Springer Series in Optical Sciences 227

Barry R. Masters

Superresolution Optical Microscopy

The Quest for Enhanced Resolution and Contrast

EXTRAS ONLINE

Springer

Springer Series in Optical Sciences

Volume 227

Springer Series in Optical Sciences is led by Editor-in-Chief William T. Rhodes, Florida Atlantic University, USA, and provides an expanding selection of research monographs in all major areas of optics:

- lasers and quantum optics
- ultrafast phenomena
- optical spectroscopy techniques
- optoelectronics
- information optics
- applied laser technology
- industrial applications and
- other topics of contemporary interest.

With this broad coverage of topics the series is useful to research scientists and engineers who need up-to-date reference books.

More information about this series at http://www.springer.com/series/624

Barry R. Masters

Superresolution Optical Microscopy

The Quest for Enhanced Resolution and Contrast

Springer

Barry R. Masters
Previously, Visiting Scientist
Department of Biological Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA

Previously, Visiting Scholar
Department of the History of Science
Harvard University
Cambridge, MA, USA

Additional material to this book can be downloaded from http://extras.springer.com.

ISSN 0342-4111 ISSN 1556-1534 (electronic)
Springer Series in Optical Sciences
ISBN 978-3-030-21690-0 ISBN 978-3-030-21691-7 (eBook)
https://doi.org/10.1007/978-3-030-21691-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Epigraph

On curiosity in science

> "I have no special talents. I am only passionately curious."
>
> —Albert Einstein

On risk taking in science

> "The fact that in his discussions he sometimes went astray, as for example in his hypothesis concerning radiation quanta, should not be held against him, since nothing really new can be achieved, even in the exact sciences, without a venture."
>
> —Max Planck, 1913, writing in support of the election of Einstein to the Prussian Academy of Sciences in Berlin

On creativity in science

> "On looking back to this event, I am impressed by the great limitations of the human mind. How quick are we to learn, that is, to imitate what others have done or thought before. And how slow to understand, that is, to see the deeper connections. Slowest of all, however, are we in inventing new connections or even in applying old ideas in a new field."
>
> —Frits Zernike, Nobel Lecture, December 11, 1953

On mathematics and science

> "It is especially the general insight, however, which gains very much by the discovery of the asymptotic development. It shows that physical intuition combined with experimental ability may go far towards elucidating the main characteristics of phenomena, but that only an adequate mathematical treatment can give a satisfactory final solution."
>
> —Frits Zernike, 1948

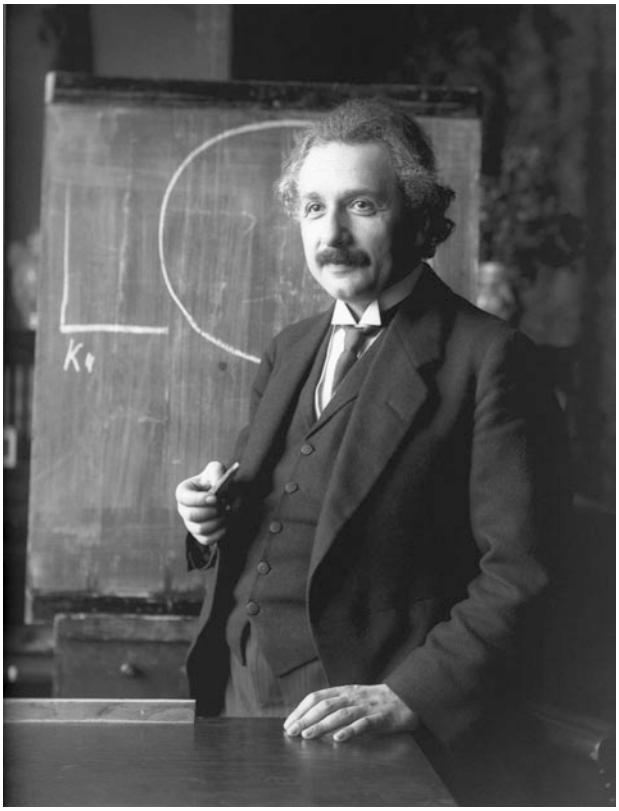
On science and society

> "I maintain the only goal of science is to alleviate the hardships of human existence."
>
> —Bertolt Brecht

Preface

"The pursuit of knowledge for its own sake, and almost fanatical love of justice, and the desire for personal independence—these are the features of Jewish tradition which makes me thank my stars that I belong to it."

—Albert Einstein, *The World as I See It*



Albert Einstein in Vienna, 1921

The goal of this book is to present a critical, comprehensive, coherent, and clear summary of knowledge about techniques that enhance the resolution and image contrast provided by optical microscopes and to critically evaluate and compare this knowledge. In part this book also seeks to provoke critical inquiry and to develop a skeptical approach to science, in general, and to microscopy, in particular. My hope is to stimulate critical thinking and discussion concerning questions about the assumptions, limitations, problems, and achievements of a variety of microscopic techniques. I strive for an understanding of these microscopic techniques that is based on physics, chemistry, and biology. The physics behind each microscope technique is critical to optimal application to address a specific question about a specific function or structure in a specific specimen. Since light interacts with the specimen it is important to understand this interaction, as well as the photophysics, photochemistry, and photobiology that occur during the interaction, and the acute and long-term effects of the light–specimen interaction.

To achieve the stated goal of the book (outlined in the previous paragraph) researchers must have requisite knowledge about the physics, chemistry, and biology of the specimen. In addition, researchers must avail themselves of the appropriate information about light sources, optical and mechanical components of the instrument, light detectors, signals and noise, digital signal processing, statistics, sampling theory, and algorithms used to form, manipulate, quantitate, transform, and interpret the microscopic image.

