatest challenges of solid state physics. Later many other so-called unconventional superconductors were found, usually having a small T_c. A notable exception is the recent discovery of superconductivity with T_c up to 55 K in iron pnictides [10–14].
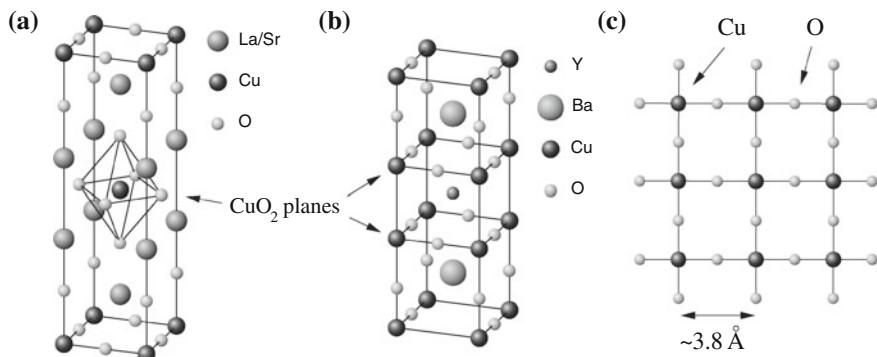
Fig. 12.2 Crystal structures of (**a**) $La_{2-x}Sr_xCuO_4$ and (**b**) $YBa_2Cu_3O_{7-\delta}$ cuprate superconductors. (**c**) Single CuO_2 plane of the cuprates formed by Cu ions arranged in an approximately planar square lattice and connected via oxygen ions. The lattice parameter of the square lattice is determined by the size of Cu and O ions and varies only slightly among the cuprates

In the following we concentrate on the basic properties of the superconducting cuprates, usually labeled, due to their high transition temperature, as high temperature superconducting cuprates (HTSC) [2, 3].

The superconducting cuprates possess rather complex layered crystal structures related to the cubic perovskite structure of, e.g., $BaTiO_3$. Two examples—$La_{2-x}Sr_xCuO_4$ and $YBa_2Cu_3O_{7-\delta}$—are presented in Fig. 12.2a, b. The common element of these structures is the CuO_2 plane depicted in Fig. 12.2c. It is formed by an approximately planar square lattice of copper ions bound together via oxygen ions. The lattice parameter of the Cu sublattice is determined by the length of the Cu-O bonds and equals about 3.8 Å. The direction perpendicular to the CuO_2 planes is commonly labeled as the c-axis direction while the planes themselves are usually referred to as the ab-planes, following the crystallographic convention for tetragonal/orthorhombic crystal structures. One of the important parameters of the cuprate superconductors is the number of CuO_2 planes per unit cell. $La_{2-x}Sr_xCuO_4$ shown in Fig. 12.2a is a single-layer cuprate while $YBa_2Cu_3O_{7-\delta}$ belongs to the so-called bilayer cuprates, where pairs of closely spaced CuO_2 planes enclosing the intrabilayer regions are separated by the thicker interbilayer regions. Series of several adjacent CuO_2 planes within a unit cell are realized in Bi-, Tl-, and Hg-based cuprate families. It is found that T_c first increases with the number of the planes per unit cell, achieving a maximum value for three planes, this is followed by a decrease. The number of the CuO_2 planes per unit cell and the composition of the spacing layers determine the c-axis lattice parameter which varies greatly among the cuprates. For the two materials presented in Fig. 12.2a, b it is equal to $c = 13.2$ Å ($La_{2-x}Sr_xCuO_4$, the distance between the planes being $c/2$) and $c = 11.7$ Å ($YBa_2Cu_3O_{7-\delta}$), respectively. In contrast, $Bi_2Sr_2CaCu_2O_{8+x}$, another often studied cuprate superconductor, has a bilayer structure with a much larger value of the c-axis lattice parameter $c = 30.9$ Å (the distance between the bilayer blocks being $c/2$).
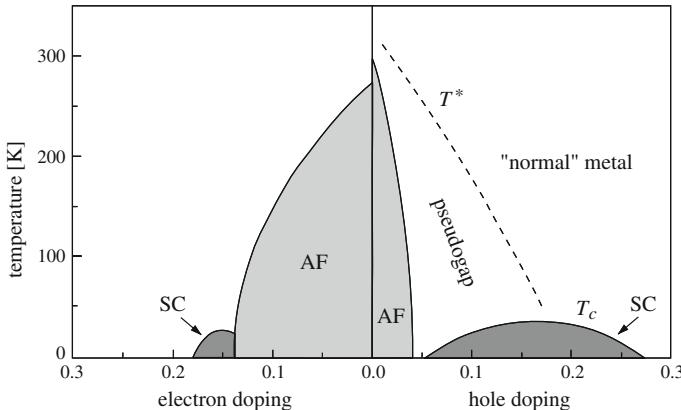
Fig. 12.3 Schematic phase diagram of electron-doped (*left*) and hole-doped cuprates (*right*). The scales correspond to $Nd_{2-x}Ce_xCuO_4$ and $La_{2-x}Sr_xCuO_4$. Adapted from Ref. [15], Copyright (2003) by the American Physical Society

The essential physics of the superconducting cuprates is believed to be hosted by the CuO_2 planes. The crucial parameter here is the number of charge carriers in the CuO_2 planes. This number is controlled either by changing the ratio of different-valence ions of the material as, e.g., the ratio La^{3+}/Sr^{2+} in $La_{2-x}Sr_xCuO_4$, or by varying the oxygen content as, e.g., in $YBa_2Cu_3O_{7-\delta}$. In the majority of cuprates this so-called doping introduces holes into the CuO_2 planes. A representative of the opposite category of electron-doped cuprates is, e.g., $Nd_{2-x}Ce_xCuO_4$.

As can be observed in the schematic phase diagram in Fig. 12.3, the behavior of cuprates is strongly doping-dependent. The undoped cuprates, where the formal valence of Cu ions in the CuO_2 planes is 2+, are antiferromagnets below the Néel temperature of several hundred kelvins. With the introduction of mobile holes, which eventually destroy the antiferromagnetic order, the cuprates switch to the highly debated pseudogap phase and further (at temperatures above T_c) to a normal state phase, which, however, still shows a fairly anomalous behavior. On the very right side of the phase diagram the usual Fermi liquid behavior is observed. The super-conducting phase of $d_{x^2-y^2}$-wave symmetry of the order parameter is found in the doping range of 0.05–0.25 holes per Cu site in the CuO_2 planes, with maximum T_c at the doping level of about 0.15 holes per Cu site. As already presented in Fig. 12.1 the maximum transition temperature is strongly material dependent and often much higher than that of conventional superconductors. For example, the widely known superconductor $YBa_2Cu_3O_{6.95}$ has the transition temperature of $T_c = 92$ K and the magnitude of the gap separating the ground state from the excited states is about $\Delta_{max} = 30$ meV at $T \ll T_c$. The superconducting dome in the phase diagram usually serves as a reference when characterizing the doping of a particular sample. One distinguishes three doping regimes: underdoped, optimally doped (with the high-est T_c) and overdoped. The most mysterious is the underdoped regime where the superconducting phase is preceded by the pseudogap phase when cooling down the
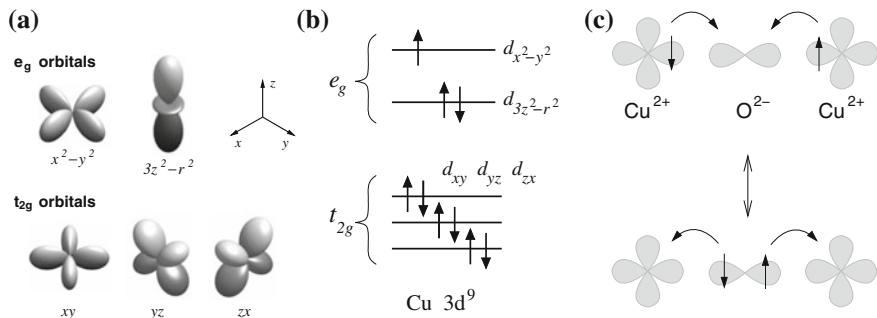
Fig. 12.4   a Shapes of $3d$ orbitals. **b** Level splitting in a crystal field of cubic symmetry and occupied orbitals in the $3d^9$ electronic configuration of Cu^{2+}. **c** Cartoon representation of one of the possible superexchange processes on a Cu–O–Cu bond. The holes residing originally at the copper ions in $3d^9$ configuration meet in the bond-oriented p orbital of oxygen and return back. Since the process is allowed only for antiparallel spins of the holes, it leads to an antiferromagnetic interaction

material. Although the excitations seem to be gapped in this phase, similarly as in case of a superconductor, the phase does not exhibit long-range coherence and superconductivity. Several microscopic scenarios have been proposed for this pseudogap state [16], but its origin is still to be understood.

Let us conclude this paragraph by briefly outlining the microscopic mechanisms particularly important for cuprates. In contrast to the conventional superconductors, where the mechanism of superconductivity involves phonons as the mediators of the pairing interaction among electrons, the dominant part of cuprate physics originates in the electron subsystem itself. The interactions with lattice vibrations seem to play a minor role [3]. The relevant electronic states close to the Fermi level are derived from the valence $3d$ states of Cu ions and partially also from the p states of O ions. The energy levels corresponding to the five $3d$ orbitals of Cu (see Fig. 12.4a) split due to the crystal field generated by the negative charges of the neighboring oxygens. A schematic picture of the $3d^9$ valence shell of Cu^{2+} ion is presented in Fig. 12.4b. The t_{2g} orbitals and e_g orbital of $3z^2 - r^2$ symmetry are completely filled, while the planar e_g orbital of $x^2 - y^2$ symmetry contains a single electron (or, equivalently, a single hole), which can to some extent move to neighboring ions. Therefore, mainly the $3d_{x^2-y^2}$ orbital participates in the states near the Fermi level.

Based on the band structure, the cuprate superconductors should be metallic in the whole doping range. However, in reality, the undoped cuprates are strongly correlated insulators of charge-transfer type [17], where a large energy penalty prevents the charge transfer processes Cu^{2+}–$O^{2-} \rightarrow Cu^{1+}$–$O^{1-}$, i.e., the motion of the holes. Since the presence of two holes at the same Cu site is disfavored by a strong Coulombic repulsion, the Cu ions in the undoped CuO_2 plane retain the $3d^9$ configuration and the holes move to neighboring sites only virtually like in the superexchange process depicted in Fig. 12.4c. Such processes lead to antiferromagnetic interactions among the spins of the Cu ions.

The doping introduces additional holes, and the insulating state with localized holes eventually melts down. The doped carriers can travel through the lattice of

Cu ions, but their motion and the interactions among them are strongly affected by the inherent antiferromagnetic nature of the system. The theoretical approaches to treat this difficult situation can be roughly divided into two classes, according to the physical limit they start with. One possibility is to start with the antiferromagnetic Mott insulator limit. This is the case of the theory developed by Anderson and collaborators [18, 19] and coined RVB theory. Approaches of the second class start with a metallic system, corresponding to highly doped cuprates, and reflect the antiferromagnetic nature by, e.g., coupling of the holes to antiferromagnetic spin waves (for a review see [20, 21]). Despite numerous successes on both sides, the ultimate theory of HTSC—in the sense of the BCS theory for conventional superconductors and its extensions—has not been formulated yet.

12.2 Experimental Setup of the Far-Infrared Spectral Ellipsometry

Optical characteristics of materials, usually in the form of the frequency-dependent dielectric function or the optical conductivity, are a valuable source of information about the electronic structure and elementary excitations such as lattice vibrations. During an optical measurement, the system is probed by photons of various energies. Whenever the energy is such that the photons induce a symmetry allowed electronic transition, resonate with a particular lattice vibration or create an excitation of other kind, a characteristic absorption feature appears in the spectra. The wide range of experimentally available photon energies which extends over several orders of magnitude enables us to explore a large number of microscopic phenomena. The energy scale of the pairing which leads to superconductivity belongs to the far infrared spectral range; for instance, the energy gap 2Δ of about 60 meV in some of the cuprates corresponds to about $500\,\mathrm{cm}^{-1}$. Therefore, it may be expected that the most relevant information to the microscopic mechanisms of cuprates from optical measurements is obtained in this spectral range.

The conventional experimental technique consists in the measurement of the normal-incidence reflectivity followed by Kramers-Kronig analysis to obtain the dielectric function or, equivalently, the complex conductivity. The analysis is necessary for absorbing materials since the information about the phase of the reflected light is lost. For the Kramers-Kronig analysis the reflectivity has to be known on a wide spectral range and extrapolated toward zero and infinite frequency.

The analysis of the reflectivity is complicated for samples which have high reflectivity (close to 1) such as the superconducting cuprates. In addition, problems with the normalization—the determination of the absolute intensity—reduce the accuracy of the dielectric function obtained. In such cases it is very convenient to employ the technique of spectral ellipsometry. This technique is able to measure almost directly the complex dielectric function with high accuracy and reproducibility of the results. For each frequency it gives two ellipsometric angles simply related to the complex dielectric function of the sample. This is in contrast to the reflectivity measurement
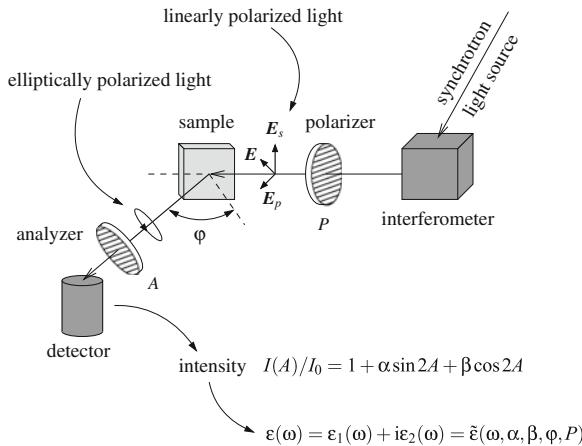
Fig. 12.5 Schematic sketch of a far-infrared ellipsometric setup. The intense light from the synchrotron passing through the interferometer is linearly polarized by the polarizer and after the reflection at the sample surface, its polarization state is analyzed. For simplicity we have omitted some of the elements present in a real setup. The complex dielectric function is determined using the measured values of α and β characterizing the oscillations of the detected intensity as function of the analyzer angle A and known values of the angle of incidence φ and the polarizer angle P

where only a single number for each frequency is obtained. In addition, ellipsometry is a self normalizing technique so that it does not suffer from the slow temporal variations of the intensity of the light source, which are difficult to avoid.

The high-quality HTSC samples suitable for optical measurements are typically flux-grown single crystals with a small size of a few millimeters. Combined with the lack of intense conventional sources of the far-infrared radiation, this makes the optical measurements on cuprates particularly challenging. Therefore the measurements are best performed using synchrotron light sources [22]. Their brilliance is several orders of magnitude higher than that of the conventional infrared light sources such as glowbars and the spectral range extends to the far infrared region. The synchrotron radiation emitted by electrons moving at ultra-relativistic velocities is concentrated into a narrow cone in the direction of the electron motion and is thus naturally collimated. This is another advantage, particularly for an ellipsometric measurement, which requires the angle of incidence to be well-defined.

A simple ellipsometric setup utilizing a synchrotron source is sketched in Fig. 12.5. Synchrotron beam passes through an interferometer which provides the spectral resolution in the experiment. Because of its importance in the far-infrared spectroscopy, and in particular because of its relative complexity compared to the use of a monochromator, we give later in this paragraph a detailed discussion of the interferometer-based approach. The light from the interferometer is linearly polarized using a polarizer. The polarizers need to be of a good quality with a high degree of polarization, having rather large and homogeneous area. In the far-infrared range, wire grid polarizers, free standing or evaporated onto thin transparent substrates,

may be used. The degree of polarization as high as 99.9 % may be achieved. The polarizer angle P is such that the polarized electric field has non-zero components perpendicular to the plane of incidence (s-component, $\boldsymbol{E}_s$) and in the plane of incidence (p-component, $\boldsymbol{E}_p$). In the case of a non-depolarizing sample, the reflection at the sample surface may be described by two complex Fresnel coefficients, r_s and r_p for the s- and p-components of the electromagnetic wave, respectively. The reflected light becomes, in general, elliptically polarized, and it is the analysis of this ellipse that gave the technique its name. The polarization state of the reflected beam is analyzed using a rotating analyzer by varying the analyzer angle A. In the simple case shown, the intensity turns out to have the form

$$I(A) = I_0(1 + \alpha \sin 2A + \beta \cos 2A). \tag{12.1}$$

The coefficients α, β and the constant component I_0 may be obtained by Fourier analysis of the intensity $I(A)$ recorded by the analyzer. They determine the complex ratio of the Fresnel coefficients r_p and r_s,

$$\rho = \frac{r_p}{r_s} = \tan \Psi \mathrm{e}^{i\Delta}, \tag{12.2}$$

which is the fundamental result of the ellipsometric measurement. For an ideal semi-infinite sample, the dielectric function can be reconstructed out of this quantity using the formula

$$\varepsilon(\omega) = \sin 2\varphi \left[1 + \left(\frac{1-\rho}{1+\rho} \right) \tan 2\varphi \right], \tag{12.3}$$

where φ is the angle of incidence indicated in Fig. 12.5. In a real situation of a measurement on a HTSC, the analysis is more complex. For these strongly absorbing samples the ellipsometric determination of the dielectric function is most precise in the grazing incidence geometry with the angle of incidence around $\varphi \approx 85°$. This not only brings enhanced surface sensitivity (requiring to take a special care of the surface quality) but also polarization dependent diffraction effects which become very pronounced in the far-infrared range even for millimeter-sized samples [23]. Next, the HTSC samples are highly anisotropic—the conductivity in the ab-plane being much larger than that along the c-axis. Even in the most isotropic samples like YBa$_2$Cu$_3$O$_{7-\delta}$ the difference is an order of magnitude. Therefore, the c-axis data can suffer from the leakage of the ab-plane response and vice versa and a great care has to be taken to separate the two contributions.

Let us return back to the problem of obtaining the spectral resolution using an interferometer device. As already mentioned, one of the major problems of the far-infrared ellipsometry is the small light intensity. The use of dispersive elements, such as usual diffraction-grating based monochromators, would lead to a large decrease of the intensity and an alternative method of spectral decomposition is thus required. The problem can be solved in the same way as in the Fourier transform infrared spectroscopy (FTIR), i.e., by using interference phenomena to obtain the information
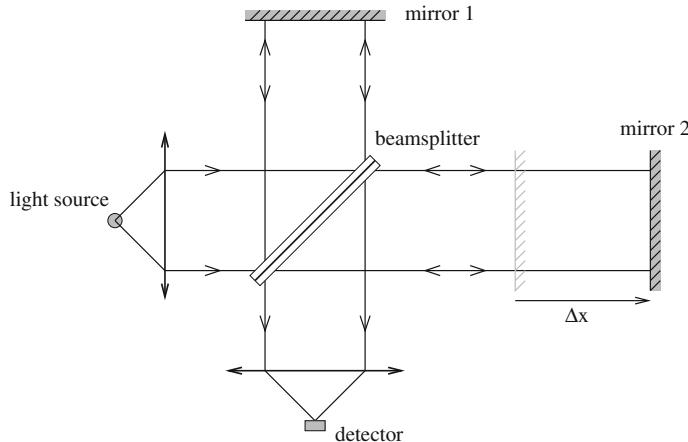
Fig. 12.6 Schematic sketch of an interferometer of Michelson type with a beamsplitter (a semi-transparent mirror). The light from the source passes to the detector either through reflection at the fixed mirror 1 or at the moving mirror 2

about the spectral distribution of the measured light [24]. The interferometric method uses the whole intensity of the measured beam and is thus very convenient for infrared applications. The advantage comes at the expense of high-frequency data sampling and their successive computer processing. Thanks to the fast development of computers in the last decades such tasks are feasible now and FTIR spectrometers are widely used.

In the rest of this paragraph, we give an elementary discussion of the principle of Fourier spectroscopy. For the sake of simplicity, we will analyze a setup depicted in Fig. 12.6 based on a Michelson interferometer. More sophisticated variants of this device can be used in actual FTIR spectrometers.

Let us first assume, that the source emits coherent monochromatic light described by the electric field $E(r, t) = E_0 \exp[i(k \cdot r - \Omega t)]$, where $k = 2\pi\nu$ and ν is the wavenumber of the light. This light is incident on the beamsplitter, where the beam is partially reflected onto the mirror 1 with a fixed position, and partially transmitted to the moving mirror 2. After reflections at the respective mirrors the light is transmitted/reflected to the detector where the interference determined by the optical path difference $2\Delta x$ occurs. Both interfering beams underwent two equivalent reflections and one transmission, so that they should have the same intensity. Neglecting the loss of intensity, the electric field in front of the detector can be expressed as

$$E(\Delta x, t) = \frac{1}{2}[E_0 \cos \Omega t + E_0 \cos(\Omega t - 4\pi\nu\Delta x)]. \qquad (12.4)$$

The detected intensity is proportional to the time average of the squared electric field, $I = c\varepsilon_0\langle E^2 \rangle$, and depends on the shift Δx of the mirror 2 as

$$I(\Delta x) = c\varepsilon_0 \frac{E_0^2}{4} [1 + \cos(4\pi\nu\Delta x)]. \qquad (12.5)$$

Now we generalize the above result by introducing the spectral density $I_S(\nu)$ of the source. For example, a monochromatic source emitting light with the wavenumber ν_0 is described by the spectral intensity $I_S(\nu) = (c\varepsilon_0 E_0^2/2)\delta(\nu - \nu_0)$. By summing up the spectral contributions, we arrive at the following formula for the intensity at the detector:

$$I_D(\Delta x) = \int_0^\infty I_S(\nu)\frac{1}{2}\left[1 + \cos(4\pi \nu \Delta x)\right]\,d\nu. \qquad (12.6)$$

Note that the maximum intensity I_{Dmax} is obtained with $\Delta x = 0$. We rearrange the above formula by introducing a new quantity $I'(\nu)$ given by

$$I'(\Delta x) = I_D(\Delta x) - \frac{1}{2}I_{Dmax} = \int_0^\infty I_S(\nu)\cos(4\pi \nu \Delta x)\,d\nu. \qquad (12.7)$$

The quantity $I'(\Delta x)$ is usually called the interferogram and equals the cosine Fourier transform of the spectral intensity. At this point, the advantage of the Fourier spectroscopy compared to methods involving a monochromator becomes clear. While the latter methods decompose the analyzed beam into many quasi-monochromatic channels to be measured separately, here all the monochromatic components contribute at the same time and the intensity measured by the detector is therefore comparable to the intensity of the entire incoming beam. This enables us to analyze even very weak signals.

The spectral intensity $I_S(\nu)$ is obtained from the interferogram by means of the cosine transform of $I'(\Delta x)$. A complete reconstruction would require to measure $I'(\Delta x)$ for an infinite range of Δx. The motion of the moving mirror is however limited by the construction of the interferometer. As will be shown below, this limits the resolution of the spectrometer. Assuming, that Δx is bound to the interval $\langle -\Delta x_{max}, +\Delta x_{max}\rangle$, we perform the cosine transform via

$$I'_{CT}(\nu) = \int_{-\Delta x_{max}}^{+\Delta x_{max}} I'(\Delta x)\cos(4\pi \nu \Delta x)\,d\Delta x, \qquad (12.8)$$

which—after insertion of $I'(\Delta x)$ from (12.7)—becomes

$$I'_{CT}(\nu) = \frac{1}{2}\int_0^\infty I_S(\nu')\left[\int_{-\Delta x_{max}}^{+\Delta x_{max}} \cos(4\pi \nu \Delta x)\cos(4\pi \nu' \Delta x)\,d\Delta x\right]d\nu'. \qquad (12.9)$$

The inner integral is easily evaluated, leading to

$$\frac{1}{2}\int_0^\infty I_S(\nu')\left\{\text{sinc}\left[4\pi(\nu - \nu')\Delta x_{max}\right] + \text{sinc}\left[4\pi(\nu + \nu')\Delta x_{max}\right]\right\}\Delta x_{max}\,d\nu'. \qquad (12.10)$$

If the range Δx_{max} of the mirror motion is large enough, we may use the limit

$$\lim_{\Delta x_{\text{max}} \to \infty} \frac{\sin[4\pi(\nu - \nu')\Delta x_{\text{max}}]}{4\pi(\nu - \nu')} = \frac{1}{4}\delta(\nu - \nu'). \qquad (12.11)$$

Then $I'_{\text{CT}}(\nu) = (1/8)I_{\text{S}}(\nu)$ and the spectral intensity of the source is recovered. In the case of Δx_{max} not sufficiently large, $I'_{\text{CT}}(\nu)$ is a convolution of the original spectral intensity $I_{\text{S}}(\nu)$ and the function $(1/2)\Delta x_{\text{max}} \text{sinc}(4\pi\nu\Delta x_{\text{max}})$. The spectral features become broader and the resolution is limited to about $(\Delta x_{\text{max}})^{-1}$. The disturbing side maxima of the sinc function are removed by the so-called apodization, which is a multiplication of the interferogram by a suitable function smoothing out the step of $I'(\Delta x)$ at $\Delta x = \pm\Delta x_{\text{max}}$.

Further complications are related to the discrete sampling of I' on the Δx interval. According to the Nyquist theorem, there is a limit of the highest frequency that can be captured with a discrete sampling and higher frequencies become "aliased" into the measured frequency range. In our case the limiting wavenumber is $1/(2 \times \text{the step in } \Delta x)$. It is thus necessary to perform high frequency sampling of the detector signal while moving the interferometer mirror. This is usually not a serious problem in infrared but becomes difficult for shorter wavelengths.

12.3 Optical Response of Superconducting Cuprates

The experimentally observed electromagnetic response of the cuprates is highly anisotropic. This is a natural consequence of the layered structure consisting of highly conductive CuO_2 planes separated by weakly conducting spacing layers. In case of radiation polarized in the CuO_2 planes, we talk about the in-plane or ab-plane response. Similarly, the response to radiation polarized perpendicularly to the CuO_2 planes is denoted as the out-of-plane or the c-axis response.

The optical response reflects the microscopic processes induced by the electromagnetic field of incident photons. It has the potential to capture the nature of the electronic structure as well as the dynamics of the ions forming the crystal lattice. In this chapter we concentrate on two contributions dominant in the infrared part of the spectra—the response of free charge carriers and that of lattice vibrations (phonons)—and, in particular, on the influence of the transition into the superconducting state on these two contributions.

The understanding of the electronic contribution to the spectra is complicated by the fact that it is a complex quantity reflecting contributions of many different types of scattering events which the electrons participate in. This includes scattering on impurities and excitations such as lattice vibrations or spin waves as well as the mutual interaction among the electrons. As a result, the electronic part of the spectra contains the Drude peak broadened by the scattering and possibly accompanied by further absorption features. The c-axis response is, in addition, influenced by interesting multilayer effects.

Somewhat simpler are manifestations of infrared-active phonons. The absorption by lattice vibrations leads to sharp peaks in the spectra. Their profiles can be mostly well fitted by the Lorentz oscillator model. The phonon line is then characterized by three quantities: the frequency of the vibration, the linewidth indicating the degree of damping, and the oscillator strength as a measure of the coupling of the phonon to the electromagnetic field.

In the following two paragraphs, we discuss the characteristics of the in-plane and out-of-plane optical response of cuprates. Since only selected aspects of this response are covered here, an interested reader should also consult [25] for a detailed review and [26] for a review of the electrodynamics of the strongly correlated materials in general.

12.3.1 In-plane Optical Response

The in-plane optical response reflects mainly the microscopic processes occurring within the CuO_2 planes. The analysis of the superconducting-state profiles of the in-plane optical conductivity suggests [27, 28] that the charge carriers are coupled to spin excitations with the characteristic energy of several tens of millielectronvolts as observed in neutron scattering experiments.

Before dealing with the specifics of cuprates, let us first discuss general aspects of the effect of the superconducting transition on the optical conductivity of charge carriers. A typical situation is sketched in Fig. 12.7. Above the transition temperature T_c, we start with a normal metallic response (Fig. 12.7a). The low energy response of free carriers generates a Drude-like peak broadened by various scattering processes. In the superconducting state the charge carriers undergo pairing and condense into a macroscopically coherent state. The condensate is characterized by an undamped response manifesting itself as a δ-peak at zero frequency in the real part of the optical conductivity. Deep below T_c, the scattering mechanisms can only have an effect above the frequency of $2\Delta/\hbar$ as it costs the energy of 2Δ to break an electron pair. Here Δ is the excitation energy gap characterizing the superconductor. The zero-frequency δ-peak in the real part of the conductivity is accompanied by a component $\sim 1/\Omega$ in the imaginary part, following from the Kramers-Kronig relations. In this context, the electron density is usually divided into the superfluid (condensate) part n_s characterized by the undamped (singular) response σ_{cond} and a normal part n_n providing the regular dissipative response σ_{reg} at finite frequencies. The conductivity can then be written as

$$\sigma(\Omega) = \sigma_{\mathrm{cond}}(\Omega) + \sigma_{\mathrm{reg}}(\Omega) = \frac{\pi n_s e^2}{m^*} \delta(\Omega) - \frac{n_s e^2}{i \Omega m^*} + \sigma_{\mathrm{reg}}(\Omega), \quad (12.12)$$

where m^* is the effective mass of the electrons. For an s-wave superconductor at $T \ll T_c$ (Fig. 12.7b), there is no absorption in the spectral range up to 2Δ, therefore $\mathrm{Re}\,\sigma_{\mathrm{reg}}(\Omega)$ is essentially zero in this range. In the case of a d-wave superconductor
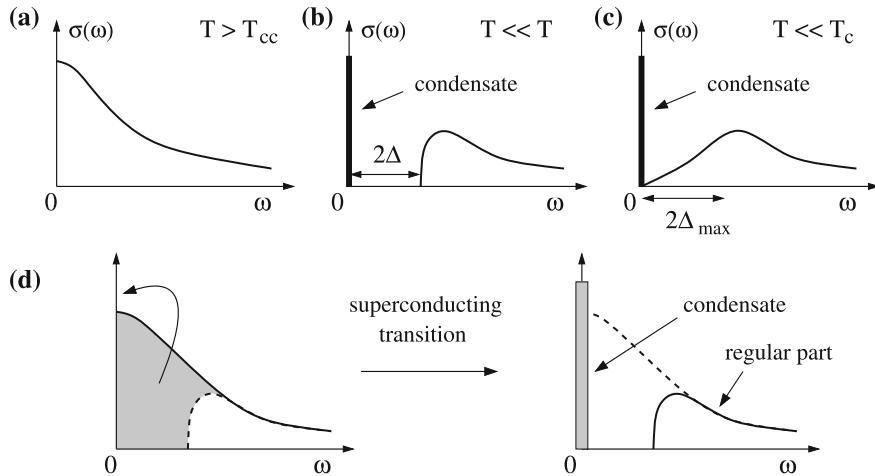
Fig. 12.7 Schematic profiles of the real part of the conductivity of a metallic system in the normal state (**a**), an s-wave superconductor deep below the transition temperature (**b**), and a d-wave superconductor deep below the transition temperature (**c**). In the superconducting state, the conductivity contains the singular condensate contribution. (**d**) Illustration to the Ferrell-Glover-Tinkham sum rule

at $T \ll T_c$, the superconducting gap spans the range $[-\Delta_{\max}, +\Delta_{\max}]$ and the excitations do not have a finite-frequency threshold as in an s-wave superconductor. Nevertheless, the conductivity below $2\Delta_{\max}$ is reduced with a typical profile shown in Fig. 12.7c. In both cases, if the temperature is elevated to a value comparable to T_c, there will be a significant number of thermally excited charge carriers leading to a restoration of the Drude peak in σ_{reg}.

Since the zero-frequency δ-peak is not directly accessible, its spectral weight, which is proportional to n_s, has to be obtained by analyzing the imaginary part of the conductivity or by measuring the magnetic penetration depth that is also determined by n_s. A practical tool is also provided by the so-called sum rules [29] obeyed by the electronic contribution to the conductivity. For example, the partial sum rule for the normal state states that

$$\int_0^W \mathrm{Re}\,\sigma(\Omega)\,\mathrm{d}\Omega = \frac{\pi n e^2}{2m^*}. \tag{12.13}$$

The integration is performed up to the electronic bandwidth W and n, m^* are the electronic density and the effective mass of the conduction band, respectively. When integrating up to very high energies reaching the X-ray range (which is rarely possible due to experimental constraints), all the electrons—both valence and core—successively contribute to the electromagnetic response and n, m^* are then the total electronic density and the true electronic mass m_e. The relevant sum rule for the superconducting case is the Ferrell-Glover-Tinkham sum rule [30], which can be used to obtain

the condensate weight at zero frequency by comparing the finite-frequency profiles of the real part of the conductivity above and below T_c. Because of the temperature independence of the expression on the right-hand side of (12.13) with $W \to \infty$ (spectral weight conservation) the condensate weight is equal to the missing spectral weight at finite frequencies:

$$\int_{0^+}^{\infty} [\mathrm{Re}\,\sigma(\Omega, T > T_c) - \mathrm{Re}\,\sigma(\Omega, T < T_c)]\,\mathrm{d}\Omega = \frac{\pi n_s e^2}{2m^*}. \tag{12.14}$$

As illustrated in Fig. 12.7d, the major part of the condensate spectral weight is collected from the region of the dip in the regular part of $\mathrm{Re}\,\sigma$ below 2Δ that appears below T_c. There are also other redistributions of the spectral weight between the zero frequency peak and the high-energy part of the spectra, but mostly they are less apparent [25].

Representative experimental data of the in-plane optical response of optimally doped HTSC are presented in Fig. 12.8. It can be seen that the electronic contribution to the conductivity is fairly high and the phonon peaks less visible. Several
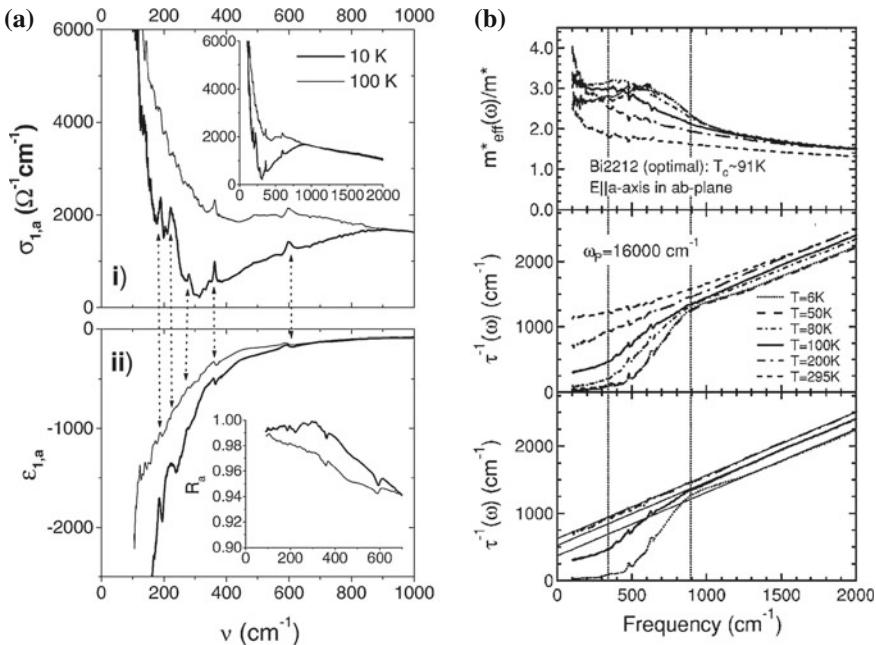
Fig. 12.8 a Real part of the (*i*) conductivity along the *a*-axis and (*ii*) the dielectric function obtained by spectral ellipsometry on optimally doped $YBa_2Cu_3O_{6.95}$ ($T_c = 91.5\,K$) at 100 K (*thin solid line*) and 10 K (*thick solid line*). Reprinted from Ref. [31], Copyright (2002), with permission from Elsevier. **b** Frequency dependent effective mass and scattering rate obtained from the optical data of an optimally doped $Bi_2Sr_2CaCu_2O_{8+x}$. In the *bottom* panel, the high-frequency parts of selected spectra are fitted by lines. Reprinted figure with permission from Tu et al., Ref. [32]. Copyright (2002) by the American Physical Society

peaks marked in Fig. 12.8a correspond to phonons or to more complex collective excitations consisting of lattice vibrations accompanied by a large electronic contribution. A comparison of $T = 100$ K and $T = 10$ K curves in Fig. 12.8a shows the depression of the real part of the conductivity associated with the opening of the superconducting gap. The temperature evolution is consistent with the picture of a d-wave superconductor as sketched in Fig. 12.7c. Some details of the interpretation will be given further in this paragraph. The inset of the bottom panel of Fig. 12.8a shows the reflectivity calculated using the ellipsometric data. The corresponding values are very close to 1, which clearly demonstrates the difficulties of the determination of the dielectric function using reflectivity measurements and the advantages of ellipsometric measurements, which enable to obtain reliably even fine details of the spectra, if diffraction effects can be eliminated.

The optical response is often analyzed in terms of two frequency dependent quantities of the extended Drude model [33]—the scattering rate (inverse lifetime) $\tau^{-1}(\Omega)$ and the effective mass $m^*(\Omega)$. The conductivity is expressed as

$$\sigma(\Omega) = \frac{\varepsilon_0 \Omega_{\mathrm{pl}}^2}{\tau^{-1}(\Omega) - i\Omega m^*(\Omega)/m_e} \ , \qquad (12.15)$$

where Ω_{pl} denotes the plasma frequency. In case of frequency independent τ^{-1} and m^*, (12.15) coincides with the well-known Drude model. Example data of these quantities obtained on optimally doped $Bi_2Sr_2CaCu_2O_{8+x}$ samples are presented in Fig. 12.8b. The interesting quantity here is the scattering rate, which exhibits a totally different behaviour above and below $1000\,\mathrm{cm}^{-1}$. Above $1000\,\mathrm{cm}^{-1}$, it increases linearly with increasing frequency. This trend may be continued to lower frequencies for $T > T_c$ but with the onset of superconductivity, a gap appears in the spectra below $1000\,\mathrm{cm}^{-1}$.

The in-plane data can be interpreted within a model where the charge carriers are coupled to antiferromagnetic spin waves, i.e., they emit and absorb these waves while moving in the crystal. The antiferromagnetic waves, which can be directly observed in neutron experiments, have a characteristic energy of several tens of millielectronvolts. For example, in $YBa_2Cu_3O_{6.92}$, the energy density of the spin waves peaks at about $E_{\mathrm{mode}} \approx 40$ meV [34]. For a detailed discussion of the in-plane infrared spectra see [35, 36] and references therein. Here we only mention the main results. The model predicts the following features: (i) A gradual onset of the real part of the conductivity and $1/\tau$ around E_{mode}; (ii) A knee feature in the spectra of $1/\tau$ around $E_{\mathrm{mode}} + \Delta_{\mathrm{max}}$ (and the corresponding maximum in the second derivative of the spectra); (iii) An overshoot of the spectra of $1/\tau$, i.e., $1/\tau(T \ll T_c)$ above $1/\tau(T \approx T_c)$, in the frequency region around $E_{\mathrm{mode}} + 2\Delta_{\mathrm{max}}$. The predicted positions of the spectral features of 320, 560 and $880\,\mathrm{cm}^{-1}$ (obtained using $E_{\mathrm{mode}} = 40$ meV) and $\Delta_{\mathrm{max}} = 30$ meV are in reasonable agreement with the data. Finally, the Drude peak visible in the data originates in those parts of the Brillouin zone near the Fermi surface, where the superconducting energy gap is close to zero.

12.3.2 Optical Response Along the c-Axis

If the probing radiation is polarized along the *c*-axis, a fairly different response
is found than that in the in-plane case. The *c*-axis conductivity is much smaller
than the in-plane conductivity and is strongly material-dependent. In the normal
state, the dc *c*-axis conductivity of the most three-dimensional high-T_c super-
conductor $YBa_2Cu_3O_7$ is about $400\,\Omega^{-1}cm^{-1}$, that of a highly two-dimensional
$Bi_2Sr_2Ca_2CuO_8$ is four orders of magnitude lower [25]. Due to the relatively low
electronic contribution to the conductivity compared to the in-plane response, the
most apparent features of the *c*-axis spectra are sharp phonon peaks. In this para-
graph we concentrate on the interesting case of cuprates with two CuO_2 planes per
unit cell such as $YBa_2Cu_3O_{7-\delta}$, where further complexity results from multilayer
effects.

The basic doping and temperature trends of the *c*-axis conductivity are shown in
Fig. 12.9 displaying the *c*-axis conductivity of $Y_{0.86}Ca_{0.14}Ba_2Cu_3O_{7-\delta}$ samples with
oxygen content given by $\delta = 0.4$ (slightly underdoped), $\delta = 0.3$ (optimally doped),
and $\delta = 0.1$ (strongly overdoped) [37]. By comparing the conductivity spectra for
different dopings and temperatures we can observe the following trends: At low
temperatures, the conductivity in a region below a certain doping dependent energy
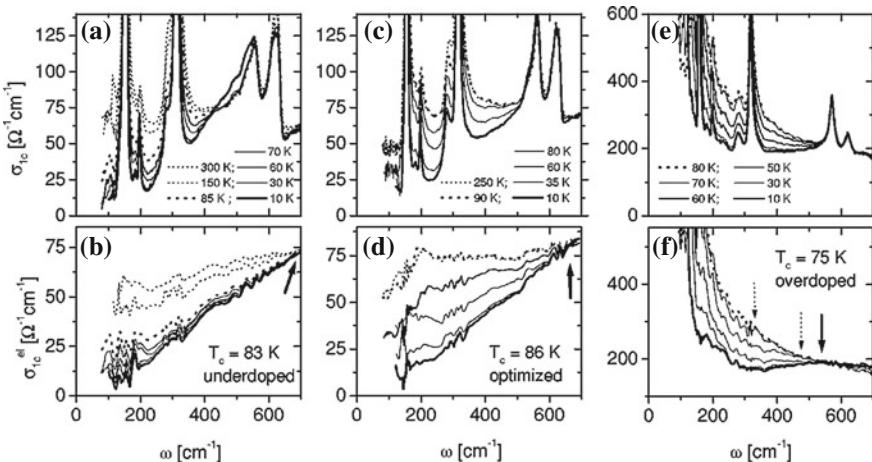
Fig. 12.9 Doping- and temperature-dependent trends observed in the *c*-axis conductivity obtained
by far-infrared ellipsometry measurements on $Y_{0.86}Ca_{0.14}Ba_2Cu_3O_{7-\delta}$ crystals. Data measured on
underdoped (**a**, **b**), optimally doped (**c**, **d**), and overdoped (**e**, **f**) samples are presented. The *upper
panels* show the complete conductivity data, in the *bottom panels*, the phonon peaks are subtracted.
Reprinted figure with permission from Bernhard et al., Ref. [37]. Copyright (1999) by the American
Physical Society

is suppressed. In the overdoped and optimally doped regime, this energy is approximately two times the superconducting gap amplitude and the suppression starts at the transition temperature T_c. In the underdoped regime, the pseudogap phenomenon causes a similar suppression that starts already above T_c. The superconducting transition and the emergence of the superconducting gap leads to an additional decrease of the conductivity. Note also, that the normal state conductivity profile and the dc-conductivity at the highest temperatures are strongly doping dependent. In the spectral region shown, the normal state conductivity is an increasing, approximately constant and decreasing function of energy for the underdoped, optimally doped and overdoped case, respectively.

A closer look at the temperature dependence of the c-axis spectra reveals two interesting spectral features appearing in the superconducting state. The first one is localized between the phonon peaks at 320 and 560 cm^{-1}. It is fairly pronounced in strongly underdoped samples, in the moderately underdoped sample of Fig. 12.9a it shows up around 500 cm^{-1}. Originally it was attributed to a phonon, its position and spectral weight, however, are strongly doping dependent so that it has to be of electronic origin. The second feature is a less pronounced mode around 1000 cm^{-1}, best visible for optimally-doped and overdoped samples [38, 39]. An important observation related to these features is that they are specific to cuprates with two or more CuO$_2$ planes per unit cell, see, e.g., [40].

The lower mode is clearly visible around 400 cm^{-1} in Fig. 12.10, where we show the c-axis conductivity spectra of an underdoped YBa$_2$Cu$_3$O$_{6.5}$ sample (two CuO$_2$ planes per unit cell) for a wide range of temperatures. The electronic contribution to the spectra obtained by subtracting that of the phonon modes is presented in Fig. 12.10b. The mode is visible already at temperatures somewhat above T_c though with a small magnitude. It rapidly gains spectral weight upon transition to the superconducting state. The presence of the mode also influences the surrounding phonon peaks leading to the so-called phonon anomalies. The phonon at 320 cm^{-1} is the most susceptible to this influence.

A sophisticated theory explaining the electronic modes on a microscopic level was developed recently [42]. It involves a microscopic model with the electronic coupling between the CuO$_2$ planes. Here we recall an earlier and simpler interpretation of the lower mode around 400 cm^{-1} in terms of the transverse plasmon of an effective medium based model of the superlattice of the CuO$_2$ planes. This phenomenological approach was first introduced by van der Marel and Tsvetkov [43] and later extended by Munzar et al. [44] to explain the observed phonon anomalies. In the limit of weakly coupled CuO$_2$ planes, the phenomenological approach is consistent with the fully microscopic one of [42].

The model is introduced in Fig. 12.11 for the case of YBa$_2$Cu$_3$O$_{7-\delta}$ with two CuO$_2$ planes per unit cell. A generalization of the following considerations to the case of cuprates with more CuO$_2$ planes per unit cell is straightforward. Within this model the CuO$_2$ planes are represented by homogeneously charged planes, separated by intrabilayer (bl) and interbilayer (int) spacing regions. The electric field applied along the c-axis induces current densities j_{bl} and j_{int} flowing in the intrabilayer and interbilayer regions, respectively. The currents carry charge between the CuO$_2$
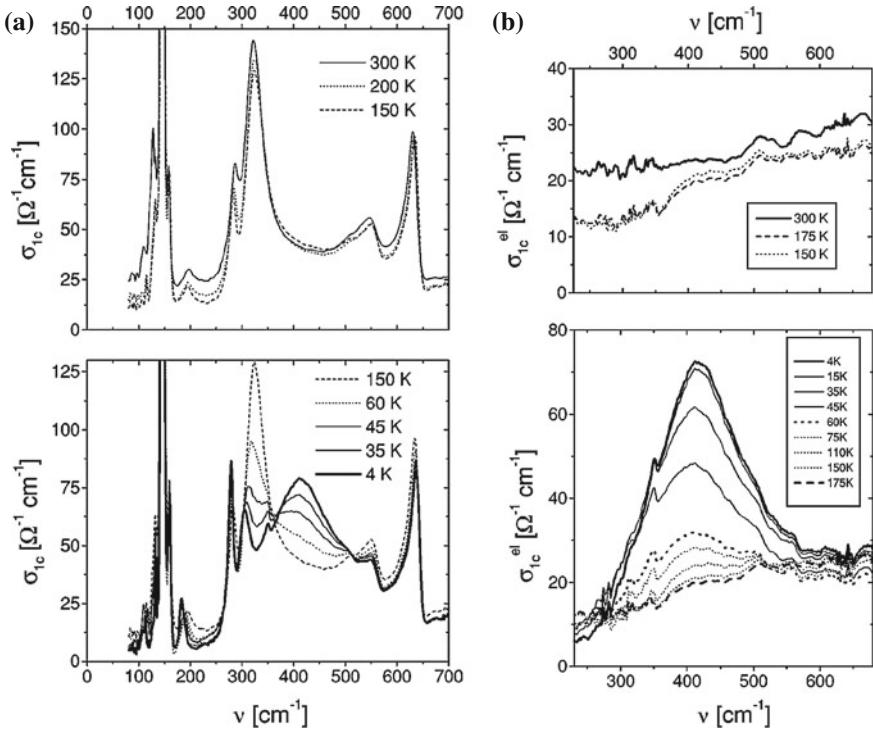
Fig. 12.10 a Temperature dependence of the real part of the *c*-axis conductivity measured on a strongly underdoped YBa$_2$Cu$_3$O$_{6.5}$ crystal with $T_c = 52$ K. **b** Temperature dependence of the electronic contribution obtained by subtracting the contributions of the phonon modes. Reprinted figures with permission from Bernhard et al., Ref. [41]. Copyright (2000) by the American Physical Society

planes. The resulting excess charge densities $\pm\rho$ then modify the external electric field producing the local electric fields E_{bl} and E_{int}. To account for these effects quantitatively and to derive a formula for the macroscopic *c*-axis conductivity, we describe the response of the intrabilayer and interbilayer region by the local conductivities σ_{bl} and σ_{int} such that the current densities are given by $j_{bl} = \sigma_{bl} E_{bl}$ and $j_{int} = \sigma_{int} E_{int}$. At this point we divide the processes contributing to the conductivity according to their characteristic energy. We keep the contribution of the low-energy processes with the characteristic energy below 10^2 meV related to the superconductivity in $\sigma_{bl/int}$ explicitly. Next we assume, that the high-energy interband processes (characteristic energy 10^0–10^1 eV) merely provide an effective medium with the relative permittivity ε_∞. The dielectric function then takes the form

$$\varepsilon = \varepsilon_\infty + \frac{i\sigma}{\varepsilon_0 \Omega} \tag{12.16}$$

for both local and macroscopic response. A realistic value of ε_∞ is about 5.
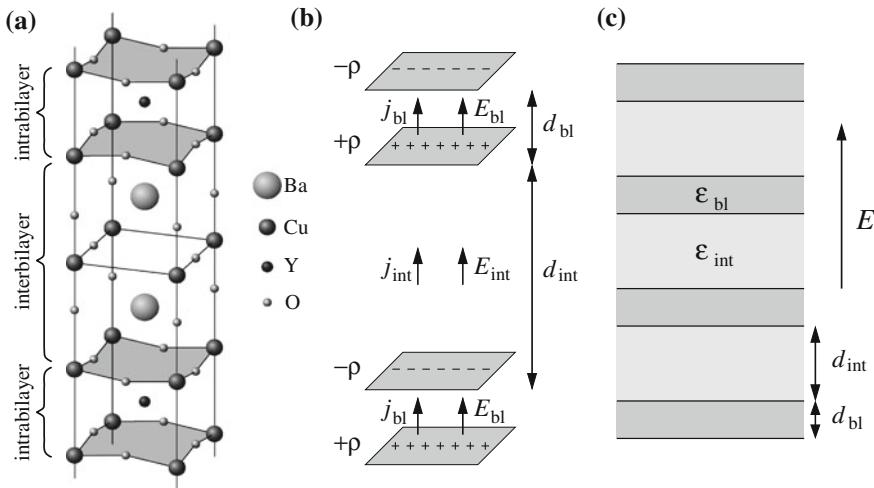
Fig. 12.11 **a** Layered crystal structure of bilayer cuprate YBa$_2$Cu$_3$O$_{7-\delta}$ and **b** its model representation by a system of charged conductive sheets. The structure is divided into the (intra)bilayer regions between the closely spaced CuO$_2$ planes and interbilayer regions between these pairs of the CuO$_2$ planes. The local fields and currents are labeled accordingly. **c** Representation of YBa$_2$Cu$_3$O$_{7-\delta}$ as a multilayered composite consisting of two constituents of thicknesses d_{bl} and d_{int} respectively

The macroscopic c-axis conductivity is defined as the ratio of the macroscopic current density to the macroscopic electric field. Assuming, that the fields and currents in the intrabilayer and interbilayer regions are homogeneous, we get the macroscopic (averaged) current density as $\langle j \rangle = (d_{bl} j_{bl} + d_{int} j_{int})/d$ and the macroscopic electric field as $\langle E \rangle = (d_{bl} E_{bl} + d_{int} E_{int})/d$, where d_{bl} and d_{int} are the thicknesses of the intrabilayer and interbilayer regions respectively, and $d = d_{bl} + d_{int}$ is the total thickness of one period of the multilayer. The resulting formula for the macroscopic conductivity thus reads

$$\sigma = \frac{\langle j \rangle}{\langle E \rangle} = \frac{d_{bl} \sigma_{bl} E_{bl} + d_{int} \sigma_{int} E_{int}}{d_{bl} E_{bl} + d_{int} E_{int}}. \tag{12.17}$$

The local electric fields differ from each other due to the fact that the planes are charged. Considerations based on elementary electrostatics lead us to the relation

$$E_{bl} - E_{int} = \frac{\rho}{\varepsilon_0 \varepsilon_\infty}. \tag{12.18}$$

The Coulomb field due to the charge density ρ is screened by the higher-energy processes providing ε_∞ as discussed above. The charge accumulated on the planes is determined by the currents and obeys the continuity equation

$$\frac{\partial \rho}{\partial t} = j_{\text{int}} - j_{\text{bl}}, \quad \text{in frequency domain} - i\Omega\rho = \sigma_{\text{int}} E_{\text{int}} - \sigma_{\text{bl}} E_{\text{bl}}. \quad (12.19)$$

By eliminating ρ from (12.18) and (12.19) and inserting into (12.17) we arrive at the multilayer formula

$$\frac{d}{\sigma_\infty + \sigma} = \frac{d_{\text{bl}}}{\sigma_\infty + \sigma_{\text{bl}}} + \frac{d_{\text{int}}}{\sigma_\infty + \sigma_{\text{int}}} \quad (12.20)$$

with $\sigma_\infty = -i\Omega\varepsilon_0\varepsilon_\infty$. In terms of the local dielectric functions, the macroscopic dielectric function is given by

$$\frac{d}{\varepsilon} = \frac{d_{\text{bl}}}{\varepsilon_{\text{bl}}} + \frac{d_{\text{int}}}{\varepsilon_{\text{int}}}. \quad (12.21)$$

This formula is equivalent to the result of the effective medium theory for the multilayer depicted in Fig. 12.11c and the electric field applied along the c-axis (for an introduction to the effective medium theory, see e.g. [45]).

We address the implications of the mixing formulas (12.20) or (12.21) for the macroscopic c-axis conductivity of cuprates below T_c by considering, for the sake of simplicity, a superconducting electromagnetic response of the intrabilayer regions and a normal response of the interbilayer regions, $\varepsilon_{\text{int}} = \varepsilon_\infty$.[1] This roughly corresponds to the actual physical situation in cuprates. Explicitly, we take the conductivity of the intrabilayer regions of the pure condensate form $\sigma_{\text{bl}}(\Omega) = i\varepsilon_0\Omega_{\text{bl}}^2/\Omega$, i.e., neglect σ_{reg} in (12.12). The corresponding dielectric function is $\varepsilon_{\text{bl}} = \varepsilon_\infty - \Omega_{\text{bl}}^2/\Omega^2$. It is easy to show, that the resulting macroscopic c-axis conductivity has a resonance (singularity) at the frequency $\Omega_p = \Omega_{\text{bl}}\sqrt{d_{\text{int}}/d\varepsilon_\infty}$, which for realistic values of the parameters gives $\Omega_p \approx 400\,\text{cm}^{-1}$. Physically, the undamped condensate response at zero frequency was shifted to a finite frequency due to charging of the planes resulting in a plasmon-like collective mode. It is usually denoted as a transverse Josephson plasmon, because the plasma oscillations are perpendicular (transverse) to the direction of propagation of the electromagnetic wave and because the multilayer of alternating normal and superconducting regions can be described as a superlattice of Josephson junctions. Note that the simple effective-medium picture describes correctly the stack of very thin layers ($d_{\text{bl}} \approx 0.33\,\text{nm}$, $d_{\text{int}} \approx 0.84\,\text{nm}$).

We conclude this chapter by commenting on the interesting interplay between the collective electronic mode around $400\,\text{cm}^{-1}$ and the lattice vibrations of nearby frequencies, clearly visible in Fig. 12.10a. It is essentially a local-field effect, where the charging of the CuO_2 planes modifies the electric fields acting on the ions and thus influences lattice vibrations. From Eqs. (12.18) and (12.19) we can extract the frequency dependent ratio of the local fields

---

[1] When dealing with the experimental data, one can successfully fit the spectra using the multilayer formula with both σ_{bl} and σ_{int} involving a superconducting condensate contribution (much weaker in σ_{int}) and phonon contributions.
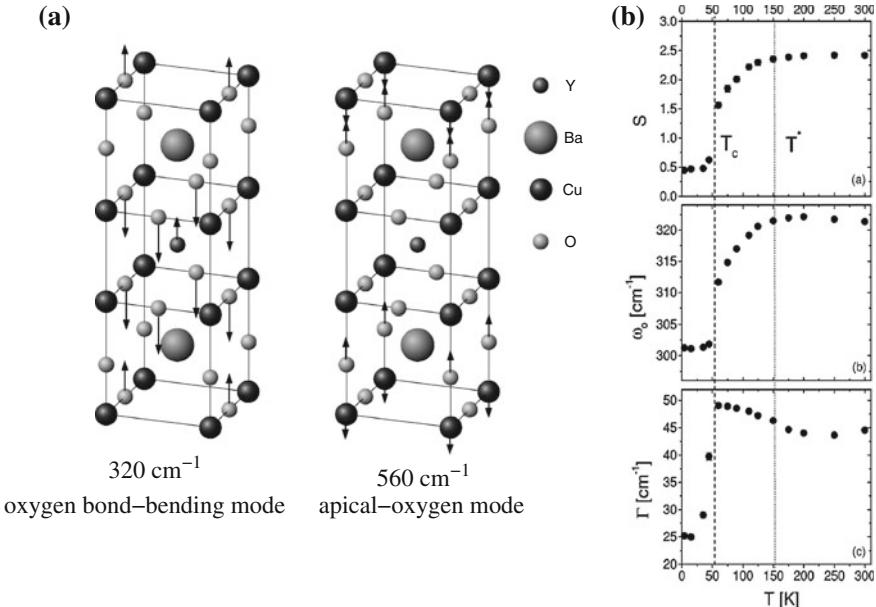
Fig. 12.12 **a** Polarization diagrams for the two phonon modes that exhibit the most pronounced anomalies [46]. The *left* and the *right* diagram correspond to the $320\,\mathrm{cm}^{-1}$ and the $560\,\mathrm{cm}^{-1}$ mode, respectively. **b** Oscillator characteristics of the oxygen bond-bending mode as obtained from the ellipsometric data measured on underdoped $YBa_2Cu_3O_{6.5}$ with $T_c = 52\,\mathrm{K}$. Shown are the oscillator strength S, the eigenfrequency Ω_0 and the half-width Γ. Reprinted figure with permission from Bernhard et al., Ref. [41]. Copyright (2000) by the American Physical Society

$$\frac{E_{\mathrm{bl}}}{E_{\mathrm{int}}} = \frac{\sigma_\infty + \sigma_{\mathrm{int}}}{\sigma_\infty + \sigma_{\mathrm{bl}}} = \frac{\varepsilon_{\mathrm{int}}}{\varepsilon_{\mathrm{bl}}}. \tag{12.22}$$

This ratio is quite different from unity in the multilayer cuprate, where the intrabilayer region becomes superconducting and the interbilayer region exhibits a normal response. In the simplest case we considered above, the ratio is equal to $E_{\mathrm{bl}}/E_{\mathrm{int}} = (1 - \Omega_{\mathrm{bl}}^2/\varepsilon_\infty \Omega^2)^{-1}$ having a singularity at $\Omega_s = \Omega_{\mathrm{bl}}/\sqrt{\varepsilon_\infty}$, i.e., very close to the frequency Ω_p of the plasmon peak, and being negative below Ω_s. From the phonon polarization diagrams in Fig. 12.12a it may be deduced, that the oxygen-bending mode with the frequency of $320\,\mathrm{cm}^{-1}$ is driven mainly by the local field acting on the planar oxygens $[\approx (E_{\mathrm{bl}} + E_{\mathrm{int}})/2]$ while the mode at $560\,\mathrm{cm}^{-1}$ involving vibrations of the ions in the interbilayer region is driven by the interbilayer local field. The changes of the spectral weight of the two phonons below T_c can be interpreted as follows. In the frequency range of the $320\,\mathrm{cm}^{-1}$ phonon, E_{bl} and E_{int} have the same sign in the normal state above T_c ($\Omega_{\mathrm{bl}} = 0$), but opposite signs below T_c (in the presence of Ω_{bl}). As a consequence, the magnitude of the local field driving the phonon decreases upon entering the superconducting state and this leads to the decrease of the spectral weight. The appearance of nonzero Ω_{bl} below T_c similarly

leads to a decrease of the field E_{int} driving the $560\,\mathrm{cm}^{-1}$ phonon and to the corresponding decrease of the spectral weight of the phonon. For a careful discussion and for the model equations see [44, 47].

Several other interesting effects have been discovered in the c-axis conductivity, for example those related to the pseudogap phenomenon and its connection to the superconductivity [39, 48], or the changes induced by strong magnetic fields [49, 50], but they are beyond the scope of the present introductory chapter.

12.4 Conclusion

Infrared ellipsometry proved to be a powerful technique for the studies of numerous peculiar properties of cuprates. In particular, electronic excitations and lattice vibrations in both normal and superconducting state have been examined with an exceptional precision and accuracy. Recent advances in the ellipsometric instrumentation overcome the difficulties encountered in measurements of small single-crystalline samples in the far-infrared range.

References

1. J.G. Bednorz, K.A. Müller, Z. Phys. B **64**, 189 (1986)
2. J.R. Schrieffer (ed.), *Handbook of High-Temperature Superconductivity* (Springer, New York, 2007)
3. N. Plakida, *High-Temperature Cuprate Superconductors* (Springer, Berlin, 2010)
4. H.K. Onnes, Comm. Phys. Lab. Univ. Leiden, Nos. 119, 120, 122 (1911)
5. H. London, F. London, Proc. Roy. Soc. (London) **A149**, 71 (1935)
6. H. London, F. London, Physica **2**, 341 (1935)
7. V.L. Ginzburg, L.D. Landau, Zh Eksp, Teor. Fiz. **20**, 1064 (1950)
8. J. Bardeen, L.N. Cooper, J.R. Schrieffer, Phys. Rev. **108**, 1175 (1957)
9. G.F. Sun, K.W. Wong, B.R. Xu, Y. Xin, D.F. Lu, Phys. Lett. A **192**, 122 (1994)
10. Y. Kamihara, T. Watanabe, M. Hirano, H. Hosono, J. Am. Chem. Soc. **130**, 3296 (2008)
11. H. Takahashi, K. Igawa, K. Arii, Y. Kamihara, M. Hirano, H. Hosono, Nature **453**, 376 (2008)
12. R.H. Liu, G. Wu, T. Wu, D.F. Fang, H. Chen, S.Y. Li, K. Liu, Y.L. Xie, X.F. Wang, R.L. Yang, L. Ding, C. He, D.L. Feng, X.H. Chen, Phys. Rev. Lett. **101**, 087001 (2008)
13. J. Paglione, R.L. Greene, Nat. Phys. **6**, 645 (2010)
14. D.C. Johnston, Adv. Phys. **59**, 803 (2010)
15. A. Damascelli, Z. Hussain, Z.X. Shen, Rev. Mod. Phys. **75**, 473 (2003)
16. M.R. Norman, D. Pines, C. Kallin, Adv. Phys. **54**, 715 (2005)
17. S. Maekawa, T. Tohyama, S.E. Barnes, S. Ishihara, W. Koshibae, G. Khaliullin, *Physics of Transition Metal Oxides* (Springer, Berlin, 2004)
18. P.W. Anderson, Science **235**, 1196 (1987)
19. P.W. Anderson, P.A. Lee, M. Randeira, T.M. Rice, N. Trivedi, F.C. Zhang, J. Phys.: Condens. Matter **16**, R755 (2004)
20. M. Eschrig, Adv. Phys. **55**, 47 (2006)
21. A.V. Chubukov, D. Pines, J. Schmalian, in *Superconductivity: Conventional and Unconventional Superconductors*, vol. 2, ed. by K.H. Bennemann, J.B. Ketterson (Springer, Berlin, 2008)

22. C. Bernhard, J. Humlíček, B. Keimer, Thin Solid Films **455–456**, 143 (2004)
23. J. Humlíček, C. Bernhard, Thin Solid Films **455–456**, 177 (2004)
24. H. Kuzmany, *Solid State Spectroscopy* (Springer, Berlin, 2009)
25. D.N. Basov, T. Timusk, Rev. Mod. Phys. **77**, 721 (2005)
26. D.N. Basov, R.D. Averitt, D. van der Marel, M. Dressel, K. Haule, Rev. Mod. Phys. **83**, 471 (2011)
27. D. Munzar, C. Bernhard, M. Cardona, Physica C **312**, 121 (1999)
28. J.P. Carbotte, E. Schachinger, D.N. Basov, Nature **401**, 354 (1999)
29. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957)
30. M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, New York, 1996)
31. C. Bernhard, T. Holden, J. Humlíček, D. Munzar, A. Golnik, M. Kläser, Th Wolf, L. Carr, C. Homes, B. Keimer, M. Cardona, Solid State Commun. **121**, 93 (2002)
32. J.J. Tu, C.C. Homes, G.D. Gu, D.N. Basov, M. Strongin, Phys. Rev. B **66**, 144514 (2002)
33. A.V. Puchkov, D.N. Basov, T. Timusk, J. Phys.: Condens. Matter **8**, 10049 (1996)
34. J. Rossat-Mignod, L.P. Regnault, C. Vettier, P. Bourges, P. Burlet, J. Bossy, J.Y. Henry, G. Laperot, Physica B **180–181**, 383 (1992)
35. P. Cásek, C. Bernhard, J. Humlíček, D. Munzar, Phys. Rev. B **72**, 134526 (2005)
36. J. Chaloupka, D. Munzar, Phys. Rev. B **76**, 214502 (2007)
37. C. Bernhard, D. Munzar, A. Wittlin, W. König, A. Golnik, C.T. Lin, M. Kläser, Th Wolf, G. Müller-Vogt, M. Cardona, Phys. Rev. B **59**, R6631 (1999)
38. M. Grüninger, D. van der Marel, A.A. Tsvetkov, A. Erb, Phys. Rev. Lett. **84**, 1575 (2000)
39. Yu. Li, D. Munzar, A.V. Boris, P. Yordanov, J. Chaloupka, Th. Wolf, C.T. Lin, B. Keimer, C. Bernhard. Phys. Rev. Lett. **100**, 177004 (2008)
40. S.V. Dordevic, E.J. Singley, J.H. Kim, M.B. Maple, Seiki Komiya, S. Ono, Yoichi Ando, T. Rõõm, Ruxing Liang, D.A. Bonn, W.N. Hardy, J.P. Carbotte, C.C. Homes, M. Strongin, D.N. Basov. Phys. Rev. B **69**, 094511 (2004)
41. C. Bernhard, D. Munzar, A. Golnik, C.T. Lin, A. Wittlin, J. Humlíček, M. Cardona, Phys. Rev. B **61**, 618 (2000)
42. J. Chaloupka, C. Bernhard, D. Munzar, Phys. Rev. B **79**, 184513 (2009)
43. D. van der Marel, A. Tsvetkov, Czech. J. Phys. **46**, 3165 (1996)
44. D. Munzar, C. Bernhard, A. Golnik, J. Humlíček, M. Cardona, Solid State Commun. **112**, 365 (1999)
45. J. Humlíček, this volume
46. R. Henn, T. Strach, E. Schönherr, M. Cardona, Phys. Rev. B **55**, 3285 (1997)
47. A. Dubroka, D. Munzar, Physica C **405**, 133 (2004)
48. A. Dubroka, M. Rössle, K.W. Kim, V.K. Malik, D. Munzar, D.N. Basov, A.A. Schafgans, S.J. Moon, C.T. Lin, D. Haug, V. Hinkov, B. Keimer, Th Wolf, J.G. Storey, J.L. Tallon, C. Bernhard, Phys. Rev. Lett. **106**, 047006 (2011)
49. K.M. Kojima, S. Uchida, Y. Fudamoto, S. Tajima, Phys. Rev. Lett. **89**, 247001 (2002)
50. A.D. LaForge, W.J. Padilla, K.S. Burch, Z.Q. Li, A.A. Schafgans, K. Segawa, Y. Ando, D.N. Basov, Phys. Rev. B **79**, 104516 (2009)

Chapter 13
Real-Time Ellipsometry for Probing Charge-Transfer Processes at the Nanoscale

Maria Losurdo, April S. Brown and Giovanni Bruno

Abstract Nanoscale charge transfer is important to both the frontier of fundamental science and to applications in molecular electronics, photonic, electronic, optical, imaging, catalysis, sensing devices, photovoltaics, and energy savings and storage. For many of those applications, plasmonic metal nanoparticles are coupled with molecules and/or semiconductors, where nanoparticles act as an electron-bridge. Metal nanoparticles experience charge transfer either by a hopping mechanism involving transient charging of the nanoparticle and/or by electron storage and delocalization among/in the nanoparticles. This electron transfer affects the electron density in the metal, and the plasmon resonance, and therefore, can be detected spectroscopically. This chapter discusses examples of exploitation of spectroscopic ellipsometry as a real time research tool that advance description and understanding of charge transfer phenomena involving (i) chemisorption of metals on semiconductor surfaces, (ii) growth of plasmonic nanoparticles on polar semiconductors, (iii) coupling plasmonic nanoparticles to graphene, and (iv) charge transfer between plasmonic nanoparticles and biomolecules, activating sensing processes.

M. Losurdo (✉) · G. Bruno
Institute of Inorganic Methodologies and of Plasmas, IMIP, National Council of Research, CNR, IMIP-CNR, via Orabona 4, 70126 Bari, Italy
e-mail: maria.losurdo@ba.imip.cnr.it

G. Bruno
e-mail: giovanni.bruno@ba.imip.cnr.it

A. S. Brown
Electrical and Computer Engineering Department, Duke University, Durham, NC, USA
e-mail: abrown@ee.duke.edu

13.1 Introduction

Charge transfer mediated by nanoscale materials is important to both the frontier of fundamental science and to applications in molecular electronics, sensors, photonics, electrocatalysis, and light energy conversion.

The building blocks for such photonic/electronic nanodevices are single or small groups of (usually organic) molecules as well as semiconductors and ultrathin metal layers and nanoparticles, because of their high surface areas and other size-dependent properties. All these nanoscale entities have in common the capability of charge storage and transfer.

As an example, for metal nanoparticles passivated by a dielectric organic layer, the particles exhibit a sub-attofarad (aF, 10^{-18} F) molecular capacitance [1, 2]. A review of the fundamentals of nanoscale charge-transfer processes is given in Ref. [3].

In a unifying simple scheme, a semiconductor, a metal nanoparticle and a molecule can all function as a donor, acceptor and/or a bridge for electrons (see scheme in Fig. 13.1).

Plasmonic metal nanoparticles are interesting because of the localized surface plasmon resonance (LSPR) phenomenon, which consists in coupled resonant oscillations of electron density and an evanescent electromagnetic field (collectively known as plasmons) that are excited near the particle surface by the incident light of specific wavelengths. LSPR leads to characteristic extinction (absorption plus scattering) bands that may span through UV, visible and near-IR parts of the energy spectrum [4]. Explorations of the coupling of light and charge via localized surface plasmons have led to the discovery that plasmonic excitation can influence macroscopic flows of charge *(coupling plasmons to electron flow)* and, conversely, that charging and discharging events of a nanoparticles can change the plasmonic excitation (*coupling electron flow to plasmons*) (see also Chaps. 5 and 6 in this book). This has recently captured significant interest because of the possibility to induce large optical changes in devices such as smart windows and displays. Moreover, such plasmon-electron tunable materials have potential applications in signal transduction and processing, where electronically tunable plasmonic waveguides are emerging as a potential
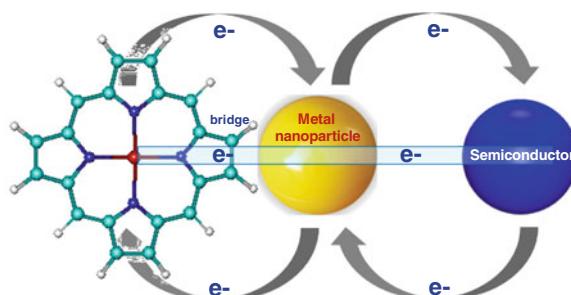
Fig. 13.1 Scheme representing the electron transfer between a semiconductor dot and a molecule bridge by a metal nanoparticle

solution to the miniaturization of metallic interconnects in integrated circuits, and to the design of responsive materials for thermal imaging and spectroscopy [5].

Semiconductor-metal nanocomposites are interesting because they are widely employed in photocatalysis, where the metal in contact with the semiconductor greatly enhances the overall photocatalytic efficiency [6], and in photovoltaics, where metal nanostructures are used to improve absorption in photovoltaic devices, permitting a considerable reduction in the physical thickness of solar photovoltaic absorber layers, and yielding new options for solar-cell design [7].

Recent studies have shown that semiconductor-metal nanocomposites exhibit shifts in the Fermi level to more negative potentials [8]. Such a shift in the Fermi level improves the energetics of the nanocomposite system and enhances the efficiency of the interfacial charge-transfer process.

A well-known example of metal nanoparticles deposited on semiconductor surfaces with catalyzed interfacial charge-transfer process is given by the TiO_2-silver (TiO_2-Ag) and TiO_2-gold nanocomposites that have been shown to store photogenerated electrons and discharge them on demand, enabling application of such nanocomposites in optics, microelectronics, catalysis, sensors, information storage, and energy conversion [9–12].

Interestingly, the dynamics of this charging process can be monitored by optical spectroscopies through variation of the absorption coefficient as schematized in
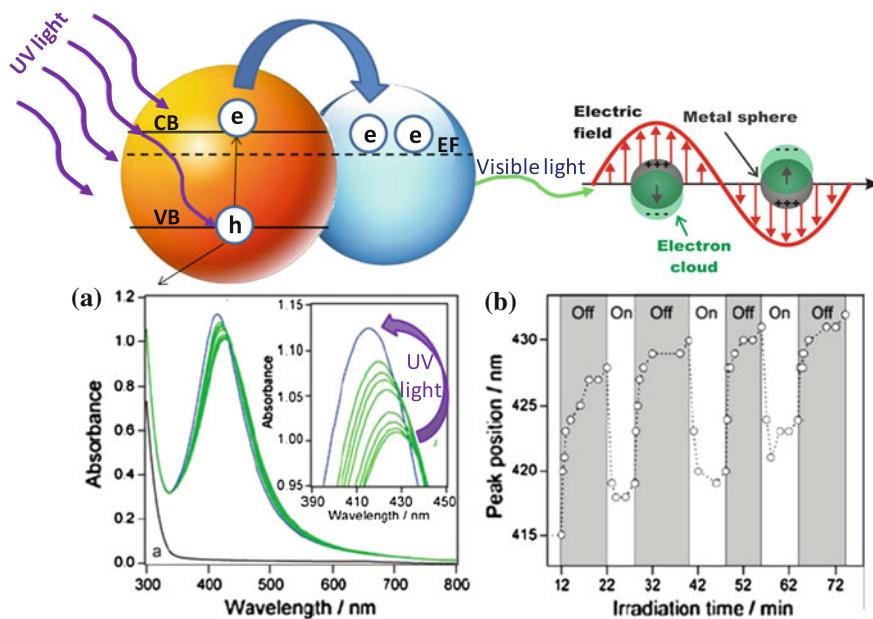
Fig. 13.2 a Absorption spectrum of TiO_2-Ag nanocomposite showing the LSPR of Ag blue-shifting under UV-light irradiation. **b** Shift in the plasmon peak during on-off cycles of light irradiation. The scheme of the charge transfer from the TiO_2 to the Ag nanoparticles during light on is also schematized. [Adapted with permission from Ref. [6]. Copyright (2011) American Chemical Society]

Fig. 13.2. The absorption spectrum of the TiO_2-Ag nanocomposite exhibits a localized surface plasmon resonance (LSPR) absorption band at 430 nm due to the Ag nanoparticles. When the system is irradiated with UV light a blue shift to 410 nm in the plasmon absorption band occurs. At light *off*, the plasmon absorption reverts to a position close to the original absorption at 430 nm. This occurs because under UV excitation, charge separation takes place within the TiO_2 semiconductor. While the holes are scavenged by the surrounding medium, the electrons are injected into the silver core (Fig. 13.2), and the electron transfer continues until the two systems achieve Fermi level equilibration [10].

The interesting aspect is the ability to charge the metal core with electrons under UV excitation and discharge them on demand and probe this electron variation optically. The charging and discharging effects in TiO_2-Ag nanocomposites can be monitored from the shift in maximum plasmon absorption. The shift in the plasmon absorption is quick and can be reproduced during repeated charge and discharge cycles.

Molecule-metal nanocomposites obtained by the assembly of nanosized metal particles with functionalizing molecules is a field of great fundamental and technological interest because charge transfer between photoactive molecules and nearby metallic nanostructures produce modulated optical and electrochemical properties that are desirable for applications including electrocatalysis [13], photoenergy conversion [14], and molecular sensing, including surface enhanced Raman spectroscopy sensing (SERS) [15].

In the case of hybrid assemblies having metal nanoparticles as the core, the (C) electron transfer depends critically on the size and shape of the nanoparticles and the distance between the photoactive molecule, and it is a process that competes with many others including (A) intermolecular interactions of the photoexcited fluorophore bound to a gold nanoparticle, (B) energy transfer, and (D) emission from the chromophores bound on the metal nanoparticles, as summarized in Fig. 13.3.

As an example, the nonmetallic property of ultrasmall metallic particles can also be utilized to capture electrons from an excited sensitizer and thus mediate a photoinduced electron-transfer process. Controlled charging of the Au nanoassembly, e.g. by applying a potential to the Au nanoparticles, enables one to modulate the excited-state interaction between the gold nanocore and surface-bound fluorophores, and therefore modulate the emission of light from the fluorophores: as schematized in Fig. 13.4, charging the Au nanoparticles with an electron may lead to emission of light, while electron transfer from the molecule to a positively charged Au nanoparticles may actually suppress emission of light [17].

Therefore, an understanding of the interactions between metallic nanoparticles and photoactive molecules and/or semiconductors is crucial for the realization of applications.

The question we address in this chapter concerns spectroscopic ellipsometric methods to probe in real time those interface charge transfer phenomena. Changes in the electronic distribution of nanoparticles perturb the optical excitation of the SPR band or the interband or intraband transitions, allowing the dynamics of the electrons and the lattice to be revealed by probing changes in the SPR band absorption.
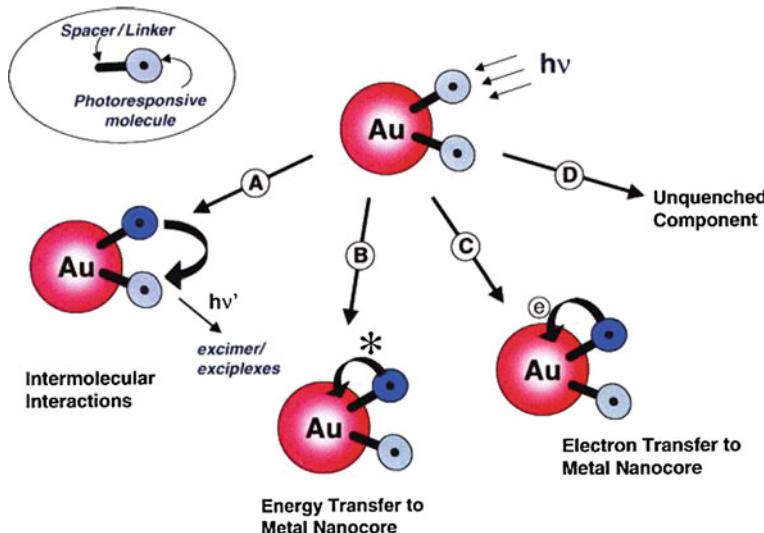
Fig. 13.3 Photoexcitation of the chromophore bound to gold nanoparticles followed by its deactivation via energy transfer, electron transfer, and intermolecular interactions [Adapted with permission from Ref. [16]. Copyright (2003) American Chemical Society]

In this chapter we explore spectroscopic ellipsometry for monitoring *electronic transfer processes* at interfaces of various metal-based nanocomposites. The method relies on the fact that, for some metals, the transfer of electronic charge results in optical changes of the absorption spectrum of the metal, as described in the next paragraph. Owing to its noninvasive character, ellipsometry is a versatile in situ real-time technique for studying a large variety of surface and interface processes, such as nanoparticles nucleation and growth as well as the deposition of materials, chemisorption of molecules, redox reactions, (bio)sensing, alloying, and electrochemical processes involving small metal nanoparticles that exhibit localized surface plasmon modes. In particular, we analyze the use of ellipsometry in probing interface phenomena in the following three nanoscale systems:

(1) *Semiconductor-ultrathin metal film,* which allows us also to demonstrate the applicability of ellipsometry in monitoring chemisorption on semiconductors. We show how real-time spectroscopic ellipsometry can be usefully exploited to probe fast chemisorptions kinetics of metals on semiconductor surfaces and the growth of metal nanoparticles on semiconductors. The chemical-kinetic analysis of the kinetic ellipsometric profile monitored directly on the surface of the semiconductor during chemisorptions of the target specie highlights phenomena of charge-transfer-induced adsorption.

(2) *Semiconductor-metal nanoparticles*, which serves also as an example of the applicability of real-time ellipsometry in tuning the localized surface plasmon resonance of metal nanoparticles. The optical response of metallic nanoparticles
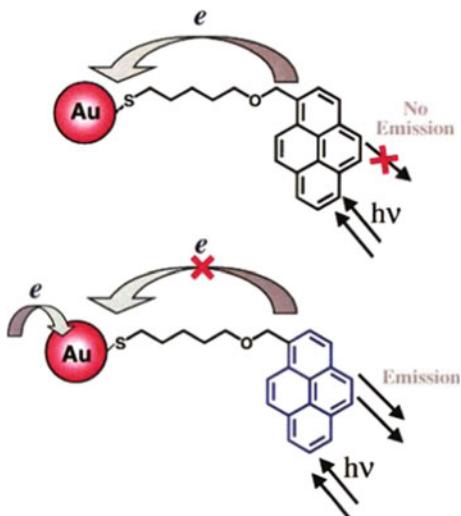
Fig. 13.4 Modulation of photoinduced electron transfer between excited chromophore and gold nanoparticles. [Adapted with permission from Ref. [16]. Copyright (2003) American Chemical Society]

with sizes of a few tens of nanometers can lead to localized surface plasmon resonance (LSPR) resulting from the coupling of the collective excitation of the conduction electrons with the incident light. The frequency at which the maximum response due to the charge oscillation in the metallic nanoparticle occurs is defined as the localized surface plasmon resonance (LSPR) frequency [4].

(3) *Metal nanoparticle-molecule*, which provides a good example of the applicability of spectroscopic ellipsometry in chemo and bio-sensing. Through the ellipsometric probing of the variation of the plasmon resonance of plasmonic metal nanoparticles functionalized with optically active molecules, we also extend this ellipsometric analysis to electron transfer driving the self-assembling of monolayers (SAMs).

This is highly interdisciplinary, involving the chemistry of surfaces, optics and physics: it is shown how spectroscopy and the optical concept of plasmonics can serve a better understanding of chemical surface and interface kinetics and dynamics also involving biomolecules, in order to realize a novel class of sensing and energy conversion, molecular opto-electronic switching and storage devices.

13.2 Real-Time Spectroscopic Ellipsometry

Real-time ellipsometry refers to in situ measurements performed during materials growth or surface modifications. It requires performing the measurements at all accessible wavelengths simultaneously, which enables parallel processing of all
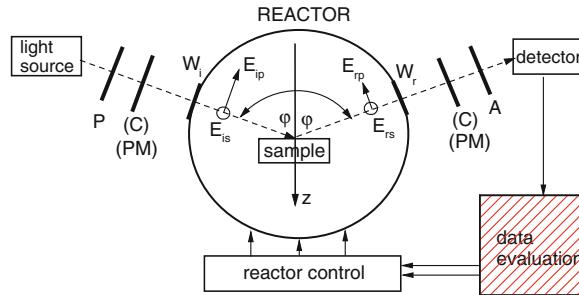
Fig. 13.5 Schematic configuration for closed-loop control using spectroscopic ellipsometry measurements. A pair of windows, W_i and W_r, enables the incident and reflected polarized light to enter and exit the reactor

wavelengths, and which is essential for many real-time applications at high speed while retaining the precision and accuracy necessary to follow the time scale of the dynamics of interest at the sample surface, e.g. it should be possible to acquire data in the 10 ms–1 s time scale.

Standard ellipsometric setups can be adapted for monitoring and controlling processes in real-time. The schematic of the assembly is shown in Fig. 13.5. For real-time studies, there are considerations that pertain to the reactor itself, to the ellipsometer itself and to their interface which need attention. As an example, most reactors have a rotating sample disk for homogeneity of samples and are equipped with strain-free window ports, between the polarizer (P) and optional compensator (C) and sample for incident light and the sample, optional compensator and analyzer for the reflected light beam, for optical access to the growth interface. A fast variant uses phase modulators (PM) in either of the light beams. Typically, the windows are located relatively far away from the substrate and are purged with carrier gas, so that deposition on the windows is avoided. Furthermore, the windows can influence the polarization state of light and have to be properly characterized. As for sample rotation, to obtain useful results with an ellipsometer it is essential that the rotation of the sample is synchronized to the rotation of the appropriate element of the ellipsometer [18].

Spectroscopic reflection ellipsometry has been developed and applied as a powerful tool for the study of thin film growth and surface modification. The most widely used instrument designs for real-time spectroscopic ellipsometry are the rotating-element ellipsometers, in which a polarizer (RPE), analyzer (RAE), or compensator (RCE) on the polarization generation or polarization detection arm of the instrument is rotated continuously as a function of time, and the phase modulated (PME) instruments, which have the advantage of no moving parts (see Chaps. 1 and 2 of this book). An overview of ellipsometer designs, characteristics, calibration and procedures is given by Collins et al. in [19]. Here, we just summarize the basic considerations for real-time applications. The first key issue for real-time spectroscopic ellipsometry is the data acquisition mode, which determines the measurement interval and the precision. The advantage of the rotating-element ellipsometer design is that the

readout time of the photodiode array used as the detector, $\sim$5 ms, is shorter than the rotational period accessible to the rotating element ($\sim$50–100 ms). As a result, the photodiode array can be operated in an integrating mode for fully parallel data acquisition, in contrast to a subsequent phase-modulation ellipsometer design that employs serial data acquisition. A rotating-polarizer multichannel ellipsometer (RPE) allows to acquire (Ψ,Δ) points at 64 spectral positions with an acquisition time of 25 ms [20]. Conversely, with phase modulated ellipsometers (PME), spectra in (Ψ,Δ) with 32 spectral positions have been collected at 16 ms [21]. The second issue concerns the spectral range/ resolution: both types of spectroscopic ellipsometers have made available the very broad range 1.5–6.5 eV (825–190 nm) with wavelength detection bandwidths of 2 nm at 190 nm and 11 nm at 825 nm. The third key issue is measurement accuracy. The rotating-polarizer (RPE) and rotating-analyzer (RAE) configurations exhibit disadvantages stemming from their inability to measure spectra in the fourth component of the 4×1 (real) Stokes vector of the light beam reflected from the sample. Because RPE and RAE measured cosΔ, their accuracy was poor when $\Delta \sim 0°$ and $\pm 180°$, leading to poor accuracy when measuring processing of dielectric materials with $\Delta \sim 0°$ or dielectric materials on metallic and semiconductor substrates for which Δ periodically crosses $0°$ and $180°$ as a function of wavelength. The accuracy was also lower in PME when $\Psi \sim 45°$ (when measuring at the polarizer, modulator and analyzer angles of $45°$, $0°$ and $45°$, i.e., the so-called configuration II) or when $\Delta \sim \pm 90°$ (when measuring at the polarizer, modulator and analyzer angles of $0°$, $45°$ and $45°$, i.e., the so-called configuration III). However, nowadays with the advancement of instrumentation, both issues have been solved. The single-rotating-compensator (RCE) configuration was developed to overcome the limitations of the rotating polarizer through measurement of spectra in the full Stokes four-vector, providing the sign of the ellipticity for the polarization state of the reflected beam. In contrast to RPE and RAE systems, which are only sensitive to three of the four Stokes parameters, RCE systems measure the sign and the cosine of the relative phase Δ and, thereby, uniquely determine Δ, which removes the previously encountered loss in data where $\Delta \sim 0$ or 180 [22]. Nevertheless, RCE was limited in its ability to characterize anisotropic thin films. Therefore, a dual-rotating compensator multichannel ellipsometer with the ability to extract spectra in the full 4×4 (real) Mueller matrix of the sample in a single optical cycle of the two synchronized compensators [23] was developed for real-time ellipsometry.

The Mueller matrix can over-determine the quantities that define the non-diagonal (complex) Jones matrix, the latter providing a mean for analysis of anisotropic media, even when partial depolarization occurs. Many designs have been proposed, and demonstrated, for Mueller polarimeters and ellipsometers, and in addition to the previously mentioned one, a Mueller polarimeter based on modulation and analysis of light polarization by ferroelectric liquid crystal cells has also been developed. It provides non-resonant polarization modulation that can be easily adapted to a CCD array coupled to a dispersion grating in order to do fast real-time spectroscopic measurements. In this way the acquisition of a complete Mueller matrix with 1 nm resolution in the range from the visible to the near infrared can be obtained in approximately 1 s. The design of the polarimeter was based on an objective criterion that allowed

optimizing the accuracy of the measurements. As a result, we measured a complete Mueller matrix with relative error lower than 0.5 %. Whenever the polarimeter is used as an ellipsometer, it is possible to profit the information given by the 16 Mueller matrix coefficients to enhance the accuracy and precision of the measured ellipsometric angles Ψ and Δ [24].

Description of Mueller Matrix ellipsometry and examples of its application are given in Chaps. 2 and 9 of this book.

For the real-time studies in this chapter, the samples did not show in-plane or out-of-plane anisotropy; this was verified by measuring some samples by Mueller ellipsometry and by ellipsometry at various angles of incidence and angle of in-plane rotation. Therefore, under the isotropy assumption, standard phase modulated ellipsometry has been applied in real-time acquiring a spectrum with 64 wavelengths per second.

13.3 Rationale for Probing Charge-Related Phenomena with Spectroscopic Ellipsometry

The key concept is the dielectric function of the metal.

The dielectric function can be written as a combination of interband transitions, $\varepsilon_{IB}(\omega)$, accounting for the response of the d-electrons, and a Drude term, $\varepsilon_D(\omega)$ considering the free conduction electrons only

$$\varepsilon(\omega) = \varepsilon_{IB}(\omega) + \varepsilon_D(\omega) \tag{13.1}$$

The optical probing of the electron transfer processes by ellipsometry is based on the fact that the conduction electrons dominate the response of metals at optical wavelengths, and the Drude theory [25] describes their dielectric function $\varepsilon(\omega)$ throughout the visible part of the spectrum according to

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\gamma\omega} \tag{13.2}$$

where γ is a phenomenological damping constant and equals the plasmon bandwidth for the case of a perfect free electron gas. The damping constant γ is related to the lifetimes of all electron scattering processes in the bulk material that are mainly due to electron-electron, electron-phonon, and electron-defect scattering, i.e., $\gamma = 1/\tau$ where τ is the mean relaxation time of conduction electrons.

The bulk plasma frequency, ω_p, is given by

$$\omega_p^2 = \frac{Ne^2}{m \cdot \varepsilon_0} \tag{13.3}$$

where e is the electron charge, m is the effective mass, ε_0 is the permittivity of vacuum, ε_∞ is the high-frequency contribution from interband transitions, and N is the conduction electron concentration.

Therefore, the real, $\varepsilon_1(\omega)$, and imaginary, $\varepsilon_2(\omega)$, parts of the Drude dielectric function are given by

$$\varepsilon_1(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + \gamma^2}$$

$$\varepsilon_2(\omega) = \frac{\omega_p^2 \cdot \gamma^2}{\omega(\omega^2 + \gamma^2)} \tag{13.4}$$

Therefore, both the real and imaginary parts of the dielectric function depend on the conduction electron concentration, N, and changes to the particle charge alter N and thereby the dielectric function by an amount $\Delta\omega_p$ given by:

$$\varepsilon_1(\omega) = \varepsilon_\infty - \frac{\omega_p^2 + \Delta\omega_p^2}{\omega^2 + \gamma^2} \tag{13.5}$$

In the case of metal nanoparticles showing a localized surface plasmon resonance (LSPR), the frequency of the LSPR (ω_{LSPR}) is proportional to the bulk plasmon frequency (ω_p), and, hence, to the square root of the free carrier concentration in the nanoparticle volume

$$\omega_{LSPR} \propto \omega_p \propto \sqrt{N} \tag{13.6}$$

Thus, as we inject charge into the nanoparticle, the SPR peak shifts to higher energy. The optical effects due to electron charging can be dramatically enhanced by altering the particle geometry and/or by supporting the nanoparticles on semiconductors or charged substrates; as an example the surface plasmon band shift, caused by the electron injection, is larger for nanorods than for small spheres. The effect of geometry can be taken into account by introducing a shape-dependent depolarization factor, L.

Within the dipole approximation [3], the absorption modes of a small metal particle of arbitrary shape consist of several plasmon modes that fulfil the condition

$$(1 - L)\varepsilon_m + L\varepsilon_1(\omega) = 0 \tag{13.7}$$

where ε_m is the dielectric constant of the non-absorbing matrix embedding the nanoparticle, normally equated to the square of the refractive index. Introducing Eq. (13.7) into Eq. (13.5) yields

$$(1 - L)\varepsilon_m + L\varepsilon_\infty - L\frac{\lambda^2}{\lambda_p^2}\left(1 + \frac{\Delta N}{N}\right) = 0 \tag{13.8}$$
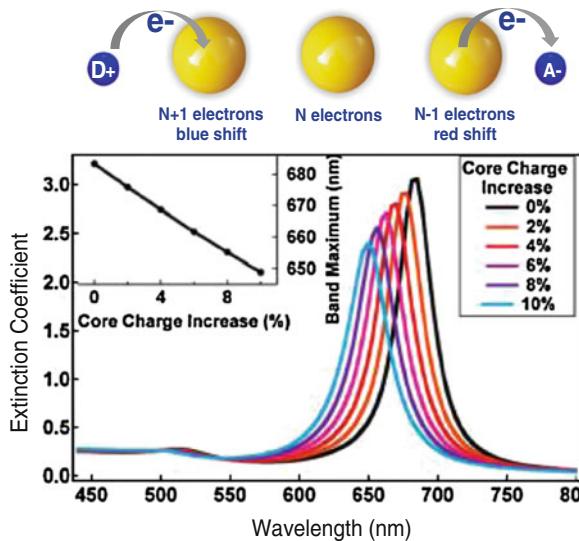
Fig. 13.6 Calculated scattered light spectra (dipole approximation) of charged gold nanorods. Variation of the plasmon resonance of noble metal nanoparticle as a function of the increase of electron concentration ΔN; the sketch schematizes the blue-shift or red-shift of the plasmon resonance by accepting or donating an electron from a donor (D+) or to an acceptor (A−) molecule

Here, λ is the measured surface plasmon peak wavelength and λ_p is the bulk metal plasma wavelength. Provided $\Delta N/N \ll 1$, we can simplify this as

$$\lambda = \lambda_p \left(1 - \frac{\Delta N}{2N}\right) \sqrt{\varepsilon_\infty + \left(\frac{1}{L} - 1\right)\varepsilon_m} \qquad (13.9)$$

By indicating

$$\lambda_0 = \lambda_p \sqrt{\varepsilon_\infty + \left(\frac{1}{L} - 1\right)\varepsilon_m} \qquad (13.10)$$

the plasmon peak position for the case of no electron injection, we can relate the wavelength shift to the variation of electron concentration as:

$$\Delta\lambda = \lambda - \lambda_0 = -\frac{\Delta N}{2N}\lambda_p \sqrt{\varepsilon_\infty + \left(\frac{1}{L} - 1\right)\varepsilon_m} \qquad (13.11)$$

When N is increased, the SPL band is blue-shifted, and there is little change in bandwidth and a slight damping. This is due to the increase in imaginary component of the dielectric function when N is increased as shown in Fig. 13.6.

Thus, changes in particle charge can be followed spectroscopically, also by ellipsometry.

The surface plasmon resonance is sensitive to the number of injected electrons, although the example below develops some limiting considerations. As an example, for a gold ellipsoid with transverse semiaxes of 2 nm and a major semiaxis with a length of 10 nm, for which the number of gold atoms is $\sim$9800, the injection of just one electron would yield a blue shift of $\sim$0.05 nm according to Eq. (13.11), which is lower than the resolution of current spectrometers. However, the differential change in the intensity around the surface plasmon peak as a result of the blue shift can still be significant [26] to be detected.

The influence of charge density on the properties of localized surface plasmons in metallic nanoparticles was first qualitatively studied by Henglein and Mulvaney in 1991 [27]. Almost ten years later, Mulvaney provided quantification and evidence of the capability of probing optically charge transfer by combining ellipsometry measurements of the SPR variation with electrochemical measurements [28]. They showed the absorption spectra of silver nanoparticles as a function of the potential applied to the electrode. The silver nanoparticles plasmon resonance blue-shifted from 430 to 393 nm and became narrower and more intense when a potential of -2 V was applied. Upon application of a positive potential, the plasmon peak red-shifted back towards its original position [28].

The relationship between electron flow and plasmon resonance can be used to understand the broad range of accessible plasmon resonance wavelengths by selecting materials with different electron concentrations; this is illustrated in Fig. 13.7, which shows that plasmon resonances are tunable from the ultraviolet to infrared by varying the electron density from 10^{23} to 10^{19} cm^{-3}.

One of the challenges in achieving even larger SPR shifts in metals is that for most metals, the high initial electron concentration n$_i$ ($\sim 10^{23}$ cm^{-3}) implies that large net flows of electrons are needed to induce large changes in the SPR maximum (see Eq. 13.11). These large net flows are limited by the constraints of material stability and the stability of the medium in which the nanoparticles are suspended. Therefore, materials with lower electron concentrations are extremely attractive because a similar net flow of electrons can lead to larger relative changes in electron concentration (and therefore plasmon resonance), as depicted in Fig. 13.6. This opens up new possibilities for highly tunable plasmoelectronic devices in which the flow of small amounts of current induces a large plasmonic response.

13.4 Gallium: A Non-Noble Plasmonic Metal with Non-Linear Optical Properties

In this chapter we discuss examples involving the non-noble plasmonic metal gallium (Ga), which has been reported to possess non-linear optical properties [29]. Recently, researchers have been increasingly drawn to the enhanced nonlinear effect of phase transitions in polyvalent metals, which can have different charges, such as gallium
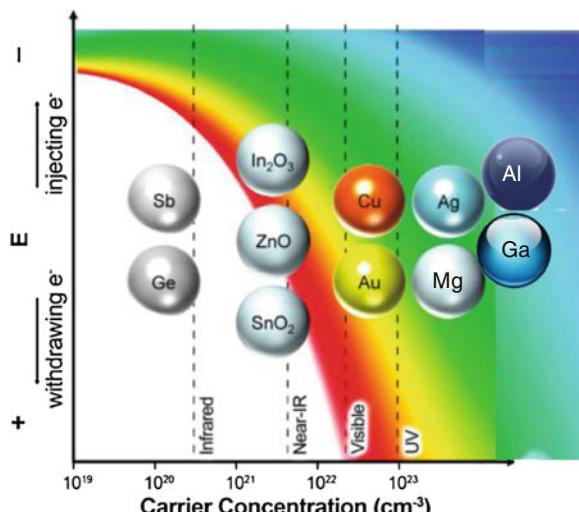
Fig. 13.7 Coupling electron flow to plasmonic excitation. By either injecting or "withdrawing" electrons, one can tune the plasmon resonance of a material. The degree to which the resonance can be manipulated is strongly dependent upon the material's carrier concentration. Materials that have a higher initial electron concentration (i.e., metals) are generally limited to smaller shifts in plasmon resonance, as the injection or removal of even large numbers of electrons results in a small relative change in the electron concentration (see Eq. 13.11 for justification). Conversely, materials with low carrier concentrations typically have plasmon resonances in the infrared, but much larger relative changes in electron concentration and, therefore, plasmon resonance are possible. [Adapted with permission from Ref. [5]. Copyright (2012) American Chemical Society]

(Ga). The significant difference in optical properties between the two phases of Ga— α and metallic—make it a very suitable material for plasmonic switches.

Gallium is a trivalent element of group IIIA, like the much more studied aluminium. However, little is known about its optical and plasmonic properties. Therefore, we have applied spectroscopic ellipsometry to investigate its optical properties at the nanoscale, which are reported herein.

Gallium has a unique combination of a high refractive index, several structural phases with widely different optical properties, a phase transition at very accessible temperatures close to room temperature, and most importantly, a susceptibility to light-assisted structural transformations. When Ga is prepared in the form of nanoparticles, it exhibits a huge optical nonlinearity via light-assisted surface metallization with switching times in the picosecond–microsecond range.

Gallium is known for its polymorphism; Ga(II), Ga(III) and liquid gallium, in terms of their optical and electronic properties, are very close to an ideal free-electron metal [30, 31] and, hence, applicable to plasmonics.

Figure 13.8 shows the Ga phase diagram for bulk Ga and for Ga nanoparticles [32]; α-gallium, the stable 'ground-state' phase, has a unique structure in which molecular and metallic properties coexist: some inter-atomic bonds are strong covalent bonds, forming well-defined Ga_2 dimers (molecules), and the rest are metallic bonds. α-Ga
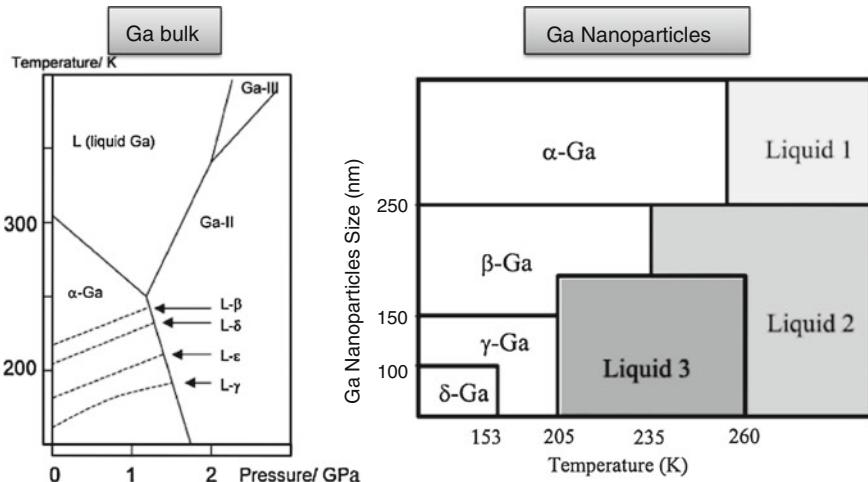
Fig. 13.8 Phase diagram for bulk gallium (*left panel*) and for gallium nanoparticles (*right panel*) showing liquid gallium, Ga(III) and Ga(II), α-gallium, and the metastable phases β, δ, ε, γ; for the Ga nanoparticles phase transitions also depend on nanoparticles' size

has a very low melting point of 29.88 °C. Interestingly, for nanoparticles, phase transitions also depend on nanoparticles' size.

The gallium nonlinearity is associated with a light-induced structural phase transition in the common form of α-gallium in the solid phase. It is believed that in the nanosecond-microsecond regime of optical excitation, the mechanism of the phase transition is predominantly non-thermal. Optical excitation is highly localized and destabilizes covalent bonding within the crystalline structure of α-gallium, thus provoking a surface assisted transition to a more metallic, more reflective metastable phase. The structural phase transition drives a considerable change in the electronic, and in particular, in the optical properties of the material, also across a very broad spectral range [32]. The light-induced structural phase transition in a confined geometry like that of Ga nanoparticles, is achieved more easily by bringing the material near to the phase-transition temperature, thus increasing the magnitude of the optical nonlinearity.

The band structure, shown in Fig. 13.9, has been investigated by Hunderi et al. [30] and it shows a number of parallel bands, most notably around the Γ point along **Δ** and **T.** Another set is found around the **Z** point along T' and **U** and a third from the M point along **W.** All these bands will give transitions in the range 1.25–2.4 eV.

The dielectric functions of liquid gallium and α-Ga are different, and at a wavelength of 1.55 μm the difference $|\varepsilon_{liquid}| - |\varepsilon_\alpha| \approx 180$. The reflectivity of α-Ga is more like that of semiconductor materials, in contrast with the high reflectivity of the metallic phases.

The dielectric function of gallium has been measured by spectroscopic ellipsometry in a broad experimental range of 0.7–6.5 eV by undercooling a continuous and very smooth gallium film. The determined spectra are shown in Fig. 13.10.
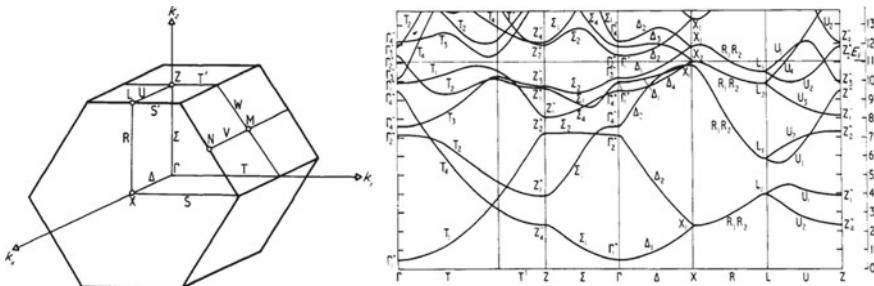
Fig. 13.9 Brillouin zone of gallium and its energy bands along major symmetry lines (re-arranged from Ref. [32])
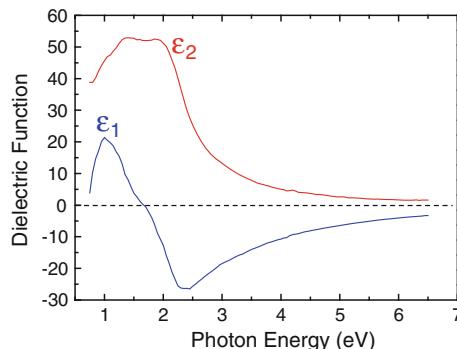
Fig. 13.10 The dielectric function of gallium at 300 K

Evaporation of gallium onto substrates tends to form nanoparticles that are interesting for plasmonics [33] and UV-SERS applications [34] because they are characterized by a LSPR that can be tuned from the UV to the VIS and NIR by tailoring the nanoparticle size. As an example, the DDA method (Discrete Dipole Approximation) predicts that Ga nanoparticles with a diameter of 20 nm have a LSPR at approximately 3.4 eV, as shown in Fig. 13.11, showing that gallium is a possible candidate to cover the UV part of the spectrum differently from gold and silver.

In this regime of Ga nanoparticles, real-time spectroscopic ellipsometry has been demonstrated to probe non destructively, in situ accurately and thereby tune the surface plasmon resonance of metal Ga nanoparticles deposited on various transparent and absorbing substrates. Spectroscopic ellipsometry, as it monitors changes in the polarization state of a linearly polarized beam reflected at oblique angles of incidence, is therefore a tool capable of investigating the full optical behavior of a plasmonic sample in the far field, allowing the pseudodielectric function, $\langle\varepsilon\rangle$, of the sample to be determined directly from SE measurements, for both the parallel and perpendicular measurement configurations. As an example, Fig. 13.12 shows the real-time evolution of $\langle\varepsilon_2\rangle$, during deposition of Ga nanoparticles with increasing size on an Al_2O_3 substrate. The pseudodielectric function is an effective dielectric
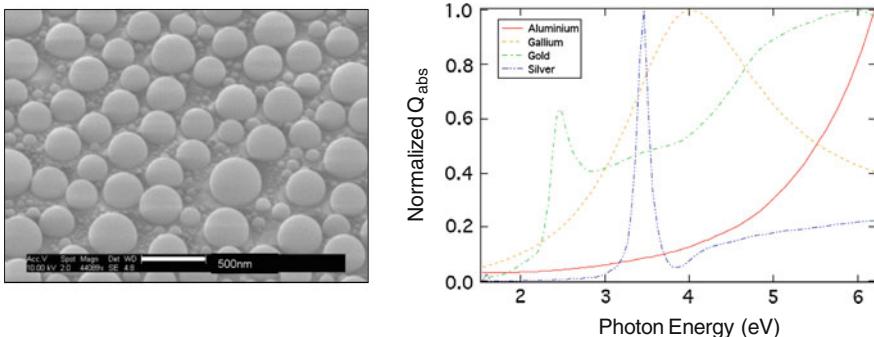
Fig. 13.11 Scanning electron microscopy image and DDA calculations of the normalized absorption efficiency (Qabs) of isolated spheres (R = 20 nm) made of aluminium, gallium, gold and silver respectively [*Courtesy of F. Moreno Department of Applied Physics. University of Cantabria. 39005 Santander (Spain)*]
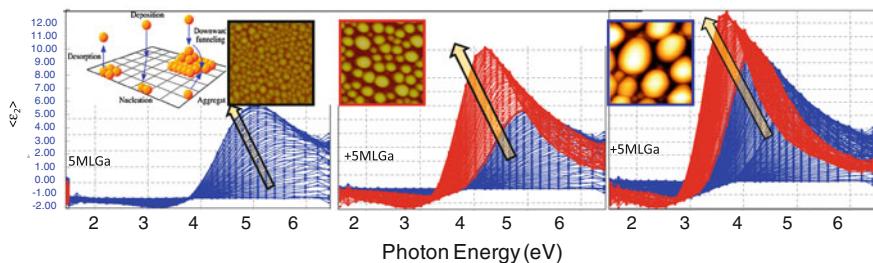
Fig. 13.12 Real-time spectra of the imaginary part $\langle \varepsilon_2 \rangle$ recorded during deposition of Ga nanoparticles on Al_2O_3. Spectra are shown every 10 s of deposition. The spectra continuously redshift with increasing Ga deposition time, i.e., increasing of the nanoparticles size, as indicated by the atomic force microscopy $1 \mu m \times 1 \mu m$ images (each panel reports deposition of 5 MLs = monolayers of Ga)

function describing the optical behavior of the system assuming a simple two-phase sample/ambient optical model. The asymptotic behavior of the NP layer on top of the substrate follows that of the bare Al_2O_3 dielectric function. A distinct resonance emerges just as the deposition begins, which results from incident photons coupling into plasmon modes of the Ga nanoparticles. The plasmon peak continuously redshifts as the deposition time increases, which correlates to an increasing average particle size. The imaginary part, $\langle \varepsilon_2 \rangle$, of the pseudo-dielectric function appears to be the simple sum of the substrate and of plasmonic peaks, whose position and intensity depend on the nanoparticles geometry and can be, therefore, tailored at the desired wavelength. Real-time data provide a means by which the deposition can be monitored and the plasmon resonance tuned to specific photon energy without prior knowledge of the particle size or distribution.

Beyond enabling plasmon resonance tuning of metallic nanoparticle ensembles, real-time monitoring of their collective plasmon resonance can also be applied while modifying the nanoparticles' surrounding environment such as for gas sensing or thermal treatment to understand the optical manifestation of physical modifications, as shown in the next paragraphs.

13.5 Real-Time Ellipsometric Monitoring of Metal Layers Chemisorption on Semiconductors: The Role of Electron Transfer Between the Metal Adsorbate and the Surface

This paragraph discusses the use of spectroscopic ellipsometry to monitor in real time the chemisorption and formation of ultrathin metal layers on semiconductors. The unique capability of real-time ellipsometry to be sensitive to surface coverage even at the monolayer level allowed highlighting and modeling charge transfer phenomena occurring between the metal and the semiconductor, and how they determine the metal surface coverage.

From the scientific point of view, thin metallic adlayers on semiconductor surfaces have attracted research interest as they are model systems for highly correlated 2*d* and 1*d* electron gases [35]. The investigation of chemisorption of ultrathin layers of metals on semiconductors is also of interest for applications such as catalysts, since such ultrathin metallic films can exhibit unexpected, but often magnificent, macroscopic quantum catalysis, whose magnitude increases superlinearly with the decrease of film thickness [36].

Furthermore, metal adlayers are also important to grow good quality semiconductors, as they can act as surfactants. As an example, metal adatoms on a semiconductor prefer to incorporate and diffuse between the adlayer and the substrate. This mechanism, which is called adlayer enhanced lateral diffusion (AELD) becomes activated already at rather low temperatures, thereby enabling the growth of materials having a high melting temperature even at modest temperatures. The materials systems for which this mechanism is important are the group-III nitrides (GaN, InN, AlN, and their alloys) [37]. In the field of III-nitrides epitaxial growth, recent experimental studies indicated that the presence of an In or Ga adlayer leads to a smooth surface morphology and a better crystal quality of InN and GaN nitrides.

The adsorption and desorption kinetics of metal adlayers are generally investigated in the "standard" framework of the Langmuir theory [38].

Indeed, it is quite often neglected that in the case of semiconductors, chemisorption occurs by electron transfer between the adsorbate and the surface [39].

Unlike the Langmuir's approach, Wolkenstein's theory of chemisorption [40] takes into account these electronic interactions between the semiconductor surface and the adsorbate and their effect on the adsorptivity of semiconductors.

The Wolkenstein theory considers both chemical equilibrium between the surface and the gas phase and electronic equilibrium at the semiconductor surface.

The cardinal feature of Wolkenstein's theory is that localized electronic states are created in the semiconductor band gap by the chemisorbed species. These states serve as traps for electrons or holes (acceptor-like or donor-like states, respectively), depending on their nature. The adsorbed species, depending on electron transitions between those states and semiconductors bands, may be chemisorbed on the semiconductor surface in three ways:

(1) "weak" chemisorption involving a neutral adsorbate species: in this case free carriers (electron or holes) from the substrate do not participate in the adsorption process;
(2) "strong acceptor chemisorption" occurring when an electron from the surface is captured by the adsorbate species and denoted as CeL (where eL denotes the free electron participating from the substrate);
(3) "strong donor chemisorption" occurring when a hole is captured by the adsorbate species and denoted as CpL (pL is the free hole from the substrate).

The possibility of those different types of chemisorptions stems from the ability of the chemisorbed species to draw or donate a free electron and/or hole from/to the substrate lattice (L), with a consequent variation in surface band bending. A diagram representing these three forms of chemisorption is shown in Fig. 13.13.

Here we show how time resolved ellipsometry can be used to monitor and model the adsorption of metals on semiconductor surfaces, metal adlayer formation and charge transfer between the metal and the semiconductor, discussing the example of gallium (Ga) on gallium nitride (GaN), but its validity can be extended to other III-nitrides, III-V (e.g. GaAs, InSb) and II-VI (e.g. ZnO) semiconductors.

Figure 13.14 shows the real and imaginary parts of the pseudodielectric function of a GaN bulk template grown by hydride vapour epitaxy; it is the pseudodielectric function because it includes the effect of the surface roughness. GaN has a band gap of 3.4 eV clearly visible as onset of the absorption in $\langle \varepsilon_2 \langle$ (see Fig. 13.14b) and by the corresponding Van Hove singularity in the spectrum of $\langle \varepsilon_1 \rangle$ (see Fig. 13.14a). When a monolayer or bilayer of Ga (whose dielectric function has been reported in Fig. 13.10) is deposited on GaN, an increase of $\langle \varepsilon_2 \rangle$ is experimentally observed above the bandgap of GaN, as shown by the experimental spectra in Fig. 13.14d acquired during the deposition of Ga on GaN in a molecular beam epitaxy system, and also supported by the simulation shown in Fig. 13.14c, which shows the variation of the
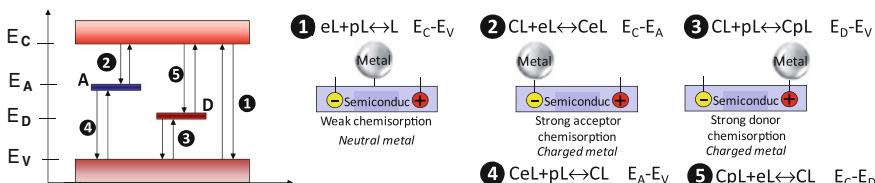
Fig. 13.13 Scheme of chemisorption of a metal on a semiconductor surface involving charge transfer between states of the semiconductor and the metal charge; A (acceptor level)—particle with an affinity for a free electron; D (donor level)—particle with an affinity for a free hole
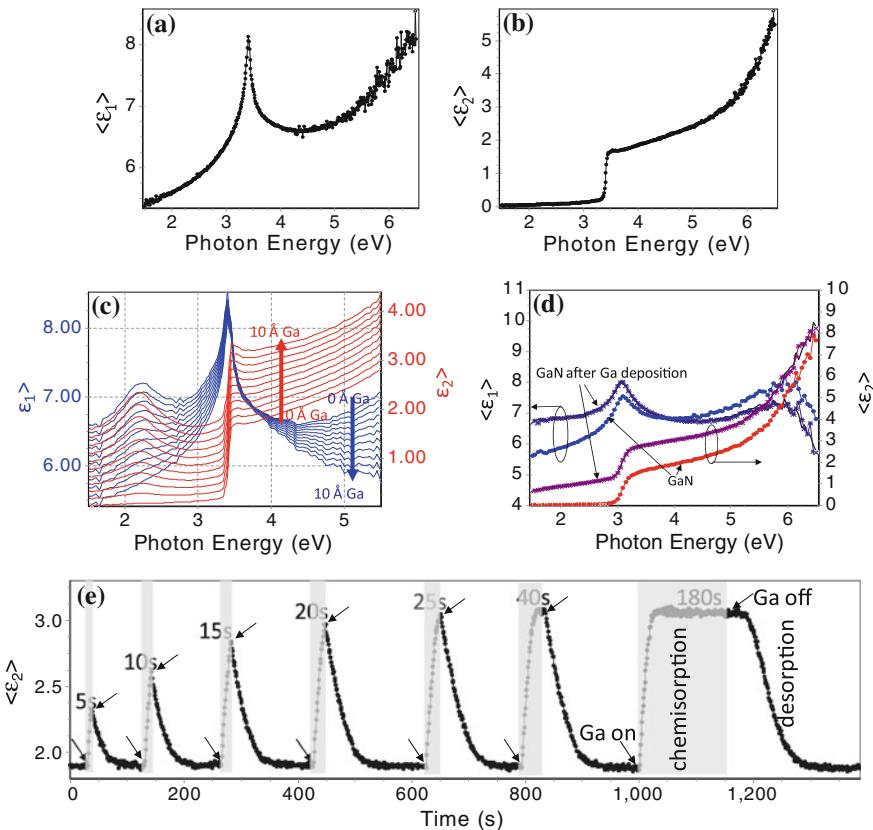
Fig. 13.14 a real, $\langle \varepsilon_1 \rangle$, and **b** imaginary, $\langle \varepsilon_2 \rangle$ parts of the pseudodielectric function of GaN. **c** Simulation of the variation of the GaN pseudodielectric function during chemisorption of 1–10 Å of Ga on GaN; the various curves are for an increment of the Ga thickness of 1 Å at the temperature of 680 °C. **d** GaN experimental spectra measured at 680 °C before and after deposition of two monolayers of Ga. **e** Variation in the imaginary part, $\langle \varepsilon_2 \rangle$, of the GaN pseudodielectric function probed at the photon energy of 4 eV during Ga adsorption when the Ga shutter is on (*shadow area*), and its desorption at the Ga shutter off, at a constant Ga flux of 9.63×10^{-8} Torr and at a temperature of 680 °C. The various times are the timing of the Ga shutter aperture to flux Ga on GaN; this determines different amounts of Ga chemisorbed

GaN pseudodielectric function by depositing 0–10 Å of Ga with an increment of 1 Å to show the sensitivity of the ellipsometric monitoring. From Fig. 13.14 c–d it can be inferred that the chemisorption of few monolayers of Ga on GaN can be monitored by ellipsometry through the temporal variation of the pseudodielectric function. Although spectra have been acquired during Ga chemisorption, in Fig. 13.14e only the data at the probing photon energy of 4 eV are shown. Any probing photon energy above 3.4 eV (the bandgap of GaN) is extremely sensitive to the chemisorption of ultrathin layers and/or cluster because the light penetration depth is limited to few

Angstrom due to the GaN absorption. Specifically, Fig. 13.14f shows the variation of the imaginary part of the GaN pseudodielectric function, $\langle \varepsilon_2 \rangle$, during chemisorption of different amount of Ga and its desorption. The various curves in Fig. 13.14e are obtained by chemisorbing different amounts of Ga controlled by the time the Ga shutter is open. Therefore, the data in Fig. 13.14 show the suitability and sensitivity even of single wavelength ellipsometry in monitoring the deposition of few Angstrom of Ga.

Noteworthy, GaN has the wurtzite structure, which is characterized by faces, i.e., crystallographic planes with different charge. The most common +c (0001) and –c (000-1) planes have fixed negative and positive polarization charges, respectively, at the surface, as schematized in Fig. 13.15, while the m-plane (1-100) is neutral. During chemisorption, electronic charge transfer between the GaN semiconductor and the chemisorbed species may occur, modifying the electronic structure at the surface and in the adjacent space charge region, which in return affects the chemisorption process in a nonlinear manner. This concept of electron transfer between the adsorbate and the surface has been previously considered for adsorption on ZnO [41].

The Ga adlayer adsorption/desorption kinetics can be monitored for the different planes with different charge of GaN recording in real-time the variation in the real, $\langle \varepsilon_1 \rangle$, and imaginary, $\langle \varepsilon_2 \rangle$, parts of the semiconductor (GaN) pseudodielectric function, as shown in Fig. 13.15. The different profiles for the different planes of GaN with different surface charge provide an indication that electron transfer is driving the chemisorption of Ga on GaN.

Specifically the chemisorption rate follows the order

$$(0001) \text{ negative charge} > (000-1) \text{ positive charge} > (1-100) \text{ neutral}$$

This is consistent with the electropositive character of Ga, which is an electrophilic specie and, hence, possibly chemisorbs according to the scheme in Fig. 13.13 as strong acceptor chemisorption.

In order to explain the effect of the different surface charge on the metal chemisorption, the real-time ellipsometric profiles of $\langle \varepsilon_2 \rangle$ were recorded at various temperatures and time of Ga fluxing and they were analyzed to determine the critical coverage for Ga depending on the GaN surface charge, and they are shown in Fig. 13.16.

The critical thickness occurs at an increase, $\Delta \langle \varepsilon_2 \rangle \sim 0.7$ and $\Delta \langle \varepsilon_2 \rangle \sim 1.2$ with respect to bare GaN for the N-polar and Ga-polar surfaces. The Ga surface coverage and/or equivalent thickness can be estimated from the modeling of ellipsometric spectra knowing the dielectric function of GaN and Ga as described in Ref. [42], and by applying a simple 2-layer model (GaN/Ga/air). Thus, for Ga a critical coverage of 4.48 Å is determined, corresponding to 2.5 ML (ML = monolayer) on the negatively charged (0001) surface, and of 3.2 Å corresponding to 1.5 ML on the positively charged (000-1) surface [42, 43]. Thus, also the critical coverage thickness, which is lower for the positively charged (000-1) surface than for the negatively charged (0001) one supports a chemisorption driven by surface charge and charge transfer between the semiconductor and the metal adsorbate.
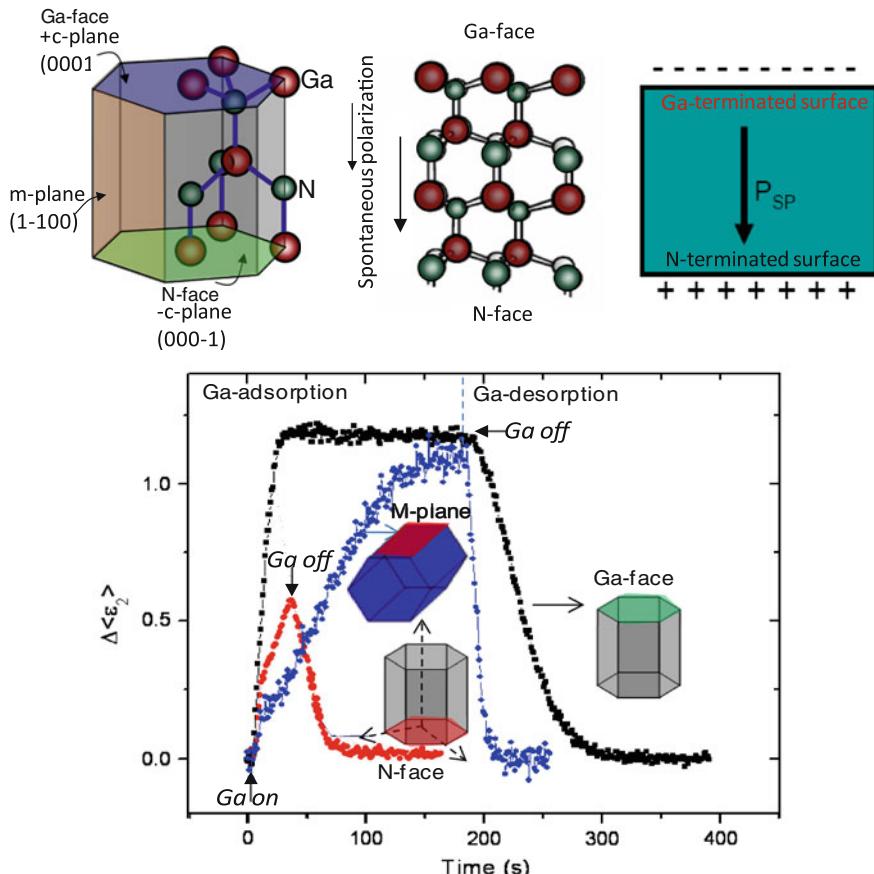
**Fig. 13.15** GaN Wurtzite structure and main crystallographic planes with their fixed polarization charge. Variation in the imaginary part, $\langle\varepsilon_2\rangle$, of the GaN pseudodielectric function probed at the photon energy of 4 eV during Ga adsorption and its desorption at the Ga shutter Off, at a constant Ga flux of 9.63×10^{-8} Torr and at a temperature of 680 °C for the +c (0001) Ga polar face, -c(000-1) N-polar face and for the m- (1-100) non polar surfaces. The initial point of all ellipsometric profiles corresponds to bare GaN surface. $\langle\varepsilon_2\rangle$ increases during Ga chemisorption and decreases during Ga desorption

Data in Fig. 13.16 also show that the critical Ga thickness decreases with increasing temperature, an observation that can be explained by a decrease in the adsorption rate constant and an increase in desorption rate constant with the increase in the temperature, indicating that the adsorption on those charged surfaces is not a thermally activated process. This is another indication that the driving force for the chemisorption of Ga metal on GaN comes from surface charge.

Another interesting feature revealed by the direct sampling of the surface modification by real-time ellipsometry is the delay existing between the Ga OFF-time and the onset of $\langle\varepsilon_2\rangle$ decrease, i.e., of desorption. This delay is highlighted by the
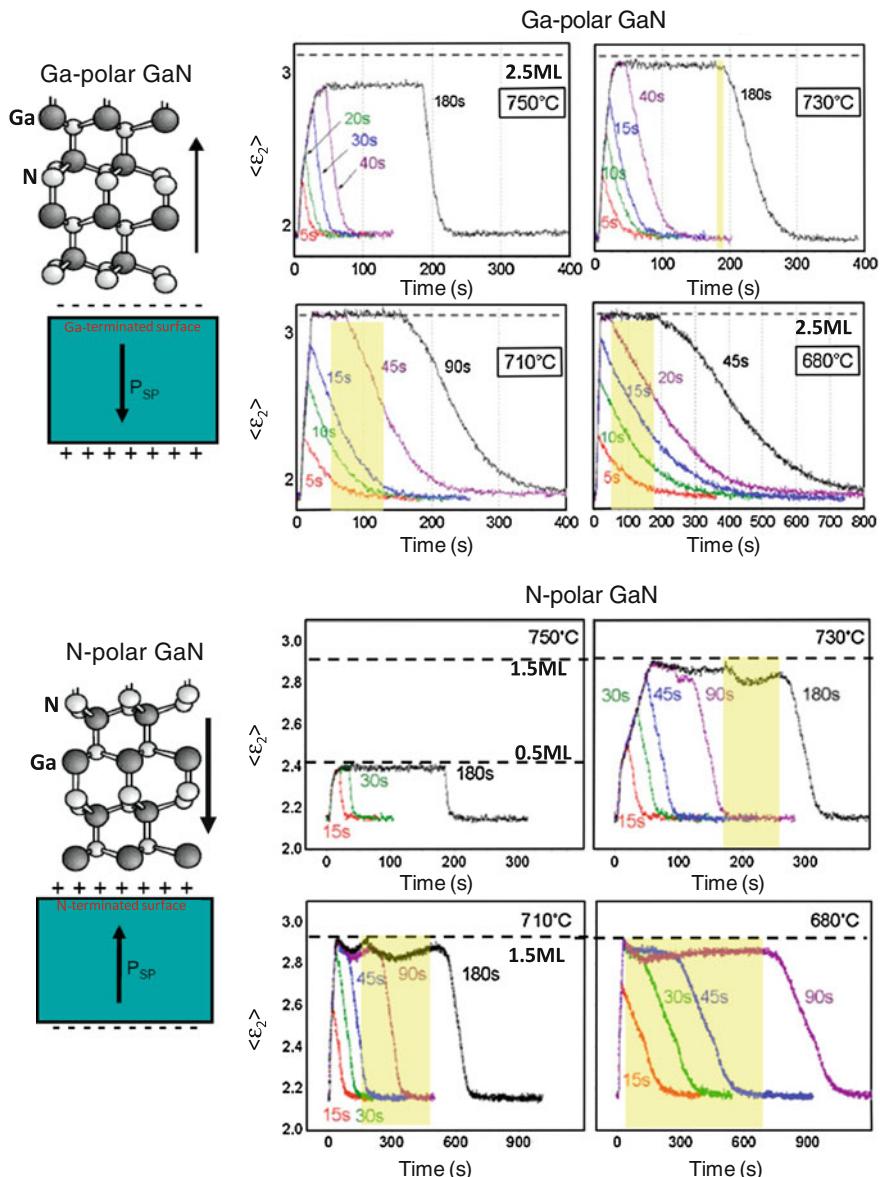
Fig. 13.16 Variation in the imaginary part, $\langle\varepsilon_2\rangle$, of the GaN pseudodielectric recorded at the sampling photon energy of 4 eV function during Ga adsorption and its desorption from N-polar and Ga-polar GaN, at a constant Ga flux of 9.63×10^{-8} Torr for temperatures of 680, 710, 730 and 750 °C and Ga pulse times in the range 5–90 s as indicated in the plots. The dotted lines represent the critical Ga coverage level reached for the various surfaces. The shadow regions indicate the delay between the Ga OFF time and the starting of $\langle\varepsilon_2\rangle$ decrease, i.e., desorption. The sketches representing the polarization surface charge for Ga-polar and N-polar GaN are also shown

yellow-shadow areas for the 180 s/730 °C, 180 s/710 °C and 90 s/680 °C profiles for the N-polar surface and for the 180 s/730 °c, 90 s/710 °C and 45 s/680 °C profiles for the Ga-polar surface. This delay is an additional indication that a strong (charged) form of chemisorption is involved.

If we focus on the analysis of one of those adsorption/desorption profiles, e.g. the 45 s Ga at 680 °C for the Ga-polar surface, we see that the adsorption and S-shaped desorption profiles, that include the observed delay time, cannot be fitted to a Langmuir profile. Only considering electron transfer *(e-)* from the semiconductor to the neutral adsorbed gallium Ga*L, forming a negatively charged surface species, Ga-eL according to the reaction schematized in Fig. 13.13, a consistent analysis of all the measured adsorption/desorption profiles can be achieved. Specifically, a neutral form of Ga, N^0, from the gas phase approaches the surface and captures a free electron (indicated as eL in the chemical reactions below) at the GaN surface, becoming a chemisorbed negatively charged Ga− (denoted as Ga-eL) in the chemical reaction below. This is the charged form of chemisorption, referred to as the "strong" form in Wolkenstein's theory.

Therefore, it can be calculated that the time dependence of the Ga concentration on the surface will be the sum of the neutral N^0 and charged N^- gallium according to:

$$N(t) = N^0(t) + N^-(t) \tag{13.12}$$

where

$$N^0(t) = \frac{N_\infty^0}{(\tau_2 - \tau_1)\,\tau} \left\{ \tau_2\,(\tau_2 - \tau_1) \left[1 - \exp\left(-\frac{t}{\tau_2}\right)\right] + \tau_1\,(\tau_2 - \tau) \cdot \left[1 - \exp\left(-\frac{t}{\tau_1}\right)\right] \right\}$$

and

$$N^-(t) = \frac{N_\infty^-}{\tau_2 - \tau_1} \left\{ \tau_2 \left[1 - \exp\left(-\frac{t}{\tau_2}\right)\right] - \tau_1 \left[1 - \exp\left(-\frac{t}{\tau_1}\right)\right] \right\} \tag{13.13}$$

where τ_1 and τ_2 are complex functions of τ ($1/\tau = k-1$ is the desorption probability), $\tau 0$ and $\tau -$ as reported in Ref. [44].

During desorption, the electronic equilibrium is maintained through two balancing processes: the discharging of charged Ga and desorption of neutral Ga. This desorption mechanism can be summarized by the reaction in Fig. 13.17 which leads to the following expression for the Ga concentration as a function of time:

$$N(t) = N^* - \frac{N^* \cdot \langle pL \rangle^* \cdot \left\{1 - \exp\left[(\langle pL \rangle^* - N^*) \cdot k_2 \cdot t\right]\right\}}{\langle pL \rangle^* + N^* \exp\left[(\langle pL \rangle^* - N^*) \cdot k_2 \cdot t\right]} \tag{13.14}$$

where N* is the Ga concentration impinging on the surface and $\langle pL \rangle^*$ is the number of surface site where Ga is adsorbed.

This charge-transfer based analysis of chemisorption can be understood considering that for an ionic polar surface to be stabilized, the surface charge density must be
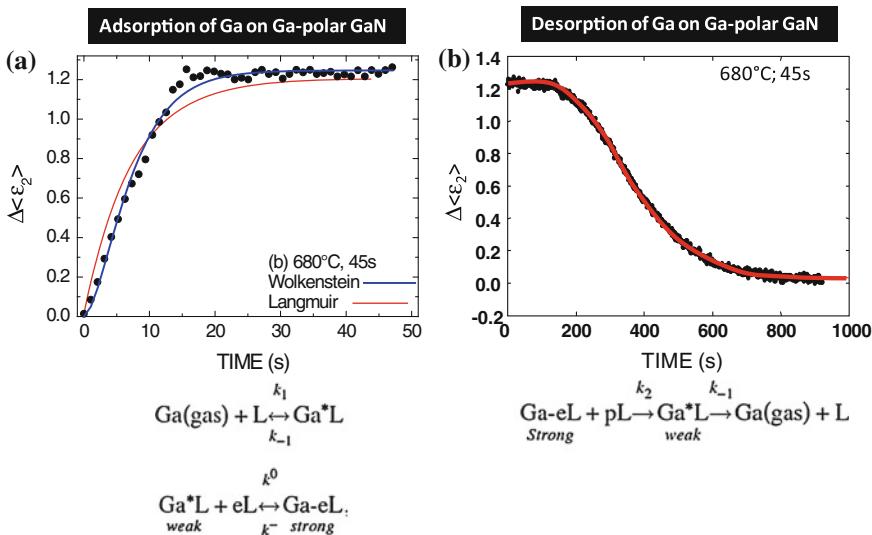
Fig. 13.17   a Fit according to the Langmuir and Wolkenstein theories for Ga adsorption for profile at 680 °C—45 s, and **b** fit according to Wolkenstein theory for Ga desorption

modified so as to balance the polarization electric field, and the charge compensation is operated by the Ga surface metallization.

The final result of this analysis is that fitting all the ellipsometric real-time profiles obtained as a function of temperature to the appropriate electronic model, the kinetic constant and activation energy for metal layers adsorption/desorption on semiconductor surfaces can be obtained. In the specific case of Ga chemisorption on GaN, an activation energy $E_a = 2.85$ eV for Ga desorption from the Ga-polar GaN has been determined.

Thus, this analysis demonstrates that it is possible to transform an ellipsometric input into a chemical-kinetic output.

From a technological perspective, it indicates a way of tailoring the band bending of semiconductor surfaces to tune the critical coverage and metallization of surfaces.

13.6 Charge Transfer in Substrate Supported Ga Nanoparticles: Implications for Plasmonics

13.6.1 Ga Nanoparticles Deposited on SiC

The properties of supported metal nanoparticles differ substantially from homogeneous films or bulk material, since spatial confinement of the electronic wave function in metallic nanoparticles leads to the surface plasmon resonance and to the unusual property of quantized double-layer charging effects [45]. This charging characteristic
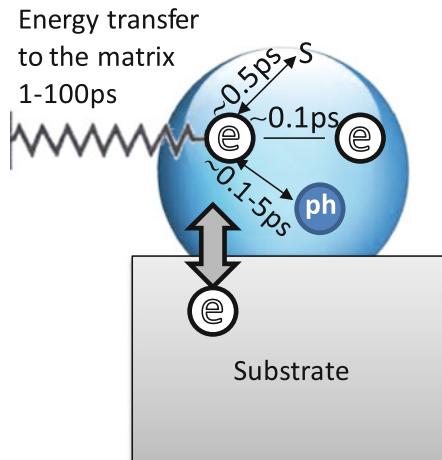
Fig. 13.18 Sketch of the relaxation processes in a metallic nanoparticle including electron–electron (e–e) scattering, electron–phonon (e–ph) coupling, electron-surface (e-S) scattering (S = surface), energy transfer from the particle to the surrounding medium through the phonon–phonon (p–p) coupling and charge transfer between the nanoparticles and the substrate

of metal nanoparticles allows semiconductor-metal nanocomposites to be effective in facilitating photocatalytic processes, in catalysis, sensors, information storage, and energy conversion. When nanoparticles are supported by substrates, the substrate can act as source or drain for excited electrons leading either to energy gain or loss for the electron gas inside the nanoparticles. Accordingly, a modified electron relaxation behavior in supported nanoparticles might substantially influence the photochemistry on their surfaces and the chemical reactivity in general. This perturbation of electron gas in the nanoparticle can induce chemical surface reactions that are inhibited under equilibrium conditions.

Therefore, it is important to investigate the electron relaxation and charge transfer phenomena in supported nanoparticles.

The charge transfer at the interface between nanoparticle and substrate is also relevant for single electron tunnelling devices. In this context supported nanoparticles are of special interest since their small capacitance leads to nonlinear current–voltage characteristics even at ambient temperature.

The coupling strength between the nanoparticles and the substrate determines the charge transfer probability and therefore influences the electron relaxation dynamics [46] and the mechanisms for damping of a collective resonance in metallic nanoparticles. The plasmon can also decay due to a transfer of energy involving electron-hole pairs, electron-surface scattering, phonon scattering and electron-electron-scattering or reemission of photons (radiation damping, which might be altered by finite size effects, depending on the mean free path of the electrons [47]. Different time scales [48] are involved in those relaxation processes, as schematized in Fig. 13.18.

A convenient way to probe the electron storage in a metal nanoparticle is by monitoring its plasmon resonance frequency and amplitude. The addition of electrons to silver and gold nanoparticles causes a blue shift in the absorption spectrum due to the increasing surface plasmon frequency of the electron gas, as explained in paragraph 3. Corresponding changes of the real $\Delta\langle\varepsilon_1\rangle$ and imaginary part $\Delta\langle\varepsilon_2\rangle$ as a function of time, induced by the charge transfer are the physical quantities of interest to understand the electron dynamics as also shown in Ref. [48].

Here we demonstrate that spectroscopic ellipsometry is a well suitable probe to monitor in real time the charge transfer between substrate and Ga nanoparticles, and to investigate its effect on the dynamics of the surface plasmon resonance and on the growth of the nanoparticles itself.

We show this by the example of Ga nanoparticles on polar SiC substrates. Fundamentally, SiC is intriguing because it is a polar semiconductor, like GaN discussed in the previous paragraph.

The non-centrosymmetric structure and the partial ionicity of the bonds, which results from the different Pauling electronegativity of gallium and nitrogen, yield a net spontaneous polarization field along the *c*-axis of $+0.043$ C/m^2. Therefore, the Si-polar SiC(0001) and C-polar SiC(000-1) are characterized by different surface charge of opposing sign, with the Si-polar surface having a positive charge and the C-polar surface a negative one (see scheme in Fig. 13.19).

When Ga nanoparticles are grown on the C-face or the Si-face of SiC, a different surface plasmon dynamics is observed by real-time spectroscopic ellipsometry monitoring as shown in Fig. 13.18. Indeed, for equivalent Ga nanoparticles size, the position and amplitude of the plasmon resonance vary for the opposing polarity surfaces, despite the constant refractive index for the SiC substrate for the two faces. In order to better emphasize the difference in NPs plasmon evolution on the opposing polarity surfaces, the energy (wavelength) of the longitudinal plasmon mode was extracted from the real-time data and plotted as a function of total Ga dosage (in MLs) in Fig. 13.18c. The Si-surface exhibits faster Ga plasmon evolution compared to the C-surface, implying that blue-shifted plasmon resonances are observed for the C-polar SiC.

In keeping with the opposite signs of the surface polarization charge (shown in the inset of Fig. 13.19), we measured by X-ray photoelectron spectroscopy (XPS) larger barrier potentials at the C-polar surface than at the Si-polar surface, namely 1.81 and 0.72 eV, respectively [49]. This larger barrier indicates a much larger positive upward band bending at the Ga/C-polar SiC interface, implying a larger electron transfer to the Ga nanoparticles than at the Ga/Si-polar interface. Noteworthy, as a consequence of the charge transfer between the Ga nanoparticles and the SiC surfaces, the nanoparticles grow with a different shape on the two surfaces, as revealed by atomic force microscopy (AFM) measured at the end of the growth. Since the two polished faces of SiC have the same dielectric function and, hence, refractive index, and similar surface roughness and morphology, the only difference being the sign of the polarization, we can infer that the different density and size of the NPs assemblies and their plasmonic behavior is driven by the surface charge of the opposing polarity surfaces. The larger charge transfer between the C-face SiC (with negative surface polarization charge)
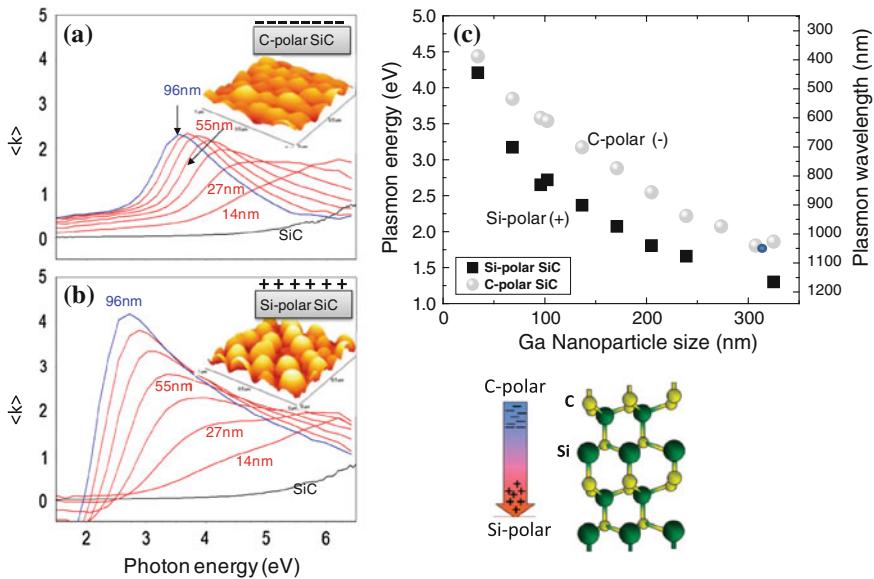
Fig. 13.19 Real-time evolution of the ellipsometric spectra of the pseudoextinction coefficient, $\langle k \rangle$, recorded during Ga NP formation on **a** C-polar and **b** Si-polar SiC (spectra are shown every 20 s with corresponding nanoparticles size –measured aposteriori); only the longitudinal SPR is visible in the investigated photon energy range. **c** Comparison of the plasmon energy peak position as a function of size for the two SiC polarities, which the inset shows having different surface charges. The corresponding final atomic force microscopy images are also reported

and Ga determines a stronger coupling for the Ga NPs/C-face SiC and, consequently, the metal cluster surface mobility is reduced while the contact area between the metal and the substrate is increasing, leading to flatter nanoparticles [50].

The impact of the interface charge transfer on the nanoparticle morphology and, therefore, on the plasmon resonance can be explained by the Young-Lippmann theory of electrowetting [51], which relates the wetting and, hence, the geometry of nanoparticles through the contact angle θ to the interface potential, U, according to

$$\cos\theta = \cos\theta_Y + f\left(U - U_{pzc}\right) \tag{13.15}$$

where the Young contact angle, θ_Y, is defined by $\cos\theta_Y = \frac{\sigma_{SiC} - \sigma_{SiC/Ga}\text{interf}}{\sigma_{Ga}}$ with σ_{SiC}, $\sigma_{Si/Ga}$, and σ_{Ga} being the surface free energies of the SiC semiconductor solid, SiC semiconductor/gallium interface, and of gallium, respectively. According to the Young-Lippman theory, the larger U_{pzc} potential measured at the interface with C-polar SiC favors wetting of Ga NPs on those surfaces, decreasing their surface mobility, and ultimately increasing the NPs density and decreasing their diameter, as schematized in Fig. 13.20.

(+) surface charge $\Longleftrightarrow$ Low U_{pzc} $\Longleftrightarrow$ High θ (-) surface charge $\longleftrightarrow$ High U_{pzc} $\longleftrightarrow$ Low θ

Fig. 13.20 Scheme of the effect of interface charge on the geometry of a semiconductor-supported nanoparticles, i.e. contact angle varies with Upzc (interface potential) strength (see text Eq. (13.1))

Therefore, the real-time monitoring by spectroscopic ellipsometry allows not only tuning the plasmon resonance wavelength/energy, but also explaining the different shape of nanoparticles forming on surfaces with the same refractive index.

13.6.2 Ga Nanoparticles Deposited on Graphene/SiC

As a latest development, graphene has also been reported to be a good electron transfer platform [52], and epitaxial graphene has been grown on both Si- and C-faces of SiC by the sublimation method [53]. Therefore, we also used spectroscopic ellipsometry to detect charge transfer between Ga NPs and graphene/C-face SiC and its impact on the plasmon resonance. Figure 13.21 compares the Ga nanoparticles plasmon resonance dynamics for SiC (see Fig. 13.21a) and graphene/SiC (see Fig. 13.21b); spectra are plotted with the same time resolution (i.e., every 1 s). The LSPR peak for Ga on graphene is observed to dampen and broaden with growth and is slightly blue-shifted (365 nm) in comparison to that on the SiC (390 nm) (see Fig. 13.21d). In the context of the Drude model, the plasmon peak becomes narrower and its intensity increases with increasing particle size due to the increased lifetime of the collective plasmon oscillation resulting from reduced electron-surface scattering relaxations (i.e., the 1/R effect, where R is the radius of the NPs). This can be clearly seen in the evolution of the LSPR for Ga nanoparticles on SiC. Conversely, broadening with nearly constant amplitude is observed for Ga nanoparticles on graphene, despite the NPs enlargement.

In the absence of a better analytical model, we can describe the surface plasmon resonance due to an ensemble of N_v absorption oscillators/unit volume (i.e., the Ga NPs) with a Lorentzian oscillator of the form

$$\varepsilon(\omega) = 1 + \frac{N_v e^2}{\varepsilon_0 m} \frac{1}{\omega_p^2 - \omega^2 + i\beta\omega} \tag{13.16}$$

The damping coefficient β mainly depends on electron- phonon interaction in all materials and is independent of size and distance distribution of the Ga NPs. The
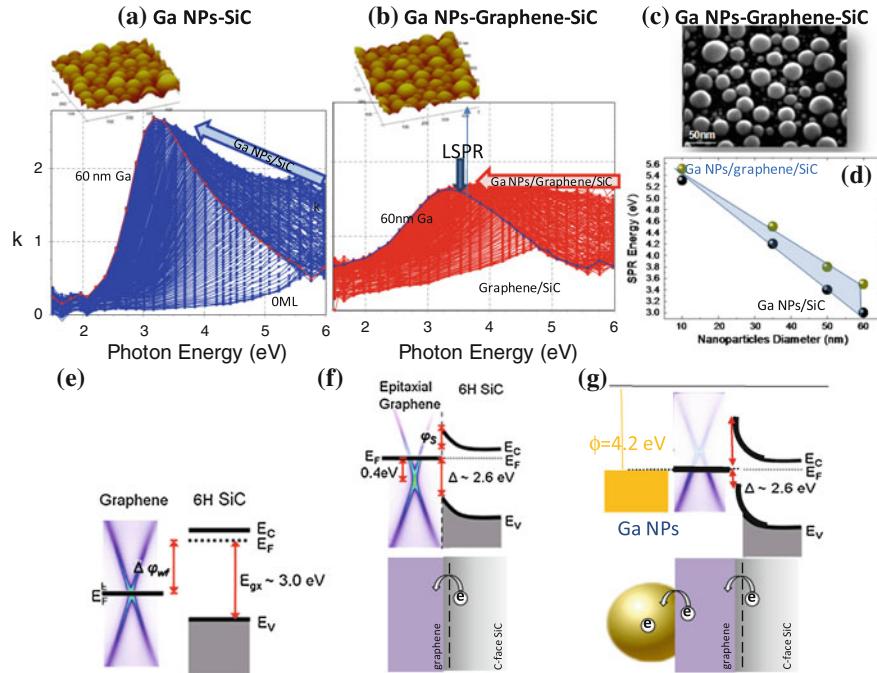
Fig. 13.21 Real-time ellipsometric spectra of the plasmon resonance of gallium nanoparticles grown directly on C-face SiC (**a**) and on graphene/SiC (**b**); the corresponding atomic force microscopy images of the nanoparticles morphology are shown, indicating that they are very similar. A representative tilted scanning electron microscopy image of the same Ga nanoparticles on graphene is also shown in (**c**). The blue shift of the SPR energy as a function of the nanoparticles size for the Ga/graphene/SiC compared to Ga/SiC is shown in (**d**). The sketch at the bottom shows (**e**) the bands and Fermi level EF for the isolated graphene and SiC, (**f**) the alignment of SiC bands with graphene, inducing the upward band bending due to the electron transfer from SiC to graphene lowering its work function and n-doping of graphene, and (**g**) the further electron transfer to gallium nanoparticles that have a work function of 4.2 eV

plasmon resonance is centered around $\omega_p = \frac{Ne^2}{\varepsilon_0 m}$, with N being the number of electrons per unit volume, m their effective mass and ε_0 the vacuum permittivity. We infer from Eq. (13.16) that an increase of ω_p by an enhancement of the electron density in the nanoparticles leads to a blueshift of the resonance.

It is important to point out the role of the SiC substrate in determining the electron transfer to the metal NPs. Specifically, the electron doping of graphene induced by SiC (i.e. charge transfer from SiC to graphene) causes a decrease of the work function of the graphene/SiC ensemble, which becomes approximately 4.0–4.2 eV [53]—lower than the work function of Ga (4.2 eV) and, therefore, activating transfer of electrons from graphene to Ga to ensure Fermi level equilibration between the nanoparticles and graphene/SiC, as schematized at the bottom of Fig. 13.21.

Thus, real-time spectroscopic ellipsometry allows investigating plasmonic coupling also on novel systems such as graphene/metal nanoparticles

13.7 Molecule-Metal Nanoparticle Interfacial Charge Transfer Absorption: Application to Nanoparticles Functionalization and Self-Assembled Monolayers—The Working Principle of a Molecular Sensor

Metal nanoparticles provided a new opportunity to study interfacial charge-transfer with molecules because of their size-dependent properties differing from the same metal in the bulk state. Among other characteristics described in the previous paragraphs and chapters, metal nanoparticles, because of their high surface-to-volume ratio, have many sites for adsorption of molecular donors/acceptors, so that the probability of charge-transfer absorption is increased relative to comparable adsorption on a bulk material.

As an example, for a spherical nanoparticle, the fraction, f_s, of atoms, N, on the surface is $f_s = 4/N^{2/3}$, and some numbers are given in the table in Fig. 13.21, depending on the size of nanoparticles, taking the well known gold as an example. From the electronic-optical point of view, also discreteness of the electronic states becomes important as the size of the metal nanoparticles decreases. Depending on the size of the nanoparticle, there may be a gap $E_{g,N}$ between filled and empty electronic levels as schematized in Fig. 13.22, (known as the Kubo gap) [54], given by $E_{g,N} = 4E_F/3N$, where E_F is the Fermi level of the bulk metal and N is the number of valence electrons. Some typical values for the most investigated gold nanoparticles are also given in the table of Fig. 13.21. This means that exciting nanoparticles will also lead to interband

N atoms	D_{sphere} (nm)	$E_{g,N}$ (meV)*	f_s
10^1	0.69	680	1.0
10^2	1.48	68	0.86
10^3	3.20	6.8	0.4
10^4	6.88	0.68	0.19
10^5	20	0.027	0.063
6×10^{23}	2.7cm	0	5×10^{-8}

* Calculated using E_F=5.1eV (for gold)
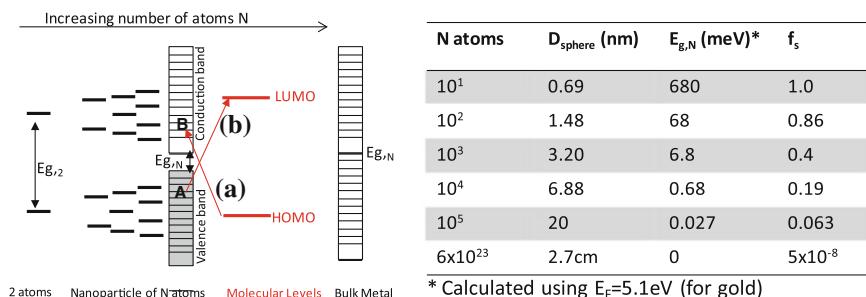
Fig. 13.22 Energy level scheme for a metal-molecule system. The evolution of the electronic levels from metal atoms to nanoparticles to bulk shows the decrease in the gap, $E_{g,N}$, between filled and empty states with the increase of number of atoms N. The (a)-*arrow* indicates the molecule-to-metal charge transfer, while the (b)-*arrow* indicate the charge transfer from the metal to the excited molecular level. The table summarizes number of atoms, diameter, gap and fraction of surface atoms for gold nanoparticles

transitions that can be size dependent [55], allowing the possibility to tailor the bandgap of nanoparticles to HOMO-LUMO levels of molecules in order to activate electron transfer, as schematized in Fig. 13.21. Specifically, considering a molecular redox couple in contact with a metal nanoparticle, light absorption may induce

(a) molecule-to-metal electron transfer: electron transfer from the reduced form of the molecule (Red), i.e., a molecular electron donor, to an empty energy level (B) of the metal nanoparticle, ie., Red-B( ) → Ox-B(e-)
(b) metal-to-excited molecule electron transfer: electron transfer from an occupied level A(e-) of the metal to the molecular electron acceptor (Oxidized, Ox) Ox-A(e-) → Red-A( )

If the metal shows plasmonic characteristics, the plasmon resonance of metal nanoparticles has important consequences when molecules are near the nanoparticles. When incident light excites coherent oscillations of the free electrons in metallic nanoparticles, inducing localized surface plasmon resonance (LSPR), a strong optical near-field decays exponentially away from the metal nanoparticle surface with length scales on the order of 10–200 nm [56] The molecule interacts with this electromagnetic field and absorption of photons can occur, leading to electronic excitations and vibrational excitations like the enhancement of Raman signals of nearby molecules, i.e., the SERS (surface enhanced Raman spectroscopy) effect [57, 58].

Due to the strong field-enhancement effects, surface plasmons can couple strongly to the molecular electronic transitions to form hybrid molecular-plasmon states. The LSPR can also interact with the excited states of the molecules, which results in phenomena such as shortening of molecular radiative lifetime, energy transfer, charge transfer, and mediation of molecular energy redistribution [59].

The electron transfer involving interband transitions and the plasmon induced electromagnetic enhancement are the basis of the chemical mechanism (CM) and electromagnetic mechanism (EM) of the SERS of molecules.

The EM is due to local electric fields in the surroundings of the metal nanoparticles which are enhanced due to the surface plasmon excitation, leading to more intense electronic transitions in molecules located near the nanoparticles.

The CM mechanism, also called the "charge transfer model" considers that the Fermi level of the metal lies between the molecular ground state and one or more of the excited states of the molecule, so that transitions from the Fermi level to the excited state of the molecule, which is photoreduced, or from the molecule ground state to the Fermi level (in this case the molecule is photo-oxidized) can be involved (see Fig. 13.22).

Here we show the use of real-time spectroscopic ellipsometry in real time to probe the interaction Ga nanoparticles-molecule, which is the basis of sensors and of SERS mechanisms.

For optical sensors, a major advantage is that ellipsometry, being a reflection technique, is sensitive to changes occurring on the sensor surface or in its near-surface region. Furthermore, the real-time advantage implies that continuous measurements are possible. As far as concerns sensitivity, most of research ellipsometers have a resolution better than 0.001° in Δ and Ψ, which corresponds to thickness changes
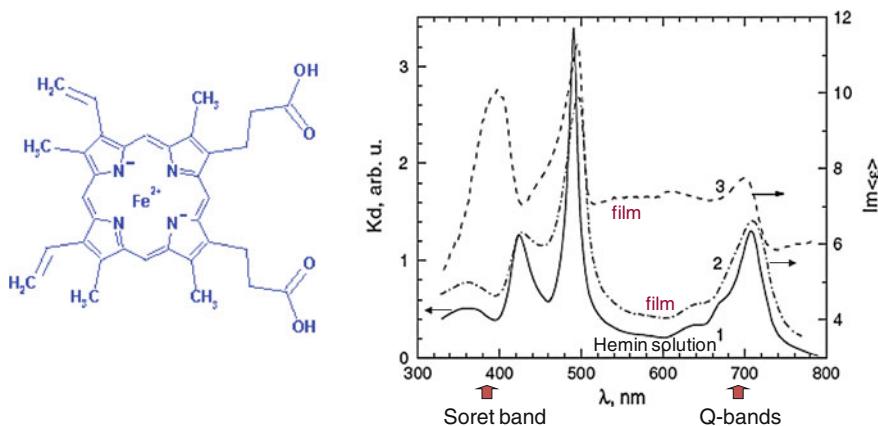
Fig. 13.23 Spectra of the molar extinction coefficient (*left*) and of the imaginary part of the dielectric function (*right*) of, respectively, hemin in solution and as self assembled monolayers. The hemin formula is also shown

of 0.0003 nm of an organic layer on a silicon surface in air and aqueous ambient. Experimentally, detection limits of 0.01 ng/ml for proteins [60] and of 10 ppm for gas sensor [61] have been reported in the approximation of a thin layer. On the other hand, exploiting the surface plasmon resonance of nanoparticles in SPR-based sensors, the intensity change and wavelength shift based on conventional SPR are limited to a refractive index unit (RIU) of 5×10^{-5} RIU and 2×10^{-5} RIUM, respectively, which are insufficient to monitor low concentrations of small biomolecular analytes. An improvement is achieved by SPR ellipsometry measuring the phase modulation of the reflected light, achieving 10^{-7} RIU.

One of the molecular systems that have been extensively exploited to fabricate sensors and to study molecule-nanoparticle charge transfer is alkanethiol due to the formation of robust self-assembled monolayers (SAMs) on gold nanoparticles surfaces [62–64]. Another class of molecular wires, porphyrin-based systems, has been studied due to high charge mobilities along the molecule. Porphyrins are aromatic, heterocyclic macrocycles, which typically have intense absorption maximum in the visible regime (see Fig. 13.23) [65, 66]. Porphyrins coupled to metallic nanostructures have been shown to possess properties of charge transfer, plasmon-enhanced electrical conduction, and electrocatalyic activity. Among porphyrins, hemin (iron(III) protoporphyrin IX chloride), which is often used as a model compound for the investigation of heme c-cofactor that is present in hemoproteins, has the Soret band at ~400 nm (3.4 eV) and four distinct Q-bands in the range from 515 to 646 nm (see Fig. 13.23). The relative intensity ratio of the Soret and Q-bands also depends on the interaction with the support, as shown in Fig. 13.22 by the comparison of spectra in solution and in the solid state. The hemin absorption can overlap significantly with the plasmon band of Ga nanoparticles, and this is important because when photoactive molecules possess absorption near the frequency of the LSPR, coupling is expected to be strong
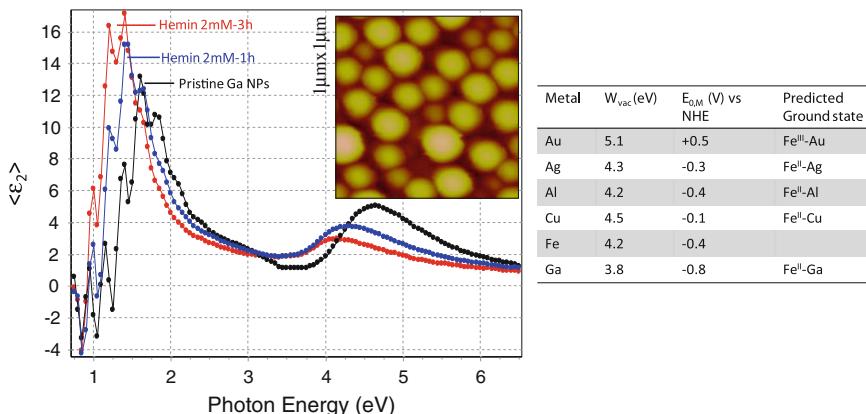
**Fig. 13.24** Ellipsometric spectra of the imaginary part, $\langle\varepsilon_2\rangle$, of pristine Ga nanoparticles and after various times of 1 and 3 h of functionalization by hemin 2 mM solution. The table reports for some common metals of interest, the work function, W_{vac}, the redox potential with respect to the hydrogen standard electrode and the iron predicted state when iron is coupled to the other metal. The inset shows the $1 \mu m \times 1 \mu m$ AFM topography of the Ga nanoparticles
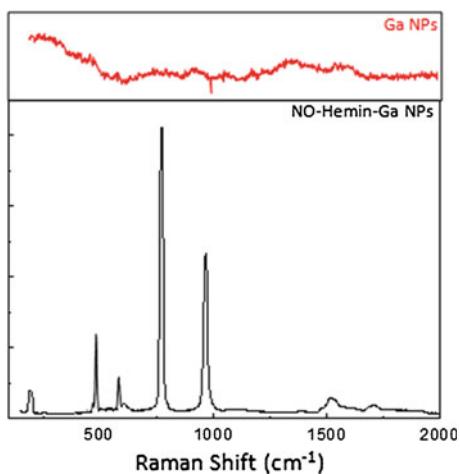
[67, 68]. Therefore, we discuss here hemin attached to Ga nanoparticles because of the tunability of their plasmon resonance.

Ga nanoparticles with the longitudinal (LO) SPR at 1.8 (below the onset of hemin absorption, i.e., the Soret band of hemin) and the transverse (TO) SPR mode at 4.8 eV (i.e., above the hemin absorption) are considered. Figure 13.24 shows the ellipsometric spectra of Ga nanoparticles at various time of dipping in a functionalizing solution of hemin. Coordinating molecules such as hemin causes the LO surface plasmon peak to shift to lower energy, and to increase in intensity. The increased low-energy absorption is due to light-induced electron transfer from the Ga nanoparticle to the adsorbate molecular layer derived from the direct interaction between the porphyrin π and Ga sp-orbitals, in other words, the partial charge transfer between the porphyrin and Ga.

In hemin, the charge transfer involves the metallic center Fe(III), and it can be explained by considering the work function and bulk particle reduction potentials for gallium and for Fe(III) also compared to other plasmonic metals, which are given in the table included in Fig. 13.24.

The adsorption of iron derivative (E° typically +0.4 V), like hemin, onto Au (E° +0.5 V) yields adsorbed Fe(III); conversely if, as in the present case, Au is replaced by Ga Nanoparticles (E°−0.8 V) [69] the molecular adsorbed is Fe(II) because of the reducing potential of Ga compared to Fe(III). This supports the electron transfer from the Ga nanoparticles to the hemin, with a net decrease of the electron concentration in the Ga nanoparticle volume, and, hence of the plasma frequency, of the Ga nanoparticles, leading to the red-shift of the plasmon resonance and increase of the absorption, as discussed in paragraph 2.
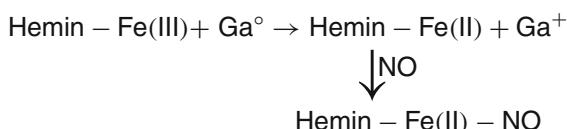
This electron transfer and plasmon coupling yield a strong resonant enhancement of the Raman signal of hemin. Indeed, we were able to observe the SERS spectra of hemin functionalizing the Ga nanoparticles. Bands characteristic of totally symmetric A_{1g} modes at $755\,cm^{-1}$ associated with B_{1g} modes of hemin have become dramatically enhanced by chemisorption on Ga nanoparticles (see Fig. 13.25).

A peculiarity of hemin is that it is known to react with nitric oxide (NO) resulting in substantial changes in the metal coordination environment and electronic structure, serving therefore as a source of NO optical sensing. The formation of the axially bound NO-hemin adduct affects modes localized on the chelating ring and that these modified modes can serve as markers for the presence of the adduct. It has been predicted that excitation of charge-transfer bands should allow resonance of iron-nitrogen stretching modes because both the transition moments of these bands and the vibrational modes are perpendicular (or *z*-polarized) relative to the porphyrin plane [70]. Consequently, the band observed at 475 and $577\,cm^{-1}$ in hemin exposed to NO must have charge-transfer character. The shift to $200\,cm^{-1}$ of the Fe(II)-imidazole stretch, which is expected between 200 and $250\,cm^{-1}$ also suggests an appreciable weakening of the Fe-imidazole bond because of the electron transfer from Fe to NO [71].

Once Ga nanoparticles have reduced Fe(III) to Fe(II) state, the reduced Fe(II) readily binds the NO molecule that has an unpaired electron: NO bonds to Fe(II) with a greater affinity than to Fe(III). Upon formation of the NO-Fe bond, electrons form an electron pair and the FE-NO nitrosyl complex develops a stable closed-shell character, as schematized by the sequence below

$$\text{Hemin} - \text{Fe(III)} + \text{Ga}° \rightarrow \text{Hemin} - \text{Fe(II)} + \text{Ga}^+$$
$$\downarrow \text{NO}$$
$$\text{Hemin} - \text{Fe(II)} - \text{NO}$$
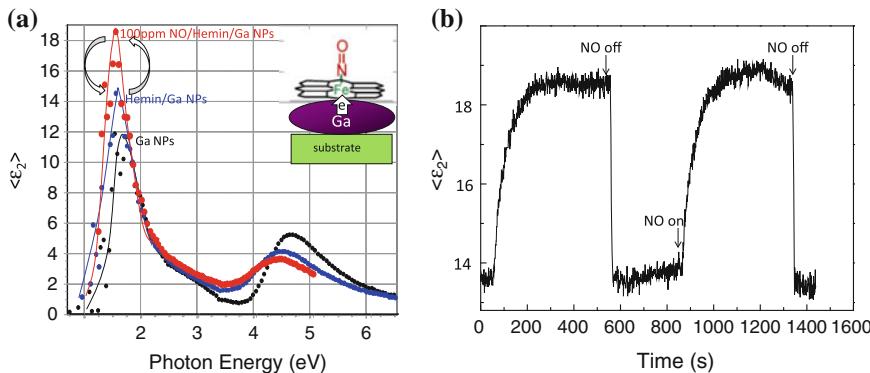
Fig. 13.26 a Ellipsometric spectra of the imaginary part, $\langle\varepsilon_2\rangle$, of pristine Ga nanoparticles and after functionalization by hemin solution 2 mM and further exposure to 100 ppm of NO. The inset sketches the electron transfer from the Ga nanoparticles to reducing Fe(III) to Fe(II) that, therefore, binds NO. **b** $\langle\varepsilon_2\rangle$ time evolution monitored at the maximum of the plasmon resonance ($\sim$1.6 eV) upon cycles of exposure to NO gas

Therefore, electrons are transferred from the Ga Nanoparticles to the NO binding, causing a further red-shift of the Ga NPs plasmon resonance, as shown in Fig. 13.26a. The energy shift and the amplitude variation of the plasmon resonance detected by spectroscopic ellipsometry provide enough variables and sensitivity to sensing the NO gas optically. Fig. 13.26b shows the variation of $\langle\varepsilon_2\rangle$ during cycles of exposure to NO and the reversibility of the ellipsometric signal by flowing in and out a gas cell the NO gas.

The plasmon resonance energy and amplitude can change by refractive index or electron transfer, however, the change in refractive index by NO is negligible, and therefore the observed variation can be ascribed to electron transfer activated by process in Eq. (13.16). On the basis of these ellipsometric observations, we can assume that the formation of the axially bound NO-hemin adduct affects modes localized on the chelating ring due to the charge transfer inducing variation of the coordination geometry of Fe center, and that these changes can serve as sensor for the presence of the NO, demonstrating how the charge transfer and its impact on the plasmon resonance, which can be probed by ellipsometry, constitutes a diagnostic methodology to study and develop sensors.

13.8 Conclusions and Outlook

The electrodynamic coupling of photons and electrons via localized surface plasmons has been intensively studied for over 100 years. Nevertheless, an appreciation for the material design considerations needed to fully harness this coupling has only begun to emerge.

Investigations on the charging of nanoparticles have led to the discovery that the plasmon resonance (and therefore the light absorption and scattering) is sensitive to charging events. These two sets of phenomena, which form the core of plasmo-electronic effects, suggest that a complete description of the electronic behavior and plasmonic response of a material requires knowledge of not only nanoparticle shape, size, and dielectric environment, but also of charge.

The examples presented in this chapter show the unique properties of real-time spectroscopic ellipsometry in monitoring interface phenomena at the nanoscale which involve charge transfer from/to an ultrathin layer and or nanoparticle of a metal to/from semiconductors and/or organic molecules.

From a more fundamental perspective, highlighting these plasmoelectronic materials provide new insights into the coupling of light with charge in complex heterogeneous materials. As highlighted by the studies presented herein, significant progress has been made in recent years in understanding the coupling of plasmons with net flows of charge carriers, which has resulted in improvement in growth methodologies for semiconductors, in novel plasmonics systems growth and in sensors with improved performance.

In spite of the significant progress, however, we feel that the field of plasmoelectronics is yet in its infancy and that fundamental insights as well as a full exploration of potential applications are yet to be revealed.

From the ellipsometric perspective, real time spectroscopic ellipsometry is now well established as interactive diagnostic tool, and is making a major contribution to the study of surfaces and interfaces phenomena in general, and of nanostructure processing in particular. Significant advantages of real time spectroscopic ellipsometry have been demonstrated: all pressure ranges during growth processes are accessible; transparent as well as opaque substrates can be monitored; insulators can be studied without the problem of charging effects, and, therefore, no substrate/sample specific preparation is required; buried interfaces are accessible; sub-second temporal resolution for kinetic studies can be achieved.

Enabling nanoscale materials real-time analysis under realistic conditions is a critical need. Ellipsometry provides a unique capability to measure in real time under realistic environments; therefore, ellipsometry real time application has paramount opportunities to speed the rate at which new information regarding the chemical and physical behavior of nanomaterials could be gained thus enabling optimization of manufacturing processes.

Therefore, the real-time spectroscopic ellipsometric approach can now be generalized to an extensive array of device configurations utilizing organic as well as inorganic compounds, on thin films, single crystals, and polycrystalline substrates, producing a variety of molecule-nanoparticle combinations with unique and novel properties. The extension of this work with other optical active molecules, chemistries, and surfaces provides an unlimited number of variables to create varied hybrid device structures with diverse property sets.

In the next few years, a major effort is expected in the exploration of laboratory best practices for modelling data and the harmonization of preliminary stage development of standards for ellipsometry characterization, as there is now a wealth of

high-quality experimental data available on nanostructures. A vital element in the development of effective and validated protocols (or standard practices) for application of ellipsometry especially to nanomaterials is the conductance of interlaboratory studies, commonly referred to as round robin testing.

These studies are necessary to provide estimates of measurement precision at the nanoscale, and to ensure that modelling procedures generate accurate and meaningful results while avoiding potential artefacts.

References

1. S. Chen, R.S. Ingram, M.J. Hostetler, J.J. Pietron, R.W. Murray, T.G. Schaaff, J.T. Khoury, M.M. Alvarez, R.L. Whetten, Science **280**, 2098 (1998)
2. A.C. Templeton, W.P. Wuelfing, R.W. Murray, Acc. Chem. Res. **33**, 27 (2000)
3. D.M. Adams, J. Phys. Chem. B **107**, 6668 (2003)
4. S.A. Maier, *Plasmonics: Fundamentals and Applications* (Springer, New york, 2007)
5. S.C. Warren, D.A. Walker, B.A. Grzybowski, Langmuir **28**, 9093 (2012)
6. A. Takai, P.V. Kamat, ACS Nano **5**, 7369 (2011)
7. H.A. Atwater, A. Polman, Nat. Mater. **9**, 205 (2010)
8. A. Wood, M. Giersig, P. Mulvaney, J. Phys. Chem. B **105**, 8810 (2001)
9. M. Jakob, H. Levanon, P.V. Kamat, Nano Lett. **3**, 353 (2003)
10. V. Subramanian, E.E. Wolf, P.V. Kamat, J. Am. Chem. Soc. **126**, 4943 (2004)
11. A. Wood, M. Giersig, P. Mulvaney, J. Phys. Chem. B **105**, 8810 (2001)
12. T. Hirakawa, P. V. Kamat, Langmuir **20**, 5645 (2004)
13. R.W. Murray, Chem. Rev. **108**, 2688 (2008)
14. J.B. Asbury, E. Hao, Y. Wang, H.N. Shosh, T. Lian, J. Phys. Chem. B **105**, 4545 (2001)
15. F.J. Adrian, J. Chem., Phys. **77**, 5302 (1982)
16. K.G. Thomas, P.V. Kamat, Acc. Chem. Res. **36**, 888 (2003)
17. P.V. Kamat, S. Barazzouk, S. Hotchandani, Angew. Chem., Int. Ed. **41**, 2764 (2002)
18. D. E. Aspnes, Phys. Stat. Sol. (b) 242, 2551 (2005)
19. R.W. Collins, I. An, J. Lee, J.A. Zapien, Multichannel ellipsometry, in *Handbook of Ellipsometry*, ed. by H.G. Tompkins, E.A. Irene (Springer, Heidelberg, 2005)
20. J.A. Zapien, R.W. Collins, R. Messier, Rev. Sci. Instrum. **71**, 3451 (2000)
21. D. Daineka, D. Kouznetsov, P. Bulkin, G. Girard, J.-E. Bour'ee, B. Drevillon, Eur. Phys. J. Appl. Phys. **28**, 343–346 (2004)
22. J. Lee, P.I. Rovira, I. An, R.W. Collins, Rev. Sci. Instrum. **69**, 1800 (1998)
23. C. Chen et al., Appl. Surf. Sci. **253**, 38–46 (2006)
24. E. Garcia-Caurel, A. De Martino, B. Drévillon, Thin Solid Films **455–456**, 120 (2004)
25. N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Harcourt College Publisher, Philadelfia, 1976), pp. 1–20
26. C. Novo, A.M. Funston, A.K. Gooding, P. Mulvaney, J. Am. Chem. Soc. **131**, 14664 (2009)
27. A. Henglein, P. Mulvaney, T. Linnert, Faraday Discuss. **92**, 31 (1991)
28. R. Chapman, P. Mulvaney, Chem. Phys. Lett. **249**, 358 (2001)

29. N.I. Zheludev, Contemp. Phys. **43**, 365 (2002)
30. O. Hunderi, R. J. Ryberg, Phys. F Met. Phys. **4** 2084, 2096 (1974)
31. V. Albanis, S. Dhanjal, N.I. Zheludev, P. Petropoulos, D.J. Richardson, Opt. Express **5**, 157 (1999)
32. X.F. Li, G.T. Fei, X.M. Chen, Y. Zhang, K. Zheng, X.L. Liu, L.D. Zhang, Europhys. Lett. **94**, 16001 (2011)
33. P.C Wu, T.H. Kim, A.S. Brown, M. Losurdo, G. Bruno, H. O. Everitt, Appl. Phys. Lett. **90**, 103119 (2007)
34. P.C. Wu, C.G. Khoury, T.H. Kim, Y. Yang, M. Losurdo, G.V. Bianco, T. Vo-Dinh, A.S. Brown, H.O. Everitt, J. Am. Chem. Soc. **131**, 12032 (2009)
35. J.R. Ahn, P.G. Kang, K.D. Ryang, H.W. Yeom, Phys. Rev. Lett. **95**, 196402 (2005)
36. D. Eom, S. Qin, M.Y. Chou, C.K. Shih, Phys. Rev. Lett. **96**, 027005 (2006)
37. J. Neugebauer, T.K. Zywietz, M. Scheffler, J.E. Northrup, H. Chen, R.M. Feenstra, Phys. Rev. Lett. **90**, 056101 (2003)
38. K.J. Laidler, *Chemical Kinetics* (Pearson Education, New York, 1987)
39. Th. Wolkenstein, O. Peshev, J. Catal. **4**, 301 (1965)
40. T. Wolkenstein, *Electronic Processes on Semiconductor Surfaces During Chemisorption* Chap. 3, (Consultants Bureau, New York, 1991), pp. 83–124
41. W. Göpel, G. Rocker, J. Vac. Sci. Technol. **21**, 389 (1982)
42. S. Choi, T.H. Kim, H.O. Everitt, A. Brown, M. Losurdo, G. Bruno, A. Moto, J. Vac. Sci. Technol. B **25**, 969 (2007)
43. S. Choi, T.H. Kim, P.C Wu, A. Brown, H.O. Everitt, M. Losurdo, G. Bruno, **27**, 107 (2009)
44. G. Bruno, M. Losurdo, T.H. Kim, A.S. Brown, Phys. Rev. B **82**, 075326 (2010)
45. A.C. Templeton, W.P. Wuelfing, R.W. Murray, Acc. Chem. Res. **33**, 27 (2000)
46. W. Pfeiffer, C. Kennerknecht, M. Merschdorf, Appl. Phys. A **78**, 1011 (2004)
47. J.Y. Bigot, V. Halte, J.C. Merle, A. Daunois, Chem. Phys. **251**, 181 (2000)
48. S. Link, M.A. El-Sayed, J. Phys. Chem. B **103**, 8410 (1999)
49. J.R. Waldrop, R.W. Grant, Y.C. Wang, R.F. Davis, J. Appl. Phys. **72**, 4757 (1992)
50. P.C. Wu, M. Losurdo, T.H. Kim, M. Giangregorio, G. Bruno, H.O. Everitt, A.S. Brown Langmuir **25**, 924 (2009)
51. F. Mugele, J.C. Baret, J. Phys. Conden. Matter R705 (2005)
52. X. Zuo, S. He, D. Li, C. Peng, Q. Huang, S. Song, C. Fan, Langmuir, **26**, 1936–1939 (2010)
53. C. Berger, Z. Song, X. Li, X. Wu, N. Brown, C. Naud, D. Mayou, T. Li, J. Hass, A.N. Marchenkov, E.H. Conrad, P.N. First, W.A. de Heer, Science **312**, 1191 (2006)
54. R.L. Johnston, Phil. Trans. R. Soc. Lond. A **356**, 211 (1998)
55. M. Losurdo, M.M. Giangregorio, G.V. Bianco A.A. Suvorova, C. Kong, S. Rubanov, P. Capezzuto, J. Humlicek, G. Bruno, Phys. Rev. B **82**, 155451 (2010)
56. U. Kreibig, M. Vollmer, *Optical Properties of Metal Clusters*, vol. 25(Springer, Berlin, 1995)
57. P.C. Lee, D. Meisel, J. Phys. Chem. **86**, 3391 (1982)
58. S. Nie, S.R. Emory, Science **275**, 1102 (1997)
59. U. Kreibig, M. Gartz, A. Hilger, H. Hovel, Mie-plasmon spectroscopy: a tool of surface science, in Fine Particles *Science and Technology* ed. by E. Pelizzatti (Kluwer Academic Publishers, Boston, 1996), pp. 499–516
60. R.M. Ostroff, D. Hopkins, A.B. Haeberli, W. Baouchi, B. Polisky, Clin. Chem. **45**, 1659 (1999)
61. S. Zangooie, R. Bjorklund, H. Arwin, Sens. Actuators B **43**, 168 (1997)
62. H. Nakanishi, K.J.M. Bishop, B. Kowalczyk, A. Nitzan, E.A. Weiss, K.V. Tretiakov, M.M. Apodaca, R. Klajn, J.F. Stoddart, B.A. Grzybowski, Nature **460**, 371 (2009)
63. G. Bruno, F. Babudri, A. Operamolla, G.V. Bianco, M. Losurdo, M.M. Giangregorio, O. Hassan Omar, F. Mavelli, G.M. Farinola, P. Capezzuto, F. Naso, Langmuir **26**, 8430 (2010)
64. M.M. Giangregorio, M. Losurdo, G.V. Bianco, A. Operamolla, E. Dilonardo, A. Sacchetti, P. Capezzuto, F. Babudri, G. Bruno, J. Phys. Chem. C **115**, 19520 (2011)
65. K. Burrell, D.L. Officer, P.G. Plieger, D.C.W. Reid, Chem. Rev. **101**, 2751 (2001)
66. A. Tsuda, A. Osuka, Science **293**, 79 (2001)

67. I. Carmeli, I. Lieberman, L. Kraversky, Z.Y. Fan, A.O. Govorov, G. Markovich, S. Richter, Nano Lett. **10**, 2069 (2010)
68. A. Yoshida, N. Kometani, J. Phys. Chem. C **114**, 2867 (2010)
69. C. Creutz, B.S. Brunschwig, N. Sutin, Chem. Phys. **324**, 244 (2006)
70. A. Desbois, M. Lutz, R. Banerjee, Biochemistry **18**, 1510–1518 (1979)
71. G. Smulevicht, T.G. Spiro, J. Phys. Chem. **89**, 5168 (1985)

Chapter 14
Polarimetric and Other Optical Probes for the Solid–Liquid Interface

Kurt Hingerl

Abstract The rise of nanoscience and nanotechnology has induced a remarkable change in the science of liquid surfaces and interfaces: from a macroscopic and thermodynamic oriented approach the field is now steadily moving towards a microscopic and atomistic understanding. However, the need to measure, monitor, understand and ultimately control (electro-) chemical and physical processes occurring at liquid interfaces and surfaces cannot been simply fulfilled by transferring the highly successful electron beam or ion beam techniques into the liquid environment due to the limited free path of electrons or ions. Optical techniques, especially polarimetric techniques, have experienced a renaissance in the surface science of the liquid–solid interface and will become even more important in the future. Despite the penetration depth of light is for all materials at least of the order of 100 nm or above, surface sensitive optical probes have been developed, capable of monitoring processes on the nanosecond scale with thickness resolution of less than a monolayer. On the first sight, however, the major disadvantage of optical probes, being of indirect nature compared to electron microscopy or spectroscopy, proves then to be a fascinating feature, because all polarimetric, linear or nonlinear intensity and sum frequency, or scattering optical probes require modeling the full system for understanding the interactions between the liquid and solid at the interface. In the following we will demonstrate, after an introduction to electrochemistry, with a few selected examples the power of polarimetric techniques for understanding the liquid interface and review prior work especially for the electrochemical interface by polarimetric probes. The importance of correlation measurements, mainly current voltage measurements and scanning probe techniques for a thorough understanding is finally highlighted.

K. Hingerl (✉)
Center for Surface- and Nanoanalytics, Johannes Kepler University,
Altenbergerstr. 69, 4040 Linz, Austria
e-mail: kurt.hingerl@jku.at

14.1 Introduction

When the properties of liquids and solutions are considered, attention is normally focused on the bulk of the phase, and the properties of the system at its boundaries are ignored. These boundaries, however, exhibit significant effects and understanding their influence is a prerequisite for exploiting technological important processes. The recent nanoscience and nanotechnology research aiming at a microscopic and atomistic understanding of the liquid–solid interface has fostered a different approach to liquid surfaces and interfaces: instead of working with thermodynamic relations, microscopic approaches for measuring, monitoring and ultimately controlling the liquid–solid interface have increased the use and the development of in-situ techniques, scanning probe techniques, as well as optical ones. Furthermore much of the current knowledge about the structure and properties of electrode-electrolyte interfaces has been obtained from a combination of electrochemical in-situ and ex-situ spectroscopic methods. Ex-situ electron- based techniques, such as Auger electron, and X-ray and UV photoelectron spectroscopies, low-energy electron diffraction, have been used extensively to characterize electrodes before and after electrochemical measurements especially with single-crystal metal surfaces (for applications of ex-situ electron-based techniques to the study of electrochemical interfaces, see Ref. [1]). The interpretation of results obtained with these powerful ultra high vacuum (UHV) analytical methods, however, particularly in the case of post-electrochemical examination, must be performed with caution. This is due to possible changes in the surface resulting from the loss of potential control and subsequent exposure to vacuum and electron and photon beams. UHV compatible surface analysis methods refer to those experimental techniques that give chemical information on surfaces to a depth of 2–3 nm and even less. These analysis methods like X-ray photoelectron spectroscopy (XPS) or ultraviolet photoelectron spectroscopy (UPS) or scanning Auger electron spectroscopy (AES) can be combined with ion sputtering, which removes atomic layers from the surface and makes it thereby possible to measure concentration variations over a depth of tens or even hundreds of nanometers.
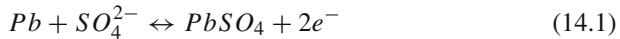
The lack of suitable surface science measurements for the liquid–solid interface leads to the use of optical and X-ray techniques. The optical ones will be discussed in the subsequent sections of this chapter with a special emphasis on linear polarization optical probes. It is out of the scope of this chapter to review the experimental realizations of spectroscopic ellipsometry and related techniques. In the following we will refer to other chapters of this book (especially Chap. 1) and to previous books, especially the 1987 edition of *Ellipsometry and Polarized Light* by Azzam and Bashara [2] for a general introduction to spectroscopic ellipsometry (SE) and Müller Matrix polarimetry (MMP) and to the different chapters of the *Handbook of Ellipsometry*, edited by Tompkins and Irene [3]. For further information on reflectance difference/anisotropy spectroscopy (RDS/RAS) the reader should consult the chapter by Zahn on this technique in the book [4]. An introduction to Raman spectroscopy can be found in the books [5, 6].

Just to name a few applications, electrochemical deposition and corrosion, batteries, charge storage and charge transfer, etching of semiconductors, etc. are major commercial processes, which benefit from a basic understanding of the liquid–solid interface. To start with the first, corrosion reaction rates are strongly dependent on surface adsorption, oxidation, and (electrolyte/surface) segregation. These phenomena can drastically modify the chemical composition of surfaces relative to that of the bulk and consequently influence the reactivity and the corrosion behavior of metals. From a kinetic point of view, three types of corrosion reactions, depending on the rate-limiting step, have been identified [7]: (a) Corrosion controlled by the kinetics of a charge-transfer reaction, either anodic or cathodic, at the metal-electrolyte interface. An example is the corrosion of steel in acids. (b) Corrosion controlled by the rate of mass transport of the oxidizing agent or of anodic reaction products. A well-known example is the corrosion of steel in neutral, aerated solution. (c) Corrosion controlled by the properties of the passive film. The reaction is then under anodic control and the average corrosion rate is often quite small. Stainless steel in aqueous solutions behaves in this manner. Therefore, there is a clear need for in-situ compatible techniques, because only they can identify the relevant time scales, the rate limiting steps, surface phase transitions or a present hysteresis behavior.
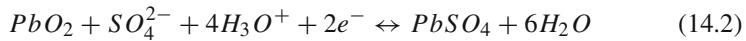
An industrial realization of electrochemical deposition in microelectronics is the damascene process, i.e. depositing/interlaying copper. This is becoming more and more important as the number of interconnect levels for logic devices has substantially increased due to the large number of transistors that are now interconnected. Connected with a change in the wiring material from aluminium to copper and from the silicon dioxide to new low-k materials a performance enhancement comes at a reduced cost via damascene processing that eliminates processing steps (through silicon vias). In damascene processing, in contrast to subtractive aluminium technology, the dielectric–insulating-material is deposited first as a blank film, which is then patterned and etched leaving holes or trenches. In single damascene processing, poly-crystalline copper is deposited in the holes or trenches surrounded by a thin barrier film resulting in filled vias or wire lines. In dual damascene technology, both the trench and the via are fabricated before the deposition of copper resulting in formation of both the via and line simultaneously, further reducing the number of processing steps. A thin barrier film, called copper barrier seed, is necessary to prevent copper diffusion into the dielectric. As the presence of excessive barrier film competes with the available copper wire cross section, formation of the thinnest continuous barrier represents one of the greatest ongoing challenges in copper processing [8]. The structural description of these systems at a molecular level can be used to control reactions at interfaces.

The final application discussed in the introduction is the direct conversion of electric energy into chemical energy, either by primary (single use) or secondary batteries (chargeable) or fuel cells. The oldest form of rechargeable battery is the lead—acid battery, which is displayed via the cell diagram $Pb(s)/PbSO_4//PbO_2/PbSO_4$. The nomenclature is the following: left to the double forward slash // the *anode* is indicated, where *oxidation takes place*. The first right forward slash represents the
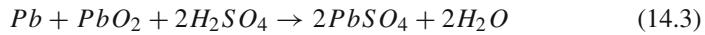
liquid–solid interface between Pb(s) and the oxidation product $PbSO_4$ at the anode. To the right of the double forward slash // the cathode reaction is given. The respective chemical decharging reaction at the anode, establishing for this half cell a potential difference of -0.36 V, is written as

$$Pb + SO_4^{2-} \leftrightarrow PbSO_4 + 2e^- \qquad (14.1)$$

The cathode is indicated to the right of the double forward slash, with its liquid–solid interface between the oxidation product PbO_2, which is reduced to $PbSO_4$ by the decharging reaction, establishing for this half cell $+1.68$ V,

$$PbO_2 + SO_4^{2-} + 4H_3O^+ + 2e^- \leftrightarrow PbSO_4 + 6H_2O \qquad (14.2)$$

Summing both, providing an electromotive force of 2.04 V the total reaction yields

$$Pb + PbO_2 + 2H_2SO_4 \rightarrow 2PbSO_4 + 2H_2O \qquad (14.3)$$

where the arrow points to the right when decharging, and to the left, when charging the battery. The lead acid battery is rather old and also heavy for the amount of electrical energy it can supply. Furthermore, corrosion of the solid PbO_2 electrode limits the number of charging cycles.

The field of the liquid–solid interface is so rich and broad and interdisciplinary—and treated by physicists, chemists, materials scientists and biologists—such that it is impossible to cover all aspects of electrochemistry. On the other hand, physicists and optical engineers are hardly trained in the foundations of electrochemistry. Therefore this chapter repeats in the next section definitions from electrochemistry (Sect. 14.2) and the nature of and models for the electrolyte interface (Sect. 14.3). After giving the basics, the subsequent section deals with adsorption processes (Sect. 14.4) and in a rather short presentation non-equilibrium processes are reviewed in Sect. 14.5. Here the focus is mainly on optical techniques and the kinetic information gained from them. However, the major progress arose from traditional electrochemical correlation measurements, mainly techniques as cyclovoltammetry (CyV) and electrochemical impedance spectroscopy (ECIS), but also atomic force microscopy (AFM) and especially electrochemical scanning tunneling microscopy (ECSTM), which are all reviewed in Sect. 14.6. Finally, in Sect. 14.7 we review linear polarization optical techniques and their findings. We summarize and provide an outlook in Sect. 14.8, where a recommendation and warning is issued on the use of SE and related techniques: all the optical techniques mentioned above yield a wealth of indirect information since they require modeling the (effective) dielectric function and thereby the processes occurring at the interface. Therefore, polarization optical techniques can, on the one hand, provide otherwise unaccessible data, on the other hand can also easily lead to erroneous conclusions.
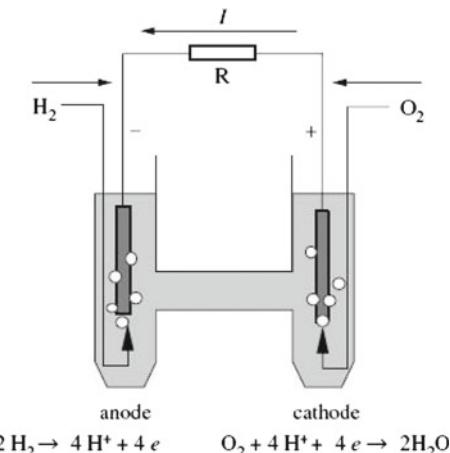
Fig. 14.1 H_2/O_2 fuel cell. At the anode (*left* electrode) oxidation occurs (supply of electrons) which generate an external (negative) current. At the cathode the supplied material (oxygen) is reduced. The indicated anode reaction is a shorthand notation for $4H_2O + 2H_2 \leftrightarrow 4H_3O^+ + 4e^-$, the indicated cathode reaction is a shorthand notation for $O_2 + 4H_3O^+ + 4e^- \leftrightarrow 6H_2O$. Reproduced with permission by EPFL Press from Ref. [7]

14.2 Basics of Electrochemistry
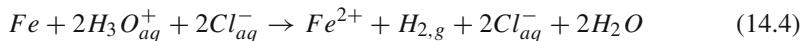
An electrochemical cell consists of two electrodes in contact with an electrolyte and connected by an electric conductor. Corrosion cells, batteries and fuel cells are examples of electrochemical cells. As an illustration, in Fig. 14.1 an H_2/O_2 fuel cell (defined by an external flow of gases providing chemical energy, which is transformed into electrical energy) is shown. The anode is defined as the electrode, where oxidation takes place, this is for the case of Fig. 14.1, the one where hydrogen is oxidized (electrons are taken off the hydrogen) to $H^+(H_3O^+)$. Oxidation means that electrons are provided to the external wire, the ionized hydrogen is soluted in the electrolyte and the electrons are leaving the electrode and flowing through the wire and provide a current. At the cathode oxygen is reduced (i.e. electrons are taken up) and reacts to H_2O. For each two moles of H_2 four electrons are provided and with one mole of O_2, two moles of water are produced.

Because chemical energy and electrical energy can be (in principle) fully converted into each other, the process of producing current by material flow and subsequent reaction(s) can be reversed by producing an electrical power and splitting a chemical compound into its parts; this is called electrolysis. In an electrolysis cell, the anode is the positive pole, whereas in a fuel cell or battery it is the negative one.
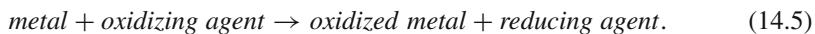
Faradays law states: when n_i moles of a given substance react, a proportional electric charge Q passes across the electrode-electrolyte interface: $Q = nFn_i$, where F is the Faraday constant $F = 96,485\,\text{C/mol}$ and the charge number n, without units, expresses the stoichiometric coefficient of the electrons in the equation for the

electrode reaction. For example, for the reaction describing the anodic dissolution of iron Eq. (14.4), $n = 2$.

In addition, electrochemical processes occur (often!) without the presence of an explicitly present second electrode: for corrosion, e.g. an oxidation-reduction (redox) reaction between the metal and an oxidizing agent present in the environment takes place. Therefore the oxidation of the metal is inseparably linked to the reduction of the oxidizing agent. If metallic iron is dropped into a bottle of an acidic (chlorinated) water, the positive charge of the H_3O^+ is transferred to the Fe, which is dissolved in the water.

$$Fe + 2H_3O_{aq}^+ + 2Cl_{aq}^- \rightarrow Fe^{2+} + H_{2,g} + 2Cl_{aq}^- + 2H_2O \qquad (14.4)$$
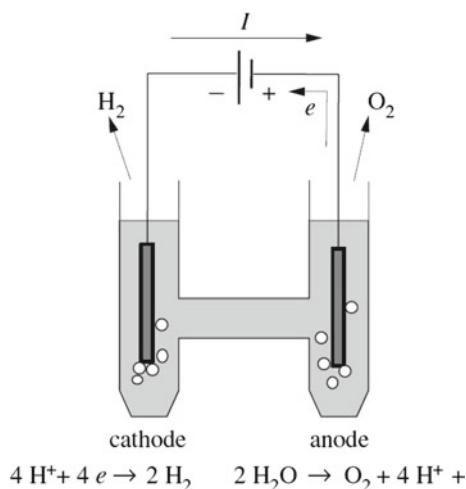
with the oxidizing agent as the solvated proton H_3O^+. The state of chlorine is unaffected, therefore the reaction can also be written as $Fe + 2H_{aq}^+ \rightarrow Fe^{2+} + H_{2,g}$. This reaction is reversible. In neutral or alkaline conditions oxygen becomes the oxidizing agent for iron forming rust: $4Fe + 3O_2 + 2H_2O \rightarrow 4FeOOH$. This rust (salt) is partially dissolved and the reaction becomes *thereby irreversible*. In a more general form, both corrosion reactions can be written in the form

$$metal + oxidizing\ agent \rightarrow oxidized\ metal + reducing\ agent. \qquad (14.5)$$

Despite on a first view, the schematic fuel cell and the electrolysis cell shown in Figs. 14.1 and 14.2 function independently of the electrode material, it is important which materials (metals) are used as electrodes, because the surface structure as well as catalytic activities and kinetic effects drastically change the technically achievable efficiency. The open circuit voltage, which builds up in the fuel cell reaction in Fig. 14.1 is $+1.229$ V, if oxygen and hydrogen are bubbled over platinum electrodes.

Fig. 14.2 Electrolysis cell for water. Again, at the anode (now the *right* electrode) oxidation occurs, producing besides electrons also H^+ (H_3O^+). Reproduced with permission by EPFL Press from Ref. [7]

cathode

anode

$4\ H^+ + 4\ e \rightarrow 2\ H_2$ $2\ H_2O \rightarrow O_2 + 4\ H^+ +$

The equilibrium potential or reversible potential of an electrode is defined as the difference in electric potential between the metal and the solution at equilibrium. It is named E_{rev} under standard conditions ($p = 1013$ mbar, $T = 25\,^\circ$C) E^0. Potentials cannot be measured absolutely, only potential differences between two electrodes in an electrochemical cell are measurable. By convention, a scale of standard electrode potentials, i.e. the open circuit voltage, which builds up between two different chemical electrodes due to their ionic solution in water, has been defined by arbitrarily assigning the value of zero to the equilibrium potential of the hydrogen electrode under standard conditions $2H^+ + 2e^- = H_2$ (normal (or standard) hydrogen electrode- *nhe, she*). A list of standard potentials of electrode reactions can be found e.g. in Ref. [9]. It is clear that these potentials depend on ambient conditions (p, T) as well as on the pH value and possible additional solvents. Taking into account the solvation energy of protons, the energy gained by the reaction of the gaseous proton with the electron, and the energy gained by forming H_2, a relation between *nhe* and the work function can be derived. For hydrogen, the vacuum level is 4.44 eV above the *nhe* level.

Having an electrochemical cell with two different electrodes in contact with a common electrolyte, the free reaction enthalpy ΔG_0 can be expressed by the number (of moles) of electrons n, and the amount and activity of the reactants and reaction products, called B_i. The relation between the electrode potential, the temperature and the stochiometric coefficients as a function of the activity (concentration) of the ions is called Nernst's equation. (Its derivation is repeated to facilitate for the reader without electrochemistry background the concepts and nomenclature). A chemical reaction is usually denoted with the stoichiometric coefficients ν_i (positive for reaction products and negative for reactants):

$$\sum_i \nu_i B_i = 0 \tag{14.6}$$

The Gibbs free energy of reaction ΔG is defined through the chemical potentials of the participating reactants and reaction products μ_i. The chemical potential of a species i, present in a mixture, can be separated into two parts (by using the model for ideal gases), where the first one depends only on temperature and pressure and is denoted by a superscript 0 and the second one depends on the activity a_i, which itself depends on the concentration and can be approximated by it. In the case of a pure substance, μ_i is equivalent to the molar free energy.

$$\mu_i = \mu_i^0 + RT \ln(a_i) \tag{14.7}$$

Using the equilibrium reaction constant $K = \ln(\prod a_i^{\nu_i})$ the Gibbs free energy of reaction ΔG can be expressed with $\Delta G^0 = \sum \nu_i \mu_i^0$ as

$$\Delta G = \sum_i \nu_i \mu_i = \Delta G^0 + RT \ln(K) \tag{14.8}$$

For an electrochemical cell the electrons that are produced at the anode are capable of supplying electrical work (negative sign, because work performed outside the system) $W = -nFE_{rev}$ and pass via the external conductor to the cathode, where they participate in a reduction reaction. Therefore this electrical work has to be taken into account in the Gibbs free energy of reaction ΔG

$$\Delta G = \sum_i v_i \mu_i = -nFE_{rev} \tag{14.9}$$

which allows us to calculate the reversible electrical work. Under standard conditions these values are listed.

$$\Delta G^0 = \sum_i v_i \mu_i^0 = -nFE^0 \tag{14.10}$$

We have now set the basis for deriving Nernst's equation (Eq. 14.13), which allows the calculation of the equilibrium potential as a function of concentration (respectively activity) and temperature: Writing Eq. 14.6 explicitly with oxidation and reduction products and the used H_2 and writing also the equations for ΔG and E_{rev} leads to

$$\sum_i v_{i,ox} B_{i,ox} + \frac{n}{2} H_2 = \sum v_{i,red} B_{i,red} + nH^+ \tag{14.11}$$

$$\Delta G = \Delta G^0 + RT \ln \frac{a_{H^+}^n \prod a_{i,red}^{v_{i,red}}}{p_{H_2}^{\frac{n}{2}} \prod a_{i,ox}^{v_{i,ox}}} = -nFE_{rev} \tag{14.12}$$

Solving for E_{rev} and setting at standard conditions $a_{H^+} = 1$ and $p_{H_2} = 1$ atm leads to the following equation. If the activities are (in good approximation) replaced by the concentrations, these can be calculated after measuring the potential difference.

$$E_{rev} = E^0 + \frac{RT}{nF} + \ln \frac{\prod a_{i,ox}^{v_{i,ox}}}{\prod a_{i,red}^{v_{i,red}}} \tag{14.13}$$

For silver, Ag, the concentration of Ag^+ ions in aqueous electrolyte can be calculated at $T = 25°C$ using the fact that metallic Ag has by definition a concentration (activitiy) of 1 and using the tabulated standard electrode potential E^0 of 0.799 V as $E_{rev} = 0.799 + 0.059 \ln c_{Ag^+}$. When applying a certain open circuit potential across an electrochemical cell, the concentrations are determined through the potential, and the increase/decrease of the potential increases/lowers the concentrations of the Ag^+ cations in the solution. Nernst's equation immediately yields that the more noble dissolved metal ions with the higher E^0 will deposit immediately on the solid less noble metal surface. The less noble metal will immediately be dissolved in the electrolyte, yielding a cation exchange reaction $M \rightarrow M^{n+}(aq) + ne^-$.

We will use the Nernst equation later, especially when discussing underpotential deposition and surface effects driven by externally applied potentials. A discussion of

the *pH* dependence of the Nernst equation Eq. 14.13 will lead to Pourbaix diagrams and also corrosion currents can be derived (Tafel plots); however discussing these topics would be out of proportion for this review. For readers educated in electro-chemistry the above topics are well known and the other readers with an optics/solid state physics background are referred to two recent online teaching resources at http://www.doitpoms.ac.uk/ Refs. [10, 11].

In the next paragraph we focus on models of the electrode-electrolyte interface for understanding electrochemical surface science effects such as underpotential deposition and kinetic or non-equilibrium charge transfer processes, which have been investigated with optical techniques. For a tutorial discussion on the electrode-electrolyte interface we follow the book [12].

14.3 The Electrode-Electrolyte Interface

From an atomistic point of view, a metal surface could be modeled by a particle in a semi-infinite potential well with finite height towards the interface. Electrons are much more delocalized than ions and there is a finite probability density also outside the metal, already in the vacuum or in the dielectric. This is called a jellium model, because the attractive potentials of the single ion cores are smeared out. This "spilling out" of the electrons (of the order of 5 Å) leaves positive charges at the metal side of the interface and accumulates negative charge outside (shown schematically in Fig. 14.3). Integrating Poisson's equation yields the potential. If such a metal is dipped into an electrolyte without applying an external potential, the previously dissolved positive cations will preferably gather at the interface in order to minimize the electrostatic energy, and will thereby build up the additional surface charge σ and modify the potential.

For a metal-vacuum interface a charge separation is observed when a metal surface is in contact with an electrolyte. The charge distribution at the metal-electrolyte interface depends in addition to the simple scheme in Fig. 14.3 mainly on:
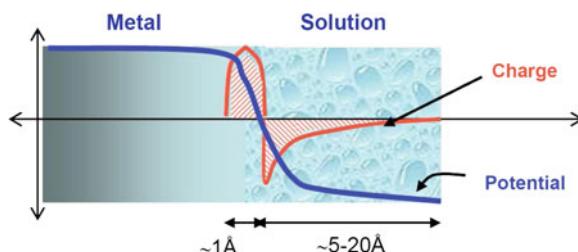
Fig. 14.3 Because electrons leak out of the surface and the ions are rigid—provided there is no dissolution—, there is a positive charge density at the inner side and a negative charge density at the outer side. The corresponding potential drops towards the electrolyte

1. the electronic properties of the solid,
2. the adsorption of water molecules and of hydrated cations,
3. the chemisorption of anions (surface excesses),
4. the externally controlled potential that is applied.