One metric of success in advancing the field of optical imaging is quantitative enhancement of resolution and/or contrast in microscopy (e.g., a new staining procedure or a new clearing protocol). Each new microscopy technical advance is important in its own right. However, I favor a second metric based on eventual outcomes that follow from the new technology. What biological questions were answered by the development of each new microscopy technique? Are these fundamental questions? Why is current microscope instrumentation incapable of being used to investigate a particular biological question? Did the advance in the development of a new form of microscopy directly result in the capability to investigate the biological specimen and to answer the previously undetermined question? Study of the history of microscopy offers multiple examples of biological questions driving the development of new imaging technologies. An important example was the independent invention of the confocal microscope by several inventors driven by the need for an instrument with the capacity to image neurons and their connections in the brain.

This book is written for students, researchers, and engineers in the life sciences, medicine, and biological engineering who plan to work with or already are working with superresolution light microscopes. While the emphasis is on the life sciences, superresolution microscopic techniques are also applicable to those engineers and scientists in the material sciences. The entire book can serve as a reference for these areas of study. A selected set of individual chapters can be used as the textbook for a one-semester undergraduate or first-year graduate course on superresolution microscopy.

The purpose of this book is to critically summarize the body of knowledge that comprises optical microscopy, to compare and contrast the various instruments, to provide a clear discussion of the physical principles that are the basis for these techniques, and to describe advances in science and medicine for which superresolution microscopes are required and are making major contributions. While the field of microscopy continues to advance rapidly, the scientific community seeks a clear, comprehensive, and critical book that presents and teaches the theories, techniques, instrumentation, and uses of superresolution microscopes to students, researchers, and engineers alike. I frame the discussion of new types of microscopes in terms of problems with existing types of microscopes and their solutions with inventions of new probes and microscopes and I state and emphasize the limitations of each technique and instrument.

This book aims to fill the significant scholarly lacunae that exist in other books that discuss superresolution imaging. First, it places a new emphasis on the specimen, which I consider to be a critical component of the microscope setup. The book's title gives equal importance to the enhancement of resolution and contrast. In the absence of appropriate contrast the image will not contain the required detail regardless of the resolution. The field of stains, dyes, and molecular probes, and in particular the development of fluorescent probes, is intimately linked to progress in microscopy and may be the rate-limiting step to progress in microscopy.

This book contains my English translation of Ernst Abbe's "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung" (Contributions to the Theory of the Microscope and the Nature of Microscopic Vision). This recent translation and my commentary on Abbe's seminal work is followed by Stephenson's paper (1877), which is discussed below.

The skeptical reader may rightfully ask just how relevant these two papers are to the modern reader of a book on *Superresolution Optical Microscopy*. Generally, Abbe's German publication is cited in many articles on superresolution microscopy. Abbe's publication is devoid of any equations or figures. It is a phenomenological description of numerous experiments that Abbe devised and performed. While Abbe promised a sequential publication with the detailed mathematical analysis, his death prevented it being written.

In the literature on the resolution of optical microscopes Abbe's final paragraphs on the diffraction limit are misunderstood and therefore misquoted. Inherent in his publication is a major error in his understanding of the process of image formation. Decades later Abbe admitted this error and offered a correction. A component of the man's genius was his ability to devise what are now called Abbe's Experiments and to use them to demonstrate his theory of image formation in the microscope based on diffraction. When Abbe arrived in London to make his demonstrations at the Royal Society, he learned that the Society's members were not meeting at the time of his visit; therefore, he made his demonstration to those members who were available. Stephenson, the Treasurer of the Royal Microscopical Society, was present for these demonstrations and published a paper that included drawings of his observations of the diffraction patterns in the back focal plane of the microscope objective. His drawings show specific grid patterns as the object and the various

diffraction patterns that resulted from various combinations of diffraction orders entering the microscope objective. Masks were used to block specific diffraction orders and to transmit others. Since Stephenson was a witness to Abbe's experiments I think that his publication with his drawings is of high pedagogical value to the understanding of Abbe's theory. In addition, it complements Abbe's publication on image formation in the microscope by providing the reader with visual descriptions in contrast to Abbe's verbal ones.

I also include my English translation of Helmholtz's full paper titled "Die theoretische Grenze für die Leistungsfähigkeit der Mikroskop" (On the Theoretical Limits of the Optical Capacity [Resolution] of the Microscope), so that the reader can understand how two disparate lines of physical reasoning, one by Abbe and the other by Helmholtz, could result in the same limitation to resolution of the optical microscope. Helmholtz, who worked independently of Abbe, published his own theory of the resolution limit of the microscope. Both these authors independently arrived at the sine condition in optics and the minimal resolution of the light microscope.

This book describes and discusses in depth several topics of basic physics that are typically not discussed in sufficient detail in other works on superresolution optical microscopy. A comprehensive description of Bessel and Airy beams as illumination sources for light sheet microscopy is more extensive than that generally available. The physics of the spiral phase plate and vortex beams, which in STED microscopes form an annular beam of light with zero intensity at its center, is presented with details that are typically lacking in similar books.

Specific Features of the Book

Why publish another book on *Superresolution Optical Microscopy*? Perhaps a list of the book's specific features will help answer this question.

1. The text is both a teaching and reference work. An important feature is the clear presentation of the physics and its mathematical formulation, which forms the basis of each of the microscopy techniques that enhance the resolution and/or the contrast. The book can be used in a consistent and integrated manner to teach the physics, mathematics, instrumentation, chemistry, biology, and spectroscopy of microscopic techniques and instruments.
2. The text clearly develops and discusses all the important equations related to superresolution microscopy. The assumptions, approximations, and instrument limitations are clearly explained. Additionally, the physical meanings of this mathematical description are clearly stated. There are two parallel approaches used in the exposition of the book: rigorous mathematical analysis and description of phenomena with clear and physically correct explanations of superresolution microscopy.