The zone of charge separation near the interface is called *electric double layer*, or simply "double layer". Depending on the applied potential, the charge of the metal can either be positive or negative relative to the electrolyte. The structure of the double layer therefore varies with the potential and with the chemical nature of the ions present. Figure 14.4 gives a schematic representation of the double layer region, for the specific case of excess negative charge at the metal surface. Water is a dipolar molecule. Therefore, when it adsorbs, it preferentially orients itself according to the charges located at the surface. For the case of a negative potential at the metal surface, the first few water layers are immobile and oriented to minimize the electrostatic energy ($d \simeq 2$ nm). Within the electrolyte, ions are dressed by solvation shells of polarized water molecules, which screen the fields originating from the ions. The ions with their solvation shell carry a net charge. Consequently, a space charge layer builds up near the solid surface, with a characteristic decay length depending on the ion concentration in the electrolyte ($\simeq 2$–20 nm). At least for the larger lengths, the electrical properties of this zone can be treated in a continuum approximation.

The Stern layer consists (i) of water molecules (see Fig. 14.4, for this charge state with the hydrogens oriented towards the metal), (ii) cations, which have kept their
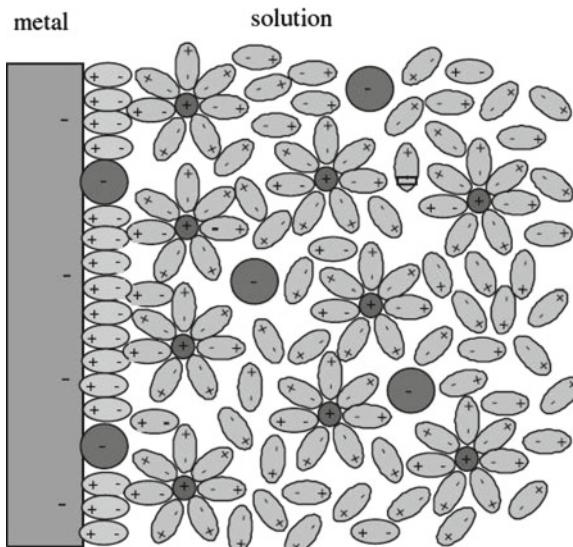
Fig. 14.4 Schematic model of the electric double layer; cations usually carry a solvation shell in the electrolyte, whereas anions usually do not. The specific absorption of anions, despite the fact that the metal is negatively charged, is also shown schematically. This specific adsorption can arise due to strong chemical binding. Reprinted with permission from EPFL from Ref. [7]

solvation shell and (iii) anions in a chemisorbed state directly at the interface (or in the liquid, usually without solvation shells). Electrochemists call these adsorbed species "specifically adsorbed ions" to distinguish between species that adsorb directly on the solid surfaces and those ions that keep their solvation shell. Examples of "specifically adsorbed ions" are Cl^-, Br^-, or $SO_4^=$. When these species are adsorbed on the surface, they form structures similar to the reconstructions known from ultra high vacuum. Depending on specific adsorption the potential changes: Without specific adsorption the potential follows a curve as the one shown in Fig. 14.3. The first few nanometers are called (rigid, inner) Helmholtz (or Stern) layer. The outer region is called (diffuse) Guoy Chapman layer and is characterized by an exponential decay of the potential. For the case of specific adsorption the potential curve varies rather drastically: Due to the presence of negatively charged ions very close to the interface, there is a huge potential drop towards the border of the inner Helmholtz layer, showing even a minimum, and then an exponential rise. For the determination of the spatial dependence of the potential the reader is referred to Sect. 3.2 in Ref. [13]. The thickness of the Stern-layer is ≈ 0.3 nm, its static dielectric constant less than that of free water since the water molecules in that layer cannot rotate freely as in the bulk of the water. To the right the boundary of the Stern-layer, the potential decays exponentially, this region is called diffuse layer. The decay length depends on the concentration of ions in the electrolyte solution.

From a thermodynamic point of view Gibb's thermodynamic potential, which already depends on the electrical potential φ and the interface charge density σ, (and in case of preferred absorbed species with a chemical potential μ_i on the surface excess Γ_i) is in the presence of surfaces amended with a term describing the surface tension ("energy") $X = \gamma A$. We use quotation marks for naming the addition "energy", because the term represents a linear combination of surface deformation energy and the energy for breaking the bonds. The differential of γ can be written as (please note that the same symbol also appears on the right side, but disappears there for liquids, where strain cannot exist):

$$d\gamma(T, \varepsilon_{kl}, \mu_i, \varphi) = -\frac{S}{A}dT + \sum_{kl}(\tau_{kl} - \gamma\delta_{kl})\varepsilon_{kl} - \sum_{i}\Gamma_i d\mu_i - \sigma d\varphi \quad (14.14)$$

For a surface, this equation shows the interrelation between the specific surface free energy and the temperature, the surface stress, the surface excesses and the surface charge. The definition of the charge density is herein only formal in the sense that from a thermodynamic point of view we know nothing about the actual distribution of the charge. It acquires its meaning only within a model (Helmholtz, Guoy Chapman, etc.) in which the metal charge and the ionic charge are separated along an interface.

The Gouy Chapman capacity C_{CG} for a parallel plate capacitor is defined as the ratio of the charge density σ at the interface divided by the potential ϕ; for a not constant field respectively a nonlinear potential decay C_{CG} is obtained by the derivative of the charge density at the interface with respect to the potential at $z = 0$,

$C_{GC} = \partial\sigma/\partial\phi$.

$$C_{CG} = \varepsilon\varepsilon_0 \cosh \frac{Ze\phi(0)}{2k_BT} \qquad (14.15)$$

At low electrolyte concentrations, up to about a 10^{-3} M solution, the Gouy-Chapman theory agrees quite well with experimental values of the double layer capacity for nonadsorbing electrolytes. The Gouy-Chapman capacitance, which was derived by assuming Boltzmann statistics for the charges, has a minimum at the potential where the charge is zero. The (DC) potential, where the minimum occurs, is called point of zero charge. For the tightly bound species within the Stern layer modifications are needed. This is seen experimentally by plotting the measured capacity (and this is the only measurable quantity) versus the calculated C_{CG} with varying electrolyte concentrations. At higher concentrations, the experimental values follow an equation for an equivalent circuit of series connection of two capacitors:

$$\frac{1}{C_{tot}} = \frac{1}{C_{CG}} + \frac{1}{C_H} \qquad (14.16)$$

where the Helmholtz capacity C_H is almost independent of the electrolyte concentration. The current models for the origin of the Helmholtz layer discuss either electronic effects or structural origins coming from the periodic modulation not taken into account by the jellium/continuum model [12]. In the absence of specific adsorption, the potential of zero charge (pzc) can be measured as the potential at which the Gouy-Chapman capacity has a minimum. If this value is independent of the electrolyte concentration, specific adsorption does not occur. The pzc coincides with the maximum of the surface tension, which can be measured directly for liquid metals.

14.4 Adsorption

Adsorption is the adhesion of atoms, ions, or molecules from the liquid phase to a surface. This can happen either by electrostatic interaction (e.g. cations adsorb on a negatively biased electrode) or by chemical reaction (chemisorption) or usually weaker physical (van der Waals) bonds. The exact nature of bonding depends on details of the species involved, but the adsorption process is generally classified as physisorption (characteristic of weak van der Waals forces, with a lower heat of adsorption ~50 meV/atom) or chemisorption (characteristic of covalent bonding, with a higher one up to 500 meV/atom). The adsorption process creates a film of the adsorbate on the surface of the adsorbent. Whenever adsorption occurs, the free energy of the system is reduced by an amount which is called free energy of adsorption $\Delta G_{ads} = \Delta H_{ads} - T\Delta S_{ads}$. Because the structure gets more ordered, the second term is always positive and $\Delta H_{ads} = -Q_{ads}$ must be negative, which means that the heat of adsorption is exothermic and a positive quantity. The heat of adsorption can be varied by changing the concentration in the solvent and can be determined by

measuring the coverage θ as a function of temperature. The underlying adsorption isotherm models are called Langmuir, Frumkin, etc. For details see Ref. [14]. Without going into details of chemisorption, the coverage θ can be measured by the overall charge transfer e.g. in cyclic voltammetry (Sect. 5.2.1.1 in Ref. [15]).

The electrochemical solvent is always present at the interface; if specific adsorption occurs, the interaction of the cation or anion with the electrode has to be greater than that of the solvent. Dissolved cations carry—in general—a hydration shell, therefore they are unable to approach the surface any closer than the distance defined by the radius of their hydration shell. Anions are not hydrated in general. Certain anions such as chloride or other halides are capable of chemisorbing onto the surface, replacing adsorbed water molecules. This may even take place, if the metal is negatively charged.

Beside the adsorption of halides another well investigated adsorption process is termed "underpotential deposition" (upd), when a single (maximum double) layer of a less noble metal (or hydrogen) is attracted to a more noble electrode and adsorbs there. As seen from Eq. (14.13), solution into the solvent is preferred at more positive potentials, deposition would occur only at more negative potentials. However, it is experimentally observed that a monolayer of a less noble metals gets deposited on noble ones already at potentials higher than the calculated equilibrium potential. For example, copper atoms in a solution attach at a bar of silver kept at the same electric potential than the solvent. This is explained atomistically by an enhanced attractive interaction between the first layer of the adsorbate and the substrate. Experimentally the upd effect is detectable, by changing (i.e. reducing) the potential of the substrate electrode rather slowly (mV/s) and detecting the current. At a more positive potential than determined by Nernst's Equation 14.13, the Ag substrate atoms are not dissolved, because the potential is still below the equilibrium potential. Lowering the potential should yield a current increase beginning at the Nernst potential E_{rev}; nevertheless, current peaks are already observed at higher voltages.

14.5 Electrochemical Kinetics

Thermodynamic principles can help to explain a corrosion situation in terms of the stability of chemical species and reactions associated with corrosion processes. However, thermodynamic calculations cannot be use to predict corrosion *rates*. Principles of electrode kinetics have to be used to estimate these rates. Because the field of electrochemical kinetics is extensive, here only the most important aspects are introduced.

There is a basic distinction on the rate limiting cases: (A) When the electrode reaction is controlled by electrical charge transfer at the electrode, then the Butler-Volmer equation is valid, which is discussed below. (B) If the mass transfer to/from the electrode surface from/to the bulk electrolyte limits the kinetics, then diffusion in the electrolyte plays the major role. For the first case, most often, several electrode reactions take place simultaneously at the metal-electrolyte interface. Such systems

are referred to as "mixed electrodes" [7]. The important quantity for kinetics is the activation overpotential η, defined as $\eta = (E - E_{eq})$.

$$i = i_0 \cdot \left\{ \exp\left[\frac{\alpha_a n F}{RT}\eta\right] - \exp\left[-\frac{\alpha_c n F}{RT}\eta\right] \right\} \qquad (14.17)$$

where i is the measured electrode current, i_o the exchange current density, α_c the dimensionless cathodic charge transfer coefficient and α_a the dimensionless anodic charge transfer coefficient, both describing the behavior of the mixed electrodes. The factors in the exponents beside η are the inverse Tafel coefficients. The values Tafel coefficients depend on the precise mechanism of the electrode reactions, which often consist of several elementary reactions. The Butler-Volmer equation describes the charge-transfer kinetics in a global, mechanism-independent fashion, and can be applied without detailed knowledge of the single steps. The three kinetic quantities α_a, α_c and β can be easily measured by polarizing the cell and measuring the resulting current. Evaluating the Tafel lines and the Evans diagram [7] provides these macroscopic kinetic quantities. Despite there have been recent approaches to calculate these kinetic quantities with ab inito techniques, possible in-situ measurements will contribute to understanding irreversible reactions at mixed electrodes.

14.6 Correlation Measurements

14.6.1 Cyclic Voltammetry

Even very small currents (fA) can be measured with electrometeramplifiers very precisely. To obtain an estimation, if on a surface with an area of approximately $1\,\mathrm{cm}^2$ each atom changes its charge state with one electron per second, currents between $100\,\mu\mathrm{A}$ and a few milli amperes result. Since the time of Faraday, time resolved charge transfer reactions were used to study non-steady-state electrochemical processes. In contrast to steady state polarization methods, these transient techniques are able to distinguish reaction phenomena with different time constants. Typically, a variation in either the current or the potential at the working electrode is imposed and then the response as a function of time or frequency is recorded. For corrosion measurements, sometimes potential step or current step measurements are used obtaining information on (a) the mass transport of the reacting species in the electrolyte, (b) the charging of the double layer and (c) the potential drop in (series resistance of) the electrolyte as discussed more extensively in Ref. [7]. With cyclic voltammetry (CyV) the electrode potential is varied periodically with a voltage change between $10\,\mathrm{mV/s}$ up to a few V/s. Depending on the reversal voltage and the voltage change different physical effects can be detected.

As a tutorial example [15] we choose a system where electrolysis is present, Pt in KOH. With Pt electrodes the hydrogen generation, the electrochemical double
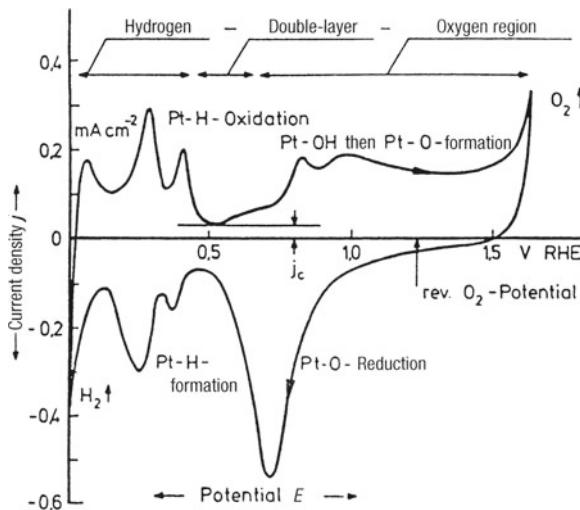
Fig. 14.5 Cyclic voltammogramm of polycrystalline Pt in 1M KOH The solution was purged at $T = 20°$ C with N_2 and the scan rate was $100\,mV/s$. Reproduced with permission by Wiley-VCH from Ref. [15]

layer and also the oxygen generation can be detected, as depicted in Fig. 14.5, taken in de-aerated 1 molar KOH ($pH = 14$) with a potential variation of $100\,mV/s$. For electrolysis measurements, it is advisable to use two electrodes of the same material. Pt itself is inert and does not dissolve in the electrolyte. We start with the discussion at $0\,V\,RHE = -0.812\,V_{SHE}$, which is at $pH = 14$ at the edge of the stability region of water. Further negative potentials will immediately result in hydrogen production $2H^+ + 2e^- \rightarrow H_2$ and a large cathodic (negative) current. Increasing the voltage at the electrode a positive current results displaying three peaks between 0 and 0.5 V. The first one at $\simeq 0.1$ V is assigned to the oxidation of the immediately before produced hydrogen, (which allows by varying the potential variation rate to determine its diffusion constant) and the other two at $\simeq 0.3$ V and at $\simeq 0.4$ V indicate the decomposition of a Pt-H adlayer. At $\simeq 0.45$ V the double layer starts to form and will stay intact for at least $\simeq 1.3$ V. The exact value between hydrogen and oxygen evolution is 1.223 V, but there is some small kinetic hindrance detected. The minimum of the current j_c indicates the onset of the double layer formation. At $\simeq 0.8$ V, and $\simeq 0.95$ V a shoulder and two peaks can be detected, which are assigned to different adsorbate layers involving oxygen (for details see Ref. [15], $Pt + OH^- \rightarrow Pt - OH + e^-, 2Pt - OH \rightarrow Pt_2 - O + H_2O$). From $RHE \geq 1.6$ V, oxygen evolution starts. Reducing the potential yields a cathodic current with only one negative peak at $\simeq 0.75$ V, showing a few hundred mVs overvoltage, indicating the dissolving of the oxygen adlayer. At $\simeq 0.8$ V an adsorbate layer assigned to atomic hydrogen starts to form $Pt + H_2O + e^- \rightarrow Pt - H + OH^-$.

This example shows the power of CyV for detecting hydrogen and oxygen evolution, but also for detecting equilibrium and non-equilibrium adsorption states. For a more detailed discussion, especially highlighting differences in the electrolysis between single crystalline faces of a metal, the reader is referred to Ref. [15]. For the analytical modeling of CyV and the determination of rate constants, and also for derived DC and AC electrochemical techniques (electrochemical impedance spectroscopy as the major AC technique), the reader is referred to Chaps. 9–11 of Ref. [16]. The assignment of some of the adsorbate states is plausible, but still hypothetical, and therefore additional surface science techniques, mainly scanning probe techniques, have been exploited.

14.6.2 Scanning Probe Techniques

Originally electrochemical microscopy has been developed by decreasing the size of the electrodes using already commercially available systems as a scanning electrochemical microscope (SECM). SECM involves the use of a mobile ultramicroelectrode probe (or scanning droplet cell) to investigate the activity and/or topography of an interface on a localized scale. The basic idea of this method is to position a small drop of electrolyte with a capillary on the investigated surface. The wetted area acts as a working electrode and the glass capillary contains the counter and reference electrodes, which are electrolytically connected to the surface through the drop. This resulted in an effective, spatially resolved, in-situ investigation by all standard electrochemical techniques such as cyclic voltametry, electrochemical impedance spectroscopy, current transients of potentiostatic pulse steps etc. In the next paragraph we discuss electrochemical scanning tunneling microscopy. The information obtained by this kind of microscopy is equivalent to CyV and related techniques, but on a scale of $\sim 1\,\mu$m. By measuring faradayic currents, different grains of a polycrystalline material can be visualized, but the resolution is still far above atomic levels.

Electrochemistry gained a new impetus by using single crystalline surfaces and thereby questions of the atomic ordering and the atomic structure of the liquid–solid interface were raised. Scanning probe microscopy forms images of surfaces using a physical probe that scans the specimen as sensor for an action (current, force, capacitance, etc.). An image of the surface is obtained by mechanically moving the probe in a raster scan of the specimen, line by line, and recording the sensor signal as a function of position. The resolution varies depending on the sensor principle, but some techniques reach atomic resolution due to the ability of piezoelectric actuators to execute motions with a precision and accuracy at parts of Å level. The data are typically obtained as a two-dimensional grid of data points, visualizing the sensor signal in false color as a computer image.

Despite atomic force microscopy, chemical force microscopy, near-field scanning optical microscopy can be operated in liquids, we focus here on electrochemical scanning tunneling microscopy (ECSTM). For the development of the scanning tunneling

microscope its inventors, Gerd Binnig and Heinrich Rohrer, earned the Nobel Prize in Physics in 1986. Seven years later Prof. Itaya transferred the STM into water [17]. As the distribution of anions or cations changes with the potential of the electrode, it is necessary to control the reaction on the electrode. The potentials of the *two* working electrodes (the sample and the tip) are controlled independently against a reference electrode. In this case, the tunneling bias voltage is the difference between the potential of the working electrode and the coated tip (to suppress faradaic currents) with a free apex. A counter electrode is used to complete the current-carrying circuits with the working electrodes. By using these four electrodes, the electro-chemical reaction is controlled precisely by the external voltage, and the surface in liquid can be observed. It also became clear how to suppress the faradayic current and measure (mainly) the tunnel current: by coating all the tip except the apex. The tunneling current is exponentially decaying with the distance of the tip from the surface and contains in addition information about the (surface) density of states and the work functions of the tip as well as of the sample. Increasing the distance by approximately 1 Å reduces the current approximately to one third.

A tutorial example taken from Ref. [18] shows in Fig. 14.6a the comparison between a CyV of $Au(111)$ surface in 0.1M H_2SO_4 and an ECSTM image, where during the scan the potential was varied. In Fig. 14.6b an ECSTM image is shown, where during the scan the potential was varied from 0.65 to 0.8 V. Each adsorbate layer is providing faradayic currents, which are detected by CyV. These adsorbate layers can simultaneously change the surface reconstruction. In this specific example a gold-sulfur adlayer forms and within a narrow voltage region this yields an imme-
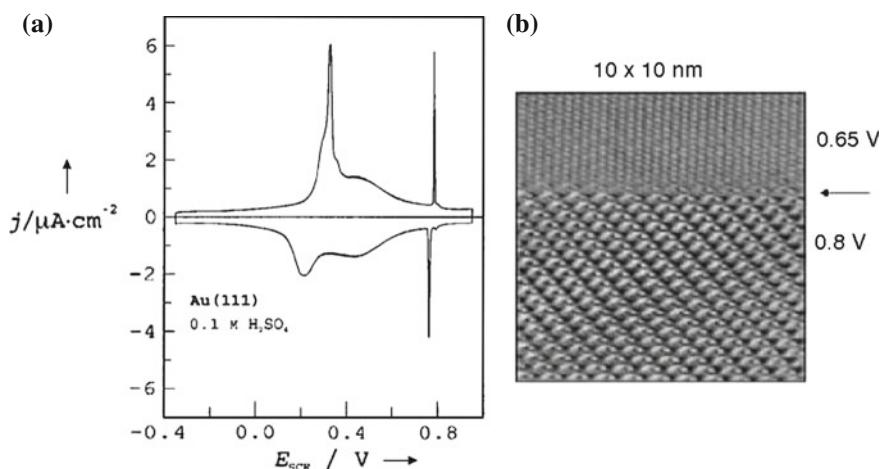
Fig. 14.6 Part **a** of the figure shows cyclic voltammetry of a $Au(111)$ surface in 0.1M H_2SO_4. In part **b** the ECSTM image is shown, where during the scan the potential was varied from 0.65 to 0.8 V, yielding an immediate change in the surface reconstruction from atomic Au (above) to an adsorbed layer comprising sulfur. Reproduced with permission by Bunsengesellschaft from Ref. [18]

diate change in the surface reconstruction from atomic Au (upper part of Fig. 14.6b) to an adsorbed layer comprising sulfur (lower part).

14.7 The Use of Polarization Optical Probes for the Liquid–Solid Interface

Despite this chapter reviews polarization optical probes or *epioptic* probes [4], we start, in order to issue a warning, with mentioning three major disadvantages of all optical probes due to their indirect nature:

- The large penetration depth of optical radiation into condensed matter makes the isolation of a surface or interface contributions difficult. Special experimental efforts and techniques, partially supported by *ab-initio* calculations, are necessary to overcome this limitation and to interpret the measured signals in terms of physical processes.
- The optical response provides usually an indirect information, compared to electron microscopy or scanning probe techniques. Only with a model for the optical response the requested physical or chemical parameters, such as composition, surface excess, stress and resulting strain, interface energy as equilibrium quantities or kinetic parameters such as diffusion coefficients, can be deduced. This modeling is often controversial and leads to scientific discussions, on the other hand it fosters a thorough understanding of the optical and quantum mechanical effects and often leads to unexpected conclusions. Especially the advent of high power computing equipment made these modeling procedures possible.
- The quest for the true dielectric function of the material as well as the structural dispersion relation. To explain this point in more detail, we discuss the example of a metal or semiconductor quantum dot. *If the dielectric function does not change with the size of the inclusion or dot on the surface*, i.e. no quantization effects are occurring, solving the electrodynamic (ED) Maxwell equations is the main difficulty to interpret the measurements in terms of geometrical parameters. Because we physicists and chemists are used to use the Fresnel formulas as "ad-hoc" approximations, even when the system is neither stratified nor homogeneous, effective media are introduced and their properties are deduced and discussed. However, this procedure can be totally wrong if the dielectric function changes and this is well known for metal particles due to the scattering at the surface [19] as well as for semiconductors of nanometer-sized particles [20]. These two different origins fuel scientific discussions and rather seldom consensus can be reached.[1]

---

[1] Especially these discrepancies make the discussions exciting and the indirect nature of the optical probes as well as the possible controversial origin provide a fascinating playground for physics.

These shortcomings are complemented by major advantages of the optical probes: The material damage and contamination associated with charged particle beams is eliminated; in-situ, in-process, and real time experiments are possible and transparent liquids can be used for studying the liquid surface or interface. Insulators can be studied without the problem of charging effects and also buried interfaces are accessible owing to the large penetration depth of the optical radiation. There is no influence nor a time limitation due to *steric hindrance*, which inhibits the use of scanning probes in industrial process technology. In addition, epioptic probes offer micron lateral resolution and picosecond temporal resolution.

Optical techniques, especially reflection and absorption measurements, were the oldest and one of the few in-situ methods for investigating electrodes. The substantial progress made in this area over the past two decades has been partly due to technological advancements in spectroscopic instrumentation in general, including detection systems, the advent of Fourier transform techniques, and the availability of synchrotron radiation. Beside intensity based techniques, the only physical methods that can monitor the surface processes in liquids are polarimetric probes. For the purpose of controlling physical parameters as thickness, growth rate, diffusion or alloy compositions only optical techniques (including X-ray) are suited to be applied simultaneously with coating/deposition/liquid phase epitaxy, because they are non-invasive and non-destructive and light can be guided through any transparent liquid onto the substrate.

Absorption measurements are mainly carried out by either using transparent and conducting electrodes as indium tin oxide, or by attenuated total reflection (ATR), where light experiences multiple internal reflections in the crystal of high refractive index adjacent to the electrolyte with lower one. The evanescent field, which extends into the electrolyte, probes also the species in the electrolyte and provides thereby chemical information. The major disadvantage of ATR is, however, that the working electrode has to be rather thin and can hardly be monocrystalline. For a collection of monographs in the area of in-situ UV-VIS and IR spectroscopy as applied to the study of electrochemical interfaces, see the book [21]. Additionally, nonlinear optical techniques as second harmonic generation and some frequency generation are applied to the liquid–solid interface; for a survey we recommend the recent article Ref. [22] and the book chapter Ref. [23]. Raman spectroscopy (RS) has also been used to investigate electrolytes and electrodes. Its major advantage for aqueous systems is that the penetration depth of the carrier is usually much larger than for the IR radiation, but this does not fully compensate the usually weak intensity of the Raman process of a thin overlayer. Surface sensitivity can only be obtained with surface enhanced Raman spectroscopy (SERS). Despite the causes for the sensitivity of SERS are still discussed (probably a combination of field enhancement, resonant excitation and charge transfer), the enhancement factor of 10^6 for SERS in scattering efficiency over normal Raman scattering shows its potential to identify molecules and complexes at the surface, combined with polarization information the orientation and steric alignment are accessible. For Raman spectroscopy we refer the reader to different reviews: The use of RS for measurements of dissociation rate to differentiate

activity and concentration has been discussed in the book chapter [24], vibrational spectra are discussed in the review [25], and SERS is reviewed in Ref. [26].

Ellipsometry has also been used since 1960, but mainly in its single wavelength form. In its *spectroscopic* form right now (and even more in the future) it experiences a renaissance due to its nondestructive and real time nature. Despite the penetration depth of light is for all materials at least of the order of 100 nm or above, surface sensitive optical probes have been developed, capable of monitoring processes on the nanosecond scale with thickness resolution of less than a monolayer. On the first sight the major disadvantage of the indirect nature of optical probes, compared to electron microscopy or spectroscopy proves then to be a fascinating feature, because all polarimetric, intensity, linear or nonlinear or scattering optical probes require modeling the full system for understanding the interactions between the liquid, solid and interface properties. In this chapter we limit ourselves to report on the electrochemical interface, where either in-situ Spectroscopic Ellipsometry (SE), in-situ infrared Spectroscopic Ellipsometry (IRSE), or in-situ Reflectance Difference (Anisotropy) Spectroscopy—(RDS) have been used.

14.7.1 Electroreflectance with Polarized Light

An overview of polarimetric probes applied to the liquid–solid interface should start with measurements of Cardona, Shaklee and Pollak, who employed—to our knowledge—for the first time polarized electroreflectance (ER) measurements on the Si (110) face [27], with light polarized parallel to the [001] or to the [1$\bar{1}$0] direction. Their main intention was to observe in the spectra transitions due to critical points in the joint density of bulk states for the semiconductor under study. Despite the Si bulk has cubic symmetry and should provide polarization insensitive response, ER spectra for the Si (110) sample orientation should depend on the direction of polarization of the incident light, while that of a (111) and a (100) surface should not. The measurements of the Si (110) surface for the two normal modes of light polarization (electric field vector parallel to [001] and [1$\bar{1}$0] direction) are shown in Fig. 14.7 for a DC bias of $+3$ V, and in Fig. 14.8 for a DC bias of -3 V. Already these authors found anisotropy: for light polarized along the [001] direction the structures are much stronger than for light polarized along [1$\bar{1}$0]. They also found a dependence of the ER spectra on the bias voltage. Despite the labeling of the peaks has changed due to a better understanding of the band structure in the last 45 years, (E_0' is called E_1, etc.) and some of these transitions are assigned in addition to excitonic states [28], the sensitivity of a polarization optical probe on the bias voltage has been demonstrated uniquely. By varying the bias from $+3$ V to -3 V, the signatures get inverted, which is only possible due to the existence of an inversion layer beneath the native oxide of silicon. So, this inversion is a typical bulk effect, especially, because the penetration depth of light is smaller than the space charge region. However, it is also currently not clear, why the relative changes in ER as well as the width of the transitions are different for the different bias, as well as for different polarizations.
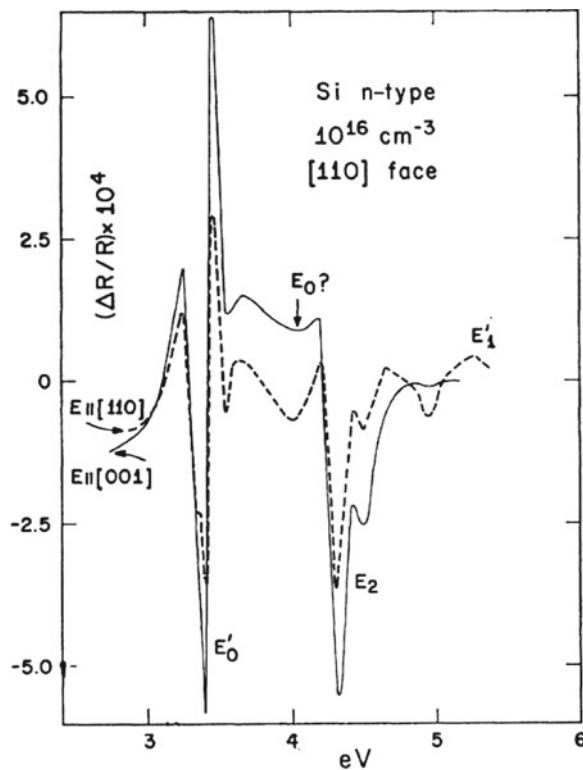
Fig. 14.7 Electroreflectance spectrum of n-type silicon (110) face at room temperature in water. The electric field of the light was either polarized along [001] or [1$\bar{1}$0]. The DC bias voltage is +3 V, the AC rms modulation is 4 V. Reproduced with permission by the APS from Ref. [27]

The measurements were performed on a semiconductor with a thin oxide layer separating the electrolyte from the metallic region and so possible surface states will not play a major role. For the metal-liquid interface, it is an important question, whether surface states, which have been observed by a variety of experimental techniques on well-prepared single crystal surfaces in UHV, exist at metal-liquid interfaces. Such states are highly localized at the metal surface. Since they are extremely sensitive to surface contamination in UHV it might be expected that they would not exist at a metal-liquid interface. Early electroreflectance (ER) measurements by Dieter Kolb and coworkers [29] revealed anisotropic polarization dependence in reflected light from noble metal/electrolyte interfaces $Cu(110)$ and $Ag(110)$ that were interpreted in terms of contributions from interband transitions. In the cited article ER spectra for normal incidence as well as for oblique angles are shown and discussed.

The dependence of ER (the change comes, as before, from the AC modulation, which is in this example 0.5 V, i.e. $\frac{\Delta R}{R} = \frac{R_{V_{max}} - R_{V_{min}}}{R_{V_{avg}}}$) on the angle (the indicated angle ϕ_1 is the one in air) is shown in Fig. 14.9. The data were taken in a goniometric
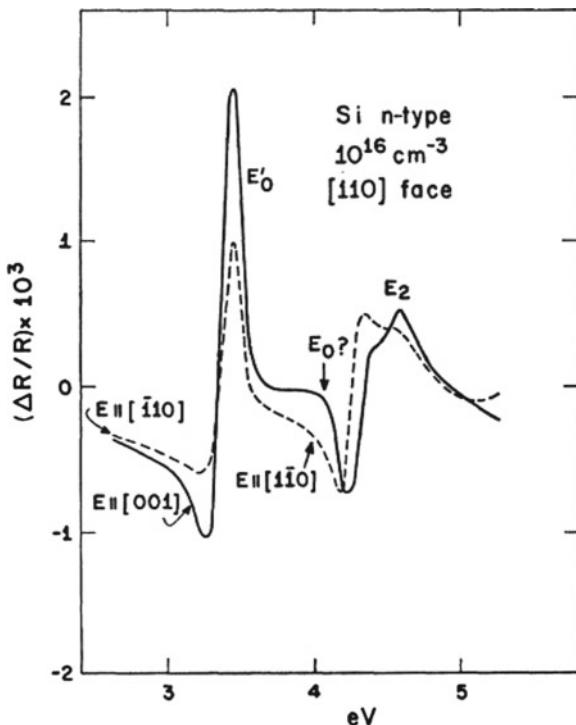
Fig. 14.8 Electroreflectance spectrum of n-type silicon (110) face at room temperature in water. The electric field of the light was either polarized along [001] or [1$\bar{1}$0]. The DC bias voltage is −3 V, the AC rms modulation is 4 V. Reproduced with permission by the APS from Ref. [27]

setup with a parallel glass plate covering an electrolyte. It can be clearly distinguished that for these high symmetry surfaces the usual argument of working with p- polarization close to Brewster angle to enhance the sensitivity drastically is not correct; for strongly absorbing substrates there are changes in the intensity of the s- as well as the p- polarized light.

Changes in the reflectances can have various origins, when the potential is modulated: (a) either a layer is adsorbed (chemisorbed) at the surface, or (b) the electronic density at the surface varies or (c) the ER spectra are influenced by (e.g. periodic) corrugations, where the steps induce, due to the periodic AC modulation, a rather high local electric field, which supplies momentum for plasmon excitation. The authors discuss the assignment to special transitions in the Brillouin zone (e.g. for Au the 2.5 eV structure is probably related to the interband transition and also the 3.5 eV peak). For silver the $\frac{\Delta R}{R}$ data at 3.8 eV can be explained mainly in terms of a free electron model (plasmonic resonance at 3.8 eV).

If electroreflectance monitors, as indicated above, interband transitions, this can only be due a finite depletion or accumulation region at the surface. Therefore we use these data to discuss the possible origins of the $\frac{\Delta R}{R}$ features. For this discussion
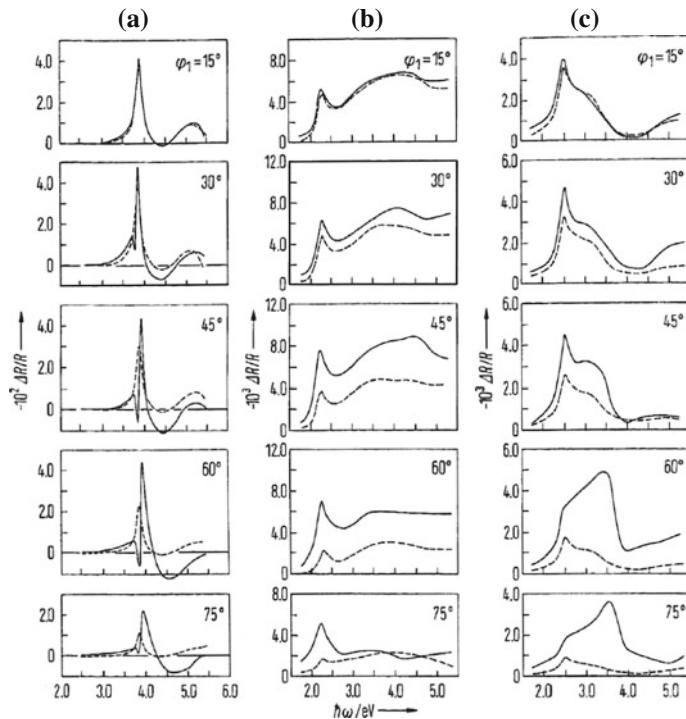
Fig. 14.9 Electroreflectance spectra of Ag(111), Cu(111) and Au(l11) electrodes in 0.5 M $NaClO_4$, as a function of the angle of incidence ϕ_1. *Full line*: p-polarization, *dashed line*: s-polarization. **a** Ag(111) with a DC bias voltage of -0.25 V, the AC peak to peak modulation is 0.25 V. **b** Cu(111) with a potential and modulation as for Ag(111). **c** Au(111) with a potential and modulation as for Ag(111). Reproduced with permission by Oldenbourg Verlag from Ref. [29]

we have to keep in mind that each optical transition occurs from a ground state to an excited state and that the strength of this quantum mechanical dipole transitions constitutes the origin of the dielectric function.

- The first possible origin are "bulk like" interband transitions. In semiconductors, the energy of the ground state is modulated by the applied potential and the depletion (accumulation) region is much wider than the penetration depth. The higher lying excited states are less influenced by the external potential than the ground state and therefore "bulk like" interband transitions with a slightly modified (by the potential modulation) transition energy occur.
- For metals with an electronic density of 10^{23} cm^{-3} the depletion or accumulation region will only be of the orders of maximum a few Å. In this limit one does not speak any more in terms of depletion regions and a single electron picture, but of Friedel oscillations in the many particle picture. Furthermore, in this model the electron density varies around the impurities or perpendicular to surfaces and when applying external potential modulation the ground state of the many particle

Hamiltonian will experience a different energy shift than the excited states. Density functional theory calculations start by assuming a *coherent* many particle wave-function (respectively density) which extends over all the computational cell.

- The difficulty to grasp conceptually and also analytically the physics of a many particle Hamiltonian led Shockley and Tamm to a simplified model of the crystal potential in one dimension. It can be shown analytically (Sect. 6.1 in Ref. [30]) that the one-dimensional single-electron Schrödinger equation gives two qualitatively different types of solutions. The first type of states extends into the crystal and has Bloch character there, but decays exponentially when reaching into the neighboring vacuum (liquid). The second type of states decays exponentially both into the vacuum (liquid) and the bulk crystal. These types of solutions correspond to states with wave functions localized close to the crystal surface and are either called Shockley or Tamm states. It can be shown that the energies of these states all lie within the band gap. As a consequence, in the crystal these states are characterized by an imaginary wavenumber (perpendicular to the interface) leading to an exponential decay into the bulk. In this picture an optical transition could occur between two surface states, both characterized by an imaginary wavenumber, or between a (single particle) surface state and a (single particle) "bulk" state, where the energy of the excited (single particle) state is already resonant to the bulk band structure. In the respect of one coherent wavefunction the question of transitions involving bulk and surface states, surface states and states in the liquid, etc. are meaningless; it can only be well posed, if single particle states are present.

- Adsorbates will also modify the optical response of a system, at least if there is charge transfer present or the electron density (of the ground state) varies through an applied potential. This process is called chemisorption. Physisorption is also monitorable by optical techniques, if the refractive index of the adsorbate is different from the one of the repelled liquid.

Therefore resonance structures in ER signal $\frac{\Delta R}{R}$ are only seen due to a dynamic (periodic) change. If the structures are due to modified bulk transitions, due to adsorption of a species or due to surface states is an ad-hoc open question and can only be answered by comparing the optical response to complementary kinetic measurements. However, the sensitivity of polarization optics, combined with a lock-in technique, usually allows to detect even very minor changes of $\frac{\Delta R}{R}$ down to 10^{-5}. Despite the assignment of the structures to either interband transitions, plasmonic resonances, or adsorbates in Fig. 14.9 is plausible, it should be kept in mind that an optical measurement alone can be misleading. Even on platinum, as shown before, CyV indicates adsorbate layers and optics, especially polarization optics, on the other hand, is strongly sensitive on these adsorbate layers.

For near normal incidence, however, the difference of the ER data between s- and p- polarized light is negligible. The situation changes when ER data of fcc(110) surfaces are measured as also presented for Ag and Cu in the reference above [29]. Here, we reprint these experimental findings in Fig. 14.10.

Polarized $\frac{\Delta R}{R}$ measurements of fcc(110) surfaces are published for Ag in Ref. [31]. The authors claim that the effect is due to a field induced indirect interband transition,
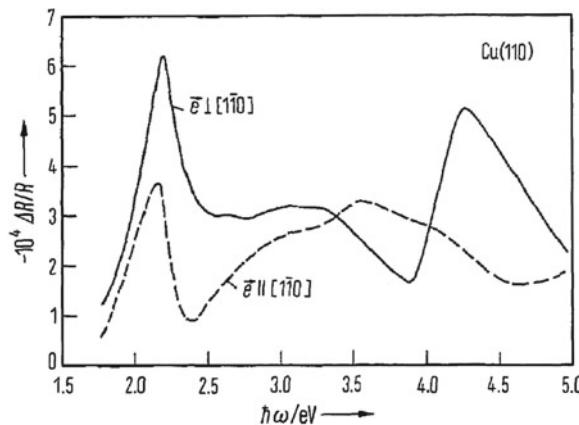
Fig. 14.10 Normal-incidence electroreflectance spectra of Cu(110) in 1 m H_2SO_4 for the electric field vector of the light being parallel (—-) and perpendicular (-) to the [1$\bar{1}$0] direction. DC potential = −0.3 V; AC (= 100 mV peak-to-peak, f = 180 Hz). Reproduced with permission by Oldenbourg Verlag from Ref. [29]

but they display the data only for one applied DC voltage. For Au and Cu, potential dependent polarized ER are given in Ref. [32]. These authors discuss the unclear origin of the ER changes.

Kolb et al. [33] obtained evidence by ER for the existence of surfaces states at Ag(100)/electrolyte interfaces by extending the ER technique into the infrared and they observed two additional ER resonances. Due to the strong potential dependence of the ER spectra they modelled the spectra by transitions within the Ag(100) surface band structure and argue that biasing the surface gives rise to shifts of the Fermi level. The onset of the signal at a certain DC threshold is assigned to a transition between two surface bands. However, as discussed before, ER results are influenced by the concentration of the electrolyte and the presence of adsorbed species and so it is unclear if this explanation can really hold.

14.7.2 Spectroscopic Ellipsometric Measurements

This chapter mainly focuses on spectroscopic ellipsometry. Single wavelength ellipsometry has been applied since the middle of the last century to the solid–liquid interface and has been reviewed in detail by Beaglehole [34] and Hamnett [35], and a decade ago by Christensen and Hamnett [36]. In the electrochemical scientific community it has been understood that ellipsometry measurements are only taken and published at He-Ne wavelength, therefore in some articles there is even no mentioning of the wavelength used. Despite this shortcoming, the sensitivity of (seldom spectroscopic) ellipsometry has been clearly shown in the old literature; and hence

most measurements have been interpreted on a kinetic basis at the He-Ne wavelength. Occasionally the single wavelength data are also interpreted in terms of changes of the real and imaginary part of the refractive index, but it is clear that hidden effects can also contribute to their changes. Here, reviewing biological systems and protein adsorption are out of the scope, therefore the reader is referred for these topics to a recent article by Arwin [37].

In the following, we will discuss four examples from the recent literature, which show spectroscopic ellipsometry measurements during (1) thin film deposition, (2) monitoring underpotential deposition, (3) the etching of semiconductors and (4) measuring the dielectric function of liquids. For two further examples, namely the spectroscopic ellipsometric of (a) monitoring the freezing of alloys (and detecting the onset of interband transitions) [38], and (b) spectroscopic ellipsometry results of the gas-liquid interface [39], we refer to these original articles. As a side remark we note that for the case of graded or continuous interfaces the usual Mc-Intyre-Aspnes three phase model (applicable for discrete dielectric function changes) is replaced by an integral expression, which is proportional to the surface excess. The underlying equations are used e.g. in Ref. [39] to determine the surface excess and are derived in the book of John Lekner [40]. This book, however, uses a different definition for the reflection coordinate system and so the usual r_p is replaced by $-r_p$.

The first example covers the use of SE for electroplating. Electrochemical thin film deposition or electrodeposition is a process in which metal ions in a solution are moved by an electric field to coat an electrode. It is analogous to a galvanic cell acting in reverse. The process uses electrical current to reduce cations of a desired material from a solution and coat a conductive substrate with a thin layer of the material, such as the Bi_2Te_3 semiconductor, as discussed by Broch et al. [41–43]. Bismuth telluride thin films promise high thermoelectric efficiencies through miniaturization and stacking. The deposition proceeds via a 1 M nitric acid electrolyte with bismuth and tellurium salts ($[Bi^{3+}] = [TeO_2^+] = 10^{-2}$ M) dissolved and using a chemical reagent to reduce the surface roughness. The deposition proceeds on a Au substrate in the cathodic region. In this work SE measurements were combined with electrochemical quartz crystal microbalance (EQCM) measurements and special emphasis was put on fast (every 25 ms) data acquisition in the wavelength range from 1.5 to 3.2 eV.

In Fig. 14.11 the time resolved deposition current as well as the frequency shift is shown during deposition and dissolution. Figure 14.11 gives the impression that the deposition is straightforward and the dissolution reaches the same endpoint. The spectroscopic ellipsometry signal, however, recorded at 2.07 eV yields a much more complex and only partly understood behavior. The trajectory for deposition (full dots) as well as the trajectory due to dissolution (empty dots) can be seen in Fig. 14.12. Even without quantifying the thickness of the film, because the two trajectories are different, one can conclude that cathodic electrodeposition proceeds trough other morphological stages than anodic film removal. Despite the kinetic process is different, the starting points/endpoints are identical. The thickness (respectively roughness) of the film can be quantified by using the measured dielectric function of the electrolyte as well as the deposited thin layer [42].
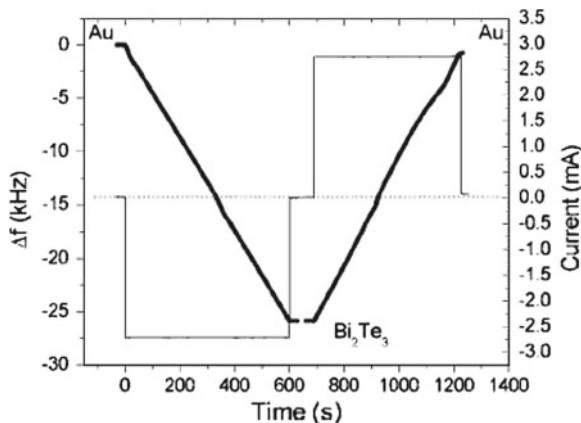
Fig. 14.11 Current cycle (*right ordinate*, *thin line*) and consecutive frequency response (EQCM) signal (*left ordinate*, *thick line*) during electrodeposition respectively stripping of Bi_2Te_3. Au marks the starting point and the end point of the cycle, where the gold substrate is film-free. Bi_2Te_3 marks the surface state, where the bismuth telluride layer covers the substrate. Reproduced with permission by AIP from Ref. [41]

The second example treats underpotential deposition (Cu UPD on Au) [44]. As already discussed before, UPD is an effect where a less noble cation adheres at a potential less negative than the equilibrium (Nernst) potential at a more noble substrate. UPD is an indication of a strong interaction between the electrodepositing metal (Cu) with the substrate Au. The Cu-Au interaction is energetically more favored than the Cu-Cu interaction in the crystal lattice. Depending on the strength of the next nearest neighbor interactions, either one or two monolayers can be deposited. It is known from cyclovoltammetry (by measuring the charge) that for the Cu-Au system only one monolayer UPD occurs. As Au substrates, polycrystalline Au films with a preferred texture (Au(111) from Arrandee company) have been used. The electrolyte used was a 50 mM H_2SO_4 + 1 mM H_2SO_4. As a reference electrode, a standard calomel electrode (SCE) was used, where the potential difference between SCE and SHE is +0.241 V. The SE data were taken in the energy range $1.7 - 4$ eV. The experimental data are shown in Fig. 14.13. At very positive potentials Cu is fully dissolved in the solutions. Approaching $E_{SCE} = 0.2$ V from the cathodic scan shows two peaks identified as "A" and "B", which are known to be related to Cu adlayers (UPD) (left side of Fig. 14.13). Decreasing the potential below the dashed line, Cu now deposits on Au due to Nernst equation (Overpotential deposition, OPD).

In the time resolved SE data acquired during the cathodic scan two thresholds can be observed. The first is related to the onset of the UP regime. Solid arrows mark the position of the "A" and "B" peaks in the CyV. The second one, very sharp, is related to the onset of the OP regime: the larger variations of the SE parameters indicate Cu bulk deposition. In the OP regime, ψ as well as Δ apparently show a strong dependence on the scan rate.
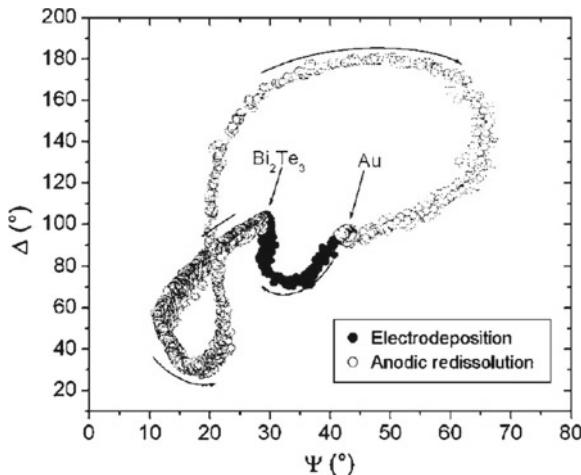
Fig. 14.12 Response of ellipsometric parameters ($\hbar\omega = 2.07\,\text{eV}$) to current cycle shown in Fig. 14.11. Au marks the starting point and the end point of the cycle, where the gold substrate is film-free. Bi_2Te_3 marks the surface state, where the bismuth telluride layer covers the substrate. The fact that the $\psi-\Delta$ trajectory takes a very different path when electrodeposition occurs (a short one from "Au" to "Bi_2Te_3") respectively when redissolution occurs (a long one from "Bi_2Te_3" to "Au") indicates morphological differences, which cannot be followed by EQCM at all. Reproduced with permission by AIP from Ref. [41]

In the UP region the overall trend of SE data is rather independent of the can rate. Upon decreasing the scan rate, one can observe a well defined increase of the absolute variations of the SE parameters. Because the change in Δ is, down to the monolayer limit, for thin films proportional to the thickness change the phase difference is more sensitive to the UPD regime than the ratios of the p- and s- reflectivities. Therefore larger Δ variations are related to the deposition of a thicker film, or, more probably, to a higher coverage. The authors assigned the stepwise behavior of ψ and Δ to the two-step "A" and "B" peaks of CyVs, which represent a submonolayer and a monolayer phase of Cu on Au(111). The authors measured SE data slightly negative at the "B" peak, where a full Cu monolayer should already be developed. The data were fitted with a three phase model with a Cu overlayer on Au, with the dielectric functions taken from Palik's handbook; the Cu adlayer thickness was determined to be 1.7 Å, which is very close to the data reported in the literature. Despite the signal-to-noise ratio of the ellipsometric data is inferior than the CyV measurements and the changes are minor, these SE data represent—different from modulated reflectometry—a static fingerprint of the liquid–solid interface. Finally, it should be mentioned that in the OPD regime, where metal deposition occurs, the SE response is very sensitive to the scan rate.

The underpotential deposition of Cu on Ru has been discussed in Ref. [45]. Besides UPD of Cu on Ru, also the adsorption of polyethylene glycol (PEG) and chloride on Ru has been investigated by SE. Two states of the Ru surface are investigated:
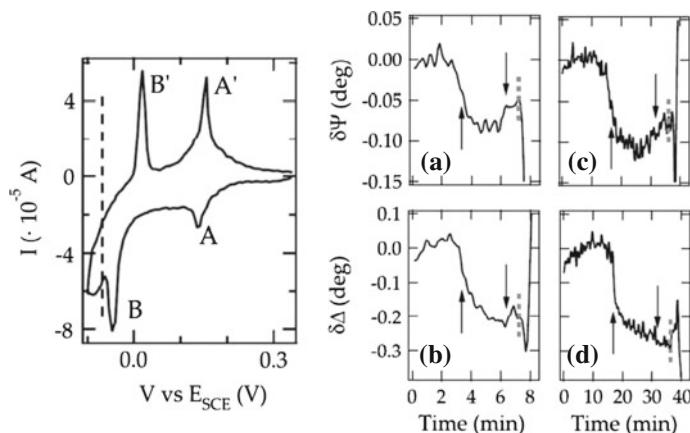
Fig. 14.13 *Left side* cyclic current-potential curve for Cu deposition on polycristalline Au substrate in $50\,mM\,H2SO4 + 1\,mM\,CuSO4$ (Scan rate: $10\,mV/s$). The *dashed line* marks the equilibrium Nernst potential for the $Cu^{2+} + 2e^- \rightarrow Cu$. *Right side* Progress of the ellipsometric parameters during cathodic scan of the potential from 0.33 to$-0.17\,V$ versus E_{SCE} at a wavelength of $632.2\,nm$. Panels A and B refers to a scan rate of $1\,mV/s$ while panel C and D refers to a $0.2\,mV/s$ scan. The *arrows* indicates the position of the UP deposition peaks. The *dashed lines* mark the equilibrium Nernst potential for the oxidation/reduction of Cu. Reproduced with permission by Wiley VCH from Ref. [44]

the activated state, where electrolytic reduction of the native oxide of Ru is electrochemically obtained and the same with the addition of a Cu UPD layer. SE data are summarized by fitting thickness of the adsorbed surface film as a function of potential or solution composition. The investigation of interactions between polyethylene glycol and chloride and the co-adsorption process is important for feature filling in microelectronic circuits [46]. The combination of SE and CyV showed that the 6 Å thick PEG/Cl layer does not block electron transport. The interaction of organic accelerating agents for Cu deposition with the surface measured through SE is described in Refs. [47, 48].

In a series of papers the industrially important process of electrochemical etching of elemental as well as III–V semiconductors has been investigated. This topic was mainly studied by Prof. Adachi and his group [49–54], and also by others [55, 56]. Comparative CyV and time resolved single wavelength ellipsometry measurements of manganese oxide films are described in Ref. [57]. Experimental data on the interaction between water and thermally oxidized porous Si are shown in Ref. [58]. It turned out for Si(111) surfaces that aqueous NH_4F and NH_4HF_2 solutions cause the removal of the silicon native oxide. After the native oxide was removed completely, as additional XPS data confirmed, the SE data yielded the spectrum of a slightly roughened surface. Further etching leads to gradual roughening of the sample surface. The roughening increases at higher solution temperature. This is due to an increased dissolution rate with increasing solution temperature. Auger measurements indicate that a $20\,\%\,NH_4HF_2$ etching provides a cleaner silicon surface than

etching with 40 %NH_4F. All etched surfaces are strongly hydrophobic, suggesting that not only HF, but also 20 %NH_4HF_2 solution can lead to stable, H-terminated silicon surfaces.

For GaAs (001), etched in either dilute or concentrated HF, it was found that both etching solutions remove the native oxide (the diluted one needs around 15 min) and attack the Ga atoms, leaving As behind. F-related peak was detected with XPS only in minor quantities on the HF-etched surface. The measured wettability of the degreased and not yet etched surface indicated that the as-degreased GaAs surface is hydrophilic, while the HF-treated surfaces are (strongly) hydrophobic. Treating GaAs after etching in an alkaline hydrazine-sulfide solution results in the formation of thin overlayers coherently bonded with the substrate. For alkaline (pH $\sim$ l2) hydrazine-sulfide solution a layer of gallium nitride forms, which probably arises through dissociative adsorption of the hydrazine molecules on the Ga-terminated (001) surface. In low alkaline solutions Ga-S bonds are formed at the Ga(001) surface. Finally, in Ref. [59] the correlation between wet chemical etchants, the Pourbaix diagram of GaAs and the related pseudo-dielectric functions, depending on the chemical pre-treatment are discussed, with the goal of measuring the Fermi-level pinning in GaAs. The authors concluded that the ability to characterize the composition of these surface layers, or determine whether the Fermi-level was essentially un-pinned, appears unfeasible by SE.

Here I just want to refer to another technique, namely Reflectance Difference Spectroscopy, described in Sect. 14.7.3. RDS is able to determine the Fermi level pinning in-situ [60]. However, to my knowledge, this technique has not been applied until now in an electrochemical environment for the purpose of determining Fermi level pinning.

In measuring the dielectric function of liquids for evaluating the dielectric function of adlayers, the knowledge of the refractive index of the liquid is a prerequisite. For water, e.g, the wavelength dependent data scatter considerably and instead of providing citations on the measurements of the different authors for the reader, a compilation of data for water and ice can be found in the following website, tabulating atmospheric data: http://reflib.wikispaces.com/.

The dielectric function of various other liquids has been determined with rotating compensator SE by Bang et al. [61]. They used three different measurement configurations: (a) the direct measurement at the air-liquid interface, (b) the prism -liquid interface, and finally the liquid-sample interface for determining the dielectric function of toluene, ethanol, methanol, methylene chloride, chloroform, ethyl acetate, hexane, tetrahydrofuran, ether and dimethylformamide. In the transparent region the dielectric function data can be fitted by a Sellmeier dispersion relation using in total 5 real parameters, which are: (1) the frequency independent dielectric constant and two oscillators with resonance frequency and oscillator strength. For the liquids mentioned previously, the resonator energy is between 6 and 10 eV. Not surprisingly, the Bruggeman effective medium theory is an excellent approximation for mixtures. In the infrared a similar study has been performed by DenBoer in Ref. [62].

14.7.3 Reflectance Difference Spectroscopy

Reflectance Difference Spectroscopy (RDS), also named reflection anisotropy spectroscopy, is a non-destructive optical probe of surfaces which is capable of operation within a wide range of environments. Because light can penetrate transparent liquids, RDS is applicable in liquid environments and also in electrochemical environments. RDS utilizes a highly accurate method of determining the polarization state of light reflected in perpendicular incidence from a surface. If the substrate has a fourfold rotational symmetry the Fresnel reflectances yield the same value for s- and p- polarized light and the RDS signal cancels. If, however, a disturbance destroys this 4-fold rotational symmetry a net signal from the difference of the phases and from the ratio of the amplitudes remains and provides a fingerprint of the disturbance. Symmetry breaking can either occur via surface reconstructions or by anisotropic bulk strain, or through surface related bulk effects. As examples, the surface reconstruction will strain the underlying bulk and built in fields on semiconductor surfaces will create anisotropy via the electro-optic effect. Symmetry breaking can also occur through magnetic effects and other perturbations of the wave functions. Therefore, despite RDS is a surface sensitive technique, the measured data should be interpreted with extreme care. In 2005 a review by Weightman and coworkers covered the use of RDS for different environments, including the liquid–solid interface [63].

Berkovits et al. [64] applied RDS for the first time in liquid for GaAs immersed in sodium sulfide solutions. They obtained two major results: The first one is the evidence of the existence of dimers at this semiconductor surface even in a liquid environment. Using these dimers signals as a probe opened up the possibility for investigating surface chemistry of semiconductor-liquid interfaces. The effect of photobreaking surface chemical bonds has been shown by RDS when investigating As-S bonds at the passivated GaAs semiconductor surface. The group of Borenstein [65] applied RDS for the first time to the liquid–solid interface of metals (Au(110)), and already compared the data with cyclic voltammetry—and as the authors write—with *"preliminary"* in-situ ECSTM measurements. This combination of setups provided a wealth of data, the cyclo-voltammetric and the RDS data are shown in Fig. 14.14. Mazine et al. monitored the phase transition between the different Au reconstructions using RDS and obtained spectra that are associated with the three surface structures. The change in RDS for different surface reconstructions is called $\delta r/r$ and is plotted in the right part of Fig. 14.14. The authors attributed the measured anisotropy and its change at different surface reconstructions to interband transitions of Au. The main features are indeed observed close to the structures related to the interband transitions, in particular around 2.5 and 3.5 eV. The surface microscopic order is expected to have an important effect on the intensities of the RDS signal at the corresponding energies.
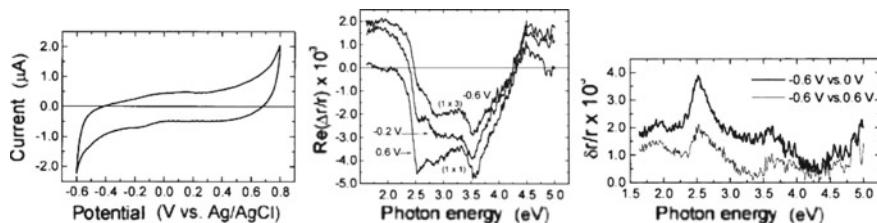
Fig. 14.14 *Left* voltammogram curve of Au(110) in 0.1 M Na2SO4 solution; sweep rate 10 mV/s. *Middle* RDS data, $\frac{\Delta \bar{r}}{\bar{r}} = \frac{4\pi i d}{\lambda} \frac{\Delta \varepsilon_0}{\varepsilon_s - 1}$ (with ε_s as the dielectric function of the Au substrate and ε_0 as the dielectric function of the overlayer) of an Au(110) surface in $0.1\,\mathrm{M}\,Na_2SO_4$ solution, for applied potentials equal to -0.6, -0.2 and 0.6 V corresponding, respectively, to a (1×3) surface, a poorly (1×2) reconstructed surface, and the (1×1) unreconstructed one. *Right* Differences between the RA spectra of the different reconstructions of Au(110) surface. Reproduced with permission by Wiley VCH from Ref. [65]

14.8 Conclusion and Outlook

The understanding and the control of the liquid–solid interface will be a major prerequisite for the solution of society related problems: generation of CO_2-free energy, maintaining a sustainable environment, using sustainable electrochemical processes, etc. and research approaches in these directions play (already) a major role. These electrochemical processes are on the one hand heavily used on an industrial scale, and on the other hand, with the exception of electric AC and DC measurements, right now hardly controlled at all. Research also focuses on improvements in the efficiency of catalysts used in important chemical processes. Improvements in existing technologies in these fields are necessary, but they will only be realized if we are able to develop new surface science techniques particularly for monitoring dynamic processes occurring in ambient conditions, especially at liquid interfaces (and also in high pressure gaseous environments). If we want to have an understanding driven development in surface technology, instead of trial and error we have to develop techniques, which allow an atomistic understanding of the reactions occurring at the liquid–solid interface. Not only solar energy conversion, but also sophisticated energy storage and delivery systems, batteries and fuel cells, corrosion science, electrodeposition and etching, electrocatalysis and others, will profit from this progress.

As long as we do not have a complete, and contradiction-free understanding of the electrochemical interface we will neither be able to apply the full potential of our understanding for the solution of the above mentioned global problems, nor to develop alternative processes converting energy and chemicals to the desired form. In the last hundred years, basically since the "Haber Bosch" process, major progress has been made in understanding surfaces in UHV, historically first of metals, then of semiconductors and recently of small molecules and polymers. The equipment developed for studying surfaces in UHV was a major enabler for the development of the modern semiconductor industry. Using first the UHV equipment for understanding semiconductor surfaces, and later, applying the concepts of surface science and

solid state physics also to interfaces, has become a success story. Compared to the multitude of surface techniques optimized for research in UHV, there are rather few techniques that can operate in ambient conditions.

In the field of chemical engineering the progress is less direct. Industrial catalysis takes place in ambient (or high pressure) conditions. Despite catalysis takes rapid progress towards the concept of "designer catalysts" by understanding the computational concepts of hyper-energy surfaces and the minimum activation energy path. Creating "designer catalysts" is still a dream and not yet realized for developing suitable active materials resulting in the wanted catalytic reaction. In catalysis the understanding, or even more important, the control of kinetic factors is central to the catalytic process. This control and hopefully widespread use can be achieved at different levels:

1. For industrial use, simple monitoring techniques (currently current–voltage based techniques) are applicable.
2. In R&D departments, techniques like CyV as well as optical reflectometry, sometimes polarization optical probes, respectively Raman techniques, are currently used on a singular basis. We expect that the progress in measurement automatization, in data taking and data evaluation by modeling, will, within a decade, transfer these slightly complex systems into a production environment, first in a fingerprint way and later with data fitting.
3. Currently the most complex techniques in university labs are based on scanning microscopy techniques, e.g. electrochemical scanning tunneling microscopy or atomic force microscopy in liquids.
4. Considered as the currently most advanced techniques researchers are developing and applying nonlinear spectroscopy systems, e.g. second harmonic and sum frequency generation. These techniques, presently measuring intensities, will be adapted towards phase sensitive (nonlinear ellipsometric) measurements. The most important advancement of these nonlinear techniques is the chance to map and understand dynamic and kinetic processes on timescales down to femtoseconds for observing the temporal behavior and relaxation of samples to follow its energy dissipation mechanisms by pump-probe studies.
5. Finally I consider also large scale research centers as a major, and currently not enough utilized resource. Especially third-generation and fourth-generation synchrotrons, using the full electromagnetic spectrum from THz to gamma rays, have played and will play a major role in understanding surface reconstructions in the liquid phase. Free electron lasers as fourth-generation sources operate with short electron bunches which has the advantage that when the length of the electron bunch is less than the wavelength of electromagnetic radiation, a coherent enhancement of the emission results. Even if water is not transparent to THz radiation, technical concepts such as coupling the long wavelength radiation either through the substrate or using waveguide structures, will allow the superposition of the signal and the pump wave. Also neutron scattering research centers will contribute enormously. The neutron scattering spectrum of H_2O can be compared to the one of D_2O. The resulting vibrational difference spectrum will show the

well known translational and vibrational bands of bulk water. Significant discrepancies arise at low frequencies indicating that the diffusive motion is strongly retarded by interactions with the surface.

What is certainly missing is the equivalent of XPS in UHV to obtain chemical sensitivity of the water-liquid covered surface. This is approached in two ways: First by using differential pumping systems precisely controlling the thickness of the liquid through its thermodynamic equilibrium vapor pressure. The other approach uses "very high energy" XPS, where the photoelectrons have such a high energy that they penetrate through the adlayer without being scattered.

Summarizing, there will be "no-one-solves-all" technique available. Only their combination and an interdisciplinary dialogue between experts in different fields aiming at a common picture of the liquid–solid interface will help to exploit dynamic processes at the liquid–solid interface in the future. After having understood the anorganic chemistry at the interface, the final aim is to apply these concepts to organic, and finally to biological processes. With creativity, new analytical tools, and an improved theoretical understanding, we can finally understand (and hopefully exploit) photosynthesis, the way how nature converts and stores energy, and master some of the upcoming social challenges with scientific and technological progress.

References

1. P.M. Sherwood, Analytica Chimica Acta **283**, 52 (1993)
2. R.M.A. Azzam, N.M. Bashara, *Ellipsometry and Polarized Light*, 2nd edn. (North Holland Elsevier, Amsterdam, 1999)
3. H.G. Tompkins , E.A. Irene, *Handbook of Ellipsometry,* 2nd edn. (Springer, Heidelberg, 2005)
4. J.F. McGilp, D. Weaire, C.H. Patterson, (eds.), *Epioptics* (Springer, Berlin, 1995), section D. Zahn
5. G. Bauer, W.R. (eds.), *Optical Characterization of epitaxial semiconductor layers* (Springer, Berlin, 1995) section W. Richter and D. Zahn
6. P.Y.Yu, M. Cardona, *Fundamentals of Semiconductors and Physics and Materials Properties*, Chap. 6, 2nd revised edn. (Springer, Berlin, 1999)
7. D. Landolt, *Corrosion and Surface Chemistry of Metals* (EFPL Press, Lausanne, 2007)
8. P.V. Zant, *Microchip Fabrication,* Chap.13, 5th edition edn. (McGraw-Hill Professional, New York, 2004)
9. J.J.A.J. Bard, R. Parsons, *Standard Potentials in Aqueous Solution* (Marcel Dekker Inc., New York, 1985)
10. http://www.doitpoms.ac.uk/tlplib/pourbaix/index.php (downloaded 5th. Nov. 2012)
11. http://www.doitpoms.ac.uk/tlplib/aqueous_corrosion/index.php (downloaded 5th. Nov. 2012)
12. E.S. Wolfgang Schmickler, *Interfacial Electrochemistry,* 2nd edn. (Springer, Berlin, 2010)
13. H. Ibach, *Physics of Surfaces and Interfaces* (Springer, Berlin, 2006)
14. A.P.G. Arthur, W. Adamson, *Physical Chemistry of Surfaces,* 6th edn. (Wiley, New York, 1997)
15. W.V.C.H. Hamann, *Elektrochemie,* 4th edn. (Wiley- VCH, Weinheim, 2005)
16. A.M.O.B. Christopher, M.A. Brett, *Electrochemistry,* 1st edn. (Oxford University Press, Oxford, 1994)
17. K. Itaya, E. Tomita, Surf. Sci. **201**, L507 (1988)
18. D. Kolb, Ber. Bunsenges. Phys. Chem. **98**, 1421 (1994)
19. U. Kreibig, Zeit. Physik D **3**, 239 (1986)

20. V. Rinnerbauer, K. Hingerl, M. Kovalenko, W. Heiss, Appl. Phys. Lett. **89**, (2006)
21. W. Kaim, A. Klein, *Spectroelectrochemistry* (Royal Society Chemistry Publishing, Cambridge, 2008)
22. F.M. Geiger, Annu. Rev. Phys. Chem. **60**, 61 (2009)
23. A. Tadjeddine, A. Peremans, Nonlinear optical spectroscopy of the electrochemical interface, in *Spectroscopy for Surface Science. Advances in Spectroscopy,*, vol. 26, ed. by R. Clark, R.F. Hester (Wiley, Chichester, 1998), p. 159
24. T.F. Young, L.F. Maranville, H.M. Smith, The structure of electrolytic, solutions, in *Raman Spectral Investigations of Ionic Equilibria in Solutions of Strong Electrolytes*, ed. by W.J. Hamer (Wiley, New York, 1959), p. 35
25. M.J. Weaver, S. Zou, Vibrational spectroscopy of electrochemivcal interfaces, some walls and bridges to surface science understanding, in *Spectroscopy for Surface Science, Advances in Spectroscopy*, vol. 26, ed. by R.J.H. Clark, R.E. Hester (Wiley, New York, 1998), p. 219
26. Z. Tian, B. Ren, Annu. Rev. Phys. Chem. **55**, 197 (2004)
27. M. Cardona, K. Shaklee, F. Pollak, Phys. Rev. **154**, 696 (1967)
28. P.H. Hahn, W.G. Schmidt, F. Bechstedt, Phys. Rev. Lett. **88**, 016402 (2001)
29. R. Kötz, D.M. Kolb, Zeitschrift für Physikalische Chemie **112**, 69 (1978)
30. H. Lüth, Solid Surfaces, Interfaces and Thin Films, Graduate Texts in Physics (Springer, Berlin, 2010)
31. T.E. Furtak, D.W. Lynch, Phys. Rev. Lett. **35**, 960 (1975)
32. R. Kofman, P. Cheyssac, J. Richard, Surf. Sci. **77**, 537 (1978)
33. D. Kolb, W. Boeck, K. Ho, Phys. Rev. Lett. **47**, 26 (1981)
34. D. Beaglehole, Experimental studies of liquid, interfaces, in *Fluid Interfacial Phenomena,* ed. by C.A. Croxton (Wiley, New York, 1986), p. 523
35. A. Hamnett, J. Chem. Soc. Faraday Trans. **89**, 1593 (1993)
36. P. Christensen, A. Hamnett, Electrochim. Acta **45**, 2443 (2000)
37. H. Arwin, Thin Solid Films **519**, 2589 (2011)
38. K. Bartel et al., J. Phys. Condens. Matter. **18**, 3535 (2006)
39. R. Teppner, K. Haage, D. Wantke, H. Motschmann, J. Phys. Chem. B **104**, 11489 (2000)
40. J. Lekner, *Theory of Reflection of Electromagnetic and Particle Waves, Developments in Electromagnetic Theory and Applications* (M. Nijhoff Publishers, Dordrecht, 1987)
41. L. Broch et al., Rev. Sci. Instrum. **78**, 064101 (2007)
42. A. Zimmer et al., Electrochim. Acta **52**, 4760 (2007)
43. A. Zimmer et al., J. Electrochem. Soc. **152**, G772 (2005)
44. M. Prato, A. Gussoni, M. Panizza, Physica Status Solidi (c) **5**, 1307–1304 (2008)
45. M.L. Walker, L.J. Richter, D. Josell, T.P. Moffat, J. Electrochem. Soc. **153**, C235 (2006)
46. M.L. Walker, L.J. Richter, T.P. Moffat, J. Electrochem. Soc. **152**, C403 (2005)
47. M.L. Walker, L.J. Richter, T.P. Moffat, J. Electrochem. Soc. **153**, C557 (2006)
48. M.L. Walker, L.J. Richter, T.P. Moffat, J. Electrochem. Soc. **154**, D277 (2007)
49. H. Rokugawa, S. Adachi, J. Electrochem. Soc. **156**, H92 (2009)
50. K. Tomioka, S. Adachi, J. Electrochem. Soc. **152**, G173 (2005)
51. N. Tomita, S. Adachi, J. Electrochem. Soc. **149**, G245 (2002)
52. S. Adachi, D. Kikuchi, J. Electrochem. Soc. **147**, 4618 (2000)
53. D. Kikuchi, Y. Matsui, S. Adachi, J. Electrochem. Soc. **147**, 1973 (2000)
54. K. Sakaino, Y. Kawabata, S. Adachi, J. Electrochem. Soc. **147**, 1530 (2000)
55. V.L. Berkovits et al., J. Electrochem. Soc. **152**, G349 (2005)
56. O.V. Sulima, A.W. Bett, J. Wagner, J. Electrochem. Soc. **147**, 1910 (2000)
57. M.H. Ubeda et al., J. Electrochem. Soc. **152**, A37 (2005)
58. S. Zangooie, R. Bjorklund, H. Arwin, J. Electrochem. Soc. **144**, 4027 (1997)
59. J. Price et al., Microelectron. Eng. **87**, 1661 (2010)
60. T. Holden et al., Appl. Phys. Lett. **70**, 1107 (1997)
61. K. Bang et al., Bull. Korean Chem. Soc. **26**, 947 (2005)
62. J.H. den Boer, G.M. Kroesen, F.J. de Hoog, Appl. Opt. **34**, 5708 (1995)
63. P. Weightman, D. Martin, R. Cole, T. Farrell, Rep. Prog. Phys. **68**, 1251 (2005)
64. V. Berkovits et al., Phys. Rev. B **54**, R8369 (1996)
65. V. Mazine, Y. Borensztein, L. Cagnon, P. Allongue, Physica Status Solidi (a) **175**, 311 (1999)

Chapter 15
Spectroscopic Ellipsometry for Functional Nano-Layers of Flexible Organic Electronic Devices

Stergios Logothetidis and Argiris Laskarakis

Abstract Flexible organic electronic devices (FEDs) will significantly improve and revolutionize several aspects our everyday life. During the last years, there are numerous advances in organic (semiconducting, conducting and insulating), inorganic and hybrid (organic–inorganic) materials that exhibit customized properties and stability, and in the synthesis and preparation methods, which are characterized by a significant amount of multidisciplinarity. The understanding of the optical and electrical properties of these materials as well as their growth mechanisms can improve the functionality and promote the performance of flexible organic electronic devices. Spectroscopic Ellipsometry (SE) is a powerful technique that can be implemented in-situ and ex-situ for the measurement and analysis of the optical response of a wide variety of materials. In this chapter, we will describe briefly some of the advances towards the implementation of SE for the study of state-of-the-art materials (flexible polymer substrates, barrier layers, transparent electrodes) for flexible organic electronics applications.

15.1 Introduction

Organic electronics is one of the most rapidly emerging sectors of the modern science and technology and it is expected to significantly improve our everyday life in the near future [1–4]. The applications for organic electronics include the devices for visualization of information, such as flexible displays (organic light-emitting diodes—OLEDs) and lighting, for generation of electricity from sunlight (organic photovoltaic

S. Logothetidis (✉) · A. Laskarakis
Lab for Thin Films-Nanosystems and Nanometrology (LTFN), Department of Physics,
Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
e-mail: logot@auth.gr

A. Laskarakis
e-mail: alask@physics.auth.gr

cells—OPVs), for electronic circuits (organic thin film transistors—OTFTs), for power generation (flexible thin film batteries), for sensing (flexible organic sensors and biosensors), for security (radio frequency identification tags—RFID), etc. [1, 4, 5]. Besides the consumer use, most of these applications will be implemented in several industrial sectors, such as in automotive industry, architecture, buildings, etc. Among the advantages of flexible electronic devices (FEDs) over the conventional electronic devices, one can mention the lightweight, thin and conformable design, the ability to be rolled when they are not in use, and the low cost [6, 7].

This field is characterized by fast paced progress concerning the efficiency and the device function. This progress is mostly determined by the innovations and advances in materials and processes, both at lab and production scale. The advances in materials include the synthesis and thin film preparation of organic semiconductors (small molecule and polymers) and transparent organic electrodes with high electrical conductivity that can be deposited in thin film form by solution-based processes (e.g. printing). Moreover, ultra barrier materials are required in order to encapsulate the active layers (organic semiconductors) and the transparent electrodes from the atmospheric gas (H_2O and O_2) molecule permeation that induces corrosion, and eventually the device malfunction [8–12]. The above functional materials can be deposited by wet, printing and vacuum processes onto flexible polymeric substrates, such as PolyEthylene Terephthalate (PET), and PolyEthylene Naphthalate (PEN) with the form of rolls, which will replace the rigid Si and glass substrates. The thin film preparation of organic semiconducting and electrode materials onto plastic substrates that have the form of web rolls will allow the cost-effective, roll-to-roll (r2r) processing in high volumes for the production of FEDs in large scale [13–15].

The performance, efficiency and lifetime of FEDs are greatly affected by the optical, electrical, and structural properties of their individual nano-layers. These should meet specific requirements, such as high optical transparency, high electrical conductivity, structural stability, ultra low atmospheric gas permeability, film-substrate adhesion, etc. [16]. These properties are defined by the materials bonding structure, and on the surface and interface nanostructure and chemistry [15]. The detailed knowledge of the optical properties of the organic semiconducting, transparent electrodes, barrier layers and flexible polymer substrates in the infrared (IR) to the visible and to the far ultraviolet (Vis–fUV) spectral region are of significant importance since it will contribute to the understanding of the bonding and electronic structure and microstructure, as well as on the optical transparency and structure-property relationships. In situ and real-time Spectroscopic Ellipsometry (SE) is a powerful, non-destructive and surface sensitive optical technique that has been used for the investigation of optical properties, the deposition rate and the growth mechanisms of inorganic and organic thin films [17–20]. The implementation of real-time SE monitoring and control to large scale production of functional thin films for numerous applications, will lead to the optimization of the materials quality and increase in production yield [15, 21]. In the following, we will provide an overview of the implementation of SE for the optical characterization of state-of-the-art materials for FEDs in order to prove the potentiality of SE as an important tool to be implemented in industrial processes for the fabrication of advanced materials for flexible

organic electronic devices in combination to thin film quality control. Also, some of the latest advances on the organic electronic materials (flexible polymeric substrates, barrier materials, electrodes and organic semiconducting materials) will be discussed in terms of their optical properties that can provide information on their structure and to other functional properties.

15.2 Flexible Organic Electronic Devices

Among the most important applications of flexible organic electronics are the organic photovoltaic cells (OPVs) and organic light emitting diodes (OLEDs). A typical structure of a flexible organic electronic device consists of a multi-layer stack consisted of specific active and passive layers [2, 22, 23]. The active materials are the organic semiconductor layers that perform the light-emitting (in the case of OLEDs) or charge-generating (in the case of OPVs) functionalities [15, 22]. These are deposited and encapsulated onto flexible polymeric substrates, such as PET or PEN.

Organic photovoltaics (OPVs) (Fig. 15.1) will be used for solar energy harvesting and conversion. It is expected that the 1 % of the global energy up by 2015 and the 10 % until 2020 will be generated by OPVs [5, 24, 25]. However, their cost should be reduced by a factor of 10 and their efficiency to reach the values of the inorganic PVs ($\sim$22.7 % on 780 cm^2 area for crystalline Si) [25]. The most efficient kind of OPV is the bulk heterojunction device that comprises of two organic materials (an electron acceptor and a polymer) sandwitched between the anode and the cathode [24, 25]. The state-of-the-art materials for OPVs are blends of poly(p-phenylene vinylene) derivatives (PPVs) or poly(3-hexylthiophene) (P3HT) (for exciton generation and hole transport) and the organo-soluble fullerene derivative (PCBM) (for electron transport) [15, 26, 27]. These conjugated molecules are electronically active because of their highly polarizable π-systems [28, 29]. In order to reach power conversion efficiencies of 5 % (1 cm^2 area) advanced device architectures are developed including many steps between photon absorption and electricity generation [28, 30, 31].
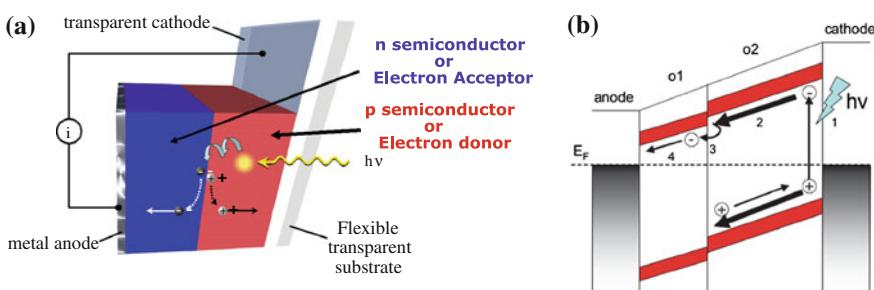
Fig. 15.1   a Schematic representation of an OPV device, **b** the four-step operation principle of a bilayer OPV device [15]
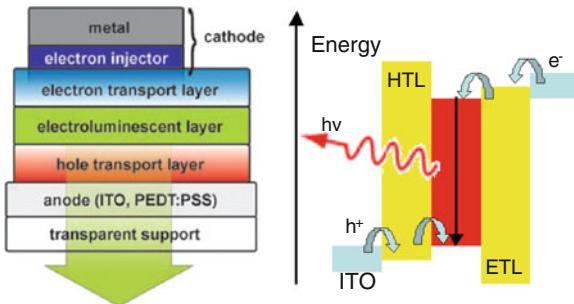
Fig. 15.2 Schematic energy-level diagrams of a modern multilayer OLED

During the photon absorption, an electron is excited from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). This electron-hole pair (exciton) relaxes with a binding energy between 0.1 and 1.4 eV (in contrast to inorganic semiconductors which is a few meV) [28]. The bound excitons must migrate to an interface where there is a sufficient chemical potential energy drop to drive dissociation into an electron-hole pair that spans the interface across the donor (material with low electron affinity) and acceptor (high electron affinity). After dissociation, each charge carrier must be transported through the device to the appropriate contact while avoiding traps and recombination. The control of the percolating pathways, in combination with broad band light absorption, long live excited states and high charge carrier mobilities are essential for high efficiency, whereas blend morphology is critical [5, 27, 32, 33].

Another major application of FEDs are the flexible OLEDs [16, 22]. A schematic energy-level diagram of a multilayer OLED can be seen in Fig. 15.2 [15, 35]. In OLEDs, the active organic (electroluminescent-EL) layer is formed in the middle of a high work-function (ϕ1) anode and a low-work-function (ϕ2) cathode. The anode materials should allow easy hole injection and it is consisted by transparent conductive oxides (TCOs) such as Indium Tin Oxide (ITO) ($\phi = 4.7$ eV) or Poly(3,4-ethylenedioxythiophene) poly(styrenesul-fonate (PEDOT:PSS). The cathode should allow easy electron injection and it must have band gap values higher than the one of the TCO. It is consisted of metals such as Ca, Mg, and Cd, whereas the typical metals such as Ag and Al have work function values of 5.1 and 4.3 eV, respectively. By applying an external driving voltage, electrons are injected into the conduction band and holes into the valence band of a semiconducting polymer. Upon injection from the electrodes, electrons and holes self-localize to form negative and positive polarons, which travel under the apparent electric field in opposite directions. When two oppositely charged polarons meet, they can form bound electron-hole pairs (excitons), which produce photons [22, 34, 35].

Despite the advances in organic electronics materials, there are significant issues that have to be addressed since they restrict the performance and stability of the organic electronic devices. Some of the main challenges are the tuning of the organic semiconductor morphology, the increase of the charge mobility and the optimization

of the barrier response of the barrier materials that are used for the encapsulation of the device (as it will be discussed below) [1, 15].

15.3 Spectroscopic Ellipsometry for Optical Characterization

Spectroscopic Ellipsometry (SE) is a powerful technique for the measurement of the bulk materials and surfaces, nanomaterials, thin film, their thickness and crystallite size and optical constants such as refractive index or dielectric function [17–20]. SE is used for characterization of all types of materials, such as dielectrics, semiconductors, metals, organic materials, opaque, semitransparent or even transparent materials [19, 20]. This technique and its instrumentation relies on the fact that the reflection of a dielectric interface depends on the polarisation of the light while the transmission of light through a transparent layer changes the phase of the incoming wave depending on the refractive index of the material [17, 19, 20]. Applications include the accurate thickness measurement of thin films, the identification of the optical properties of materials and thin films, the evaluation of vibrational, electronic, compositional and nanostructural properties, and the optical characterization of surfaces and interfaces [17, 19, 20].

SE determines the optical properties of bulk materials and thin films as a function of the photon energy ω. The determined properties are the complex refractive index $\widetilde{n}(\omega)$ and the complex dielectric function $\widetilde{\varepsilon}(\omega)$. The complex refractive index is related to dispersion and absorption of light from the medium by the relation:

$$\widetilde{n}(\omega) = n(\omega) + ik(\omega) \tag{15.1}$$

The complex dielectric function $\widetilde{\varepsilon}(\omega)$ is actually the quantity directly related to the material properties, and is connected to the refractive index through the following equation:

$$\widetilde{\varepsilon}(\omega) = \varepsilon_1 + i\varepsilon_2 \equiv \widetilde{n}^2(\omega) = (n + ik)^2 \Rightarrow \begin{cases} \varepsilon_1 = n^2 - k^2 \\ \varepsilon_2 = 2nk \end{cases} \tag{15.2}$$

In the case of the reflection of a light beam at the planar interface between two semi-infinite optically isotropic media 0 and 1, when the light beam does not penetrate the medium (1), either due to its high absorption coefficient or its infinite thickness, we are referred to a two-phase (ambient-substrate) system, or a bulk material surrounding by medium (0). In this case the ratio of the p–, s– Fresnel reflection coefficients, namely the complex reflection ratio is the quantity measured directly by SE and it is given by the expression:

$$\widetilde{\rho} = \frac{\widetilde{r}_p}{\widetilde{r}_s} = \left| \frac{\widetilde{r}_p}{\widetilde{r}_s} \right| e^{i(\delta_p - \delta_s)} = \tan\Psi e^{i\Delta}. \tag{15.3}$$
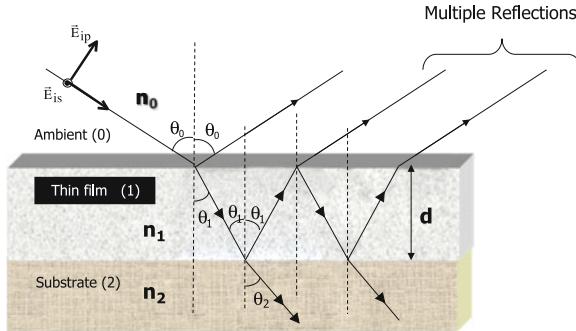
Fig. 15.3 Schematic representation of the light reflection on the surface of a thin film that is deposited onto a bulk substrate

In this expression, Ψ and Δ are the ellipsometric angles, and for a bulk material take values $0° < \Psi < 45°$ and $0° < \Delta < 180°$. From an ellipsometric measurement, the complex reflection ratio $\widetilde{\rho}$ is estimated, through the calculation of amplitude ratio $\tan\Psi$ and the phase difference Δ. From these two quantities one can extract all the other optical constants of the material. For example, the complex dielectric function of a bulk material with smooth surfaces is directly calculated by the following expression [19, 20]:

$$\widetilde{\varepsilon}(\omega) = \varepsilon_1 + i\varepsilon_2 = \widetilde{\eta}^2(\omega) = \widetilde{\varepsilon}_0 \sin^2\theta_0 \left\{ 1 + \left[\frac{1 - \widetilde{\rho}(\omega)}{1 + \widetilde{\rho}(\omega)} \right]^2 \tan^2\theta_0 \right\}, \qquad (15.4)$$

where θ is the angle of incidence of the beam, and $\widetilde{\varepsilon}_0$ the dielectric constant of the ambient medium (for the case of air $\widetilde{\varepsilon}_0 = 1$).

In the case of a film deposited on a substrate (see Fig. 15.3), we have the three-phase model where the film with thickness d [medium (1)] is confined between the semi-infinite ambient [medium (0)] and the substrate [medium (2)]. The complex reflectance ratio is defined as [19]:

$$\widetilde{\rho} = \frac{\widetilde{R}_p}{\widetilde{R}_s} \qquad (15.5)$$

$$\widetilde{R}_p = \frac{\widetilde{r}_{01p} + \widetilde{r}_{12p} e^{i2\beta}}{1 + \widetilde{r}_{01p} \widetilde{r}_{12p} e^{i2\beta}}, \qquad (15.6)$$

$$\widetilde{R}_s = \frac{\widetilde{r}_{01s} + \widetilde{r}_{12s} e^{i2\beta}}{1 + \widetilde{r}_{01s} \widetilde{r}_{12s} e^{i2\beta}}, \qquad (15.7)$$

where, $\widetilde{r}_{01i}$ and $\widetilde{r}_{12i}$ ($i = p, s$) are the *Fresnel* reflection coefficients for the interfaces between medium (0) and (1), and medium (1) and (2), respectively. These depend on film thickness due to the multiple reflections of light between the media (0)-(1)

and (1)-(2), through the phase angle β given by [19]:

$$\beta = 2\pi (d/\lambda)(n_1^2 - n_0^2 \sin^2\theta_0)^{1/2}, \tag{15.8}$$

where λ is the wavelength, θ is the angle of incidence and n_0, n_1 are the complex refraction indices of the ambient and the film, respectively. Thus, the measured quantity is the pseudodielectric function $\langle\varepsilon(\omega)\rangle (= \langle\varepsilon_1(\omega)\rangle + i\langle\varepsilon_2(\omega)\rangle)$ which contains information also about the substrate and the film thickness. As a result, the phase angle β diminishes in the energy region of high absorption, leading to $\widetilde{R}_i = \widetilde{r}_{01i} = \widetilde{r}_i$, (i = p, s) and consequently to $\langle\rho\rangle = \rho$. Thus, at the absorption bands we obtain information only of the optical properties of the bulk of the film.

The real $\varepsilon_1(\omega)$ and the imaginary $\varepsilon_2(\omega)$ parts of the dielectric function are strongly related through the well-known Kramers–Kronig relation (is based on the principle of the causality [19, 20, 36]:

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' \varepsilon_2(\omega')}{\omega'^2 - \omega^2} d\omega' \tag{15.9}$$

$$\varepsilon_2(\omega) = -\frac{2\omega}{\pi} P \int_0^\infty \frac{\varepsilon_1(\omega') - 1}{\omega'^2 - \omega^2} d\omega', \tag{15.10}$$

where, P means the principal value of the integral around the characteristic of the material electronic resonance ($\omega' = \omega$) and

$$\varepsilon_1(\omega = 0) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\varepsilon_2(\omega')}{\omega'} d\omega' \tag{15.11}$$

is the static dielectric function, the material strength (deviation from the strength of vacuum $\varepsilon_0 = 1$), that describes all losses in the whole electromagnetic spectrum in the material due to the electron absorption.

The optical response of the thin films can be deduced by the parameterization of the measured $\langle\varepsilon(\omega)\rangle$ by the use of appropriate theoretical models. One of these models is the damped harmonic oscillator (Lorenz model), which is described by the expression [17, 18, 36]:

$$\widetilde{\varepsilon}(\omega) = 1 + \frac{f\,\omega_0^2}{\omega_0^2 - \omega^2 + i\Gamma\omega}, \tag{15.12}$$

where ω is the energy of light and ω_0 is the absorption energy of the electronic transition. The constants f and Γ denote the oscillator strength and the damping (broadening) of the specific transition, respectively. The quantity $\varepsilon_1(\omega = 0)$ is given by the relation $\varepsilon_1(\omega = 0) = 1 + \omega_p^2 / \omega_0^2 = f$, which is the static dielectric constant

and represents the contribution of the electronic transition that occurs at an energy ω_0 in the NIR-Visible-UV energy region, on the dielectric function [18] and ω_p is the plasma energy.

In the case that more than one electronic transition occurs, their contribution in $\varepsilon_1(\omega = 0)$ is accounted by the summation $\varepsilon_1(\omega = 0) = 1 + \Sigma f_i$ where Σf_i describes the losses in the material in the whole electromagnetic region due to the electronic transitions. In the case of semiconducting materials, interband electronic transitions take place due to the interaction between the electromagnetic radiation (photons) and the matter (electrons). According to classical Lorentz oscillator model, the dielectric function is given by Eq. 15.12. However, one of the models that are used for the modeling of the measured dielectric spectra of amorphous semiconductors is the Tauc–Lorentz (TL) model [37]. This is based on the combination of the classical Lorentz dispersion relation and the Tauc density of states in the proximity of the fundamental optical gap ω_g. This results to an asymmetrical Lorentzian lineshape for the imaginary part $\varepsilon_2(\omega)$ of the dielectric function. The TL dispersion model is described by the following relations in which the real part $\varepsilon_1(\omega)$ is determined by the imaginary part $\varepsilon_2(\omega)$ by the Kramers–Kroning integration (see Eq. 15.13) [37]:

$$\varepsilon_2(\omega) = \begin{cases} \dfrac{A\,\omega_0\,C(\omega - \omega_g)^2}{(\omega^2 - \omega_0^2)^2 + C^2\,\omega^2}\dfrac{1}{\omega} & \omega > \omega_g \\ 0 & \omega \le \omega_g \end{cases} \tag{15.13}$$

where ω_g is the fundamental band gap energy, A is related to the transition probability, ω_0 is the Lorentz resonant energy and C the broadening term, which is a measure of the materials disorder [37].

In the following, we present some representative examples of the investigation of the optical properties of various state-of-the-art materials by SE. These materials are flexible polymeric substrates, and functional thin films used for flexible organic electronic device applications.

15.4 Optical Properties and Electronic Structure of Flexible Polymer Substrates

Among the flexible polymeric substrates that are used as substrates for the deposition of active and passive layers by wet, vacuum and printing methods, are the polyester sheets of Poly(Ethylene Terephthalate)—PET and Poly(Ethylene Naphthalate)—PEN [13, 15, 38, 39]. In addition, Polyimide (PI) and Polycarbonate (PC) are candidates to be used as substrates, whereas another potential substrate for flexible organic electronics is paper [40]. The latter has been used for printing antennas, and as a support for sensors and various types of displays. However, the paper surface is very rough compared to the other polymeric films, which limits its use. Also, its
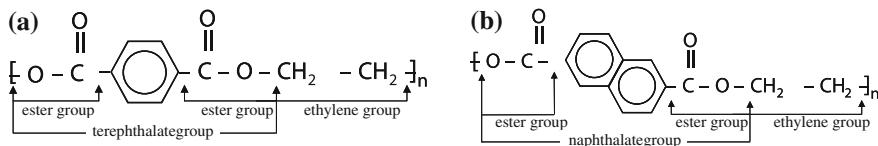
Fig. 15.4 Schematic representation of the monomer units of: **a** PET and **b** PEN

temperature capability is limited, and it is not generally compatible with solution processing processes, such as printing [40].

The potentiality of the use of PET and PEN to be used as flexible substrates for FEDs is justified by the combination of the desirable physical properties that they exhibit. These properties include easy r2r processing, good mechanical and chemical stability and reasonably good resistance to oxygen and water vapour penetration [9, 14, 41, 42] The monomer units of PET and PEN are schematically shown in Fig. 15.4 [13, 14]. The unit cell of PET, which is triclinic with a density of 1.455 g/cm^3, presents a C$_{2h}$ point symmetry, and consists of an aromatic ring and an ester function that form the terephthalate group, and by a short aliphatic chain that constitutes the ethylene segment. PEN exhibits large similarities with PET and it has a triclinic unit cell with the addition of a second phenyl ring, forming the naphthalene group.

The roll-to-roll production of these polymers in the form of flexible sheets with thickness in the range between 12 and 100 μm includes the extrusion and/or stretching process in order to achieve the necessary thickness and other desirable mechanical properties [1, 43]. During this stretching process, their macromolecular chains obtain a preferred orientation towards the stretching direction (otherwise called the Machine Direction-MD), resulting to a composite material structure consisted of oriented ("crystalline-like") regions embedded in a non-oriented (non-crystalline-like) matrix [13, 14].

On the other hand, the optical and electronic response of these materials is greatly affected by the interrelation between the optical properties of a single molecule and its local environment as well as to the structural arrangement of the monomer units and furthermore of the macromolecular chains [1, 13, 14]. Therefore, the investigation of the optical properties of PET and PEN and its relation to their bonding structure and to the degree of the macromolecular orientation can contribute to the understanding of the mechanisms that take place during the deposition of active and passive materials onto their surfaces.

Spectroscopic Ellipsometry from the IR to the Vis–fUV spectral regions is a powerful tool to investigate the optical properties of PET and PEN. The optical response functions of PET and PEN films measured in the Vis–fUV spectral region (1.5–6.5 eV) with an angle of incidence of 70° and with the plane of incidence parallel to the MD, are shown in Fig. 15.5.

The measured $\langle \varepsilon(\omega) \rangle$ of PET at energies above 4 eV, is dominated by four characteristic features. These are located at 4.15 (called as peak I) and 4.3 eV (peak II), whereas two stronger bands are centered at around 5.0 (peak III) and 6.3 eV (peak IV).
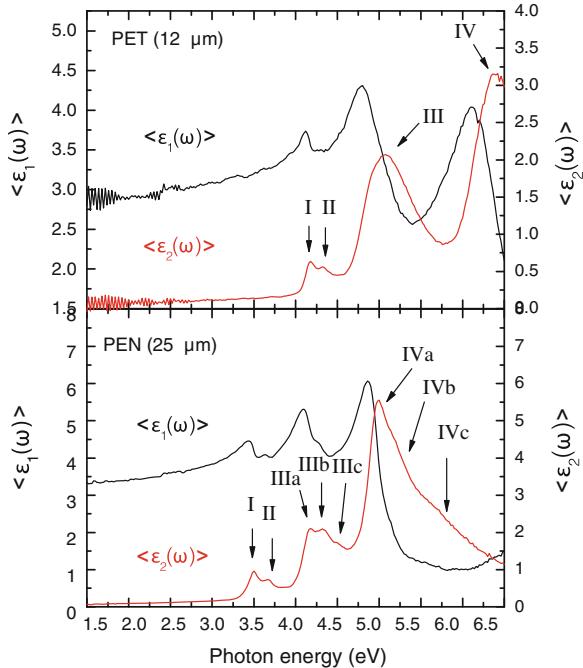
Fig. 15.5 Measured pseudo-dielectric function $\langle\varepsilon(\omega)\rangle$ of PET and PEN flexible substrates [44]

The measured $\langle\varepsilon(\omega)\rangle$ of PEN shows similar features due to the similarities in the molecular structure between the monomer units of PET and PEN. Also, an energy shift of peaks I and II to 3.4 and 3.6 eV, respectively, is observed although this shift is much more pronounced in the case of peaks III and IV, which are appeared at 4.4 and 5.1 eV, respectively, along with a characteristic split in three components [13, 14, 44].

As it has been discussed elsewhere, the peaks I and II have been attributed to the electronic transition of the non-bonded electron of the carbonyl O atom from the n state to the π^* unoccupied valence state orbital (n $\rightarrow$ π^* transition) [13, 14]. The peak III of PET, which is possibly attributed to the spin-allowed, orbitally-forbidden $^1A_{1g} \rightarrow {}^1B_{1u}$ transition, is composed by two sub-peaks of parallel polarization dependence. In the case of peak III of PEN, this is attributed to the same electronic transition and it is composed by three sub-peaks. Finally, the peak IV in both polymer substrates is composed by two sub-peaks with different polarization at 6.33 and 6.44 eV for PET (4.8 and 5.5 eV for PEN) after molecular orbital calculations, and it can be attributed to the $^1A_{1g} \rightarrow {}^1B_{1u}$ electronic transition of the paradistributed benzene (naphthanene) rings with polarization rules on the plane of the rings [1, 13, 14, 44].

Besides the electronic transitions, SE can be applied to study the optical anisotropy of PET and PEN substrates by the analysis of the $\langle\varepsilon(\omega)\rangle$ spectra measured at different angles between the plane of incidence and the MD. Their optical anisotropy is
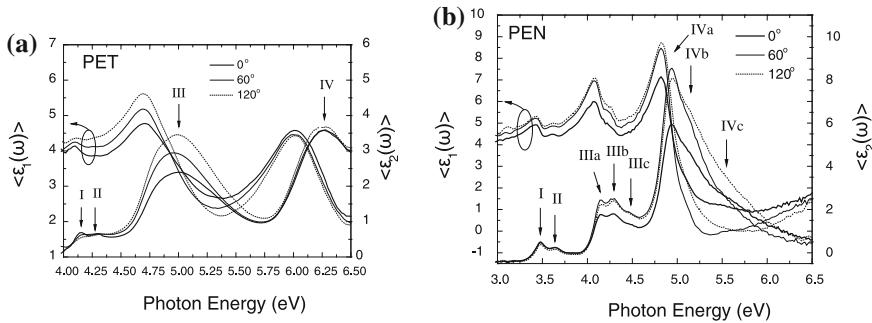
Fig. 15.6 Measured $\langle\varepsilon(\omega)\rangle$ of: **a** PET and **b** PEN flexible substrates at various angles between the plane of incidence and the MD

the result of the arrangement of macromolecular chains during their production process. This arrangement results to the formation of oriented crystalline-like regions embedded in a non-oriented non-crystalline matrix. Therefore, the polymer films can be regarded as effective dielectric media, composed by oriented and non-oriented regions, and their optical response can be represented by a general dielectric tensor containing information about both the optical response of the macromolecular chains and their orientation with respect to the MD. In order to investigate the optical anisotropy of PET and PEN, the dependence of the $\langle\varepsilon(\omega)\rangle$ with the angle between the plane of incidence and the MD (angle φ) has been analyzed. Figure 15.6 shows the $\langle\varepsilon(\omega)\rangle$ of the PET and PEN substrates measured at various φ angles [13, 14].