3. All terms and symbols used in the text are carefully defined. All mathematical and physical assumptions used in the analysis are explicitly stated.
4. While the aim of each chapter is didactic, the emphasis is on a critical, comparative analysis of each microscopic technique with a critical discussion of the inherent and current limitations of each instrument.
5. As noted, the book contains my English translation of Ernst Abbe's seminal paper. In this famous paper Abbe explained his theory of image formation in the light microscope based on diffraction. He also derived an expression for the resolution of a light microscope due to diffraction. This paper is often cited, seldom read in the original German, and usually misunderstood. My translation corrects the omissions, added sentences, misunderstandings, and the bias of Fripp who made the original English translation in 1875.
6. This volume also contains my English translation of Helmholtz's paper on the theoretical limits to optical capacity—read resolution—of the microscope titled "Die theoretische Grenze für die Leistungsfähigkeit der Mikroskop" (On the Theoretical Limits of the Optical Capacity [Resolution] of the Microscope). Hermann von Helmholtz (1821–1894), who worked independently of Abbe, developed his theory of image formation in the light microscope based on diffraction. In his paper Helmholtz derived an expression for the resolution of a light microscope due to diffraction that is equivalent to the derivation of Abbe. My translation corrects the omissions, added sentences, misunderstandings, and the bias of Fripp who made the original English translation in 1876.
7. An extensive chapter details Abbe's life, his prescient social reforms and programs, and his innovative contributions to the field of optics. I investigate what aspects of Abbe's education led to his predilection for precision measurements and his unique ability to derive and demonstrate his theory of image formation in the microscope in phenomenological terms (the famous Abbe Experiments)?
8. A comprehensive chapter describes Abbe's diffraction-based theory of image formation in the light microscope. Seminal books and publications by his students and colleagues at Zeiss Werke in Jena are discussed in great detail. I also point out the error in Abbe's theory of image formation in the light microscope and his acknowledgment of this key error decades later.
9. Abbe's creative experiments are discussed in terms of modern diffraction theory, and their pedagogical role in disseminating his theory is developed. Abbe's genius allowed him to develop simple experiments that could be performed with simple optical components and could be used to experimentally demonstrate his theory of image formation in the microscope based on diffraction.
10. The ethical dimensions of each discovery, codiscovery, or simultaneous discovery by independent individuals or groups is embedded in the discussion of each new microscope. Where necessary I point out the existence of misleading or obfuscating statements and the lack of proper citation.
11. I present tutorials on topics that are largely undeveloped in books of a similar genre. These topics and tutorials include the historical context of discovery and rediscovery in microscopy; the concepts of resolution and contrast, as well as

those of molecular spectroscopy; the physics of spiral waveplates; stimulated emission; moiré patterns; pathways to discovery and innovation in science; phase microscopy; the superresolution techniques of structured illumination microscopy (SIM) and stimulated emission depletion (STED) microscopy; the localization techniques of photoactivated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM); the attribution of credit for invention and discovery; linear and nonlinear spectroscopic effects; the quantum properties of photons; the Nyquist–Shannon sampling theorem; Fourier transforms; reciprocal space; an extensive presentation on light sheet microscopy; Bessel and Airy beams; self-reconstructing beams; and lattice light sheet microscopy. For each of the far-field (the domain of Fraunhofer diffraction) superresolution optical microscopic techniques I explain and discuss the required trade-offs, cautions, and limitations. In far-field microscopy the light from the object (specimen) is located many wavelengths from the aperture of the microscope objective. The optical system of the microscope forms an image of the object. Space restrictions required limiting the exposition to far-field light microscopy. This decision in no way limits the utility and the importance of various near-field optical microscopes[1] such as scanning near-field optical microscopes. Near-field optical microscopes depend on nonpropagating evanescent wave fields. Typically, the probe that collects the light from the object is situated at a distance less than one wavelength from the object. Other important techniques, such as photoacoustic microscopy, X-ray diffraction, and electron microscopy, are not discussed in the book.

Curiosity, Skepticism, Innovation, and Creativity in Science and Technology

Superresolution Optical Microscopy: The Quest for Enhanced Resolution and Contrast integrates topics of creativity in science, curiosity in science, and science and society with the technical innovations of superresolution light microscopy.

This book is about curiosity, skepticism, innovation, and creativity in science and technology. The brilliant developments in microscopy are the result of individuals asking the right questions and pursuing logical consequences. Science begins with curiosity and asking questions; genius is found in posing good questions. The foundations of science are curiosity and skepticism. In my initial lecture to students I tell them "Trust no one." By that dictum I mean that they should critically evaluate the evidence behind every statement. Be cautious of the statements of experts. Beware of authority. Albert Einstein proclaimed, "Blind respect

[1]Jutamulia, S. (2002). *Selected Papers on Near-Field Optics*. Bellingham: SPIE Optical Engineering Press.

for authority is the greatest enemy of truth." Murray Gell-Mann in his book *The Quark and the Jaguar* wrote, "When the Royal Society of London was founded in 1661, the fellows selected the motto *Nullius in verba*. I interpret that phrase as meaning 'Don't believe in anyone's words' and as being a rejection of the appeal to authority." *Nullius in verba* expresses the compelling desire of the fellows of the Royal Society to resist the prevalent domination of authority and to validate all statements by evaluation of facts determined by experiment. The motto was valid in 1661 and is still valid today.