The measured $\langle\varepsilon(\omega)\rangle = \langle\varepsilon_1(\omega)\rangle + i\langle\varepsilon_2(\omega)\rangle$ has been analyzed using damped harmonic Lorentzian oscillators, Eq. (15.12), in combination to two-phase (air/bulk material) model as a function of the angle φ. As it has been reported, the deter-mined values of the oscillator strength f corresponding to peaks I to IV has a harmonic azimuthal dependence, with a period of half a complete rotation (180°), due to the light interaction with oriented and non-oriented regions. The azimuthal dependence of f_I and f_{II} is similar in both PET and PEN. This is in agreement with the assignment of peaks I and II to the n $\rightarrow \pi^*$ electronic transition of the carbonyl group which is dominated by selection rules perpendicular to the macromolecular chains. Moreover, all the sub-peak components of peak III are characterized by same parallel azimuthal dependence, as a result of the $\pi \rightarrow \pi^*$ electronic transition selection rule parallel to the MD. Finally, the azimuthal dependence of f_{IV} appears different between PET and PEN.

The above support the fact that PET and PEN are characterized by biaxial optical anisotropy and their optical response can be approximated as the response of a uniaxial material, with its optic axis parallel to its surface. This is also proved by the calculation of the refractive index n at the transparent region, or more precisely at $n^2(\omega) = \varepsilon_1(\omega \approx 0)$. In the case of a uniaxial material, with its optic axis parallel to the surface, the dependence of $n(\omega = 0)$ with the angle φ is shown in Fig. 15.7 ($n_\parallel$ and $n_\perp$ are the principal values of the refractive index). In this analysis, the
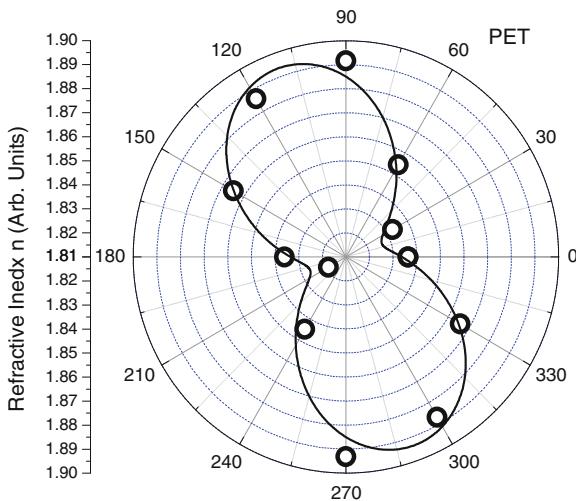
Fig. 15.7 Dependence of the determined refractive index n of PET with the angle between the MD and the plane of incidence of light

formulation $\varphi \rightarrow \varphi + \Delta\varphi$ has been used to shift from the polymer film axis system to the optic axis system. $\Delta\varphi$ represents the angle between the optic axis, (high symmetry axis corresponding to macromolecular chains), and the MD. It is clear that there is excellent agreement between the calculated values of n and the fit, justifying the approximation of PET and PEN as uniaxial materials with their optic axes parallel to their surface. Furthermore, by the azimuthal dependence of n, the angle between the optic axis and the MD can be determined. By the analysis, we obtain: $n_{\perp PET} = 1.680$, $n_{\parallel PET} = 1.730$, $\Delta\varphi_{PET} = 5.30° \pm 0.70°$, $n_{\perp PEN} = 1.805$, $n_{\parallel PEN} = 1.857$ and $\Delta\varphi_{PEN} = 18.98° \pm 1.44°$ [13, 14].

The optical response of PET and PEN films in the IR spectral region can be measured by the use of FTIRSE as it is shown in Fig. 15.8. The strong absorption bands, at 900–1800 cm^{-1}, indicate the contribution of the vibrational modes corresponding to the IR-active chemical bonds of PET and PEN and it shows several bonding vibration modes. Among the more intense characteristic vibration bands in the FTIRSE spectra of PET, one can observe the vibration modes at around 940 and 971 cm^{-1} (trans) that are attributed to the C–O stretching mode, the aromatic CH$_2$ stretching mode at 1125 cm^{-1}, and the ester mode at 1255 cm^{-1}. Also, the characteristic vibration band at 1720 cm^{-1} corresponding to the stretching vibration of the carbonyl C=O groups is shown [14, 43].

The vibration bands in PEN include the one at 1098 cm^{-1} that has been attributed to the stretching and bending modes of ethylene glycol attached to the aromatic structures of the PEN monomer units. Moreover, the characteristic band at 1184 cm^{-1} corresponds to the C–C stretching modes of the naphthalene group. The complex bands at 1335 and 1374 cm^{-1} reveal the bending mode of the ethylene glycol CH$_2$
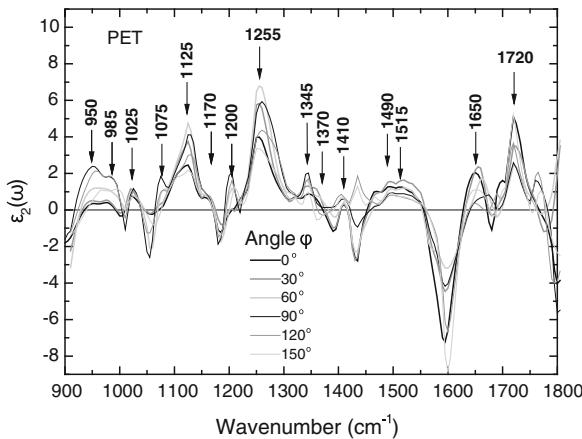
Fig. 15.8 Measured $\langle\varepsilon_2(\omega)\rangle$ of PET in the IR spectral region by FTIRSE. The measurements were performed at various angles between the MD and the plane of incidence of light

group in the *gauche* and *trans* conformations, respectively. The C=C stretching modes of the aromatic (naphthalene) ring structures of PEN are shown at 1635 cm^{-1}. Moreover, the stretching vibration of the carbonyl C=O group appears in lower energy in case of PEN (1713 cm^{-1}) than in PET (1720 cm^{-1}) [14, 43].

   In summary, the optical properties and anisotropy of stretch-oriented PET and PEN films can be investigated by SE from the IR to the Vis–fUV spectral region with detail. The analysis of the SE spectra measured at various angles between the plane of incidence and the MD can lead to the identification and assignment of the characteristic features corresponding to the different electronic transitions and molecular vibrational modes. Also, by the analysis of the refractive index n with a uniaxial model, PET and PEN substrates can be treated as uniaxial materials with the optic axis parallel to their surface. The above contribute to the better understanding of the structure-property relationships during the incorporation of PET and PEN films to the large-scale production of flexible electronic devices.

15.5 Barrier Materials for Device Encapsulation

Although the optimization of the photon and electric charge generation functionalities of the organic semiconductors is the primary scientific and technological challenge, the protection of the entire multilayer structure of the organic electronic device from corrosion is a main concern. The corrosion of the organic semiconductor materials (polymers, small molecules) takes place due to the penetration of atmospheric gas molecules (O_2, H_2O) through the device passive layers (polymer substrates, transparent electrodes) [1, 15]. Another result of the atmospheric gas molecule permeation is the organic film delamination that also leads to the failure of the organic electronic device [8–10]. The permeation is controlled by the defects of the flexible polymeric

substrate and of the barrier nano-layers. These defects are the result of the intrinsic surface roughness, the surface and structure in-homogeneities and by cracks of the inorganic nano-layers created by the bending and/or tension of the flexible polymeric substrates [8, 12, 45].

Therefore, the active layers should be encapsulated with transparent and non-conductive layers in order to sustain their structure and performance during a considerable lifetime and to ensure a stable device operation. The commercially supplied polymer films that are used as flexible substrates, such as PET and PEN, exhibit gas permeability values that are in the range of 10^{-1} to 10^{2} cm^3/m^{2}dbar (for oxygen transmission rate—OTR) and g/m^{2}d (for water vapor transmission rate—WVTR). These values are sufficient only for some food packaging applications [5, 8, 10]. On the other hand, the requirements for FED applications ask for OTR and WVTR values that are lower by another 3 orders of magnitude; that is in the range below 10^{-5} cm^3/m^{2}dbar (OTR) and g/m^{2}d (WVTR) [45]. Therefore, transparent barrier layers have to be deposited onto the flexible polymeric substrates [15].

A common approach for encapsulation materials is the deposition of inorganic thin films, as SiO$_x$ and AlO$_x$ onto the flexible polymeric substrate. The deposition of an inorganic thin film of thickness of 50 nm onto a 50 μm polymer substrate reduces the O$_2$ and H$_2$O permeation rates by two to three orders of magnitude from the permeation rate values of the uncoated plastic substrate. These thin films can be prepared onto rigid and flexible substrates by vacuum processes, such as dc or rf sputtering, as well as by electron beam evaporation [21].

Another class of materials that can be used as barrier layers includes the hybrid (inorganic–organic) barrier materials, alternatively referred as Ormocer$^{®}$ [12, 46, 47]. These can be synthesized via the sol-gel processes from organoalkoxysilanes, and they have strong covalent or ionic-covalent bonds between the inorganic and organic components. One of the main factors that affect the barrier response of hybrid polymers is the crosslinking between the inorganic and the organic components as well as its adhesion on the substrate [21, 46–48].

In the following, we will discuss on some representative examples about the real-time investigation of the optical properties of the growing barrier thin films (AlO$_x$, SiO$_x$) onto PET and PEN flexible polymeric substrates. The investigation of the optical properties has been performed in the Vis–fUV energy region (1.5–6.5 eV) by an ultra-fast multi-wavelength phase modulated SE unit that were adapted onto the UHV chamber at an angle of 70°. This system is equipped with a 32-Fiber-optic array detector for simultaneous measurements at 32 different wavelengths (MWL mode), in the energy range 3–6.5 eV.

Figure 15.9 shows the imaginary part $\langle \varepsilon_2(\omega) \rangle$ in the 3–6.5 eV region measured during the deposition of AlO$_x$ thin film onto PET flexible substrate by e-beam evaporation. The $\langle \varepsilon_2(\omega) \rangle$ of PET substrate is shown with solid circles (•), whereas the gray lines represent individual "snapshots" of the optical properties of the growing AlO$_x$ film onto the of PET substrate. The last spectrum composed of hollow squares (□) is measured after the end of the deposition process and it represents the final optical response of the AlO$_x$/PET system. The duration of the deposition process is 240 s, whereas the thickness of the AlO$_x$ film is finalized at 33 nm [49].
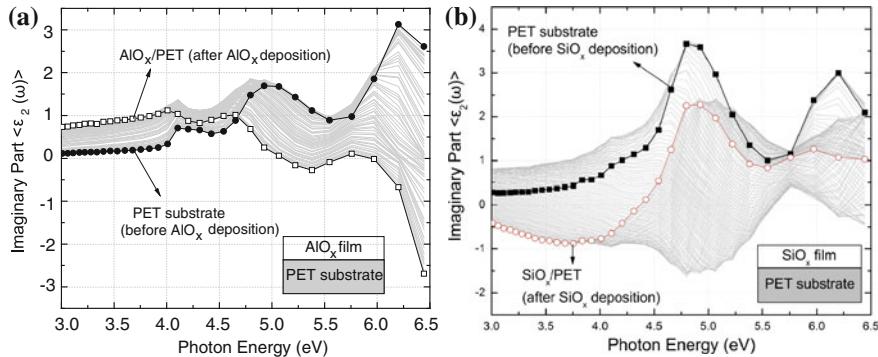
Fig. 15.9 Real-time measured $\langle\varepsilon_2(\omega)\rangle$ spectra during the deposition of: **a** AlO$_x$ and **b** SiO$_x$ thin film onto PET flexible substrate by e-beam evaporation

The determination of the optical and electronic properties of the AlO$_x$ thin films has been realized by the analysis of the measured pseudo-dielectric function $\langle\varepsilon(\omega)\rangle$ with the TL dispersion model (Eq. (15.13)) [37]. The geometrical structure that has been used for the analysis procedure includes a three-phase model that consists of air, the inorganic layer of thickness d and the flexible polymeric substrate (shown in the inset of Fig. 15.9a).

The AlO$_x$ films are characterized by a high optical transparency in the extended UV-vis energy range and due to this transparency of AlO$_x$ materials, we have obtained only a small spectral dependence on the $\langle\varepsilon(\omega)\rangle$. Figure 15.10 shows the dependence of the calculated ω_g and ω_0 values of the growing AlO$_x$ film as a function of the deposition time. The ω_g values follow an almost constant dependence with the deposition time, with a slight reduction starting from ~5.5 to ~4.8 eV. This leads to the conclusion that the AlO$_x$ films' transparency slightly decreases with increasing film
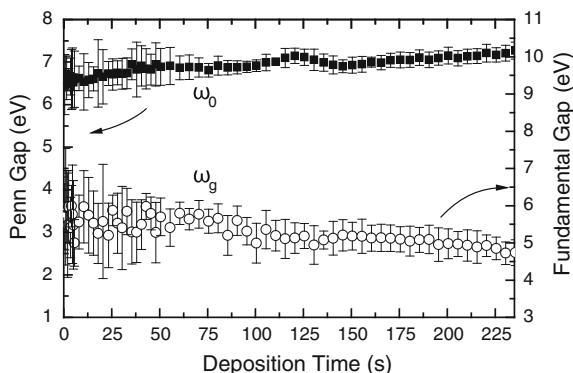
Fig. 15.10 Evolution of the calculated Penn gap ω_0 and fundamental gap ω_g of the AlO$_x$ thin film as determined from the analysis of the real-time $\langle\varepsilon(\omega)\rangle$ spectra
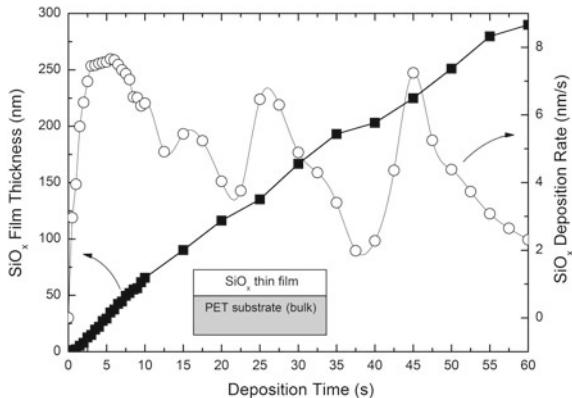
Fig. 15.11 Evolution of the SiO_x film thickness with the deposition time onto PET as determined by in-situ and real-time SE, and calculated deposition rate of the SiO_x film, which indicates an island-type growth mechanism of SiO_x onto PET [21, 49]

thickness. On the other hand, the Penn gap ω_0 of AlO_x increases only slightly from 6.5 to 7.25 eV. The calculated maximum electronic absorption (Penn gap ω_0) values are found below the 12.4 eV of the maximum electronic absorption of the stoichiometric AlO_x (where x = 1.5) [21, 49]. The optical and electronic properties of the AlO_x film have been found to be almost constant throughout the entire deposition process towards the final thickness value of $\sim$33 nm. These AlO_x films are suitable to be used as gas barriers for flexible organic electronic devices encapsulation [21, 49].

Another inorganic material that is used as a barrier coating for the encapsulation of the active layers of flexible organic electronic devices is SiO_x. Figure 15.9b shows the imaginary part $\langle \varepsilon_2(\omega) \rangle$ in the 3–6.5 eV region measured during the deposition of SiO_x thin film onto PET flexible substrate by e-beam evaporation. The determination of the optical and electronic properties of the SiO_x thin films has been realized by the analysis of the measured pseudo-dielectric function $\langle \varepsilon(\omega) \rangle$ with the TL dispersion model (see Eq. (15.13)) [37]. The geometrical structure that has been used for the analysis procedure includes a three-phase model that consists of air, the inorganic layer of thickness d and the flexible polymeric substrate. Based on the analysis of the $\langle \varepsilon(\omega) \rangle$ spectra, the evolution of the SiO_x films thickness and optical properties can be evaluated with the deposition time. The best-fit parameters for the evolution of the SiO_x film thickness d and the deposition rate (derivative of the thickness as a function of the deposition time), are shown in Fig. 15.11.

The study of the deposition rate provides information about the growth mechanism of SiO_x nano-layers and how they are affected by the chemical structure and surface properties of the substrates [21].

As it can be seen in Fig. 15.11, the deposition rate of SiO_x onto PET is higher than in the case of SiO_x/PEN. In order to enhance the information deduced from the
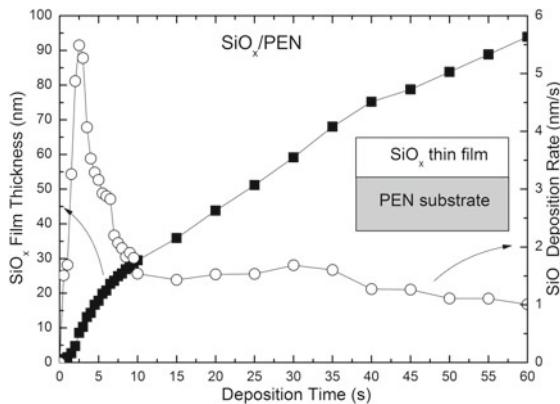
Fig. 15.12 a Evolution of the SiO$_x$ film thickness with the deposition time onto PEN substrate as determined by in-situ and real-time SE, and calculated deposition rate of the SiO$_x$ film, which shows that the growth mechanism of SiO$_x$/PEN is dominated by a layer-by-layer type growth [21, 49]

evolution of the SiO$_x$ film thickness with time, we have calculated the deposition rate d(thickness)/dt [21].

As it can be seen from Fig. 15.11, at the early stages of growth of SiO$_x$/PET, the deposition rate is increased up to 7.5 nm/s for t = 6 s, whereas at higher deposition times it follows an oscillating behavior. The oscillation of the deposition rate is characteristic of the island-type growth and defines the distinct growth stages. At the initial stages of the SiO$_x$ deposition onto PET, separate clusters are formed with a size that increases until the first 6 s with an incomplete coverage of the film. After t = 6 s, the reduction of the deposition rate indicates the coalescence stage where the clusters merge. At t > 12.5 s, the increase of the deposition rate indicates the formation of SiO$_x$ clusters that merge again. This behavior is repeated again as indicated by the oscillating of the deposition rate. On the other hand, the deposition of SiO$_x$ onto PEN follows a layer-by-layer growth mechanism. As it can be seen in Fig. 15.12, the cluster size increases in the first 2.5 s and then merge until t = 10 s. For t > 10 s the SiO$_x$ film thickness increases linearly with time and the deposition rate is almost stable, resulting in a homogeneous layer deposition [21].

Since PET and PEN are characterized by different values of roughness and peak-to-peak distance, the growth mechanism of SiO$_x$ is affected by the surface roughness of the substrate. The surface roughness (peak-to-peak distance) of PET and PEN are 1.18 nm (14 nm) and 1.75 nm (47 nm), respectively, whereas the surface roughness (peak-to-peak distance) of SiO$_x$ grown onto PET and PEN substrates, are 1.35 (44.76) and 0.95 (17.76) nm, respectively.

Finally, the above two examples clearly demonstrate the capabilities of real-time SE towards the understanding of the growth mechanisms of materials during their deposition onto polymeric substrates. This will contribute towards the optimization of the materials and processes for the manufacturing of FEDs with the desired operation,

lifetime and stability. SE has a significant potential for implementation to production lines for state-of-the-art applications as a quality control tool. Also, its robustness and high flexibility for adaptation gives numerous capabilities for the improvement of the produced structures (thin films, nanomaterials, etc.) as well as for the optimization of the production process in terms of cost and waste.

15.6 Optical Characterization of Inorganic and Polymer Transparent Electrodes

A wide variety of transparent conductive oxide (TCO) materials has been developed and investigated to be used as electrodes for organic electronic device applications. One of the most widely used materials is indium tin oxide (ITO). This is used as anode in OLEDs and in OPVs due to its excellent properties that include high electrical conductivity, piezoelectricity, easy fabrication, low cost, non-toxicity, and ultraviolet absorption [16, 50–52]. However, ITO is a brittle material, expensive and it has low abundance in comparison with other materials. For these reasons, a systematic research takes place during the last years in order to replace the ITO with other TCO materials that can be either inorganic or organic [53–55].

One of these is zinc oxide (ZnO), which is a wide direct band-gap semiconductor having the hexagonical crystal structure of wurtzite. It is a very promising material for application in FEDs since it combines several benefits such as high electrical conductivity, piezoelectricity, easy fabrication, low cost, non-toxicity, ultraviolet absorption behaviour and it is compatible with large scale applications [56]. In addition, dopants of ZnO, such as aluminum zinc oxide (AZO) and gallium zinc oxide (GZO) have been also studied during the last years, which have shown improved optical and electrical properties [16, 52]. Several efforts have been performed in order to comprehend the growth mechanisms of ZnO thin films, their functionality and combination with organic–inorganic materials as well as the effect of the deposition parameters in their optical, structural and electronic properties in order to reveal the full potential of FEDs, such as flexible OLEDs and OPVs. Figure 15.13 shows the unit cell of ZnO [52, 57].

In-situ and real-time SE has been applied to investigate the optical properties of ZnO thin films during their deposition onto rigid and polymer substrates. Figure 15.14 shows the evolution of the imaginary part of the measured pseudo-dielectric function $\langle \varepsilon_2(\omega) \rangle$ with the deposition time of the ZnO film onto a rigid (c-Si) and a flexible substrate (PET). The ZnO film has been deposited by dc magnetron sputtering onto the c-Si and PET substrates on a ceramic ZnO target (with purity of 99.995 %). The depositions were performed at room temperature (RT) and in argon atmosphere (total pressure 13.6 mTorr) by setting the pulse frequency at 100 kHz. The final thickness of the ZnO thin film after the end of the deposition process was 70 nm. The in-situ and real-time measurements were performed using an ultra-fast multi-wavelength phase modulated SE unit, adapted onto the UHV chamber at an angle of 70°. This system
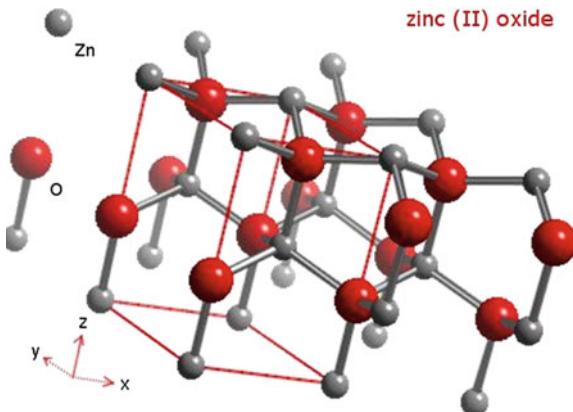
Fig. 15.13 Unit cell of Zinc Oxide (ZnO)

is equipped with a 32-fiber-optic array detector for simultaneous measurements at 32 different wavelengths (MWL mode), in the energy range 3–6.5 eV. The sampling time of MWL measurements (600 ms), is suitable for the real-time monitoring of the growth mechanisms and optical properties of the deposited ZnO thin films.

In the case of the flexible polymer substrate, in order to avoid the warping due to the heat load during the film deposition and to assure a flat surface suitable for in-situ and real-time measurements, these were fixed in a special substrate holder. The black squares (■) in Fig. 15.14 represent the $\langle \varepsilon_2 t(\omega) \rangle$ of the bare substrate (c-Si and PET), whereas the hollow circles (○) represent the $\langle \varepsilon_2(\omega) \rangle$ of the ZnO/c-Si and ZnO/PET. For the analysis of the measured $\langle \varepsilon(\omega) \rangle$, the TL model has been employed in combination to a four-phase theoretical model, consisting of PET/bottom ZnO layer/surface ZnO layer/air. SE is very sensitive to the surface quality, encouraging the use of the four phase model. The surface layer describes the effect of the surface
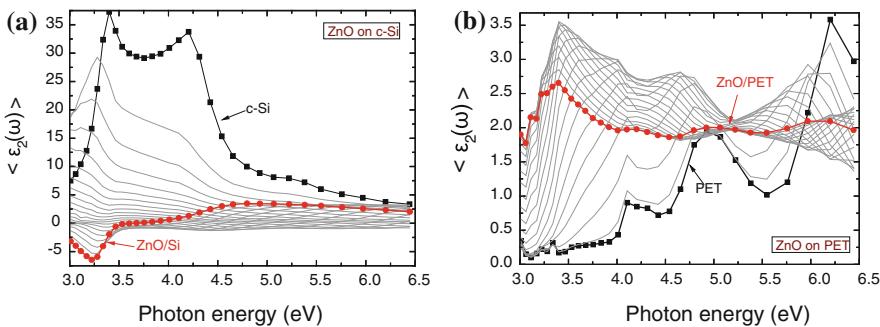
Fig. 15.14 Real-time measured $\langle \varepsilon_2(\omega) \rangle$ of ZnO film grown onto: **a** c-Si and **b** PET flexible substrate
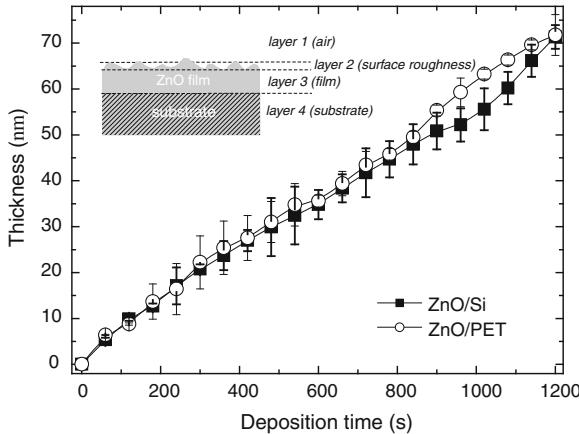
Fig. 15.15 Evolution of ZnO film thickness with the deposition time, as determined by SE [52, 57]

roughness, consisting of a mixture of ZnO and voids, and its analysis has been performed using the Bruggeman Effective Medium Approximation (BEMA).

Figure 15.15 shows the evolution of the growing ZnO film thickness with the deposition time onto the different substrates (c-Si, PET) as determined by the analysis of $\langle \varepsilon(\omega) \rangle$ using the above mentioned model. The diagrams are quite similar indicating similar, in general, growth characteristics.

Focusing on the early growth stages (first 120 s of deposition) as shown in more detail in Fig. 15.16, it is clear that the chemical structure and surface properties of the substrate affects the growth mechanism of ZnO film. The ZnO film growth can be qualitatively separated into three stages (I, II and III), according to the growth mechanisms taking place [52]. During the early stages of growth (Stage I), in the case of ZnO/PET, there is a significant increase of the apparent thickness at t = 15 s for ZnO/PET, attributed to polymer surface modification by the arrived ZnO and (neutral and charged) Ar particles on the PET surface and the formation of an interface layer (PET + ZnO). This modification includes the changing of the surface chemical bonding and the formation of clusters consisting of the arriving ZnO particles in the PET surface. This has been also justified by High Resolution Transmission Electron Microscopy (HRTEM) measurements that revealed a composite nanocrystalline overlayer of ∼10 nm on top of PET surface, including an interface layer of 1–2 nm [57]. As it can be seen in Fig. 15.16, at t = 17 s the ZnO film thickness is ∼3 nm (ZnO/PET), and ∼1 nm (ZnO/Si), whereas the deposition rates are ∼0.24 and ∼0.06 nm/s, respectively [52].

During Stage II, in the case of ZnO/Si, it is shown from Fig. 15.16 that the ZnO film is deposited directly to the aforedeposited layer, favoring the film growth (deposition rate of 0.11 nm/s). In the case of ZnO/PET, the low deposition rate (0.01 nm/s) indicates the early stages of ZnO deposition to the modified inter-face layer. Finally, during stage III the ZnO film growth is dominated by a layer-by-layer mechanism
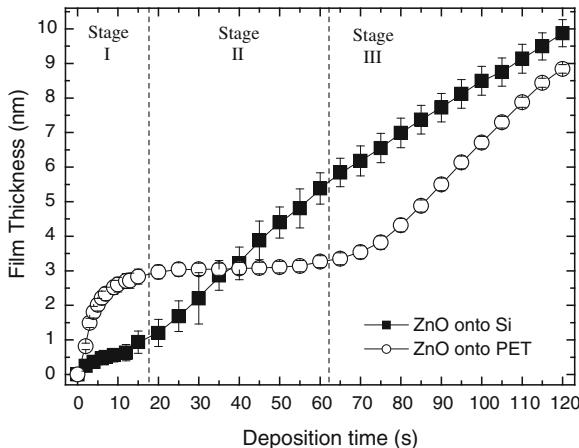
Fig. 15.16 Evolution of ZnO film thickness during the first 120 s of deposition

(deposition rates: 0.11, and 0.07 nm/s for PET, and Si, respectively), lasting until the end of deposition (1200 s), as it is shown in Fig. 15.15 [16, 52, 57].

Another material that is widely used as an anode buffer layer in FEDs is poly(3,4-ethylenedioxythiophene) poly(styrenesulfonate) (PEDOT/PSS). This is a transparent conductive polymeric material that is expected to replace the inorganic, brittle and expensive ITO and other TCO materials. This polymer is characterized by high conductivity and it is used as hole-injecting material in applications such as sensors, antistatic coatings, solar cells, etc. [16, 58]. PEDOT:PSS consists of a conducting part PEDOT, which is a low molecular weight conjugated polymer, insoluble and thus difficult to process and an insulating polymer (PSS), which is a high molecular weight polymer that gives the desirable flexibility. PSS also increases the solubility in water, making the whole system easy to process. The oligomer PEDOT segments are electrostatically attached on the PSS polymer chains [16, 59, 60] (Fig. 15.17).

The optical properties of the PEDOT:PSS films have been measured by SE in the Vis–fUV spectral region. The measured $\langle \varepsilon(\omega) \rangle$ is shown in Fig. 15.18. In order to extract quantitative information from the measured $\langle \varepsilon(\omega) \rangle$, this has been analyzed by the use of three phase geometrical model, which consists of air (ambient), the PEDOT:PSS layer (with thickness d) on top of the PET (bulk) substrate (air/PEDOT:PSS/PET substrate). For the ambient medium it has been used air in which $\varepsilon_1(\omega) = 1$ and $\varepsilon_2(\omega) = 0$ for all energy values ω.

The optical properties and the thickness of the PEDOT:PSS layer have been modeled by the use of the Tauc–Lorentz (TL) dispersion model. As it has been discussed in detail elsewhere in this model the imaginary part $\langle \varepsilon_2(\omega) \rangle$ of the dielectric function is obtained multiplying the equation of the Lorentz oscillator by the equation of the Tauc joint density of states, and the real part $\varepsilon_1(\omega)$ is determined by the imaginary part $\varepsilon_2(\omega)$ by the Kramers–Kronig integration. For the parameterization of the optical properties and thickness of the PEDOT:PSS films, we took into consideration
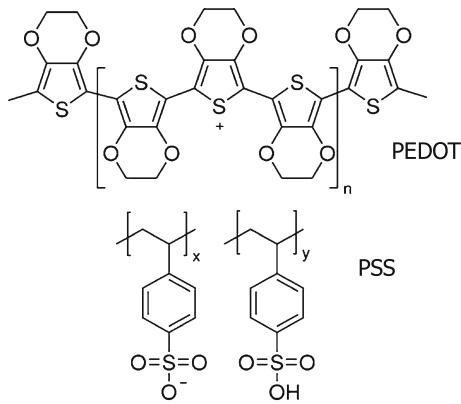
Fig. 15.17 Structure of the PEDOT:PSS
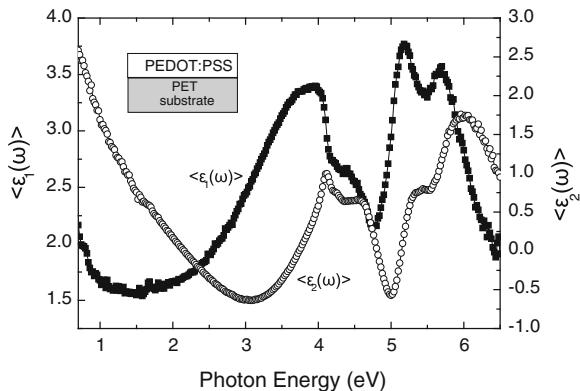
Fig. 15.18 Measured pseudodielectric function $\langle \varepsilon(\omega) \rangle$ of a representative PEDOT:PSS film that is deposited onto a PET flexible substrate by spin coating

only the characteristic optical absorptions of the $\pi - \pi^*$ transitions of the benzene rings of PSS that appear at 5.3 and 6.33 eV respectively. This modeling procedure provides in detail the thickness values of the PEDOT:PSS films. Figure 15.19 shows the determined bulk dielectric function $\varepsilon(\omega)$ of the PEDOT:PSS film in the Vis–fUV spectral region. This shows three absorptions in energies 0.47, 5.37 and 6.38 eV. The absorbance peaks at high energies can be attributed to the $\pi - \pi^*$ transitions of the benzene rings of PSS and the absorption in low energy attributed to the PEDOT [16].

Figure 15.20 shows the determined bulk dielectric function $\varepsilon(\omega)$ of the PEDOT: PSS films with various PSS weight ratios. The results show that the increase of PSS ratio in PEDOT:PSS solution leads to the reduction of ω_1 optical absorption attributed to PEDOT (decrease of oscillator's amplitude) and to a slight increase of ω_2 and ω_3 absorptions attributed to the $\pi - \pi^*$ transitions of the benzene rings of PSS [16].
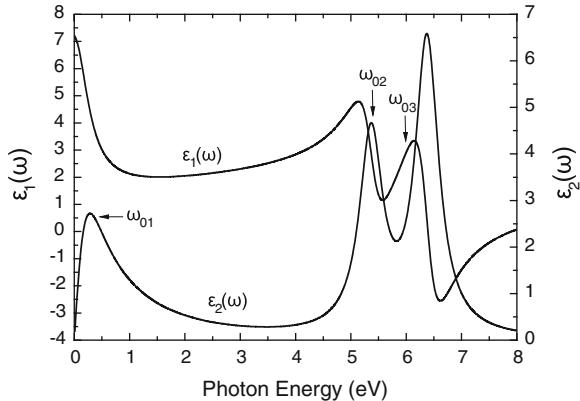
Fig. 15.19 Bulk dielectric function $\varepsilon(\omega)$ of PEDOT:PSS determined by the analysis of the $\langle\varepsilon(\omega)\rangle$, in which the individual electronic transitions that are attributed to the PEDOT and the PSS parts are distinguished [16]
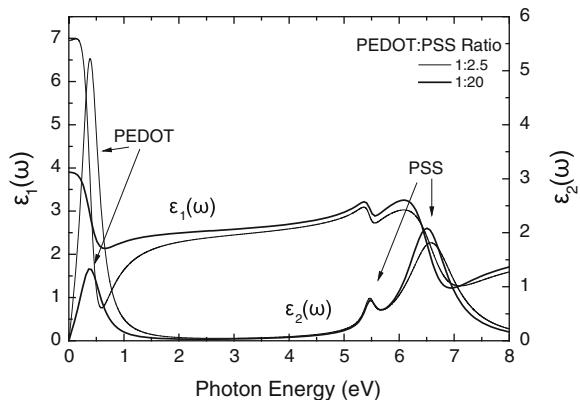
Fig. 15.20 The calculated bulk dielectric function for different PEDOT:PSS weight ratios [16]

Finally, the above two examples clearly demonstrate the capabilities of real-time SE towards the understanding of the growth mechanisms of materials during their deposition onto polymeric substrates. This nanometrology method has an enormous potential for implementation to production lines for state-of-the-art applications as a quality control tool. Also, its robustness and high flexibility for adaptation gives numerous capabilities for the improvement of the produced structures (thin films, nanomaterials, etc.) as well as for the optimization of the production process in terms of cost and waste.

15.7 Incorporation of SE to the Process Line for FEDs Fabrication

One of the major advantages of flexible organic electronic devices is their potential for fabrication flexible polymeric substrates by r2r large-scale production processes. This will offer a high throughput production combined with low cost of the final device for consumer use. Organic materials are soluble and/or solution-processable and they can deposited onto flexible polymeric substrates, which are in the form of web rolls, by solution-based methods such as printing processes, as well as by vacuum process, such as sputtering and e-beam evaporation [15].

During a large-scale process, the deposition of the individual organic (semiconductors, conductors) and functional (e.g. barrier, TCOs, electrodes, etc.) layers is realized on the moving web rolls. In this way, these processes are characterized by high volume and lower cost in comparison with batch processes, as in the case of the conventional microelectronics industry. In these r2r processes, the web rolls are run with speeds of hundreds of m/min with webs several meters wide, and are used to deposit and cure many different materials simultaneously [15].

Figure 15.21 shows such a future representative r2r process for the production of flexible organic electronic devices in which all production steps are integrated into a single production line. SE can be incorporated as a quality control tool in some or all of these production steps to ensure the quality of the deposited nano-layers and to contribute to the production of single and multi-layered stacks of organic, inorganic and hybrid layers of high quality.
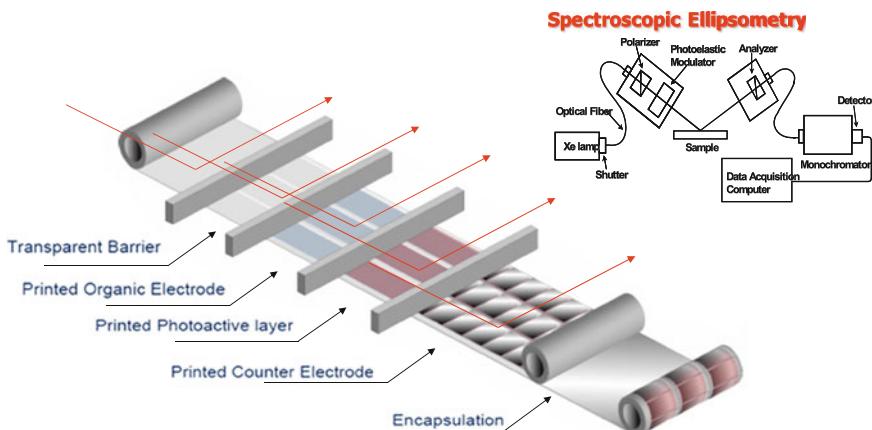
Fig. 15.21 Integration of the real-time control by SE units in the individual production steps for flexible organic electronic devices, for the quality control of the deposited nano-layers

15.8 Summary and Outlook

SE is a powerful tool for optical metrology of inorganic, hybrid and organic material for state-of-the-art applications, such as for flexible organic electronic devices. The optical constants, electronic structure, bonding vibrations, optical anisotropy can be measured and modeled accurately, with nm precision. Important parameters such as the energy band gap in organic and inorganic materials that can be used as electrodes, well as the macromolecular chain orientation in flexible polymer films can be obtained by the analysis of the measured $\langle \varepsilon(\omega) \rangle$ spectra in a wide spectral region from the IR to the Vis–fUV.

As the field of SE continues to grow with remarkable progress, more applications will be found for novel materials and film structures. The advances in the materials for organic electronics reveal new and complex structures that are essential to fulfill the requirements for efficiency, performance and stability of the whole device. This complexity requires more sophisticated modeling approaches of the optical measurements, to lead to useful information for the materials structure and optical response, as well as for the structure-property relationships.

Finally, an important aspect of the applicability of SE is its implementation to several steps of industrial processes, in order to serve as a quality control tool and to provide the ability to accurately and reproducibly measure their properties and performance at the nanoscale. Its capabilities for in-situ and real-time measurement and modeling, with measurement times in the order of ms, are an important factor that will enable this implementation.

References

1. S. Logothetidis, Polymeric substrates and encapsulation for flexible electronics: Bonding structure, surface modification and functional nanolayer growth. Rev. Adv. Mater. Sci. **10**(5), 387–397 (2005)
2. W.F. Smith, Organic electronics: self assembly is ready to roll. Nat. Nanotechnol. **2**, 77–78 (2007)
3. D.J. Gundlach, Organic electronics: Low power, high impact. Nat. Mater. **6**, 173–174 (2007)
4. A.J. Heeger, Semiconducting and metallic polymers: The fourth generation of polymeric materials. Rev. Mod. Phys. **73**, 681–700 (2001)
5. G. Dennler, C. Lungenschmied, L.C.H. Neugebauer, N.S. Sariciftci, A. Labouret, Flexible, conjugated polymer-fullerene-based bulk-heterojunction solar cells: basics, encapsulation, and integration. J. Mater. Res. **20**, 3224–3233 (2005)
6. H. Nakada, The status of development of Organic Light Emitting Diodes/organic thin-film transistors. J. Photopolym. Sci. Technol. **20**(1), 35–38 (2007)
7. H. Sirringhaus, N. Tessler, R.H. Friend, Integrated optoelectronic devices based on conjugated polymers. Science **280**(5370), 1741–1744 (1998)
8. D. Georgiou et al., Optical properties of hybrid polymers as barrier materials. Appl. Surf. Sci. **255**, 8023–8029 (2009)

9. M. Yanaka, B.M. Henry, A.P. Roberts et al., How cracks in SiOx-coated polyester films affect gas permeation. Thin Solid Films **397**, 176–185 (2001)

10. P. Roberts et al., Gas permeation in silicon-oxide/polymer (SiOx/PET) barrier films: role of the oxide lattice, nano-defects and macro-defects. J. Membr. Sci. **208**(1–2), 75–88 (2002)

11. R. Houbertz et al., Inorganic-organic hybrid materials for application in optical devices. Thin Solid Films, **442**(1–2), 194–200 (2003)

12. S. Amberg-Schwab, U. Weber, A. Burger, S. Nique, R. Xalter, Development of passive and active barrier coatings on the basis of inorganic-organic polymers. Monatshefte Für Chemie **137**, 657–666 (2006)

13. A. Laskarakis, S. Logothetidis, On the optical anisotropy of poly(ethylene terephthalate) and poly(ethylene naphthalate polymeric films by spectroscopic ellipsometry from visible-far ultra-violet to infrared spectral regions. J. Appl. Phys. **99**, 066101-1–066101-3 (2006)

14. A. Laskarakis, S. Logothetidis, Study of the electronic and vibrational properties of poly(ethylene terephthalate) and poly(ethylene naphthalate) films. J. Appl. Phys. **101**, 03503-1–053503-9 (2007)

15. S. Logothetidis, Towards the optimization of materials and processes for flexible organic electronic devices. Rev. Adv. Mater. Sci. **10**, 387 (2005)

16. S. Logothetidis, A. Laskarakis, Organic against inorganic electrodes grown onto polymer substrates for flexible organic electronics applications. Thin Solid Films **518**, 1245–1249 (2008)

17. S. Logothetidis, *Thin Films Handbook: Processing Characterization and Properties* (Academic Press, New York, 2001)

18. A. Laskarakis, S. Logothetidis, M. Gioti, Study of the bonding structure of carbon nitride films by IR spectroscopic ellipsometry. Phys. Rev. B **64**, 125419-1–125419-15 (2001)

19. R.M.A. Azzam, N.M. Bashara, *Ellipsometry and Polarized Light* (North Holland Publishing, Amsterdam, 1977)

20. H. Tompkins, E. Irene, *Handbook of Ellipsometry* (*Materials Science and Process Technology*) (William Andrew Publishing/Noyes, Norwirtch, 2005)

21. D. Georgiou, S. Logothetidis, C. Koidis, A. Laskarakis, In-situ and real-time monitoring of high barrier layers growth onto polymeric substrates. Phys. Status Solidi (c) **5**, 1–4 (2008)

22. N. Koch, Organic electronic devices and their functional interfaces. ChemPhysChem **8**, 438–1455 (2007)

23. S.A.C. Gould, D.A. Schiraldi, M.L. Occelli, Analysis of poly(ethylene terephthalate) (PET) films by atomic force microscopy. J. Appl. Polym. Sci. **65**, 1237–1243 (1997)

24. C.W. Tang, Two-layer organic photovoltaic cell. Appl. Phys. Lett. **48**, 183 (1986)

25. J. Zhao et al., 20,000 PERL silicon cells for the '1996 world solar challenge' solar car race. Prog. Photovoltaics, **5**(4), 269–276 (1997)

26. T. Erb et al., Absorption and crystallinity of poly(3-hexylthiophene)/fullerene blends in dependence on annealing temperature. Thin Solid Films **511**, 483–485 (2006)

27. N.S. Sariciftci, Plastic photovoltaic devices. Mater. Today, **7**(9), 36–40 (2004)

28. A.C. Mayer, S.R. Scully, B.E. Hardin, M.W. Rowell, M.D. McGehee, Polymer-based solar cells. Mater. Today, **10**(11), 28–33 (2007)

29. F. Zhang et al., Polymer solar cells based on a low-bandgap fluorene copolymer and a fullerene derivative with photocurrent extended to 850 nm. Adv. Funct. Mater. **15**(5), 745–750 (2005)

30. K.M. Coakley, M.D. McGehee, Conjugated polymer photovoltaic cells. Chem. Mater. **16**(23), 4533–4542 (2004)

31. S.A. McDonald et al., Solution-processed PbS quantum dot infrared photodetectors and photovoltaics. Nat. Mater. **4**(2), 138–U14 (2005)

32. V. Coropceanu et al., Charge transport in organic semiconductors. Chem. Rev. **107**, 926–952 (2007)

33. J.Y. Kim et al, Efficient tandem polymer solar cells fabricated by all-solution processing. Science **317**(5835), 222–225 (2007)

34. N. Koch et al., Evidence for temperature-dependent electron band dispersion in pentacene. Phys. Rev. Lett. **96**(15), 156803 (2006)
35. G.N. Gavrila et al., Energy band dispersion in well ordered N, N'-dimethyl-3,4,9,10-perylenetetracarboxylic dlimide films. Appl. Phys. Lett. **85**(20), 4657–4659 (2004)
36. L.V. Keldysh, D.A. Kirzhnits, A.A. Maradudin, The dielectric function of condensed systems (North-Holland Publishing, Amsterdam, 1989)
37. J.E. Jellison, F.A. Modine, Parameterization of the optical functions of amorphous materials in the interband region. Appl. Phys. Lett. **69**, 371–373 (1996)
38. K. Nomura et al., Room-temperature fabrication of transparent flexible thin-film transistors using amorphous oxide semiconductors. Nature **432**, 488 (2004)
39. R. Forest, The path to ubiquitous and low cost organic electronic appliances on plastic. Nature **428**, 911 (2004)
40. A. Laskarakis, S. Kassavetis, C. Gravalidis, S. Logothetidis, In situ and real-time optical investigation of nitrogen plasma treatment of polycarbonate. Nucl. Instrum. Methods Phys. Res. B, **268**, 460–465 (2010)
41. A.E. Tonelli, PET versus PEN: What difference can a ring make? Polymer **43**, 637–642 (2002)
42. W. Fix, A. Ullmann, J. Ficker, W. Clemens, Fast polymer integrated circuits. Appl. Phys. Lett. **81**, 1735 (2002)
43. A. Laskarakis, S. Logothetidis, Investigation of the optical anisotropy of PET and PEN films by VIS-FUV to IR spectroscopic ellipsometry. Appl. Surf. Sci. **253**, 52–56 2006
44. A. Laskarakis, M. Gioti, E. Pavlopoulou, A complementary spectroscopic ellipsometry study of PET membranes from IR to Vis-FUV. Macromol. Symp. **205**, 95–104 (2004)
45. C. Charton et al., Development of high barrier films on flexible polymer substrates. Thin Solid Films, **502**(1–2), 99–103 (2006)
46. A. Laskarakis, S. Logothetidis, D. Georgiou, S. Amberg-Schwab, U. Weber, On the influence of silicon oxide nanoparticles on the optical and surface properties of hybrid (inorganic-organic) barrier materials. Thin Solid Films, **517**, 6275–6279 (2009)
47. S. Amberg-Schwab, H. Katschorek, U. Weber, M. Hoffmann, A. Burger, Barrier properties OF inorganic-organic polymers: influence of starting compounds, curing conditions and storage-scaling-up to industrial application. J. Sol-Gel Sci. Technol. **19**(1–3), 125–129 (2000)
48. K. Haas, S. Amberg-Schwab, K. Rose, G. Schottner, Functionalized coatings based on inorganic-organic polymers (ORMOCER (R) s) and their combination with vapor deposited inorganic thin films. Surf. Coat. Technol. **111**(1), 72–79 (1999)
49. A. Laskarakis, D. Georgiou, S. Logothetidis, Real-time optical modelling and investigation of inorganic nano-layer growth onto flexible polymeric substrates. Mater. Sci. Eng. B **166**, 7–13 (2009)
50. A.N. Banerjee et al., Low-temperature deposition of ZnO thin films on PET and glass substrates by DC-sputtering technique. Thin Solid Films **496**(1), 112–116 (2006)
51. M. Suchea, S. Christoulakis, K. Moschovis, N. Katsarakis, G. Kiriakidis, ZnO transparent thin films for gas sensor applications. Thin Solid Films **515**(2), 551–554 (2006)
52. Ch. Koidis, S. Logothetidis, D. Georgiou, Growth, optical and nanostructural properties of magnetron sputtered ZnO thin films deposited on polymeric substrates. Phys. Status Solidi A, **205**, 1988–1992 (2008)
53. A.L. Dawar, A.K. Jain, C. Jagadish, H.L. Hartnagel, *Semiconducting Transparent Thin Films* (Institute of Physics Publishing, Bristol, Philadelphia, 1995)
54. H. Kim et al., Indium tin oxide thin films for organic light-emitting devices. Appl. Phys. Lett. **74**(23), 3444–3446 (1999)
55. H. Kim et al., Electrical, optical, and structural properties of indium-tin-oxide thin films for organic light-emitting devices. J. Appl. Phys. **86**(11), 6451–6461 (1999)
56. E. Fortunato, A. Goncalves, A. Marques, H. Aguas, L. Pereira, I. Ferreira, R. Martins Pimentel, Thin Solid Films **487**, 212 (2005)
57. Ch. Koidis, S. Logothetidis, A. Laskarakis, I. Tsiaoussis, N. Frangis, Thin film and interface properties during ZnO deposition onto high-barrier hybrid/PET flexible substrates. Micron **40**, 130–134 (2009)

58. M. Garganourakis et al., Deposition and characterization of PEDOT/ZnO layers onto PET substrates. Thin Solid Films, **517**(23), 6409–6413 (2009)
59. L. Schmidt-Mende, J.L. MacManus-Driscoll, ZnO—nanostructures, defects, and devices. Mater. Today **10**(5), 40–48 (2007)
60. C. Gravalidis, A. Laskarakis, S. Logothetidis, Fine tuning of PEDOT electronic properties using solvents. Eur. Phys. J. Appl. Phys. **46**, 12505 (2009)

Chapter 16
Spectroscopic Ellipsometry of Nanoscale Materials for Semiconductor Device Applications

Alain C. Diebold, Florence J. Nelson and Vimal K. Kamineni

Abstract Several years ago, the semiconductor industry began to refer to integrated circuits as nanoelectronic devices [1]. Now, most realize that nanoelectronics is the most prevalent nanotechnology. The continued decrease in device feature size has challenged spectroscopic ellipsometry (SE) with nano-films, nanowires, and nano-dots. There are many examples of the measurement of thin dielectric films [2, 3], and now there are examples of crystalline semiconductor nanowires in the form of the Fin in the transistor known as a Fin-FET [4]. The semiconductor industry is also working on materials for "beyond CMOS" devices.

16.1 Introduction

Several years ago, the semiconductor industry began to refer to integrated circuits as nanoelectronic devices [1]. Now, most realize that nanoelectronics is the most prevalent nanotechnology. The continued decrease in device feature size has challenged spectroscopic ellipsometry (SE) with nano-films, nanowires, and nano-dots. There are many examples of the measurement of thin dielectric films [2, 3], and now there are examples of crystalline semiconductor nanowires in the form of the Fin in the transistor known as a Fin-FET [4]. The semiconductor industry is also working on materials for "beyond CMOS" devices. Graphene is the most prevalent example of this. Some materials may transition between CMOS extension and Beyond

A. C. Diebold (✉) · F. J. Nelson · V. K. Kamineni
College of Nanoscale Science and Engineering, University at Albany, Albany, NY, USA
e-mail: adiebold@albany.edu

F. J. Nelson
e-mail: florencenel@gmail.com

V. K. Kamineni
GLOBALFOUNDRIES-Technology Research Group, Albany, NY12203, USA
e-mail: vimal.kamineni@globalfoundries.com

CMOS. Some integrated circuits are fabricated on silicon-on-insulator wafers. This includes extremely thin silicon-on-insulator (ETSOI). These layers are in fact crystalline silicon quantum wells (c-SI QWs). Although some predict that ETSOI will be limited to 7 nm or greater thickness in manufacturing, research devices have used 2 nm thick top silicon layers. These samples provide an interesting study in the fundamental physics. In this chapter, we discuss how SE has been applied to both graphene and c-Si QWs. One key message of this chapter is that the optical properties of nanoscale crystalline materials are strongly influenced by nanoscale properties such as quantum confinement and electron-phonon interactions. The impact of the change in phonon dispersion with nanoscale dimensions has been mostly ignored in understanding the optical properties of nanoscale materials.

16.2 Graphene

16.2.1 Electronic and Optical Properties

Graphene is a sp^2 hybridized monolayer of carbon atoms. Its bonding configuration gives rise to a band structure that may be approximated as a linear relationship between energy and momentum at the K points, or corners, of the Brillouin Zone (BZ), the so-called "Dirac points" after the relativistic equation obeyed by multi-component spinors. This linear relationship mimics that of a photon's dispersion, giving the carriers in graphene an extremely high velocity when compared to that of typical semiconductors (i.e. Si) [5]. The optical properties of graphene, like its electronic properties, are also a result of its unusual bandstructure at different points of the BZ. Graphene's optical properties between the near IR to the UV are the result of transitions between the valence π band and the conduction π* band. Near the Dirac points, which have been of most interest due to the consequences for graphene's electronic applications, symmetric transitions from the valence to conduction band result in a near-constant absorbance of ∼2.3 % of incident light from the visible into the infrared (IR) region even though graphene is only one atomic layer thick. This has been shown theoretically as well as experimentally [6–8], and is known as the fine structure constant absorbance, derived from the fundamental constant describing the coupling between light and electrons via an electromagnetic interaction. Moving away from the Dirac points, however, allows one to observe the excitonic effects that result from graphene's two-dimensional (2D) structure [9]. An absorbance peak in the UV region, due to a van Hove singularity in graphene's density of states [10], occurs at the M point of the BZ and is strongly influenced by electron-hole (e-h) interactions due to a lack of screening. These e-h effects are observable in an ellipsometric measurement and will be discussed in the Current Areas of Interest section of this work. The K point and M point absorptions are shown in Fig. 16.1 through an extended zone scheme of Ref. [11].
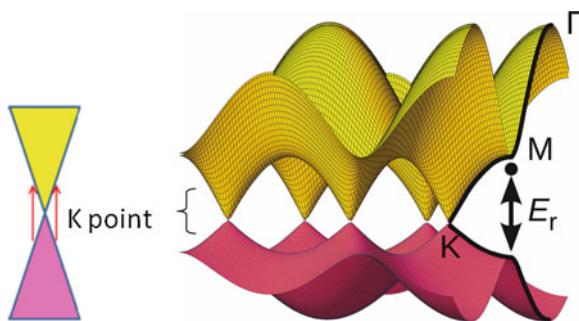
Fig. 16.1 Extended zone scheme of graphene showing the M point (*right*) in the π bands responsible for the exciton-modified UV peak, and the K point (*left*) showing symmetric excitation of valence (*pink*) to conduction (*yellow*) states responsible for the fine structure constant absorbance in the IR region. Adapted with permission from Ref. [11]. Copyright 2011 American Chemical Society
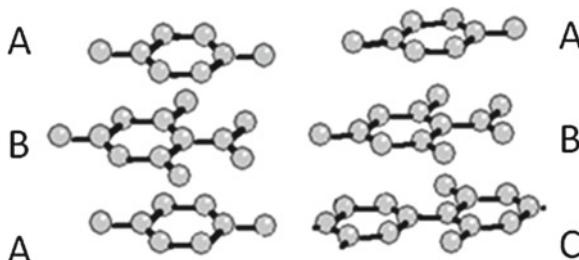
Fig. 16.2 Bernal (*left*) and rhombohedral (*right*) stacking of bulk graphite

16.2.2 Graphene Fabrication Methods and Resulting Questions for Ellipsometry

SE of thin graphitic structures has gained importance in recent years due to the 2004 discovery of graphene's unique electrical properties [5]. While graphene is a 2D material, its bulk structure of graphite is optically anisotropic and has been the subject of research and optical measurements since well before the recognition of graphene's unique band structure [12, 13]. The stacking structure of bulk graphite (Fig. 16.2) is primarily composed of a Bernal layer offset, termed "ABA," with the remaining stackings being rhombohedral , "ABC", and turbostratic (randomly misoriented layers) [14]. The stacking configuration alters the electronic band structure of the layers from the linear dispersion of one graphene monolayer in the low-energy limit of the BZ. However, turbostratic graphene has shown a return to this linear dispersion due to an electrical decoupling of the layers [15].

The early fabrication method of graphene samples consisted of applying scotch tape to bulk graphite and mechanically exfoliating several layers that could be stamped to an arbitrary substrate [5], the most notable of these being 90 or 300 nm

SiO$_2$/Si due to an interference effect that allowed one to locate few-layer-graphene (FLG) simply be optical inspection [16]. The presence of monolayer or FLG could then be verified using Raman spectroscopy from the ratio of specific phonon intensities and the number/shape of the Lorentzian oscillator components used in the data fitting process [17]. Such graphene flakes were on the order of tens of microns and were sufficient for electrical property measurements given the spatial resolution of lithographic patterning. However, standard ellipsometers possess spot sizes on the order of millimeters. An SE measurement therefore required the use of focusing probes that allowed the beam spot to fit within the area of one graphene flake. This was shown to be possible using Woollam ellipsometers [10, 18].

Due to the fact that integration of graphene into the semiconductor industry will require a large-area scalable growth process, several groups have begun to study Chemical Vapor Deposition (CVD) on metallic foils [19–21]. Hydrocarbon-based deposition on copper (Cu) foils has gained attention due to carbon's negligible solubility in the material, allowing for a surface-limiting process that produces primarily monolayer graphene which may then be transferred to secondary substrates [19]. The undulation of a foil when chucked to an SE vacuum stage inhibits the procurement of meaningful data. The ability to transfer large-area graphene to a second substrate allows for an SE measurement without the use of focusing probes.

Graphene grown by a CVD process is polycrystalline, meaning it consists of "grains" or "domains" of graphene with different orientations separated by one-dimensional grain boundaries [22]. The orientation and structure of these grain boundaries has been shown theoretically to have a profound impact on the transmission or reflection of carriers between two domains [23]. Two questions therefore arise for the area of SE concerned with such thin graphitic structures: (1) Does the polycrystalline nature of CVD graphene alter the dielectric function/complex refractive index (CRI) when compared to exfoliated or single crystal samples, and (2) what is the effect of graphene's unique 2D electronic structure on the complex refractive index? To address these questions, we first model the CRI values, n and k, of polycrystalline graphene as an isotropic layer and compare our results to previously reported values for exfoliated graphene. We then use the obtained results in a simulation study to obtain the effect of layer number on the validity of the isotropic model when compared to an anisotropic model. The differences between the two modeling approaches are gauged in terms of their magnitude when compared to that of the raw Psi (ψ) and Del (Δ) error bars.

16.2.3 Thin Graphite Films and SE of Exfoliated Graphene

One of the initial publications on the optical response of thin graphite films used a combination of reflectometry and ellipsometry to characterize the in-plane and out-of-plane CRI [24]. Few layer graphene (FLG) exfoliated films were found to exhibit birefringence when the data was analyzed with a Kramer's-Kronig (KK) consistent model on 300 nm SiO$_2$/Si. The KK relationship enforces a physically valid solution
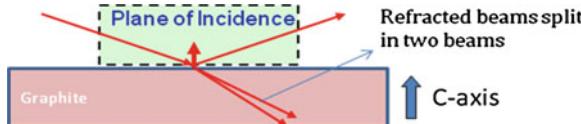
Fig. 16.3 Optical ray diagram showing birefringence of bulk graphite

that relates any value of *n* or *k* over the measured energy range (*P* is the Cauchy Principle Value):

$$n(E) - 1 = \frac{2}{\pi} P \int_{0}^{\infty} \frac{E'k(E')dE'}{E'^2 - E^2} \tag{16.1}$$

Bulk graphite is a uniaxial anisotropic material, and when the C-axis is normal to the surface the complex refractive indices are the same for the in-plane polarizations of light, x and y, and different for the out-of-plane, or z polarization (Fig. 16.3). In the simplest case of an isotropic material, reflection of light in the p (parallel) or s (perpendicular) polarization with respect to the plane of incidence may be described in terms of the 2×2 Jones Matrix, shown in (16.2). In the absence of cross-polarization, the r_{sp} and r_{ps} elements of the matrix are zero. This is the case for isotropic samples as well as anisotropic samples that are positioned in a particular geometry, such that the optic axis is oriented along an axis of symmetry.

$$\begin{pmatrix} E_p^{out} \\ E_s^{out} \end{pmatrix} = \begin{pmatrix} r_{pp} & r_{sp} \\ r_{ps} & r_{ss} \end{pmatrix} \begin{pmatrix} E_p^{in} \\ E_s^{in} \end{pmatrix} \tag{16.2}$$

For a typical ellipsometry measurement of graphene on substrate samples, the basal plane of carbon atoms is perpendicular to the optic axis, yielding the geometry that produces a Jone's matrix equivalent to that of an isotropic sample in terms of cross-polarization elements. The sensitivity of an ellipsometry measurement to anisotropy is dependent on the distance the light has to traverse the film, and therefore allow the two refracted beams to separate. In graphene, this path length is severely limited (the nominal thickness of a graphene monolayer is 3.35 Å).

One of the first publications describing SE characterization of FLG was published by the group of Novoselov and Geim [10]. FLG flakes were prepared by mechanical exfoliation of bulk graphite. The experimental setup included focusing probes capable of reducing the beam spot to a dimension that could be contained within the area of a flake. Analysis of the ψ and Δ spectra consisted of modeling the graphene layer as an anisotropic material based on the properties of bulk graphite, and yielded optical functions for *n* and *k* that presented the two main features: (1) ~2.3 % absorbance of light from the IR into the visible region, and (2) the exciton—modified absorption peak at ~4.6 eV. This work was followed in 2010 by a second SE study of exfoliated graphene which similarly incorporated the use of focusing probes for beam spot reduction, but used B-spline parameterization to fit for the CRI in order to ensure

a KK-consistent solution [18]. A more intense exciton-dominated absorption peak was observed in the latter work [18].

SE of large-area graphene does not require focusing probes. In certain instances the transfer process of graphene to a secondary substrate will yield sub-monolayer films. In this case, an effective medium approximation (EMA), to be discussed, must be incorporated to remove the effect of the ambient from the film's CRI.

16.2.4 SE of CVD Graphene

Samples prepared in a previously described method [25] were measured with a dual rotating compensator ellipsometer (RC2 - J. A.Woollam Co., Inc.). WVASE32 software allows the analysis of the raw ψ and Δ data through development of an optical dispersion model. Lorentz oscillators characterized by three parameters (energy E, amplitude, and FWHM) represent absorptions in the film at different energies for the imaginary portion of the dielectric function, $\varepsilon(E)$, with the real part of $\varepsilon(E)$ resulting from the KK relationship imposed on the regression fitting. Three Lorentz oscillators were implemented for graphene's dielectric function, corresponding to the exciton-dominated absorption peak, a Drude-like absorption at the IR end of the energy range, and a broad absorption for the region of constant optical conductivity. The resulting n and k values are shown in Fig. 16.4 for CVD graphene in comparison to those reported for exfoliated graphene on SiO_2/Si by two other groups [10, 18]. The absorbance of the CVD film was found from the relationship between n, k and film thickness for graphene in the visible region [26] and showed the fine structure constant over the visible-IR wavelengths, as well as qualitative agreement with the position of the exciton-dominated absorption peak predicted from theory [9]. Therefore SE of polycrystalline graphene shows the two main features predicted from theory and shown experimentally for exfoliated graphene, with some differences in the intensity of the van Hove peak. Differences between the data sets may be attributed to differences in the substrate on which the graphene sits (glass vs. SiO_2), the approaches to the modeling, as well as the fabrication method of the graphene itself (CVD vs. exfoliated flakes).

16.2.5 Aberration-Corrected STEM Imaging and Diffraction Pattern Analysis for Grain Size Estimation of CVD Graphene

Transmission Electron Microscopy (TEM) is an important characterization method for the observation of graphene lattice structure/defects, and in the case of CVD graphene, grain size measurement. The results of Fig. 16.4 were obtained from graphene grown using CVD process conditions which produce grain sizes of the
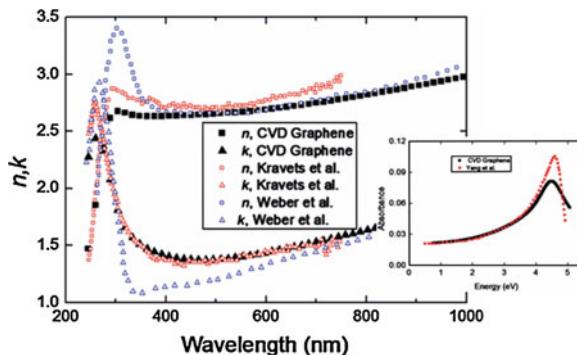
Fig. 16.4 Comparison of CVD graphene CRI to that reported for exfoliated graphene [10, 18]. The absorbance of CVD graphene was calculated from the experimental CRI values (*black*) and compared to theory (*red*) [9] in the *lower right*. Reprinted with permission from Ref. [25]. Copyright 2010, American Institute of Physics
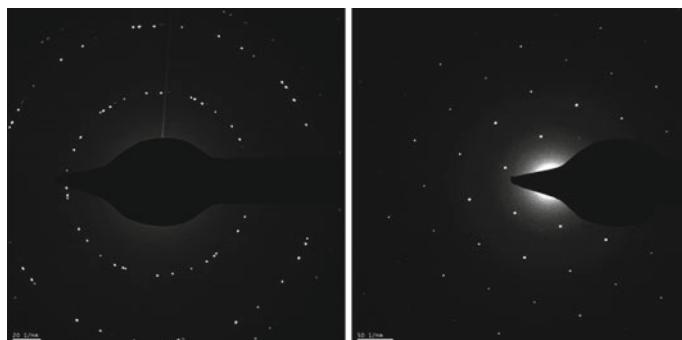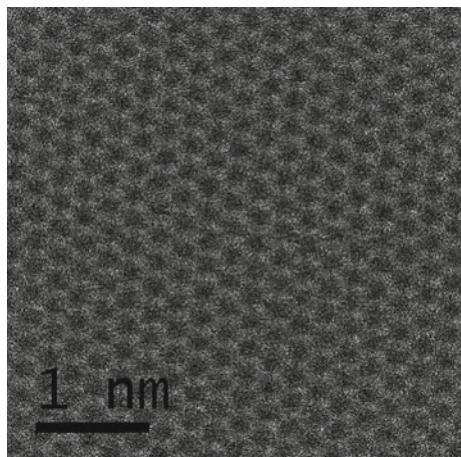
Fig. 16.5 DP's showing the difference between smaller (*left*) and larger (*right*) grain sizes in CVD graphene through the number of rotated hexagons present in the first order polycrystalline ring

order of hundreds of nanometers from Dark Field Transmission Electron Microscopy (DF-TEM), a suitable technique due to the polycrystalline nature of the films [27, 28]. The diffraction pattern (DP) of a smaller grain film will show more reflections than that of a film with larger grains, due to the multiple crystallographic orientations present within the same-size area. As shown in Fig. 16.5, the CVD graphene DP on the left has more hexagons rotated with respect to each other in the first polycrystalline ring, indicating multiple domain orientations within the measured field. The DP on the right shows a single-crystal structure, suggesting only one domain orientation, indicating larger grain size. Such imaging/diffraction can be accomplished at fairly low magnifications and is used here to determine the effect of grain size on the CRI of CVD graphene.

While grain size may be determined from imaging at low magnifications, observation of grain boundaries or lattice defects requires atomic resolution and may

Fig. 16.6 Aberration-corrected Nion UltraSTEM images of CVD graphene at 60 kV from ORNL, showing atomic resolution. The authors acknowledge Juan-Carlos Idrobo (ORNL) for his assistance in the imaging

be accomplished through aberration-corrected STEM imaging (Fig. 16.6). CVD graphene was imaged at Oak Ridge National Laboratory (ORNL) at 60 kV using a Nion UltraSTEM. A spatial resolution of 1.06 Å was obtained, allowing the direct observation of atomic positions in the graphene lattice. Such imaging allows for study of the grain boundary structure produced from different CVD growth conditions, which will affect the electronic properties of graphene devices.

16.2.6 Linear Effective Medium Approximation for Incomplete Layers

Transfer of monolayer graphene films from a Cu foil to a substrate sometimes results in graphene "patches" separated by regions of bare substrate, or sub-monolayer films. In such cases n and k will show a decrease in magnitude due to the void space in between these patches. A linear EMA has previously been reported to show a sensitivity to the percentage of graphene surface coverage [25, 29] and provides a simple tool with which to take the CRI of void space into consideration.

EMA's such as the Bruggeman and Maxwell-Garrett approximations differ in the assumptions made concerning host material as well as component geometry. For the laminar geometry of a graphene layer and the accompanying void space, the more simplistic linear EMA assumes a CRI that is a combination of the two components weighted by their volume fraction [30]. With a linear EMA, one may fit for the percentage of void space in an incomplete monolayer film, and subsequently model the optical functions for only the graphene component. Previous work has shown the proper choice of EMA to be, in some cases, material dependent. In the case of amorphous carbon a linear approximation showed better agreement with complementary EELS when comparing the sp^2 and sp^3 percentages of the films [31].

16.2.7 Graphene Wafer Mapping

SE has previously been used in conjunction with Raman spectroscopy in order to find the thickness of the graphene prior to the SE analysis. By fixing the thickness to the nominal value of 3.35 Å, *n* and *k* may be extracted from the two parameters ψ and Δ [10, 25, 29]. However the reverse process, that is using the CRI to determine thickness, should allow for the mapping of graphene on any substrate. This capability and verification of CRI values was shown through a 625 point mapping scan of a 300 mm SiO$_2$/Si wafer that had been stamped with CVD graphene in a particular layout [29]. Figure 16.7 shows the ψ map (bottom right) measured on the 300 mm wafer and the locations to which the graphene was transferred (top right) using a thermal release tape method [25]. Use of the CVD graphene CRI values from the glass substrate analysis (Fig. 16.4) resulted in the detection of a 3.3 Å film in the areas to which CVD graphene was transferred [29]. As graphene becomes a stronger candidate for device structures its incorporation into inline thickness metrology will become more important for films produced by growth processes capable of yielding large area layers.

SE mapping of graphene on a smaller scale, such as would be required for exfoliated flakes hundreds of square microns in area, is also possible. Also shown in Fig. 16.7 is a color-coded map of Δ at a wavelength of 746 nm for an area of 90 nm SiO$_2$/Si exfoliated with FLG by Graphene Industries. The optical image at the top left shows the bilayer flake, which is pointed to in the outline of the Δ map. The attractiveness of bilayer graphene lies in the fact that, unlike the monolayer system, it is possible to induce a bandgap in the bilayer structure through electrostatic doping which is necessary to produce an off-state in device structures [32].

16.2.8 Simulation Study of the Effect of Anisotropy

As the number of layers in FLG increases, the electronic structure eventually converges to that of graphite, and our previous question addressing the anisotropic demands of the modeling becomes an issue. Using the optical functions obtained for the CVD graphene on glass, a simulated graphene/SiO$_2$/Si stack generated with WVASE32 software is used to show the difference between ψ and Δ values calculated with an isotropic versus anisotropic model. The study incorporates the SiO$_2$/Si substrate for its omnipresence in device electronics. The generated data is then compared to the magnitude of the error bars in the experimental data of ψ and Δ that was measured for CVD graphene on SiO$_2$/Si. The model stacks consist of the graphene sitting on top of the standard 300 nm SiO$_2$/Si substrate, shown in Fig. 16.8. Optical functions provided in the WVASE32 software were used for the substrate layers.

The CVD *n* and *k* values in Fig. 16.4 were used for the isotropic layer model, as well as the x−y component of the refractive index for the anisotropic layer model. For the latter, the z response values were chosen based on previously reported work on graphite [33] and graphene [10]. For a c-cut uniaxial sample of highly oriented

Micron-scale mapping of exfoliated graphene   Millimeter-scale mapping of CVD graphene
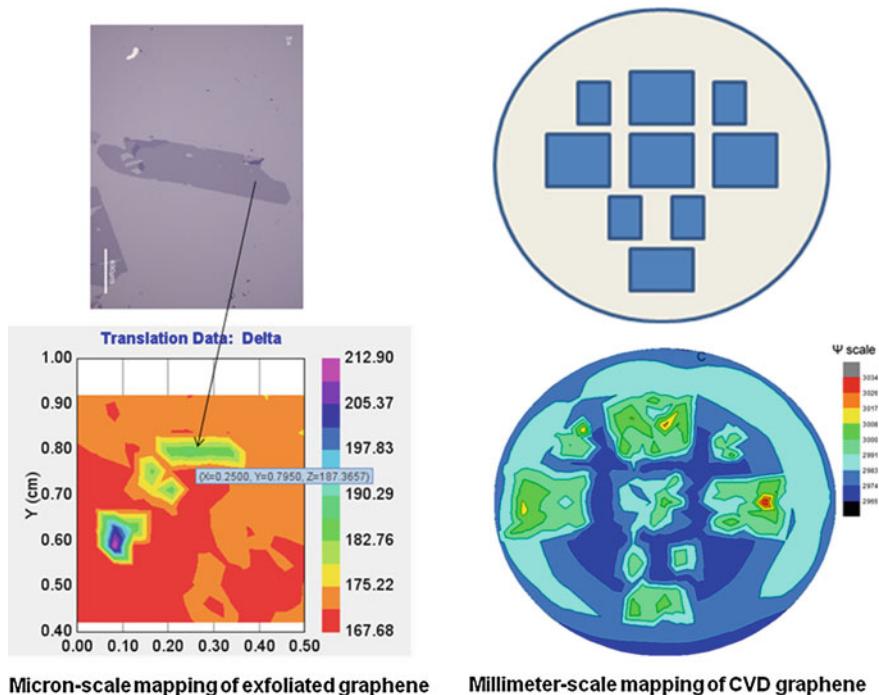
Fig. 16.7 Color-coded Δ (*bottom left*) and ψ (*bottom right*) maps of exfoliated and CVD graphene, respectively. *Top* optical image on the *left* shows the bilayer outlined in the Δ map below, 400 μm scale bar. *Top right* image shows the locations of transferred CVD graphene on the 300 mm SiO$_2$/Si wafer
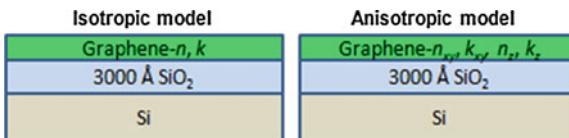
Fig. 16.8 Isotropic and anisotropic stacks for simulation of data to compare modeling of graphene layer(s)

pyrolytic graphite (HOPG), the c-axis is essentially normal to the surface and parallel to the plane of incidence during the SE measurement. This geometry prevents the sensitivity of the measurement to cross polarization. Jellison et al. determined both the ordinary (x−y) and extraordinary (z) responses of HOPG by combining two measurements, one in which the c-axis was perpendicular to the sample surface, and one in which the c-axis was parallel to the sample surface [33]. The real part of the z-response in the visible region (n_z) was determined to be within the range of 1.5. The work of Kravets et al. also yielded an n_z less than 2, with an assumed k_z value of zero [10]. Therefore, the generated data discussed below is simulated using $n = 1.5$, $k = 0$ when the graphene is modeled as an anisotropic layer.
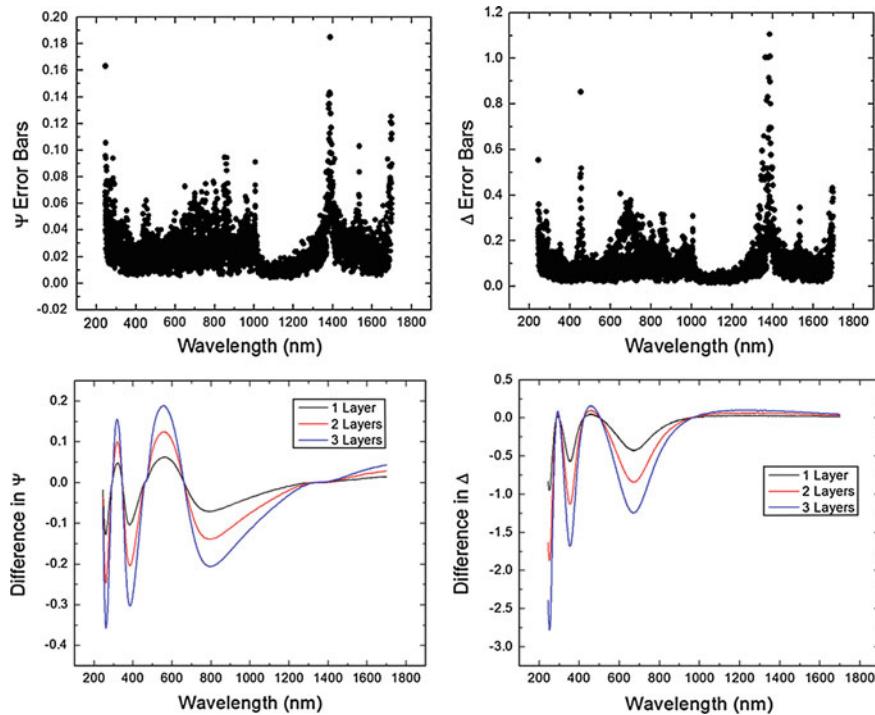
Fig. 16.9 (*Top*) Error bars for ψ and Δ (*top*) measured for monolayer CVD graphene. (*Bottom*) Difference in ψ and Δ values generated from an isotropic versus anisotropic model of graphene based on layer number. The error bars are comparable in magnitude to the differences in the modeling approach for one layer; the assumption of an isotropic film starts to break down as layer number increases

Figure 16.9 shows the differences between the isotropically and anisotropically modeled data for one to three layers of graphene, in comparison to the experimental error bar size for the measurement of monolayer CVD graphene on 300 nm SiO_2/Si. The difference in isotropic versus anisotropic modeling becomes larger as layer number increases, as one would expect while moving towards bulk graphite. The validity of isotropic modeling would therefore not hold when the difference between the models becomes significantly higher than the error bars of the raw ψ and Δ data. This begins to become an issue at approximately three layers.

We may also observe how the value of n_z directly affects the difference between the data generated from an isotropic versus anisotropic model. The larger difference between the x−y and z refractive indices is shown in the larger difference between ψ and Δ values in Fig. 16.10 for a bilayer thickness.
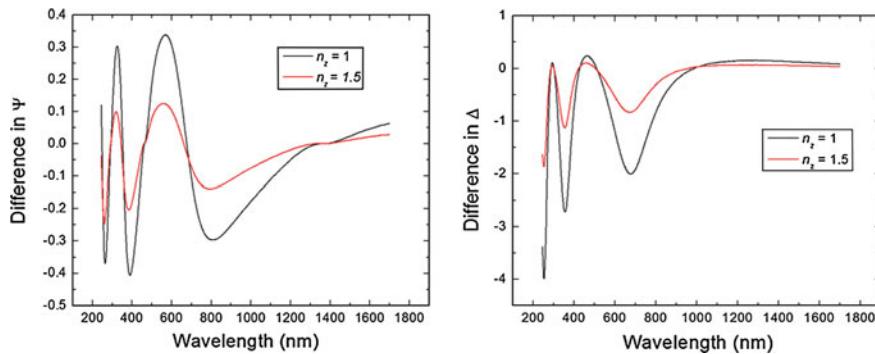
Fig. 16.10 Differences in generated ψ (*left*) and Δ (*right*) values between an isotropic and anisotropic-modeled graphene layer assuming different values for the z-polarized refractive index. The difference is more apparent for a lower value of $n(n = 1)$ due to its greater difference from the in-plane (x−y) refractive index

16.2.9 Current Areas of Interest for Graphene SE

As previously mentioned, graphene's CRI contains a UV absorption peak, the line shape and position of which is modified by excitonic effects [9]. Recent publications have shown graphene's absorbance to have an asymmetric line shape in this region [11, 34]. This is due to a Fano resonance caused by an interference effect between a discrete exciton state and the continuum band at the saddle point at the M-point of the BZ (see Fig. 16.1). The result is an asymmetric absorption profile. Modeling of SE data may account for this effect through the incorporation of an asymmetric oscillator that still maintains KK consistency. Recent modeling of CVD graphene transferred to SiO_2/Si using a PMMA-based process has incorporated this type of model and is shown in Fig. 16.11 for films of different grain size. For the purposes of this work, we define small-grain (SG) to mean domain size of several hundred nanometers and large-grain (LG) to mean domain sizes of several microns. We observe an increase in intensity of the absorbance peak for the film with smaller grains. While there is an increased density of states along grain boundaries compared to monocrystalline regions [35], greater amounts of contamination along these interfaces may also contribute to the higher absorbance peak observed.

Thermal decomposition of SiC is also a growth method of interest due to its scalability. In contrast to graphene transferred from a growth foil to a secondary substrate, graphene produced from this method sits on a "buffer layer" that forms between the SiC substrate and the graphene itself. This layer is thought to consist of a mixture of sp^2 and non-sp^2 bonding due to the sp^3 structure of the SiC. Samples of ∼4 layer graphene grown on the Si-terminated face of 6H-SiC were measured with SE. Due to the multilayer film thickness an anisotropic model was used to extract the CRI, as required by our simulation study. The resulting n and k values in the vicinity of the van Hove peak is plotted in Fig. 16.11 for comparison to CVD
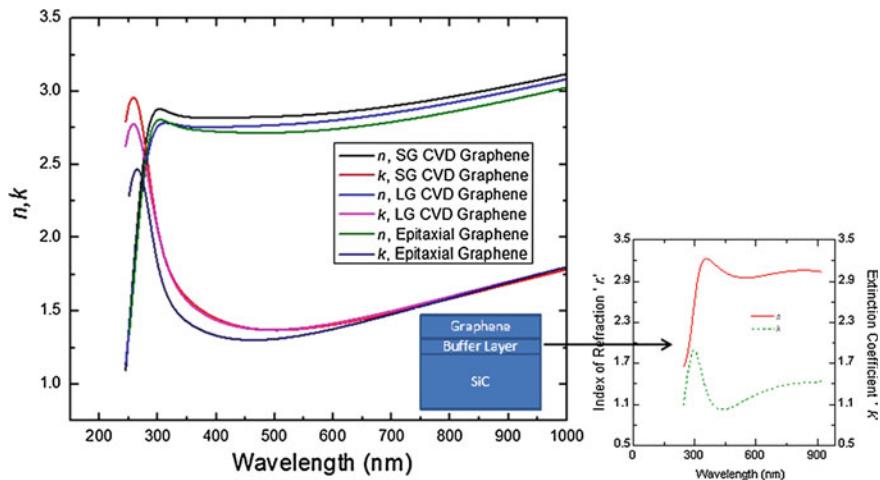
Fig. 16.11 Optical functions of small-grain (SG) and larger grain (LG) CVD graphene transferred to SiO$_2$/Si, compared to epitaxial graphene grown on 6H-SiC (optical functions of buffer layer shown on *lower right*). The optical dispersion model uses an asymmetric oscillator for the exciton-modified van Hove peak

graphene on SiO$_2$/Si. We observe a decreased intensity of the epitaxial graphene k response in comparison to the two CVD films on SiO$_2$/Si, in addition to a slight red-shift in the peak position. The latter difference may be due to the effect of the buffer layer, the main k peak of which was found to occur at a lower energy ($\sim$4.1 eV) than that of the $\sim$4.5 eV peak in graphene (shown in lower right of Fig. 16.11) [36]. Characterization of the buffer layer is important for device work, as it has been shown to possess a difference in mobility values when compared to those of the graphene region [37].

SE of FLG has evolved over the past several years due to the progress in graphene sample fabrication. The number of areas to be explored within graphene SE will increase in the coming years as growth/transfer processes are improved. The effect of different substrates on graphene's optical response is an area of current investigation and will continue to be of interest as CVD processes gain better control of layer number deposition. This work has shown the importance of SE in the characterization of "beyond CMOS" materials and highlights the necessity of consideration of the material thickness/dimensionality when studying optical properties such as anisotropy and excitonic effects on the dielectric function/CRI.

16.3 Nanoscale Single Crystal Silicon (*c*-Si) Films

Fundamental properties of dimensionally confined crystalline semiconductor materials are currently being explored by theoretical and experimental research groups. These low-dimensional structures (nanofilms, nanowires and nanodots) are

investigated for applications in transistor electronics, lasers, detectors, energy storage, solar cells, etc. The device research efforts in the CMOS industry are moving towards the use of nanoscale single crystal silicon quantum wells (*c*-Si QW) on insulator structures called extremely thin silicon-on-insulator (ETSOI) [38]. At the 22 nm node several leading semiconductor companies have reported that the thickness of the top silicon layer will be less than 10 nm. The research direction of the semiconductor industry indicates that the thickness will continue to shrink for nodes less than 22 nm. The reduced thickness of the top silicon layer impacts both device properties and the optical properties of nanostructures.

The dielectric function of *c*-Si films from 2 to 10 nm is found to have a significant dimensional dependence [39]. A majority of optical characterization of these nanoscale materials has previously been interpreted in terms of the effects of quantum confinement. In this chapter, the impact of both carrier confinement and electron-phonon interactions on the dielectric function of nanoscale films will be presented. The phonon dispersion of the *c*-Si QWs was further altered through the presence of dielectric layers above the nanoscale silicon. This points to the possibility of utilizing the elastic properties of nanostructures to alter phonon dispersion and build devices with new optical properties. In addition, the thickness dependent dielectric functions can be used for measurement of technologically important structures for the semiconductor industry. This fundamental study also has significant practical implications in the semiconductor industry.

16.3.1 Critical Points in Semiconductors

Silicon has a diamond cubic lattice having a lattice constant of $a_0 = 0.543$ nm with two atoms per primitive cell. Two interlocking face center cubic lattices displaced by ($a_0/4$, $a_0/4$, $a_0/4$) form the diamond structure. Silicon has the FCC Brillouin Zone (BZ), and the first BZ is a truncated octahedron as shown in Fig. 16.10a. The points of high-symmetry in the BZ are labeled in Fig. 16.12a and the high-symmetry lines along the [100], [110] and [111] directions from the center of the BZ (Γ-point) are labeled as Δ, Σ and Λ respectively. The relationship between the band structure and the linear optical response (dielectric function) has been the subject of considerable research [40–42]; most studies represent a theoretical understanding of the optical response at zero Kelvin [41]. The dielectric function of semiconductors have several sharp features that can be related to direct band transitions, where there is a constant energy difference between the valence band (VB) and the conduction band (CB). This constant energy difference along the various high-symmetry directions in the electronic band structure of Si is shown in Fig. 16.12b. This high joint density of states can result in sharp features in the optical absorption known as critical points (CPs) [42]. Silicon has several CPs between 1 and 6 eV as shown in Fig. 16.12c, that are related to direct transitions in the electronic band structure. The E_1 CP (3.4 eV) arises due to the transition from Λ_3 (VB) to Λ_1 (CB) along $k = (2\pi/a_0)$ $\langle 1/4, 1/4, 1/4 \rangle$ to the edge of the BZ [43]. Rohlfing and Louie [40] showed that the
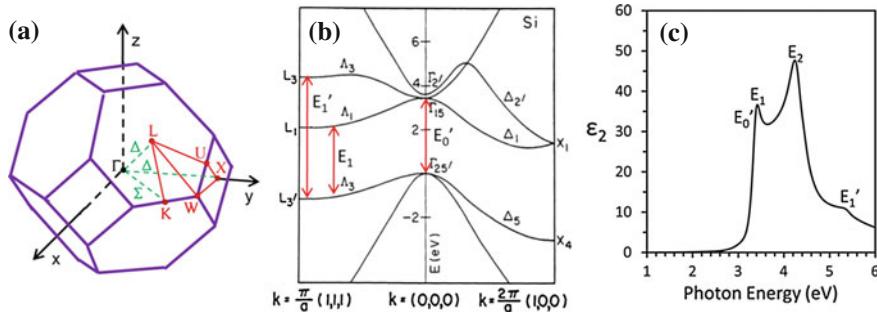
Fig. 16.12 **a** First Brillouin zone of a Si FCC lattice, **b** electronic band structure of Si calculated using $k \cdot p$ method [42] and **b** imaginary part of the dielectric function of bulk c-Si at room temperature between 1 and 6 eV. Adapted from [61] used with permission

impact of strong electron-hole (exciton) interactions result in an increased oscillator strength and energy of the E_1 CP. This excitonic direct gap transition demonstrates the strongest changes with dimensional confinement. The E_0' CP (degenerate CP of E_1) is due to transitions from the $\Gamma_{25'}$ (VB) to Γ_{15} (CB) at the center of the BZ and the E_1' CP (5.45 eV) is due to transitions from the Λ_3 (VB) to the second CB along the <111> direction. The E_2 peak in the dielectric function does not correspond to a single, well defined CP and it is due to contributions from a large range of transitions close to the edge of the BZ along the Σ and Δ direction [44, 45].

The energy and lifetime broadening of a CP can be extracted using direct space analysis. To enhance the CP structure and suppress the baseline effects a second derivative of the dielectric function is taken [45]. A Lorentzian line shape as shown in Eq. 16.3 is used to the fit the second derivative of the dielectric function of each parabolic CP.

$$\frac{d^2\varepsilon}{dE^2} = \left\{ \begin{array}{ll} \mu\,(\mu+1)\,A e^{i\beta}\left(E - E_g + i\Gamma\right)^{-\mu-2}, & \mu \neq 0 \\ A e^{i\beta}\left(E - E_g + i\Gamma\right)^{-2}, & \mu = 0 \end{array} \right\} \tag{16.3}$$

where A is the amplitude, β is the phase angle, E_g is the threshold energy, Γ is the broadening and μ is the order of singularity. The value of μ is based on the type of CP and it is 1, 1/2, 0 and $-1/2$ for discrete excitons, one-dimensional (1D), two-dimensional (2D) and three-dimensional (3D) one-electron transitions, respectively [39].

16.3.2 Carrier Confinement and Electron-Phonon Interactions

At the nanoscale dimensions both carrier confinement and electron-phonon interactions play a significant role on the dielectric function of semiconductor materials.

In the case of *c*-Si QWs the electrons in the CB due to optical excitation are confined between the buried-oxide layer and the top dielectric layer. This carrier confinement results in a change in the dielectric function with dimension. Also, the direct gap transitions or critical point (E_1 CP) in *c*-Si QW were shown to blue shift with increased dimensional confinement [39, 46–48]. Primarily these changes at nanoscale dimensions are related to quantum confinement of carriers, and disregard the role of change in phonon dispersion. Here we will show experimental evidence of both electron confinement and electron-phonon interaction effects on the direct transitions in nanoscale *c*-Si QWs. The phonon modes are modified in nanostructures and further altered when embedded within materials with large elastic mismatch [49]. In the case of bulk semiconductors the effect of electron-phonon interactions on the temperature dependence of the indirect band gap and the direct transitions in bulk semiconductors was reported by Cardona et al. [43, 45, 50, 51]. The phonon modes in *c*-Si QWs have strong dimensional dependence and also vary with the surrounding dielectric layers. To study the effect of phonon modes on the direct transitions due to dimensional changes, a temperature dependent study can be performed on *c*-Si QWs. To further demonstrate that the dielectric function will change with the surrounding layers the Si QWs can be embedded within different dielectric layers. Here we will focus on the thickness dependent carrier confinement and the role of electron-phonon interactions on the optical properties of nanostructures.

16.3.3 Low Temperature Measurements to Determine Average Phonon Energy

The energy and lifetime broadening of direct transitions in semiconductors have temperature dependence. This temperature dependence is primarily due to the contributions from the electron-phonon interactions and the change in band structure due to thermal expansion. Two types of electron-phonon interactions contribute to the temperature dependence—the Debye-Waller term and Fan or "self-energy" term. The Debye-Waller terms are obtained from the second-order electron-phonon Hamiltonian and the Fan terms are the first-order electron-phonon interactions taken in second-order perturbation theory [52]. The changes in the CP energy with temperature are related to the thermal expansion of the lattice as one term and the combined effect of the electron-phonon interactions [43] are accounted for by the other term in Eq. 16.4.

$$\left(\frac{\partial E_g}{\partial T} \right)_p = \left(\frac{\partial E_g}{\partial T} \right)_{thermal\ exp} + \left(\frac{\partial E_g}{\partial T} \right)_V \qquad (16.4)$$

The thermal expansion is described by Eq. 16.5 which results in a constant term in Eq. 16.4 assuming the bulk modulus (B) and coefficient of thermal expansion (α_L) are thickness independent.

$$\left(\frac{\partial E_g}{\partial T}\right)_{thermal\ exp} = -3B\alpha_L\left(\frac{\partial E_g}{\partial T}\right)_T \qquad (16.5)$$

The average interaction frequency '$K_B\theta/\hbar$' can be extracted by fitting the energy shifts due to both thermal expansion and electron-phonon interactions by a phenomenological Bose-Einstein statistical equation [45] as shown in Eq. 16.6.

$$\left(\frac{\partial E_g}{\partial T}\right)_p = E_B - a_B\left(1 + \frac{2}{e^{\theta/T} - 1}\right) \qquad (16.6)$$

The impact of optical and acoustic phonons on the conduction and valence band states has been reported by Cardona using temperature dependent electron-phonon spectral function along the various high-symmetry lines [43]. These theoretical calculations of the electron-phonon interaction terms for bulk silicon indicate that both optical and acoustic phonon contribute to the temperature dependence of the energy and lifetime broadening of the direct gap transitions.

16.3.4 Experimental Data for ETSOI

Nanoscale silicon films can be fabricated from bulk 300 mm SOI wafers (700 Å c-Si/1400 Å oxide/bulk Si). The two major steps that are performed to achieve the desirable thickness (∼2 to 10 nm) of the nanoscale silicon film are: (1) A wet oxidation step to consume at a faster rate most of the silicon layer by forming a silicon dioxide layer that is etched in dilute HF (100:1), and (2) to get a high quality surface on the top silicon layer a dry oxidation step and etching in dilute HF (300:1). The thicknesses of the ETSOI wafers were monitored using an in-line spectroscopic ellipsometer (KLA-Tencor Aleris) at each thinning step. The quality of the single crystal silicon in the QWs is maintained after thinning and can be verified by imaging using a high-resolution transmission electron microscopy (HRTEM) [39]. It is important to characterize the strain in the c-Si QWs as it has been shown to shift the E_1 CP energy [53]. The c-Si QWs developed no measurable strain during the thinning and it was determined by using Raman spectroscopy and high-resolution x-ray diffraction (HRXRD) [48]. We can investigate the change in the dielectric response in c-Si QWs by changing the top dielectric layers. The wafers are fabricated with a native oxide (∼1 nm), thicker thermal oxide and hafnium oxide (∼10 nm) layers to show a change in the phonon dispersions. To demonstrate the impact of changes in phonon dispersions on the E_1 CP energy and broadening a temperature dependent SE study of the nanoscale silicon films with the native oxide is performed. The temperature dependent measurements are performed on a dual rotating compensator ellipsometer (RC2 - J. A.Woollam Co., Inc.) equipped with a variable temperature cryostat that can go up to 4 K under liquid helium. Care should be taken during modeling low-temperature SE measurements to include a layer of ice on the sample surface.
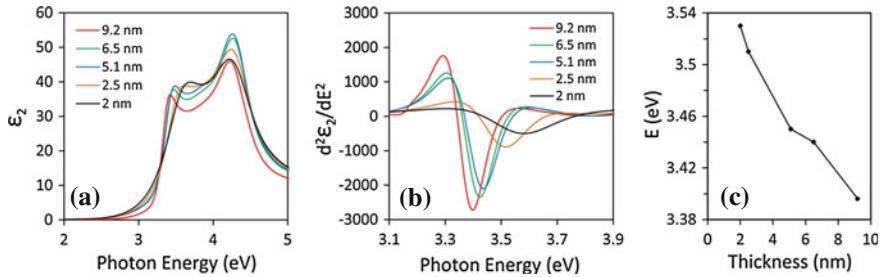
Fig. 16.13 **a** Imaginary part of the dielectric function of *c*-Si QWs at room temperature, **b** second derivative of the imaginary part of the dielectric function and **c** *blue* shift in the energy of the E_1 CP with decreasing thickness. Adapted from [61] used with permission

The thickness dependence of the dielectric function in *c*-Si films with a native oxide surface studied at room temperature is shown in Fig. 16.13a between 2 and 5 eV as measured by SE. The second derivative of the imaginary part of the dielectric function near the E_1 CP and the blue shift in the energy extracted from direct space analysis with the scaling of thickness are shown in Figs. 16.13b, c, respectively. It can be seen in Fig. 16.13a that the E_1 CP exhibits maximum shifts in energy and lifetime broadening with change in thickness and the E_2 CP does not have a significant energy shift but shows changes in oscillator strength. It has been reported by Jellison et al. that the oscillator strength of the E_2 CP is impacted by the surface preparation [54].

The dielectric function of the Bulk Si and nanoscale *c*-Si films is extracted at a series of temperatures between 4 and 300 K. The E_1 CP and the E_0' CP are both clearly separated (no longer degenerate) at low-temperatures in Bulk Si and nanoscale silicon films due to difference in the thermal coefficient of critical point energy shifts. This clear separation in the E_0' and E_1 CP energy are shown in Fig. 16.14 as a function of temperature in the second derivative of the real part of the dielectric function. This is in good agreement with what has been reported by several researchers who have studied the temperature dependence of the CP energies of Bulk Si [45, 54, 55]. The size effects more strongly influence the blue shift in the E_1 CP than the E_0' CP as can been seen for 2.5 nm *c*-Si QW Fig. 16.14b.

The change in optical response with temperature for 9, 7, and 2 nm *c*-Si films extracted using SE are shown in Fig. 16.15a–c. In the case of bulk Si and the 9 nm *c*-Si film the oscillator strength of the E_2 CP is stronger than the E_1 CP at all temperatures. However, in the case of 7 nm or less *c*-Si films with the native oxide surface layer the E_1 CP oscillator strength increases above that of the E_2CP.

The temperature and thickness dependence of the E_1 CP energy extracted using the direct space analysis is shown in Fig. 16.16. All the nanoscale films follow a similar trend with temperature as observed in bulk Si. This shift in energy with temperature is fitted by Eq. 16.6 to extract the average phonon frequency (θ) and is shown in the inset of Fig. 16.16. Lautenschlager et al. reported that the stronger the contribution of acoustic phonons on the E_1 CP the lower the average phonon frequency [43]. It is evident from the average phonon frequency that the contributions of acoustic phonons
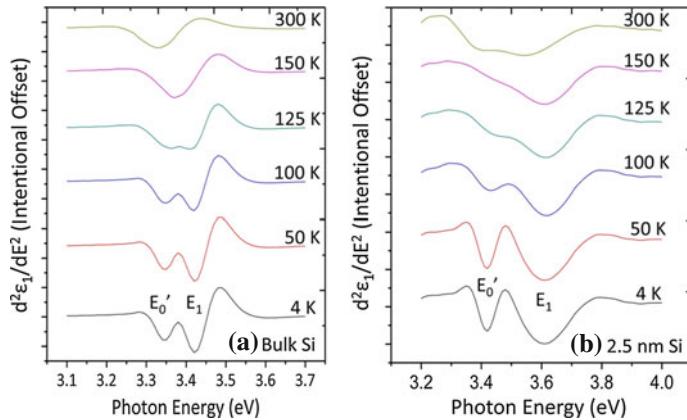
Fig. 16.14 Second derivative of the temperature dependent real part of the dielectric function for **a** bulk Si and **b** 2.5 nm *c*-Si film. Adapted from [61] used with permission
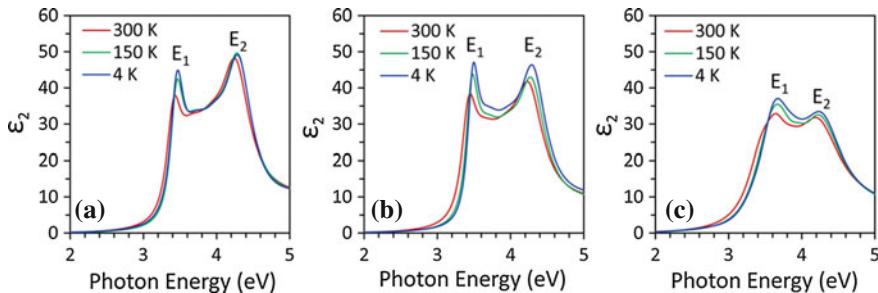
Fig. 16.15 Temperature dependent imaginary part of the dielectric function (ε_2) for **a** 9 nm, **b** 7 nm and **c** 2 nm of *c*-Si QW [56]. Adapted from Ref. [61] used with permission

to the electron-phonon interactions increase with decreasing *c*-Si QW thickness. The increased contribution of acoustic modes curves due to change in dimension can be interpreted based on recent theoretical modeling studies published by Hepplestone and Srivastava using the adiabatic bond charge method for the lattice dynamics of silicon nanofilms [57]. Their work suggests that the lowest lying optical phonon (LOP) modes are difficult to populate in nanofilms of Si and thereby increasing the acoustic phonon mode contribution with decrease in thickness. The contributions from all these phonon states effectively broaden the lifetime of the direct band gap transitions and these modes should be observable using Brillouin Scattering [58]. Thus this study clearly indicates the importance of the electron-phonon interactions in determining the optical properties of nanoscale solids.
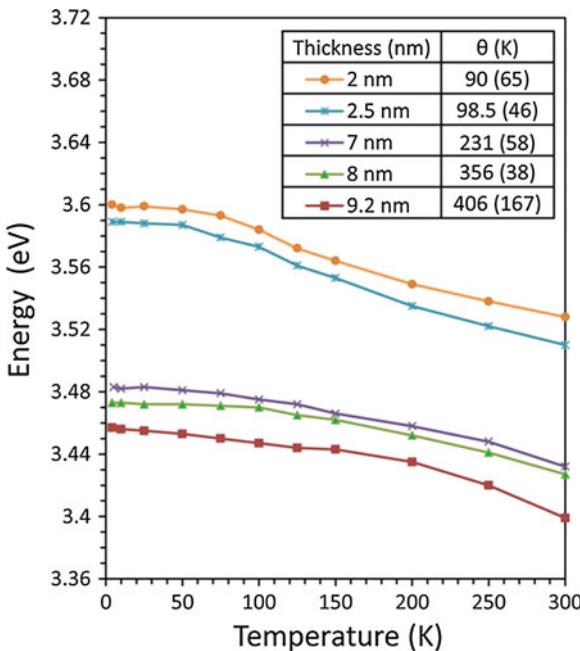
Fig. 16.16 The thickness and temperature dependence of the E_1 CP energy of c-Si QWs (Inset Table: Average phonon frequency with 90 % confidence limits). Adapted from Refs. [56, 61] used with permission

16.3.5 Impact of Surrounding Dielectric Layers on ETSOI Optical Properties

At the nanoscale dimensions, we provide here experimental evidence that the optical properties of the c-Si layer changes when confined between various dielectric layers of differing stiffness. We study the three sets of wafers previously mentioned: ETSOI with top layers of native oxide ($\sim$1 nm), thicker thermal oxide, and hafnium oxide ($\sim$10 nm). In Fig. 16.17 we show differences in imaginary part of the refractive index of 5 nm c-Si QWs each with the three different surface layers. Although the SiO_2 covered c-Si QW (5.7 eV) barrier depth is higher than the HfO_2 covered c-Si QW (2.7 eV), all the three samples with nearly identical thickness should have the same amount of carrier confinement at room temperature. We interpret this change in the optical properties as a representative of the effect of changes in the acoustic phonon modes in the three samples due to the different Young's moduli of HfO_2 ($\sim$370 GPa) and SiO_2 ($\sim$75 GPa).