The history of science has many examples of creativity and obstacles to it. On the one side, there is prejudice, bias, popular opinion, authority, collective thinking and consensus, and the credibility of "what is known." On the other side, there are creative, courageous, and insightful ideas that promote the formulation of questions, models, and theories that can be subject to experimental tests. In the field of physics we are familiar with the works of Ludwig Boltzmann, Albert Einstein, and Richard Feynman. In the field of medicine we are familiar with the work of Stanley B. Prusiner on prion diseases such as Creutzfeldt–Jakob disease and bovine spongiform encephalopathy. These scientists were skeptical of authority and group thinking and, despite the difficulties of working against the tide of "what is known," they succeeded in advancing science. Boltzmann struggled to achieve acceptance of his kinetic theory of gases that was based on the reality of atoms. Ernst Mach and Wilhelm Ostwald were the authority figures who denied the existence of atoms. Albert Einstein received the Nobel Prize for his theoretical work on the photoelectric effect. His seminal work on relativity was considered too speculative by the Nobel Committee and most of his contemporary physicists.

In the sections of this book devoted to Ernst Abbe's theory of image formation in the microscope and Fritz Zernike's invention of phase contrast microscopy I will return to these questions of creativity and innovation in science. The history of the microscope is replete with examples of simultaneous inventions that resulted in enhanced contrast, enhanced resolution, or both. As I discuss these inventions I will briefly point out instances of simultaneous independent invention. I demonstrate that Abbe, Helmholtz, and Rayleigh independently approached the question of microscopic resolution from different directions; however, they derived similar mathematical expressions to explain the limit to resolution. In many cases their publications offered proper attribution to the prior and simultaneous work of others.

The Importance of Responsible Conduct of Research

> "I maintain the only goal of science is to alleviate the hardships of human existence."
>
> —Bertolt Brecht

The social process of science, which is built on trust, is exciting, stimulating, and a source of new knowledge and understanding of the physical world. Scientists' goals

include mentoring students, generating knowledge, and communicating and sharing this knowledge. Scientists are responsible for the promotion, protection, and advancement of science. As stated in Shamoo and Resnik (2009), "There is a growing recognition among scientists, government officials, research institutions, and the public that ethical conduct is essential to scientific research. Ethical conduct is important to foster collaboration, cooperation, and trust among scientists, to advance the goals of research, to fulfill scientists' social responsibilities, and to avoid or minimize damaging scandals resulting from unethical or illegal behavior" (Shamoo and Resnik, *Responsible Conduct of Research*, second edition, 2009). I have developed a course on the responsible conduct of research and have taught it at the Massachusetts Institute of Technology (MIT) and at universities in other countries. Appendix B titled "Responsible Conduct of Research" summarizes the standards of research ethics that are the content of my courses. I have made my course presentation of "Responsible Conduct of Research" available as a pdf on the book's online website: https://storage.googleapis.com/springer-extras/zip/2020/978-3-030-21690-0.zip.

I thank the anonymous reviewers for their helpful comments and suggestions. I thank Professor William T. Rhodes for his critical editing of the manuscript and his discussions. I am solely responsible for any errors in the book. I am also grateful to Sam Harrison, my Editor at Springer Nature, for his suggestions and help in the production of this book.

Cambridge, USA Barry R. Masters

Contents

About the Author

Professor Barry R. Masters has received a Ph.D. from the Weizmann Institute of Science, Israel, an MS degree from Florida State University, and a BS from the Polytechnic Institute of Brooklyn, New York. Professor Masters was Visiting Scientist in the Department of Biological Engineering at the Massachusetts Institute of Technology, Visiting Scholar in the History of Science Department at Harvard University, Visiting Professor in the Department of Ophthalmology at the University of Bern, and Professor in Anatomy and Cell Biology at the Uniformed Services University of the Health Sciences. He was a Visiting Professor in Japan, India,

Australia, Switzerland, Taiwan, Turkey, Wales, England, P. R. China, and Egypt. He is a Fellow of the American Association for the Advancement of Science (AAAS), the Optical Society of America (OSA), and the International Society for Optics and Photonics (SPIE). From 1999 to 2000 Professor Masters was an AAAS Congressional Science and Engineering Fellow and served as Legislative Assistant in the US Congress. He is a recipient of the Vogt Prize for Ophthalmic Research (highest Swiss award for research in ophthalmology) with Professor Dr. Matthias Böhnke in 1999 for their work on "Confocal Microscopy of the Cornea." Professor Masters has published 87 refereed papers, 143 book chapters and proceedings articles, 105 abstracts, and 244 book reviews. He is the editor or author of 10 books, among them *Noninvasive Diagnostic Techniques in Ophthalmology*, *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*, and *Handbook of Biomedical Nonlinear Optical Microscopy* (with Professor T. C. So). He wrote the essay "What Is Light?" (translated into 15 languages) for the International Commission for Optics (ICO) to inaugurate the International Year of Light, 2015 (http://e-ico.org). He lectures worldwide on the Responsible Conduct of Research, biomedical ethics, critical thinking, and biomedical photonics. His scholarly interests include developments in the fields of in vivo microscopy of the human eye and skin, biomedical imaging and spectroscopy, fractal analysis of branching vascular patterns, ethics in science and medicine, the history of 19th- and 20th-century physics, science education, and art and science.

Part I
Development of Microscopes

Introduction

The book begins with a stimulating set of simple questions: What is resolution and what is contrast in an image? What does a microscope do? What is an "ideal" microscope? Why can the unaided eye not observe microscopic objects? Why were microscopes invented?

What is resolution? There are a variety of definitions, all of which are confounded by aberrations and noise. One definition of the resolution of an optical microscope is the capability to image two adjacent point objects (point sources of light) as two separate objects. In 1896 Rayleigh posed the question: When are two distinct point sources of light imaged as two points or blurred due to the finite numerical aperture of the lens? If the separation distance of two point objects is below the resolution of the optical microscope, then the two point objects are imaged as a single object. The field of physical optics provides a variety of definitions of resolution; it differs for incoherent light, coherent light, and for telescopes and microscopes. Different authors provided different definitions of resolution. Rayleigh noted the effects of noise on resolution and stated that the definition of resolution is dependent on the optical system. Another confounding factor is the effect of optical aberrations on resolution.