Another compelling piece of evidence is the small red shift versus decreasing thickness of the E_1 CP of hafnium oxide covered c-Si QWs as shown in Fig. 16.18. This red shift is unexpected, and indicates that the electrons and holes excited at the E_1 CP of the hafnium oxide covered c-Si QWs are more strongly influenced
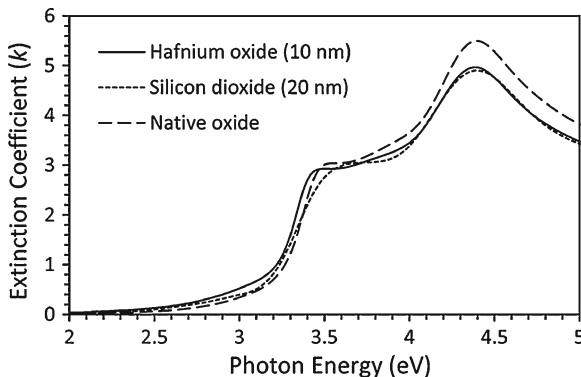
Fig. 16.17 Imaginary part of the refractive index of *c*-Si QWs (∼5 nm) with native oxide, 20 nm SiO₂, and 10 nm HfO₂ with a SiO₂ interfacial layer. Adapted from [61] used with permission
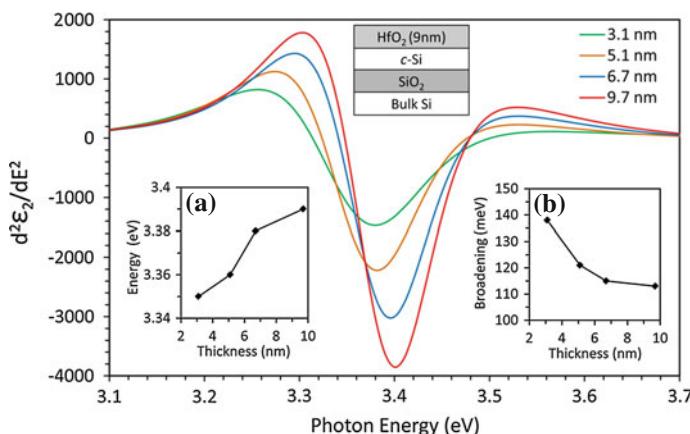
Fig. 16.18 Second derivative of the imaginary part of the dielectric function of *c*-Si QW with a HfO₂ surface layer. Inset: **a** energy and **b** lifetime broadening (Γ) of the E_1 CP extracted using direct space analysis. Adapted from [56] used with permission

by electron-phonon interactions than quantum confinement of carriers. The lifetime broadening of the E_1 CP increases for thinner *c*-Si, due to change in the electron-phonon interactions induced by changes in the phonon modes of thin films. This provides further proof of the importance of electron-phonon interactions on the optical response of nanoscale single crystal films. We would like the reader to take note that the changes in energy of the E_1 CP reported here are much larger than the surface preparation dependent changes in the dielectric function reported by Aspnes et al. [59, 60].

There are several fundamental and practical implications of the dimensional changes in the dielectric function. One of the major practical implications of model-

ing the top silicon thickness in ETSOI with bulk optical constants leads to inaccurate thickness measurements as the optical properties are thickness and surface layer dependent. Also, the carrier transport in devices can be improved by surrounding the *c*-Si QWs with selective dielectric films to lower electron-phonon scattering and thereby enhancing the mobility. We suggest that the electron-phonon interactions and quantum confinement effects will significantly impact semiconductor nanowires and nanodots. This section suggests that the optical properties of dimensionally confined materials can be selectively altered through the design of structures with dielectric layers having appropriate elastic properties. This chapter points to new avenues into the fabrication of dimensionally confined structures with desired optical properties.

16.4 Implications of Nanoscale Properties on SE Research & Development and Education

The examples of graphene and c-Si QWs illustrate the challenging nature of optical measurement of nanoscale materials and structures. This chapter presents some evidence that the dielectric function/CRI of nanoscale materials can be altered by surrounding materials. The two examples provided in this chapter were examples of the influence of surrounding materials on direct gap transitions having an excitonic nature. This leads to several suggestions for research, development, and educations (R, D & E) for SE of nanoscale materials.

- Careful use of effective medium approximations shows great promise for measuring nanoscale material dimensions for embedded materials and incomplete layers. Thus EMA theory and application should receive attention from R, D & E.
- Optical properties of nanoscale materials are influenced by both quantum confinement and changes in phonon dispersion. Exploration of these combined effects is just beginning making it an important area for R, D & E.
- The success of the Rigorous Coupled Wave method for optical measurement of the dimensions of features of ordered arrays of lines, holes, and more complicated structures (i.e., scatterometry or optical critical dimension) in semiconductor manufacturing provides a new approach for measurement of nanoscale materials in other fields.
- R & D has yet to address challenging systems such as partially ordered materials systems such as the fingerprint pattern observed for block co-polymers.
- Transfer of advances in research requires application of new optical modeling ideas to a variety of materials in development before they can be used in manufacturing applications.
- Application of Muller Matrix Spectroscopic Ellipsometry to fully capture the anisotropic nature of nanoscale materials

References

1. G.D. Hutcheson, The first nanochips. Sci. Am. **290**(4), 48–55 (2004)
2. V.K. Kamineni, J. Hilfiker, J. Freeouf, J. Fielden, S. Consiglio, R. Clark, G.J. Leusink, A.C. Diebold, Extension of Far UV spectroscopic ellipsometry studies of high-κ dielectric films to 130 nm. Thin Solid Films **519**, 2894 (2011)
3. M. Di, E. Bersch, S. Consiglio, R. Clark, T. Zhang, P. Tyagi, G. Leusink, T. Kaack, A.C. Diebold, Spectroscopic ellipsometry characterization of high-κ metal gate stacks with V_t shift layers. Thin Solid Films **519**, 2889 (2011)
4. P. Leray, G.F. Lorusso, S. Cheng, N. Collaert, M. Jurczak, S. Shirke, Accurate and reliable optical CD of MuGFET down to 10 nm, metrology, inspection, and process control for microlithography XXI, in *Proceedings of SPIE*, 2007 ed. by C.N. Archie, vol. 6518, p. 65183B
5. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, A.A. Firsov, Electric field effect in atomically thin carbon films. Science **306**(5695), 666 (2004)
6. L.A. Falkovsky, Optical properties of graphene. J. Phys. Conf. Ser. **129**, 012004 (2008)
7. K.F. Mak, M.Y. Sfeir, Y. Wu, C.H. Lui, J.A. Misewich, T.F. Heinz, Measurement of the optical conductivity of graphene. Phys. Rev. Lett. **101**, 196405 (2008)
8. R.R. Nair, P. Blake, A.N. Grigorenko, K.S. Novoselov, T.J. Booth, T. Stauber, N.M.R. Peres, A.K. Geim, Fine structure constant defines visual transparency of graphene. Science **320**, 1308 (2008)
9. L. Yang, J. Deslippe, C.H. Park, M.L. Cohen, S.G. Louie, Excitonic effects on the optical response of graphene and bilayer graphene. Phys. Rev. Lett. **103**, 186802 (2009)
10. V.G. Kravets, A.N. Grigorenko, R.R. Nair, P. Blake, S. Anisimova, K.S. Novoselov, A.K. Geim, Spectroscopic ellipsometry of graphene and an exciton-shifted van Hove peak in absorption. Phys. Rev. B **81**, 155413 (2010)
11. D.-H. Chae, T. Utikal, S. Weisenburger, H. Giessen, K.v. Klitzing, M. Lippitz, J. Smet, Excitonic fano resonance in free-standing graphene. Nano Lett. **11**, 1379 (2011)
12. J.C. Charlier, J.P. Michenaud, First-principles study of the electronic properties of simple hexagonal graphite. Phys. Rev. B **46**, 4531 (1992)
13. J.G. Carter, R.H. Heubner, R.N. Hamm, R.D. Burkhoff, Optical properties of graphite in the region 1100 to 3000 Å. Phys. Rev. **137**(2A), 639 (1965)
14. J.C. Charlier, J.P. Michenaud, Tight-binding model for the electronic properties of simple hexagonal graphite. Phys. Rev. B **44**, 13237 (1991)
15. S. Latil, S.V. Meunier, L. Henrard, Massless fermions in multilayer graphitic systems with misoriented layers: Ab initio calculations and experimental fingerprints. Phys. Rev. B **76**, 201402 (2007)

16. P. Blake, E.W. Hill, A.H. Castro Neto, K.S. Novoselov, D. Jiang, R. Yang, T.J. Booth, A.K. Geim, Making graphene visible. Appl. Phys. Lett. **91**, 063124 (2007)

17. A.C. Ferrari, J.C. Meyer, V. Scardaci, C. Casiraghi, M. Lazzeri, F. Mauri, S. Piscanec, D. Jiang, K.S. Novoselov, S. Roth, A.K. Geim, Raman spectrum of graphene and graphene layers. Phys. Rev. Lett. **97**, 187401 (2006)

18. J.W. Weber, V.E. Calado, M.C.M. van de Sanden, Optical constants of graphene measured by spectroscopic ellipsometry. Appl. Phys. Lett. **97**, 091904 (2010)

19. X. Li, W. Cai, J. An, S. Kim, J. Nah, D. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S.K. Banerjee, L. Colombo, R.S. Ruoff, Large-area synthesis of high-quality and uniform graphene on copper foils. Science **324**, 1312 (2009)

20. A. Reina, X. Jia, J. Ho, D. Nezich, H. Son, V. Bulovic, M.S. Dresselhaus, J. Kong, Large area, few-layer graphene films on arbitrary substrates by chemical vapor deposition. Nano Lett. **9**, 30 (2009)

21. S. Bae, H. Kim, Y. Lee, X. Xu, J.S. Park, Y. Zheng, J. Balakrishnan, T. Lei, H.R. Kim, Y. Il Song, Y.J. Kim, K.S. Kim, B. Ozyilmaz, J.H. Ahn, B.H. Hong, S. Iijima, Roll-to-roll production of 30-inch graphene films for transparent electrodes. Nat. Nanotechnol. **5**(8), 574–578 (2010)

22. X. Li, C.W. Magnuson, A. Venugopal, J. An, J.W. Suk, B. Han, M. Borysiak, W. Cai, A. Velamakanni, Y. Zhu, L. Fu, E.M. Vogel, E. Voelkl, L. Colombo, R.S. Ruoff, Graphene films with large domain size by a two-step chemical vapor deposition process. Nano Lett. **10**(11), 4328–4334 (2010)

23. O.V. Yazyev, S.G. Louie, electron transport in polycrystalline graphene. see: http://arxiv.org/abs/1007.1703

24. W. Wang, M. Balooch, C. Claypool, M. Zawaideh, K. Farnaam, Combined refl ectometry-ellipsometry technique to measure graphite down to monolayer thickness. Solid State Technol. **52**, 18 (2009)

25. F.J. Nelson, V.K. Kamineni, T. Zhang, E.S. Comfort, J.U. Lee, A.C. Diebold, Optical properties of large-area polycrystalline chemical vapor deposited graphene by spectroscopic ellipsometry. Appl. Phys. Lett. **97**, 253110 (2010)

26. M. Bruna, S. Borini, Optical constants of graphene layers in the visible range. Appl. Phys. Lett. **94**, 031901 (2009)

27. P.Y. Huang, C.S. Ruiz-Vargas, A.M. van der Zande, W.S. Whitney, S. Garg, J.S. Alden, C.J. Hustedt, Y. Zhu, J. Park, P.L. McEuen, D.A. Muller, Imaging grains and grain boundaries in single-layer graphene: an atomic patchwork quilt. Nature **469**, 389 (2011)

28. K. Kim, Z. Lee, W. Regan, C. Kisielowski, M.F. Crommie, A. Zettl, Grain boundary mapping in polycrystalline graphene. ACS Nano **5**(3), 2142 (2011)

29. F. Nelson, V. Kamineni, T. Zhang, E. Comfort, J.U. Lee, A.C. Diebold, Spectroscopic ellipsometry of CVD graphene. ECS Trans. **35**(3), 173 (2011)

30. H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (Wiley, Chichester, 2007)

31. J. Lee, R.W. Collins, V.S. Veerasamy, J. Robertson, Analysis of amorphous carbon thin films by spectroscopic ellipsometry. J. Non-Cryst. Solids **227–230**(1), 617 (1998)

32. K.F. Mak, C.H. Lui, J. Shan, T.F. Heinz, Observation of an electric-field-induced band gap in bilayer graphene by infrared spectroscopy. Phys. Rev. Lett. **102**, 256405 (2009)

33. G.E. Jellison Jr, J.D. Hunn, H.N. Lee, Measurement of optical functions of highly oriented pyrolytic graphite in the visible. Phys. Rev. B **76**, 085125 (2007)

34. K.F. Mak, J. Shan, T.F. Heinz, Seeing many-body effects in single- and few-layer graphene: observation of two-dimensional saddle-point excitons. Phys. Rev. Lett. **106**, 046401 (2011)

35. J. Cervenka, M.I. Katsnelson, C.F.J. Flipse, Room-temperature ferromagnetism in graphite driven by two-dimensional networks of point defects. Nat. Phys. **5**, 840 (2009)

36. F. Nelson, A.C. Diebold, A. Sandin, D.B. Dougherty, D.E. Aspnes, J.E. Rowe, Optical and structural characterization of epitaxial graphene on 6H-SiC(0001)-Si by spectroscopic ellipsometry, auger and STM, in *Proceedings of the 39th Conference on the Physics and Chemistry of Surfaces and Interfaces (PCSI-39)*, 2012 (In preparation)

37. T. Hofmann, C.M. Herzinger, J.L. Tedesco, D.K. Gaskill, J.A. Woollam, M. Schubert, Terahertz ellipsometry and terahertz optical-hall effect. Thin Solid Films **519**(9), 2593 (2011)
38. A. Majumdar, X. Wang, A. Kumar, J.R. Holt, D. Dobuzinsky, R. Venigalla, C. Ouyang, S.J. Koester, W. Haensch, Gate length and performance scaling of undoped-body extremely thin SOI MOSFETs. IEEE Electron Device Lett. **30**(4), 413 (2009)
39. J. Price, A.C. Diebold, Spectroscopic ellipsometry characterization of ultrathin silicon-on-insulator films. J. Vac. Sci. Technol. B **24**(4), 2156 (2006)
40. M. Rohlfing, S.G. Louie, Electron-hole excitations and optical spectra from first principles. Phys. Rev. B **62**(8), 4927 (2000)
41. M. Cardona, L.F. Lastras-Martinez, D.E. Aspnes, Comment on ab initio calculation of excitonic effects in the optical spectra of semiconductors. Phys. Rev. Lett. **83**(19), 3970 (1999)
42. P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996)
43. P. Lautenschlager, P.B. Allen, M. Cardona, Phonon-induced lifetime broadenings of electronic states and critical points in Si and Ge. Phys. Rev. B **33**(8), 5501 (1986)
44. E.O. Kane, Band structure of silicon from an adjusted Heine-Abarenkov calculation. Phys. Rev. B **146**(2), 558 (1966)
45. P. Lautenschlager, M. Garriga, L. Vina, M. Cardona, Temperature dependence of the dielectric function and interband critical points in silicon. Phys. Rev. B **36**(9), 4821 (1987)
46. X. Zhao, C.M. Wei, L. Yang, M.Y. Chou, Quantum confinement and electronic properties of silicon nanowires. Phys. Rev. Lett. **92**(23), 236805 (2004)
47. M.I. Alonso, I.C. Marcus, M. Garriga, A.R. Goni, Evidence of quantum confinement effects on interband optical transitions in Si nanocrystals. Phys. Rev. B **82**, 045302 (2010)
48. A.C. Diebold, J. Price, Observation of quantum confinement and quantum size effects. Phys. Status Solidi A **205**(4), 896 (2008)
49. A.A. Balandin, E.P. Pokatilov, D.L. Nika, Phonon engineering in Hetero- and nanostructures. J. Nanoelectron. Optoelectron. **2**, 140 (2007)
50. M. Cardona, T.A. Meyer, M.L.W. Thewalt, Temperature dependence of the energy gap of semiconductors in the low-temperature limit. Phys. Rev. Lett. **92**(19), 196403 (2004)
51. M. Cardona, Electron-phonon interaction in tetrahedral semiconductors. Solid State Commun. **133**, 3 (2005)
52. P.B. Allen, M. Cardona, Temperature dependence of the direct gap of Si and Ge. Phys. Rev. B **27**(8), 4760 (1983)
53. R. Lange, K.E. Junge, S. Zollner, S.S. Iyer, A.P. Powell, K. Eberl, Dielectric response of strained and relaxed $Si_{1-x-y}Ge_xC_y$ alloys grown by molecular beam epitaxy on Si(001). J. Appl. Phys. **80**(8), 4578 (1996)
54. G.E. Jellison, F.A. Modine, Optical constants for silicon at 300 and 10K determined from 1.64 to 4.73 eV by ellipsometry. J. Appl. Phys. **53**(5), 3745 (1982)
55. A. Daunois, D.E. Aspnes, Electroreflectance and ellipsometry of silicon from 3 to 6 e V. Phys. Rev. B **18**(4), 1978 (1824)
56. V.K. Kamineni, A.C. Diebold, Electron-phonon interaction effects on the direct gap transitions of nanoscale Si films. Appl. Phys. Lett. **99**, 151903 (2011)
57. S.P. Hepplestone, G.P. Srivastava, Lattice dynamics of silicon nanostructures. Nanotechnology **17**, 3288 (2006)
58. N. Lou, J. Groenen, G. Benassayag, A. Zwick, Acoustics at nanoscale: Raman-Brillouin scattering from thin silicon-on-insulator layers. Appl. Phys. Lett. **97**, 141908 (2010)
59. L. Mantese, K.A. Bell, U. Rossow, D.E. Aspnes, Evidence of near-surface localization of excited electronic states in crystalline Si. J. Vac. Sci. Technol. B **15**(4), 1196 (1997)
60. D.E. Aspnes, L. Mantese, K.A. Bell, U. Rossow, Coherence effects and time dependences of the optical response of surfaces and interfaces of optically absorbing materials. Phys. Status Solidi B **220**, 709 (2000)
61. V.K. Kamineni, *Electron-Phonon Interactions and Quantum Confinement Effects on Optical Transitions in Nanoscale Silicon Films* (Dissertation submitted to the University at Albany, SUNY, 2011)

Chapter 17
Ellipsometry of Semiconductor Nanocrystals

Peter Petrik and Miklos Fried

Abstract In this chapter we make an attempt to give a comprehensive overview on the optical modeling of layer structures that accommodate or are entirely composed of semiconductor nanocrystals. This research field is huge both in terms of the theories of effective dielectric functions and applications. The dielectric function of single-crystalline semiconductors can be determined on high quality reference materials. The accuracy of the reference data depends mostly on the numerical or experimental elimination of the surface effects like oxides, nanoroughness, contamination, etc.

17.1 Introduction

Most techniques used for the preparation of semiconductor films (e.g. chemical vapor deposition, evaporation, sputtering, pulsed laser deposition, etc.) result in polycrystalline materials with a structure that strongly depends on the parameters of the sample preparation. Therefore, dielectric function references can only be used in a parameterized form. The dielectric function of poly- and nanocrystalline[1] thin films can be derived from the single-crystalline dielectric functions using parameteriza-

P. Petrik (✉) · M. Fried
Research Institute for Technical Physics and Materials Science (MTA-MFA),
Hungarian Academy of Sciences, Konkoly Thege Rd. 29-33, 1121 Budapest,
H-1525 Budapest, Hungary
e-mail: petrik.peter@ttk.mta.hu

M. Fried
e-mail: fried.miklos@ttk.mta.hu

[1] Very recently, a definition of nanomaterial has been given by the European Commision (http://ec.europa.eu/environment/chemicals/nanotech/index.htm): "*A natural, incidental or manufactured material containing particles, in an unbound state or as an aggregate or as an agglomerate and where, for 50 % or more of the particles in the number size distribution, one or more external dimensions is in the size range 1–100 nm.*"

tion of the critical point features or using the effective medium theory by "mixing" these single-crystalline components with voids, polycrystalline or amorphous components. Finally, when the effective dielectric functions are calculated for a given nanocrystalline structure in a given layer, the vertical change of the composition can be measured by defining multi-layer models or models with analytical depth profiles.

17.2 Dielectric Function of Semiconductors

Dielectric functions have been measured and tabulated for a wide range of semiconductors (Si, Ge, SiC, SiGe, AlAs, AlGaAs, AlN, AlSb, BN, CdTe, CdS, CIGS, GaAs, GaN, GaP, GaSb, HgTe, InGaAs, InAs, InN, InP, InSb, PbS, PbSe, PbTe, ZnS, ZnSe, ZnTe, just to mention some relevant materials, found in a commercial database). Most of these data have been published [9, 14, 60, 63], are available in the databases shipped with commercial ellipsometers, or on community websites [97]. Accurate reference dielectric functions of single-crystalline semiconductors are crucial when aiming at sub-nanometer precision. Even the dielectric functions of the purest and most perfect single-crystalline material, the microelectronic grade silicon shows differences between different authors (Fig. 17.1). Using the various references of Fig. 17.1 for a native oxide-covered single-crystalline silicon wafer, differences of ≈ 0.2 nm in the oxide thickness are obtained.[2] The major problem is how surface imperfections (oxide layer, nanoroughness, contamination) are taken into account.

Even amorphous dielectric functions have been tabulated as reference data [14, 41], although their dielectric functions strongly depend on the preparation [41, 81, 110]. A relevant example is the dielectric function of the ion implantation-amorphized and relaxed silicon [41]. The dielectric functions of the two kinds of amorphous silicon differ significantly (Fig. 17.2).

Most of the reference dielectric functions of semiconductors are available in a parameterized form as well. The use of analytical functions has several advantages. First of all, non-idealities, like in the case of polycrystalline materials can be derived and taken into account using these parameterizations. Furthermore, measurements at wavelengths other than those given in the data file of the reference dielectric function can be interpolated and extrapolated with higher accuracy.

There are a range of approaches from using simple Lorentz oscillators through quantum mechanical considerations (model dielectric function [1], generalized oscillator model [22]) to empirical formulas like the generalized critical point model [66].

---

[2] Using the photon energy range of 2–5 eV covered by all the references.
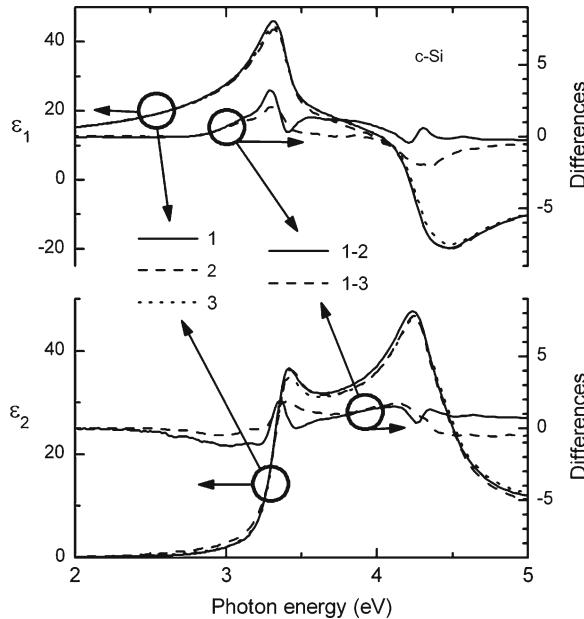
Fig. 17.1 Real (ε_1) and imaginary (ε_2) parts of the dielectric functions of single-crystalline silicon from Refs. [60] (*line 1,*) [14] (*line 2,*) and [63] (*line 3*). Differences for lines 1–2 and 1–3 are also plotted

17.3 Analytical Models

Dielectric responses of all materials follow some relatively simple rules and can be understood in microscopic terms. Here we consider these rules for both homogeneous and heterogeneous materials. Examples of generic dielectric functions for single-line absorption processes, metals and semiconductors are given in Fig. 17.3 and discussed individually below [53].

17.3.1 Drude

Simple metals are materials whose dominant characteristic is electronic charge that is able to move more or less freely through the material. In the Drude model, the equation of motion of any single carrier in an external field is

$$F = -eE exp(-i\omega t) = ma = md^2x/dt^2 = -\omega^2\, m\, \Delta x\, exp(-i\omega t), \quad (17.1)$$

where its position is assumed to be given by x(t) $=$ x$_0$ $+$ Δx exp($-$i ωt). Solving this equation for Δx and substituting the result in Maxwell's equations leads to the Drude expression,
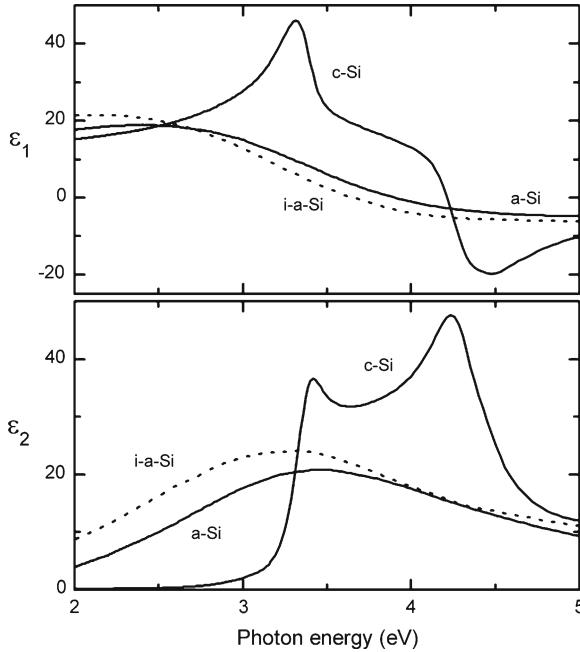
Fig. 17.2 Real (ε_1) and imaginary (ε_2) parts of the dielectric functions of single-crystalline silicon (c-Si, [60]), deposited amorphous silicon (a-Si, [14]) and ion implantation-amorphized silicon (i-a-Si, [41])

$$\varepsilon = 1 - 4\,\pi\,ne^2/m_e\omega^2 = 1 - \omega_p{}^2/\omega^2, \qquad (17.2)$$

where $\omega_p = 4\pi ne^2/m_e$ is the plasma frequency, n and m_e are the electron density and electron mass, respectively. The Drude model works quite well for most metals and for free-carriers in semiconductors. Note that we derived this dielectric function from purely classical-mechanical perspective of point charges accelerating in accordance with Newton's laws, i.e., the motion in an electric field.

More realistic calculations include electron loss due to scattering, in which case,

$$\varepsilon = 1 - \omega_p{}^2/\omega(\omega + i/\tau), \qquad (17.3)$$

where τ is the lifetime. The lifetime is an important parameter in poly- or nanocrystalline metal films of finite grain sizes. It is determined approximately as

$$1/\tau = 1/\tau_0 + 2v_f/d, \qquad (17.4)$$

where v_f is the Fermi-velocity and d is the diameter of grain. The relevant length scale is the electron mean free path $v_f\tau_0$, which in noble metals is on the order of
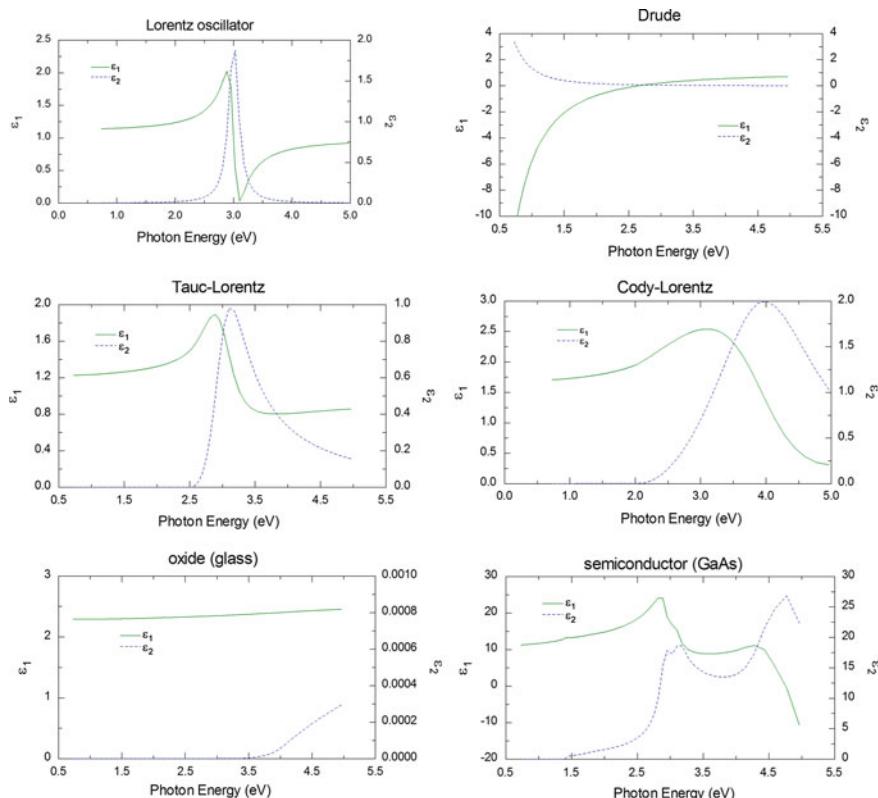
Fig. 17.3 Examples of generic dielectric functions for single-line absorption processes (Lorentz, Tauc-Lorentz and Cody-Lorentz oscillator), Drude metal, semiconductor (GaAs), and insulator (glass) [Drawn by the WVASE software of the Woollam Co., Inc.]

a few dozen nanometers. Thus for Cu, Au or Ag the dielectric function is affected when grain sizes drop below these values.

Another important application of the Drude model is to describe the optical contribution of free-carriers in heavily doped semiconductors. To the lowest order correction is made simply by adding the Drude expression to the ordinary dielectric function measured for undoped material. Note the material properties encoded: electron density and electron mass in the plasma frequency and grain size in the lifetime.

The calculations can be extended by adding a restoring term $-K(x - x_0)$ to the force equations, where K is the force constant. The equation then approximately describes the bound charge that dominates the optical response of dielectrics, oxides and amorphous semiconductors. In this case ε becomes

$$\varepsilon = 1 - \omega_p^2/(\omega^2 - \omega_0^2 + i\omega/\tau), \tag{17.5}$$

where $\omega_0^2 = K/m$ is the resonance frequency squared. Resonance leads to strong absorption in the refractive index at longer wavelengths. The above expression also applies to excitons, impurity transitions and lattice vibrations.

An important aspect of the above derivations is generality; because the same fundamental considerations (mass, restoring forces, loss) apply to all charges, all dielectric response functions have the same form. Consequently, any dielectric response for any material can be represented as a superposition of terms of the above form (spectral representation).

The same principles also apply to semiconductors, although the spectra are now more complex. Crystalline semiconductors are characterized by relatively sharp structures, whereas amorphous semiconductors generally have only a single broad peak. The transition between amorphous and crystalline behavior in semiconductors is ill-defined, depending in part on the technique used to observe it, but usually considered to extend approximately from 2 to 50 nm. On the other hand, optical excitations in insulators are usually so well localized that size effects cannot be seen.

A realistic model becomes even more important when one wishes to simulate spectra from alloys, such as $Al_xGa_{1-x}As$ or Si_xGe_{1-x}, which contain critical points in the optical spectrum that vary continuously with composition x. The best fits can be obtained when the critical points are modeled with the use of one or more Lorentz oscillators. For each oscillator, the peak photon-energy, width and amplitude are fit as a function of x, allowing the composite dielectric function to be calculated as a function of x. This approach works well near the critical point but breaks down at small photon energies, where the absorption coefficient becomes small [130].

17.3.2 Lorentz

The Lorentz model is a classical model where a negatively charged electron is bound to a positively charged atomic nucleus with a spring. Thus single-electron "two-level atoms" ($Ne = Na$) are assumed for simplicity.

Each electron is assumed to be bound to its atom with a resonant frequency of ω_0 in the near-ultraviolet range; as a result, there is no free-electron or Drude component to the dielectric function. If light is shone, the AC electric field of the light $E = E_0 exp(i\omega t)$ will induce dielectric polarization in the x direction. The Lorentz model assumes a physical model in which the electron oscillates in viscous fluid. In this case, the position of the atomic nucleus is fixed, since the mass of the atomic nucleus is far larger than that of the electron. If we use Newtons second law, the physical model is expressed as:

$$m_e d^2 x/dt^2 = -m_e \Gamma dx/dt - m_e \omega_0^2 x - eE_0 exp(i\omega t), \qquad (17.6)$$

where m_e and e show the mass and charge of the electron, respectively. In Eq. (17.6), the first term on the right represents the viscous force of the viscous fluid. In general,

the viscous force is proportional to the speed of an object when the speed is slow. The Γ in Eq. (17.6) represents a proportional constant of the viscous force, known as the damping coefficient. (It is high in amorphous semiconductors.) The second term on the right expresses that the electron moved by the electric field of light is restored according to Hooks law $F = -KFx$, and ω_0 shows the resonant frequency of the spring $\omega_0^2 = KF/m_e$. The last term on the right shows the electrostatic force $F = qE$. Eq. (17.6) represents the forced oscillation of the electron by the external AC electric field. By this forced oscillation, the electron oscillates at the same frequency as the AC electric field [i.e., exp (iωt)]. Thus, if we assume that the solution of Eq. (17.6) is described by the form $x(t) = a \cdot exp(i\omega t)$, the first and second derivatives of $x(t)$ are given by $dx/dt = ia\omega\, exp(i\omega t)$ and $d^2x/dt^2 = -a\omega^2\, exp(i\omega t)$, respectively. By substituting these into Eq. (17.6) and rearranging the terms, we get

$$a = -(eE_0/m_e)(1/(\omega_0^2 - \omega^2 + i\Gamma\,\omega)). \qquad (17.7)$$

On the other hand, if the number of electrons per unit volume is given by Ne, the dielectric polarization is expressed as $P = -eN_e x(t)$. From $x(t) = a\, exp(i\omega t)$, we obtain $P = -eN_e a\, exp(i\omega t)$. By substituting $P = -eN_e a\, exp(i\omega t)$ and $E = E_0\, exp(i\omega t)$ into $\varepsilon = 1 + P/\varepsilon_0 E = 1 + \chi$, where χ is the dielectric susceptibility $\chi \equiv P/\varepsilon_0 E$, we obtain the dielectric constant as follows:

$$\varepsilon = 1 + (e^2 N_e/\varepsilon_0 m_e)(1/(\omega_0^2 - \omega^2 + i\Gamma\,\omega)) \qquad (17.8)$$

This equation represents the Lorentz model. If we multiply by $(\omega_0^2 - \omega^2 - i\Gamma\,\omega)$ both the numerator and the denominator of Eq. (17.8), we get

$$\varepsilon_1 = 1 + (e^2 N_e/\varepsilon_0 m_e)(\omega_0^2 - \omega^2)/((\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2)), \qquad (17.9)$$
$$\varepsilon_2 = (e^2 N_e/\varepsilon_0 m_e)(\Gamma\,\omega/((\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2)). \qquad (17.10)$$

In actual data analysis, we commonly express the Lorentz model using the photon energy E_{ph}:

$$\varepsilon = 1 + \Sigma_j A_j/(E_{ph0j^2} - E_{ph^2} + i\Gamma j E_{ph}) \qquad (17.11)$$

In Eq. (17.11), the dielectric function is described as the sum of different oscillators and the subscript j denotes the jth oscillator. In general, A in Eq. (17.11) is called the oscillator strength.

17.3.3 Tauc-Lorentz and Cody-Lorentz

The Tauc–Lorentz model has been employed to model the dielectric function of amorphous materials ([65]; for a review see [26]) and of transparent conductive oxides [52, 126]. The shape of ε_2 peaks calculated from the Lorentz model is completely symmetric. However the ε_2 peaks of amorphous materials generally show

asymmetric shapes. In the Tauc-Lorentz model [65], therefore, ε_2 is modeled from the product of a unique bandgap of amorphous materials (Tauc gap [134]) and the Lorentz model.

The Tauc gap E_g of amorphous materials is given by the following equation [134]:

$$\varepsilon_2 = A_{tauc}(E_{ph} - E_g)^2/E_{ph}^2 \qquad (17.12)$$

The ε_2 of the Tauc-Lorentz model is expressed by multiplying ε_2 of Eq. (17.11) by Eq. (17.12) [65]:

$$\varepsilon_2 = AE_0\Gamma(E_{ph} - E_g)^2/((E_{ph^2} - E_0^2)^2 + \Gamma^2 E_{ph^2})E_{ph}) \quad (E_{ph} > E_g) \quad (17.13a)$$

$$\varepsilon_2 = 0 \qquad (E_{ph} \leq E_g) \qquad (17.13b)$$

The ε_1 of the Tauc-Lorentz model can be derived by using the Kramers–Kronig relations [53, 65] form Eq. (17.13). Although the equation for ε_1 is rather complicated, the dielectric function of Tauc-Lorentz model is expressed from a total of five parameters $\varepsilon_1 (\infty)$, A, Γ, E_{n0}, Eg. Fig. 17.4 shows (a) the dielectric function and (b) the n-k spectra of an amorphous silicon (a-Si) calculated from the Tauc-Lorentz model [65]. The values of the analytical parameters in this calculation are A = 122 eV, Γ = 2.54 eV, E_{n0} = 3.45 eV, E_g = 1.2 eV and $\varepsilon_1(\infty)$ = 1.15. It can be seen from Fig. 17.4a that $\varepsilon_2 = 0$ at $E_n \leq E_g$ and the ε_2 peak position is given by E_{n0}. The A and Γ of the Tauc-Lorentz model represent the amplitude and half width of the ε_2 peak, respectively, similar to the Lorentz model.

So far, the dielectric function of amorphous materials has also been described using other models including the Cody–Lorentz model [36], Forouhi–Bloomer model [38], MDF theory [2], and band model [74].

The Tauc-Lorentz and Cody-Lorentz [36] dispersion types are primarily designed for modeling amorphous materials. The main difference between the two types is how they model absorption at photon energies slightly larger than the energy gap. In this region the Tauc-Lorentz model follows the Tauc law formula while the Cody-Lorentz follows the Cody formula:

Tauc Absorption Formula: $\varepsilon_2 (E) \approx [(E - Eg)^2/E^2]$
Cody Absorption Formula: $\varepsilon_2(E) \approx (E - Eg)^2$

The Cody-Lorentz type also includes an Urbach absorption term. The fit parameters for the Tauc-Lorentz are the Amplitude (*Amp*), Broadening (*Br*), Center Energy (*Eno*), and Band Gap (*Eg*). The Cody-Lorentz oscillator adds *Ep* (transition energy where absorption changes from Lorentzian to Cody), and *Et* (transition energy where absorption changes from Cody behavior to Urbach behavior). It has been shown that the Cody model provides superior fitting to experimental spectra, compared with the Tauc model [24, 25].
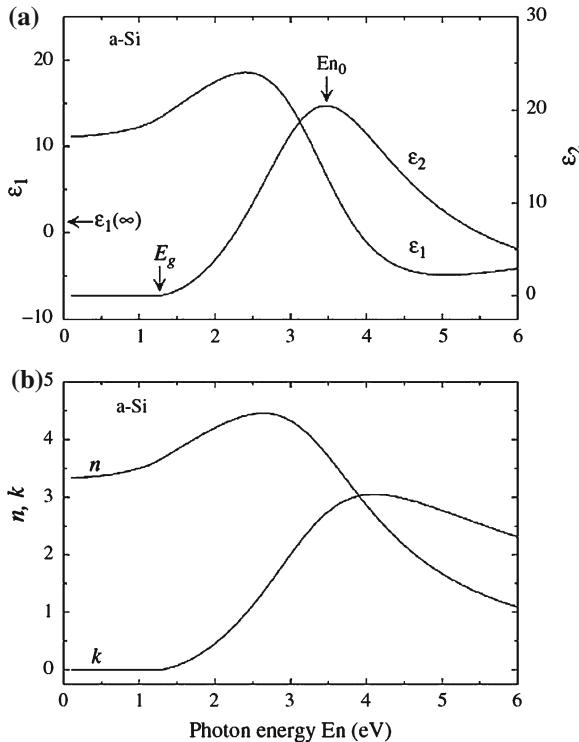
Fig. 17.4 (a) Dielectric function and [36] n k spectra of amorphous silicon (a-Si) calculated from the Tauc–Lorentz model [53]. Reprinted with permission from Wiley, Fujiwara, Spectroscopic Ellipsometry: Principles and Applications. Copyright Wiley, New York, 2007

17.4 Dielectric Function of Poly- and Nanocrystalline Semiconductors

17.4.1 Effective Medium Models

Materials composed of phases much smaller than the radiation wavelength but large enough to retain their bulk properties can generally be modeled using the effective medium approximation (EMA) [12, 20, 72, 132, 133]. Most poly- and nanocrystalline materials fulfill this requirement or at least can properly be modeled using effective medium theories. The success of this approach is shown by the fact that the first publication of the most frequently used self-consistent Bruggeman effective medium approximation (B-EMA) (Ref. [20]) has been cited more than 3,000 times (note that most articles using the B-EMA don't cite the original paper any more, so the influence of this approach is even much greater that the above number of citations would suggest).

Limitations of both increasing (finite wavelength [12, 32, 133]) and decreasing (finite size [35, 98, 105]) component sizes have been extensively studied. Increasing component size leads to scattering- [37], critical dimension- [6–8, 61, 62], patterned wafer- [89] and photonic crystal-based [68, 76, 95, 96] theories and applications, whereas decreasing component size leads to finite-size effects [35, 98], consequently, to the use of fine- grained reference materials [104] or to the parameterization of the dielectric function [1, 60, 94, 105, 127] (also combined with EMA [112]).

Using effective medium models microstructural information can be obtained by the volume fraction of the components, e.g. the amorphous/(nano)crystalline ratio as a function of growth conditions [23, 57, 58, 83, 84, 91, 92], during deposition [26, 56] or annealing [85, 106]. Using both nanocrystalline (nc-Si) and single-crystalline (c-Si) components the nanocrystalline nature (the size of the nanocrystals or the amount of grain boundaries per unit volume) of polycrystalline silicon films can be determined [5, 105]. The same approach can be used for porous silicon [44, 114].

The success of effective medium models is due to their robustness. In most cases, the samples can be fitted with a relatively small number of parameters, avoiding parameter correlations, allowing a quick evaluation. The main disadvantage of the method is that the dielectric function of the components can rarely be described by bulk references. Furthermore, in most cases isotropic behavior, i.e. not oriented grain boundaries are assumed, which might not be right in some cases. Usually the self-consistent Bruggeman method is used, but the type of effective medium model (i.e. Maxwell-Garnett or B-EMA) may depend on the volume fraction ratios.

Being a robust method, EMA is well suited for in situ ellipsometry [18, 26, 69, 85, 98, 106]. This way, the formation of nanocrystals, the transition between amorphous and crystalline structures can be followed, as well as the evolution of nucleation and surface roughness layers.

17.4.2 Analytical Function-Based Models

Analytical models can be used when EMA models fail (e.g. the dielectric functions of the components can not be obtained from bulk or single-phase thin film references), when the dielectric function has to be determined as a smooth function of the wavelength, or when a direct connection with the electronic band properties is investigated determining the change of band gap, stress, or broadening of critical points (a recent, sophisticated example on CdTe photovoltaics is given in Ref. [75]). The analytical functions used for polycrystalline materials can be derived from the formulas described above for single-crystalline semiconductors.

The Lorentz oscillator is widely used when the EMA models don't work well enough [33] or simply to determine a smooth dielectric function that can further be analyzed using EMA [35]. In this references ion implantation has been used as a way to create damage and a vanishing long range order in a controlled way. Ion implantation is

a versatile and effective method to prepare reference samples for the investigation of the dielectric function of various disordered lattice structures [33, 35, 39–43, 45, 46, 78–80, 94, 104, 107–109, 111, 112, 120, 127].

The change of characteristic critical point features can be analyzed using the (second or third) derivative method [13, 116]. In this approach one assumes that the crystal-related structure is directly connected to the lineshape of the critical point features. The vanishing long range order (e.g. as a result of ion implantation-caused disorder) causes the sharp critical point features to broaden, which can be characterized by second or third derivative amplitudes of the pseudo-dielectric function. The change of the derivative amplitudes as a function of fluence can be connected to the size of the single ion tracks. Neglecting a possible surface layer like oxide or nanoroughness doesn't cause a significant error in the derivatives, because these layers are practically non-dispersive in the spectral range around the critical point energies.

The generalized oscillator model uses standard analytical line shapes for the interband critical point features [11, 22, 28, 73, 117]:

$$\varepsilon(E) = A e^{i\varphi} (E_{CP} - E - i\Gamma)^{\mu}, \text{ (genosc)}$$

where A, E_{CP}, Γ, and ϕ are the amplitude, the critical point energy, the broadening, and the excitonic phase angle, respectively. Using this equation critical points of different dimensionalities can be described by adjusting the value of μ. This method allows the analysis of most important critical point features like the energy positions or the broadenings as a function of temperature [73], disorder [117] or other process parameters.

The dielectric function of numerous semiconductors (e.g. Si, Ge, CdS, and CdTe) have been parameterized by Adachi using the Kramers-Kronig transformation. The model dielectric function (MDF) calculated using this method has the advantage that it provides a good fit using a set of dedicated oscillators with a few parameters for each critical point [1]. This method has been successfully applied for disordered Si [3, 71, 127], for III–V semiconductors [34] as well as for nanocrystals in porous Si [118]. Similar parameterization can be applied for ferroelectric and dielectric layers [47], but most importantly, it has been demonstrated that it can be applied for nanocrystals in silicon nitride as well [17]. The fit of the MDF on a fine-grained polycrystalline silicon reference is shown in Fig. 17.5.

When studying nanocrystalline semiconductors, the dielectric functions range from the nearly single-crystalline to the nearly amorphous. In the latter case the parametrizations are simpler, because the critical point features disappear from the spectrum, having only one broad peak. The most successful phenomenological model is the Tauc-Lorentz (TL) model suggested by Jellison, which combines a Tauc gap with a Lorentzian oscillator [65]. An important feature of the model is that a Kramers-Kronig consistent analytical formula is given for ε_1, which makes the calculation of the model accurate and fast. This model has later been refined through a more accurate description of the line shape at the gap energy using the Cody parameterization,
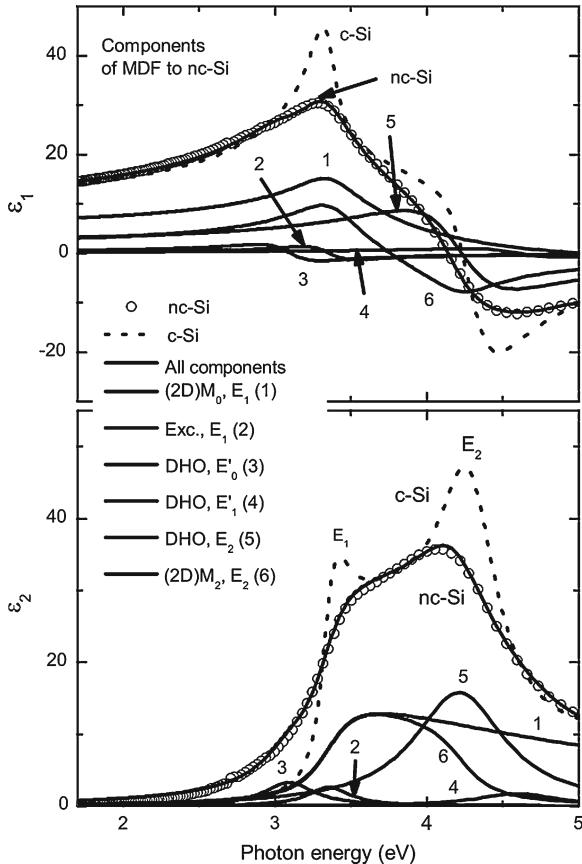
Fig. 17.5 Lineshapes of the oscillators of the model dielectric function fitted on a nanocrystalline polysilicon reference from Ref. [64]. Reprinted with permission from Petrik et al. [118]. Copyright 2009, American Institute of Physics

resulting in the Cody-Lorentz model [36]. The TL models have successfully been applied for extremely small embedded crystals, carbon based thin films [21, 77, 113], nanocrystalline diamond [82] as well as for amorphous SiC [81].

Finally, a general approach that provides the largest flexibility to describe the critical point features was suggested by Johs et al. [66]. A detailed analysis of the method termed 'Generalized critical point model' can be found in Ref. [28]. In this approach the line shape of the dielectric function around the critical points are composed of four Gaussian-broadened polynomials. The fit parameters are the so called control points of the polynomials, their position, curvature, connection and the broadening energy, and amplitude values of the critical point. Because of the large number of fit parameters this model can only be used in a very careful and systematic way to fit each critical point. For nanocrystalline materials coupling or fixing the most parameters may be needed to avoid cross correlations [119].

17.4.3 Quantum Confinement

Several papers studied silicon nanocrystals embedded in insulator matrix by theoretical calculations [101, 125] or experimentally [30, 31, 55, 99]. Ögüt et al. [101] calculated quasiparticle gaps, self-energy corrections, exciton Coulomb energies, and optical gaps in Si quantum dots from first principles using a real-space pseudopotential method. The calculations were performed on hydrogen-passivated spherical Si clusters with diameters up to 2.7 nm (800 Si and H atoms). They showed that (i) the self-energy correction in quantum dots is enhanced substantially compared to bulk, and is not size independent as implicitly assumed in all semiempirical calculations, and (ii) quantum confinement and reduced electronic screening result in appreciable excitonic Coulomb energies. They fitted the calculated data to a power law of the diameter as d^{-a}, and found $a = 0.7$. The calculated optical gaps were in very good agreement with absorption data [55].

Proot et al. [125] calculated the electronic structure of spherical silicon crystallites containing up to 2058 Si atoms. They predicted a variation of the optical band gap with respect to the size of the crystallites in very good agreement with available experimental results [55]. They also calculated the electron-hole recombination time which is of the order of 10^{-4}–10^{-6} s for crystallites with diameters of 2.0–3.0 nm. Their results were applied to porous silicon for which they confirmed that a possible origin of the luminescence is the quantum confinement. The calculated optical gap varies from 5 to 1.6 eV (size range from 0.8 to 4.3 nm) following approximately a $d^{-1.39}$ law where d is the crystallite diameter. (From a simple effective mass approximation one would expect a d^{-2} law. The lower exponent 1.39 shows that the exact nature of the bands has to be taken into account, in particular for the complex conduction band.)

Ding et al. [30, 31] studied Si nanocrystals (nc-Si) with different sizes embedded in SiO2 matrix having been synthesized with various recipes of Si ion implantation. The influence of nanocrystal size on optical properties, including dielectric functions and optical constants, of the nc-Si has been investigated with spectroscopic ellipsometry. The optical properties of the nc-Si are found to be well described by the four-term Forouhi-Bloomer model [38]. A strong dependence of the dielectric functions and optical constants on the nc-Si size is observed. For the imaginary part of the dielectric functions, the magnitude of the main peaks at the transition energies E_1 and E_2 exhibits a large reduction and a significant redshift in E_2 depending on the nc-Si size. A band gap expansion is observed when the nc-Si size is reduced. The band gap expansion with the reduction of nc-Si size is in good agreement with the prediction of first-principles calculations based on quantum confinement [101]. They fit their results with the $E_g(D) = E_{g0} + C/D^n$ formula, where D is the nanocrystal size in nm, $Eg(D)$ is the band gap in eV of the nanocrystal, $E_{g0} = 1.12$ eV is the band gap of bulk crystalline Si, $C = 3.9$, and $n = 1.22$.

17.5 Modeling of Layer Structure and Composition

17.5.1 Single layers

Even when investigating single layers, in most cases not only the bulk layer is included in the model, but the surface roughness [10, 70] and the nucleation layer [69] is also taken into account. The systematic improvement of the fit quality during the introduction of interface layers into the optical model has been investigated by numerous authors (see Refs. [10, 44, 121, 136], just to mention a few examples). The formation of the nucleation layer and the surface roughness has also been extensively investigated using in situ spectroscopic ellipsometry [18, 19, 26, 51, 69].

The dielectric function of the bulk layer is modeled as described in the previous chapter. When using single-layer models either the sample is close to perfect (e.g. prepared by molecular beam epitaxy, atomic layer deposition or other methods that can create nearly perfect layers), or the non-ideal interfaces are not taken into account. In the latter case the error caused by not considering the interfaces will be transferred to the determined layer parameters as errors (to both the layer thickness or the dielectric function). The modeling of the interfaces leads us to the following chapters on vertically inhomogeneous structures.

17.5.2 Ultrathin nc-Si RTSE Measurements

Nguyen et al. [98–100] studied the evolution of thin film Si nanostructure during growing and etching processes that yields thin layer nanocrystallites. They performed real-time spectroscopic ellipsometry (RTSE) measurements during the growth of nanocrystalline silicon (nc-Si:H) or a-Si:H nanoclusters by plasma enhanced chemical-vapor deposition on chromium at $250\,°C$. They focused on the regime when the film consists of isolated nanocrystallites and intervening void volume. In this regime, the observed three-dimensional growth behavior allowed to associate the crystallite size with the physical thickness of the film. The RTSE measurements were self-contained in that they provided not only microstructural information, including film thickness and volume fraction, but also the effective optical functions of the film. From this combination of results, the optical functions the Si crystallites, themselves, could be deduced by mathematically extracting the influence of the void-volume fraction on the effective optical functions. A critical-point (CP) analysis of E_1 transitions visible near 3.3 eV in the crystallite optical functions provided information on the electronic properties as a continuous function of crystallite size. Over the physical thickness range in this experiments (max. 25 nm), the transition energy and phase deduced in the CP analysis were constant (at the single-crystal values), while the optical gap and broadening parameter decreased with increasing thickness. (See Fig. 17.6) This behavior was consistent with a finite-size effect (quantum confinement).
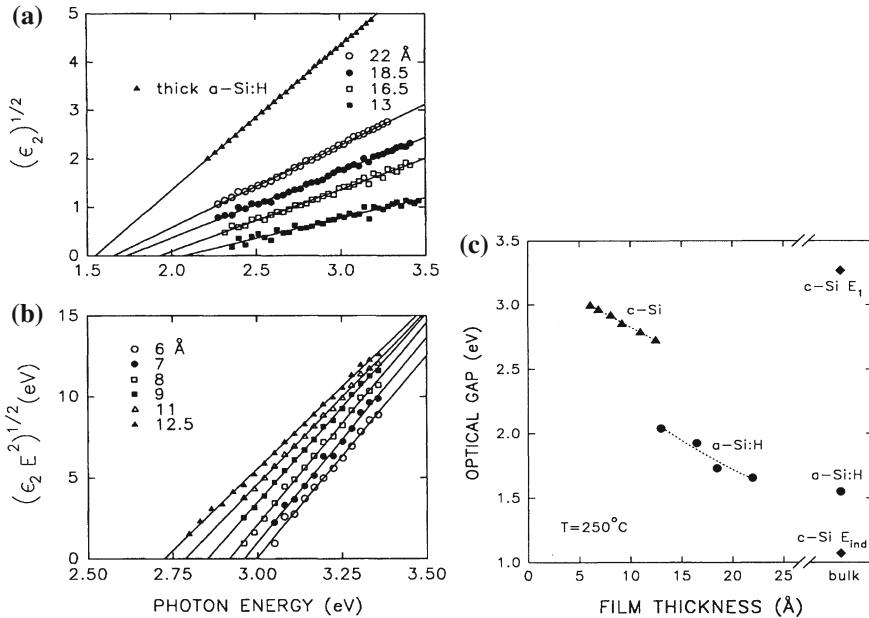
Fig. 17.6 Optical gap determinations for ultrathin **(a)** a-Si:H and **(b)** nc-Si films at 250 °C deduced from RTSE measurements. The optical gaps are obtained by extrapolating observed linear plots to $\varepsilon_2 = 0$; **c** optical gaps plotted versus thickness for ultrathin a-Si:H and nc-Si films at 250 °C, from **a** and **b**. Values for thick a-Si:H and bulk c-Si are provided at the *right*. Reprinted with permission from Nguyen et al. [100]. Copyright 1995, The American Physical Society

The etching [99] was performed with thermally generated atomic H in order to avoid plasma damage. RTSE was applied to characterize the evolution of the thicknesses of near-surface nc-Si:H and underlying a-Si:H layers, as well as their optical properties. Using the end-point detection capability of RTSE, they could terminate etching to obtain an ultrathin (<1.5 nm), single-phase layer of Si nanocrystallites ∼i. e., with no detectable a-Si:H. These crystallites were densely packed on the substrate compared to those of conventional, single-phase nc-Si:H prepared by PECVD [98].

17.5.3 Superlattices

In case of superlattices one should assume a periodic layer structure [129]. This can simplify the optical model because the properties (mainly thicknesses) of the layers within one period only have to be defined once, providing only the number of periods as a required additional parameter [4]. Equivalently, in case of a fewer number of layers it can be considered as a coupling of the repeated parameters [128]. Even if the preparation conditions suggest a periodic structure, the structures are usually
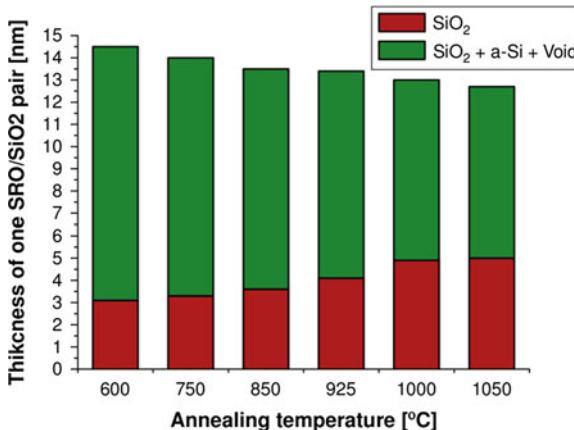
Fig. 17.7 Evolution of sublayer thicknesses in a layer stack of $10\times$([silicon rich oxide]/SiO2). Silicon rich oxide (SRO) was modeled using an effective medium mixture of SiO_2, a-Si, and voids. The graph shows the increase of the SiO_2 sublayer as a function of the annealing temperature. Reprinted with permission from Agocs et al. [4]. Copyright 2011 Elsevier
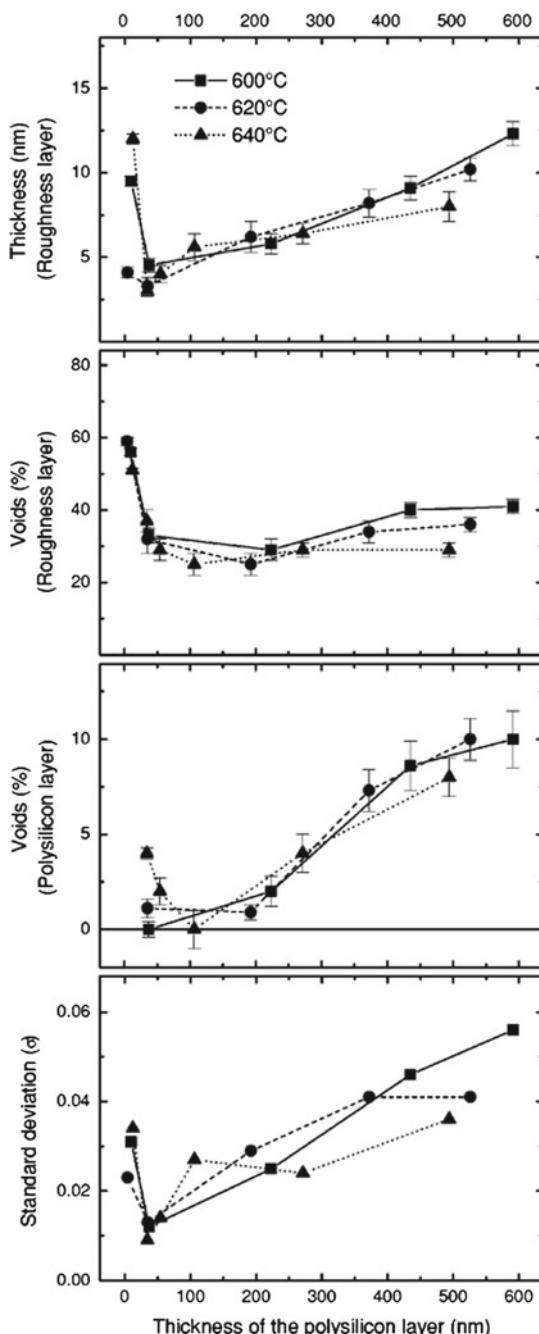
not ideal on the scale of sensitivity of ellipsometry. First, the surface layer usually has to be fitted independently because its contact with air modifies its properties (due to oxidation [4] or hydrocarbon contamination [131]). Second, the layer quality usually deteriorates with each new prepared layer [90] (caused by e.g. the interface roughness), which results in a vertical inhomogeneity over the whole stack. Of course, the ideality of the superlattice can be estimated from the fit quality. As sublayers in superlattice stacks are usually very thin deposited layers, bulk-like references can be used in very rare cases, and the dielectric function of each stack has to be fitted as well. Fitting the dielectric function even structural changes can be followed during e.g. oxidation (Fig. 17.7 from Ref. [4]).

17.5.4 Inhomogeneity in Vertical and Lateral Dimensions

Most methods of thin film preparation result in vertically inhomogeneous layers, at least on the scale of the precision of ellipsometry. An ideal thin film for ellipsometry would have atomically sharp interfaces and a vertical refractive index uniformity of better than about 10^{-4}. These requirements can rarely be met. Usually, at least an interface layer of nucleation [28] and a surface roughness layer have to be taken into account [53]. For example, in case of polysilicon, most layer properties depend strongly on the layer thickness (Fig. 17.8, [105]).

Neither the amorphous layers created by ion implantation have sharp interfaces. However, in this case the shape of the profile can be parameterized based on the theory of ion implantation. This method reduces the number of fit parameters significantly,

Fig. 17.8 Selected relevant fit parameters of an optical model as a function of thickness and deposition temperature used to measure low pressure chemically vapor deposited polycrystalline silicon. The optical model consists of a surface roughness layer, a polysilicon bulk layer, and a buried oxide layer (both the surface roughness layer and the polysilicon bulk layer are a Bruggeman effective medium composition of single-crystalline Si, fine-grained polycrystalline Si and voids). Reprinted with permission from Petrik et al. [105]. Copyright 2000, American Institute of Physics
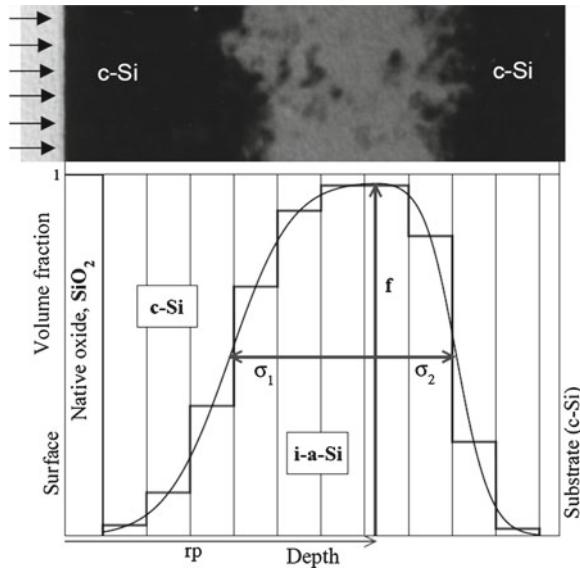
Fig. 17.9 Optical model for the ellipsometric measurement of the Gaussian damage profile created by 100 keV Ar ions implanted into single-crystal Si. σ_1 and σ_2 are the standard deviations of the coupled half-Gaussian distribution functions. R_p and f are the peak position and the height parameter of the damage profile, respectively. The fifth parameter of the optical model is the thickness of the surface oxide layer (Reprinted with permission from Petrik et al. [115]. Copyright 2008, John Wiley and Sons.)

because only four key parameters describing the Gaussian profile need to be fitted [42, 107, 115] (Fig. 17.9).

Depth profiles of voids in Si created for gettering at a depth as large as 400 nm can sensitively be determined using spectroscopic ellipsometry [54, 111]. Although the penetration depth of light is significantly less than 400 nm in a significant part of the spectrum (at the critical point energies and even below, down to about 2 eV), a high sensitivity on the void profiles has been demonstrated because of the large optical contrast between the dielectric function of Si and voids. In Ref. [54] a sophisticated parameterization of the depth profiles using a combination of normalized error functions is suggested and successfully demonstrated.

Because the penetration depth of light at the most sensitive parts of the spectrum (i.e. at the photon energies of the critical points—in Si about 3.4 and 4.2 eV) is only in the range of 5–10 nm, high sensitivity structural characterization using ellipsometry taking into account the line shape of the dielectric function at the critical points cannot be performed for deep implanted profiles. However, using wedge etching [16] buried profiles can be brought to the sample surface allowing sensitive characterizations [48, 59]. The bevels are made after ion implantation by pulling the sample out gradually from the anodization liquid, growing an oxide with laterally increasing thickness. In the experiment of Ref. [59] the wedge length and maximum height were 50 mm and
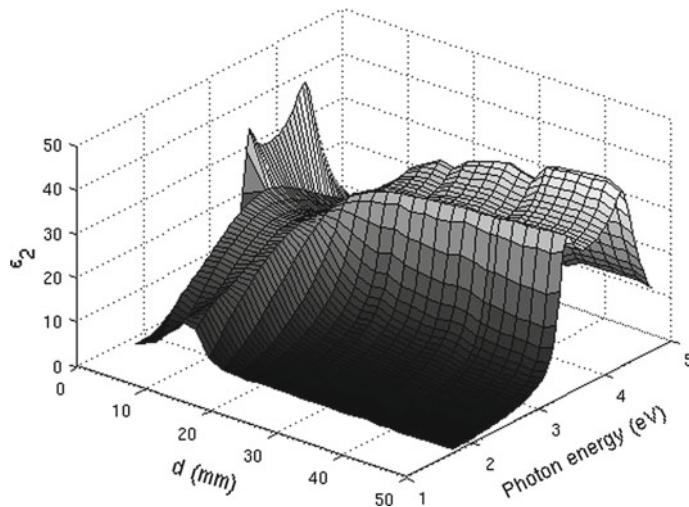
Fig. 17.10 Imaginary part of the pseudo-dielectric function of Si implanted with 100 keV Xe through a wedge mask created by anodic oxidation [59]. *d* denotes the lateral position on the sample. The oxide thickness decreased from 180 to 0 nm from $d = 50$ to $d = 5$ mm. The oxide was removed after ion implantation, prior to the ellipsometric measurement
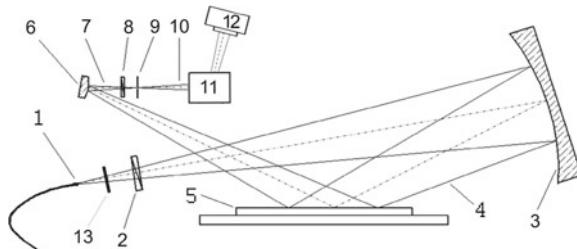
Fig. 17.11 Side view of the optical arrangement of the divergent light source mapping spectroscopic ellipsometer. (*1*) point source; (*2*) polarizer; (*3*) spherical mirror; (*4*) non collimated beam; (*5*) sample; (*6*) cylindrical mirror; (*7*) corrected beam; (*8*) analyzer; (*9*) pinhole; (*10*) beam after pinhole; (*11*) corrector-disperser optics; (*12*) ccd detector; (*13*) rectangular (*narrow*) aperture

200 nm, respectively. Finally, the oxide is removed by conventional HF etch. It has to be taken into account that the growth of a unit of SiO_2 consumes only about 0.45 unit of Si. So in the above case after etching the 200 nm SiO_2 the thickness of the removed Si is about 90 nm, resulting in a slope of about 1×10^{-4} degree, a slope so small that is impossible to produce by other techniques. Using this technique even the plotting of the pseudo-dielectric function as a function of lateral position shows clearly the locations where the damage peak reaches the surface [115].

The second or third derivative analysis described in chapter "Analytical function-based models" combined with Rutherford backscattering spectrometry (RBS) and numerical track modeling can even be used to measure the 3D evolution of the ion

tracks and the remaining nanocrystalline regions during ion implantation [116]. The broadening of the critical point features measured by ellipsometry as a function of fluence is proportional to the remaining non-damaged surface and therefore shows an exponential dependence on the fluence. With RBS we can measure the relative damage as a function of depth. In the numerical simulation we can define a vertical distribution of damage and a track size. The dependence of relative damage on the fluence and depth (obtained from the RBS and SE measurements) can be fitted using the parameters of the numerical model, so the parameters like the track size and the depth distribution can be obtained [116].

The complex optical models used for depth profiling often require the fitting of numerous model parameters. In this case a crucial question is whether the result is the global minimum? The final step in the fitting procedure is usually a Levenberg-Marquardt gradient search algorithm that finds the minimum based on the change of fitting error when changing a parameter locally. Therefore, this method can not guarantee to find the global minimum. There is a range of techniques developed to find the global minimum (random global search, genetic algorithms, neural networks, simulated annealing, hill climbing, etc., see Refs. [122–124]). Even a simple global search and a gradient fit from the best sets of parameters found by the global search has been proven to be an effective way of avoiding to get into local minima [107].

Lateral mapping is usually performed by moving the samples stage or (in case of larger substrates) moving the ellipsometer head over the sample [11]. Imaging ellipsometry provides a concept of measuring in different lateral positions simultaneously. Both microscopic [15, 137, 138] and macroscopic [49, 50, 67, 87, 88] concepts have been demonstrated. The size of the measured area and the resolution ranges from several microns to several millimeters. For divergent light source non-collimated beam macroscopic ellipsometric configurations there is theoretically no limit of the maximum size—the resolution can be defined as the viewing angle of a pixel on the CCD camera. The concept is shown in Fig. 17.11 [49] demonstrating a new design, in which all the polarizing optical components can be small—the only component that scales with the mapping size is the spherical mirror.

References

1. S. Adachi, Phys. Rev. B **38**, 12966 (1988)
2. S. Adachi, *Optical Properties of Crystalline and Amorphous Semiconductors: Materials and Fundamental Principles* (Kluwer Academic Publishers, Norwell, 1999)
3. S. Adachi, H. Mori, Phys. Rev. B **62**, 10158 (2000)
4. E. Agocs, P. Petrik, S. Milita, L. Vanzetti, S. Gardelis, A.G. Nassiopoulou, G. Pucker, R. Balboni, M. Fried, Thin Solid Films **519**, 3002 (2011)
5. E. Agocs, P. Petrik, M. Fried, A.G. Nassiopoulou, Mater. Res. Soc. Symp. Proc. **1321**, 367–372 (2011). doi:10.1557/opl.2011.949
6. R. Antos, I. Ohlidal, D. Franta, P. Klapetek, J. Mistrik, T. Yamaguchi, S. Visnovsky, Appl. Surf. Sci. **244**, 221 (2005)

7. R. Antos, J. Pistora, I. Ohlidal, K. Postava, J. Mistrik, T. Yamaguchi, S. Visnovsky, M. Horie, J. Appl. Phys. **97**, 053107 (2005)
8. R. Antos, J. Pistora, J. Mistrik, T. Yamaguchi, S. Yamaguchi, M. Horie, S. Visnovsky, Y. Otani, J. Appl. Phys. **100**, 054906 (2006)
9. H. Arwin, D.E. Aspnes, J. Vac. Sci. Technol. A **2**, 1316 (1984)
10. D.E. Aspnes, J.B. Theeten, F. Hottier, Phys. Rev. B **20**, 3292 (1979)
11. D.E. Aspnes, in *Handbook on Semiconductors*, vol. 2, ed. by M. Balkanski (North-Holland, Amsterdam, 1980), p. 109
12. D.E. Aspnes, Thin Solid Films **89**, 249 (1982)
13. D.E. Aspnes, S.M. Kelso, C.G. Olson, D.W. Lynch, Phys. Rev. Lett. **48**, 1863 (1982)
14. D. E. Aspnes, in *Handbook of Optical Constants of Solids*, ed. by E. D. Palik (Academic, New York, 1985)
15. Y.M. Bae, B-K Oh, W. Lee, W.H. Lee, J-W Choi, Biosensors and Bioelectronics **20**, 895 (2004)
16. A. Balazs, L. Hermann, J. Gyulai, Phys. Stat. Sol. (a) **29**, 105 (1975)
17. P. Basa, P. Petrik, M. Fried, L. Dobos, B. Pécz, L. Tóth, Physica E-Low-Dimens. Syst. Nanostruct. **38**, 76 (2007)
18. A. Bonanni, D. Stifter, A. Montaigne-Ramil, K. Schmidegg, K. Hingerl, H. Sitter, J. Cryst. Growth **248**, 211 (2003)
19. A. Bonanni, K. Schmidegg, A. Montaigne-Ramil, H. Sitter, K. Hingerl, D. Stifter, J. Vac. Sci. Technol. B **21**, 1825 (2003)
20. D.A.G. Bruggeman, Ann. Phys. (Leipzig) **24**, 636 (1935)
21. J. Budai, Z. Toth, A. Juhasz, G. Szakacs, E. Szilagyi, M. Veres, M. Koos, J. Appl. Phys. **100**, 043501 (2006)
22. M. Cardona, *Modulation Spectroscopy*, Suppl. 11 of Solid State Physics, ed. by F. Seitz, D. Turnbull, H. Ehrenreich (Academic, New York, 1969)
23. M.F. Cerqueira, M. Stepikhova, M. Losurdo, M.M. Giangregorio, A. Kozanecki, T. Monteiro, Opt. Mater. **28**, 836 (2006)
24. G.D. Cody, B.D. Brooks, B. Abeles, Sol. Energy Mat. **8**, 231–240 (1982)
25. R.W. Collins, K. Vedam, Optical properties of solids, in *Encyclopedia of Applied Physics*, vol. 12, ed. by G.L. Trigg (VCH, New York, 1995), p. 285
26. R.W. Collins, Joohyun Koh, H. Fujiwara, P.I. Rovira, A.S. Ferlauto, J.A. Zapien. C.R. Wronski, R. Messier, Appl. Surf. Sci. **154–155**, 217–228 (2000)
27. R.W. Collins, Joohyun Koh, A.S. Ferlauto, P.I. Rovira, Yeeheng Lee, R.J. Koval, C.R. Wronski. Thin Solid Films **364**, 129 (2000)
28. R.W. Collins, A.S. Ferlauto, in *Handbook of Ellipsometry* (William Andrew, Norwich, 2005), p. 93
29. P. Aryal, J. Chen, Z. Huang, L.R. Dahal, M.N. Sestak, D. Attygalle, R. Jacobs, V. Ranjan, S. Marsillac, R.W. Collins, *37th IEEE Photovoltaic Specialists Conference (PVSC 2011)*, Seattle, WA;19–24 June 2011; Category number CFP11PSC-ART; Code 89752, Article number 6186402, pp. 002241–002246
30. L. Ding, T.P. Chen, Y. Liu, C.Y. Ng, S. Fung, Phys. Rev. B **72**, 125419 (2005)
31. L. Ding, T.P. Chen, Y. Liu, M. Yang, J.I. Wong, Y.C. Liu, A.D. Trigg, F. R. Zhu M.C. Tan, S. Fung, J. Appl. Phys. **101**, 103525 (2007)
32. W.G. Egan, D.E. Aspnes, Phys. Rev. B **26**, 5313 (1982)
33. W.G. Egan, D.E. Aspnes, Phys. Rev. B **26**, 5313 (1982)
34. Aleksandra B. Djurisic, Aleksandar D. Rakic, Paul C. K. Kwok, E. Herbert Li, Martin L. Majewski, J. Appl. Phys. **85**, 3638 (1999)
35. G.F. Feng, R. Zallen, Phys. Rev. B **40**, 1064 (1989)
36. A.S. Ferlauto, G.M. Ferreira, J.M. Pearce, C.R. Wronski, R.W. Collins, X. Deng, G. Ganguly, J. Appl. Phys. **92**, 2424 (2002)
37. D. Franta, I. Ohlidal, Opt. Commun. **248**, 459 (2005)
38. A.R. Forouhi, I. Bloomer, Phys. Rev. B **34**, 7018–7026 (1986)

39. M. Fried, T. Lohner, E. Jároli, G. Vizkelethy, G. Mezey, J. Gyulai, M. Somogyi, H. Kerkow, Thin Solid Films **116**, 191 (1984)
40. M. Fried, T. Lohner, J.M.M. de Nijs, A. van Silfhout, L.J. Hanekamp, Z. Laczik, N.Q. Khanh, J. Gyulai, J. Appl. Phys. **66**, 5052 (1989)
41. M. Fried, T. Lohner, W.A.M. Aarnink, L.J. Hanekamp, A. van Silfhout, J. Appl. Phys. **71**, 5260 (1992)
42. M. Fried, T. Lohner, W.A.M. Aarnink, L.J. Hanekamp, A. van Silfhout, J. Appl. Phys. **71**, 2835 (1992)
43. M. Fried, A. van Silfhout, Phys. Rev. **B49**, 5699 (1994)
44. M. Fried, T. Lohner, O. Polgar, P. Petrik, E. Vazsonyi, I. Barsony J.P. Piel, J.-L. Stehle, Thin Solid Films **276**, 2223 (1996)
45. M. Fried, T. Lohner, P. Petrik, *Ellipsometric Characterization of Thin Films in Handbook of Surfaces and Interfaces of Materials: "Solid Thin Films and Layers"*, vol. 4, ed. by H. S. Nalwa(Academic Press, San Diego, 2001) pp. 335–367
46. M. Fried, P. Petrik, T. Lohner, N.Q. Khánh, O. Polgár, J. Gyulai, Thin Solid Films **455–456**, 404 (2004)
47. M. Fried, P. Petrik, Zs. E. Horváth, T. Lohner. Appl. Surf. Sci. **253**, 349 (2006)
48. M. Fried, N.Q. Khan, P. Petrik, Physica Status Solidi C-Curr. Top. Solid State Phys. **5**, 1227 (2008)
49. M. Fried, G. Juhász, C. Major, P. Petrik, O. Polgár, Z. Horváth, A. Nutsch, Thin Solid Films **519**, 2730 (2011)
50. M. Fried, G. Juhasz, C. Major, A. Nemeth, P. Petrik, O. Polgar, C. Salupo, Lila R. Dahal, R. W. Collins, Mater. Res. Soc. Symp. Proc. 1323 (2011) doi:10.1557/opl.2011.820
51. H. Fujiwara, J. Koh, C.R. Wronski, R. W. Collins. Appl. Phys. Lett **70**, 2151 (1997)
52. H. Fujiwara, M. Kondo, Phys. Rev. B **71** 075109–1-10 (2005)
53. H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (Wiley, New York, 2007)
54. W. Fukarek, J.R. Kaschny, J. Appl. Phys. **86**, 4160 (1999)
55. S. Furukawa, T. Miyasato, Phys. Rev. B **38**, 5726 (1988)
56. M.M. Giangregorio, M. Losurdo, A. Sacchetti, P. Capezzuto, G. Bruno, Thin Solid Films **511–512**, 598 (2006)
57. M.M. Giangregorio, M. Losurdo, A. Sacchetti, P. Capezzuto, F. Giorgis, G. Bruno, Appl. Surf. Sci. **253**, 287 (2006)
58. M.M. Giangregorio, M. Losurdo, G.V. Bianco, P. Capezzuto, G. Bruno, Thin Solid Films **519**, 2787 (2011)
59. J. Gyulai, G. Battistig, T. Lohner, Z. Hajnal, Nucl. Instrum. Methods Phys. Res. B **266**, 1434 (2008)
60. C.M. Herzinger, B. Johs, W.A. McGahan, J.A. Woollam, W. Paulson, J. Appl. Phys. **83**, 3323 (1998)
61. H.-T. Huang, W. Kong, F.L. Terry, Jr., Appl. Phys. Lett. **78**, 3983 (2001)
62. H.-T. Huang, F.L. Terry Jr., Thin Solid Films **455/456**, 828 (2004)
63. G.E. Jellison, Opt. Mater. **1**, 41 (1992)
64. G.E. Jellison Jr, M.F. Chisholm, S.M. Gorbatkin, Appl. Phys. Lett. **62**, 3348 (1993)
65. G.E. Jellison, F.A. Modine, Appl. Phys. Lett. **69**, 371 (1996)
66. B. Johs, C.M. Herzinger, J.H. Dinan, A. Cornfeld, J.D. Benson, Thin Solid Films **313–314**, 137 (1998)
67. G. Juhasz, Z. Horvath, C. Major, P. Petrik, O. Polgar, M. Fried, Phys. Status Solidi C-Curr. Top. Solid State Phys. **5**, 1081 (2008)
68. K. Kertész, Z. Bálint, Z. Vértesy, G. Márk, V. Lousse, J.P. Vigneron, M. Rassart, L.P. Biró, Phys. Rev. E **74**, 021922 (2006)
69. J. Koh, Y. Lu , S. Kim, J.S. Burnham, C.R. Wronski, R.W. Collins, Appl. Phys. Lett. **67**, 2669 (1995)
70. J. Koh, Y. Lu , C.R. Wronski, Y. Kuang, R.W. Collins, T.T Tsong, Y.E. Strausser, Appl. Phys. Lett. **69**, 1297 (1996)

71. K. Kurihara, S. Hikino, S. Adachi, J. Appl. Phys. **96**, 3247 (2004)
72. W. Lamb, D.M. Wood, N.W. Ashcroft, Phys. Rev. B **21**, 2248 (1980)
73. P. Lautenschlager, M. Garriga, M. Cardona, Phys. Rev. B **36**, 4813 (1987)
74. J. Leng, J. Opsal, H. Chu, M. Senko, D.E. Aspnes, Thin Solid Films **313–314**, 132–136 (1998)
75. J. Li, J. Chen, M.N. Sestak, R.W. Collins, IEEE J. Photovolt. **1**, 187 (2011)
76. C.-H. Lin, H.-L. Chen, W.-C. Chao, C.-I. Hsieh, W.-H. Chang, Microelectron. Eng. **83**, 1798 (2006)
77. S. Logothetidis, Diam. Relat. Mater. **12**, 141 (2003)
78. T. Lohner, G. Mezey, E. Kótai, F. Pászti, L. Királyhidi, G. Vályi, J. Gyulai, Nucl. Instrum. Methods **182/183**, 591 (1981)
79. T. Lohner, G. Mezey, E. Kótai, F Pászti, A. Manuaba, J. Gyulai, Nucl. Instrum. Methods **209/210**, 615 (1983)
80. T. Lohner, M. Fried, J. Gyulai, K. Vedam, N. Nguyen, L.J. Hanekamp, A. van Silfhout, Thin Solid Films **233**, 117 (1993)
81. T. Lohner, Z. Zolnai, P. Petrik, G. Battistig, J.G. Lopez, Y. Morilla, A. Koos, Z. Osvath, M. Fried, Phys. Status Solidi C-Curr. Top. Solid State Phys. No. **5**, 1374 (2008)
82. T. Lohner, P. Csíkvári, N.Q. Khánh, S. Dávid, Z.E. Horváth, P. Petrik, G.Hárs, Thin Solid Films **519**, 2806 (2011)
83. M. Losurdo, G. Bruno, D. Barreca, E. Tondello, Appl. Phys. Lett. **77**, 1129 (2000)
84. M. Losurdo, M.F. Cerqueira, M.V. Stepikhova, E. Alves, M.M. Giangregorio, P. Pinto, J.A. Ferreira, Physica B **308–3010**, 374 (2001)
85. M. Losurdo, F. Roca, R. De Rosa, P. Capezzuto, G. Bruno, Thin Solid Films **383**, 69 (2001)
86. M. Losurdo, M. Giangregorio, P. Capezzuto, G. Bruno, R. De Rosa, F. Roca, C. Summonte, J. Pla, R. Rizzoli, J. Vac. Sci. Technol. A **20**, 37 (2002)
87. C. Major, G. Juhász, Z. Horvath, O. Polgar, M. Fried, I. Bársony, Physica Status Solidi C-Curr. Top. Solid State Phys. **5**, 1077 (2008)
88. C. Major, G. Juhasz, P. Petrik, Z. Horvath, O. Polgar, M. Fried, Vacuum **84**, 119 (2010)
89. H.L. Maynard, N. Layadi, J.T.C. Lee, Thin Solid Films **313/314**, 398 (1998)
90. M. Modreanu, M. Gartner, D. Cristea, Mater. Sci. Eng. C **19**, 225 (2002)
91. M. Modreanu, M. Gartner, E. Aperathitis, N. Tomozeiu, M. Androulidaki, D. Cristea, Paul Hurley, Physica E **16**, 461 (2003)
92. M. Modreanu, M. Gartner, C. Cobianu, B. O'Looney, F. Murphy, Thin Solid Films **450**, 105 (2004)
93. I. Mohacsi, P. Petrik, M. Fried, T. Lohner, J.A. van den Berg, M.A. Reading, D. Giubertoni, M. Barozzi, A. Parisini, Thin Solid Films **519**, 2847 (2011)
94. H. Mori, S. Adachi, M. Takahashi, J. Appl. Phys. **90**, 87 (2001)
95. N. Nagy, A.E. Pap, E. Horváth, J. Volk, I. Bársony, A. Deák, Z. Hórvölgyi, Appl. Phys. Lett. **89**, 063104 (2006)
96. N. Nagy, A. Deák, Z. Hórvölgyi, M. Fried, A. Agod, I. Bársony, Langmuir **22**, 8416 (2006)
97. Public community-driven database on the Nanocharm, http://www.nanocharm.org/index.php?option=com_content\&task=section\&id=15\&Itemid=122
98. H. V. Nguyen, R.W. Collins, Phys. Rev. B **47**, 1911 (1993)
99. H.V. Nguyen, Ilsin An, R.W. Collins, Yiwei Lu, M. Wakagi, C.R. Wronski, Appl. Phys. Lett. **65**, 3335 (1994)
100. H.V. Nguyen, Y. Lu, S. Kim, M. Wakagi, R.W. Collins, Phys. Rev. Lett. **74**(19), 3880–3883 (1995)
101. S. Ögüt, J.R. Chelikowsky, S.G. Louie, Phys. Rev. Lett. **79**, 1770–1773 (1997)
102. E.D. Palik (ed.), *Handbook of Optical Constants of Solids II* (Academic, New York, 1991)
103. P. Petrik, M. Fried, T. Lohner, R. Berger, L. P. Biró, C. Schneider, J. Gyulai, H. Ryssel, Thin Solid Films **313/314**, 259 (1998)
104. P. Petrik, T. Lohner, M. Fried, N.Q. Khánh, O. Polgár, J. Gyulai, Nucl. Instr. Methods B **147**, 84 (1999)
105. P. Petrik, T. Lohner, M. Fried, L.P. Biró, N.Q. Khánh, J. Gyulai, W. Lehnert, C. Schneider, H. Ryssel, J. Appl. Phys. **87**, 1734 (2000)

106. P. Petrik, W. Lehnert, C. Schneider, T. Lohner, M. Fried, J. Gyulai, H. Ryssel, Thin Solid Films **383**, 235 (2001)
107. P. Petrik, O. Polgár, M. Fried, T. Lohner, N.Q. Khánh, J. Gyulai, J. Appl. Phys. **93**(5), 1987 (2003)
108. P. Petrik, E.R. Shaaban, T. Lohner, G. Battistig, M. Fried, J. Garcia Lopez, Y. Morilla, O. Polgár, J. Gyulai, Thin Solid Films **455–456**, 239 (2004)
109. P. Petrik, F. Cayrel, M. Fried, O. Polgár, T. Lohner, L. Vincent, D. Alquier, J. Gyulai, Thin Solid Films **455–456**, 344 (2004)
110. P. Petrik, E.R. Shaaban, T. Lohner, G. Battistig, M. Fried, J. Garcia Lopez, Y. Morilla, O. Polgar, J. Gyulai, Thin Solid Films **455–456**, 239 (2004)
111. P. Petrik, M. Fried, T. Lohner, O. Polgár, J. Gyulai, F. Cayrel, D. Alquier, J. Appl. Phys. **97**, 123514 (2005)
112. P. Petrik, M. Fried, T. Lohner, N.Q. Khánh, P. Basa, O. Polgár, C. Major, J. Gyulai, F. Cayrel, D. Alquier, Nucl. Instrum. Methods B **253**, 192 (2006)
113. P. Petrik, T. Lohner, L. Égerházi, Zs. Geretovszky. Appl. Surf. Sci. **253**, 173 (2006)
114. P. Petrik, M. Fried, É. Vázsonyi, T. Lohner, E. Horváth, O. Polgár, P. Basa, I. Bársony, J. Gyulai, Appl. Surf. Sci. **253**, 200 (2006)
115. P. Petrik, Phys. Status Solidi A **205**(4), 732–738 (2008)
116. P. Petrik, N.Q. Khanh, J. Li, J. Chen, R.W. Collins, M. Fried, G.Z. Radnoczi, T. Lohner, J. Gyulai, Physica Status Solidi C-Curr. Top. Solid State Phys. **5**, 1358 (2008)
117. P. Petrik, M. Fried, Z. Zolnai, N.Q. Khánh, J. Li, R.W. Collins, T. Lohner, Mater. Res. Soc. Symp. Proc. **1123**, P05-01 (2009)
118. P. Petrik, M. Fried, E. Vazsonyi, P. Basa, T. Lohner, P. Kozma, Z. Makkai, J. Appl. Phys. **105**, 024908 (2009)
119. P. Petrik, S. Milita, G. Pucker, A.G. Nassiopoulou, J.A. van den Berg, M.A. Reading, M. Fried, T. Lohner, M. Theodoropoulou, S. Gardelis, M. Barozzi, M. Ghulinyan, A. Lui, L. Vanzetti, A. Picciotto, ECS Trans. **25**, 373 (2009)
120. P. Petrik, Z. Zolnai, O. Polgar, M. Fried, Z. Betyak, E. Agocs, T. Lohner, C. Werner, M. Röppischer, C. Cobet, Thin Solid Films **519**, 2791 (2011)
121. P. Pintér, P. Petrik, E. Szilágyi, Sz. Kátai, P. Deák. Diam. Relat. Mater. **6**, 1633 (1997)
122. O. Polgár, M. Fried, T. Lohner, I. Bársony, Surf. Sci. **457**, 157 (2000)
123. O. Polgár, P. Petrik, T. Lohner, M. Fried, Appl. Surf. Sci. **253**, 57 (2006)
124. O. Polgar, M. Fried, N. Khanh, P. Petrik, I. Barsony, Physica Status Solidi C-Curr. Top. Solid State Phys. **5**, 1354–1357 (2008)
125. J.P. Proot, C. Delerue, G. Allan, Appl. Phys. Lett. **61**(16), 1948 (1992)
126. P.I. Rovira, R.W. Collins, J. Appl. Phys. **85**, 2015–2025 (1999)
127. K. Tsunoda, S. Adachi, M. Takahashi, J. Appl. Phys. **91**, 2936 (2002)
128. M. Serényi, T. Lohner, P. Petrik, C. Frigeri, Thin Solid Films **515**, 3559 (2007)
129. C. Simbrunner, Tian Li, A. Bonanni, A. Kharchenko, J. Bethke, K. Lischka, H. Sitter. J. Cryst. Growth **308**, 258 (2007)
130. P.G. Snyder, J.A. Woollam, S.A. Alterovitz, B. Johs, J. Appl. Phys. **68**, 5925 (1990)
131. E. Strein, D. Allred, Thin Solid Films **517**, 1011 (2008)
132. D. Stroud, Phys. Rev. B **12**, 3368 (1975)
133. D. Stroud, F.P. Pan, Phys. Rev. B **17**, 1602 (1978)
134. J. Tauc, R. Grigorovici, A. Vancu, Phys. Stat. Sol. **15**, 627–637 (1966)
135. M. Vaupel, U. Stoberl, Nanocharm Newslett. 3, 4 (2009) http://www.nanocharm.org
136. E. Vazsonyi, E. Szilagyi, P. Petrik, Z.E. Horvath, T. Lohner, M. Fried, G. Jalsovszky, Thin Solid Films **388**, 295 (2001)
137. U. Wurstbauer, C. Röling, U. Wurstbauer, W. Wegscheider, M. Vaupel, Peter H. Thiesen, Dieter Weiss, Appl. Phys. Lett. **97**, 231901 (2010)
138. Y. Zhang, Y. Chen, G. Jin, Appl. Surf. Sci. **257**, 9407 (2011)

Chapter 18
Spectroscopic Ellipsometry for Inline Process Control in the Semiconductor Industry

Stefan Zollner

Abstract Building high-performance silicon CMOS or III/V devices requires many structural process steps, where a physical layer (dielectric, epitaxial semiconductor, gate electrode, or metal contact) is first deposited across the wafer and then partially removed by a masked or blanket etch process. Whenever possible, the deposition and etch processes are verified by a thickness metrology step. Typically, a latest-generation CMOS process flow contains about 100 thickness measurements, at least during the development of the technology. The metrology method of choice in the semiconductor industry is spectroscopic ellipsometry, because it is fast, nondestructive, and capable of measuring product wafers in small areas (30–50 μm beam diameter) set aside in the scribe grid. This chapter will describe several typical applications of ellipsometry in CMOS and III/V device manufacturing and address capabilities and limitations and how future basic research on optical properties of materials can benefit the industry.

18.1 Introduction

Building a high-performance complementary metal-oxide-semiconductor (CMOS) processor (see Fig. 18.1) with eleven layers of metals requires about 75 photolayers [9]. Many of these are dopant implant layers and therefore only require after develop inspection (ADI), overlay, and critical dimension (CD) metrology for optical lithography. However, there are about 40 structural process steps, where a physical layer (dielectric, epitaxial semiconductor, gate electrode, or metal contact) is first deposited across the wafer and then partially removed by a masked or blanket etch process (wet or dry). Whenever possible, the deposition and etch processes are verified by a thick-

---

S. Zollner (✉)

Department of Physics, New Mexico State University, MSC 3D, P.O. Box 30001,
Las Cruces, NM 88003-8001, USA
e-mail: zollner@nmsu.edu
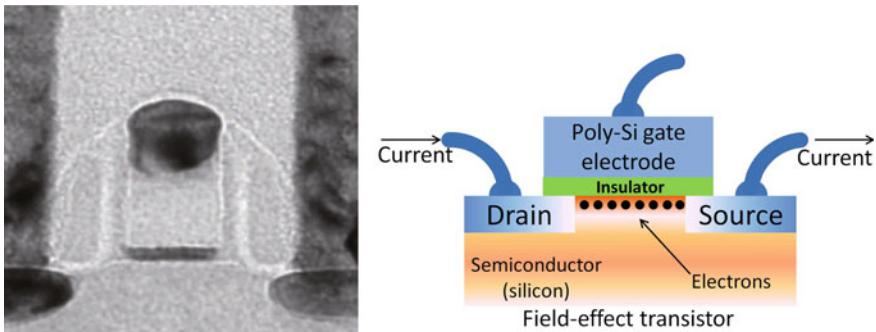
Fig. 18.1 *Left* transmission electron micrograph (TEM) of an N-channel metal-oxide-semi-conductor (NMOS) field effect transistor. In the center of the image, from the *bottom*, we see the transistor's Si channel (*grey*), followed by the metal oxide gate dielectric as a *thin bright line*, a *dark* transition metal nitride acting as the thin metal gate, followed by a polycrystalline silicon gate (*grey*), then the *dark* NiSi gate electrode contact. The gate is encapsulated on both sides by a thin nitride spacer (*grey*), a thin L-shaped oxide spacer (*bright*), and a final nitride spacer (*grey*). Above the gate, we see the dual stress nitride liner etch stop (*grey*). On both sides, there are tungsten plugs (*dark*) with transition metal nitride barriers (not visible) landing on the NiSi contacts to the silicon current electrodes. PMOS devices are similar, but have silicided $Si_{1-x}Ge_x$ source-drain stressors. The gate width of this device is on the order of 30 nm. From Ref. [9]. © 2009 The Japan Society of Applied Physics. *Right* schematic of the same device

ness metrology step, at least during process development (before transferring the process to manufacturing).

A (pre) measurement after deposition determines the initial thickness of the layer just deposited and (if possible) the thicknesses of buried layers deposited previously. A (post) measurement after the etch measures the thicknesses of the remaining layers, see Fig. 18.2. Sometimes it is desirable to etch away most, but not all of the last layer (underetch). In many cases, the last layer needs to be removed completely without etching into the layer underneath. Sometimes, the last layer and some of the layer underneath need to be removed (overetch). Ellipsometry is capable of determining the amount of under-/overetch as well as potential roughening of the surface (lateral variations in etch rate). Such measurements are usually performed at several locations on a wafer (to check for across-wafer variations of deposition and etch rates). Also, one usually measures 2–3 wafers per lot with a predetermined sampling strategy, since many process tools have multiple chambers or chucks (with potential variations in deposition or etch rates). Typical measurement results along with the accuracy of the measurement and allowed tolerances are listed in a *process assumptions* document for the specific technology. During manufacturing, a lot is put *on hold* if the measurement result is outside of the specified range. Sometimes, a manufacturing (deposition or etch) tool is restricted from processing additional lots, until the measurement discrepancy is understood and (if needed) the manufacturing tool repaired or tuned to avoid further misprocessing.
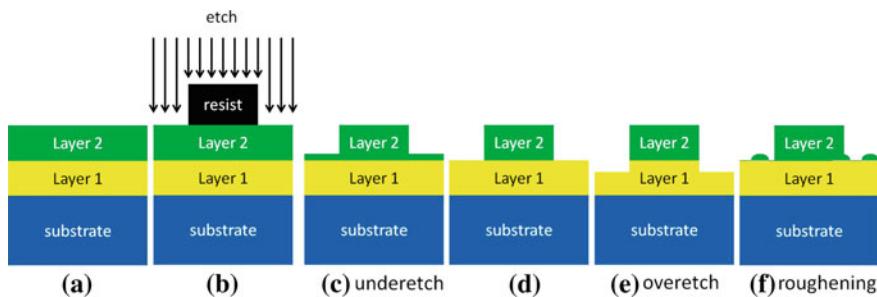
Fig. 18.2   **a** Two layers were deposited on a substrate, followed by **b** patterning a line of photoresist (*black*). The *arrows* indicate exposure to a directional dry etch. **c** Underetch: a thin residual layer 2 remains after the etch. **d** Layer 2 is removed completely, without etching into layer 1. **e** Overetch: layer 2 and a portion of layer 1 were removed. **f** Nonuniform portions of layer 2 remain, leading to a roughening of the surface. Ellipsometry measurements are usually not performed on these nano-electronic devices, but on scaled-up (much larger) structures to accommodate the 30 by 50 μm spot size of the focusing ellipsometer

All measurement results are loaded in the factory data warehouse, together with process data (e.g., photolithography exposure dose and focus, etch end point times) and in-line electrical test results, which allows searching for correlations to improve yield and performance. Results can also be used for advanced process control. For example, if the thickness measurement is below target (but still within specifications), the wafer can move on to the next process step, but the deposition time for the next wafer is automatically adjusted to bring the next thickness closer to target.

A typical 32 nm CMOS process flow contains about 100 thickness measurements. This includes about 15 step height measurements using profilometry or atomic force microscopy (AFM, mostly in back-end metallization), 80 ellipsometry measurements, and a handful of others, such as x-ray diffraction, x-ray reflectivity, x-ray fluorescence, or experimental inline surface science techniques (mostly to verify composition and thicknesses in the high-k metal gate stack).

The thickness metrology method of choice in the semiconductor industry is spectroscopic ellipsometry, because it is fast, non-destructive, and capable of measuring product wafers in small areas (30–50 μm beam diameter) set aside in the scribe grid between functional chips. Also, ellipsometry is usually capable of determining several layer thicknesses in a single measurement. The beam focusing optics for inline ellipsometers are quite complex (often involving reflective optics for the widest spectral range) and therefore limited to a single angle of incidence (around 70–75°). Older inline ellipsometers had a spectral range of 250–800 nm, which made them less sensitive to ultrathin oxides (where best accuracy is achieved near the E_2 peak of Si). More recently, the spectral range was extended to 190 or even 150 nm by using two light sources, a deuterium lamp in addition to the common xenon short-arc lamp. Measurements in the vacuum-ultraviolet (VUV) near 150 nm require purging of the ellipsometer with nitrogen or evacuating the wafer chamber (more time-consuming). In comparison with research ellipsometers, inline ellipsometers often
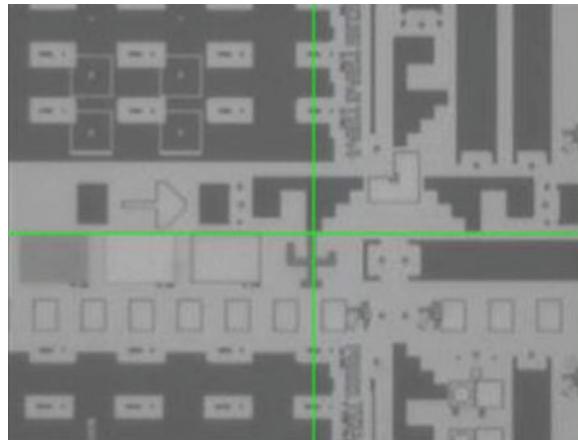
Fig. 18.3 Optical image of a scribe grid with metrology fields. The features in the center (the saguaro cactus, the staircases, and the arrow) are unique alignment marks to find the metrology fields (100 μm wide by 70 μm tall) to the *left* of the cactus

have very limited software options. A leading manufacturer only allows the following models: tabulated optical constants, effective medium approximation, Lorentz oscillator [18].

A typical CMOS semiconductor wafer factory (fab) will employ a dozen or more spectroscopic ellipsometers with a typical acquisition cost of USD 500 k to 1 M. This makes matching of the ellipsometry tools (to 0.2 Å oxide thickness) within the same fab or between fabs and sharing of ellipsometry models and recipes (including optical constants) a high priority. While pattern recognition (see Fig. 18.3) with white light can be simple for some classes of films, it can be a challenge for ultrathin gate oxides or for thick films, where small relative thickness variations can cause a strong change in feature reflectivity. Often, spectroscopic ellipsometry is combined with other measurement techniques in the same tool, for example broad-band dual-beam reflectance, single-wavelength ellipsometry, or wafer bow stress measurements using laser deflection.