The historical development of microscopes from the early single-lens microscope to microscopes constructed following Ernst Abbe's theory of image formation based on diffraction is the content of Part I. The timeline of these developments suggests not a continuous progression in design and construction, but a pattern of advancement and regression and further advancement in the development of microscopes. Prior to the work of Abbe, empirical design and manufacturing techniques yielded microscope objectives of variable quality. Parallel with the development of microscopes were advances in the manufacture of glass, the development of dyes and stains, the development of the microtome to slice a fixed

specimen into extremely thin sections, new illumination sources, and improved microscope objectives that reduced optical aberrations.

Part I provides my English translation of Abbe's seminal paper on the theory of the microscope, my commentary on his publication, and a discussion of his experiments that he used together with his public lectures to promote his theory of the microscope. I reprint with my commentary Stephenson's publication together with his figures that follow from his being a witness to Abbe's demonstration experiments in London. Stephenson's eye-witness account in London of Abbe demonstrating his experiments is the only account of this historical event that is written by someone who observed the demonstrations. Abbe's theory of 1873 never appeared in mathematical form during his lifetime. It was misunderstood, misinterpreted, and poorly accepted by the physics community as well as the majority of microscopists for a variety of reasons that are discussed in Part I. In fact, Abbe's theory contained a major error and Abbe only corrected it decades later. In Appendix A Abbe's words about his error and its eventual correction are quoted. After Abbe's death his colleagues published his lecture notes on his mathematical theory as well as further analysis of Abbe's experiments on the role of diffraction in image formation in the light microscope. The parallel theories of Helmholtz and Rayleigh were derived from different independent approaches yet resulted in equivalent equations for the dependence of resolution on wavelength and aperture.

In addition to my translation of Abbe's seminal paper of 1873, "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung" (A contribution to the theory of the microscope and the nature of microscopic vision), I also translated Helmholtz's 1874 paper "Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope" (On the theoretical limits of the optical capacity [resolution] of the microscope). Both these translations were made to correct the omissions and incorrect statements that comprised Fripp's earlier translations from German into English.

The history of microscopy is the story of empirical developments, theoretical developments, and mechanical, electrical, computer, and optical engineering. The resulting devices had major impacts on basic sciences and on medicine. There were many examples in which specific communities of researchers had minimal communication with other research communities. The development of microscopes to observe the eye and to measure its optical properties is centered on the ophthalmology community. Other interesting examples indicate the inventive role of individuals. The invention of the phase contrast microscope by Zernike is an amazing story. The phase contrast microscope became ubiquitous in hospitals and biological laboratories and was a critical tool in the study of cells and cellular processes in cell culture. Another interesting example is the independent invention of confocal microscopes by three individuals located in three different countries. They shared one common goal: how best to image thick, highly scattering specimens of the brain. This is an example of how a question, in this case how to image neurons and their multiple connections in the brain and thus improve our understanding of brain function, was the driving force behind technological innovation.

Other examples of the development of new forms of microscopes were driven by the aim to enhance the resolution and/or contrast of microscopes.

In independent theoretical publications that appeared within one year of each other both Abbe and Helmholtz worked out the limits of optical resolution for an optical microscope. This limit was historically called the diffraction limit of resolution. We now know that this equation is only an estimate of resolution; resolution is also affected by noise, optical aberrations, and the shape of the point spread function. These physicists produced formulas that expressed resolution of an optical microscope in terms of the wavelength of incident light and what is today called the numerical aperture of the microscope objective. In the conventional light microscope, diffraction limits the resolution to 200–300 nm in the lateral dimension. Following the works of Abbe and Helmholtz many physicists have explained the resolution limit of an optical microscope in terms of the diffraction of light from the finite aperture of the microscope objective. Therefore, diffraction is the fundamental cause of the resolution limit of an optical microscope.

The study of biological structures requires a spectrum of microscopic techniques from the ultrahigh resolution of high-energy electron microscopes to a variety of optical microscopes that use many different techniques to enhance contrast. Many cell organelles and supermolecular assemblies are below the classical resolution limits of conventional light microscopes. There is a gap in our imaging techniques between the high resolution of the electron microscope and the diffraction-limited resolution of modern optical microscopes. While electron microscopes offer high resolution, they are not suitable to investigate the dynamics of cellular processes.

To successfully investigate the complex, time-dependent structure and functional changes in biology a variety of techniques are required, each one with its inherent limitations and unique advantages. Complex biological processes will ultimately be understood by appropriate parallel and interdisciplinary approaches that involve different levels of biological scale (atoms, molecules, cells, tissues, organs, and populations of organisms).

New types of microscopy will pave the way to observing living processes in ever-finer detail and will enable study of the dynamics of living cells at the molecular level. Today, microscopy is a vibrant field of invention, technological development, and commercial product development, and all this activity is driving new advances in biology, medicine, and materials science. Concomitantly, future advances in microfabrication and nanofabrication, high-resolution photolithography, and our understanding of materials at the atomic and molecular level are dependent on improving the resolution and the contrast of microscopes.

Chapter 1
Connections Between Light, Vision, and Microscopes

In this chapter I look at the historical development of our conceptions of light and our understanding of the human visual system and their connection to the development of microscopes. Microscopes are instruments that can be used to visualize objects that are not observable with the naked eye. How is our understanding of light and our knowledge of the structure and function of the human visual system related? How is the development of microscopes related to our understanding of light?

Light and vision provide a major source of our sensory knowledge of the external world. People have undertaken work to see the heavens in greater detail and to see objects on smaller and smaller scales. Inquiry begins with observations and questions. The cycle of observations and questions and explanations, and more questions, and more observations have stimulated people to try to understand the nature of light, vision, and the microscopic world. The development of the light microscope is closely associated with improved understanding of light: the propagation of light, the interaction of light and matter, and our ability to manipulate the properties of light. But modern advances in the fields of geometrical optics and physical optics—and further advances in our understandings of coherence theory, nonlinear optics, and quantum optics—have all impacted the development of new types of microscopes.