Finally, it is well known that semiconductor surfaces age with time through growing of native oxides or adsorption of surface overlayers. In the semiconductor industry with its extensive use of wet chemistry for wafer processing, this is known as atmospheric molecular contamination (AMC). For reliable thin-film thickness measurements, adsorbed films need to be removed inside the ellipsometer before the measurement. This can be achieved by heating the wafer with a hot plate or by localized exposure to an intense light source. Within a few seconds, adsorbed surface overlayers can be removed, significantly improving the accuracy and repeatability of the thickness measurement.

Table 18.1 Recommended references for optical constants and dispersion of materials used in the semiconductor industry

Material	Reference	Type	Comments
Si (undoped, crystalline)	[10, 11]	Tabulated	"Woollam silicon"
Si (implanted, annealed)	[20]	Tabulated	Drude contribution
Si (implanted, not annealed)	[6, 19]	Tauc-Lorentz	similar to a-Si
Si (poly-crystalline)	[40]	Gen. oscillator	Depends on anneal conditions
Si (amorphous)	[6, 19]	Tauc-Lorentz	Large variability
SiO_2	[10, 11]	Tabulated	"Woollam SiO_2"
Silicon nitride	[3, 21, 37]	Tauc-Lorentz	Large variability
Silicon oxynitride	[5]	Tauc-Lorentz	Large variability
$Si_{1-x}Ge_x$ alloys	[39]	Tabulated	x-dependent (pseudomorphic)
$Si_{1-y}C_y$ alloys	[41]	Tabulated	y-dependent (pseudomorphic)
TiN	[15]	Tabulated	Highly variable
Ge	[23]	Tabulated	In progress
GeO_2	[1, 2, 4, 14, 25]	Tabulated	
GaAs	[36]	Tabulated	
GaAs native oxide	[35]	Formula	Fitted to tabulated data

18.2 Legacy CMOS Applications

Up to the 90 nm node,[1] the only front-end materials to be measured were silicon (undoped or implanted and annealed single-crystalline, poly-crystalline, or amorphous), SiO_2, and silicon nitride. Table 18.1 shows recommended references for the optical constants of these materials.

For undoped single-crystalline Si and SiO_2, accurate optical constants were determined from a multiple-wafer analysis (consisting of several Si wafers with different oxide thicknesses). Assuming that the Si and SiO_2 optical constants do not vary with thickness, this set of wafers allows unambigious determination of the optical constants [10, 11]. A slight variation of the SiO_2 optical constants is possible, since some oxides are denser than others. Within the effective medium approxation, this can be modeled as a mixture of SiO_2 and voids.

Amorphous Si (a-Si) is described with a Tauc-Lorentz model with variable parameters [6, 19]. Polycrystalline Si (poly-Si) produced by annealing of a-Si films, by ion implantation followed by annealing, or by high-temperature chemical vapor deposition have optical constants combining elements of both crystalline and amorphous Si. Peaks in the spectra are usually broadened and sometimes shifted. To describe the optical constants of poly-Si, one first fits the tabulated optical constants of Si with various oscillator shapes [40] and then allows variations of the oscillator parameters to achieve a good fit. Depending on the crystallinity of poly-Si, large variations in these parameters are possible. Free-carrier effects due to *in situ* doping

---

[1] The node length is the dynamical random access memoriy (DRAM) half pitch of transistors for the technology.

or ion implantation followed by annealing lead to a Drude contribution (a Lorentz oscillator with zero energy) and a broadening and red-shifting of the interband peaks in the spectra. Unless the dopant concentration is near or above the solid-solubility limit, such free carrier effects do not need to be considered within the limited spectral range of inline ellipsometers (to 900 nm in the IR).

Silicon nitride is an amorphous material with a large band gap. It is used in the semiconductor industry for many purposes, since it has a high dielectric constant and is chemically distinguishable from SiO_2. By changing the Si to N ratio, these properties are tunable. The optical properties of silicon nitride depend strongly on the growth conditions. This material is best described using the Tauc-Lorentz model, where the parameters need to be varied over a broad range [3, 21, 37]. Inline ellipsometry usually measures the nitride thickness and its refractive index at 633 nm.

While layers used in nano-electronic device fabrication should be as smooth as possible, some surface roughness is unavoidable. Surface roughness is modeled within the effective medium approximation as a mixture of the underlying material and voids.

18.3 Applications to Recent CMOS Gate-First Integrations

For the 90 and 65 nm CMOS nodes, new materials had to be introduced in front-end semiconductor manufacturing to keep meeting the scaling requirements of Moore's Law. To avoid physical tunneling of charge carriers through the gate oxide, the equivalent (SiO_2) oxide thickness (EOT) had to be decreased while keeping the physical oxide thickness constant (near 2–3 nm). This was achieved by incorporating nitrogen into the gate oxide, which increases the static dielectric constant. Gate oxide metrology no longer just requires measurement of an oxide thickness, but also the determination of the nitrogen content. Since the N content has a strong impact on the band gap of silicon oxynitride (SiON), VUV ellipsometry has become essential for gate oxide metrology [5, 13]. It is obvious that VUV ellipsometry is also more sensitive to ultrathin gate oxides (with 1–2 nm thickness) [5].

The second class of new materials comprises alloys of Si with Ge and C. The mobility of the Si channel (see Fig. 18.1) can be increased by partially etching away the Si in the source-drain regions, followed by selective epitaxial growth of pseudo-morphic $Si_{1-x-y}Ge_xC_y$ for the PMOS or $Si_{1-y}C_y$ for the NMOS transistor. The dielectric functions of such alloys are shown in Figs. 18.4 and 18.5. To maximize the stress in the channel, to improve the quality of the source-drain (silicide) Ohmic contacts, and for other process-related reasons, these layers may not necessarily be uniform in composition. The source-drain alloy stressors also may be *in situ* doped during the epitaxial growth. For silicon-on-insulator (SOI) CMOS technology, ellipsometry can measure the depth of the source-drain divot and the thickness of the remaining silicon over the buried oxide after etching, similar to Fig. 18.2e. After epitaxial growth, the optical contrast between silicon and the semiconductor alloy allows the measurement of the epitaxial layer thickness and the average composition,
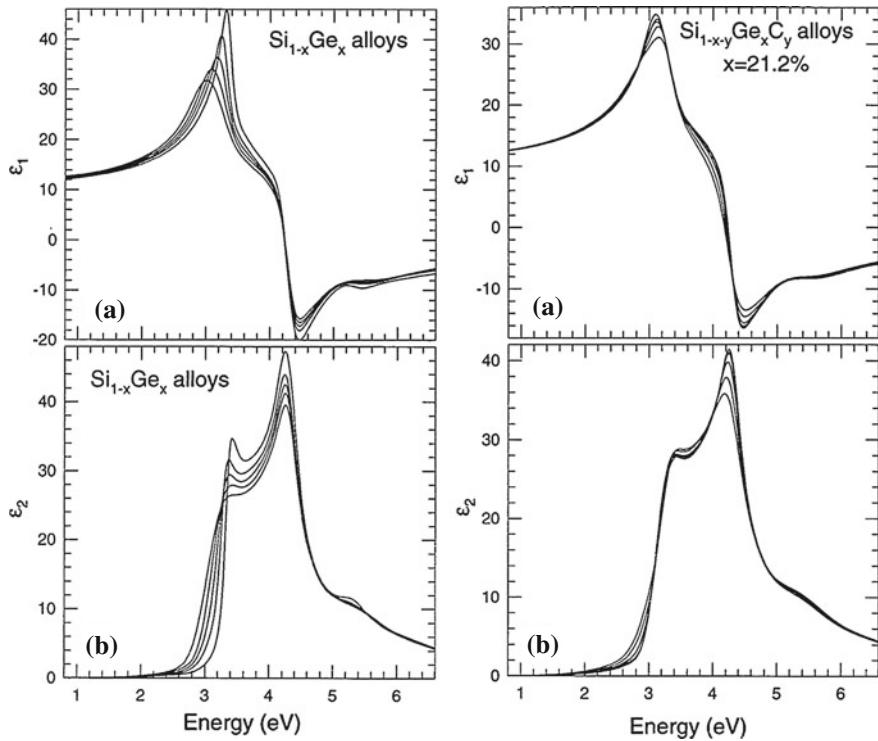
Fig. 18.4 Real (**a**) and imaginary (**b**) parts of the dielectric function of $Si_{1-x}Ge_x$ ($x = 0, \ldots, 0.275$) and $Si_{1-x-y}Ge_xC_y$ ($x = 0.212$, $y = 0, \ldots, 0.013$) alloys grown pseudomorphically on Si. Reprinted with permission from Ref. [39]. © 2000, American Institute of Physics

see Fig. 18.6, since the dielectric function of the $Si_{1-x-y}Ge_xC_y$ alloys as a function of composition is reasonably well known [39, 41]. This application of ellipsometry is much more complex than determining the thickness of an SiO_2 layer on Si. The factory needs to grow pseudomorphic alloy layers with discrete compositions (determined by high-resolution x-ray diffraction, a new type of inline metrology tool for CMOS wafer fabs) and build a database of dielectric functions for $Si_{1-x-y}Ge_xC_y$ alloys, which are then used to determine thickness and composition from ellipsometric spectra. (For this application, the term *composition* means the *effective* Ge content. While Ge increases the lattice constant of Si, inclusion of C decreases it. The *effective* Ge content achieves the same stress as $Si_{1-x-y}Ge_xC_y$ in the absence of C, i.e., $y = 0$.) For SOI technologies, numerical fitting of the remaining Si thickness (about 10 nm), the $Si_{1-x}Ge_x$ alloy thickness (about 50 nm), and the effective Ge content is often numerically unstable. Layer swapping (switching the thicknesses of the Si and $Si_{1-x}Ge_x$ layers during the fit) is common. The inline ellipsometry recipe needs to have precisely defined starting values and parameter limits for the fit to be successful.
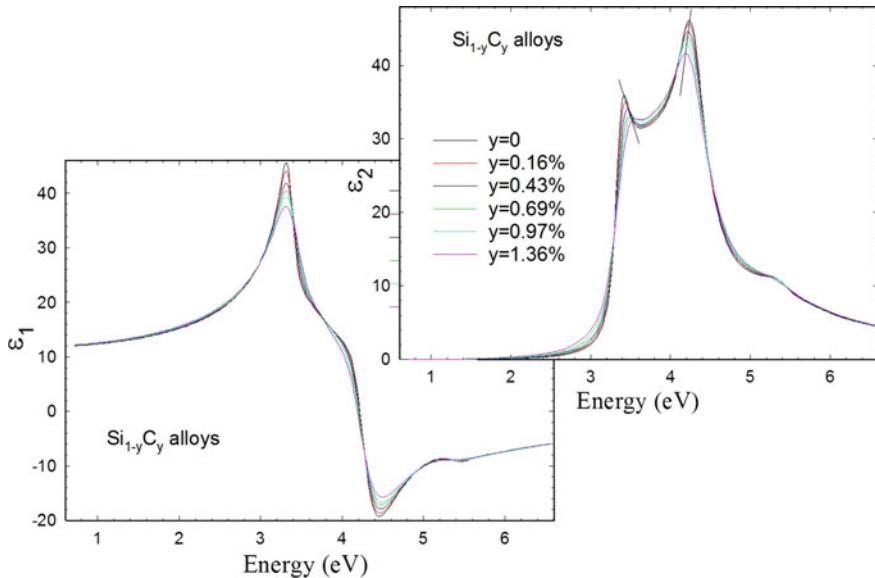
Fig. 18.5 Real and imaginary parts of the dielectric function of $Si_{1-y}C_y$ alloys pseudomorphically grown on Si ($y = 0, \ldots, 0.136$). The E_1 and E_2 critical points shift to higher and lower energies, respectively, with increasing carbon content. The dependence of ε_1 on y is barely noticeable in the visible. Nevertheless, it leads to interference oscillations for $Si_{1-y}C_y$ alloys on Si, which allow the determination of the alloy thickness and an estimation of the carbon content. Data from Ref. [41]
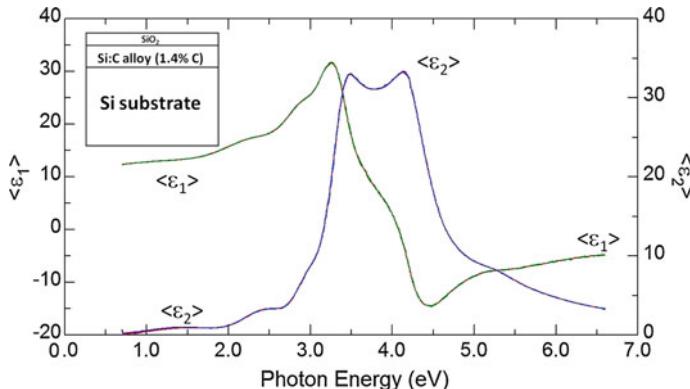
Fig. 18.6 Pseudodielectric function for a $Si_{1-y}C_y$ alloy with 102 nm thickness and $y = 1.4\%$, grown pseudomorphically on Si. A 2.1 nm thick native oxide is also assumed. Interference oscillations are clearly visible below the E_1 critical point at 3.4 eV

The most difficult application of inline ellipsometry is in the gate stack module of CMOS devices. The gate stack consists of a thin metal oxide as the gate dielectric, a transition metal nitride (TiN, TaN, etc) metal gate, and a poly-Si gate to allow
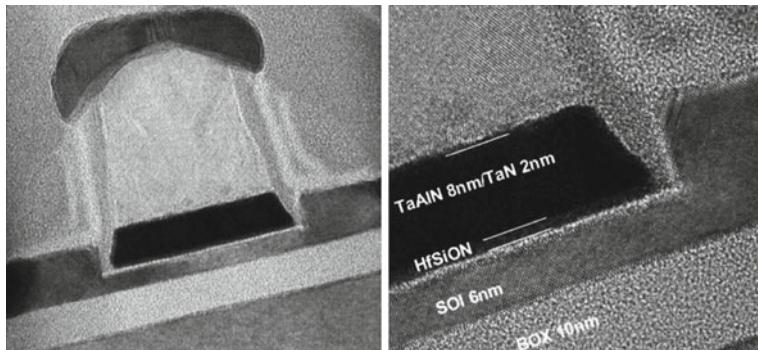
Fig. 18.7 Transmission electron micrograph of a metal gate stack, consisting of (from the *bottom*) Si substrate (*grey*), buried oxide (BOX, *bright*), top silicon (SOI channel, *grey*), thin metal oxide gate dielectric (HfSiON, *bright*), recessed TaAlN/TaN metal gate (*dark*), poly-Si gate (*bright*), and silicide gate electrode contact (*grey*). On both sides of the gate, raised Si source-drains were grown by selective epitaxy. © 2010 IEEE. Reprinted with permission, from Ref. [33]

conventional silicide Ohmic contacts, see Fig. 18.7. Tuning the threshold voltage of transistors requires adjustment of the work function of the metal gate. In conventional CMOS devices, this is achieved by implanting dopants into the poly-crystalline silicon gate. For metal gates, workfunction adjustment requires fine-tuning the composition (doping with N or Al) of the gate oxide and the metal gate electrodes using proprietary (and often poorly understood) techniques [33].

The most basic metal oxide gate dielectric is HfO_2, which has been characterized thoroughly using VUV ellipsometry [5, 32]. Thick HfO_2 layers grown at high temperatures (550 °C) crystallize in the monoclinic phase, which is thermodynamically stable [26]. Such films show a sharp increase of the absorption at the band edge near 5.8 eV along with a crystal-field splitting (monoclinic) absorption peak at 6 eV, see Fig. 18.8. Thick films grown at lower temperatures (350 °C) using metal-organic chemical vapor deposition (MOCVD) or atomic layer deposition (ALD) and also very thin films crystallize in the higher-symmetry tetragonal or cubic phase or they are amorphous. They also may contain a significant amount of carbon or other residues from the precursor. In such films, the absorption peak at 6 eV is absent and the main absorption peak at 7 eV is much broader. Such films also have significant absorption below the band gap. The refractive index of films produced at higher temperatures also tends to be higher, indicating a higher film density. The band gap of HfO_2 can be reduced by doping with Ti or Zr, see Fig. 18.9, which in turn increases the low-frequency dielectric constant important for electrical performance [12, 31]. The work function of the metal gate deposited over the metal oxide gate dielectric also can be tuned by doping with proprietary methods.

To avoid reaction of HfO_2 with the Si substrate, the HfO_2 layer needs to be grown on a chemical oxide starting surface. This chemical oxide can be formed by cleaning the wafer in a solution consisting of deionized water, hydrogen peroxide,
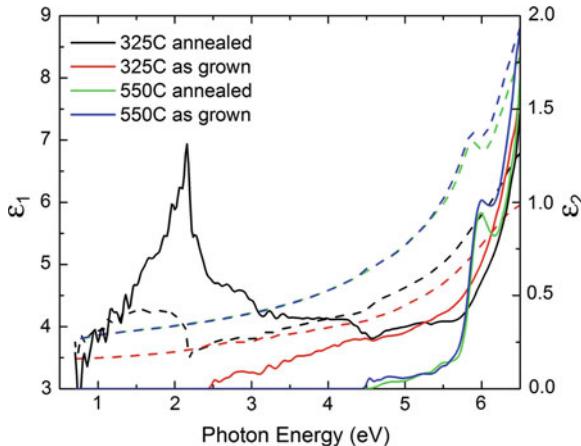
Fig. 18.8 Real (*dashed*) and imaginary (*solid*) parts of the dielectric function of 20 nm thick HfO$_2$ films grown by MOCVD at 325 and 550 °C using tetrakis diethylamido hafnium (TDEAH), as grown and after annealing at 900 °C for 60 s. The peaks at 2 eV in the data for the annealed film grown at 325 °C are probably an error in the measurement caused by the lack of surface sensitivity at low energies. Data from Ref. [26]
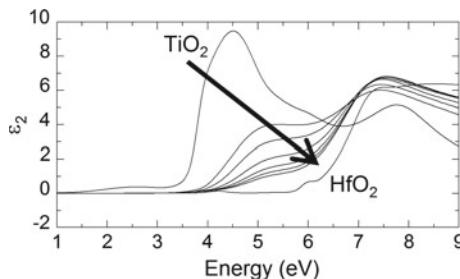
Fig. 18.9 Imaginary part of the dielectric function of Hf$_{1-x}$Ti$_x$O$_2$ grown by atomic layer deposition at 300 °C for various compositions x. With increasing Ti content x, the band gap decreases. Reprinted with permission from Ref. [31]. © 2005, American Institute of Physics

and hydrochloric oxide [31]. Similar oxide or oxynitride interfacial layers can also be formed in a conventional gate oxide diffusion furnace.

Gate oxide metrology begins with a VUV ellipsometry measurement of the oxide interfacial layer. Our goal is to determine the thickness and the composition of this layer. The latter can be obtained from the VUV absorption coefficient. Since the oxide intefacial layer is extremely thin (<5 Å), it may not be possible to determine thickness and dielectric function independently. However, it is still possible to detect deviations in the ellipsometric data (such as Fourier coefficients as a function of photon energy) due to process drifts.

Once the thickness of the interfacial oxide has been established, the HfO$_2$ (or similar) gate oxide is deposited, followed by another ellipsometry measurement

to determine thickness and composition of the metal oxide. The metal oxide may consume all or part of the SiO_2 interfacial layer, forming a silicate. This is known as interfacial layer *scavenging* [7] and significantly reduces the total effective oxide thickness. Since the dielectric function of HfO_2 is strongly process-dependent, each factory needs to build its own database of optical constants for commonly used gate oxides. Separate measurements may be needed for NMOS and PMOS, since different workfunctions of the metal gate may require different compositions of the metal oxide. Details depend on the integration scheme. Since it often takes weeks or even months from gate oxide deposition to the first reliable electrical data with copper contacts, an early detection of thickness or compositional variations in the gate sector is crucial. It should also be noted that the channel (on which the gate oxide is deposited) may not be pure silicon. It may contain a thin buried SiGe channel with a Si cap to overcome PMOS short-channel effects, which creates more complications for the gate oxide ellipsometry measurement [34]. A SiGe channel with 6 nm thickness and 55 % Ge, buried under a Si cap with 2 nm thickness, has been suggested [7, 8]. More exotic channels containing pure Ge or III/V compound semiconductors are actively investigated by universities and research institutions.

On top of this complex channel/oxide stack, a variety of transition metal gate electrodes (TiN, TaN, TaAlN, TaC, etc) can be deposited to achieve gate control of the transistor [33]. On the same wafer, several electrode materials may be needed to form NMOS and PMOS devices with different threshold characteristics. Even for the same nominal material (say, TiN) materials properties (grain size, composition, texture) may vary over a broad range. The dielectric function of such transition metal nitrides displays a metallic Drude term at low photon energies and broad interband transitions at higher energies [15]. The Drude parameters (carrier concentration and relaxation time) are process-dependent, requiring optical characterization of each transition metal nitride film used by the factory.

Because of all these complexities, inline ellipsometry applications on metal gates are rare. However, we need to keep in mind that the integration scheme may allow measurements on only parts of the gate stack. For example, the metal gate and poly-Si gate can be deposited on SiO_2 in the shallow trench isolation (STI) areas, which are easier to model than the complex buried SiGe channels.

On top of the transition metal gate, a layer of amorphous or poly-crystalline silicon is grown. Historically, the total gate height (transition metal plus poly-Si) has been about 80–100 nm, but it is desirable to reduce this height to better fit the tungsten plug contacts to the source-drain regions between narrowly spaced gates. Exact control of the gate height is important not only for uniform electrical performance, but also to allow tight control of the gate patterning. Modern lithography tools have a very shallow focal range. Therefore, the height of the gate stack is extremely important for good patterning and control of critical dimensions. Ellipsometry can, of course, determine both the silicon thickness and its roughness, if the optical properties of the transition metal gate are known.

Finally, the gate stack is finished with deposition of a thin SiO_2 and silicon nitride cap. These layers are needed as a hard mask to avoid growth of the epitaxial source-

drain semiconductors on top of the gate. Tight process control of this sacrificial hard mask can be aided with ellipsometry measurements.

18.4 Thickness Metrology Requirements in the International Technology Roadmap for Semiconductors (ITRS)

As described in the previous section, the increasing complexity of the transistor device structure and the use of many novel materials have made front-end semiconductor metrology much more challenging than in older CMOS generations. Advancements in research needed to continue scaling of CMOS circuits (to follow Moore's Law) have long been described in the International Technology Roadmap for Semiconductors (ITRS), see http://www.itrs.net [16], previously known as the National Technology Roadmap for Semiconductors (NTRS). Since 1994, the NTRS (ITRS) has had a section on metrology, which describes measurement needs, including layer thicknesses, measurement precision, and potential measurement techniques, such as spectroscopic ellipsometry, x-ray reflectometry, etc. The NTRS and ITRS have long listed the *equivalent gate oxide thickness* (EOT) and the metal barrier thickness as critical parameters for front-end and back-end thickness metrology. More recently, the top silicon thickness has been added as a third critical parameter for silicon-on-insulator (SOI) technologies.

The 1997 NTRS [24] specified a 40–50 Å thick gate oxide with a 3σ precision of 4 % (about 2 Å) for the 250 nm DRAM half-pitch node. For this quarter-micron technology, the gate oxide was a thin SiO_2 layer. Measuring its thickness on a smooth Si channel was quite straightforward, since only a single thickness had to be determined. (Spectrocopic ellipsometry, unlike x-ray reflectometry, is not usually able to distinguish between an oxide overlayer, a different surface overlayer, or surface roughness.) Single-wavelength, multiple-wavelength, and spectroscopic ellipsometers could all be used for this measurement. Even spectroscopic reflectometry was reasoably accurate. The UV spectral range was not particularly important in this technology node for such a thick gate oxide.

The 2001 ITRS edition [17] listed a gate dielectric EOT of about 15 Å and a 3σ precision requirement of 0.05 Å for the gate oxide thickness measurement of 130 nm DRAM half-pitch logic devices. For this 130 nm technology, the gate oxide was nitrided SiO_2 (or silicon oxynitride) over a bulk Si channel. The EOT refers to an SiO_2 dielectric with the same capacitance. With recent improvements in accuracy of spectroscopic ellipsometers (by emplyoying a compensator or similar phase-shifting element) and extensions of the spectral range into the quartz- and vacuum-ultraviolet (to take advantage of the low penetration depth of Si near 4.2 eV), this task can be achieved with good (relative) precision, if the oxide layer is uniform and its refractive index is known. In practice (since both assumptions are not true), one looks for changes in the thickness for a constant refractive index. Such process changes include

a change in thickness or a change in the refractive index (for example a deviation in the nitrogen content or its distribution throughout the oxide film).

But what does an ellipsometry measurement of such an ultrathin gate oxide actually mean? Let's assume that the gate oxide consists of a 15 Å thick nitrided SiO_2. Most certainly, this gate oxide has a graded refractive index. Let's cut the gate oxide into N layers (maybe $N = 15$ layers with a thickness of 1 Å each). The thickness of layer j (from 1 to N) is d_j, its complex refractive index is $\tilde{n}_j$ and the angle of refraction in layer j is ϕ_j. Following Jellison's notation[18] for the Abelés transfer matrix formalism, the diagonal elements of the Jones matrix are given by

$$r_p = \frac{M_{21,p}}{M_{11,p}} \quad \text{and} \quad r_s = \frac{M_{21,s}}{M_{11,s}}. \tag{18.1}$$

If the layers are sufficiently thin and the vacuum wavelength λ sufficiently large, then the phase factors

$$b_j = \frac{2\pi d_j}{\lambda} \tilde{n}_j \cos \phi_j \ll 1 \tag{18.2}$$

are very small, allowing an expansion of the products of the transfer matrices to first order in b_j. The characteristic matrices become

$$M_p = P_{0,p} \begin{pmatrix} 1 & -i \sum_{j=1}^{N} \frac{b_j \cos \phi_j}{\tilde{n}_j} \\ i \sum_{j=1}^{N} \frac{\tilde{n}_j b_j}{\cos \phi_j} & 1 \end{pmatrix} P_{N+1,p} \tag{18.3}$$

$$M_s = P_{0,s} \begin{pmatrix} 1 & i \sum_{j=1}^{N} \frac{b_j}{\tilde{n}_j \cos \phi_j} \\ i \sum_{j=1}^{N} \tilde{n}_j b_j \cos \phi_j & 1 \end{pmatrix} P_{N+1,s} \tag{18.4}$$

with the transfer matrices for the ambient ($j = 0$) and substrate ($j = N + 1 = $ sub) defined as

$$P_{0,p} = \frac{1}{2} \begin{pmatrix} 1 & \frac{\cos \phi_0}{\tilde{n}_0} \\ -1 & \frac{\cos \phi_0}{\tilde{n}_0} \end{pmatrix} \quad \text{and} \quad P_{0,s} = \frac{1}{2} \begin{pmatrix} 1 & \frac{1}{\tilde{n}_0 \cos \phi_0} \\ 1 & \frac{-1}{\tilde{n}_0 \cos \phi_0} \end{pmatrix} \tag{18.5}$$

$$P_{N+1,p} = \begin{pmatrix} \frac{\cos \phi_{\text{sub}}}{\tilde{n}_{\text{sub}}} & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad P_{N+1,s} = \begin{pmatrix} \frac{1}{\tilde{n}_{\text{sub}} \cos \phi_{\text{sub}}} & 0 \\ 1 & 0 \end{pmatrix}. \tag{18.6}$$

It is not clear to this author how an ellipsometric measurement of the Jones ratio $\rho = r_p/r_s$ with characteristic matrices given by (18.3) and (18.4) allows the determination of the average refractive index

$$\tilde{n}_{\text{ave}} = \left(\sum_{j=1}^{N} \tilde{n}_j d_j \right) / d \quad \text{with} \quad d = \sum_{j=1}^{N} d_j \tag{18.7}$$

of the gate oxide and its thickness d (or a combination of both), even if the off-diagonal elements in the characteristic matrices are very small (for an ultrathin gate oxide). This problem needs to be resolved to apply ellipsometry measurements to nanostructures and ultrathin films.

The gate oxide measurement becomes even more difficult with more advanced CMOS technologies. The 2011 ITRS edition [16] clearly states that device architecture has become a distinguishing element. There is no longer a "standard" planar CMOS architecture used by most device manufacturers. Reverse engineering of product samples or announcements from the manufacturers list SOI and bulk CMOS devices, gate-first and gate-last integrations, or even FINFETs for future generations. The gate oxide has become a metal oxide grown over a nitrided SiO_2 interfacial layer. The optical constants of the metal oxide depend on crystallinity, thickness, and various doping mechanisms to affect the work function of the metal gate for the NMOS and PMOS transistors. The complexities of process integration usually cause the gate oxide to have several layers with different compositions. Ideally, it is desirable to disassemble the gate oxide atom by atom to determine its composition and properties. In theory, this is possible using the *atom probe* technique, but extremely time-consuming and not yet quite reproducible. Clearly, spectroscopic ellipsometry is not able to yield the same amount of information as the atom probe. Nevertheless, the ITRS still lists a single thickness of 8 Å and a precision of 0.03 Å for the 22 nm half-pitch technology. Ellipsometry can certainly measure the thickness of an ultrathin oxide on silicon with high precision (assuming a constant refractive index), but since the result of the measurement depends on dozens of process inputs, it is difficult for the process engineer to decide how to determine the root cause of changes in the measurement results (i.e., how to address process drifts). For the year 2025 (the 8 nm half-pitch node), the thickness and 3σ precision of the gate oxide are specified as 5 and 0.02 Å, respectively. It is not clear how this oxide thickness can be measured over a rough semiconductor channel with varying composition (silicon-germanium, GaAs, etc), with spectroscopic ellipsometry, x-ray reflectometry, or any other inline technique.

The second thickness parameter listed in the ITRS metrology section is the top Si thickness for SOI devices. The optical constants of this ultrathin silicon layer depend on thickness, strain, and other process parameters. Details of this ellipsometry measurement are described in the chapter by Diebold et al. in this volume.

The third critical thickness parameter in the ITRS metrology section is the metal barrier thickness for contacts and planar interconnects. Typically, spectroscopic ellipsometry is not the method of choice for metal thickness measurements. X-ray fluorescence is much better suited for this purpose [38]. Nevertheless, work in the author's laboratory addresses applications of spectroscopic ellipsometry to metals used in semiconductor devices.

Table 18.2 Doping and layer content profile of a typical InGaP double heterojunction bipolar transistor (DHBT)

Layer	Material	Doping	Concentration (cm^{-3})	Thickness (nm)	Function
11	InGaAs	n+	$>10^{19}$	100	Emitter contact
10	GaAs	n+	5×10^{18}	120	Contact buffer layer
9	InGaP	n	3×10^{17}	40	Emitter
8	GaAs	p+	5×10^{19}	70	Base
7	GaAs	n	3×10^{16}	30	Collector
6	GaAs	n+	2×10^{18}	5	Dopant spike for DHBT
5	InGaP	n	3×10^{16}	10	DHBT collector
4	GaAs	n	3×10^{16}	155	Collector layer
3	GaAs	n	7.5×10^{15}	400	Collector layer
2	GaAs	n+	5×10^{18}	1000	Subcollector
1	AlAs (?)	?	?	30	Substrate isolation
0	GaAs	?	?	NA	Semi-insulating substrate

Compare Ref. [28]. See also Ref. [42]

18.5 InGaP Heterostructure Bipolar Transistors

Over the past decade, III/V devices have dominated the market for high-frequency front end electronics in mobile handsets [22, 28–30]. A particularly popular device is the InGaP heterojunction bipolar transistor (HBT), used as a power amplifier. A typical InGaP HBT epitaxial layer stack is shown in Table 18.2 and Fig. 18.10. The GaAs substrate (layer 0) is electrically and chemically isolated by a barrier. For ellipsometry modeling purposes, we assume that this barrier is AlAs (layer 1). Electrical connections to the collector are made through a highly doped thick GaAs subcollector (layer 2). For improved electrical properties, the collector has a very complex structure, including a thin GaAs collector (layer 7) separated from an InGaP barrier (layer 5) by a thin GaAs dopant spike (layer 6) on top of two GaAs layers with low doping (layers 3 and 4). Layer 8 is a thin GaAs base with high p-type doing. The InGaP emitter (layer 9) is chemically isolated from the emitter contact (highly-doped n-type InGaAs layer 11) by a thick GaAs contact buffer (layer 10). From this epi wafer, an InGaP HBT device is produced by a double-mesa etch, see Fig. 18.10.

Qualification of a new epi reactor (from an existing supplier) for producing device quality wafers can take up to a year. Qualification of a new supplier can take much longer. Epi suppliers rely on characterization techniques such as double-crystal x-ray diffraction (XRD, to determine lattice constant, strain, thickness, and alloy content), Hall measurements (for carrier concentration), and photoluminescence (PL, for band gap, doping, strain, ordering in InGaP, and composition). Secondary ion mass spectrometry (SIMS) also provides information on alloy composition, doping, and thickness. Matching epi wafers from different suppliers or different reactors requires identical procedures used in the interpretation of characterization data. For example, XRD and PL data must be analyzed using identical expressions relating
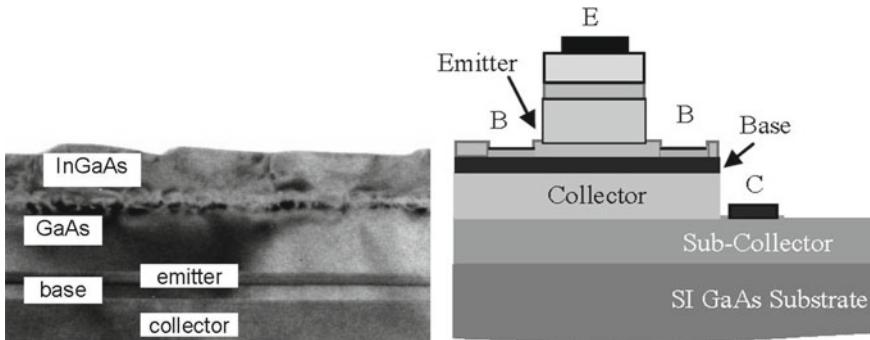
Fig. 18.10 *Left* transmission electron micrograph for an InGaP HBT epi wafer, from Ref. [42]. Non-exclusive © GaAs Mantech, Inc. *Right* schematic for an InGaP heterojunction bipolar transistor (HBT) produced by a double-mesa etch (compare Fig. 18.2) from the epi wafer described in Table 18.2. *E*, *B*, and *C* denote the contacts to the emitter, base, and collector, respectively. From Ref. [29]

the measurements to materials properties such as composition and strain. In the case of SIMS, identical calibrations standards must be used. In addition to the matching using physical measurements, electrical matching is required, since subtle epi interface mismatch can create significant discrepancies in transistor characteristics and RF module performance [30].

Ellipsometry seems like a natural fit for epitaxial materials matching and process monitoring in the compound semiconductor industry. Highly accurate reliable thin-film metrology tools (ellipsometers) are offered by several instrument suppliers. Their use is widely accepted in silicon wafer manufacturing. However, there are significant challenges. For example, we know the optical constants for pure Si and GaAs with high accuracy, but not for Ge, see Table 18.1. Similarly, SiO_2 optical constants are well known over the complete spectral range from the near-IR to the vacuum-UV, but not those of GeO_2 or of native oxides on compound semiconductors. Epitaxial structures for III/V manufacturing usually contain many different layers (see Table 18.2) including semiconductor alloys, whose optical constants depend on alloy composition, doping, and ordering in complicated ways due to the shifts and broadenings of critical points. Another issue is the following: GaAs wafer fabs often utilize obsolete four-inch wafer processing tools for manufacturing. Obsolete ellipsometry tools often cannot measure accurately below 250 or 300 nm. Just this range (near the E_2 critical point) is most useful for thickness determination of thin oxide or roughness overlayers.

Despite all these difficulties, ellipsometry can provide valuable insights of III/V device epitaxial layer stacks such as that shown in Table 18.2. Let's discuss this stack in detail: The top layer in the epi stack is a 100 nm thick graded InGaAs layer with a low band gap and high doping. This layer is needed to provide a low Schottky barrier between the emitter of the InGaP HBT device and the Ohmic contact metal. (Both high doping and low band gap reduce the specific contact resistivity.) This material
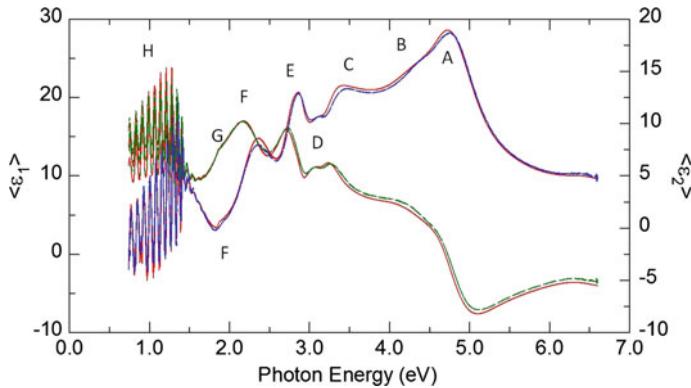
Fig. 18.11 Real (*green*) and imaginary part (*blue*) of the pseudodielectric function for the epi stack shown in Table 18.2, including layers 0–9 and a native oxide (after layers 10 and 11 were removed by etching). The characters *A–H* indicate various features of the spectra as discussed in the text. The *red solid lines* are the result of a model calculation using the layers and thickness similar to those of Table 18.2

is often rough and of poor crystal quality (heavily dislocated, see Fig. 18.10) and has no significant impact on the electrical characteristics of the HBT. An ellipsometry measurement of the entire stack with layer 11 intact mostly provides information about In content, layer 11 thickness, and surface roughness, none of which are very important for epiwafer screening. If layer 11 is removed by etching, similar arguments can be made about the thick GaAs emitter contact buffer (layer 10). An ellipsometry measurement of the stack ending with layer 10 (after layer 11 has been removed by etching) only reveals information about layer 10 thickness and native oxide thickness (or roughness).

Let's instead discuss ellipsometry measurements of the epiwafer in Table 18.2, where layers 10 and 11 have been removed by etching, see Fig. 18.11. We find that a wealth of information is obtainable. The maximum A in $\langle \varepsilon_2 \rangle$ is the E_2 critical point of InGaP. The height of this maximum determines the surface roughness or the thickness of the native oxide (modeled as a mixture of native oxide on InP and GaP following Ref. [35]). The shoulder B is related to the E_2 peak of GaAs (layers 6–8), which leaks through the InGaP cap layer. From the strength of this shoulder, the thickness of the InGaP layer 9 can be determined. The peaks C and D are mostly influenced by the E_1 and $E_1 + \Delta_1$ peaks of GaAs (layers 6–8). Surprisingly, the E_1 gap of InGaP does not influence the spectra very much, perhaps because layer 9 is too thin and the absorption too low in the E_1 region of InGaP. Structures E and F are related to interference effects. Their spacings (position in photon energy) are due to the total GaAs thickness between the InGaP barriers, i.e., the sum of the thicknesses of layers 6–8. The amplitudes of the interference fringes E and F are related to the magnitude of the electric fields reflected by the InGaP barriers and therefore contain information on the thicknesses of layers 5 and 9. On the other hand, the change in slope near symbol G is much too abrupt for an interference effect. This is indeed
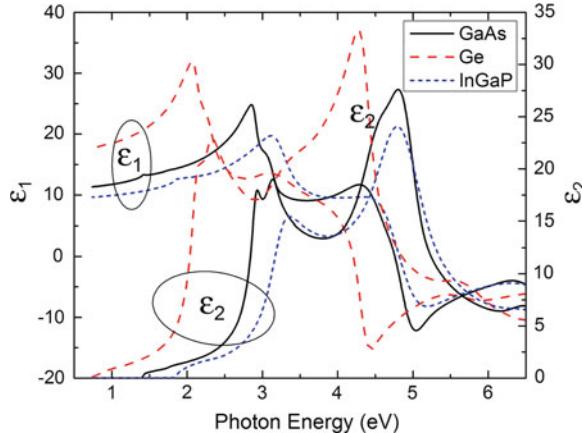
Fig. 18.12 Complex dielectric functions for Ge (*dashed*) [23], GaAs (*solid*) [36], and InGaP (*dotted*) [42]

the band gap of InGaP, which depends on the alloy composition and ordering of the InGaP alloy in layers 5 and 9. Finally, the period of the interference fringes H is given by the thickness of the GaAs collector and subcollector (layers 2–4). The amplitude of the fringes H determines the thickness of the AlAs (layer 1) substrate buffer layer. For this analysis, we used InGaP alloy optical constants similar to those published by Schubert [27] for completely disordered InGaP. We ignored the doping and confinement effects in all layers. Using this model, we are able to obtain an excellent fit to the data and we could obtain six different layer thicknesses (native oxide, layer 9, layers 6–8, layer 5, layers 2–4, layer 1). Since different layers influence the ellipsometry data in different spectral regions, we are confident that there is little correlation between these six thickness parameters. The GaAs and InGaP optical constants used for this fit are shown in Fig. 18.12.

As described in Ref. [42], several ellipsometry measurements were inserted in a commercial six-inch InGaP HBT process flow. An initial measurement as described above yields information of the InGaP emitter thickness and surface overlayers (oxide and roughness). A second ellipsometry measurement (after removal of the InGaP emitter, stopping on the GaAs base, layer 8) uses the small difference between the optical constants of the heavily p-type (carbon-doped) GaAs (layer 8) and GaAs with much lower dopant concentrations (layers 6–7) to determine the thickness of the base. A third ellipsometry measurement occurs after the removal of the base and the collector (layers 3–6). The purpose of this measurement is primarily to measure the roughness of the subcollector. In manufacturing, it was sometimes found that As precipitates (defects in the epi stack) caused excessive surface roughness (haze), which was easily detectable with ellipsometry. Figure 18.13 shows plan-view SEM image of a rough GaAs surface after a wet etch.
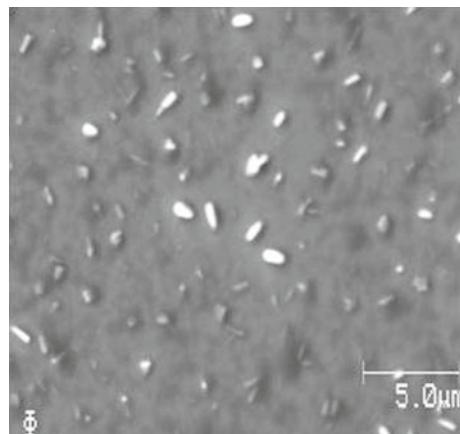
Fig. 18.13 Plan-view scanning electron micrograph (SEM) of a rough compound semiconductor surface after a wet etch. See Ref. [42]

18.6 Conclusion and Outlook

For about 15 years, since the mid-1990s, spectroscopic ellipsometry has been an integral part of process control for the semiconductor industry, for manufacturing of CMOS and bipolar silicon transistors as well as for III/V compound semiconductor devices. At first, the applications were very simple, such as the measurement of a rather thick oxide or silicon nitride on a single-crystalline silicon substrate. As integration schemes have become more complex (involving the use of new materials) and tolerances have become tighter to meet the scaling requirements of Moore's Law, ellipsometry models also have become more sophisticated to meet the metrology requirements. Similarly, instruments were improved significantly to achieve better accuracy and precision and a wider spectral range.

In many cases, inline metrology applications have exceeded our knowledge of optical constants of materials. This is especially true for manufacturing of compound semiconductors (where many different III/V alloys and their native oxides are used) and for the complicated metal gate stack in modern CMOS devices. As other industries (especially solar cells, solid-state lighting, flexible electronics) become mature, we can expect that ellipsometry applications will become crucial for them as well.

References

1. M.R. Baklanov, K.K. Svitashev, L.V. Semenenko, V.K. Sokolov, Opt. Spektrosk. **39**, 362 (1975). [Opt. Spectrosc. (USSR) **39**, 205 (1975)]
2. P. Boháč, L. Jastragic, D. Chvostová, V. Železný, Vacuum **41**, 1466 (1990)
3. C.S. Cook, T. Daly, R. Liu, M. Canonico, Q. Xie, R.B. Gregory, S. Zollner, Thin Solid Films **455**–**456**, 794 (2004)
4. G.G. Devyatykh, E.M. Dianov, N.S. Karpychev, S.M. Mazzavin, V.M. Mashinskii, V.B. Neustruev, A.V. Nikolaichik, A.M. Prokhorov, A.I. Ritus, N.I. Sokolov, A.S. Yushing, Kvant. Electron. (Moscow) **7**, 1563 (1980) [Sov. J. Quantum Electron. **10**, 900 (1981)]
5. N.V. Edwards, in *2003 International Conference on Characterization and Metrology for ULSI Technology*, ed. by D.G. Seiler, A.C. Diebold, T.J. Shaffner, R. McDonald, S. Zollner, R.P. Khosla, E.M. Secula (American Institute of Physics, Melville, 2003), p. 723
6. A.S. Ferlauto, G.M. Ferreira, J.M. Pearce, C.R. Wronski, R.W. Collins, X. Deng, G. Ganguly, Thin Solid Films **455**–**456**, 388 (2004)
7. J. Franco, B. Kaczer, G. Eneman, J. Mitard, A. Stesmans, V. Afanas'ev, T. Kauerauf, Ph.J. Roussel, M. Toledano-Luque, M. Cho, R. Degraeve, T. Grasser, L.-Å. Ragnarsson, L. Witters, J. Tseng, S. Takeoka, W.-E. Wang, T.Y. Hoffmann, G. Groeseneken, in *2010 IEEE International Electron Devices Meeting*, p. 4.1.1
8. J. Franco, B. Kaczer, G. Eneman, Ph.J. Roussel, T. Grasser, J. Mitard, L.-Å. Ragnarsson, M. Cho, L. Witters, T. Chiarella, M. Togo, W.-E. Wang, A. Hikavyy, R. Loo, N. Horiguchi, G. Goeseneken, in *2011 IEEE International Electron Devices Meeting*, p. 18.5.1
9. B. Greene, Q. Liang, K. Amarnath, Y. Wang, J. Schaeffer, M. Cai, Y. Liang, S. Saroop, J. Cheng, A. Rotondaro, S.-H. Han, R. Mo, K. McStay, S. Ku, R. Pal, M. Kumar, B. Dirahoui, B. Yang, F. Tamweber, W.-H. Lee, M. Steigerwalt, H. Weijtmans, J. Holt, L. Black, S. Samavedam, M. Turner, K. Ramani, D. Lee, M. Belyansky, M. Chowdhury, D. Aimé, B. Min, H. van Meer, H. Yin, K. Chan, M. Angyal, M. Zaleski, O. Ogunsola, C. Child, L. Zhuang, H. Yan, D. Permana, J. Sleight, D. Guo, S. Mittl, D. Ioannou, E. Wu, M. Chudzik, D.-G. Park, D. Brown, S. Luning, D. Mocuta, E. Maciejewski, K. Henson, E. Leobandung, in *Proceedings of the 2009 Symposium on VLSI Technology*, p. 140
10. C.M. Herzinger, B. Johs, W.A. McGahan, W. Paulson, Thin Solid Films **313**–**314**, 281 (1998)
11. C.M. Herzinger, B. Johs, W.A. McGahan, J.A. Woollam, W. Paulson, J. Appl. Phys. **83**, 3323 (1998)
12. R.I. Hegde, D.H. Triyoso, S.B. Samavedam, B.E. White, J. Appl. Phys. **101**, 074113 (2007)
13. J.N. Hilfiker, B. Singh, R.A. Synowicki, C.L. Bungay, in *Metrology, Inspection, and Process Control for Microlithography XIV*, ed. by N.T. Sullivan, SPIE Proceedings, vol. 3998 (SPIE, Bellingham, 2000), p. 390
14. Y.Z. Hu, J-Th. Zettler, S. Chongsawangvirod, Y.Q. Wang, E.A. Irene, Appl. Phys. Lett. **61**, 1098 (1992)
15. J. Humlíček, A. Nebojsa, J. Hora, M. Stráský, J. Spousta, T. Šikola, Thin Solid Films **332**, 25 (1998)
16. International Technology Roadmap for Semiconductors, 2011 Edition, http://www.itrs.net
17. International Technology Roadmap for Semiconductors, 2001 Edition, http://www.itrs.net
18. G.E. Jellison Jr, in *Handbook of Ellipsometry*, ed. by H.G. Tompkins, E.A. Irene (Springer, Heidelberg, 2005), p. 237
19. G.E. Jellison Jr, F.A. Modine, Appl. Phys. Lett. **69**, 371 (1996)
20. G.E. Jellison Jr, S.P. Withrow, J.W. McCamy, J.D. Budai, D. Lubben, M.J. Godbole, Phys. Rev. B **52**, 14607 (1995)
21. G.E. Jellison Jr, F.A. Modine, P. Doshi, A. Rohatgi, Thin Solid Films **313**–**314**, 193 (1998)
22. J. Li, C. Cismaru, P. Zampardi, A. Wu, E. Babcock, M. Sun, K. Stevens, R. Ramanathan, in *The International Conference on Compound Semiconductor Manufacturing Technology*, Austin, May 14–17, 2007, p. 177
23. A.A. Medina, L.S. Abdallah, S. Zollner, (unpublished).

24. National Technology Roadmap for Semiconductors, 1997 Edition (Semiconductor Industry Association, San. Jose)
25. L. Pajasová, Czech. J. Phys. B **19**, 1265 (1969)
26. J. Schaeffer, N.V. Edwards, R. Liu, D. Roan, B. Hradsky, R. Gregory, J. Kulik, E. Duda, L. Contreras, J. Christiansen, S. Zollner, P. Tobin, B.-Y. Nguyen, R. Nieh, M. Ramon, R. Rao, R. Hegde, R. Rai, J. Baker, S. Voight, J. Electrochem. Soc. **150**, F67 (2003)
27. M. Schubert, V. Gottschalch, C.M. Herzinger, H. Yao, P.G. Snyder, J.A. Woollam, J. Appl. Phys. **77**, 3416 (1995)
28. A. Shimukovitch, P. Sakalas, P. Zampardi, M. Schroter, A. Matulionis, Phys. Status Solidi RRL **4**, 335 (2010)
29. M. Sun, P. Zampardi, communication at the 9th Topical workshop on heterostructure micro-electronics (2011)
30. M. Sun, P. Zampardi, C. Cismaru, L. Rushin, in *The International Conference on Compound Semiconductor Manufacturing Technology*, Tampa, 18–21 May 2009, p. 11.2
31. D.H. Triyoso, R.I. Hegde, S. Zollner, M.E. Ramon, S. Kalpat, R. Gregory, X.-D. Wang, J. Jiang, M. Raymond, R. Rai, D. Werho, D. Roan, B.E. White, P.J. Tobin, J. Appl. Phys. **98**, 054104 (2005)
32. D. Triyoso, R. Liu, D. Roan, M. Ramon, N.V. Edwards, R. Gregory. D. Werho, J. Kulik, G. Tam, E. Irwin, X.-D. Wang, L.B. La, C. Hobbs, R. Garcia, J. Baker, B.E. White, P. Tobin, J. Electrochem. Soc. **151**, F220 (2004)
33. O. Weber, F. Andrieu, J. Mazurier, M. Cassé, X. Garros, C. Leroux, F. Martin, P. Perreau, C. Fenouillet-Béranger, S. Barnola, R. Gassilloud, C. Arvet, O. Thomas. J.-P. Noel, O. Rozeau, M.-A. Jaud, T. Poiroux, D. Lafond, A. Toffoli, F. Allain, C. Tabone, L. Tosti, L. Brévard, P. Lehnen, U. Weber, P.K. Baumann, O. Boissiere, W. Schwarzenbach, K. Bourdelle, B.-Y. Nguyen, F. Bœuf, T. Skotnicki, O. Faynot, in *2010 IEEE International Electron Devices Meeting*, p. 3.4.1
34. B. Winstead, W.J. Taylor, E. Verret, K. Loiko, D. Tekleab, C. Capasso, M. Foisy, S.B. Samavedam, IEEE Electron Device Lett. **28**, 719 (2007)
35. S. Zollner, Model dielectric functions for native oxides on compound semiconductors. Appl. Phys. Lett. **63**, 2523 (1993)
36. S. Zollner, J. Appl. Phys. **90**, 515 (2001)
37. S. Zollner, E. Apen, in *Characterization and Metrology for ULSI Technology 2000*, ed. by D.G. Seiler, A.C. Diebold, T.J. Shaffner, R. McDonald, W.M. Bullis, P.J. Smith, E.M. Secula (American Institute of Physics, Melville, 2001), p. 532
38. S. Zollner, R.B. Gregory, M.L. Kottke, V. Vartanian, X.-D. Wang, D. Theodore, P.L. Fejes, J.R. Conner, M. Raymond, X. Zhu, D. Denning, S. Bolton, K. Chang, R. Noble, M. Jahanbani, M. Rossow, D. Goedeke, S. Filipiak, R. Garcia, D. Jawarani, B. Taylor, B.-Y. Nguyen, P.E. Crabtree, A. Thean, in *2007 International Conference on Frontiers of Characterization and Metrology*, ed. by D.G. Seiler, A.C. Diebold, R. McDonald, C.M. Garner, D. Herr, R.P. Khosla, E.M. Secula. AIP Conference Proceedings, vol. 937 (2007) p. 337
39. S. Zollner, J. Hildreth, R. Liu, P. Zaumseil, M. Weidner, B. Tillack, J. Appl. Phys. **88**, 4102 (2000)
40. S. Zollner, R. Liu, J. Christiansen, W. Chen, K. Monarch, T.-C. Lee, R. Singh, J. Yater, W.M. Paulson, C. Feng, in *Amorphous and Microcrystalline Silicon Technology-1998*, ed. by R. Schropp, H.M. Branz, M. Hack, I. Shimizu, S. Wagner (Materials Research Society, Pittsburgh, 1998), p. 957
41. S. Zollner, J.P. Liu, P. Zaumseil, H.J. Osten, A.A. Demkov, Semicond. Sci. Technol. **22**, S13 (2007)
42. S. Zollner, M.G. Sadaka, N.V. Edwards, C.S. Cook, Q. Xie, H.T. Le, J. Hildreth, A.S. Morton, in *2002 GaAs MANTECH Conference*, paper 8l

Chapter 19
Thin Film Applications in Research and Industry Characterized by Spectroscopic Ellipsometry

Denis Cattelan, Céline Eypert, Marzouk Kloul, Mélanie Gaillet, Jean-Paul Gaston, Roland Seitz, Assia Shagaleeva and Michel Stchakovsky

Abstract Spectroscopic ellipsometry is one of the most accurate and reliable optical techniques to characterize a large variety of materials and thin films on any substrate. Because today's devices such as solar cells, light emitting devices and thin film transistors are formed by complex multilayers structures, including optical anisotropy, absorbing and graded materials, the correct use of ellipsometry requires a combination of the proper choice of hardware for spectra acquisition and of models for data analysis. This chapter aims at giving an industrial overview of applications of thin films whose characterization can be performed by spectroscopic ellipsometry. After giving some insights into the various addressed technology markets and a brief explanation of the application fields, examples of thin film structures characterization by spectroscopic ellipsometry are presented exploiting commercial available

D. Cattelan (✉) · C. Eypert · M. Kloul · M. Gaillet · J.-P. Gaston · R. Seitz · A. Shagaleeva · M. Stchakovsky
HORIBA Jobin Yvon S.A.S., Avenue de la Vauve–Passage Jobin Yvon, 45002-91120, CS Palaiseau, France
e-mail: denis.cattelan@horiba.com

C. Eypert
e-mail: celine.eypert@horiba.com

M. Kloul
e-mail: marzouk.kloul@horiba.com

M. Gaillet
e-mail: melanie.gaillet@horiba.com

J.-P. Gaston
e-mail: jean-paul.gaston@horiba.com

R. Seitz
e-mail: roland.seitz@horiba.com

A. Shagaleeva
e-mail: assia.shagaleeva@horiba.com

M. Stchakovsky
e-mail: michel.stchakovsky@horiba.com

spectroscopic ellipsometers that differentiate for research and industrial applications. The examples cover the latest scientific and industrial applications in the areas of optoelectronics, microelectronics and telecommunications, photovoltaics, optical coatings and chemistry and biology engineering.

19.1 Introduction

Thin film technology is pervasive in many applications, including microelectronics, photovoltaics, optoelectronics, flat panel display technology, nano- and biotechnology, optics, and protective and decorative coatings.

The development and optimization of thin films technologies involve and result from compromising process specifications, substrate material limitations, designed film properties and final cost. Technically, the development of thin film devices depends on the capability to selectively and controllably deposit thin films with thicknesses ranging from a few Angströms to a few micrometers.

Spectroscopic ellipsometry (SE) is an optical characterization technique, mainly used to measure the thickness, from a few Angstroms to several microns, and optical constants of very thin films. It turns out that few tools are as accurate, economical, and fast as ellipsometers for nanometer-scale dimensions. For example, a spectroscopic ellipsometers can measure the 8-nm-thick gate oxide of a metal-oxide semiconductor field effect transistor (MOSFET) with a precision of better than 0.01 nm.

The complex dielectric function, ε, identifies the optical properties and directly relates to the electronic properties of the material (e.g. see Chap. 1 of this book). Through measurements of the complex dielectric function, SE can provide detailed knowledge of the electronic properties of the material such as the band gap, the interband critical point energies, and the absorption coefficient. This in turn provides insight into material properties such as the composition, the degree of crystallinity, the hydrogen content, and many other physical properties such as the doping density that affect the electronic structure.

Although ellipsometry has been around for nearly a century and has been applied to problems in the microelectronics industry since the beginning of 1960s, the rapidly advancing technology requiring thinner films, multiple film stacks, films with complex and variable stoichiometries, and better precision of thickness and refractive index has been the driving force for advancements in instrumentation, software and analysis procedures. Therefore, spectroscopic ellipsometers have evolved over the past twenty years, from research instruments to fully automatic characterization tools for in-line control in fabs. This development has gone hand-in-hand with the explosion of technological breakthroughs in microelectronics and nanotechnology. Today, commercially available ellipsometers provide spectral coverage from the vacuum ultraviolet to the infrared, and the angle of incidence can be controlled by computer-driven stepper motors with a precision of better than 0.01°. The new and continuous developments of spectroscopic ellipsometers including imaging, modeling capabilities in Mueller Matrix, microspots and improved spatial resolution

provide the opportunity to fulfill the needs for three-dimensional and nanostructures characterization. Moreover, advanced software, models based on effective medium theories and/or new model for parameterization of optical constants permit sophisticated analysis of complex materials systems. Once analysis of a particular system becomes routine, software appropriate for technician use is available, including mapping material properties capabilities. As a result, despite the largest application for spectroscopic ellipsometry continues to be the semiconductor industry, in which the ability to characterize ever-smaller and more-complex architecture is critical to development and production monitoring, ellipsometry can now be applied usefully to a large variety of new industrial applications. Those include light emitting diodes-LEDs lighting, multilayered films for solar cells and organic electroluminescent displays, thin magnetic layers of computer read-write heads, photoresists for integrated circuits, transparent conducting oxides (TCOs) on glass for thermally efficient coatings, and quantum well superlattices for high-speed electronic materials.

This chapter aims at giving an industrial overview of thin film applications and their characterization performed by spectroscopic ellipsometry. After giving some insights into the various technology markets and a brief explanation of the application fields, examples of thin film structures characterized by spectroscopic ellipsometry are presented. The examples cover the latest scientific and industrial applications in the areas of optoelectronics, microelectronics and telecommunications, photovoltaics, optical coatings and chemistry and biology engineering.

19.2 Optoelectronics: Focus on LEDs

19.2.1 Market Trends

The light emitting diodes (LEDs) industry entered a rapid growth stage in 2010 with an impressive expansion in LED manufacturing and new facility projects. Digitimes Research projects that the global high-brightness LED industry will see year-on-year growth of 13.4 % in 2012, with output rising to US$10.1 billion. The global LED market is projected to grow rapidly to $14.8bn by 2015, driven by the increasing demand for more efficient and larger electronic displays for TVs, lighting, notebooks and mobile handsets.

The main challenges for general LED lighting include reducing overall production costs and increasing efficiency and lifetimes. LED lighting is already used in a wide variety of applications as shown in Fig. 19.1, such media-tablets and displays, signs, back-lighting of TV/LCD displays, traffic signals, automotive (dashboard and external lights), and for architectural displays. In terms of applications for high-brightness LEDs, large-size LED backlights will remain the biggest sector with 34.1 %, with lighting trailing behind with 16.7 %.

Gallium Nitride based (GaN) LEDs drive the market and now account for 81 % of the total market.
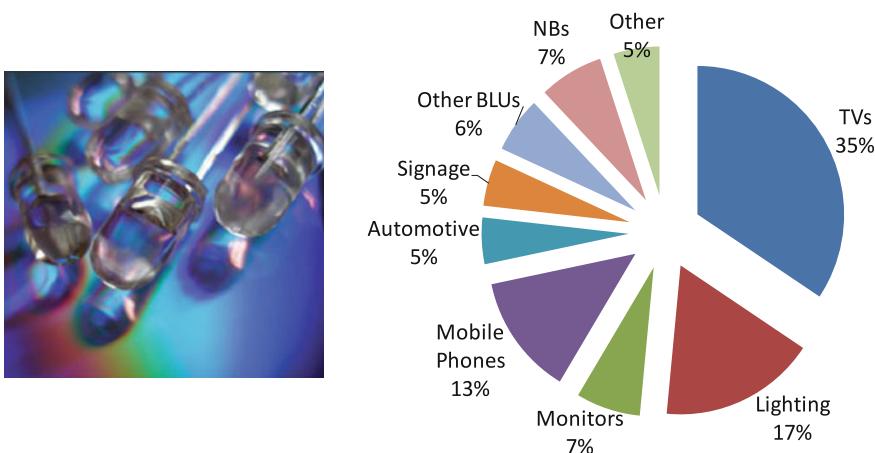
Fig. 19.1 LED lighting and their sectors of applications

A semiconductor LED is a solid-state device that emits incoherent light in a narrow spectral range when a forward bias is applied. The wavelength of the emitted light is dependent on the energy bandgap (E_g) of the material used in the active region of the device. A LED consists of a p–n junction with a multiple-quantum well (QW) active region and carrier-confined layers [1].

Nowadays the most common LEDs are based on III-nitrides materials, i.e., GaN, InGaN, AlGaN emitting over 150 lm of white, blue, or green light.

The LED structure is normally formed on a lattice-matched, or nearly lattice-matched, substrate with a low dislocation density. However, in the case of III-nitride semiconductors, sapphire (which has about 16 % lattice mismatch with GaN) is the most common substrate of choice owing to the fact that high quality GaN bulk lattice-matched substrates are not readily available or extremely expensive.

A typical GaN-based UV LED structure, shown in Fig. 19.2, has a thin low-temperature nucleation layer (often referred to as a *buffer layer*) of GaN or AlN, which is used to accommodate the lattice mismatch with sapphire, and an n-type AlGaN contact layer followed by the active, top p-type and p-contact layers. The active region typically consists of confinement layers (n-type and p-type) with either a single or multiple QWs between them. The alloy compositions of the AlGaN well and GaN barrier layers and their thicknesses are chosen based on the desired emission wavelength. The confinement layers are used to introduce an energy barrier on either side of the wells to prevent carriers—electrons and holes—from escaping the QWs without radiatively recombining. The p-contact layer is highly doped for ohmic contact formation. Most of the III-nitride-based UV LED structures are grown by metalorganic chemical vapor deposition (MOCVD), although other methods such as molecular beam epitaxy (MBE) and hydride-vapor phase epitaxy (HVPE) have also been used to a less extent [2].
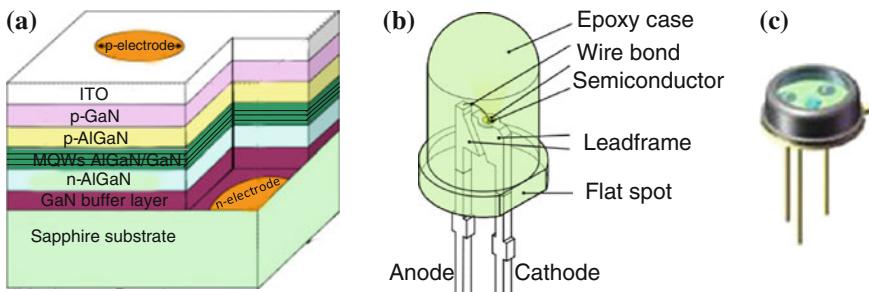
Fig. 19.2 a Structure of UV LED based on III-nitrides semiconductors. **b** Scheme of a typical 365 nm Water Clear UV LED structure. **c** Image of a 280 nm metal package deep UV LED

19.2.2 Ellipsometry Characterization

The performance of a LED depends on the design and on material properties as well as on the precise control of each layer of the LED thin film structure. The characterization by spectroscopic ellipsometry aims at determining thin film thickness and optical constants of the contact and buffer layers, and of the multi-quantum well (MQW) layers. The composition of alloys $Al_xGa_{1-x}N$ included in the MQW can also be determined.

The two examples below illustrate the non-destructive characterization of LED structures carried out by spectroscopic ellipsometry. The first example intends to develop a blue LED and we show the determination of thicknesses and optical constants in the NIR/visible range using the UVISEL Spectroscopic Phase Modulated Ellipsometer. The second example shows an industrial application where the automatic Auto SE Ellipsometer, intended as a metrology tool (as opposed to analytical tools) with pattern recognition function, is used for quality control of the LED during its production process.

The UVISEL Spectroscopic Phase Modulated Ellipsometer uses photoelastic devices to perform the polarization modulation without any mechanical movement, resulting in improved signal-to-noise ratio from FUV to NIR and no insensitive regions [3] (see also Chap. 2 of this book). Measured data are expressed as (Is, Ic), which are functions of the main ellipsometric parameters (ψ, Δ), according to the equations:

$$Is = \sin 2\,\psi \sin \Delta$$
$$Ic = \sin 2\,\psi \cos \Delta$$

The UVISEL is a research ellipsometer optically designed to perform measurements with high accuracy and high resolution on a wide spectral range from FUV to NIR (145–2100 nm). The light is analyzed by a grating monochromator that directs sequentially each individual wavelength into the detector.
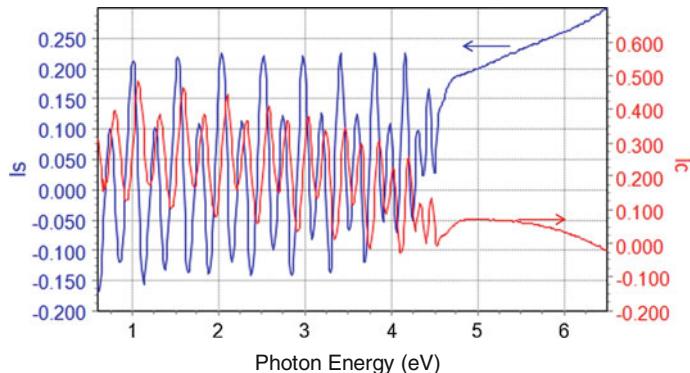
Fig. 19.3 Ellipsometric measurements of a sapphire/AlN/AlGaN sample

The Auto SE, the automatic version of the MM-16 ellipsometer, allows full automatic analysis of thin film samples with simple push button operation. The innovative liquid crystal modulation technology used provides an ellipsometer that is very fast, accurate, compact and simple to operate. It acquires full spectral ellipsometric data at high resolution and calculates the complete 16-element Mueller Matrix (see Chap. 2) in less than 2 s [4].
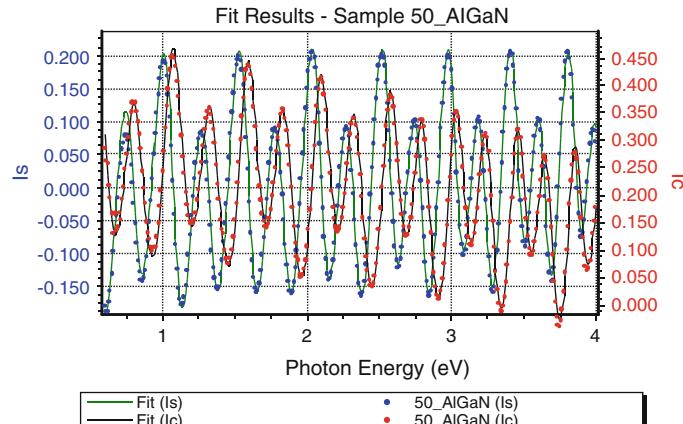
19.2.2.1 Example of Ex-Situ Characterization of a Multilayer Structure Used in a III-Nitrides-Based LED

Spectroscopic ellipsometry was used to determine the thickness and the optical properties of AlN and AlGaN layers in a Sapphire substrate/AlN/Al$_x$Ga$_{1-x}$N multilayers structure. The ellipsometric measurements were collected at an angle of incidence of 70° in the spectral range 0.6–6.5 eV (equivalent to 190–2066 in nm), and they are shown in Fig. 19.3.

Two regions can be clearly distinguished in the spectra:

- The spectral region from 0.6 to 4.5 eV, which corresponds to the transparent range of materials, exhibits interference fringes stemming from multiple reflections at the substrate/film interface. This spectral region enables accurate determination of film thicknesses.
- The region from 4.5 to 6.5 eV is characterized by the absence of interference fringes due to the Al$_x$Ga$_{1-x}$N semiconductor material absorption. Therefore, this spectral range mainly provides information on the sample surface. It also enables the calculation of the bandgap of the Al$_x$Ga$_{1-x}$N using the Tauc Plot method [5]. The boundary between those two regions ($\sim$4.5 eV) relates to the onset of absorption of the Al$_x$Ga$_{1-x}$N layer, and we can determine a bandgap of approximately

Overlayer	2.1 nm
$Al_xGa_{1-x}N$	570.3 nm
AlN	669.5 nm
Sapphire substrate	330 μm

Fig. 19.4 Best-fit model with thicknesses results

Fig. 19.5 Fitting results of the multilayer structure sapphire/AlN/AlGaN

4.5 eV, from which information on the $Al_xGa_{1-x}N$ composition can also be inferred.

The model depicted in Fig. 19.4 has been used to fit the experimental data in Fig. 19.3. A roughness surface layer, modeled by 50 % $Al_xGa_{1-x}N$ + 50 % void, improves significantly the goodness of the fit (described by the χ^2 value).

Figure 19.5 shows the excellent agreement ($\chi^2 = 0.77$) between the experimental data and the corresponding fit in the spectral range 0.6–4.0 eV dominated by the interference system.

The fitting process determines simultaneously the AlN and AlGaN thicknesses and the corresponding optical constants. Parameterization of the optical constants of the AlN and $Al_xGa_{1-x}N$ layers is needed because they are crystallinity and composition dependent. Therefore, the optical constants have been determined using dispersion formulae, specifically Lorentzian oscillators, which are included in the DeltaPsi2™ software, expressed as:

$$\varepsilon = 1 + \frac{(\varepsilon_s - 1)\,\omega_t{}^2}{\omega_t{}^2 - \omega^2 + i\Gamma_0\omega}$$
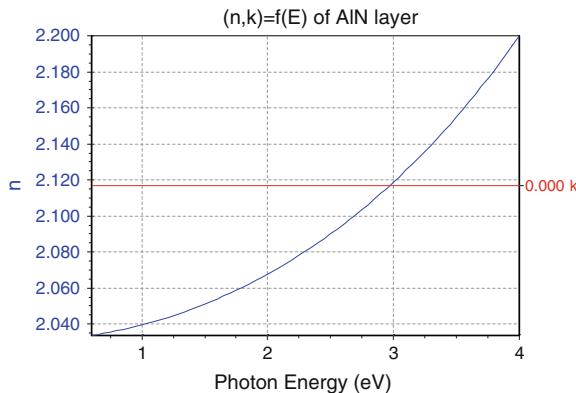
Fig. 19.6 Optical constants of the AlN layer
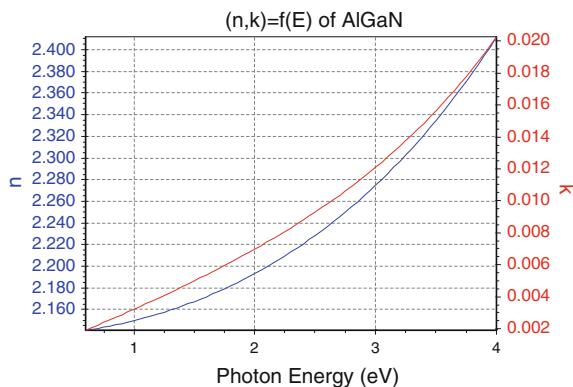
Fig. 19.7 Optical constants of the $Al_x Ga_{1-x}N$ layer

The Lorentz formula is based on the classical theory of the interaction between light and matter and it is used to describe the frequency dependent polarization due to bound charges. ω_t represents the energy of the resonant oscillator representing the main optical transition and, hence, absorption of the material. Γ_0 (in eV) is the broadening of the oscillator also known as the damping factor. For the AlN layer, which has a bandgap above 6 eV, the extinction coefficient, k, is equal to 0 in the analyzed range, implying $\Gamma_0 = 0$. For the $Al_x Ga_{1-x}N$ material, k increases with the increase of energy. In this case, the Γ_0 parameter has a value different from 0.
The optical properties calculated for the AlN and $Al_x Ga_{1-x}N$ layers are shown in Figs. 19.6 and 19.7:
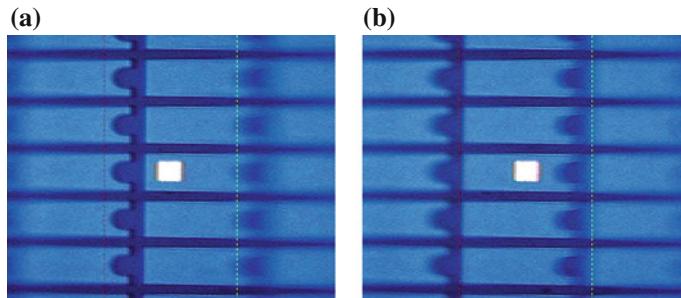
Fig. 19.8 Ellipsometric beam inside the LED pattern: **a** *left* part of the pattern; **b** *right* part of the pattern

19.2.2.2 Example of In-Line Process Control of LED Production

The transfer of the measurement technique from the laboratory scale up to industrial production scale under fast moving inline conditions represents a large challenge particularly in terms of determining the optical properties of a complex layer stack. This example illustrates the use of ellipsometry for quality control of a GaN based light-emitting diode production process. The structure includes a thick photoresist layer used to form an etching mask on top of GaN deposited on a Sapphire substrate. Ellipsometry is used regularly in the optical characterization of new photoresists and for process control of photoresists in production lines.

The Auto SE ellipsometer was used to characterize the thin film structure directly inside a pattern in the wavelength range 440–850 nm. Using the MyAutoView vision system coupled to the microspot of the Auto SE, it is a straightforward procedure to easily locate the measurement spot on the sample, as shown in Fig. 19.8.

Using the spot size $100 \times 100\,\mu$m, different locations within the pattern can be measured, as shown in Fig. 19.8, with the corresponding ellipsometric measurements reported in Fig. 19.9. The ellipsometric measurements overlap quite well, indicating that the layers were deposited uniformly.

The model used to describe the sample is represented in Fig. 19.10, and it also takes into account the anisotropy of the sapphire substrate. Specifically, the sapphire, Al_2O_3, substrate is considered uniaxial and the ordinary (O) and extraordinary (E) reference dielectric function present in the software library are used. The model includes a rough GaN film that is modeled using two layers: a homogeneous bottom GaN layer of 7200 nm (layer 1), and a top rough graded GaN overlayer of 222 nm (layer 2) composed of a mixture of GaN and photoresist (indicated as PR), which simulates an interface layer that forms between the GaN active layer and the photoresist. Ontop the GaN there is the thick photoresist layer of $2.55\,\mu$m (layer 3), with a thin roughness layer (3.4 nm) at the surface of the photoresist film (layer 4). The determined thicknesses displayed in the model are expressed in nanometers.

(a) Is = f(Ê) - Two experimental data inside one pattern

(b) Ic = f(Ê) - Two experimental data inside one pattern

— Is – measurement 1
— Ic – measurement 1

— Is – measurement 2
— Ic – measurement 2
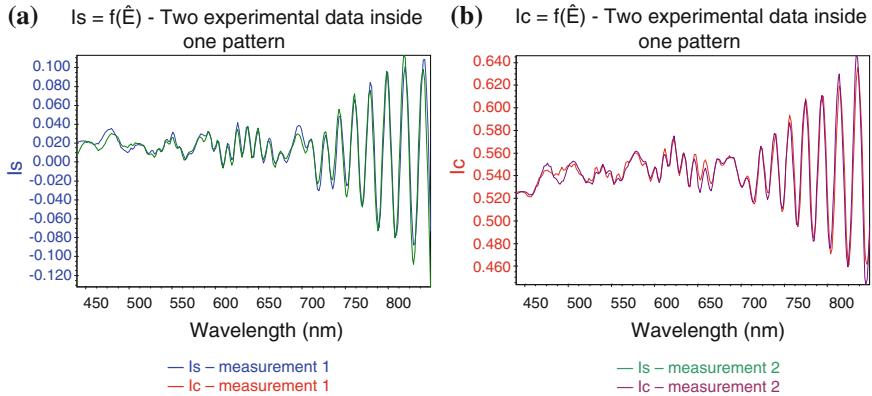
Fig. 19.9 Ellipsometric spectra taken inside the **a** *left* part of the LED pattern; **b** *right* part of the pattern, as in Fig. 19.8

Fig. 19.10 Model with thicknesses and layers' volume fractions (%) results

In the measured region the GaN can be considered transparent since its optical gap of 3.4 eV (365 nm) is outside the range. Therefore, the optical constants, i.e., the dielectric function, ε, of both the GaN and Photoresist layers have been modeled using a single Lorentz oscillator dispersion equation, given by

$$n^2 = \varepsilon = 1 + \frac{(\varepsilon_s - 1)\omega_t^2}{\omega_t^2 - \omega^2}$$

where n is the refractive index, and ωt is the frequency of the oscillator describing the main optical absorption transition.

Figure 19.11 shows the good agreement between the experimental data (dots) and the corresponding fit (line) with a fit goodness of $\chi^2 = 1.60$. It can be observed that the photoresist may have some absorption peak around 700 nm. For the purpose of thickness determination, this absorption could be neglected to keep the dispersion formula simple. Furthermore, a damping of the fringes is observed, which can be due to a significant roughness of the GaN layer.
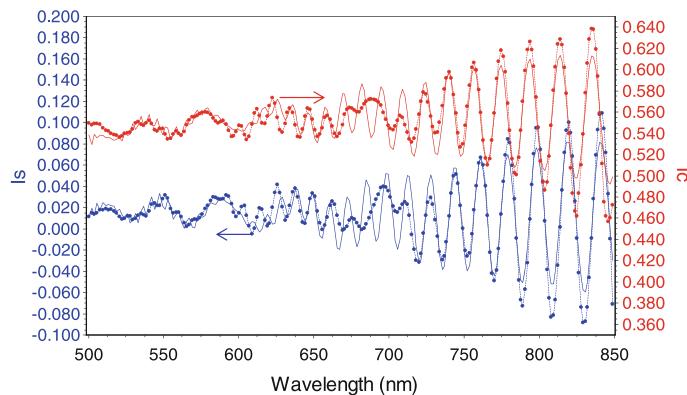
Fig. 19.11 Fitting results for the sapphire/GaN/Photoresist structure
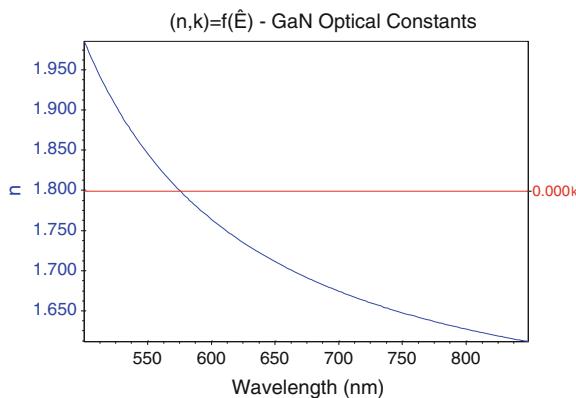
Fig. 19.12 Optical constants of the GaN layer

Figures 19.12 and 19.13 show the refractive index, n, determined, respectively, for the GaN layer and for the Photoresist layer for the two in-line measured spots in Fig. 19.8

19.3 Photovoltaics

19.3.1 Market Trends

The last decade has seen the photovoltaic (PV) technology emerging as a potentially major technology for power generation worldwide. Despite the economic crisis, the PV market has continued to grow by almost 15 % in 2009 compared to 2008 and
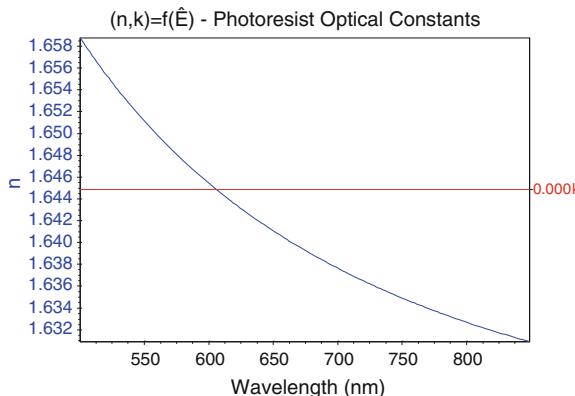
Fig. 19.13 Optical constants of the photoresist

the total power installed in the World raised by 45 % up to 22.9 GW. Compared to the growth rate for new photovoltaic installations in 2010, 2011 was not as strong for the industry. Newly added capacity in 2011 is estimated to have reached 23.8 gigawatts (GW), but 2012 is not expected to see as much growth. Predictions from market analysts say that the year should see around 25 GW installed. Overall, new sales markets and additional applications for solar energy mean that newly installed photovoltaic capacity is predicted to rise by an annual average of 18 percent up to 2015. The European Union contributes to over 60 % of the global cumulative capacity while Japan and the US are following behind. By 2020, photovoltaic electricity can become a mainstream energy source in Europe and cover up to 12 % of the electricity demand [6].

About pricing and technology, 2011 was characterized by rapid capacity growth combined with weaker than expected demand in some key countries. This led to a severe over supply of products throughout the supply chain which has forced prices down significantly throughout the year. There are already cases in Europe of modules selling for less than €0.80 per watt (/W), while in China, prices have been as low as €0.70/W. Overall, module prices reached €0.65/W in March 2012. The huge supply of attractively priced c-Si modules means that crystalline will continue to hold a majority share of the market. Specifically, crystalline silicon (c-Si) modules represent 85–90 % of the global annual market today. c-Si modules are subdivided in two main categories: (i) single crystalline (sc-Si) and (ii) poly-crystalline (pc-Si). Thin films currently account for 10–15 % of global PV module sales. Various thin-film technologies are currently being developed to reduce the mass of light absorbing material required to create a solar cell.

The thin film technologies shares is also growing (see Fig. 19.14) due to lower costs, and advantages such as flexibility, lighter weight, and feasibility of integration. The main light absorbing materials in thin film photovoltaic cells are:

- Cadmium telluride (CdTe)
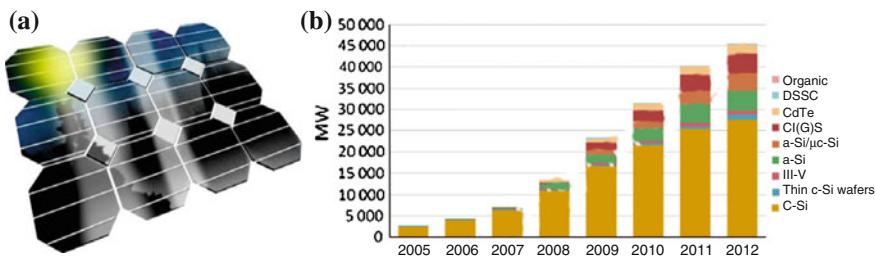
Fig. 19.14 **a** Example of photovoltaic module. **b** Trends over the years of the main thin film solar cells technologies

- Copper indium gallium selenide (Ci(G)S)
- Gallium arsenide multijunction
- Dye sensitized solar cells (DSSC)
- Organic materials
- Amorphous silicon (a-Si) and nanocrystalline silicon (nc-Si) or microcrystalline silicon (μc-Si)

Emerging technologies encompass advanced thin films and organic cells, which are about to enter the market via niche applications.

Novel PV concepts aim at achieving ultra-high efficiency solar cells via advanced materials and new conversion concepts and processes. They are currently the subject of basic research.

The performance of a solar cell is measured in terms of its efficiency at turning sunlight into electricity.

The efficiency/production cost relationship estimated in 2011 for the main thin films technologies listed above is shown in Fig. 19.15, indicating that major solution are required to improve efficiency and reduce cost of organic photovoltaics to become competitive.

With the aim of achieving further significant cost reductions and efficiency improvements, R&D is predicted to continuously progress in improving existing technologies and developing new technologies. It is expected that a broad variety of technologies will continue to characterize the PV technology portfolio, depending on the specific requirements and economics of the various applications. Figure 19.16 gives an overview of the different PV technologies and concepts under development.

19.3.2 Ellipsometry Characterization

In thin film solar cells, the thickness of each layer and its uniformity represent a critically important point in the evaluation of the overall performance of a multilayer PV stack. Furthermore, the photon energy dependence of the refractive index and absorption coefficient and their uniformity are also critically important, as these
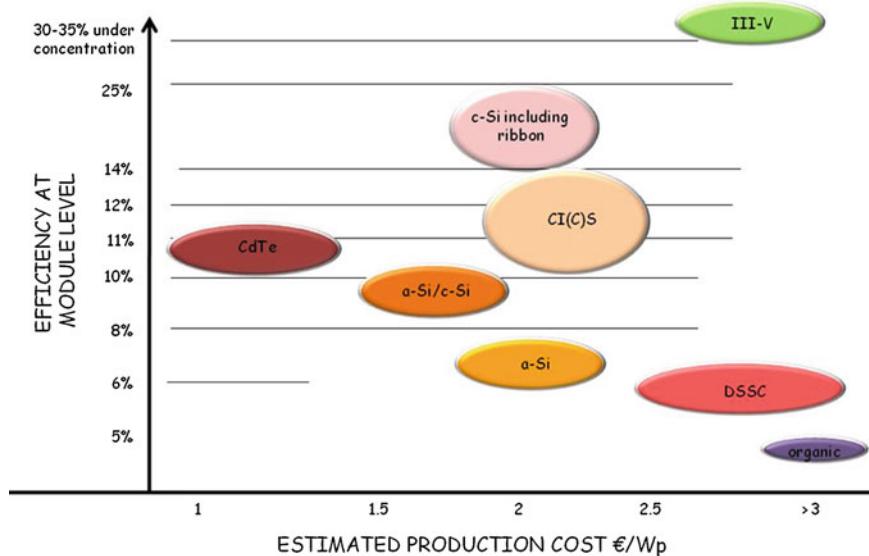
Fig. 19.15 Production costs and efficiency of the main thin film solar cells as at 2011 [*Source* Yole Development Report 2011].
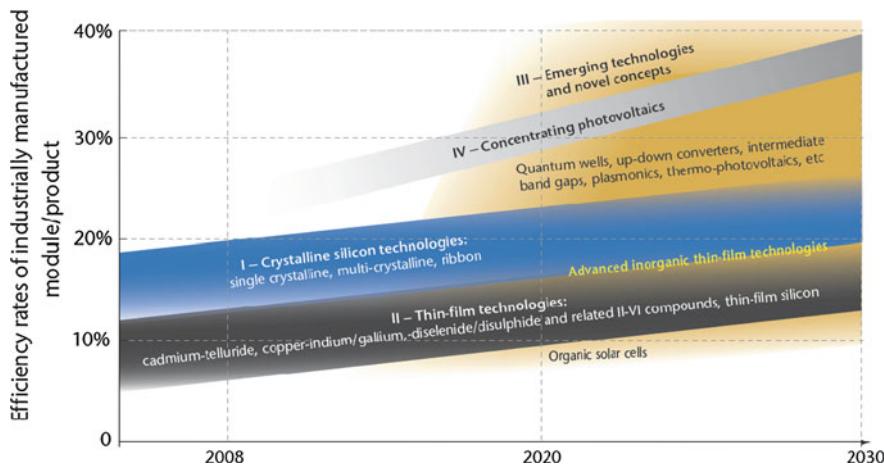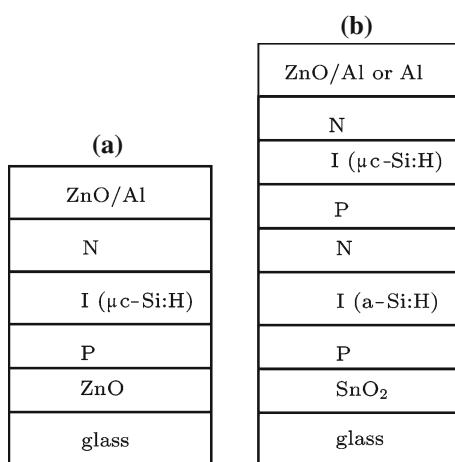

Fig. 19.16 Photovoltaic technology status and prospects [7]

spectra, together with thicknesses, enable the calculation of the optical quantum efficiency of the PV device. Analysis of the optical spectra of the dielectric function for each layer constituting a solar cell can also give fundamental optical parameters such as optical gap, density, composition, strain, grain size, and defect density. Therefore, SE is an ideal technique for measuring film thickness and roughness for multilayers solar cells.

Fig. 19.17 Scheme of the sin-
gle junction microcrystalline
silicon solar cell (**a**) and the
amorphous and microcrys-
talline silicon tandem solar
cell (**b**)

(a)

ZnO/Al
N
I (μc-Si:H)
P
ZnO
glass

(b)

ZnO/Al or Al
N
I (μc-Si:H)
P
N
I (a-Si:H)
P
SnO$_2$
glass

Here applications are described in the two major thin film technologies of micro-
crystalline silicon and of ZnO layers used as transparent conductive oxide (TCO) to
replace the common indium tin oxide (ITO).

Transparent Conductive Oxides (TCOs) are an essential part of thin-film silicon
solar cells. They have to exhibit good electrical (high conductivity) and optical (high
transmittance) properties. In addition to these characteristics, they also have to scat-
ter light at the TCO/cell interface in order to increase the effective absorption of
light within the active layer of the cell. This aspect is especially important in the
case of microcrystalline silicon (μc-Si:H) thin-film solar cells, because of their rel-
atively low optical absorption coefficient in the red and near-infrared (NIR) spectral
range. Hydrogenated microcrystalline silicon (μc-Si:H) has attracted great attention
for use as a long wavelength absorbing material for the bottom cell in double-junction
thin film solar cells, where the top cell is normally hydrogenated amorphous silicon
(a-Si:H). This cell structure is normally called 'micromorph' tandem solar cell. The
schematic structure of solar cells coupling ZnO with μc-Si:H are shown in Fig. 19.17.
The thickness of the microcrystalline bottom solar cells as well as of the top amor-
phous cell and of the back reflector are important to optimize the performance of the
tandem solar cells.

Therefore, ellipsometric characterization aims at establishing optical properties for
all the major thin film PV material components at a research level, while at a manu-
facturing level the aim is the analysis of thickness, optical properties and uniformity
over full scale panels.

Measurements are performed using the HORIBA Jobin Yvon UVISEL spectroscopic
phase modulated ellipsometer. Ellipsometric data were acquired at an angle of inci-
dence of 70°, across the extended spectral range from 0.6 to 6.5 eV (190–2100 nm).
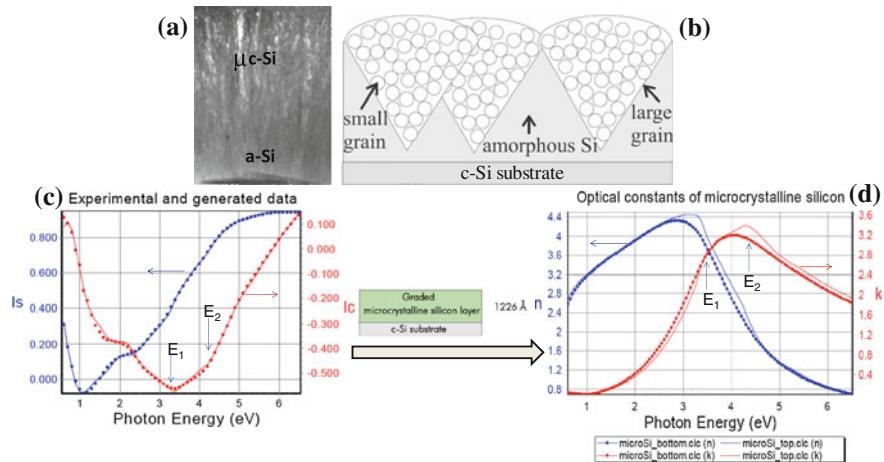
Fig. 19.18 a Transmission electron microscopy image of a μc-Si layer. **b** Scheme showing the gradient in crystallinity from the *bottom* amorphous layer to the *top* microcrystalline layer that characterizes μc-Si. **c** Experimental data and **d** fit results according to the sketched model with thicknesses results of a microcrystalline silicon layer. The derived graded optical constants of the microcrystalline silicon are also shown in (**d**)

19.3.2.1 Example of Characterization of Microcrystalline Silicon

This example has been chosen because it allows to discuss the description of a non-homogeneous material that also does not have a standard reference dielectric function, since the optical properties of μc-Si depend on the grain size and on doping level. Therefore, a parameterization of the optical properties has to be used. Microcrystalline silicon layers, generally composed of nanometer-scale crystalline Si grains with amorphous silicon phase at the grain boundaries, have been applied widely to fabricate a-Si:H/μc-Si tandem solar cells as shown in Fig. 19.17. The microcrystalline silicon layer is inhomogeneous in depth and, therefore, it exhibits an optical gradient (Fig. 19.18) that can be accurately detected by spectroscopic ellipsometry. Figure 19.18 shows the experimental ellipsometric spectra acquired for a μc-Si layer grown on a crystalline Si substrate, the optical model used to analyze spectra and the results of the analysis. The model includes a graded layer used to specify different values of crystallinity at the bottom and at the top of the microcrystalline layer. The optical constants of the microcrystalline layer were parameterized using a Tauc-Lorentz dispersion formula [8].

To minimize the numerical parameters, the microcrystalline silicon dielectric function is modeled by considering three optical transitions of the E_1, E_2 critical points of silicon and of the fundamental bandgap. Figure 19.18 shows the good agreement between the experimental spectra (dots) and the model calculations (lines) on the whole spectral range with a fit goodness $\chi^2 = 0.81$. Noteworthy, analyzing the shape of the determined optical constants for the microcrystalline layer, the bottom

roughness	27 Å
ZnO	508 Å
82.2% ZnO + 17.8% void	302 Å
c-Si substrate	

layer of the gradient (dots in Fig. 19.18d) shows in the extinction coefficient, k, spectrum a single broad peak due to the formation of the a-Si:H phase, as indicated by the scheme and TEM picture in Fig. 19.18. At the top of the gradient, the appearance of two sharp features at 3.4 and at 4.2 eV for the top layer (line in Fig. 19.18d) indicates that the grains of microcrystalline silicon become larger towards the top surface, as confirmed by the transmission microscopy analysis. These two optical transitions E_1 and E_2 arise from the direct interband transitions: specifically, the E_1 critical point at 3.4 eV is due to the transition from Λ_3 (valence band) to Λ_1 (conduction band) in the first Brillouin zone, whereas the E_2 critical point at 4.3 eV originates from many transitions along the X and Σ directions (see also Chaps. 16 and 17 of this book). Thus, as a result of the ellipsometric analysis, the thickness and non-homogeneous optical constants of the microcrystalline layers can be determined.

19.3.2.2 Example of Ellipsometric Characterization of Transparent Conductive Electrodes (TCO)

Among all TCOs studied recently, zinc oxide (ZnO), which has an optical band gap of 3.4 eV, has emerged as one of the most promising materials due to its optical and electrical properties, its high chemical and mechanical stability and, owing to its abundance in nature, it is a lower cost material compared to the commonly used indium tin oxide (ITO). Here we provide the example of the ellipsometric analysis of the optical properties of a ZnO layer deposited on c-Si.

Crystalline ZnO is an anisotropic uniaxial material with a hexagonal wurtzite crystal structure. However, most of the ZnO thin films for solar cells grown by sputtering or plasma enhanced chemical vapor deposition (PECVD) are polycrystalline, although highly preferentially oriented, and hence most of the time do not show significant anisotropy. Therefore, an isotropic model is sufficient to describe the ZnO optical constants. On the other hand, often due to the deposition process and interaction/diffusion of oxygen, ZnO thin films may present a gradient in density and/or conductivity. In order to represent the in depth non-homogeneity of the ZnO layer), a graded layer and/or a three layer model is used (Fig. 19.19). For many ZnO samples the density increases from the c-Si interface (1st layer) to the top (2nd layer); the sample also exhibits a surface roughness. Therefore, the modeling of SE spectra was performed using a multilayer and multicomponent Bruggemann effective medium approximation (BEMA) (ZnO + voids); the dielectric function, ε, of the ZnO was parameterized by a dispersion equation that combines the Drude model to describe the contribution of the free carrier concentration to the ZnO conductivity, with a
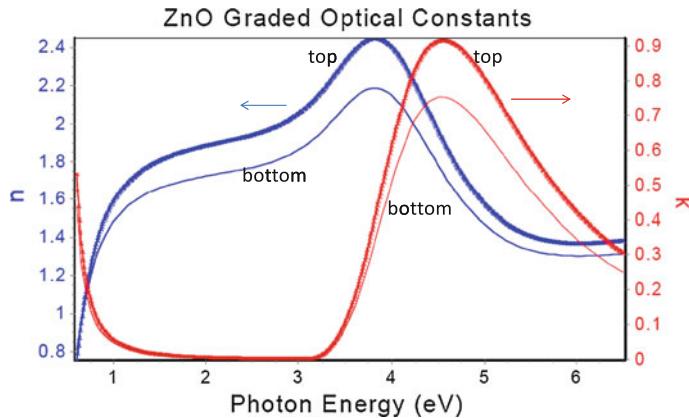
Fig. 19.20 Refractive index, n, and extinction coefficient, k, of ZnO *bottom* (*thin line*) and *top* (*thick line*) layers

double Lorentzian oscillator to describe the contribution of the band edge and of the interband transition, i.e.,

$$\varepsilon(E) = \varepsilon_\infty - \frac{\hbar^2 \omega_p^2}{(\hbar\omega)^2 + i(\hbar\omega_\tau)(\hbar\omega)} + \sum_{i=1}^{2} \frac{f_i(\hbar\omega)}{(\hbar\omega_i)^2 - (\hbar\omega)^2 + i\gamma_i(\hbar\omega)}$$

where ε_∞ is the high-frequency dielectric constant, ω_i is the resonance frequency, γ_i is the damping factor of the ith Lorentzian oscillator with strength f, while ω_p is the plasma frequency and ω_τ the collision frequency in the Drude model. The refractive index, n, and extinction coefficient, k, as obtained from fitting the spectra are shown in Fig. 19.20. The k-spectra for the bottom and top layers shows an onset of the absorption at 3.3 eV and a characteristic peak around 4 eV, which correspond to the absorption edge of the direct band gap semiconductor ZnO. The difference between the onset and peak energies observed here and that reported for single crystalline ZnO may be interpreted in terms of the Burstein-Moss shift, which is a blue-shift of the bandgap with increasing doping level, after partially filling the conduction band of ZnO with electrons. The contribution of those electrons to the ZnO conductivity is seen in the increase of k at the low-energy side of the range, which is described by the Drude theory.

Ultimately, the ellipsometric analysis and the Drude component enables determination of the scattering time, resistivity, and sheet resistance of ZnO layers, providing a contactless tool for control in-line of ZnO conductivity.
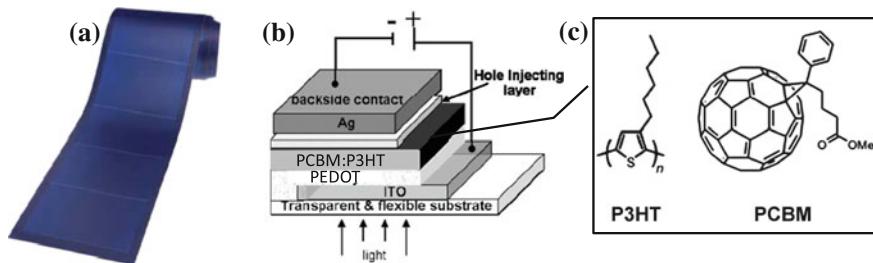
Fig. 19.21 a Flexible solar cell. **b** Scheme of a flexible organic solar cell. **c** Formula of the organic P3HT donor and PCBM acceptor in the OPV active layer

19.3.2.3 Example of Characterization of Flexible Solar Cells

Lately, research efforts have moved toward the production of flexible substrates to take advantage of the semiconducting, conducting, and light-emitting properties of organics (polymers, oligomers) and hybrids (organic-inorganic composites) through novel synthesis and self-assembly techniques. The next technological revolution will be the mass production of plastic-based, printable organic semiconductors on flexible substrates (Fig. 19.21), which will further reduce production prices and open new fields of applications.

The use of flexible substrates offers significant advantages such as light weight, thinness, robustness, and the ability to be rolled. In addition, plastic-based substrates-coupled with new and smart deposition processes-open up the possibility of cost-effective and high volume roll-to-roll processing, which is attractive to optical-coating, and solar-module industries where cost-efficient production and high-quality transparent coatings are the drivers.

Efficiency and lifetime have reached commercially acceptable levels and flexible modules with roughly 2 % power conversion efficiency, and over 1 year lifetimes are already on the market. Mitsubishi Chemical Holdings have attained laboratory-scale cell efficiencies of 10 % [9]. Among the thinnest reported substrates for organic photovoltaics (OPV) devices are 125-μm plastic foils. Indeed, to answer the questions "how light?" and "how flexible"? OPV can be, recent work has demonstrated ultrathin, flexible and compliant OPV devices constructed on only 1.4 μm thick poly-ethylene terephtalate (PET) substrates with a 4.2 % power conversion efficiency [10]. Accurate control of the film thickness is mandatory to achieve a fabrication process that gives repeatable film properties and remains cost-effective. Because of the nature of the data provided by SE, and its ability to provide them quickly, ellipsometry is a preferred nondestructive thin-film metrology technique for fast multiple film-thickness and refractive-index measurements.

Organic polymer solar cells are built from thin films (typically 100 nm) of organic semiconductors such as polymers and small-molecule compounds like polypheny-lene vinylene, copper phthalocyanine (a blue or green organic pigment) and carbon fullerenes and fullerene derivatives such as PCBM. Energy conversion efficiencies

achieved to date using conductive polymers are low compared to inorganic materials, with the highest reported efficiency approaching 10 % [10]. However, these cells could be beneficial for some applications where mechanical flexibility and disposability are important.