What is light? The answer is still a work in progress (Adams and Hughes, 2019; Masters, 2015; Ronchi, 1955, 1957, 1979). For those who dream of developing a new type of microscope I suggest they learn modern theories of light and its interaction with matter. There is also much to be learned by reading primary sources (i.e., the original papers on the development of our knowledge of light). Of the plethora of sources related to the history of optics I have selected a subset that I have found to be both stimulating and useful, and I share them with the reader. Of course, there is no substitute for reading the primary sources of philosophers and inventors—their original books and publications. Other useful resources are popular textbooks and handbooks that critically compile and evaluate known knowledge of

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_1

a past time. These include the German volumes of Winkelmann's *Handbuch der Physik* (Winkelmann, 1906).

From their works we learn that the early Greek philosophers thought the goal of the study of optics was to explain vision and the appearance of objects. However, at the end of the 17th century the major aims changed from understanding the visual process to the nature of light and the understanding of its properties such as reflection, refraction, and diffraction. An excellent secondary source for this information is Smith's recent book *From Sight to Light, the Passage from Ancient to Modern Optics* (Smith, 2015).

Let us begin in the early part of the 11th century with Alhazen, called the father of optics due to his seminal studies and writings on both optics and the structure and function of the human eye. His Latin name is a transcription of the Arabic name Ibn al-Haytham. He experimented with mirrors and lenses and studied reflection and refraction. Alhazen wrote his *Kitab al-Manazir* (Book of Optics) between 1011 and 1021. This influential treatise was translated into Latin at the end of the 12th century and printed in 1572 as the *Opticae Thesaurus* (Optics Treasure) (Sabra, 1989). Books I–III discuss Alhazen's theories of the rectilinear propagation of light, color, his theory of visual perception, and errors in vision and their causes. The last four books that comprise the *Opticae Thesaurus* elucidate his theories of reflection and refraction of light and his studies on vision produced by reflected and refracted light rays. For a century Alhazen's treatise was the standard source of knowledge on optics and visual perception.

In 1604 Kepler proposed a theory of image formation in the human eye. But it was the Jesuit Christoph Scheiner (1573–1650) who actually demonstrated Kepler's proposal. Scheiner cut out some of the tissue at the back part of an excised human eye and several animal eyes, and when he observed the image projected through a hole in the retina he viewed an inverted image (Wade, 1998; Werner and Chalupa, 2004, 2014).

The history of optics in Western thought is replete with many connections to other sciences (e.g., the analogies between light and sound; Darrigol, 2012; Park, 1997). Claudius Polemaeus, known in English as Ptolemy (AD 90–168), worked in Alexandria and is well known for his comprehensive treatise on astronomy titled *Almagest*. But Ptolemy also wrote the important work *Optics*, which today exists as a 12th-century Latin version of an earlier Arabic translation that was made from the Greek original. Ptolemy's *Optics* contained his explanations of visual perception: size and shape, color, and binocular vision (Smith, 1996).

The first edition of Newton's *Optics*, a masterful classic of physical science, was published in 1730 (Newton, 1952). In the 17th century Christiaan Huygens made significant contributions to physics, especially to our understanding of the wave nature of light based on what is now called Huygens' principle or construction (Andriesse, 2005; Bell, 1947).

At the beginning of the 19th century there was a dramatic development: the ether became predominate in theoretical physics, and waves replaced rays (Buchwald, 1985, 1989, 1994; Chappert, 2007). The creative innovation of Michael Faraday introduced the new concept of the electromagnetic field. The next important step, as

discussed by Darrigol (2000), was in the unification of electric and magnetic phenomena in terms of Maxwell's equations, which represent a critically important milestone in the development of mathematical physics and exemplify the unity of physics. Darrigol also explains how developments in electrodynamics provided the foundation of Einstein's special theory of relativity. Other developments, described in the historiography of the development of electromagnetic theory and its experimental realizations by Heinrich Hertz, provide further insights into the nexus of theory and experiment in late 19th-century physics (Buchwald, 1985, 1994; Masters, 2011). Albert Einstein made fundamental contributions to our understanding of the nature of light and the idea of wave–particle duality; his innovative theoretical work bridged the gap between radiation in space and radiation–matter interactions (Masters, 2012).

From our discussion of the history of our understanding of light we proceed to the history of our understanding of vision, and then we are in a position to discuss microscopes. The human visual system composed of the eye and the brain is remarkable in its complexity and functionality (Werner and Chalupa, 2004, 2014). A good source of knowledge on the computational approach to biological vision is the book *Seeing* (Frisby and Stone, 2010). For many decades the standard work on the human visual system was Hermann von Helmholtz's *Treatise on Physiological Optics* (Helmholtz, 1909). Recently, Artal published a comprehensive, modern review on image formation in the living human eye and a two-volume set on human visual optics (Artal, 2017). Comparison of the works of 1909 and 2017 illustrates the progress made in our understanding of the human visual system.

The history of the study of the human visual system is fascinating and contains many connections to the historical developments of our understanding of the nature of light and its interaction with matter. I have selected six books that I believe provide the reader with a comprehensive history of this study: Lindberg (1976), Park (1997), Ronchi (1955, 1957, 1979), and Wade (1998). This list is augmented by the publication *From Sight to Light* (Smith, 2015).

The history of the microscope is connected to the human eye. But why should we believe what we see in the microscope? Are the images just artifacts? How can we verify that microscope images are real and free from artifacts? What is the relation between the observer and the object seen in the microscope (Schickore, 2007)? These confounding and important questions persist to this day.