These devices differ from inorganic semiconductor solar cells in that they do not rely on the large built-in electric field of a PN junction to separate the electrons and holes created when photons are absorbed. To achieve efficient organic solar cells, donor and acceptor molecules are mixed in the photoactive layer to form a so-called bulk heterojunction. Due to molecular interactions, a certain degree of phase separation between donor and acceptor domains arises, which is necessary to achieve efficient charge extraction within the absorber layer. However, the mechanism that induces the phase separation is not fully understood and gaining detailed information about the molecular arrangement within these blend layers is quite challenging. Therefore, we show the use of spectroscopic ellipsometry to infer information about composition and phase separation of organic blend layers.

The active region of an organic device consists of two materials (see Fig. 19.21b), one which acts as an electron donor and the other as an acceptor. When a photon is converted into an electron hole pair, typically in the donor material, the charges tend to remain bound in the form of an exciton, and are separated when the exciton diffuses to the donor-acceptor interface. The short exciton diffusion lengths of most polymer systems tend to limit the efficiency of such devices. Nanostructured interfaces, sometimes in the form of bulk heterojunctions, can improve performance. The most successful OPV active material system consists of a bulk heterojunction that is formed by the p-type semiconductor (electron donor) poly(3-hexylthiophene) (P3HT) with the n-type semiconductor (electron acceptor) [6,6]-phenylC61-butyric acid methyl ester (PCBM), as schematized in Fig. 19.21. The distribution of the constituents in the blend film plays an important role for efficient charge extraction toward the electrodes. Ideally, more p-type material (polymer) should be located at the interface of the hole collecting anode and more n-type (fullerene) material should be at the electron collecting cathode facilitating collecting of charges from the active layer to the electrodes. Concerning P3HT:PCBM blends deposited by spin coating, a lateral phase separation can take place leading to the formation of PCBM clusters. Vertical phase separation can also occur due to the surface energy difference of the two components and of the surface energy of the substrate. The distribution of the P3HT and PCBM can dramatically affect the final OPV device performance. Therefore, it is important to detect non destructively and or infer the distribution of the polymer and fullerene constituents.

We demonstrate here the use of ellipsometry as a powerful and sensitive metrology for monitoring blend morphology, phase separation as well as crystallinity.

The UVISEL spectroscopic phase modulated ellipsometer has been used to characterize three organic samples including P3HT/c-Si, PCBM/c-Si and P3HT:PCBM/c-Si. Ellipsometric measurements were performed at an angle of incidence of 70°, across the spectral range 190–2100 nm.

The goal for the characterization of the first two samples was to find the film thickness and optical constants of P3HT and PCBM materials. As shown in Figs. 19.22 and
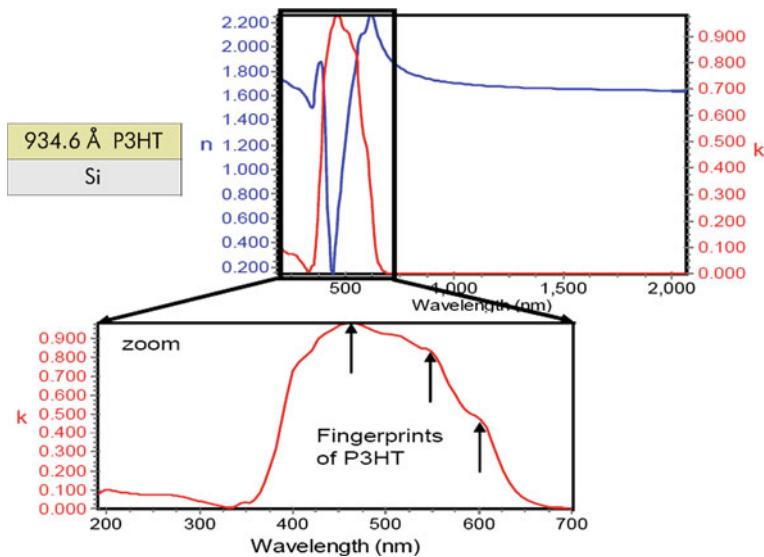
**Fig. 19.22**   Fit model and optical constants of P3HT
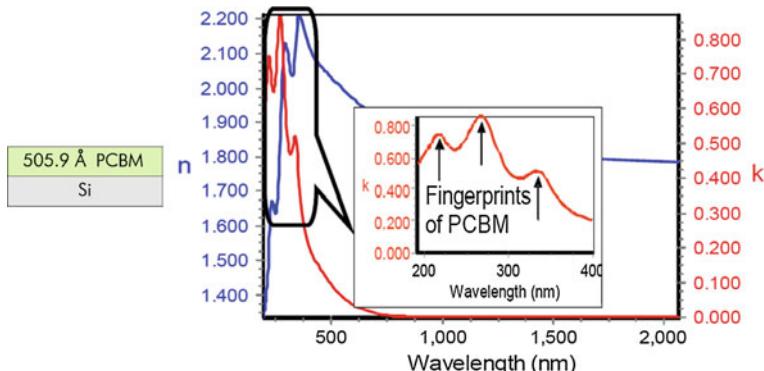
Fig. 19.23   Fit model and optical constants of PCBM

19.23, a simple isotropic one-layer model was used to derive the P3HT and PCBM optical constants modeled using a dispersion formula.

Two different analysis protocols were used to model the P3HT:PCBM blend.

The first protocol analysis is based on the use of the effective medium theory (see Chap. 3 of this book), allowing to mix pure P3HT and PCBM in a layer to represent the blend P3HT:PCBM structure (Fig. 19.24). In the second protocol, the blend is considered as a, homogenous material, represented optically through a dispersion equation (Fig. 19.25).
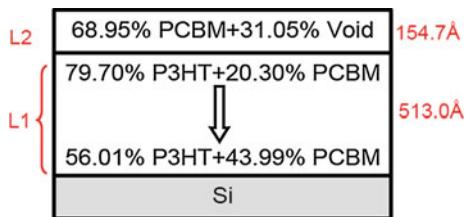
Fig. 19.24 Best-fit model based on EMA for the analysis of the P3HT:PCBM blend with thicknesses results
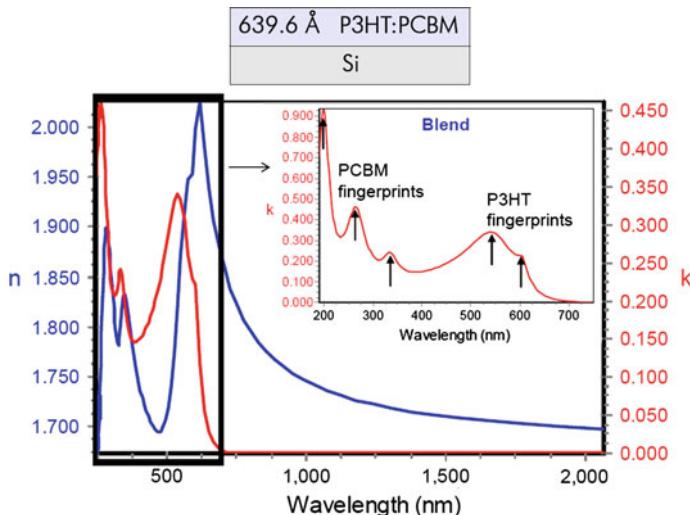
Fig. 19.25 Model based on the use of a dispersion model with thicknesses results and derived P3HT:PCBM Blend Optical Constants. The inset shows the fingerprints of P3HT and PCBM visible in the extinction coefficient, k, measurements

Analysis protocol I: use of the effective medium theory

The model used to represent the blend P3HT:PCBM is an effective medium approximation, i.e. a mixture of the two constituents, as shown in Fig. 19.24. The best-fit is based on a bi-layer model. The layer 2 (L2) represents the surface roughness layer consisting of PCBM and void (i.e., air). The layer 1 (L1) is a linear gradient, with P3HT concentration increasing whereas PCBM decreasing towards surface layer. The individual P3HT and PCBM optical constants are fixed to those previously determined using the previously discussed samples 1 and 2, whereas the volume fraction percentages are varied to model blend index changes. Therefore, the fit parameters are the volume fractions of the P3HT and PCBM in the graded layer, i.e., at the bottom and at the top of the gradient, and the thickness of the graded and surface layer.

This analysis protocol results in detecting gradients in the phase separation of the two components.

The best-fit model in Fig. 19.24 shows that those blends are characterized by donor-enriched top regions (high P3HT content), consistently with literature reports on this OPV system.

Analysis protocol II

The P3HT:PCBM blend has been approximated as a homogenous material, represented optically through one single new dispersion function. Specifically, the blend optical response has been modeled by the Tauc-Lorenz (TL) oscillator model using five TL oscillators to describe the interband absorptions above the energy band gap of the blend. Noteworthy the fingerprints of constituents are visible in effective optical spectra of P3HT:PCBM blend. As it can be seen in Fig. 19.25, the optical response of the blend films in the VIS-UV spectral range includes five optical absorption peaks. The first optical absorption at 2.05 eV to the singlet excitonic transition of the P3HT conjugated polymer, whereas the transition at 2.24 eV is due to the coupling of excitons with phonons. The other three peaks at higher energies are originated from the PCBM. The advantage of the latter protocol II is that it provides insights into blend film composition and crystallinity through the monitoring of fingerprints of each component (peak amplitude, broadening, etc.)

Thus, for the P3HT:PCBM blend layer, protocol I and protocol II are complementary and allow accurate characterization of the variations of the material composition as well as optical constants.

19.4 Information and Communication Technologies

19.4.1 Memories and Data Storage

The data storage industry is currently growing at a healthy rate. A large range of memory technologies is developed for data storage (Fig. 19.26) including non-volatile, flash, molecular, resistive and magnetoresistive, and super resolution optical data storage. Hard disk drives (HDD) and solid-state storage media, such as flash memory USB devices and cards, are facing very high growth (Fig. 19.26). The demand is so high that there might be a shortage of Solid State Drives (SSD), which unlike HDD contain no moving parts and instead use flash memory and are therefore less susceptible to physical shock and vibration than HDD.

Regarding resistive memory, the information is stored as a variation of electrical resistivity of a material, without involving the storage of electrical charges. The different types of physical mechanisms that are used for varying the memory resistivity utilize different materials: fast phase-changing material for Phase-Change RAM (PCRAM) or metal oxides for memory (OxRAM).

Phase change memory is a term used to describe a class of nonvolatile memory (NVM) devices that exploit the ability of certain materials to rapidly change phase,
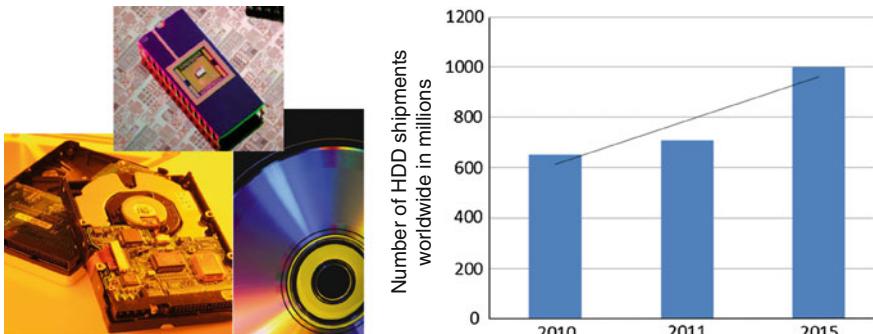
Fig. 19.26 Hard disk drive, example of "Non-Volatile Memory Chips" and CD storage devices and total worldwide HDD shipments increased by 16.9 % over 2009, to 651.32 million units in 2010, and this is expected to rise to 708.81 million this year and go to close to 1 billion units by 2015

between two stable physical states, e.g., localized laser induced heating of a thin layer to cause a phase transition from the crystalline to amorphous state.

This transformation results in optical reflectance differences detectable by spectroscopic ellipsometry.

The large optical contrast between the amorphous and crystalline states is the key property for optical read/write data storage media and has been applied widely in compact disk (CD), digital versatile disk (DVD), and Blu-ray disk technologies.

Although the optical contrast of phase change materials has found wide applications in optical storage media, the large change in electrical resistivity accompanying the transition from the amorphous to the crystalline state is only now becoming available in novel nonvolatile random access memory devices. Such phase change random access memory (PCRAM) devices have the potential to compete with flash memory as nonvolatile storage media and have switching speeds high enough to replace dynamic random access memory (DRAM) devices for use in general computing. Consequently, PCRAM could become a key enabling technology for universal data storage devices, combining roles as fast-switching DRAM and nonvolatile memory [11].

Phase change materials have to meet several requirements to be suitable for data storage applications. These properties can be summarized as follows:

- large optical contrast between the amorphous and crystalline states (optical read/write memory),
- large resistance change between the amorphous and crystalline states (PCRAM),
- high-speed phase transition (fast switching),
- long thermal stability of the amorphous phase (data retention),
- several possible cycles between the amorphous and crystalline states (device integrity), and
- excellent chemical stability.

The PCRAM technology uses a class of materials known as chalcogenides, which are alloys that contain an element in the oxygen/sulphur family of the periodic table.
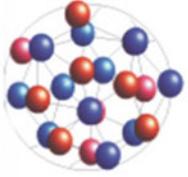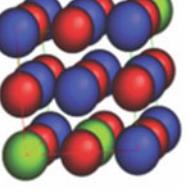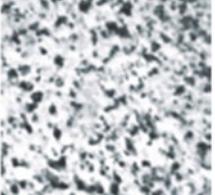
Fig. 19.27 Phase change chalcogenides exhibit a reversible phase change between the amorphous phase and the crystalline phase

Germanium-Antimony-Tellurium, GeTe-Sb$_2$Te$_3$, commonly referred as Ge-Sb-Te (GST) alloys are the phase-change materials of choice for optical memory devices used in rewritable optical memory devices such as CDs and DVDs.

As illustrated in Fig. 19.27, in the amorphous phase, there is an absence of regular order to the crystalline lattice. In this phase, the material demonstrates high resistivity and low reflectivity. In contrast, in the polycrystalline phase, the material has a regular crystalline structure and exhibits high reflectivity and low resistivity.

Spectroscopic ellipsometry is able to determine the key dimensional and optical parameters for accurate production control of PCRAM devices on multiple wafers in a single batch run.

This example discusses automatic multi-measurements (9-points) acquired according to a pre-set mapping grid, as shown in Fig. 19.28 and analyzed according to the model also schematized in Fig. 19.28 to check for homogeneity in thickness and optical properties.

The three layer model used (Fig. 19.28) includes a SiO$_2$ layer, a GeSbTe layer and a rough GeSbTe overlayer (thicknesses are in Angstroms) on a substrate of crystalline silicon (a reference file present in all ellipsometry softwares was used for c-Si) . The rough overlayer is modelled by a mixture of 50 % voids and 50 % GeSbTe using the effective medium approximation.

The Tauc Lorentz formula [8] was used to model the optical constants of the GeSbTe layer. The fitting parameters were the GeSbTe film thickness, the surface roughness thickness and the Tauc-Lorentz parameters. The determined optical constants are shown in Fig. 19.29, which put in evidence an amorphous state for GeSbTe.

3	F	29.7 ☐ %	F GST-tl_t.dsp	50.00 % ×	☐ void.dsp	50.00 % × ×
2	F	1412.7	F GST-tl_t.dsp			× ×
1	F	994.5	☐ sio2.dsp			× ×
			C-si_isa.ref			×
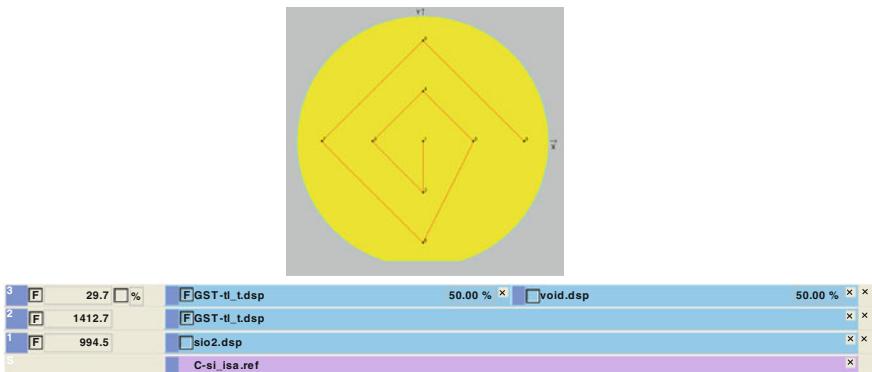
Fig. 19.28 Automatic nine points mapping grid acquisition and model used to fit spectra in the various points with thicknesses results
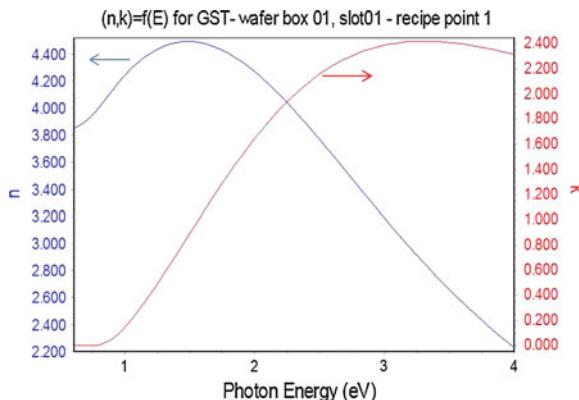

Fig. 19.29 Optical constants of refractive index, n, and extinction coefficient, k, derived for GST

The thickness determined in the various points by the three-layers model are summarized in the table in Fig. 19.30.

19.4.2 RF and Antennas Components

19.4.2.1 Example of Characterization of PZT Thin Films

In recent years, we have become accustomed to using radiofrequency (RF) wireless technology for telephone calls, internet access and navigation. Going forward, RF will be a basic enabling technology for multiple applications, from high-speed

Step	X	Y	L1 Thickness [Å]	L2 Thickness [Å]	L3 Thickness [Å]
1	0.000	0.000	1004.193	1421.032	33.754
2	0.000	-40.000	994.180	1363.040	33.824
3	-40.000	0.000	980.985	1361.749	33.238
4	0.000	40.000	989.311	1357.792	31.663
5	40.000	0.000	995.164	1359.618	32.947
6	0.000	-80.000	987.104	1225.173	36.590
7	-80.000	0.000	975.974	1225.283	32.078
8	0.000	80.000	984.611	1215.035	35.194
9	80.000	0.000	992.662	1224.952	31.957

Fig. 19.30 Mapping results of the thickness on the nine points of the wafer as in Fig. 19.27

data transmission to networks of tiny autonomous sensors and in-body therapeutic devices. Within the radiofrequency microelectromechanical systems (RF MEMS) research community, there is great interest in RF MEMS switches for creating low loss phase shifters for phased array antennas for communication and radar systems. In this field, the material of interest is lead zirconate titanate (PZT) thin films for improved device performance and actuation voltages compatible with most CMOS electronics.

Recently, the use of ferroelectric thin films of the $PbZr_xTi_{1-x}O_3$ (PZT) family of functional materials in microsystems technology has drawn remarkable interest [12]. PZT films provide a direct electro-mechanical coupling and are useful for memory, piezoelectric and pyroelectric devices . PZT thin films have advantage over bulk materials: they can be directly deposited on platinized silicon to allow direct integration with electronics. Due to the applications of PZT in pyroelectric devices, studies on PZT stability with changes in temperature are very important. For actuation, the most significant PZT material property of interest is the piezoelectric stress constant for which PZT has an effective, transverse piezoelectric stress constant nearly an order of magnitude larger than AlN [13]. The ferroelectric and dielectric properties of PZT have been extensively investigated, while few studies exist on their optical properties. Indeed, the optical constants, e.g., the refractive index and extinction coefficient, have great importance for waveguiding.

Every application of MEMS requires a different thickness of the functional film, leading to the challenge of manufacturing film of the required thickness. Furthermore, the broad applications of PZT films, and *inter alia* the growing interest in graded refractive index films for applications in optical devices, make it imperative to study the depth profile of optical properties of thin films throughout a single layer and an entire coating. Additionally, information on the homogeneity of the films and the physical properties resulting from different processing methods represents crucial knowledge, especially because gradients in optical properties have been reported for sputtered PZT films, and chemical composition gradients were reported for some
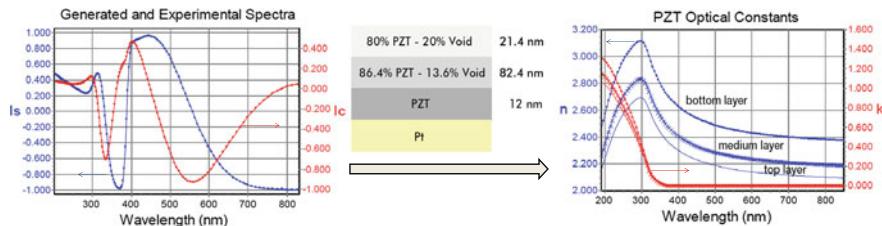
Fig. 19.31 Experimental spectra (*dots*) and fit results (*line*) according to the sketched model with thicknesses results yielding the graded optical properties of refractive index, n, and extinction coefficient, k, for PZT shown in the right panel

PZT films [14].Variation in chemical composition throughout the film thickness due to the inhomogeneity results in variation in physical properties of the films and lowers the performance of electromechanical systems. Thus, knowledge of the compositional gradients within a film allows identification and further optimization of thin-film performance and applications in piezo- and ferro-devices. Therefore, here we present ellipsometry for determining thickness and optical gradient in PZT films.

The phase modulated spectroscopic ellipsometer (PMSE) was to determine the optical constants of PZT materials, used as PZT actuator for RF switches in mobile phone applications. Ellipsometric measurements were collected at an angle of incidence of 70° across the spectral range 190–850 nm. PZT thin films were grown directly on Platinum 100 nm (Pt)/Ti/SiO$_2$/c-Si substrate.

Due to the Pt thickness, a pseudo 3-phase model—air/PZT/Pt—was used (Fig. 19.31). The PZT coating has been described using a 3 layers stack (Fig. 19.31) where the density of the PZT decreases from the bottom to the top surface. The model assumed that the PZT sample consists of three layers on the Pt substrate, i.e., (1) Interface layer, (2) PZT film, and (3) top layer with surface roughness—an effective medium approximation mixture of the void (air) and the PZT material. The PZT optical constants have been determined using a dispersion formula. Specifically, the Cody–Lorentz CLO and Lorentz oscillators were applied to characterize the complex dielectric functions of the PZT thin films and Pt thin film. The explanation of motivation of choosing these oscillators to describe the complex dielectric functions of our materials can be found in Ref. [15].

Experimental data and model fit are presented in Fig. 19.31, where film is considered to be isotropic.

This non homogeneity in depth results in graded optical constants as shown in Fig. 19.31.

The variation in refractive index with thickness for PZT is more likely due to different compositions, such as a variation in Zr/Ti ratio and/or directional increase in Pb from the substrate to the film surface accompanied by high stress. Therefore, the detection of this non-homogeneity by ellipsometry can be used to reveal the influence of sample preparation conditions on film homogeneity.
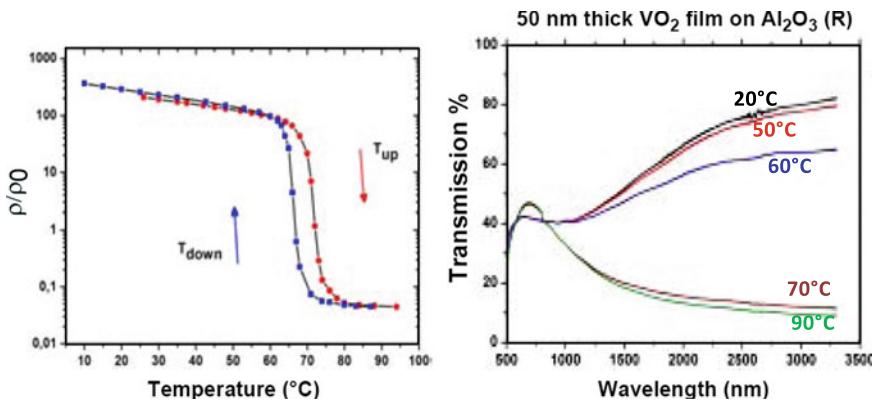
Fig. 19.32 VO$_2$ SMT transition: electrical (*left*) and optical (*right*) properties changes upon temperature variation

19.4.2.2 Example of Characterization of VO$_2$ Thin Films

The development of new types of ultra-fast switches operating in the RF-microwave and optical domains is based on the use of a class of materials undergoing fast, reversible phase transitions from a semiconducting state to a metallic one (Semiconductor Metal Transition / SMT). An example of such a material is vanadium dioxide (VO$_2$).

VO$_2$ exhibits very fast SMT transition that can be triggered by different external excitations such as temperature change, optical excitation or charge injection. During the phase transition the electrical resistivity of the VO$_2$ thin film can decrease by several orders of magnitude. Optically the material is transparent in the semiconductor state and highly reflective in the metallic state for a large spectral range (from 1 mm up to THz frequencies), as shown in Fig. 19.32. These remarkable properties can be exploited for optical micro mirrors, modulators, switches and microwave waveguides to overcome the current problems encountered in communication systems (consumption, power handling, slow switching and integration) [16].

Here it is described how to characterize the optical properties of VO$_2$ including its phase change , being the transition induced by a temperature change.

A HORIBA Scientific UVISEL Spectroscopic Phase Modulated Ellipsometer with a temperature controlled cell has been used to characterize the optical properties VO$_2$ deposited on a sapphire substrate (Al$_2$O$_3$) (Fig. 19.33). The change in the optical constants (n, k) at various temperatures is characteristic of the SMT exhibited by VO$_2$. Ellipsometric measurements were performed at a range of temperatures from room temperature to 90 °C using an angle of incidence of 70° and spectral range 190–2100 nm.

The model used to describe the sample is represented in Fig. 19.34. It includes a thin roughness layer (58 Å) at the surface of the VO$_2$ film; this surface roughness

Fig. 19.33 UVISEL 2 with a temperature controlled cell
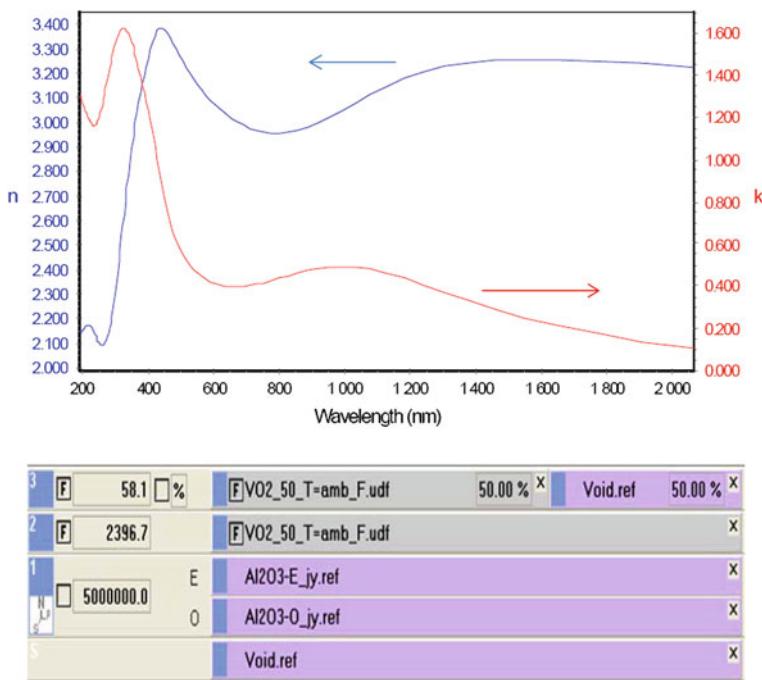
Fig. 19.34 Best-fit model with thicknesses results and optical constants of VO_2

improves the goodness of fit from $\chi^2 = 1.3$ to $\chi^2 = 0.2$. As the sapphire is a transparent substrate the model also includes the calculation of backside reflection contribution (voids on the back side), and it also takes into account the anisotropy of the sapphire for which the ordinary (O) and extraordinary (E) optical constant are introduced.
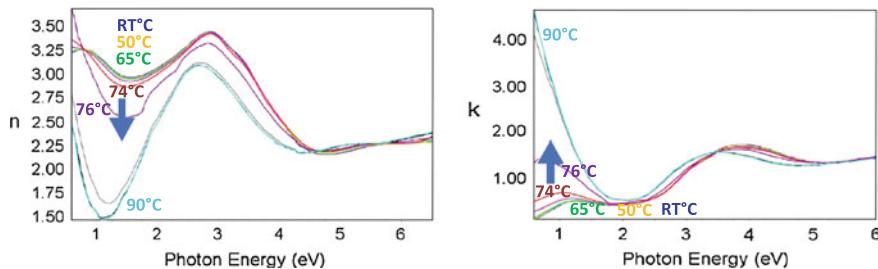
Fig. 19.35 Evolution of the VO$_2$ refractive index, n, (*left panel*) and extinction coefficient, k, (*right panel*) as a function of temperature

The optical constants of the VO$_2$ were determined using the multiple Tauc Lorentz [8] component dispersion formula, and the derived refractive index, n, and extinction coefficient, k, for the bulk VO$_2$ layer with a thickness of approximately 2398 Å are also shown in Fig. 19.34.

It has also been investigated the temperature dependence of the optical properties from ambient to 90°C and the results are shown in Fig. 19.35.

It can be seen that the refractive index decreases and the extinction coefficient increases in the near infrared region, for temperatures above 76 °C. This behavior is typical of metals. Therefore, using spectroscopic ellipsometric data in the NIR, one can infer the presence of the semiconducting or metallic state of VO$_2$.

19.5 Optical Coatings and Functional Coatings

19.5.1 Market Trends

Continuous advancements have given rise to new business and market opportunities in the optical coatings industry. The total global market for optical coatings is estimated to rise from 4.6 billion in 2010 to 5.7 billion in 2015, with a compound annual growth rate of 4.3 %, according to the report Optical Coatings: Technologies and Global Markets [17].

Optical coatings are used in a wide range of sectors namely abrasion, chemical and environmental resistance, low and high temperature operation, biological protection, climate control, EMI (electro-magnetic interference) shielding, optical filters and fibers applications.

The major end-use areas are optical components (mirror, lenses, etc), TV and computer screens, laser optics, flat panel displays, transportation and medical industries. Lighting technology is on the verge of a significant technological transition as incandescent lighting is replaced. Optical coatings and filters will play a role in producing new lighting options. Flat screens on televisions, cell phones, computers and other displays are becoming ubiquitous, and optical coating technology can

extend their usefulness in various lighting settings. Coatings are usually deployed to improve a device; thus, cost is a limiting factor. Most common are antireflection coatings, which are typically low-margin high-volume-produced for lenses and other optical components. Optical coatings, will continue to be important in the military, which now emphasizes remote sensing and observation in warfare tactics. In terms of architecture, coatings are applied to windows. Growth in 'green' window coating is expected to increase. Solar energy is another green revenue opportunity for coatings. Optical coatings are a key component in lasers, adding that, as lasers evolve, very complex coatings will be in greater demand. Because much funding is going into next-generation telescopes, advances in optics and coatings likely will be needed in these efforts.

A variety of materials and composites can be exploited as spectrally selective coatings depending on the application. Thin coatings (less than $1 \mu m$) of oxides, nitrides, oxynitrides and sulfides of Si can be used as selective infrared emitters. Inorganic coatings of SiO, Si_3N_4 and polymer coatings are also suitable for high infrared emitting applications [18].

All most all the energy related applications of selective coatings desire a high emittance at a particular frequency or within a particular frequency range. Therefore, determining the optical properties of the various categories of coatings is of paramount importance.

19.5.2 Ellipsometry Characterization

Here three different examples of the use of ellipsometry for the industrial control of optical coatings, of hard coatings for cutting tools and of electrochromic coatings [19].

19.5.2.1 Example of Characterization of TiN/AlN Hard Coatings

The interest in thin films of metal nitrides is growing rapidly for such diverse applications as wear resistant coatings on cutting tools, selective transmission coatings on architectural glass, and as diffusion barriers in integrated circuits. Among the various transition metal nitrides, processing, properties and applications of titanium nitride (TiN) have been investigated extensively. Titanium nitride crystallizes in the B1 NaCl structure and exists as a solid solution containing nitrogen in the range 37.5/50 at.%. Hard coatings such as TiN and its variants including titanium carbon nitride (TiCN), titanium aluminium nitride (TiAlN or AlTiN), and titanium aluminum carbon nitride (TiAlCN) are used to improve the performance of cutting tools, moulds and dies.

It is well established that (Ti,Al)N coating enhances wear and oxidation resistance of tools, while overcoming the shortcomings of TiN and TiCN coatings. Recently, TiAlN coatings have attracted more and more attentions because of the optical properties. Compared with tradition cermet material,TiAlN coating has good thermal stability and excellent oxidation resistance at high temperature in air
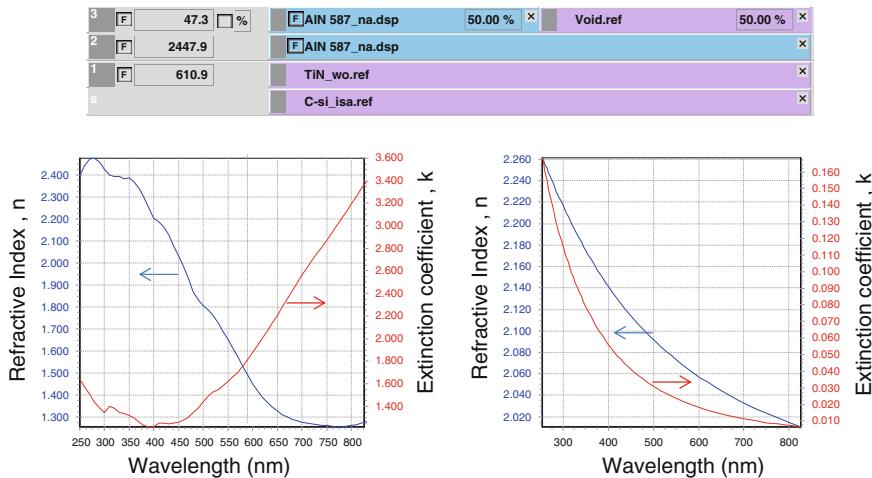
Fig. 19.36 Model with thicknesses results and optical constant of TiN (*left panel*) and optical constants of AlN (*right panel*)

atmosphere and has become a favorable candidate for the applications of solar selective absorbers [20].

A nanoscale bilayer TiN/AlN deposited on crystalline silicon, c-Si, was investigated by the UVISEL spectroscopic ellipsometer to determine thicknesses and optical constants of the various layers. The TiN material exhibits a high absorption coefficient that limits the thickness measurement by ellipsometry. In this example, it was still possible to determine a TiN thickness of 611 Å.

A simple model, shown in Fig. 19.36, represented the sample and it included a rough overlayer modeled using the effective medium approximation of 50 % AlN+50 % voids.

The optical constants of TiN (Fig. 19.36) were directly taken from the reference library of the DeltaPsi2 software whereas the AlN optical constants were calculated using a Lorentz oscillator dispersion formula.

19.5.2.2 Example of Characterization of Electrochromic Thin Films

Electrochromic (EC) materials have been finding several applications in smart windows, anti-dazzling rear view mirrors, switchable motorcycle helmet, for contrast enhancement in some emissive display devices and in non-emissive large area colour displays for information advertisement [21].

The refractive index (n), extinction coefficient (k), and thickness (d) are important parameters for predicting the performance of a film in an optical system. Spectroscopic ellipsometry is the ideal technique to measure these parameters simultaneously, very accurately and non destructively.
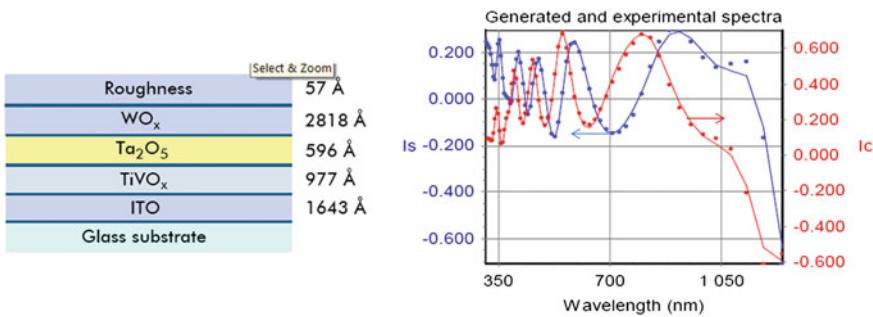
Fig. 19.37 Model with thicknesses results used to obtain the best-fit (*lines*) to the experimental spectra (*dots*) of the *right panel*

A typical EC device comprises two EC films separated by a transparent ion conductor, and this three-layer stack is positioned between transparent electrical conductors. One of the EC electrodes usually consists of tungsten oxide, WOx. The optical absorption can be modulated—reversibly and persistently—by shuttling charge between the two EC films, which is effected by applying a voltage between the transparent electrodes (e.g. ITO).

Here it is discussed the characterization of the multilayer device sketched in Fig. 19.37 exploiting the UVISEL spectroscopic ellipsometer. Indeed, the multilayer sample was characterized in several steps measuring layers step by step, to achieve a better precision in the found results.

The optical constants have been determined using:

- A combined Lorentz oscillator and Drude term for the ITO layer
- An absorbing Lorentz oscillator for the Ta_2O_5 layer
- The W-oxide film was homogeneous and could be well represented by a Tauc–Lorentz parameterization as well as the $TiVO_x$ layer.

The excellent agreement to the experimental data, evident from Fig. 19.37, was obtained for fitted values of the W-oxide layer thicknesses equal to 2828 A, and for its surface roughness (represented by EMA mixture of 50 % voids+50 % WOx) equal to 57 Å, respectively, when the corresponding thickness numbers for the Ta_2O_5, TVO_x and ITO layers were, respectively, 596 Å, 977 Å and 1643 Å. $TiVO_x$ and WO_x materials both exhibit an absorption peak in the FUV range around 400 nm, while Ta_2O_5 exhibits a continuous absorption along the measured spectral range. The optical gap Eg $= 3.11 \pm 0.03$ found for the WO_x layer is in agreement with the literature data. The optical constants derived for the four oxide layers are shown in Fig. 19.38.
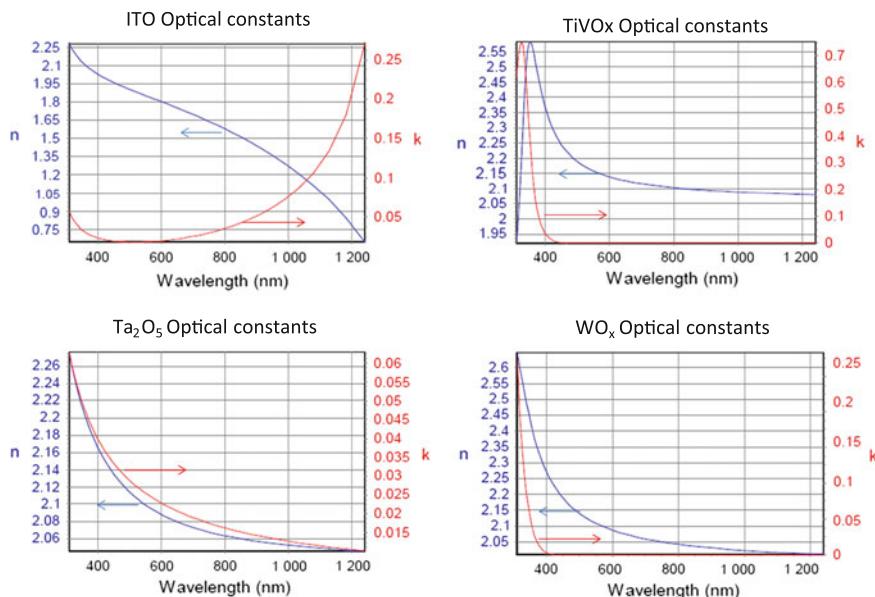
Fig. 19.38 Optical constants, namely, the refractive index, n, and the extinction coefficient, k, for ITO, $TiVOI_x$, Ta_2O_5 and WO_x thin films

19.6 Chemistry and Biology Engineering

The sensitivity (down to a few picometers thickness), the speed of measurement (down to a few seconds) and the possibility to make in situ measurements at solid/liquid interfaces make ellipsometry a reliable and attractive tool for applications in the life science area. Spectroscopic ellipsometry can be used in situ at solid/liquid interfaces thus allowing measurements to be performed in a physiologically-like environment. The non-invasive character and the possibility to perform fast measurements in situ allows to monitor adsorption kinetics with a time resolution sufficient to resolve the adsorption of a single macromolecular layer [22]. Therefore, a large number and variety of applications of ellipsometry are found in biology and biochemistry as well as on organic layers in general. For instance, biomolecular recognition processes, like antigen-antibody binding, hormon-recept binding and enzymatic reactions on surfaces can be in principle monitored in real time.

The formation of ultrathin layers of biologic molecules on well characterized surfaces is attracting growing scientific and technological interest. In this field, a key concept is that of Self-Assembled Monolayer (SAM) to be opposed to the formation of defined layers of adventitious molecules. SAMs are well defined assemblies formed by the adsorption of molecular constituents from the solution or the gas phase onto the surface of solids or even liquids. Obeying to the interplay, often very subtle, between molecule-substrate and molecule-molecule interactions, the adsorbates tend to organize spontaneously (self-assembly) into somewhat spectacular networks. In

fact, when substrates of high morphologic quality are used (e.g. if they present large, atomically flat terraces) these networks may display two-dimensional order properties over very long ranges. The thickness being typically a few nanometers, SAMs represent the most elementary form of a nanometer scale biological thin-film material. The most extensively studied class of SAMs is represented by far from the adsorption of alkanethiols [23], proteins and DNA.

The main applications of ellipsometry for chemistry and biology engineering are:

- The measurement of thickness of organic layers deposited on surfaces. The corresponding application of interests for chemistry and biology are: surface functionalization, binding reactions, surface coverage, self-assembly processes, biomolecule orientation.
- The interface studies in liquid environment

19.6.1 Ellipsometry Characterization of Ultra Thin Surfactant Films

Non-ionic surfactants find widespread industrial application in detergency, dispersion stabilization, foaming and emulsification. They are also commonly used as stabilizers of thin films including wetting films. Therefore, the equilibrium and dynamic adsorption of surfactants at the air-water and solid-water interfaces are critical information. A common method of studying surfactant adsorption at the air-water and solid-water surfaces is to use a pressure balance technique and interpret the surface tension as function of bulk concentration. However, this method is rather imprecise since even small errors in the tension values may cause large errors in the derived surface excess concentrations. Additionally, the restricted amount of information provided by such technique limits the understanding of the adsorption process. Spectroscopic ellipsometry is a powerful technique for thin layer characterization and has the required sensitivity to probe thin film at the nanoscale. Additionally, ellipsometry is not limited to information related to the amount of adsorbed surfactant and process knowledge can be gained by determining both thickness and optical constants of the film.

However, film thickness at the nano-scale remains a challenging application for spectroscopic ellipsometry for two main reasons: the reliability of the extracted film parameters depends on the sensitivity of the ellipsometric data in picking up both film thickness and optical constants information. Since the ellipsometric data are collected in a limit-detection regime, any error carried in these data will be passed to the predicted film thickness and optical constants. As a consequence, in most of ultrathin film ellipsometric data analysis, the calculation of the film thickness is done by using predetermined values for the refractive index. The thickness is calculated from the assumption that the film is homogeneous and a refractive index is kept constant and estimated either from literature or measured by another technique (i.e. Abbe refractometer).
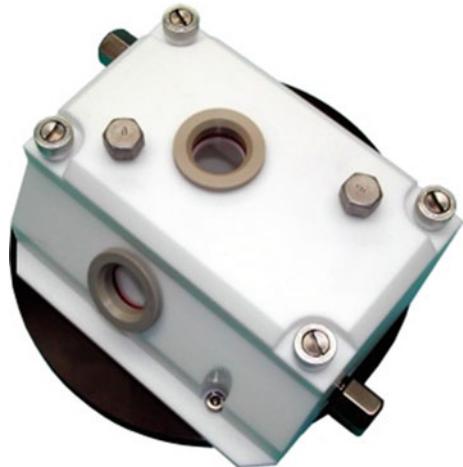
Fig. 19.39 Liquid cell for ellipsometry

Membrane-derivatized liquid-solid interfaces either on particles or on flat surfaces offer an extensive repertoire of chemical functionality. However, the interaction between bilayers vesicles and mineral surfaces does not always produce the desired spreading, which is a single lipid bilayer on the support. Here, we demonstrate the excellent sensitivity of the UVISEL phase modulated spectroscopic ellipsometer to monitor the adsorption process of two non-ionic surfactants, Pentaethylene glycol monododecyl ether ($C_{12}E_{25}$) and Dioctadecyldimethylammonium bromide (DODAB) at silica-water and air-water interfaces, respectively.

19.6.1.1 Example of Characterization of Adsorption of $C_{12}E_{25}$ at the Silica-Water Interface

In order to characterize accurately the adsorption process of $C_{12}E_{25}$, the pseudo substrate consisting of 30 nm of thermal grown oxide deposited on c-Si was first measured into a liquid cell environment filled with water, shown in Fig. 19.39. Measurements were performed by the UVISEL at an angle of 70° over a spectral range of 260–600 nm. A film thickness of 31.95 nm was found for the SiO_2 layer.
The equilibrium adsorption of $C_{12}E_5$ was monitored at a wavelength of 400 nm with a time resolution of 200 ms. Figure 19.40a shows the equilibrium adsorption derived from the UVISEL measurements. Results show clearly that the critical surfactant concentration (cmc) is reached around 480 s from which the adsorbed amount of surfactant can be deduced. A one layer model with a Lorentz oscillator formula was applied to measure simultaneously the surfactant film thickness and the refractive index. The thickness of the surfactant was found equal to 3.59 nm and the optical con-

(a)   **(b)**
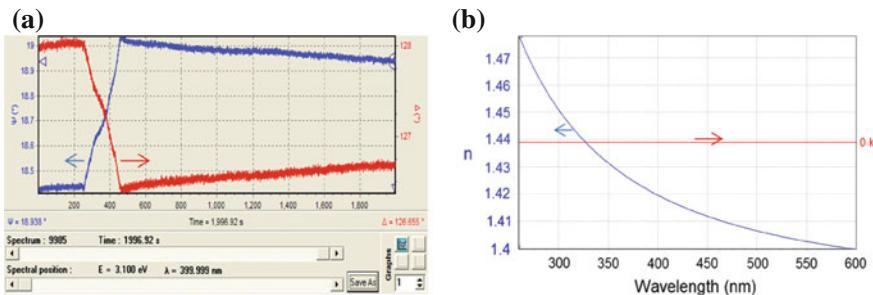
Fig. 19.40 a Adsorption kinetics of $C_{12}E_{25}$ on SiO_2 monitored at the wavelength of 400 nm.
b $C_{12}E_{25}$ optical constants

stants derived across the full spectral range after the critical surfactant concentration
is reached are shown in Fig. 19.40b.

19.6.1.2 Example of Characterization of Adsorption of DODAB at the Air-Water Interface

The cationic lipid DODAB is special due to its anti-microbial properties and the
additional possibility of direct immobilization of oppositely charged biomolecules.
These properties justify the systematic studies of DODAB bilayer adsorption on sili-
con wafers. In order to characterize the adsorption process of the DODAB surfactant
at the air-water interface, the angle of incidence of the UVISEL was changed to 54°,
close to the Brewster angle for pure water substrates, and ellipsometric data were
collected from 260 to 600 nm.
As previously, measurements on pure water were first collected and used as the
pseudo substrate in the ellipsometric model. The film thickness and the optical con-
stants of the DODAB surfactant were also extracted by using a one layer model with
a Lorentz oscillator formula. Figure 19.41a shows a good fit agreement between the

(a)   **(b)**
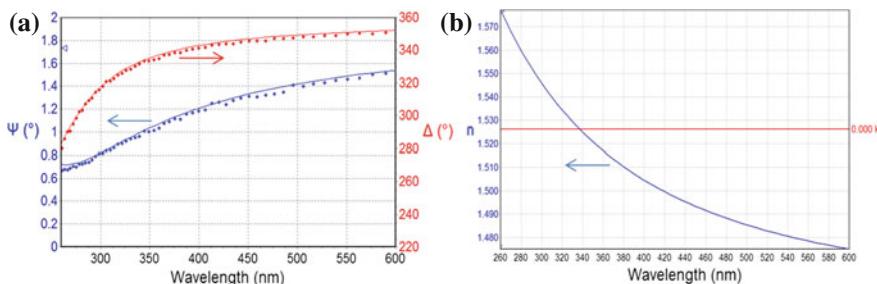
Fig. 19.41 a Fit agreement (*lines*) for the ellipsometric parameters Ψ and Δ measured (*dots*) for
the DODAB film. **b** DODAB optical constants derived from the fit analysis

collected data (dots) and the model (solid line). The thickness found for the DOBAD film was 1.44 nm and its optical constants are displayed in the Fig. 19.41b.

19.7 Conclusions

For the accurate and non destructive characterization of thickness and optical constants of thin film structures, spectroscopic ellipsometers are the instruments of choice. They are well suited for scientific research and industrial control in a production environment. The ellipsometers have considerably evolved these last ten years serving the technological advances in microelectronic, flat panel display, photovoltaic, optoelectronic and biotechnology. Ellipsometry is now an important metrology tool in several different industries and an important technique for industrial research on new materials, devices and processes.

References

1. http://en.wikipedia.org/wiki/Light-emitting_diode
2. A. Khan, K. Balakrishnan, T. Katona, Ultraviolet light-emitting diodes based on group three nitrides. Nat. Photon. **2**, 77–84 (2008). doi:10.1038/nphoton.2007.293
3. B. Drevillon, J. Perrin, R. Marbot, A. Violet, J.L. Dalby, Rev. Sci. Instrum. **53**, 969 (1982)
4. E. Garcia-Caurel, A. De Martino, B. Drevillon, Thin Solid Films **455–456**, 120–123 (2004)
5. G.E. Jellison, F.A. Modine, Appl. Opt. **36**(31), 8190–8198 (1997)
6. EPIA (http://www.setfor2020.eu)
7. Technology Roadmaps *Solar photovoltaic energy*, International Energy Agency
8. G.E. Jellison Jr., Thin Solid Films **206**, 294 (1991)
9. M.A. Green, K. Emery, Y. Hishikawa, W. Warta, E.D. Dunlop, Prog. Photovolt. Res. Appl. **20**, 12–20 (2011)
10. M. Kaltenbrunner, M.S. White, E.D. Glowacki, T. Sekitani, T. Someya, N.S. Sariciftci, S. Bauer, Nat. Commun. **3**(770), 1–7 (2012). doi:10.1038/ncomms1772
11. T. Siegrist, P. Merkelbach, M. Wuttig, Annu. Rev. Condens. Matter Phys. **3**, 215 (2012)
12. A. Dejneka, I. Aulika, V. Trepakov, J. Krepelka, L. Jastrabik, Z. Hubicka, A. Lynnyk, Opt. Express **17**, 14322 (2009)
13. M.-A. Dubois, P. Muralt, Sens. Actuators **77**, 106–112 (1999)
14. I. Aulika, S. Corkovic, A. Bencan, S. D'Astorg, A. Dejneka, Q. Zhang, M. Kosec, V. Zauls, J. Electrochem. Soc. **156**, G217–G225 (2009)
15. I. Aulika, A. Dejneka, V. Zauls, K. Kundzins, J. Electrochem. Soc. **155**, G209 (2008)
16. F. Dumas-Bouchaiat, C. Champeaux, A. Catherinot, J. Givernaud, A. Crunteanu, P. Blondy, Mat. Res. Soc. Symp. Proc. **1129**, 275–286 (2009)
17. http://www.photonics.com/Article.aspx?AID=41242
18. S. Wijewardane, D.Y. Goswami, Renew. Sustain. Energy Rev. **16**, 1863 (2012)
19. http://www.coatingsworld.com/contents/knowledgecenter/7876
20. M. Du, L. Hao, J. Mi, F. Lv, X. Liu, L. Jiang, S. Wang, Solar Energy Mater. Solar Cells **95**, 1193–1196 (2011)
21. http://www.horiba.com/scientific/products/ellipsometers/application-notes/
22. G. Poste, C. Moss, Prog. Surf. Sci. **2**, 139–232 (1972)
23. M.M. Giangregorio, M. Losurdo, G.V. Bianco, A. Operamolla, E. Dilonardo, A. Sacchetti, P. Capezzuto, F. Babudri, G. Bruno, J. Phys. Chem. C **115**, 19520–19528 (2011)

Chapter 20
Ellipsometry and Correlation Measurements

Rados Gajic and Milka Jakovljevic

Abstract Ellipsometry, in general, is one of the most sensitive spectroscopic techniques nowadays in both macro- and nano-scale research. Advantages like absolute measurements (without need for references), high precision, non-destructive character, easy real-time monitoring and many possible applications make Imaging Ellipsometry and Spectroscopic Ellipsometry unavoidable tools in modern research practice. Still, one of the main disadvantages of ellipsometry concerns the fact that an optical model is required in order to define constituents, the structure and morphology of each sample. In most cases, a good optical model that properly describes the structure is not possible to make without additional information like the number of layers, their structure, surface roughness or type of interfaces. For this reason we need additional measurements that corroborate ellipsometric ones in order to construct a valid optical model. We call these measurements Correlation Measurements and they are the topic of this communication. We describe the most frequently used ones like SEM, AFM, STM, Raman or FTS, and we also mention new trends that combine spectroscopic or imaging ellipsometers with one or more correlation techniques in the same instrument.

20.1 Introduction

In the last 40 years, ellipsometry (both imaging ellipsometry (IElli) and especially spectroscopic ellipsometry (SE)) has become one of the most commonly used techniques for optical characterization of surfaces, interfaces and thin films particularly

R. Gajic (✉) · M. Jakovljevic
Institute of Physics, University of Belgrade, Pregrevica 118,
P.O. Box 68, 11080 Belgrade, Serbia
e-mail: rgajic@ipb.ac.rs

M. Jakovljevic
e-mail: mjakovljevic@ipb.ac.rs

in the area of micro- and nano-electronics. SE has been and is used to measure film thickness and optical dielectric function of dielectrics, semiconductors and metallic thin films.

Ellipsometry is a non-destructive, non-invasive and fast technique that could be used even in-situ. Ellipsometry, as discussed in details in Chap. 1, measures the change of polarization state of linearly polarized light upon reflection from a sample, obtaining two independent parameters, Ψ and Δ. Ellipsometry determines the complex amplitude ratio ρ defined as the ratio of the amplitude reflection coefficients r_p and r_s given as $\rho = r_p/r_s = \tan(\Psi) \cdot \exp(i\Delta)$, where indices p and s stand for the p- and s-polarizations [1–3]. In SE the (Ψ, Δ) pair is measured by changing the wavelength of light. In most cases the spectral range covers the ultraviolet/visible (UV/VIS) range but the infrared (IR) range has also been widely used.

One of the advantages of SE consists in performing absolute measurements (there is no need for a reference sample as in spectrophotometry and optical constants can be directly determined without the Kramers-Kronig relations from (Ψ, Δ)). Other advantages include high precision, nondestructive measurements, a variety of applications that can be addressed (as described in the previous chapters), fast and real-time monitoring, a wide spectral range and the possibility to characterize anisotropic samples using generalized ellipsometry (see Chap. 10). One of the main features of ellipsometry is its precision. Namely, this technique is very sensitive to any changes of layer thicknesses even below 1 nm with a sensitivity of 0.1 Å. It turns out that ellipsometry is a promising technique for nanomaterial characterization. For an overview of the various ellipsometry strategies for the measurement and analysis of nanometric films, metal nanoparticles and nanowires, semiconductor nanocrystals, and submicron periodic structures the reader should take a look at [4]. Ellipsometry is a very sensitive technique but the crucial problem is, both for macro and nanosize samples, that one often does not know what has been measured so sensitively (stated by E. A. Irene) (see Chap. 1).

The inherent disadvantage of ellipsometry stems from the fact that it requires an optical model that defines the number of constituents, the structure and morphology of each constituent and the overall structure and morphology of the sample. For the correct interpretation of the sensitive and accurate ellipsometry measurements in most cases it is necessary to have a good optical model that describes the structure and sometimes it is not possible to make one without additional information like the number of layers, their structure, surface roughness, interfaces etc.

It is unlikely that a single technique will solve the problem of rational characterization of nanomaterials; hence, there is a general consensus that a multi-diagnostic approach has to be used and that platforms for simultaneously assessing multiple properties have to be addressed. Therefore, different characterization techniques including scanning electron microscopy (SEM), transmission electron microscopy (TEM), atomic force microscopy (AFM), scanning tunneling microscopy (STM), Raman spectroscopy, Fourier transform infrared (FTIR) spectroscopy, photoluminescence (PL) spectroscopy, X-ray photoelectron spectroscopy (XPS), X-ray reflec-

tivity, and many others, are being used to corroborate ellipsometry. We refer to those additional methods as correlation measurements or techniques.

For correlation measurements the numerous imaging techniques help us to visualize the structures whereas the analytical techniques give us direct additional morphological and structural information that cannot be obtained by ellipsometry only and identify materials properties and changes at a nanoscale. Therefore, correlation techniques are necessary for proper interpretation and extraction of information based on ellipsometry data.

20.2 Correlation Measurement Techniques

Correlation techniques can be tentatively divided into imaging (microscopies) and analytical techniques (spectroscopies or different surface analytical techniques). The imaging techniques can involve light, electrons, ions or scanning probes.

A common approach is to use light as a probe in the so-called near-field scanning microscopes (SNOM/NSOM). The other important light microscope is the confocal laser scanning microscope (e.g. take a look at Olympus LEXT OLS LEXT 4000) with a lateral resolution close to 100 nm. Still, the most widely used techniques for the characterization of nanomaterials are electron microscopes, such as scanning electron and transmission electron microscopies (SEM and TEM). The second most important imaging methods are numerous scanning probe techniques like scanning tunneling microscopy (STM), atomic force microscopy (AFM), lateral (shear) force microscopy (LFM), conductive atomic force microscopy (c-AFM), two-pass AFM methods, electric force microscopy (EFM), magnetic force microscopy (MFM), and Kelvin probe microscopy (KPM). On the other hand, the analytical methods comprise of different spectroscopic techniques like Raman spectroscopy, photometric methods like Fourier transform spectroscopies (FTS/FTIR), attenuated total reflection (ATR), photoluminescence spectroscopy (PL), and surface analytical techniques such as Auger electron spectroscopy (AES), x-ray photoelectron spectroscopy (XPS) and X-ray reflectometry (XRR), energy dispersive X-ray spectroscopy (EDS or EDX, EDAX)—usually combined with SEM or TEM, low-energy electron diffraction (LEED)—sometimes as an accessory to the low-vacuum STM/AFM systems or electron energy loss spectroscopy (EELS)—also often combined with electron microscopes. In order to set the correct model for the ellipsometric analysis to increase accuracy of our determined parameters, the proper correlation techniques are involved. In most cases people are limited by the experimental techniques available in laboratories. In general, ellipsometry is used for characterization of thin films and nanostructures. Usually, ellipsometry is combined with at least one imaging/analytical (structural/compositional) technique. Statistics based on Web of Science [5] still show a clear difference between the correlation techniques used for thin films and nanostructures in conjunction with SE. It seems that ellipsometry used for the characterization of 3D nanostructures, including nanowires, nanoparticles, quantum dots, fullerenes and dentrimers, is much more often combined with different correla-

tion methods compared to thin film analysis. For instance, in the case of nanostructure, in more than 50 % of cases ellipsometry is supported by a microscopy technique for visualizing their size, geometry and spatial distribution. Since it is almost an impossible task to describe all these techniques in detail in this chapter we will mention just the most important and often used ones.

20.2.1 Imaging Correlation Techniques

Transmission electron microscopy (TEM) is a microscopic technique whereby a beam of electrons is transmitted through an ultra-thin specimen. An image forms by the interaction of the electrons transmitted through the specimen with the specimen, which is magnified and detected by a sensor such as a CCD camera. TEMs are capable of imaging at a significantly higher resolution than light microscopes, owing to the small de Broglie wavelength of the electrons. The best TEM systems have an atomic resolution. One of the disadvantages of TEM is the complex sample preparation.

The scanning electron microscopy (SEM) images a sample by scanning it with a high-energy beam of electrons in a raster scan pattern. The electrons interact with the atoms that make up the sample producing signals that contain information about the sample's surface topography, composition, and other properties such as electrical conductivity. Comparing to TEM, SEM usually has a smaller spatial resolution of a few nm in high vacuum. For conventional imaging in SEM, specimens must be electrically conductive, at least at the surface, and electrically grounded to prevent the accumulation of electrostatic charge at the surface. Metal objects require not much special preparation for SEM except for cleaning and mounting on a specimen stub. Nonconductive specimens tend to charge when scanned by the electron beam, and especially in secondary electron imaging mode, this causes scanning faults and other image artifacts. They are therefore usually coated with an ultrathin coating of electrically conducting material (like Au or graphite), deposited on the sample either by low-vacuum sputter coating or by high-vacuum evaporation. Nonconducting specimens may be imaged uncoated using specialized SEM instrumentation such as the Environmental SEM (ESEM) or field emission gun (FEG) SEMs operated at low voltage and low vacuum. The resolution of SEM systems can be increased using the field emission guns (FEG SEM) of the cold-cathode type using tungsten single crystal emitters or the thermally assisted Schottky type. FEG SEM can have a resolution even below 1 nm approaching the TEM resolution.

AFM is one of many different scanning probe methods (SPM) and the best known member in the numerous scanning force microscopy family are electrical force microscopy (EFM), lateral force microscopy (LFM), Kelvin probe microscopy (KPM), magnetic force microscopy (MFM), and conductive-AFM (c-AFM). AFM is a very high-resolution type of scanning probe microscopy, with demonstrated resolution of the order of fractions of a nanometer, more than 1000 times better than the optical diffraction limit. AFM is one of the foremost tools for imaging, measuring,

and manipulating matter at the nanoscale. The information is gathered by measuring the force between the probe and the surface. The AFM resolution is limited by the probe geometry. Typical maximum resolutions in the plane are of a few nm and with specific tips, 1 nm or even atomic (UHV systems) resolution is possible. In standard use AFM gives high resolution topography of the sample.

STM is a scanning probe microscopy based of the concept of quantum tunneling between the tip and the conductive surface of the sample. For STM, good resolution is considered to be 0.1 nm lateral resolution and 0.01 nm depth resolution. With this resolution, individual atoms are routinely imaged and manipulated. The STM can be used not only in ultra-high vacuum but also in air, water, and various other liquid or gas environments, and at temperatures ranging from near zero Kelvin to a few hundred degrees Celsius. The STM probes the local density of states (LDOS) of a material measuring the tunneling current.

20.2.2 Analytical Correlation Techniques

A powerful spectroscopic method used for nano-characterization is Raman spectroscopy. Raman spectroscopy is used to observe low frequency symmetric excitations (phonons, plasmons, magnons,...) in a sample. It relies on inelastic scattering, or Raman scattering, of monochromatic light, usually from a laser in the visible, near infrared, or near ultraviolet range. The laser light interacts with molecular vibrations, phonons or other excitations in the system, resulting in the inelastic scattered photons below and above the Rayleigh line (laser line) called Stokes and Anti-Stokes lines. The shift in energy gives information about the excitations in the system. Infrared spectroscopy yields similar, but complementary, information since the selection rules for IR and Raman active modes are different. Raman spectroscopy as a simple, quick and noninvasive technique is a useful tool for characterization. However, the main limitation of Raman spectroscopy for applications in the domain of nanotechnologies is its low spatial resolution. Namely, even with the best confocal microscopes, a spatial resolution better than 500 nm is hardly obtainable. Recently new more sensitive Raman techniques are appearing like surface enhanced Raman spectroscopy or surface enhanced Raman scattering (SERS) and Tip Enhanced Raman Scattering (TERS). SERS exploits concentrated electromagnetic energy in metallic nanostructures to significantly enhance the Raman scattering cross section. The enhancement factor can be as much as $10^{10} - 10^{11}$ meaning that single molecules can be detected.

FTS/FTIR is a photometric technique where dispersion elements like diffraction gratings in monochromators are replaced with interferometers (the Michelson's intereferometer is mostly used). After reflection or transmission through the sample, the interferogram is recorded followed by the Fourier transform of the interferogram signal resulting in the reflectance or transmittance spectrum. FTS is, similar to Raman spectroscopy, a noninvasive technique whose advantage concerns a high spectral resolution and high wavenumber accuracy in the whole spectral range

from the FAR IR to the UV range. Because of low spatial resolution Raman spectroscopy and FTS cannot be used directly for characterization of nanosize particles but they enable an indirect determination of nanoparticles due to the confinement effect. For instance phonon lines broaden and shift for particles less than 50 nm in size. The fitting of the line shape provides information on the average size of particles [6].

GXR is a powerful technique to measure the absolute thickness without "a priori" structural modeling of a multi-layer stack. GXR applies to any type of material: semiconductor, metallic, dielectric or organic. GXR and SE measurement can be performed quasi-simultaneously at exactly the same sample spot. Analysis of complementary data from both techniques with the same physical model leads to unprecedented accuracy in measurement results. GXR is certainly the technique of choice for absolute thickness measurement of new materials. At this energy, every material displays the same refractive indices (n being 1 and k being 0), simplifying considerably the structural model for data analysis. In the case of complex multi-layers, accurate information can be obtained on the periodicity of stacks as well as on the interlayer structures. SE combined with GXR is the winning choice for the following applications: high K dielectrics (ZrO_2, nitrided oxides), thin ONO layers, thin optical coatings, Si and SiGe thin epitaxial structures, thin metallic and organic layers, and coatings for micro-lithography. SOPRALAB GES5E-GXR SE is an example of combining two non-destructive characterization techniques on the same instrument: Spectroscopic Ellipsometry and Grazing X-Ray Reflectometry [7].

The higher the energy of electromagnetic radiation, the deeper we can look into the structure of nanomaterials. The classical application of X-rays is elastic diffraction used for the analysis of crystal structures and grain size in crystalline materials. The beam of X-rays hits a sample and scatters into many different directions. By analysis of angles and intensities of the scattered beams, a 3D picture of the density of electrons can be reconstructed. The electron density reveals information about the mean position of the atoms in the crystal, their chemical bonds, disorder and other information. Still, XRD has some limitations when it comes to nano-domains. As the size of the nanocrystal decreases, the broadening of the diffraction peak increases. At some point, the broadening is so high that it can merge with other peaks and information is lost. Grain size can be extracted from XRD spectra by using the Scherrer's formula. Still, it is important to be aware that the Scherrer's formula gives a lower limit of particle size. The XRD peaks depends on the "crystalline coherent domain" size which does not always correspond to the "grain" or "particle" size obtained by microscopies, since the plane defects can break the structure symmetry.

X-ray photoelectron spectroscopy (XPS), also known as electron spectroscopy for chemical analysis (ESCA), is a technique in which x-rays are employed to generate photoelectrons. Analysis of the energies of these electrons is used to identify and quantify the elemental composition and chemical state of the material being analyzed. XPS spectra are obtained by irradiating a material with a beam of X-rays while simultaneously measuring the kinetic energy and number of electrons that escape from the top 1–10 nm of the material being analyzed. XPS requires ultra-high vacuum
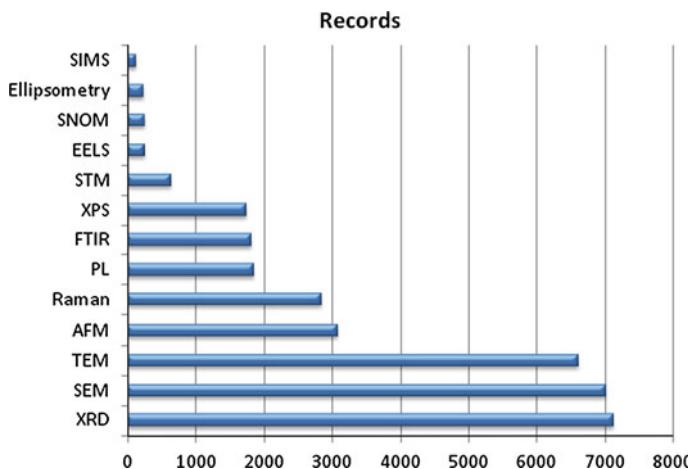
Fig. 20.1 Different techniques used in nanomaterial characterization (*Data source* ISI Web of Science)

conditions. XPS is used to analyze the surface chemistry of a material before or after some treatments, for example: fracturing, cutting or scraping in air or UHV to expose the bulk chemistry, ion beam etching to clean off some of the surface contamination, exposure to heat to study the changes due to heating, exposure to reactive gases or solutions, exposure to ion beam implant, exposure to ultraviolet light. XPS is used to determine: elemental composition of the surface (top 1–10 nm usually), empirical formula of pure materials, elements that contaminate a surface, chemical or electronic state of each element in the surface and uniformity of elemental composition across the top surface.

Secondary ion mass spectrometry (SIMS) is a technique used in materials science and surface science to analyze the composition of solid surfaces and thin films by sputtering the surface of the specimen with a focused primary ion beam and collecting and analyzing ejected secondary ions with a mass spectrometer. SIMS can be used in a scanning mode and a spatial image of surface chemical composition can be constructed [8]. SIMS is the most sensitive surface analysis technique, being able to detect elements present in "the parts per billion" range.

The ISI Web of Science has been used to map the techniques used for characterization of nanomaterials. There are 76,419 "nano" records. The main techniques used for their characterization are shown in Fig. 20.1.

There are 230 "nano + ellipsometry" records. In those papers, ellipsometry was corroborated with other techniques used for "nano" characterization, and the percentage of papers corroborating ellipsometry with different characterization techniques is presented in Fig. 20.2. The statistics shows both the importance of imaging and analytical techniques in nanomaterial characterization.
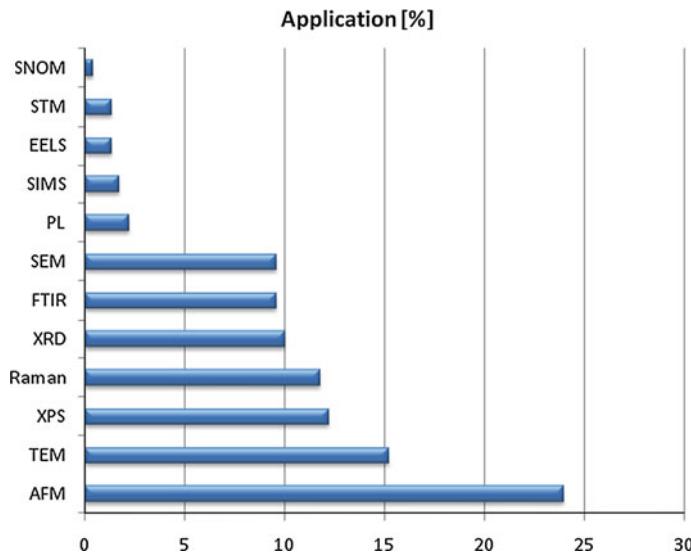
Fig. 20.2 Percentage of the techniques used in correlation with ellipsometry (ISI Web of Science)

20.3 Examples of Ellipsometry and Correlation Measurements

20.3.1 AFM for Surface Roughness Analysis and Visualization of Particle Size

One of the most common output of ellipsometry modeling and analysis concerns surface roughness. Namely, from the measured pair (Ψ, Δ) a pseudo-dielectric function $\langle \varepsilon \rangle$ is obtained from an optical model that typically assumes an ideally flat surface for the sample. When roughness is present, the pseudo-dielectric function depends on the size and complexity of the roughness and could differ significantly from the bulk dielectric function of the sample (see Chap. 4 in this book). The pseudodielectric function $\langle \varepsilon \rangle$ is the dielectric function obtained directly from the measured values (Ψ, Δ) assuming a perfectly flat substrate of infinite thickness (bulk sample) as in Fig. 20.3.

It is estimated that when surface roughness exceeds 30 % of measurement wavelength, the errors cannot be tolerable. Therefore, an estimation of the sample roughness is an important issue. Figure 20.4 shows a real sample with surface roughness and the corresponding effective medium model based on surface roughness and bulk layers. Parameters, d_s and f_{void} are the roughness thickness and volume fraction of the voids, respectively [1]. The calculation of the SE (Ψ, Δ) spectra of c-Si based on the effective medium optical model from Fig. 20.4, with the surface roughness layer modeled as a Bruggeman effective medium approximation (BEMA) of

Fig. 20.3 Structure with a surface roughness and corresponding pseudo-dielectric function $\langle \varepsilon \rangle$ given as in the *right panel* [1] (for $\langle \varepsilon \rangle$ definition, see Chap. 1 of this book)


$$\langle \varepsilon \rangle = \varepsilon_t = sin^2\theta_i \left[1 + tan^2\theta_i \left(\frac{1-\rho}{1+\rho} \right)^2 \right]$$
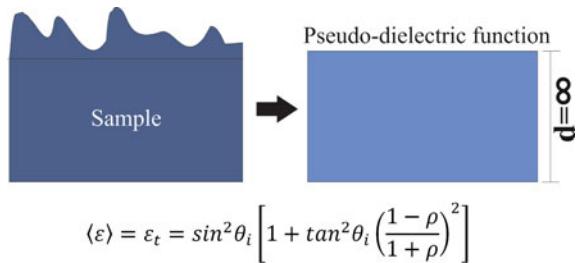
Fig. 20.4 Modeling of surface roughness
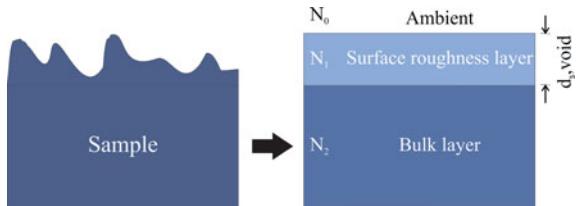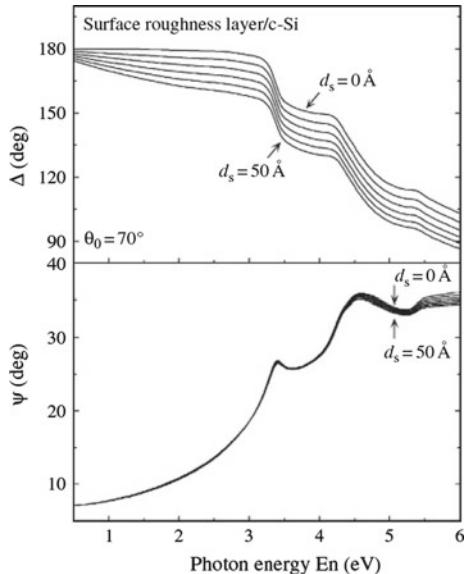
Fig. 20.5 (Ψ, Δ) spectra of c-Si based on the optical model for d_s between 0 and 50 Å and $f_{void} = 0.5$ at the incident angle of 70° [1]

50 %c-Si + 50 %voids (see Chap. 4) is given in Fig. 20.5. As a result, Fig. 20.5 puts in evidence that even a small surface roughness causes large changes of Δ confirming the importance of the effect of surface roughness in ellipsometry.

The simplest way to estimate surface roughness is by the use of AFM as Fig. 20.6 shows. 3D scanning laser confocal microscopy with a smaller spatial resolution comparing to AFM can also be used (www.olympus.com) for determination of roughness.
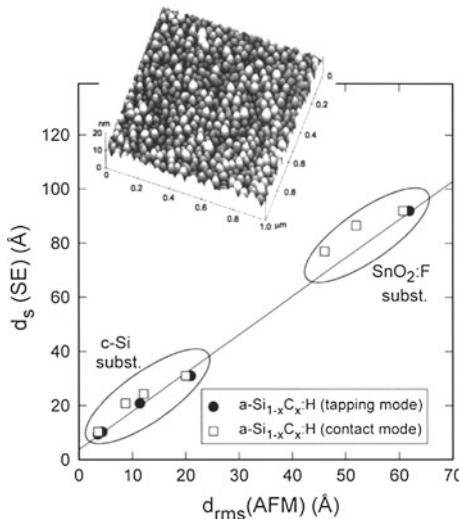
Fig. 20.6 Correlation between root-mean-square (ms) surface roughness deduced from atomic force microscopy images (see *inset* for example) and the roughness thickness from real-time SE measurements at the end of the depositions for a set of eight a-Si$_{1-x}$C$_x$:H thin films prepared under different conditions. Substrates were smooth c-Si (*lower data set*) and rough SnO$_2$:F (*upper data set*). Two different AFM instruments were used, one operated in contact mode (*squares*) and the other in tapping mode (*solid circles*). The *solid line* is a fit to the data points obtained in tapping mode, since these yield larger roughness thicknesses for the four samples measured by both instruments, presumably due to the higher resolution in tapping mode. The form of the *solid line* correlation is ds(SE) = 1.5 drms(AFM) + 4Å. [Reprinted with permission from Ref. [9]. Copyright Elsevier Limited (1998)]

There are many reports in literature demonstrating the correlation between the surface roughness measured by ellipsometry and AFM, and an example is shown in Fig. 20.6: the roughness layer thicknesses (*ds*) on a-Si$_{1-x}$C$_x$:H films deduced by real-time SE correlate with those measured ex situ by atomic force microscopy (AFM) in tapping mode according to the relationship: $d(SE) = 1.5 \, drms(AFM) + 4Å$ [9].

As another example, SE and AFM were used to study the film thickness and the surface roughness of both 'soft' and solid thin films. 'Soft' polymer thin films of polystyrene and polystyrene–ethylene/butylene–styrene block copolymer were prepared by spin-coating onto planar silicon wafers [10]. SE parameters were obtained by the Cauchy model using a two-layer model with planar boundaries between the layers. The smooth surfaces of the prepared polymer films were confirmed by AFM. There is good agreement between AFM and ellipsometry in the 80–130 nm thickness range. In the same communication, Si surfaces obtained by anisotropic chemical etching were investigated as an example of a randomly rough surface. To define roughness parameters by ellipsometry, the top rough layers were treated according to the Bruggeman effective medium approximation. The combined use of both meth-

Fig. 20.7 A direct comparison of thickness measurements of homogeneous soft polymer films obtained by combining SE and AFM. [Reprinted with permission from Ref. [10]. Copyright John Wiley and Sons (2007)]
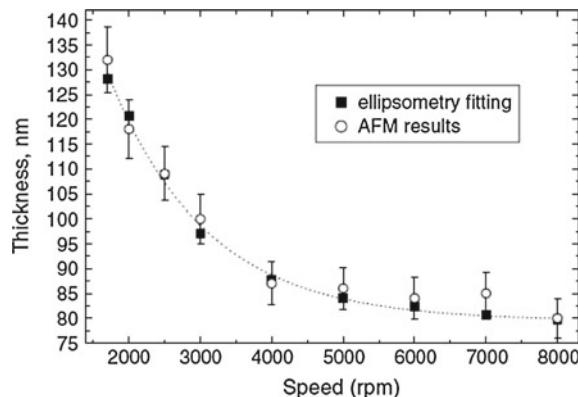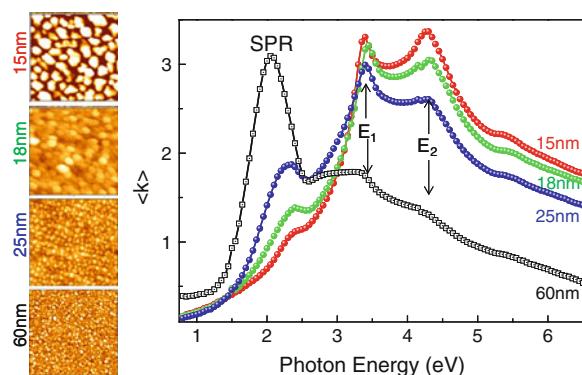

Fig. 20.8 Pseudo-extinction coefficient $\langle k \rangle$ of Au nanoparticles on the Si substrate as a function of the particle size. E_1 and E_2 are the critical points of c-Si. At the *bottom* are the $500 \times 500\,nm$ AFM images. [Reprinted with permission from Ref. [11]. Copyright Elsevier Limited (2011)]

ods appears to offer the most comprehensive route to quantitative surface roughness characterization of solid films as in Fig. 20.7.

One of the promising fields of application of ellipsometry in recent years is plasmonics. By applying ellipsometry to plasmonic structures it is possible to infer nanoparticle size and plasmonic properties [11]. Here we report the exploitation of SE for gold nanoparticles directly sputtered on the Si(111) substrate. It turned out that the Si/Au nanoparticle system is challenging for ellipsometry because of $AuSi_x$ interface layers. Figure 20.8 shows an example of the temporal evolution of pseudo-extinction coefficient $\langle k \rangle$ as a function of Au sputtering time i.e. the Au nanoparticle size. The corresponding AFM measurements corroborate the ellipsometric analysis providing nanoparticles' size. The surface plasmon resonance (SPR) is red-shifted whereas their amplitudes increase in correlation with an increasing particle size, as confirmed by the AFM measurements. Therefore, SE in combination with AFM provides means to monitor the formation of Au nanoparticles on Si and characterizes their plasmon properties [11].
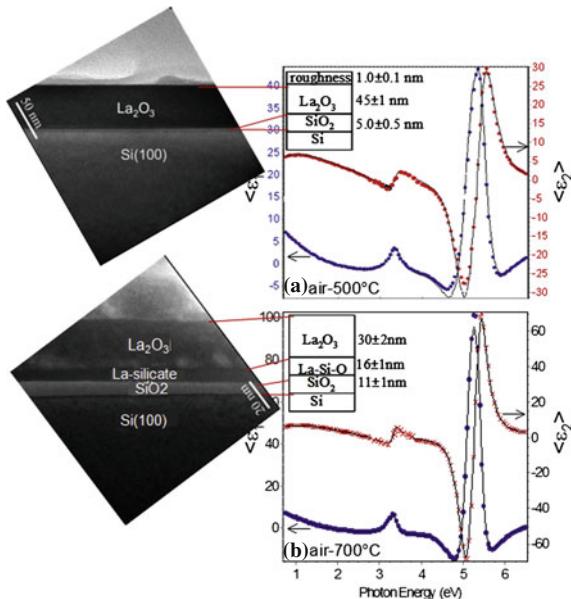
Fig. 20.9 Experimental (*dots*) and best-fit (*lines*) ellipsometric spectra of the real, $\langle\varepsilon_1\rangle$, and imaginary, $\langle\varepsilon_2\rangle$, parts of the pseudodielectric function of the samples annealed in air for 1 h at **a** 500 °C and **b** 700 °C. The best-fit models showing the layered structure of films are also reported

20.3.2 SEM/TEM Corroborating Ellipsometry

An example of TEM measurements corroborating ellipsometry is given in Fig. 20.9, reporting thickness measurement of nanosize interface layers. The example shows La_2O_3 thin films deposited on a Si(100) substrate, which have gained attention as a potential high-permittivity dielectric material in complementary metal-oxide-semiconductor (CMOS) devices. The complexity of this system rises from the interface reactivity resulting in nanometric interfaces layers of different composition, including a SiO_2 interface layer and a silicate $LaSiO_x$ interface layer. Determining accurately the thickness of the interface layers is mandatory for an accurate determination of the optical functions of the La_2O_3 film. Figure 20.9 shows the experimental and best-fit ellipsometric spectra of the real $\langle\varepsilon_1\rangle$ and imaginary $\langle\varepsilon_2\rangle$ parts of the pseudodielectric function for the samples heated in air at two different temperatures of 500 and 700 °C, which resulted in different interfaces layers [12]. In a first attempt, thickness values as determined by TEM analysis have been used as input for the ellipsometric model to derive the optical properties of the silicate interface layer as well as of the La_2O_3 films. Ellipsometric measurements clearly support a multi-layer structure for the annealed films. The best-fit models show the layered structure of the films. In particular, a SiO_2 interface layer with a thickness of 5 nm is found for the 45 nm La_2O_3 film upon 500 °C annealing in air, while the best-fit of the
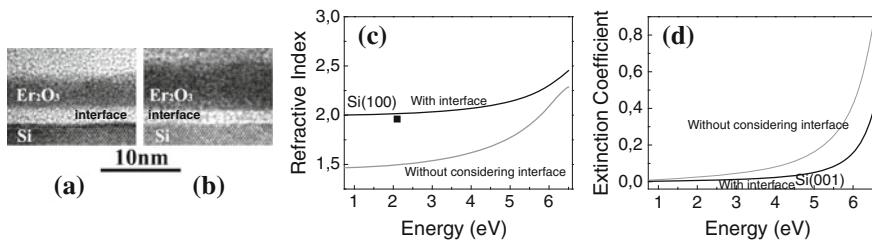
Fig. 20.10 **a**, **b** TEM images of Er_2O_3 ultrathin layers grown on Si(100) used to accurately calculate layer thicknesses and to extract the spectra of the refractive index and of the extinction coefficient of Er_2O_3 layers from ellipsometric analysis. For (**a**) the interface and Er_2O_3 thicknesses are 1.8 and 4 nm, respectively; for (**b**) the interface and Er_2O_3 thicknesses are 1.5 and 6.7 nm, respectively. Spectra derived from ellipsometric modeling for the refractive index (**c**) and extinction coefficient (**d**) of the Er_2O_3 layer in (**a**) considering (and not considering) in the model the interface layer. The *square dot* indicates a reference value reported in literature for the Er_2O_3 layer

sample annealed at 700 °C has been achieved with two interface layers, i.e., a 11 nm thick SiO_2 interface layer followed by a 16 nm thick layer with a higher refractive index, which has been attributed to the formation of lanthanum silicate, and a La_2O_3 thickness of 30 nm. Once the ellipsometric analysis approach is validated with TEM analysis, and the optical properties of the La_2O_3 films are known, the ellipsometric analysis can be routinely used for measuring thicknesses, interface reactivity, and as an optimization process for growth of La_2O_3 films on Si.

The use of TEM in corroborating SE measurements of ultrathin films with a thickness <10 nm is also shown for the thin Er_2O_3 layers grown on Si(100) [13]. Without the corroborating TEM analysis, and simply relying on fitting ellipsometric data using a simple Er_2O_3 layer whose optical function was unknown and parameterized with a Lorentzian oscillator, overestimated total film thickness was obtained. Accurate analysis was instead achieved by determining the thickness of the SiO_2 interface layer and the Er_2O_3 ultrathin layer by corroborating TEM analysis. These thickness values were used as input in a two-layer model to derive the correct optical properties of the interface layer as well as the Er_2O_3 films as reported in Fig. 20.10.

20.3.3 XPS and Ellipsometry

Based on the statistics shown in Fig. 20.2, XPS is one of the most used correlation analytical compositional techniques corroborating ellipsometry. Here this is discussed for the chemical and physical analysis of the $Si(SiO_2)/Au(Ag)$ interfaces. The first example concerns gold (Au) nanoparticles sputtered at room temperature on a silicon substrate [5]. The Au nanoparticles dielectric function has to be parameterized and determined by the fitting to ellipsometric spectra, and its accuracy and reliability depend on the model used in the analysis. Therefore, in order to reduce unknown parameters about the sample structures, the proper ellipsometric model was based on
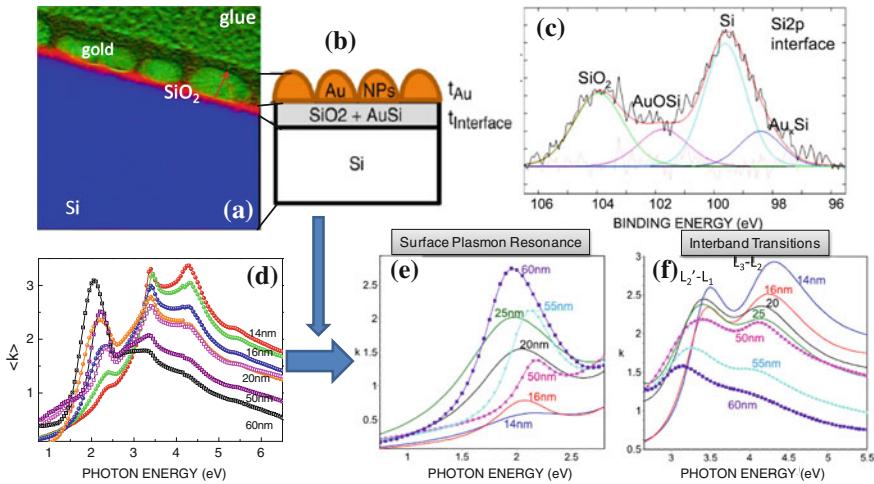
Fig. 20.11 **a** TEM image, **b** best-fit model, **c** XPS spectrum of the Si2p photoelectron core level, **d** experimental spectra and **e**, **f** best fit results for gold nanoparticles sputtered on Si

information obtained by XPS and TEM as shown in Fig. 20.11. The TEM provides information about the thickness of the interface and of the Au nanoparticle layer while XPS gives the chemical composition of the interface. Specifically, Fig. 20.11a shows the HR-TEM image of Au NPs sputtered on Si, which put in evidence the formation of an extended, amorphous, and rough alloyed interface layer between the Au NPs and the Si substrate. The interface layer is a mixture of AuSi and SiO$_2$, as confirmed by EFTEM and by the XPS spectrum of the Si 2p photoelectron peak shown in Fig. 20.11c. The thickness of the interface layer was obtained by HR-TEM and the ratio of the SiO$_2$/AuSi was determined from XPS and both were used as input and fixed in the ellipsometric analysis. In particular, the Si 2p core level in Fig. 20.11c shows main peak component at a binding energy (BE) of 99.7 eV due to the Si substrate and at BE of 103.8 eV due to the interface SiO$_2$, together with additional peaks at BE of 98.2 and 101.5 eV, which are indicative of interface inter-mixing, i.e., gold silicide AuSi and its oxide. On the basis of those structural-chemical information, the best-fit model sketched in Fig. 20.11b was used to model the ellipsometric spectra shown in Fig. 20.11d. Being the composition and thickness of layers defined by those correlation techniques, the only fit variables are the parameters of the ensemble of Lorentzian oscillators used to describe the plasmon resonance and interband transitions of gold according to the equation:

$$N^2 = (n + ik)^2 = \varepsilon = \varepsilon_1 + i\varepsilon_2 = \varepsilon_\infty + \sum_j \frac{A_j \omega_j^2}{\omega_j^2 - \omega^2 - i\gamma\omega}$$

where ε_∞ is the high-frequency dielectric constant, ω_j, γ_j, and A_j are the frequency, width and strength of the j oscillator.

These correlating techniques are used to build a model in order to derive the Au nanoparticle dielectric function. Similar analysis was used for the Ag/(SiO_2, SiN_x, TiO_2) films on Si(100) substrates [14].

20.3.4 Ellipsometry of Uniaxial Samples Corroborated by Polarized Reflectometry for the True In-Plane Response

As mentioned before, the dielectric function of isotropic, bulk samples can directly be measured using SE. In anisotropic materials, the dielectric function varies with the propagation direction of light and the ellipsometry data analysis using conventional Fresnel equations becomes rather difficult. A special mathematical formalism is necessary for the data calculation of anisotropic structures. Besides that, appropriate experimental setups are required. Generalized ellipsometry comprises theory and experiment of anisotropic samples (see Chap. 10). This can be avoided if the orientation of the optical axes is known "a priori", e.g. using spectrophotometric measurements.

The uniaxial (anisotropic) response of highly oriented pyrolytic graphite (HOPG) in mid-IR was studied using spectroscopic ellipsometry in order to extract parameters of the IR active vibrations of E_{1u} (in-plane mode) and E_{2u} (out-of-plane mode) symmetry [15]. The large in-plane polarizability is seen directly as the large values of the pseudodielectric function and the influence of free electrons is detectable as the steep decrease of its real part with decreasing wavenumber (Fig. 20.12a). At the oblique incidence used in SE, both the in-plane and out-of-plane IR active modes are detectable. In order to observe the true in-plane response, the near-normal incidence
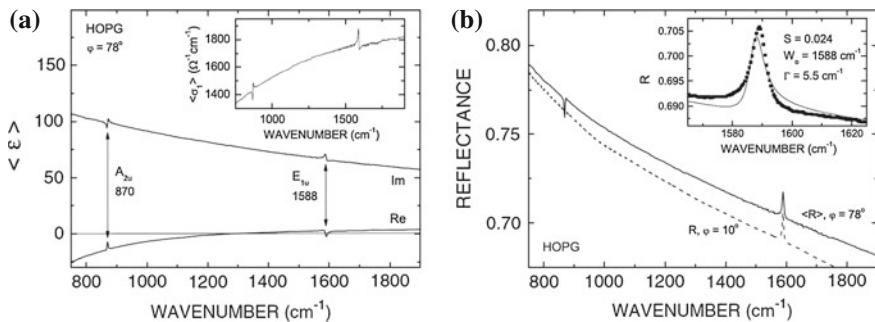

Fig. 20.12 a Real and imaginary parts of the pseudodielectric function of HOPG measured at an angle of incidence of 78 °. The *arrows* indicate the in-plane (E_{1u}) and out-of-plane (A_{2u}) vibrational modes. *Inset* the real part of conductivity computed from the imaginary part of $\langle\varepsilon\rangle$ and **b** the near-normal incidence reflectance (*dashed line*) and the reflectance spectrum calculated from the pseudodielectric function (*solid line*). The *inset* shows the measured data (*symbols*) and the best-fit Lorentzian lineshape. [Reprinted with permission from Ref. [15]. Copyright Elsevier Limited (2011)]
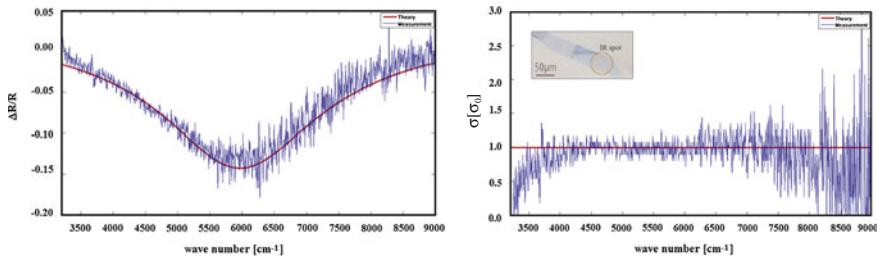
Fig. 20.13 Measured and theoretical normalized reflectance, $(R_{graphene} - R_{substrate})/R_{substrate}$ for the single layer graphene flake on the *left*; and the calculated conductivity in the units of $\sigma_0 = \pi e^2/2h$ (universal optical conductivity); on the *right*. The *inset* represents the flake and the position of the IR spot [18]

reflectivity of HOPG is measured. The results and the comparison with the pseudo-reflectance computed from the ellipsometric data are shown in Fig. 20.12b. While both phonon lines are seen in the latter spectrum, the out-of-plane vibration is absent in the measured reflectance.

20.3.5 Characterization of Single and Few Layer Graphene by IR Spectroscopy Measurements
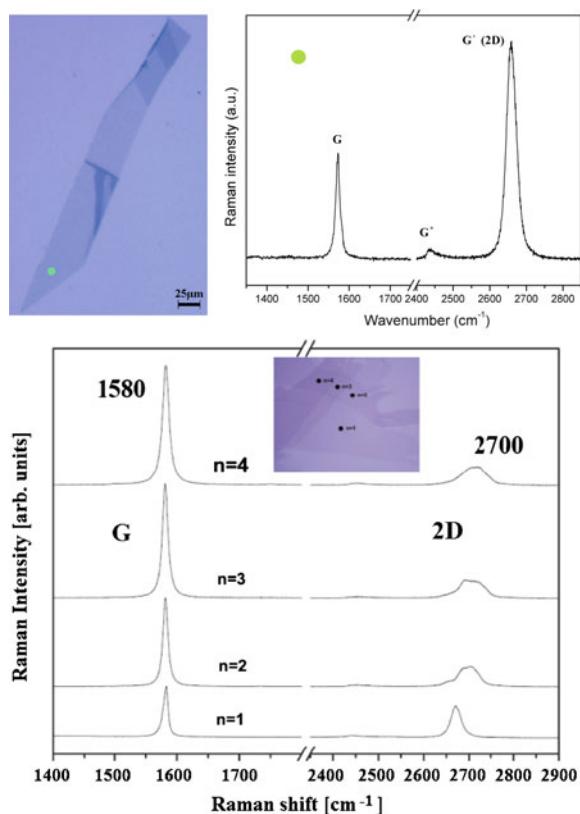
Near-normal NIR reflectance measurements can be applied for determination of the number of graphene layers and subsequently used for the ellipsometry characterization of optical parameters of graphene films (see Chap. 17). Optical conductivity (AC conductivity) is calculated from NIR reflectance measurements carried out under an IR microscope mounted on a FTIR spectrometer. In the NIR, energies are well above $2\mu_c$, where μ_c is the chemical potential, and conductivity saturates to $\sigma_0 = \pi e^2/2h$ for single layer graphene (SLG). Conductivity of few layer graphene (FLG) samples could be interpreted as parallel SLG conductors, meaning that conductivity linearly increased by σ_0 for each layer of the FLG sample. Obtained results are in agreement with both theoretical and experimental data reported by other groups [16, 17] (Fig. 20.13).

20.3.6 Raman Spectroscopy and Ellipsometry of Nanostructures and Thin Films

20.3.6.1 Raman Spectroscopy and Ellipsometry of Graphene

As a suitable example of the correlation between SE and Raman spectroscopy we will consider graphene. Graphene is the first truly 2D system with extraordinary

Fig. 20.14 Large graphene flake (15,000 μm^2) and the corresponding Raman spectra at different points of the flake indicating single and few-layer graphene as well as the Raman spectra of single and few-layer graphene as a function of number of layers [21]

physical properties [19, 20] and its characterization is of outmost importance. For this purpose Raman spectroscopy is proven to be a suitable tool for identifying the number of graphene layers as presented in Fig. 20.14. Raman spectra of single and few-layer graphene consists of two prominent peaks, the G peak represents a normal in plane mode allowed by the symmetry rules whereas the 2D one is a two-phonon mode. Both the G and 2D peaks can be used to characterize graphene flakes before the ellipsometric measurements are carried out. Namely, a symmetric shape of the 2D at 2700 cm^{-1} confirms the presence of a single graphene layer while, on the other hand, the structure of the same peak for few-layer graphene (up to 5 layers) can be exploited for determination of the numbers of layers. It is worth mentioning that the frequency of the G mode can also be used for the same purpose but it did not turn out as a reliable method because of possible impurities and charging of the flakes. When the whole Raman spectrum of a graphene flake is recorded in a single scan it is possible to compare the intensities of the G and 2D peaks. In the case of single-layer graphene the 2D peak is much higher than the G one. It is worth emphasizing that FTIR cannot be used for identifying graphene since, differently from Raman
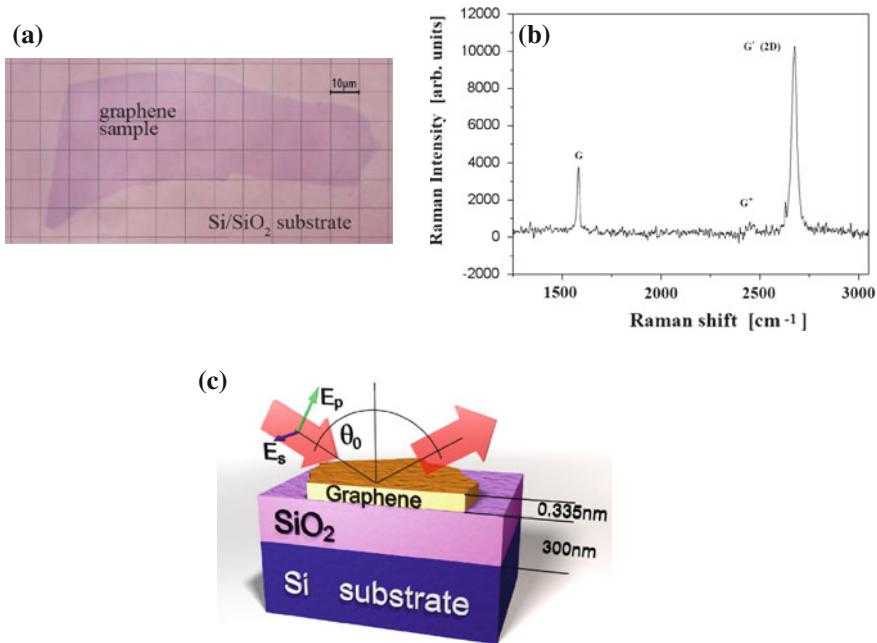
Fig. 20.15 **a** Single-layer graphene on a 300 nm thin layer of SiO_2. Optical micrography and **b** its Raman spectra; the Raman spectrum of a single-layer graphene was obtained in a single scan [23]. **c** The ellipsometric measurement configuration

spectra, graphene has no active normal modes in the IR spectrum. Therefore Raman spectroscopy represents an unavoidable method for graphene characterization.

After identification of graphene on its substrate using optical microscopy, Raman spectroscopy (see the example shown in Fig. 20.15), ellipsometric measurements can be carried out on the same flake and they are shown in Fig. 20.16. The Raman spectrum of a single-layer graphene was obtained in a single scan in order to compare the ratio I_{2D}/I_G of the 2D and G Raman peaks more precisely. For a single graphene layer $I_{2D}/I_G > 1$. More details on the Raman spectroscopy of graphene can be found in a recent review paper [22].