The human retina is sensitive to single photons. The human eye is sensitive to five decades of intensity in bright sunlight. But with slow adaptive mechanisms in the presence of low light (e.g., starlight) the human eye can exhibit a sensitivity over a dynamic range of nine decades of intensity. Adaption is the mechanism by which the eye changes (over a period of 30 min) from a scotopic (rod-driven night or low-light) system to a photopic (cone-driven daylight) system.

The human visual system is excellent for pattern recognition, motion detection, stereopsis, and color vision (Artal, 2015; Werner and Chalupa, 2014). But the human retina has limitations. It is only sensitive to light of a specific wavelength range: the visual spectrum. The finite size of the pupil (in the range of 2–8 mm) limits the resolution of the eye (i.e., the ability to see two separate objects as distinct

objects when only one object is present). When the pupil diameter is 3 mm the Airy disk subtends 1.5 arcmin. At the near point 1 arcmin is subtended by 75 μm for the human eye and 1 arcsec is subtended by 1.2 μm. The two-point resolution for the average human eye is 1 arcmin (Goodman, 1988).

The image fidelity of the eye is further limited by optical aberrations introduced by the cornea, the ocular lens, and the media within the eye. Light scattering and absorption within the eye reduce the intensity and contrast of incident light that is detected by the retina. Diffraction of light from cataracts and opacities within the ocular lens result in spurious halos of light around point sources of light.

Objects that we may wish to observe span a wide range of scales in their linear dimensions. The old literature on microscopy defines microscopic objects as those that cannot be seen with the naked human eye. The term nanoscopic as used in the modern literature on superresolution microscopes refers to the imaging of objects in the range of 1–100 nm. Microscopic objects (cells, bacteria, chromosomes) by definition have linear dimensions in the range of 1–100 μm. Nanoscopic objects (proteins, viruses, ribosomes, carbon nanotubes, quantum dots) by definition have linear dimensions in the range of 1–100 nm. How can we observe microscopic and nanoscopic objects? The content of this book is the story of the amazing quest over centuries to achieve enhanced resolution and contrast (the ability to detect an object from the background). In the context of fluorescence microscopy and in superresolution optical microscopes contrast is obtained by using fluorescent molecules that are specifically bound to specific parts of the specimen. Alternatively, phase contrast microscopy forms contrast in unstained specimens using a light microscope that converts phase differences in various parts of the specimen into intensity differences in the image. It is the story of an exciting field of human endeavor that is, like the study of light itself, a work in progress.

From our empirical experience we know that as the distance between an object and the eye decreases we can see smaller objects. Therefore, to observe ever-smaller objects why not just move them ever-closer to the eye? We are familiar with large magnifying glasses to read the small print in books and newspapers. In a movie theater the movie projector projects a greatly magnified image of the object (a single frame of the film) onto the movie screen. The magnification can be very high. Why not use that principle to observe microscopic objects? These are reasonable questions. But in the next paragraph I explain why these ideas will not work. As we shall see, resolution is not the equivalent of magnification.

To understand the answers to the previous questions it is necessary to introduce the following terms, and at this point to provide elementary definitions that in the course of the subsequent chapters are developed in greater detail. The terms are magnification, resolution, and contrast. Magnification is multiplication of the linear dimensions of an object. For example, an object that is 1 mm high and 1 mm wide with a magnification of 10× would be observed as an object 10 mm high and 10 mm wide.

Many types of magnifying instruments form parallel light and virtual images. Such instruments are characterized by angular magnification. Angular magnification, sometimes called the magnifying power, of an optical instrument is the ratio of

the size of the retinal image as observed through the instrument to the size of the retinal image as seen with the naked eye at a normal viewing distance (called the near point). The near point is the closest position of an object to the eye for which the eye can focus. Similarly, the far point is the farthest. An example is the simple magnifying glass that some people used for reading. The lens forms a virtual image of an object. This virtual image subtends a greater angle than the object and therefore it appears larger to the observer. Such a device produces an increased angular subtense of the retinal image. The increased angular subtense of the virtual image with the magnifier measured in radians divided by the angular subtense of the object at a comfortable viewing distance (250 mm for young adults) is the definition of magnifying power (MP). Magnifying power is a measure of how much larger the virtual image is relative to the object at some distance from the lens.

Another important concept is empty magnification that occurs when a barely resolvable element in the image exceeds roughly a minute of arc when viewed by the eye. This relation between the angular resolution of the eye and the angular subtense of the just-resolved detail determines whether magnification is useful or not. Empty magnification (typically above 1000× or 1500×) is magnification that is not useful since it does not result in increased resolution and therefore no additional details are observed.

Resolution (the subject of Chapter 2), a property of an optical imaging system, refers to the ability to observe small details and is often related to the ability to observe two closely separated objects as distinct. As the resolution of an imaging system is increased the separation between two small microscopic objects, such as point sources of light, can be decreased and they can still be observed as distinct objects. The resolution of the human eye limits the separation distance between two distinct objects, so that when the distance of separation is further decreased the two objects are no longer observed as distinct from each other.

The third term is the contrast of an image. Image contrast can be defined in terms of the ratio of light intensities between different areas (or pixels) in the optical plane of a microscope, such as the difference in intensity between different points of a specimen or between a specimen and the background

$$\text{Contrast} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{1.1}$$

Contrast in an image is determined by several parameters: the number of detected photons, noise, optical aberrations, and the number of pixels or picture elements per unit area.

A microscope has the capability to resolve objects that are beyond the resolution of the unaided human eye. A microscope also increases the contrast of the object over the background. Two examples are the fluorescence microscope and the phase contrast microscope. Which is more important in a microscope: resolution or contrast? Both resolution and contrast are necessary. If there is sufficient resolution in the microscope but insufficient contrast in the object, then the object cannot be observed. Contrast in the specimen permits the observer to see the object as separate

from the background. If there is sufficient contrast in the microscope but insufficient resolution, then the object may be observed without observing some details of the object. Magnification in a microscope without concomitant increased resolution is empty magnification and therefore useless because empty magnification will not provide increased resolution.