Because of ellipsometry high sensitivity (down to 1 Å) to thickness of layers, SE turned out to be a useful tool [24, 25] for the proper determination of optical parameters of graphene for many different applications. Figure 20.17 shows the complex susceptibility ($\varepsilon = 1 + \chi$), refractive index and conductivity of graphene derived from the ellipsometric measurements shown in Fig. 20.16 of the graphene flakes in Fig. 20.15. Indeed, one of the problems in SE measurements of graphene is the small size of the graphene flake of high quality. Even when the micro-spot optics (approximately $60 \times 110\ \mu m$ in our case at an incidence of 60°) is used, the ellipsometric spot size could be larger than the graphene flake rendering correct ellipsometric data analysis difficult because part of the substrate is also sampled as schematized in
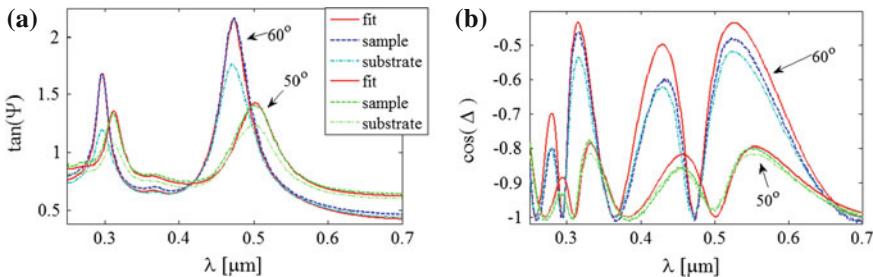
Fig. 20.16 **a** tan(Ψ) and **b** cos(Δ) of the graphene sample (*dashed line*) and of the bare substrate (*dot-dashed line*) for the two different incident angles of 60° and 50°. *Filled lines* represent fitted sample data [23]
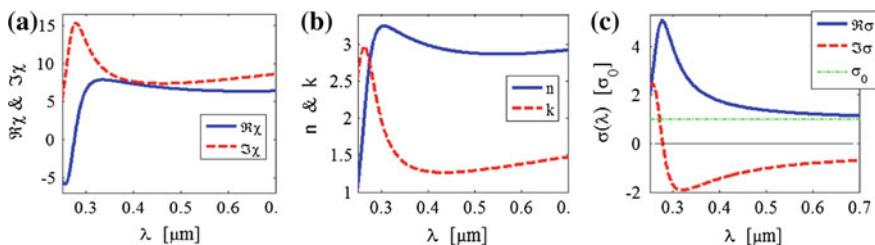

Fig. 20.17 **a** Complex dielectric susceptibility, **b** complex refractive index, and **c** complex optical conductivity of graphene derived by the measured tan(Ψ) and cos(Δ) from Fig. 20.16
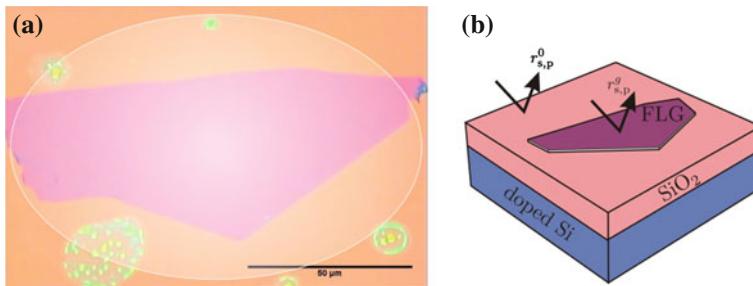

Fig. 20.18 Large few-layer graphene flake and the light beam spot obtained by optical microscopy (**a**) and the sample configuration used for the island model (**b**)

Fig. 20.18. For this reason the island-film model (see [1] and references therein and also Chap. 6 of this book) has been used in data modeling. The idea, as explained in Fig. 20.15b, is that the reflection coefficient r is a weighted average of the reflection coefficient r^0 of the bare substrate and the reflection coefficient r^g with graphene on top, $r = (1 - s)r^0 + sr^g$, (1) where s is the fraction of the surface covered by the graphene flake [26].

(a)

(b)
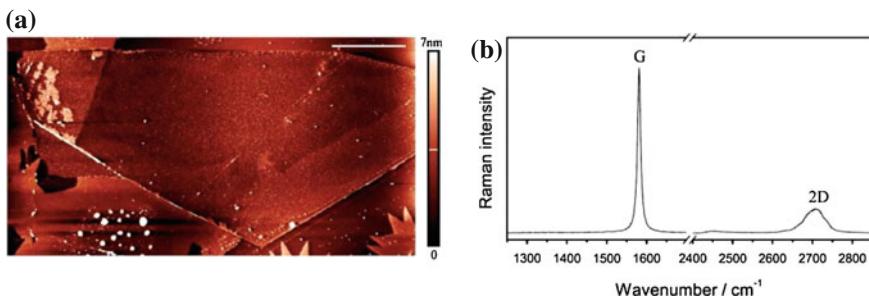
Fig. 20.19 AFM and Raman spectroscopy characterization of a graphene sample. **a** AFM tapping mode topography map of the sample. The *middle part* shows the FLG while some glue residue is seen around it. The size of the *white scale bar* in the *upper right corner* is 20 μm. **b** Raman spectrum of the sample

In order to establish the proper ellipsometric model, the surface fraction, s, covered by the graphene flake and the thickness of the graphene flake have to be known. Therefore, we used optical microscopy, AFM and Raman spectroscopy as shown in Fig. 20.18. The AFM height profile across the flake edge, shown in Fig. 20.19, gave an approximate thickness of 2 nm, while from the I_G/I_{2D} peaks intensity ratio five layers of graphene are inferred Therefore, in the ellipsometric analysis, we entered that thickness to enable a reliable determination of optical parameters of graphene (n, k) in agreement with a few previous reports [24, 25].

20.3.6.2 Raman Spectroscopy and Ellipsometry of ZnO Thin Films

As mentioned before, Raman and IR selection rules are different. If a system has a center of inversion, the normal vibrational modes cannot be both Raman and IR active. Otherwise, vibrational modes are both IR and Raman active as in case of ZnO [27]. Infrared dielectric function spectra and phonon modes of high-quality, single crystalline wurtzite ZnO films were obtained from IR and Raman studies (Fig. 20.20). The optical phonons at the Γ-point of the Brillouin zone belong to the following irreducible representation: $\Gamma_{opt} = 1A_1 + 2B_1 + 1E_1 + 2E_2$. Both the A_1 and E_1 modes are polar and split into transverse (TO) and longitudinal (LO) phonons with different frequencies. The short range interatomic forces cause anisotropy and the A_1 and E_1 modes have different frequencies. They are both IR and Raman active. The two nonpolar IR inactive E_2 modes are only Raman active, and the B_1 modes are silent modes. It turns out that the phonon mode parameters for A_1 and E_1 modes obtained by both techniques are highly consistent. Furthermore, parameters obtained by Raman spectra for A_1 and E_1 modes can be used to make fitting of the IRSE spectra easier or to confirm the quality of the fit. Also, for detection of the two E_2 modes, Raman gives complementary results. They cannot be detected by IRSE, only by Raman measurements.
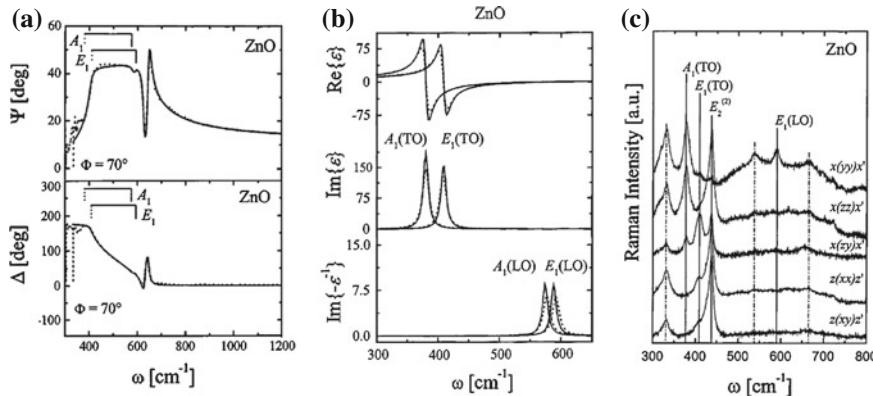
Fig. 20.20 a IR ellipsometric spectra of the ZnO bulk and film samples (*dotted lines*, experimental data; *solid lines*, best-fit spectra). The ZnO phonon modes of A_1 and E_1 symmetry are indicated by *vertical lines* (*solid lines*, TO phonons; *dotted lines*, LO phonons). **b** Spectra for the real and imaginary part of the dielectric function according to the best-fit (*solid lines*, film; *dotted lines*, bulk sample). **c** Raman spectra of the bulk ZnO sample. First order phonon modes of ZnO are indicated by the *vertical solid lines*. The *vertical dashed-dotted lines* mark the features due to multiple-phonon scattering processes. [Reprinted with permission from Ref. [27]. Copyright of the American Institute of Physics (2003)]

20.3.7 Surface Enhanced Raman Spectroscopy and Spectroscopic Ellipsometry

An unusual example of correlation measurements represents a combination of SERS, SEM and SE for the characterization of silver island films [28]. The authors utilized this multidiagnostic method to determine the dielectric functions of silver island films over a large range of sizes and morphologies from the percolation threshold down to an average particle size smaller than 5 nm as shown in Fig. 20.21. They measured films on silicon substrates with 2 and 20 nm oxide layers and compared the surface-enhanced Raman scattering properties of the films. The use of a combinatorial deposition method in creating a gradient film is crucial to reducing the uncertainties associated with multiple sequential depositions and Raman measurements. The statistical analysis from high quality SEM images, combined with SE analysis of the film dielectric functions, and scanning Raman measurements of the relative Raman enhancement on combinatorial films allow nanostructural, optical, and SERS parameters to be directly correlated (see Fig. 20.22). Optimal enhancement is observed when the interparticle spacing is minimized. The use of a combinatorial method resulted in significantly reduced uncertainties by avoiding multiple sample preparations and allowed unambiguous identification of optimal film parameters for the different substrates.
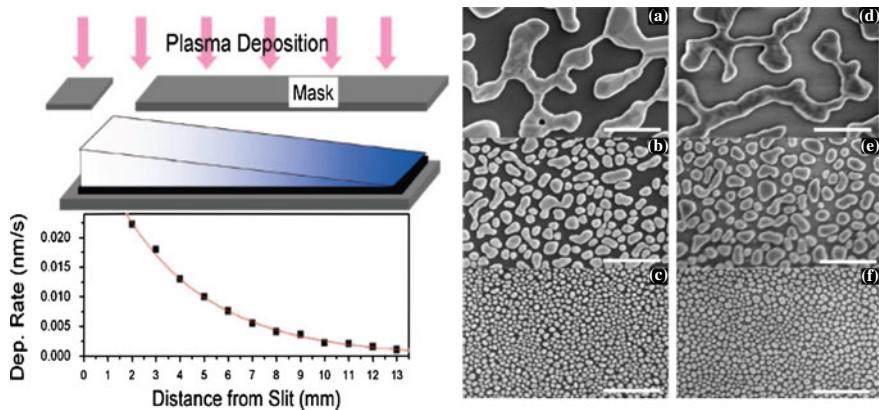
Fig. 20.21 (*Left*) Deposition of silver films by RF magnetron and (*right*) SEM images of silver nanostructures on silicon with **a–c** native oxide (2.7 nm) and **d–f** thermal oxide (19.3 nm). Images correspond to (*x*) distance from slit in mm) **a** 0, **b** 2, **c** 4, **d** 1, **e** 3, and **f**. The scale bar is 500 nm
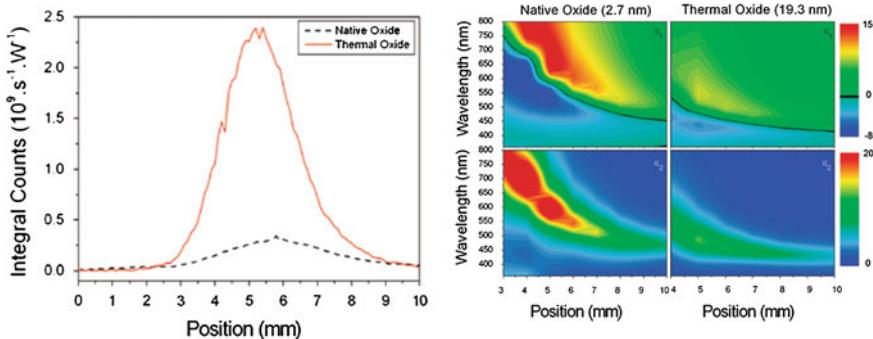
Fig. 20.22 (*Left*) Dependence of the relative Raman enhancement on the substrate and (*right*) the dielectric functions of the silver films as a function of position. The *black line* shows $\varepsilon_1 = 0$, which corresponds to the position of the plasmon resonance maximum

20.3.8 Grain Size Measured by XRD Used in Ellipsometric Modelling

The properties of certain nanomaterials depend on their synthesis. Improving existing processes and developing new ones enabled creation of many different nanocomposites. For nanocomposites, the grain size of the constituting materials and phases is important in rationalizing their properties. The dielectric function also depends on the crystallites and grain size.

Grain size of certain nanocomposites can be measured using different microscopies, but also by the analysis of XRD spectra. The broadening of a peak in a

diffraction pattern in a solid correlates to the size of the particles or crystallites by the Scherrer's equation. Here the example of silver zirconium nitride films is discussed.

Silver zirconium nitride films were deposited by unbalanced magnetron sputtering [29]. They were studied by different techniques like XRD, TEM, XPS and SE. The films consisted of nanocrystals of zirconium nitride embedded in a silver matrix. The grain size was deduced from the width of the XRD peaks using the Scherrer formula and it was found to decrease with addition of silver. The chemical and phase composition was determined from XPS measurements. The optical constants were measured using SE and the correlation between film and structure/composition was established. The ellipsometric data were fitted using a single layer model since the films were optically thick. Based on the Bruggeman effective medium theory the layer was simulated as a mixture of the constituent phases zirconium nitride and silver. The dielectric function of the constituent phases was determined by measuring the dielectric function of zirconium nitride and silver coatings deposited under the same conditions. Good fits were obtained only when the reference dielectric function for zirconium nitride corresponded to a film with a grain size equivalent to the one deduced from the Scherrer formula. XRD and ellipsometric spectra are shown in Fig. 20.23.
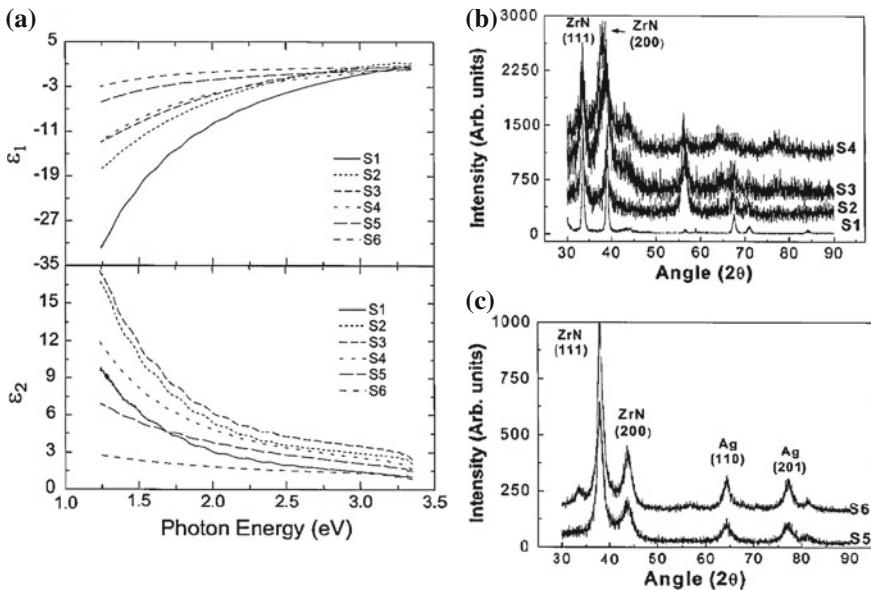

Fig. 20.23 a Real and imaginary part of the dielectric function for different samples. **b** and **c** XRD spectra for the samples used in ellipsometry [29]
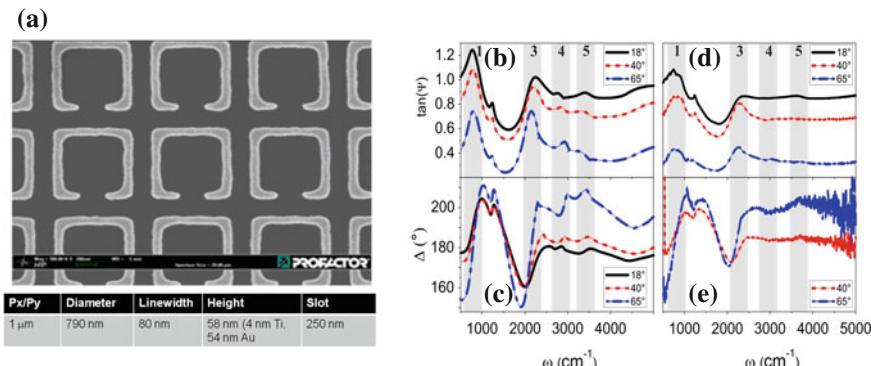
Fig. 20.24 a SEM image of golden split-ring resonators with tabulated geometric parameters. The ellipsometric parameters **b** tan(Ψ) and **c** cos(Δ) are calculated using data extracted from the SEM image. The measured ellipsometric parameters are **d** tan(Ψ) and **e** cos(Δ)

20.3.9 Spectroscopic Ellipsometry of Metamaterials

The most common artificial structures used as building blocks for metamaterials with negative effective permeability are split-ring resonators (SRRs). SRRs are cut metal rings with a non-conducting gap. The spectral positions of the multiple resonances in SRRs are interpreted using the standing wave plasmonic resonance model [30].The positions of these resonances can be measured using SE [31]. Different numerical tools are used for simulation and interpretation of the measured ellipsometric spectra, and they all use the geometrical parameters of the real structure (shape, size, orientation, thickness etc.). Therefore, visualization of SRRs structures is necessary. Since the current goal is to push the EM response to the optical regime, the size of "meta-atoms" (SRRs) is in order of a few hundred nm. A direct imaging of nanostructures is possible only by using TEM, SEM and SPM. One example of ellipsometry of golden split-ring resonators is presented in Fig. 20.24. The geometrical parameters were extracted from the SEM image of the structure (Fig. 20.24a), and the ellipsometric spectra were simulated using rigorous coupled wave analysis (RCWA) (see also Chap. 7 of this book). The simulated data are shown in Fig. 20.24b, c and are in a good agreement with the measured spectra in Fig. 20.24d, e. There are many variables like length of the rings, period of the structure, size of the gap, thickness of the layers etc. Without prior knowledge about all of them, it would be impossible to find a good qualitative agreement between measured and calculated data as it is in this example.

20.4 New Trends

In recent years, there is a new trend regarding ellipsometry and correlation techniques consisting in combining correlation techniques and ellipsometry in the same instrument. The main advantage is that both the correlation and ellipsometric measurements are done on the same sample spot, in addition to the obvious advantages of being less time consuming and allowing easier sample handling, and practically in real time.

20.4.1 Attenuated Total Reflection Combined with Spectroscopic Ellipsometry

As we mentioned above, one of the disadvantages of SE concerns the characterization of samples with small absorption coefficients. Here we describe the case of detection of the Si-H_n modes in the hydrogenated silicon, S:H, thin films where due to small absorption coefficients of Si-H_n vibrations both ellipsometry and even IR transmission spectroscopy are difficult to use. Therefore, a combination of real-time infrared attenuated total reflection (ATR) spectroscopy and SE (Fig. 20.25) has been applied to assess depth profiles of Si-H_n (n =1–2) bonding modes in Si:H thin films [32].

To evaluate the SiH_n bonding modes having low absorption coefficients IR spectroscopy performed by ATR is the most promising technique, owing to very high sensitivity realized by internal multiple reflections of the IR light inside an ATR prism. The proper data analysis procedure for real-time ATR spectra was established by combining ATR with SE measurements since in estimating the SiH_n bonding concentrations in thin Si:H layers, an accurate thickness measurement is required. Real-time SE performed in the ultraviolet /visible range provides an ideal tool for the purpose of thickness determination.
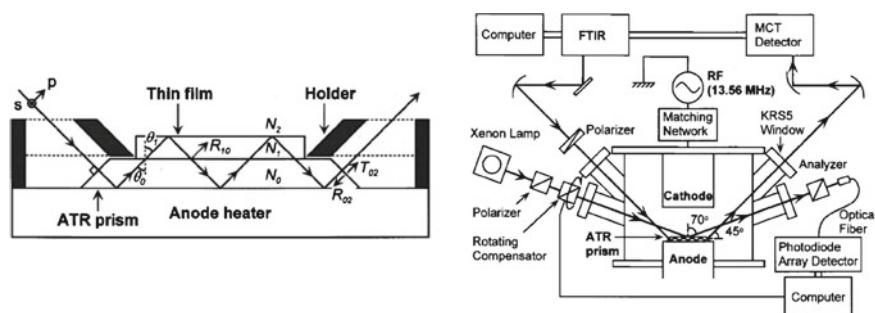

Fig. 20.25 ATR prism and the setup for the real time measurements by SE and ATR. [Reprinted with permission from Ref. [32]. Copyright American Institute of Physics (2002)]

20.4.2 Spectroscopic Ellipsometry Corroborated by Glow Discharge Optical Emission Spectroscopy

Recently Horiba Jobin Yvon successfully combined ellipsometry and Glow Discharge Optical Emission Spectroscopy (GDOES). They studied Copper Indium Selenide (CuInSe$_2$) 2 micron thick films. CuInSe$_2$ has extremely high optical absorption coefficient that allows nearly all of the available light within the terrestrial spectrum to be absorbed in the first micrometer of the material. Therefore, the total thickness of the active layers is of the order of 2 μm, resulting in the efficient use of materials without negatively impacting the conversion efficiency. The addition of controlled amounts of gallium and/or sulfur into the CuInSe$_2$ absorber layer permits the adjustment of its energy gap to provide device with higher voltage, better carrier collection and higher conversion efficiency. Various analytical techniques are applied to understand the growth mechanism, to monitor and control the optical properties and the composition, to verify the thickness of the layers during deposition, to profile the layers and to see gradients. Among the panel of techniques available, SE and Glow Discharge Optical Emission Spectroscopy stand out due to their ease of use and the complementary information they provide. Glow Discharge Optical Emission Spectroscopy is a technique that relies on the controlled erosion of a representative part of a material by dense RF plasma. It provides the elemental distribution as a function of the penetration depth it offers a nanometer depth resolution and it is ideal to detect composition changes, interface contaminations and gradients (www.horiba.com/scientific). In order to characterize photovoltaic materials, Horiba recently also combined SE with various spectroscopic methods including: Raman Spectroscopy Glow Discharge Optical Emission Spectroscopy and Plasma Profiling Time of Flight Mass Spectroscopy

20.4.3 Grazing X-Ray Reflectometry Combined with Spectroscopic Ellipsometry

SOPRALAB has developed a new approach by combining two non-destructive characterization techniques on the same instrument: Spectroscopic Ellipsometry (SE) and Grazing X-Ray Reflectometry (GXR) (Fig. 20.26). GXR is a powerful technique to measure the absolute thickness without a-priori structural modelling of a multi-layer stack. GXR applies to any type of material: semiconductor, metallic, dielectric, organic. GXR and SE measurements can be performed quasi-simultaneously at exactly the same sample location. Analysis of complementary data from both techniques with the same physical model leads to unprecedented accuracy in measurement results.

Grazing X-Ray Reflectometry (GXR) is certainly the technique of choice for absolute thickness measurement of new materials. At this energy, every material displays the same refractive indices (n being 1 and k being 0), simplifying consid-
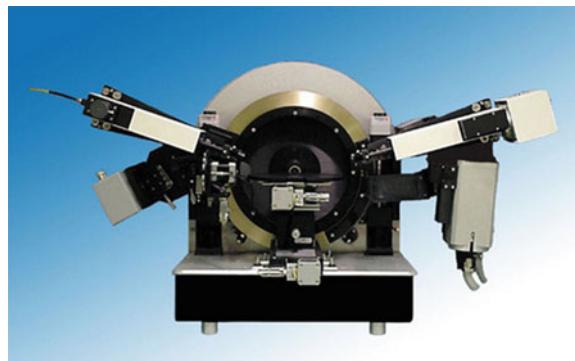
Fig. 20.26 Spectroscopic ellipsometry and grazing X-ray reflectometry

erably the structural model for data analysis. In the case of complex multi-layers, accurate information can be obtained on the periodicity of stacks as well as on the interlayer structures. SE combined with GXR is the choice for the following applications: High K dielectrics (ZrO_2, nitrided oxides), thin ONO layers, thin optical coatings, Si and SiGe thin epitaxial structures, thin metallic and organic layers, and coatings for micro-lithography.

20.4.4 Correlating Techniques in Imaging Ellipsometry

20.4.4.1 Ellipsometry and Near-Field Optics

Borchagovsky theoretically demonstrated the capability of an ellipsometric detection method for near-field optical measurements [33]. He showed that the change in the state of polarization of scattered light due to tip-protrusion interaction is sufficient for detection in most of the analyzed cases, even without resonant enhancement. This suggests that ellipsometry of scattered light may be used in near-field optics for the investigation of nanosize structures [34] experimentally demonstrated by near-field ellipsometry. They created an apertureless optical near-field scanning microscope system combining a commercially available AFM and an ellipsometer without any prior changes in design of the respective devices. In preliminary experiments, an optical resolution of about 20 nm ($\lambda/32$) has been achieved using the combined microscope. They called this technique scanning near-field ellipsometric microscopy (SNEM) (Fig. 20.27).

The AFM is placed on the sample so that the tip is over the spot illuminated by the ellipsometer. During the measurement, the AFM operates in contact mode. By scanning the tip within the evanescent field, null ellipsometry conditions are disturbed. The detector registers a change in the optical signal which is simultaneously displayed with the topography as a two-dimensional image (Fig. 20.28). It is

Fig. 20.27 Experimental setup of SNEM where P denotes a polarizer, C—compensator, A—analyzer, and D—detector. [Reprinted with permission from Ref. [34]. Copyright American Institute of Physics (2001)]
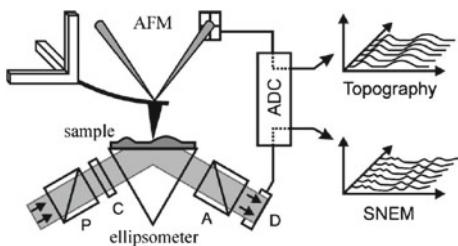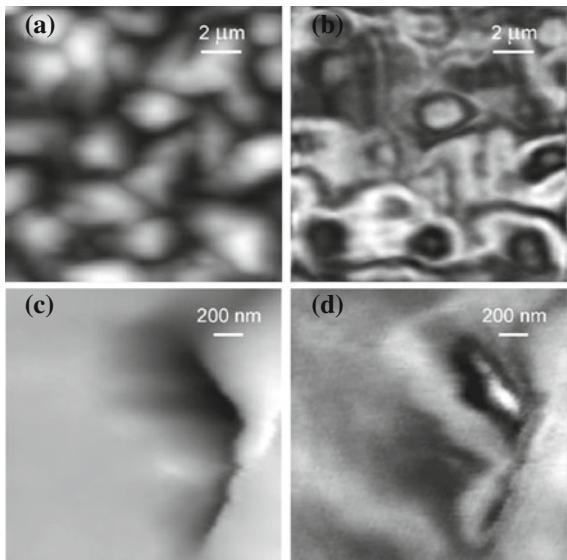
Fig. 20.28 Topography (**a**), (**c**) and SNEM (**b**), and (**d**) images of a polycrystalline film of a thermotropic liquid crystal. [Reprinted with permission from Ref. [34]. Copyright American Institute of Physics (2001)]

interesting to note that the structures were better resolved in a SNEM than in an AFM image.

20.4.4.2 Imaging Ellipsometry and Scanning Probe Microscopy

In order to add imaging to an ellipsometer, an objective and a spatially resolving detector, e.g. a sensitive CCD camera, are needed. The objective images the illuminated area of the sample onto the camera. As a consequence, areas that have different optical properties cause a different signal in the camera image. Especially those areas that are currently fulfilling the condition of the ellipsometric "Null", i.e. where the optical parameters are such that the light reflected is extinguished for that particular setting of polarizer (P), compensator (C) and analyzer (A) angles, and will appear dark in the image. Higher light intensities are incident at the detector from areas of the sample that do not meet that condition, producing brighter image regions. Changing the settings of P, C, and A it is now possible to find the Null also for these regions,

which will cause the former dark areas to appear now bright (see Fig. 20.30a). The main advantage of such an imaging ellipsometer is that the signal is not the average over an entire laser beam spot on the sample, but it is spatially resolved to show the details of the sample. With the imaging ellipsometer it is possible to get not only immediate qualitative information but it is also possible to restrict ellipsometric analysis to a particular region of interest within the field-of-view. Applying appropriate algorithms allows mapping the Nulls for the entire image. This yields a two-dimensional map of the ellipsometric data that can be transformed into a thickness map of the sample. One of the advantages of the null ellipsometry is the fact that one measures angles instead of light flux (($\Psi, \Delta) = f(P, C, A)$), partly avoiding problems due to the stability of the light source or non-linearity of the detectors. Imaging ellipsometry combines the power of ellipsometry with microscopy and overcomes the limits of classical ellipsometers. With imaging ellipsometry the size of the object under investigation may be reduced to the micrometer range, accomplishing the need to measure on micro-structured samples. For example, reactive coatings on cantilever-based microsensors can be characterized by imaging ellipsometry, being inaccessible to the large spot size of conventional ellipsometers. Simultaneously, it is possible to generate maps of ellipsometric data with spatial resolution down to a micrometer, while preserving the sub-nm film thickness resolution of a classical ellipsometer. Presently, imaging ellipsometry is combined with AFM for characterization of nanostructures and nanofilms (sometimes this combination is called scanning probe ellipsometric microscopy—SPEM). Such instruments have made commercially available in the last few years (Fig. 20.29). In general, the thickness is obtained by ellipsometry and the topography by atomic force microscopy. Imaging ellipsometry enables not only the quantitative determination of the thickness but also the optical properties with a lateral resolution down to 1 μm. By performing ellipsometric contrast pictures a real time imaging of nanofilms is a unique feature of this technique. The integration of a scanning probe microscope into the setup of an imaging ellipsometers ensures rapid evaluation of optical properties and subsequent zooming-in with sub-micron lateral resolution by means of SPM on the same field of view. A typical task of the instrument is to simplify the search for regions of interest on a sample; the preparatory work in terms of identification of thickness variations is performed by the imaging ellipsometer. This is not only timesaving; it also circumvents the artifacts produced by scanning probe microscopy (SPM). Once a region of interest has been defined, the sub-micron lateral resolution of the SPM overcomes the constraints of the imaging ellipsometry caused by the optical diffraction limit. The advantage of combining AFM and imaging ellipsometry stems from the convenience of using imaging ellipsometry to visualize thin films and surface structures, and then zoom into nanometer details with SPM on the same spot!

As an example, the SPEM analysis of few layer graphene on the n-type GaAs substrate is discussed here.

The advantage of AFM measurements is that it seems to be the most direct way to identify graphene thickness but the measurements are slow and could easily damage the surface of the sample during measurements. Still, the obtained thickness is often
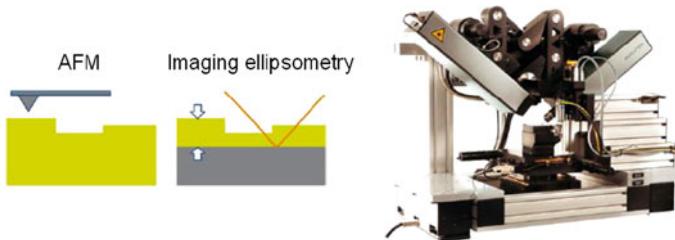
Fig. 20.29 AFM and ellipsometry (*left panel*) and an integration of AFM and IElli (http://www.accurion.com)

hard to interpret due to adsorbents on the flake under ambient conditions. Imaging ellipsometry is a powerful tool to detect and characterize graphene on any flat substrate. The results of the imaging ellipsometry measurements on a graphene flake are shown in Fig. 20.30c–e.
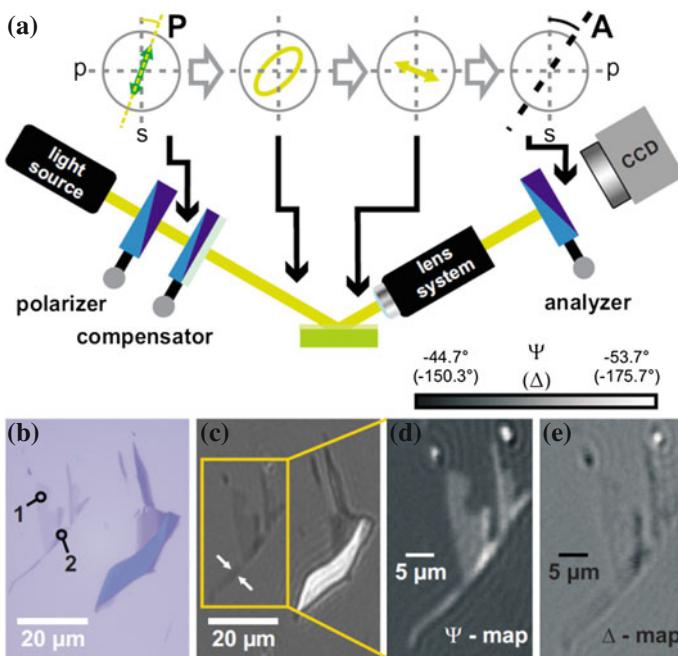
Fig. 20.30 a Imaging ellipsometry setup. The lens system mounted between the sample and analyzer allows imaging with submicron lateral resolution. **b** Optical image and **c** imaging ellipsometric intensity image of a sample on SiO$_2$/Si showing regions with graphene monolayer covering up to thin graphite. Numbers in (**b**) correspond to the layer number. **d** Ellipsometric → map and **e** the corresponding _ map of the boxed region display grapheme mono- and bilayer areas with higher resolution. Copyright American Institute of Physics (2010)
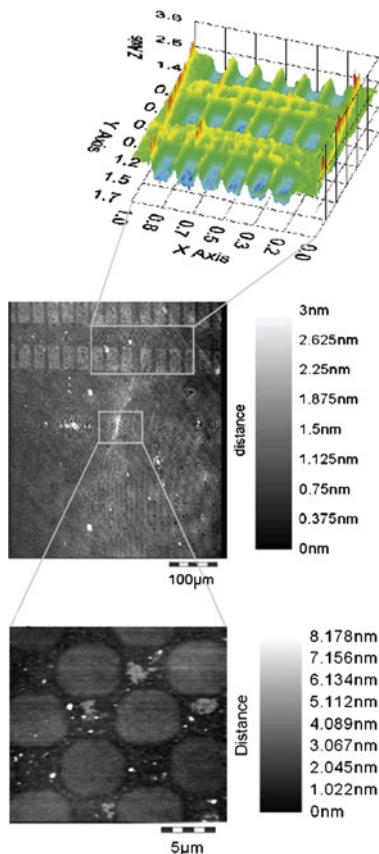
Fig. 20.31 Scanning probe ellipsometry of the micro-contact printed thiol on glass. From the *top* to the *bottom* 3D profile by the SPEM mapping mode, ellipsometric microscopic contrast image, and the thickness map recorded by AFM

Another example of the full integration of imaging ellipsometry and AFM—the scanning probe ellipsometric microscope (SPEM)—is the measurement of the micro-contact printed thiols on glass, as shown in Fig. 20.31.

The measurement has been carried out at the same spot on the sample using software-controlled sample transport between the ellipsometer and SPM with a μm accuracy Nanofilm (Accurion now) application note on scanning probe ellipsometric microscope. The following figure shows a 3D profile obtained with the SPEM mapping mode together with the ellipsometric contrast-image and a corresponding AFM image.

Interesting combinations of ATR-FTIR and imaging ellipsometry corroborated by an in situ AFM were successively used for the investigation of structural and physical properties of bilayers of the hydrocarbon and the fluorocarbon segments in bilayers of 1,2- dipalmitoyl-*sn* glycero-3-phosphocholine (DPPC) [36]. Interested

readers can take a look at a review paper on applications of imaging ellipsometry in biotechnology [37].

20.4.5 Scanning Probe Ellipsometric Microscopy and Grazing Incidence Small Angle X-Ray Scattering

Recently a new combination of grazing incidence small angle X-ray scattering (µGISAXS) and imaging ellipsometry was used for the in situ characterization of colloidal polystyrene nanoparticles [38] in HASYLAB. It turned out that this combination introduced a new versatile tool for the characterization of nanostructures. µGISAXS provides a local lateral and depth-sensitive structural characterization, and imaging ellipsometry adds the position-sensitive determination of the three-dimensional morphology in terms of thickness, roughness, refractive index, and extinction coefficient. Together µGISAXS and imaging ellipsometry enable a complete characterization of structure and morphology.

In contrast to AFM and SEM/TEM, µGISAXS allows large sample areas to be probed and yield structural information about nanosize objects. For the experimental setup the authors used a commercial surface probe ellipsometric microscope (SPEM) integrated into a GISAXS beamline as in Fig. 20.32. The SPEM comprises an imaging ellipsometer from Nanofilm Technologies-GmbH and an atomic force microscope from Surface Imaging Systems. There is obvious benefit in using imaging ellipsometry instead of optical microscopy since a position-sensitive determination of the 3D morphology in terms of thickness, roughness, refractive index, and extinction coefficient becomes possible instead of a simple picture of the sample surface. The results of these complementary techniques are presented in Fig. 20.33.
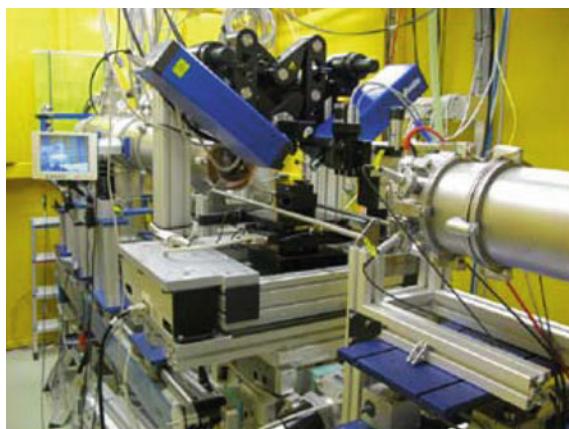
Fig. 20.32 Ellipsometer installation at the GISAXS beamline at HASYLAB (DESY, Hamburg, Germany)
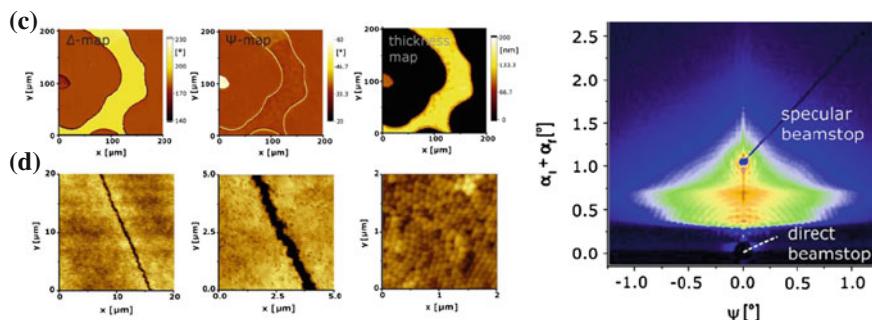
Fig. 20.33 c The Δ and Ψ maps of the dewetted polymer structure obtained with IElli and the resulting thickness map; **d** AFM topography and the two-dimensional (2D) microbeam μGISAXS pattern measured at the center of the colloidal drop at the right

20.5 Conclusion

We have emphasized the role of correlation techniques that corroborate ellipsometric measurements in order to apply proper models on experimental data, enabling a reliable extraction of optical and compositional parameters of the samples under investigation. The most frequently used correlation methods have been described and demonstrated on chosen examples. Since every paper related to ellipsometry contains one or a few corroborating techniques it is practically impossible to describe all of these methods in detail and cite all relevant papers. Therefore, in this limited communication we confined ourselves to the most important techniques in our opinion, as well as the ones we used in our investigations. Also, an overlook on the new trends in ellipsometry has been provided.

Combining one or more correlating techniques with spectroscopic ellipsometry or imaging ellipsometry assembled on the same instrument is going to be one of the most fascinating fields of research on nanomaterials from both instruments-development and application points of view to gain deeper knowledge.

References

1. H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (Wiley, New York, 2007)
2. H.G. Tompkins, E.A. Irene, *Handbook of Ellipsometry* (William Andrew Publishing, New York, 2005)
3. Accurion—nanofilm_ep3se (Online). http://www.accurion.com/nanofilm_ep3se/

4. M. Losurdo, M. Bergmair, G. Bruno, D. Cattelan, C. Cobet, A. de Martino, K. Fleischer, Z. Dohcevic-Mitrovic, N. Esser, M. Galliet, R. Gajic, D. Hemzal, K. Hingerl, J. Humlicek, R. Ossikovski, Z.V. Popovic, O. Saxl, Spectroscopic ellipsometry and polarimetry for materials and systems analysis at the nanometer scale: state-of-the-art, potential, and perspectives. J. Nanopart. Res. **11**, 1521–1554 (2009)
5. M. Losurdo, Applications of ellipsometry in nanoscale science: needs, status, achievements and future challenges. Thin Solid Films **519**, 2575–2583 (2011)
6. W. Kiefer, J. Raman Spectrosc. **38**(6), 597–796 (2007)
7. Combined metrology ges5e-gxr SOPRALAB (Online). http://www.sopra-sa.com/ges5e-gxr-pdt-1-14.php
8. Z. Guo, L. Tan, *Fundamentals and Applications of Nanomaterials* (Artech House, Norwood, 2009)
9. R.W. CollinsU, I. An, H. Fujiwara, J. Lee, Y. Lu, J. Koh, P.I. Rovira, Advances in multichannel spectroscopic ellipsometry. Thin Solid Films **313–314**, 18–32 (1998)
10. T.A. Mykhaylyk, N.L. Dmitruk, S.D. Evans, I.W. Hamley, J.R. Henderson, Comparative characterisation by atomic force microscopy and ellipsometry of soft and solid thin films. Surf. Interf. Anal. **39**, 575–581 (2007)
11. D.E. Aspnes, Expanding horizons: new developments in ellipsometry and polarimetry. Thin Solid Films **455–456**, 3 (2004)
12. L. Armelao, M. Pascolini, G. Bottaro, G. Bruno, M.M. Giangregorio, M. Losurdo, G. Malandrino, R. Lo Nigro, M.E. FragalaI, E. Tondello, Microstructural and optical properties modifications induced by plasma and annealing treatments of lanthanum oxide sol-gel thin films. J. Phys. Chem. C **113**, 2911–2918 (2009)
13. M. Losurdo, M.M. Giangregorio, G. Bruno, D. Yang, E.A. Irene, A.A. Suvorova, M. Saunders, Er2O3 as a high-K dielectric candidate. Appl. Phys. Lett. **91**, 091914 (2007)
14. E. Masetti, J. Bulir, S. Gagliardi, V. Janicki, A. Krasilnikova, G. Di Santo, C. Coluzza, Ellipsometric and XPS analysis of the interface between silver and SiO2, TiO2 and SiNx thin films. Thin Solid Films **455–456**, 468–472 (2004)
15. J. Humlicek, A. Nebojsa, F. Munz, M. Miric, R. Gajic, Infrared ellipsometry of highly oriented pyrolytic graphite. Thin Solid Films **519**, 2624–2626 (2011)
16. Z.Q. Li, E.A. Henriksen, Z. Jiang, Z. Hao, M.C. Martin, P. Kim, H.L. Stormer, D.N. Basov, Dirac charge dynamics in graphene by infrared spectroscopy. Nat. Phys. **4**, 532–535 (2008)
17. K.F. Mak, M.Y. Sfeir, Y. Wu, C.H. Lui, J.A. Misewich, T.F. Heinz, Measurement of the optical conductivity of graphene. Phys. Rev. Lett. **101**, 196405 (2008)
18. R. Gajic, A. Matkovic, U. Ralevic, G. Isic, M. Jakovljevic, B. Vasic, Dj. Jovanovic, Optical spectroscopy of single and few-layer graphene, in *XVIII Symposium on Condensed Matter Physics*, Belgrade, 13–22 April 2011, p. 41
19. A.K. Geim, K.S. Novoselov, The rise of graphene. Nat. Mater. **6**, 183–191 (2007)
20. A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, The electronic properties of graphene. Rev. Mod. Phys. **81**, 109–162 (2009)
21. R. Kostić, M. Mirić, T. Radić, M. Radović, R. Gajić, Z.V. Popović, Optical characterization of graphene and highly oriented pyrolytic graphite. Acta Phys. Polonica A **116**, 718 (2009)
22. L.M. Malard, M.A. Pimenta, G. Dresselhaus, M.S. Dresselhaus, Raman spectroscopy in graphene. Phys. Rep. **473**, 51–87 (2009)
23. A. Matkovic, U. Ralevic, G. Isic, M.M. Jakovljevic, B. Vasic, Dj. Jovanovic, I. Milosevic, D. Markovic, R. Gajic, Spectroscopic ellipsometry and Fano resonance modeling of graphene optical parameters. Phys. Scr. **T149**, 014069 (2012)
24. V.G. Kravets, A.N. Grigorenko, R.R. Nair, P. Blake, S. Anissimova, K.S. Novoselov, A.K. Geim, Spectroscopic ellipsometry of graphene and an exciton-shifted van Hove peak in absorption. Phys. Rev. B **81**, 155413 (2010)
25. J.W. Weber, V.E. Calado, M.C.M. van de Sanden, Optical constants of graphene measured by spectroscopic ellipsometry. Appl. Phys. Lett. **97**, 091904 (2010)
26. G. Isić, M. Jakovljević, M. Filipović, Dj. Jovanović, B. Vasić, S. Lazović, N. Puač, Z. Lj. Petrović, R. Kostić, R. Gajić, Spectroscopic ellipsometry of few-layer graphene. J. Nanophotonics **5**, 051809 (2012)

27. N. Ashkenov, B.N. Mbenkum, C. Bundesmann, V. Riede, M. Lorenz, D. Spemann, E.M. Kaidashev, A. Kasic, M. Schubert, M. Grundmann, Infrared dielectric functions and phonon modes of high-quality ZnO films. J. Appl. Phys. **93**, 126 (2003)
28. T.W.H. Oates, H. Sugime, S. Noda, Combinatorial surface-enhanced Raman spectroscopy and spectroscopic ellipsometry of Silver island films. J. Phys. Chem. C **113**, 4820–4828 (2009)
29. S.M. Aouadia, M. Debessai, P. Filip, Zirconium nitride/silver nanocomposite structures for biomedical applications. J. Vac. Sci. Technol. B **22**, 1134 (2004)
30. C.Y. Chen, S.C. Wu, T.J. Yen, Experimental verification of standing-wave plasmonic resonances in split ring resonators. Appl. Phys. Lett. **93**, 034110 (2008)
31. M. Jakovljević, B. Vasić, G. Isić, R. Gajić, T. Oates, K. Hinrichs, I. Bergmair, K. Hingerl, Oblique incidence reflectometry and spectroscopic ellipsometry of split-ring resonators in infrared. J. Nanophotonics **5**, 051815 (2012)
32. H. Fujiwara, M. Kondo, A. Matsuda, Depth profiling of silicon-hydrogen bonding modes in amorphous and microcrystalline Si:H thin films by real-time infrared spectroscopy and spectroscopic ellipsometry. J. Appl. Phys. **91**, 4181 (2002)
33. E.G. Bortchagovsky, Ellipsometric method in near-field recording. Surf. Interf. Anal. **27**, 482–485 (1999)
34. P. Karageorgiev, H. Orendi, B. Stiller, L. Brehmer, Scanning near-field ellipsometric microscope-imaging ellipsometry with a lateral resolution in nanometer range. Appl. Phys. Lett. **79**, 1730 (2001)
35. U. Wurstbauer, C. Röling, U. Wurstbauer, W. Wegscheider, M. Vaupel, P.H. Thiesen, D. Weiss, Imaging ellipsometry of grapheme. Appl. Phys. Lett. **97**, 231901 (2010)
36. S. Schuy, S. Faiss, N.C. Yoder, V. Kalsani, K. Kumar, A. Janshoff, R. Vogel, Structure and thermotropic phase behavior of fluorinated phospholipid bilayers: a combined attenuated total reflection FTIR spectroscopy and imaging ellipsometry study. J. Phys. Chem. B **112**, 8250–8256 (2008)
37. A. Eing, M. Vaupel, *Imaging Ellipsometry in Biotechnology* (Verlag Axel Geispeck, Gottingen, 2002)
38. V. Körstgens, J. Wiedersich, R. Meier, J. Perlich, S.V. Roth, R. Gehrke, P. Müller-Buschbaum, Combining imaging ellipsometry and grazing incidence small angle X-ray scattering for in situ characterization of polymer nanostructures. Anal. Bioanal. Chem. **396**, 139–149 (2010)

Chapter 21
Nanotechnology: Applications and Markets, Present and Future

Ottilia Saxl

Abstract The effects of nanotechnology, the application of nanoscience, which owes its early success to the IT industry, are already being felt across society. Because nanotechnology can now be applied so widely, it is called an *enabling* technology. There are few, if any, industries or products that could not benefit from the application of nanotechnology. It is not unreasonable to predict that nanotechnology will completely revolutionise industry and industrial processes within a decade. This chapter examines the economic and social potential of nanotechnology, and put figures on this potential, setting the scene for the remainder of the book which addresses how the most important issue of real time, non-destructive or invasive characterization at the nanoscale is being resolved.

21.1 Nanotechnology: Applications and Markets

21.1.1 Nanotechnology: Already of Global Economic Importance

Today we understand more and more about how the properties of everything in the world around us (including ourselves!) are defined by their component atoms and molecules. Already we are at the early stages of being able to manipulate these atoms and molecules in order to design new things with the features we want, improve medical treatments, the way we generate energy and process our food. This is nanoscience. The effects of nanotechnology, the application of nanoscience, which owes its early success to the IT industry, are already being felt across society. Because nanotechnology can now be applied so widely, it is called an *enabling* technology. For example, nanotech underpins many novel industrial developments, such as fabric treatments

O. Saxl (✉)
NANO Magazine, ION Publishing Ltd, Glasgow, UK
e-mail: ottilia.saxl@nanomagazine.co.uk

that repel stains, non-toxic paints that prevent rusting and biofouling, surface coatings that can kill bacteria and in more effective ways to deliver drugs.

We only need to think of the characteristics of a mobile phone, which today is in effect a mini computer, to appreciate just how many features can be incorporated into a product through micro and nanotechnology. A mobile phone can transmit sound, data, pictures, act as a video camera, a bank, a global positioning system, a health monitoring system, a teaching tool and an active (rather than passive) news provider. In essence, nano–enabled products are smaller, cheaper, lighter, faster, cleverer and often dramatically different to what they supercede. In theory, these new nano products should also use fewer raw materials and consume less energy to make. The developing countries are leapfrogging old technology and aiming to produce these new multifunctional products from the outset, in industries as diverse as aerospace, automotive, materials, foods, textiles and medicine, as well as in military, electronics and telecoms applications. There are few, if any, industries or products that could not benefit from the application of nanotechnology, and those that do not embrace the technology will be left behind by their competitors who do. It is not unreasonable to predict that nanotechnology will completely revolutionise industry and industrial processes within a decade.

As of March, 2011, the nanotechnology consumer products inventory contains 1317 products or product lines as shown in Fig. 21.1. Indeed, the inventory quoted suffers from the lack of a clear definition of what is a nanoproduct, and so is unlikely to be comprehensive, and should be considered only as a guide to the possible scale of market penetration.

Products were grouped according to relevant main categories (Fig. 21.2) that are loosely based on publicly available consumer product classification systems. The largest main category is Health and Fitness, with a total of 738 products. This includes products like cosmetics and sunscreens.
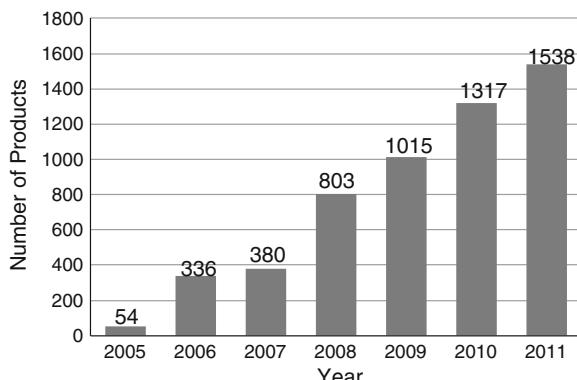
Fig. 21.1 Number of total products listed, by date of inventory update (*source* http://www. nanotechproject.org/inventories/consumer/analysis_draft/)
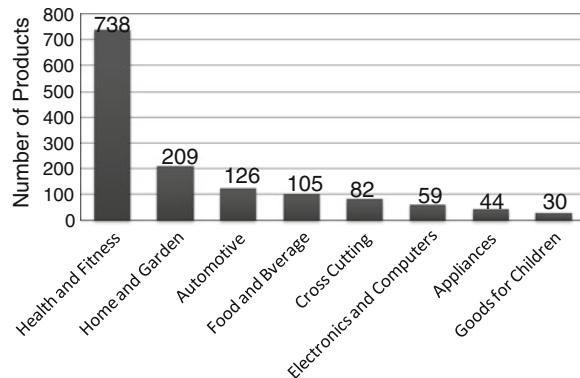
The message is clear—the potential of nanotechnology is vast. However, there are some critically important sticking points that need to be quickly overcome before the benefits of nanotechnology are fully realized. These are: repeatability, stabilisation and characterization; latter being essential to establishing the first two.

These factors have determined so far the limited number of raw materials that constitute, so far, the main nanotechnology products on the market (see Fig. 21.3)

This chapter examines the economic and social potential of nanotechnology, and put figures on this potential, setting the scene for the remainder of the book which addresses how the most important issue of real time, non-destructive or invasive characterization at the nanoscale is being resolved.
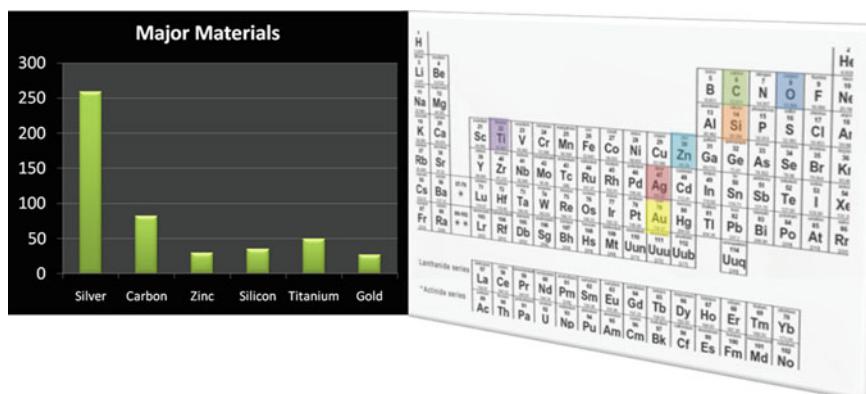
Fig. 21.3 Main elements constituting nanotechnology products on the market

21.1.2 Nano in the Home

If we take a whistle-stop tour of the nanotechnology products around us, we can usefully start in our homes. In a modern house, there could be stay-clean windows, solar collectors, dirt repellent carpets, curtains and chair coverings, anti-bacterial surfaces in the kitchen and toilet, stay-clean exterior paints, anti-corrosion coatings for radiators and gutterings, security lighting, lightweight and transparent wall and roof insulation and furniture treatments, all inside a structure made of environmentally friendly lightweight concrete.

If we examine the 'nano' behind some of these new products in a little more detail. To achieve a self-cleaning surface, there are two principal 'nano' techniques—one using super-hydrophobic and the other using super-hydrophilic materials. New structural paints can cleverly impart water and dirt repellency to whichever surface they are applied to. This is through emulating the Lotus effect, the way the leaves of the Lotus plant repel water through their special structure. Superhydrophilic surfaces on the other hand can be obtained by coating glass, ceramic tiles or plastics with nanoparticles of a semiconducting photocatalyst, titanium dioxide (TiO_2). If TiO_2 is illuminated by light, any grease, dirt and organic contaminants are decomposed and are easily swept away by rainwater. By definition, the size of a nanoparticle is below the wavelength of light, so the bonus is that the coating on the glass is therefore invisible.[1]

Even the technology for non-stick frying pans has switched to "nano" by using nanocomposites to provide heavy duty non sticking surface finishes. The idea of using silver for antimicrobial purposes has been adopted in many non-medical products. As well as in filters for air-conditioning, silver has also been introduced into washing machines and refrigerators for its antibacterial effect (Table. 21.1).

Nanoporous materials also offer superior thermal insulation. Although silica aerogels are the world champions in low thermal conductivity, their brittleness and high price have so far prevented a widespread use. Nevertheless, nanoporous materials have been developed in the form of flexible blankets or evacuated panels for better thermal management, vital in energy saving.

Table 21.1 Uses of nanotechnology products in the home

Dirt resistant facade paints
Self-cleaning coatings (windows etc)
Easy clean hygienic surfaces (in bathrooms, on food preparation surfaces)
Aerogels for lightweight insulation, and reduced flammability
Stay clean textiles (for carpets, furniture)

[1] http://www.nanoes.com/Reference/Self-Clean%20Materials.pdf

> **The Market for Nano-Enabled Homecare Products**
> According to Euromonitor, the sale of homecare products in the UK in 2010 was 3.7 billion UK£. Say the affluent world consists of only 1 billion people, then let's speculate that sales wolrdwide of homecare products to that community alone could be around 50 billion UK£, extrapolating the Euromonitor figures for the UK. If nano penetration is only 10 %, then the nano market worldwide for homecare products is already a not inconsiderable 5 billion UK£, or about 7.5 billion USD. As the world population increases, as wealth increases and the penetration of nano increases, it is not unreasonable to expect that nano influenced household products might reach 10 billion USD easily by 2015, and possibly much more.

Of course, nanotechnology represents not only a commercial opportunity but also to some, a considerable threat—and not just in product differentiation. For example, still in the area of household products, companies involved in manufacturing cleaning products could be threatened by self-cleaning coatings. Similarly detergent manufacturers might be disadvantaged by the development of dirt repellent textiles.

21.1.3 Up Close and Personal: A Nano Revolution in the Sunscreen and Cosmetic Industries

Nanoparticles serve many purposes in cosmetics: they enhance the properties and acceptability of cosmetics by providing softness, lustre, moisturizing and optical effects; they can protect the skin through sunscreens incorporating UV-filters. Most cosmetic manufacturers are extremely interested in the new properties nano brings to their formulations.

Ultra small capsules called liposomes were introduced into cosmetics as far back as 1986 by L'Oreal, a company that holds many nanotech patents. Liposomes are small containers or capsules ranging in size from below 100 nm to several micrometres. They are able to contain and deliver active ingredients such as drugs, vitamins and other cosmetic materials.

The same ingredients as found in common cosmetic formulations can be formulated in a variety of ways at the nanoscale—as nanospheres, nanocapsules, oleosomes and liposomes. These nano-based structures are important because they often improve the stability of the cosmetic and the active ingredients that they contain. Nano encapsulation techniques have led to many improvements in cosmetic products, such as the time-controlled release of scents, which means a more constant odour, as opposed to being overpowering immediately after application, but then the scent disappearing

quickly.[2, 3] Nanocapsules with a porous shell can release their content steadily over many months. If the capsule wall is rigid, it only releases the fragrance if it is ruptured (the release-on-demand principle). As only a few of the capsules burst at any one time, the effect can also last for months in this case. If not encapsulated, fragrance shelf life is short, with the scent altering or disappearing within a few weeks.

Ordinary cosmetic emulsions have droplet sizes between 100 and 100,000 nm. Nanoemulsions contain the same type of ingredients but their droplet dimensions may be as low as 10 nm. Given this small droplet size, nanoemulsions are transparent and have particular rheological properties, and because of their rheology and transparent properties, nanoemulsions are used in a number of cosmetics. When applied to skin or hair, nanoemulsions break down into their constituent ingredients.[4]

Some cosmetic products use mineral-based materials and their performance depends on particle size. In sunscreens, mineral nanoparticles (e.g. titanium dioxide and zinc oxide with particle sizes in the order of 20 nm) are efficient UV-filters. They transmit, reflect and scatter the visible part of the sun's rays while strongly absorbing in the UV region. These UV-filters consist of micron-sized aggregates, which are composed of nanosized primary particles.[5] The surface of these nanoparticles may be treated with an inert coating to improve their dispersion in the formulation and to prevent photocatalytic activity. The advantage of mineral UV-filters is that they provide broad UV-protection and usually do not cause adverse health effects.

> **The Market for Nano in Cosmetics**.
> According to Global Industry Analysts (GIA), the global cosmetics market is predicted to reach $41.4 billion by 2014.[6] Nano offers many advantages for new formulations, and few successful companies will have a range of cosmetics that do not have some kind of nano aspect to their formulation. So penetration is likely to be high, at the very least 20 %, and growing. So by 2015, nano enhanced cosmetics could be worth at least 8–10 billion USD.

21.1.4 NanoTextiles for Fashion, Work, Sport and Industry

Nano is having a major impact in textiles, an industry that is remarkable for being an early adopter of new ideas and technologies. Textiles are not only for the fashion

[2] http://www.salvona.com/19-nanosal-

[3] http://www.cosmeticsdesign.com/news/ng.asp?n=80466-naturalnano-nanotechnology-license-agreement-nanotubes

[4] http://www.pharmainfo.net/pharma-student-magazine/nanoemulsions

[5] Nanotechnology and Sunscreens, http://www.foe.org/nano_sunscreens_guide/Nano_Sunscreens.pdf

[6] http://www.cosmeticsdesign.com/Market-Trends/Global-market-for-color-cosmetics-set-to-reach-41.4billion.

conscious—they have important applications in the aerospace, automotive, construction, healthcare and sportswear industries. Already on the market are socks and leisurewear with embedded silver nanoparticles that combat odour through killing bacteria—and this capability has been extended successfully to wound dressings. Several brands of clothing, including designer labels, have incorporated self-cleaning and stain repellent nanotechnologies, very convenient for school clothes and military wear,—and, of course, the less a garment needs to be washed, the more energy and money is saved.

More glamorous applications include embedding gold and other precious metal nanoparticles into natural fabrics such as wool. The gold nanoparticles impart soft colours from pale soft greens, to browns and beiges, depending on the particle size and shape. These colours are stable, and may even provide some antibacterial properties to the fabrics, as an added bonus.

Currently, considerable research is focused on developing electrospinning techniques which produce long fibres of polymer, only nanometres in width, originally developed in an effort to emulate spider silk. The spun, polymer-based nanofibres can be 'loaded' with different additives which could be nanoparticles, enzymes, drugs or catalysts. Some combinations can be antibacterial and sprayed on to wounds as a kind of healing 'web', others can be conductive or even form filters or membranes. There are many exciting applications of the ability to spray-on a textile.

Scientists are also working on nanoelectronic devices that can be embedded into textiles to provide special support systems for individuals in dangerous professions or sports. Some garments can now provide monitoring for several life signs, including temperature, chemical sensing and they can be used for power generation and storage to enable communication with the outside world. Garments with this kind of technology can be vital for the safety of say firefighters working in dangerous situations in isolation from their colleagues, or even for skiers or their rescuers to give early warning signs of hypothermia.

Research is also ongoing into man-made nanofibres where clay minerals, carbon nanotubes or nanoparticulate metal oxides are used to impart new properties. These properties provide halogen-free, flame retardancy, increased strength and shock-absorbency, heat and UV radiation stability, and even brighter colouration. Other work is focused on the very exciting area of inkjet printing onto textiles. This is opening up many possibilities, not just for the customised or localised printing of textiles to an individual design, but inkjet techniques can be used to create flexible electronic materials, sensing materials, and even the materials of the future with printed-on display capabilities.

> **The Market for Nano in Textiles**
> Textiles include materials with an almost unimaginable spectrum of applications, and nanotechnology is further widening that spectrum. Given the many varieties and uses of textiles, and considering figures from various bodies involved in the industry, it is not unreasonable to suggest that in 2010, the global textile market was worth around 100 billion USD. It is likely that a high percentage of textile manufacturers are not taking advantage of nano, as much clothing is being produced in time-honoured ways. If the penetration of nano is only 1 % for high end applications, this still represents a 1 billion USD market, possibly 1.5 billion USD by 2015, as nano becomes more accepted and desirable, offering product differentiation opportunities.

21.1.5 Nano for Security: Anti-Counterfeit Protection and Discouraging Thieves

In a study undertaken in 2010, and reported in The Enquirer,[7] US Congress decided it was almost impossible to quantify the cost of counterfeit goods to industry, and surprisingly, that the effect was not necessarily all bad, and in the long run counterfeit goods can benefit consumers, mostly because counterfeit products are cheaper than the alternatives.

Although Congress in the US may dismiss counterfeiting as not all 'bad', this is not how commercial companies, who have built up their brands embodying quality and value over possibly many decades, view it. Most companies with known brands will fight tooth and nail to protect their brand rights and their investment in product development and promotion. Counterfeiting can also be dangerous, as in an extreme situation where aeroplane parts are replaced by counterfeit ones that are not fit for purpose.

The latest developments in nano for brand security and packaging offer what is called "smart protection", where often invisible, nano-enabled taggants provide not only an authentication of products, but can also assist in identifying the provenance of stolen goods. This security technique depends on the fact that nanoparticles are invisible. Consequently, they can impart new, unseen properties.

Quantum dots, which are fluorescent nanoparticles, have a whole range of possible applications in this industry. They are invisible until 'lit up' by ultraviolet light, and can even be made to exhibit a range of colours, depending on their composition and size. Such nanoparticles are ideal for crime prevention, where goods can be invisibly and non-destructively 'tagged'. Stolen goods can be traced by this invisible 'bar code'

---

[7] The Inquirer (http://s.tt/14vqR).

Table 21.2 Nanotech in brand and product security

Nanoparticles for security printing (e.g. invisible bar codes)
Unique recognition systems
Paper-like electronic displays which can be used to display information on a given product (history of manufacture and transport)
Magnetic nanoparticle tagging
Nanoparticle chemical markers
Ambient atmosphere nanoparticulate indicators (anti-tamper)

and also illict drugs by the fact they possess no legal nano identification. In some countries, cheap agricultural fuel can be 'laced' with harmless nanoparticles, making it easy for police to identify a stolen consignment, merely by using ultraviolet light.

Some taggants contain unique magnetic 'fingerprints' and can be used on a wide range of articles, including pharmaceutical packaging, luxury goods (such as watches and handbags), and automotive and aviation spare parts. Other, invisible, nano scale identification tags can be added to products and, using special equipment, field testers can immediately spot the real from the fake. A number of nano-based products have already been introduced including intelligent inks and nanoparticle coatings for product security (Table 21.2).

To illustrate the importance of the problem. Nanotechnology has been a key priority in the Sixth EU framework programme for RTD (FP6, 2002–2006) and this remains the case in the Seventh Framework programme (FP7, 2007–2013), with a budget of €3475 Million for the NMP programme. Key areas include security and the prevention of identity theft, counterfeiting, with especial resonance for the financial services industry, in relation to credit cards, personal identification techniques, authentication, forensics, quantum cryptography and the market for counterfeit and grey goods.

The Market for Nano in Brand and Product Security Applications.
According to a 2009 Observatory Nano report,[8] the size of the global market for brand and product security applications is estimated at between 40 and 60 billion USD. Nanotech offers an innovative approach in this technologically advanced market, so even if penetration is only 10–15 % now, it is likely to rise to at least 20 % by 2015, representing a possible market size of 8–12 billion USD.

---

[8] http://www.observatorynano.eu/project/filesystem/files/Economics_Security_final.pdf
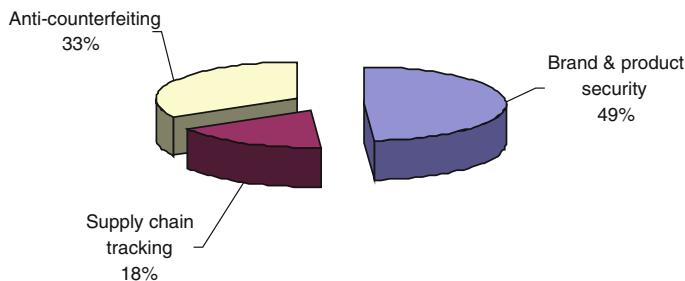
Fig. 21.4 Global brand and product security: nanotechnology implementation

21.1.6 Food for the Family: Nanotech in Food Processing, Improved Shelf Life, Flavour Enhancement, Vitamin Encapsulation and Fat Reduction

Working at the nanoscale is not new to food companies. The properties of all foods and beverages relate to their constituent components that are nanoscale in size. The manipulation of naturally occurring nanoparticles involved in the processing of dairy products, for example, has been undertaken for some time under the name of 'colloid science'. More recently, an improved understanding of techniques such as targeted delivery, has enabled food companies to deliver scents, flavours, vitamins and minerals that offer health benefits or impart new physical, visual and sensory effects to foods. This has not only helped the exponential growth in the market for nutraceuticals and other functional foods but has enabled a wide range of new food products with new tastes, flavours and textures to be developed. Other applications of nanotechnology to food manufacturing include antibacterial work surfaces, filters that can extract toxins, and packaging that provides a better barrier against contamination, or can signal when perishable contents are spoiling, by changing colour.

Whether a product is a food, a drink, a pharmaceutical drug or a cosmetic, whether it is ingested or applied, so long as it enters the bloodstream, it will produce an effect on the human organism. The line between these different groups is hard to draw. It is interesting to note that while hospitals are focused on the treatment of patients using prescription drugs, very few consider that treatment may be possible by monitoring/selecting appropriate foodstuffs—although everything we ingest is a chemical to some degree or other, as it is made up of molecules that are absorbed in the body. In fact, treatment through a professional nutritional analysis is an area that has been almost entirely ignored until now, in preference to treatment by drugs which, because of their concentration of a single chemical, are often highly toxic.

Food companies themselves are increasingly aware of the medical component of their products, and the implementation of scientific knowledge in commercial foodstuff production could have much wider implications for the health of the population

than is presently acknowledged. Table 21.3 below lists some of the applications of nanotechnology in foods, from production to ingestion.

It is only in relatively recent times that novel technologies have come under investigation as offering new functionalities and benefits as well as efficient delivery mechanisms for the food and beverage industries. For example, food researchers are gaining a greater understanding of the mechanisms of targeted delivery, with a view to optimizing the delivery of vitamins and minerals in food to benefit health; technologies related to the creation of novel physical, visual and sensory effects for competitive advantage.

The Market for Nano in the Processed Food, Beverage and Related Packaging Industries

There are many figures for the world market for processed foods. Whatever figure is used, the numbers are extraordinary. In 2005, this market was estimated at 3.2 trillion USD.[9] So for the world giants in food processing, the competition is cut throat, with everything to play for in terms of market share through increased food stability, shelf life, safety, and a better consumer experience. It is estimated that the market for alcoholic, soft and hot drinks is about 50 % of that for processed foods. So the total market, given an annual growth rate of 7.5 % can be estimated at around 7.5 trillion USD in 2010 and up to 10 trillion dollars in 2015.

If nano has a penetration of even 1 % in this market, and it may be more as the market is dominated by around 100 companies who can afford to (and need to!) adopt new technologies, then nano in the food and beverage industries could be worth at least 100 billion USD by 2015. Not an insignificant number by any means. Regarding smart, active packaging for the food and beverage industries, this has been estimated at 11.7 billion USD in 2011.[10] It is likely that nano will have an important impact in most of this market sector, because of its dependence on technology. Even estimating only a 20 % penetration, this represents over 2 billion USD now, and probably twice as much by 2015 as nano applications become increasingly important to the sector. So again, more large numbers for nano here.

Potential applications of nanotechnology includes nano-encapsulation of flavours or nutrients to suit consumer preference or health requirements; nanofilters that can remove toxins; food constituents that can be made to alter their colour; flavour modifications that can be created by using differently-'twisted' molecules (for example, the direction of chirality of a molecule may determine whether the flavour imparted is 'lemon' or 'orange'); packaging that can keep perishable contents fresher for longer, or detect when contents are spoiling and changing colour to warn consumers. In

---

[9] http://www.ers.usda.gov/AmberWaves/February05/Features/ProcessedFood.htm
[10] http://www.visiongain.com/Report/574/The-Active-Intelligent-and-Smart-Food-Drink-Packaging-Market-2011-2021

essence, the understanding of food materials and food processing at the nanoscale is increasingly recognized as vital in the creation of new and better food products, and also to minimizing waste from increasing shelf life and visual indicators of freshness. No more sell-by dates! (Table 21.3).

Table 21.3 Applications of nanoproducts in food related areas

Market sub sector	Nano applications
Food production	- Anti-bacterial food preparation surface coatings
	- Colloid stability improvements
Conservation	- Preservatives, antioxidants etc
	- Optimal environment emulation
	- Lifespan extension
	- Fridge food freshness maintenance
Packaging	- Anti-counterfeit (tracking systems, smart packaging)
	- Contamination prevention, freshness maintenance
	- Novel, brand-oriented packaging
	- Freshness/shelf life indicators
	- Speed check out enhancements
	- Improved flexibility, durability, temperature/ moisture stability, barrier, anti-microbial properties
Novel and 'Fashion' Foods	- Colour, scent, flavour, taste and texture enhancement
Health foods	- Supplement encapsulation (vitamins, minerals etc)
	- Enhanced bioavailability
	- Reduction in salts, fats and sugars
	- 'Delivery systems' (scents, flavours etc)
	- Sprays
Agriculture	- Soil remediation
	- Water purification
	- Pesticides
	- Nanosensors

Source International Nano Markets

Table 21.4 Areas where applications of nanotechnology for health foods are currently close to market or already available

Enhanced delivery of nutraceuticals and bioactive compounds in functional foods

Enhanced flavours, texture and delivery of bioactive functional ingredients

Enhanced solubility—the smaller the component particle, the more soluble

Controlled release for in-situ flavour and colour modification of products

Improved bioavailability of vitamins and minerals for medical and sporting applications

Protection of the stability of micronutrients and bioactive compounds during processing, storage and distribution

Encapsulation of fats and oils for reduced calorie products

Nano particulate salt for more flavour with less salt content

21.1.7 Getting Around: Nano in Cars, Aeroplanes, Ships, Trains etc

The automotive industry has appreciated for some time that nanotechnology can offer many benefits in this highly competitive and litigious sector. Research is taking place into applications of nanotechnology to improved safety systems from tyre blow out and brake failure warning systems to collision avoidance (Table 21.4).

Other nano benefits include improved lubricants, lighter and stronger materials for engine blocks, nanoporous filters, self-cleaning windshields, self-healing and scratch-resistant coatings, environmentally friendly corrosion protection and colour-changing paints. Major advances are also being made in the use of polymer nanocomposites for body panels as these can be made lightweight yet rugged, and in new metal nanocomposites to improve engine efficiency. Specially designed nanoparticles are presently used as fuel additives to lower consumption in commercial vehicles and reduce toxic emissions (Table 21.5).

Car manufacturers are keen to be more environmentally friendly in their manufacturing processes as well as in the final product. Investigations are underway in how nanotechnology may lead to a reduction in toxic wastes and by-products by substituting new nanomaterials for hazardous reactants and solvents or, better still, by using nanotechniques to eliminate their use altogether (Table 21.6).

The Market for Nano in the Automotive Sector.
According to Wikipedia (which cites the "2008 Global Market Data Book", Automotive News, p. 5), 71.9 million new cars were sold worldwide in 2007. This figure is probably quite similar in scale to today's sales, given the downturn, balanced by the increase in population. The sale of new trucks is about one third that of cars, making total industry sales of about 96 million new vehicles a year. The automotive sector has always been an early adopter of new technology for competitive and liability reasons. If the size of the automotive industry is estimated at a reasonable average value of 4500 USD per car, and the truck industry at twice that, namely 9000 USD per vehicle, then the value of the total industry (trucks and cars) is around 650 billion USD/yr. If the penetration of nano into automotive components is about 15 % by value, possibly more, then the worth of nano to that industry can be estimated today at about 100 billion USD, and by 2015, given a market penetration of at least 20 %, would be about 130–150 billion USD—the latter given some increased growth also in the sector value.

Table 21.5 Nanotechnologies in the food and beverage processing and related packaging industries

Organic nanoadditives

Inorganic nanoadditives

Addition of nanoparticles offering specific additional functionalities or novelty

Nanosensors for food quality control and smart packaging

Nanoparticles for toxin scrubbing and to slow down ripening

Nanocoatings and nanofilms for protecting kitchenware and foodstuffs against pathogenic bacteria

Packaging for ambient temperature maintenance

Nanosprays of bioluminescent indicators in antibacterial defence systems

Incorporation of nanosized ingredients and additives

Processing of food at the nanoscale

Nanoencapsulation of ingredients, additives and supplements (based on micelles and liposomes)

Manufactured mineral nanoparticles as additives and supplements

Incorporation of nano sunscreens and other modifications in food packaging

Table 21.6 Areas of application of nanotechnology in the automotive sector	
	Fuel cells
	Power systems
	Tyres
	Heat transfer
	Lighting and displays
	Sensors
	Coatings
	Smart materials
	Interiors
	Structural and functional materials

21.1.8 When We Get Ill: Nanomedicine, High Speed Diagnostics, Drug Delivery and Medical Devices

In the past, medical treatments have been, rather like medieval architecture, the result of adopting those techniques that worked and discarding those that didn't. Today, our improving knowledge of how the body functions at the molecular, or 'nano', level, is leading to many new and better medical techniques. For example, we know that the earlier a disease can be detected, the easier it is to remedy, but until now, early detection has been very difficult (Table 21.7).

Today we can introduce into the body specially designed nanoparticles, which are composed of tiny fluorescent quantum dots that are chemically bound to antibodies, which in turn 'bind' to diseased cells. When this happens, the quantum dots fluoresce brightly. This fluorescence can then be picked up by new, specially developed, advanced imaging systems, enabling the accurate pinpointing of a disease even at a very early stage indeed.

Table 21.7 Applications of nanotechnology in the life sciences and healthcare sector

Market area	Nanomaterials application
Diagnostics	- Diagnostic imaging, with the use of targeted imaging agents that signal the site of disease
	- In-vivo monitoring of the body's physiology using a variety of nanotechnology-based probes and sensors
	- In-vitro diagnosis that is faster, more precise and informative, building hundreds of diagnostic tests into a single device
Drug delivery	- Nanoparticles for the delivery of minute doses of highly-active drugs, or of novel theranostic products to specific sites that can be tracked and activated by conventional radiological devices
	- Nanodevices that could be used to deliver substances or drugs to specific target areas in the body, e.g. oxygen to poorly vascularized tissues
	- New "smart" nanoengineered materials that can impart new levels of performance to "conventional" medical devices and drugs
Implants and medical supplies/devices	- Biocompatible nanostructured implant surfaces and coated stents
	- Easy-clean and self-cleaning nanocoatings and antibacterial surfaces
	- Better performing and cheaper to produce artificial retinas and cochlear implants that mimic more closely nature's light/sound receptor and transmission systems
	- Novel implantable devices allowing treatment of hitherto incurable degenerative or congenital diseases
	- Nanodevices for minimally invasive surgery
	- Design of surgical and diagnostic tools with cheaper manufacturing facilitated by nanotechnology and offering improved performance, e.g. ultra-sharp nano-diamond coated scalpels for microsurgery
Tissue engineering	- Biomimetic scaffolds as pattern for tissues and organ regeneration

Source International Nano Markets

Nanotechnology is also leading to faster diagnosis. Diagnosis can be a lengthy and stressful business, often with a test sample having to be sent away for analysis, with the results taking several days at least. Ultra miniaturisation is enabling many tests to be speedily undertaken on 'lab-on-a-chip' devices, using tiny samples which can be processed and analysed so rapidly that the results can be read out almost instantaneously.

People often complain that the cure for a disease can feel almost as bad as the disease itself, as prescription drugs may have unpleasant, sometimes even fatal, side effects. This is because the body needs to be flooded with very high doses of a drug in order to ensure that a sufficient volume reaches the site of the disease. Accurate targeting of the drug can now be achieved, using specially designed drug-carrying nanoparticles. This also means that much smaller quantities of a drug are necessary, so it is less toxic to the body. The drug is then activated only at the disease site (such as a tumour) by light or other means, and the progress of the cure can also be monitored by imaging

the location and movement of the specially designed nanoparticles, which usually have a fluorescent component.

The Market for Nano in Pharmaceutical Drugs

Most commentators seem to have a figure of around a trillion dollars for the value of pharma sales worldwide. IMS Health estimate 1.1 trillion USD sales by 2014; PRLog estimate a trillion USD sales by 2012. As the pipeline for new drugs has faltered over the last decade, balanced by new opportunities in targeted drug delivery, which offers reduces toxicity while improving outcomes, the importance of nano to the pharma industry has increased dramatically. Designer nanoparticles carry drugs and release them to the therapeutic target, and the benefit to drug companies from using innovative techniques with existing drugs is that no further expensive testing is required. Furthermore, protein-based drugs, which have hitherto been difficult to deliver, are now experiencing a huge resurgence of therapeutic interest as at the nanoscale their solubility properties change dramatically.

If nano only has 5 % penetration of the pharma market, this still means a market size of around 50 billion USD—and definitely growing.

The Market for Nano in Medical Devices

According to Espigen, the US market for medical devices (the largest in the world) is estimated at 105.8 billion USD in 2011, though CIMS, the Center for Integrated Manufacturing Studies placed it at 336 billion USD in 2008. The CIMS estimate includes: implantable devices (such as pacemakers, drug pumps, stents, and joint replacements); diagnostic testing devices, including clinical blood, urine, and tissue testing; and home healthcare products and electronic monitoring devices. Speculatively, if the rest of the world together equals the size of the US market, then the total size could be around 600–700 billion USD. Devices increasingly need to be smart, in some instances, vanishingly small and body friendly, so nano has an increasing role in new developments. Estimating a penetration today of nano in devices of about 10 %, probably rising to 15 % in 2015—within a fast growing market, driven by demographic trends of an increasing ageing and wealthy population, to about 1 trillion USD in 2015, gives a modest estimate of market size of about 100 billion USD for nano-enabled devices.

21.1.9 Nano for the Environment

Nanotechnology offers some really exciting breakthrough opportunities in environmentally friendly technologies, from extracting renewable energy from the sun to

the prevention of pollution. Geoffrey Sacks, the American economist, in his 2007 BBC Reith lectures entitled **'Bursting at the Seams'**, commented: *"The fate of the planet is not a spectator sport. We live in an interconnected world, where all parts of the world are affected by what happens in all other parts"*.

There is no doubt that the pressures we are putting on the planet are leading to potentially catastrophic consequences. In the developed world, we have grown accustomed to using our car to go to the local shops, take weekend cruises and even day trips to far-flung places that might have taken three or more months to reach before air travel became commonplace. We like our vegetables and fruit out of season, and increasingly expect to eat meat at least once, if not twice a day. We also want to hoard our assets in the form of gold and diamonds. We haven't thought about the effects of these activities on the planet, which in the past could absorb our excesses, but with the ongoing destruction of the rainforest (which is responsible for 25 % of carbon emissions!) and the population of the world reaching an unsustainable 7 billion, the earth is showing clear signs of being unable to bounce back from the demands we are placing on it.

So what can we do to limit the damage and ensure a future for our children? Firstly, the bad news. The fossil fuel that oils our everyday lives is responsible for 44 % of the carbon dioxide we emit annually—and rising. The good news is that the energy from sunlight is sufficient to meet our needs *ten thousand times* over. Today, more efficient and cheaper solar energy collectors are in the process of being developed using nanotechnology; these could be deployed as small units in our homes. They work particularly well in diffuse light, so would suit less sunny climates. This would have the benefit of not sterilizing precious land (a diminishing resource for food), and quickly improve the quality of many people's lives, especially for people living in in poor quality housing or in the less developed world, where energy is hard or expensive to access.

Not only do we need new ways of generating energy, we need better ways of storing it, and nanotechnology is leading to improved, environmentally-friendly batteries and supercapacitors. We also need to reduce damage to the environment. Particularly toxic are those chemicals we use as solvents, and nanotechnology is leading to their eradication through the development new nanocoatings and nano structured surfaces that can effectively repel dirt and other contaminants. Most coatings used to prevent corrosion are toxic and seriously affect the environment. Many anti-corrosion coatings involve chromium and cadmium, deadly substances, the use of which is being limited by the EU. Of course, vehicle and component producers are keen to find alternatives, as recycling of toxic compounds is costly and unpleasant, so new smart nanocoatings are in the process of being developed that are non-toxic and highly effective.

Serious contamination of the environment with heavy metals and other pollutants are thrown into the atmosphere from the fumes and smoke being emitted from industrial processes. It is encouraging to note that most of these of these particles and gases (including carbon dioxide) can be 'scrubbed' out—and even reclaimed and reused,

using specially functionalised nanomaterials, placed in the waste gas stream. Finally, given the old adage, if you can't measure it, you can't control it, fast, accurate, in-situ and online pollution monitoring is essential. New, cheap nanosensors are being developed from techniques used in medicine, that will enable us to do this quickly, effectively and cost effectively.

There is a raft of potential applications enabled by nanotechnology, short and long-term, offering environmental benefits. Some of these are listed below, but it is a vast area.

- water purification and filtration;

- new materials offering 'more for less';

- more efficient use of material and energy and therefore reduced environmental impact;

- sensors for improved monitoring and detection capabilities;

- treatment and remediation techniques for cost-effective and specific site cleanup;

- green manufacturing to eliminate the generation of waste products;

- green energy technology for the creation of commercially viable clean energy sources;

- paints using less or water based solvent;

- more efficient fuel cells and batteries;

- fuel additives to improve fuel economy.

Nanotechnology offers opportunities in reducing and saving energy; some applications under investigation include:

- Thermoelectric nanomaterials that enable the generation of electricity from waste heat in consumer appliances, automobiles, and industrial processes;
- Higher efficiency photovoltaic cells using quantum dots connected by carbon nanotubes;
- Gases flowing over carbon nanotubes that have been shown to convert to an electrical current, a discovery with implications for novel distributed wind power;
- Multi-walled nanotubes that increase the performance of hydrogen fuel cells.

Advanced nanotechnology projects relating to energy are in storage and conversion, and in manufacturing improvements by reducing materials and process rates in energy saving e.g. by better thermal insulation, and in enhancing renewable energies sources. Specifically, this includes products such as batteries, manufacturing catalysts, fuel cells, solar cells, and strong lightweight materials.

Global Market for Nano in Environmental Applications

This is a huge market with several important sectors. Perhaps the largest is renewable energy generation, energy storage and energy saving.

Wind. The market for wind turbines is estimated to be 100 billion USD by 2013.[11] Nano composites have an increasingly important role to play in the development of stronger, tougher and lighterweight turbine blades. If the market penetration for nano composites is only 10 %, that is still a market of 10 billion USD.

Solar Energy. Recent figures[12] from the USA's leading solar trade group, the Solar Energy Industries Association (SEIA), show that in 2010, solar energy is America's fastest growing energy sector stock, growing in total market value by 67 % from 3.6 billion USD in 2009 to 6.0 billion USD in 2010, compared to the overall US GDP growth rate of only 3 percent. Nano has an increasing impact on solar energy collection. If penetration of nano in solar collectors is only 5 %, then that represents a market of about 0.3 billion USD in the United States alone, and around 0.75–1 billion USD worldwide, in a fast growing market.

Energy Storage. Energy storage technologies where nano is impacting include[13]: batteries, fuel cells (fuel cell development has increased in recent years due to an attempt to increase conversion efficiency of chemical energy stored in hydrocarbon or hydrogen fuels into electricity). Research is being conducted on harnessing the quantum effects of nanoscale capacitors to create digital quantum batteries. Although this technology is still in the experimental stage, it theoretically has the potential to provide dramatic increases in energy storage capacity.[14] According to a BCC Research 2009 report,[15] nano for environmental applications, including protection, maintenance, enhancement and remediation is predicted to be worth 21.8 billion USD by 2014.

21.1.10 Nanomaterials

Nanomaterials, and their associated manufacturing and processing technologies, are the key enablers of the nanotechnology industry, and encompass a wide range of materials. Nanoparticles serve as the "building blocks" for nanomaterials and devices. They include nanocrystalline materials such as ceramic, metal and metal

---

[11] http://windharvest.com/windmarket

[12] http://www.seia.org/cs/news_detail?pressrelease.id=1292

[13] http://en.wikipedia.org/wiki/Energy_storage

[14] http://www.technologyreview.com/computing/24265/?a=f

[15] http://www.bccresearch.com/report/nanotechnology-environmental-applications-nan039b.html

oxide nanoparticles; fullerenes, nanotubes and related structures. By virtue of their structure, nanomaterials exhibit different physical, chemical, electrical, and magnetic properties from conventional materials, which can be exploited for a variety of structural and non-structural applications. The following are the core nanomaterials offering the main market opportunities:

- Nanoparticles
- Nanocoatings
- Nanocomposites
- Carbon nanotubes
- Graphene
- Nanoporous materials
- Quantum dots
- Nanofibres
- Fullerenes
- Nanowires
- Nanocapsules.

There are many examples of nanotechnology applications in new materials. For example, polymer coatings are notoriously easily damaged, and affected by heat. Adding only 2 % of nanoparticulate clay minerals to a polymer coating makes a dramatic difference, resulting in coatings that are tough, durable and scratch resistant. This has implications for situations where a material fits a particular application in terms of its weight and strength, but needs protection from an external, potentially corrosive environment—which a reinforced polymer nanocoating can provide. Other nanocoatings can prevent the adherence of grafitti, enabling them to be easily removed by hosing with water once the coating has been applied. This has the important knock-on effect of improving urban environments, making them more attractive to bona fide citizens and less encouraging to criminals. These kinds of coatings, invented in Mexico, have been shown to work well in parts of Mexico City, transforming seedy crime-ridden neighbourhoods into increasingly respectable suburbs.

Novel nanocomposites have applications in many industries, including the aerospace and automotive industries. Carbon nanotubes are increasingly applied where toughness, electrical conductivity and flexibility are required, combined with light weight. Nanoporous materials have applications in purification techniques including blood and water and pure liquids for the electronics industries; quantum dots have applications in security, electrics and medical diagnostics; nanofibres have many applications in reinforcing; fullerenes have medical and other applications; nanowires for the electronics industries, and nanocapsules for the delivery of drugs, scents and flavours and the stabilisations of foods and perishable materials.

Scents and flavours are a surprisingly interesting and lucrative field for the application of nano encapsulation and delivery technologies. These technologies were first developed for the delivery of pharmaceutical drugs, and have now found new applications in foods and household products. Encapsulation is an ideal way to improve

the attributes and performance of a less-than-stable substance that might be affected by light or air, or have a tendency to sediment. Encapsulation gives active ingredients a longer shelf life, stability and protection from harsh processing environments so they can be delivered in a perfect state at 'the moment of consumption'! For the food industry, it is a way of delivering enhanced taste, or ensuring that daily doses of vitamins and minerals are met—this is discussed in more detail below. In household products, nano encapsulation techniques can aid in the deposition of a cleaner or polish onto a surface such as a floor or counter; they can provide long lasting scents in household fragrances, and the slow release of enzymic and other agents in washing machines and dishwashers, helping reduce energy and water use.

Carbon nanotubes are a recently discovered unique material possessing amazing electronic, thermal, and structural properties They are highly conductive both to electricity and heat, with an electrical conductivity as high as copper, and a thermal conductivity as great as diamond. They offer amazing possibilities for creating future nanoelectronic devices, circuits and computers. Carbon nanotubes also have extraordinary mechanical properties—they are 100 times stronger than steel, while only one sixth of the weight. These mechanical properties offer huge possibilities, for example, in the production of new stronger and lighter materials for military, aerospace and medical applications. Other applications include lubricants, coatings, catalysts and electro-optical devices.

The cost, purification and separation of nanotube types, constraints in processing and scaling up and assembly methods are hurdles that are now being overcome. Already there are several consumer products containing nanotubes on the market, for example, in sports equipment including tennis racquets, golf clubs and skis, where nanotubes improve the strength of the equipment but at less weight and also the ability to absorb shocks. Carbon nanotubes can also be mixed with many different materials such as polymers and included in textiles for example in the production of lighterweight bullet-proof vests, and in several military applications.

Graphene is the new nanomaterial that is generating huge excitement. It is an allotrope of carbon that can be produced in thin sheets of only one atom in thickness. Applications include commercial, aerospace and military. As graphene is currently the 'material of the moment' some space is given to potential applications are currently being explored:

Composites with higher strength to weight ratios. By adding graphene to thermosetting and other polymer composites, a stronger/stiffer composite may be possible compared with a similar weight of carbon nanotubes, with applications in wind turbine blades or aircraft components, for example.

Transistors that operate at higher frequency. Electrons in graphene move at higher speed compared to electrons in silicon, offering the potential to build high frequency transistors. Researchers are also working on lithography techniques for fabricating integrated circuits based on graphene.

Lower cost of display screens in mobile devices. Graphene is capable of replacing indium-based electrodes in organic light emitting diodes (OLEDs). These diodes are used in electronic device display screens which require low power consumption. The use of graphene instead of indium not only reduces the cost but eliminates the use of metals in OLEDs, which may also make devices easier to recycle.

Storing hydrogen for fuel cell powered cars. It has been found that graphene layers in a fuel tank increase the binding energy of hydrogen to the graphene surface, resulting in a higher amount of hydrogen storage and therefore a lighter weight fuel tank. This could help in the development of hydrogen fueled cars.

Sensors to diagnose diseases. Graphene, strands of DNA, and fluorescent molecules can be combined to diagnose diseases. Graphene has a large surface area and molecules sensitive to particular diseases can attach to the carbon atoms in the graphene. When an identical single strand DNA combines with the strand on the graphene a double strand DNA is formed that floats off from the graphene, increasing the fluorescence level, thus detecting the same DNA for a particular disease in a sample.

Ultracapacitors with better performance than batteries. These ultracapacitiors store electrons on graphene sheets, as the large surface of graphene provides an increase in the electrical power that can be stored in the capacitor. It is possible that these ultracapacitors will have as much electrical storage capacity as lithium ion batteries, but can be recharged in minutes instead of hours

Global Market for Nanomaterials

Nanoparticles	Nanoparticles are available as dry powders or liquid dispersions. Important nanoparticulate materials are simple metal oxides, such as silica, alumina, titania, zinc oxide, iron oxide, ceria, and zirconia. Silica and iron oxide nanoparticles have been in the market for about a half-century or more. Nanocrystalline titania, zinc oxide, ceria, and other oxides have entered the market more recently. The nanoparticles market is an important one, worth around 1.6 billion USD in 2007, and predicted to rise to over 20.5 billion USD by 2015.[a] In the last few years, clay nanoparticles have made their way into composites in cars and packaging materials, and this makes up a significant percentage of the current market; however, new applications are coming onto the market. For example, casings for electronic devices such as computers containing nanoparticles also offer improved shielding against electromagnetic intrusion.[b]

[a]Nanomaterials. Applications and Markets to 2015. International Nano Markets

[b]Nanomaterials, Applications and Markets to 2015. International Nano Markets

Nanocoatings	Nanocoatings such as thin films and engineered surfaces have been developed and applied across a wide range of industries for decades. The ability of controlling surface coatings at the nanoscale is of paramount importance for a large-scale industrial development of nanotechnology. Highly sophisticated surface-related properties, such as optical, magnetic, electronic and catalytic can be obtained via nanostructured coatings. For example, in the silicon-integrated circuit industry many devices rely on thin films for their operation, and control of film thicknesses approaching the atomic level is routine. Other applications on the market are "self-cleaning" coatings and optical-functional surfaces for facades, vehicles, solar cells etc. (e.g. for anti-reflective surfaces, sunshade glazing, anti-reflective coatings for instrument panels). At present, many physical and chemical methods are available for the nanofabrication of layers and coatings with nanometric control of the structural and functional features; however, the scale-up of these methods remains a major challenge. "One way" coating systems based on nanomaterials make up the bulk of this market, for example in anti-bacterial; protective and conductive coatings. However under development are "two way" systems such as shape-memory materials, hydrophobic/hydrophilic switching and thermochromic pigmented coatings that will come onto the market in the next 4–5 years. The fastest growing markets to 2015 is expected to be in interior and exterior household protection, textiles and medical markets, driven by the increased demands for protective and repellent coatings. Conductive carbon nanotube coatings are also finding market traction in the electronics and automotive sectors and this will continue to be a strong growth area. Continuing to hold the biggest market share will be military applications such as anti-corrosion coatings which also find application in the gas and oil markets. According to new study[c], global consumption of paints and coatings in 2012 is projected to be worth 120 billion USD. If a modest 10% of this market can be ascribed to nano-influenced coatings, then that market alone in 2012 is worth 6 billion USD. Given a year-on-year increase in market share of 15%, then the market for nanopaints and coatings in 2015 can be estimated at around 10 billion USD. It is likely that including other nano coatings could increase this market estimate by at least 50%.
Nanocomposites	The nanocomposites market was worth around 437 million USD in 2007, increasing to approximately 2 billion USD by 2010, and 7 billion USD by 2015. Applications in automotive, aerospace, packaging (food and drink), electronics and consumer goods is expected to account for the largest percentage of revenue by 2015

[c] Global Paint & Coatings, 2013–2018, Kusumgar, Nerlfi, & Growney

Nanoporous materials	Nanoporous materials possess unique surface, structural, and bulk properties that underline their important uses in various fields such as ion exchange, separation, catalysis, sensor, biological molecular isolation and purifications. Nanoporous materials are also of scientific and technological importance because of their vast ability to adsorb and interact with atoms, ions and molecules on their large interior surfaces and in the nanometer sized pore space. They offer new opportunities in areas of inclusion chemistry, guest-host synthesis and molecular manipulations and reaction in the nanoscale for making nanoparticles, nanowires and other quantum nanostructures.[d] According to a 2009 report[e] published by GIA, Global Market for Nanoporous/Microporous Materials Expect to Reach 3.75 billion USD by 2015
Quantum dots	The global market for quantum dots, which in 2010 was estimated to generate 67 million USD in revenues, is projected to grow over the next 5 years at a compound annual growth rate (CAGR) of 58.3 %, reaching almost 670 million USD by 2015, representing a tenfold increase.[f]
Nanofibres	Nanofibres have a variety of applications, in the electronics, energy, medical and aerospace sectors. Several types of nanofibres have been developed: polymeric, carbon, ceramic, glass, metallic and composite. Nanofibres are being developed for conductive and reinforcement applications in composites. The main revenue generating application in 2007 for nanofibres was in filtration and separation media. The nanofibres sector is expected to grow from 10.4 million USD in 2007 to 1193 million USD in 2015. The fastest growing sector is expected to be in ITC, increasing from 7 million USD in 2007 to 480 million USD by 2015. Inorganic nanowire light-emitting diodes (LEDs) have held promise as inexpensive and tiny light sources, but fabrication has been challenging. Nanofibre technology has been shown to enable on-chip light sources smaller and cheaper than LEDs, and easier to fabricate.[g] Thermoplastics play a vital role in protecting electronic devices from electromagnetic interference (EMI) and a variety of additives can be used to provide this protection, with carbon nanofibres the most promising application area. The energy sector will also benefit from carbon nanofibres growing from 7 million USD in 2007 to 480 million USD in 2015. Lithium-ion batteries incorporating nanofibres will come onto the market in the next 2–3 years. Carbon filaments are covered with platinum nanoparticles and can be made into nonwoven fabrics and used as electrode layers in a hydrogen fuel cell. The platinum acts as a catalyst, separating electrons from hydrogen atoms so they can be discharged as electricity. The carbon nanofibres transport the electrons out of the fuel cell with their large surface area of the nanofibres providing a base for the chemical reaction to occur. Also under development are thin film nanofibre sensors that can be made of polyaniline nanofibres having superior performance in both sensitivity and time response to a variety of gas vapours including, acids, bases, redox active vapors, alcohols and volatile organic chemicals.

[d] http://www.icpress.co.uk/etextbook/p181/p181_chap1.pdf
[e] http://www.strategyr.com/Nanoporous_Microporous_Materials_Market_Report.asp
[f] http://bccresearch.blogspot.com/2011/06/global-market-for-quantum-dots-to-grow.html
[g] **NANO OPTICS: Electrospun light-emitting nanofibers could replace LEDs,** http://www.laserfocusworld.com/display_article/311576/12/ARTCL/none/News/NANO-OPTICS:-Electrospun-light-emitting-nanofibers-could-replace-LED

Fullerenes	A fullerene is any molecule composed entirely of carbon, in the form of a hollow sphere, ellipsoid or tube. The fullerenes market is worth around 58.5 million USD in 2007. The energy sector is the most prominent with fullerenes used in fuel cells, solar cells and batteries. This market is expected to grow, with significant growth also expected in the ITC and automotive sectors, reaching approx. 1.8 billion USD by 2015
Nanowires	Nanowires are ultrafine wires or linear arrays of dots, formed by self-assembly, which can be manufactured from a wide range of materials. Semiconductor nanowires made of silicon, gallium nitride and indium phosphides have demonstrated remarkable optical, electronic and magnetic characteristics. For example, silica nanowires can bend light around very tight corners.
	The market for nanowires is worth approximately 34 million USD in 2007. The main market is in ITC, and this will continue to be the case with an estimated market worth 1.227 billion USD in 2015. Nanowires are finding application in electronic devices (including field-effect transistors, sensors, detectors. light-emitting diodes, and for metallic interconnects of quantum devices and nanodevices, allowing replacement of copper in computers and electronics); high-density data storage (as magnetic read heads or patterned storage media); and superconducting wires that can make electrical generators lighter and more efficient. Other applications are as bio and chemical sensors. Nanowires comprised of '1-dimensional' semiconducting materials have the greatest potential as biosensors because they can easily be incorporated into electronic circuits that allow continuous readout. Nanowires can be 'functionalized' to detect virtually any form of bi-molecular interaction that involves proteins (including antibodies, receptors and enzymes) or nucleic acids.[h]
Nanocapsules	The nanocapsules market is worth around 32 million USD in 2007. Applications in healthcare especially in the area of drug delivery, cosmetics and homecare are expected to account for the largest percentage of revenue by 2015, which is predicted to reach over 1.4 billion USD by then
Carbon nanotubes	The market for carbon nanotubes was approximately 168.5 million USD in 2008. The ITC market is likely to see the biggest penetration to 2015, with the performance enhancing properties allowing electronics manufacturers to meet demanding market needs. The incorporation of CNTs into the displays market will increase demand by 2010, with a revenue forecast in the ITC market of 1.1 billion USD approx. by 2015. While in the longer run, electronics will continue to dominate nanotube applications as broader use in semiconductors occurs, strong opportunities are also expected in CNT-based products.[i]
Graphene	According to a new technical market research report by BCC, the global market value for graphene-based products will be worth an estimated 67 million USD in 2015 and is expected to increase to 675.1 million USD in 2020, for a five-year compound annual growth rate (CAGR) of 58.7%.[j]

[h] Nanomaterials. Applications and Markets 2008-2015. International Nano Markets
[i] Global Market for Carbon Nanotubes, International Nano Markets
[j] http://www.electronics.ca/publications/products/Graphene%3A-Technologies%2C-Applications%2C-and-Markets.html

21.1.11 Finally, What are the Risks from Nanotechnology?

Some engineered nanoparticles, including carbon nanotubes, although offering tremendous opportunities also may pose risks which have to be addressed sensibly in order that the full benefits of new nanomaterials can be realized. We have all learned how to handle electricity, gas, steam and even cars, aeroplanes and mobile phones in a safe manner because we need their benefits. The same goes for engineered nanoparticles. Mostly they will be perfectly safe, embedded within other materials, such as polymers. There is some possibility that free nanoparticles of a specific length scales may pose health threats if inhaled, particularly at the manufacturing stage. Industry and government are very conscious of this, are funding research into identifying particles that may pose a hazard to health or the environment, and how these risks may be quantified, and minimised over the whole lifecycle of a given nanoparticle. There is no doubt that nanotechnology has great potential to bring benefits to society over a wide range of applications, but it is recognised that care has to be taken to ensure these advances come about in as safe a manner as possible.