Finally, I present some thoughts on the future existence of an "ideal microscope." Traditional optical microscopes exist whose resolution is sufficient for objects that have linear dimensions of a few hundred nanometers. For objects in the range of small molecules or atomic dimensions there are electron microscopes and microscopic techniques that are based on neutron diffraction or X-ray diffraction. Recent exciting innovations in microscopy are centered on the development of microscopes that can resolve objects of scale down to a few nanometers. These developments are important since they provide microscopes with a resolution that was not previously possible—in that range between those of conventional optical microscopes and electron and X-ray microscopes. Other desirable features of an ideal microscope would include the following: high image acquisition speed so that dynamic processes can be imaged, the capability to image large areas of the object (large field of view), and the development of new optical modes of contrast generation. The latter falls into the category of "probeless" contrast generation. The term probeless refers to optical methods to generate contrast in the specimen without the use of fluorescent molecules that are bound to specific parts (e.g., specific proteins) of the specimen. These fluorescent molecules are called fluorescent probes. Such new types of probeless contrast generation would mitigate the difficulties of using molecular probes such as genetically expressed fluorescent proteins that are typically overexpressed in cells.

The development of probes and the modification of the object by so-called clearing techniques is now the rate-limiting step in the development of new types of microscopes. A major problem in the microscopic imaging of fixed tissue sections is caused by light scattering within the specimen, scattering that also limits imaging penetration depth. Clearing is a technique that makes tissue specimens more transparent by incubating them with organic solvents that deplete the lipids of the specimen and reduce the refractive index heterogeneities of the specimen. This process results in significantly reduced light scattering and improved imaging penetration depth.

Active efforts in these two fields of research (the development of probes and the development of clearing techniques) are indicative of their importance in the new generation of microscopes that are advancing our knowledge of biology and medicine. The deeper our understanding of light the more possibilities we have to invent and develop new types of microscopes in our never-ending quest for instruments with enhanced resolution of the object and enhanced contrast in the image. These new scientific instruments will be used to provide further insights and understanding of the structure and function of cells, in particular, and the nervous system of organisms, in general. There is a true synergism between our understanding of light and the development of new types of optical microscopes that can be used to further our understanding of both the microscopic and the nanoscopic world.

References

Adams, C. S., and Hughes, I. G. (2019). *Optics f2f.* Oxford: Oxford University Press.

Andriesse, C. D. (2005). *Huygens: The man behind the principle*. Cambridge: Cambridge University Press.

Artal, P. (2015). Image formation in the living human eye. Annual Review of Vision Science, **1**, 1–17.

Artal, P. (2017). *Handbook of Visual Optics, Volume I, Volume 2*. Boca Raton: CRC Press.

Bell, A. E. (1947). *Christian Huygens and the development of science in the seventeenth century*. London: Edward Arnold & Co.

Buchwald, J. Z. (1985). *From Maxwell to Microphysics: Aspects of Electromagnetic Theory in the Last Quarter of the Nineteenth Century*. Chicago: The University of Chicago Press.

Buchwald, J. Z. (1989). *The Rise of the Wave Theory of Light. Optical Theory and Experiment in the early Nineteenth Century*. Chicago: The University of Chicago Press.

Buchwald, J. Z. (1994). *The Creation of Scientific Effects. Heinrich Hertz and Electric Waves*. Chicago: The University of Chicago Press.

Chappert, A. (2007). *Histoire de l'optique ondulatoire. De Fresnel à Maxwell*. Paris: Belin.

Darrigol, O. (2000). *Electrodynamics from Ampère to Einstein*. Oxford: Oxford University Press.

Darrigol, O. (2012). *A History of Optics. From Greek antiquity to the nineteenth century*. Oxford: Oxford University Press.

Frisby, J. P., and Stone, J. V. (2010). *Seeing: The Computational Approach to Biological Vision, Second Edition*. Cambridge: The MIT Press.

Goodman, D. S. (1988). *Basic Optical Instruments, in Geometrical and Instrumental Optics* . Edited by Daniel Malacara, San Diego: Academic Press.

Helmholtz, H. (1909). *Handbuch der Physiologischen Optik, Dritte Auflage*, English translation of the third edition, three volumes (1909–1911) edited by J. P. C. Southall. Washington, DC: Optical Society of America, 1924.

Lindberg, D. C. (1976). *Theories of Vision from Al-Kindi to Kepler*. Chicago: The University of Chicago Press.

Masters, B. R. (2011). Heinrich Hertz and the Foundations of Electromagnetism. OPN Optics & Photonics News. June, 30–35.

Masters, B. R. (2012). Albert Einstein and the nature of light. OPN Optics & Photonics News. July/August, 42–47.

Masters, B. R. (2015). What is light? ICO Newsletter, January 2015, Number 102, http://e-ico.org. Translations into Spanish, Latvian, Chinese, French, Greek, Armenian, Portuguese, Slovak, Hebrew, Italian, Hindi, Persian, Korean, Turkish and Arabic. Retrieved April 1, 2019.

Newton, Sir Isaac (1730). *Opticks or a Treatise of the Reflections, Refractions, Inflections & Colours of Light*, based on the fourth edition London 1730. Reprinted in 1952, New York: Dover Publications, Inc.

Park, D. (1997). *The Fire within the Eye. A historical essay on the nature and meaning of light*. Princeton: Princeton University Press.

Ronchi, V. (1955). *L'Ottica di Euclide e la scienza della visione*. Bologna: Nicola Zanichelli.

Ronchi, V. (1957). *Optics, the science of vision*. New York: New York University Press. Reprinted in 1991, New York: Dover Publications, Inc.

Ronchi, V. (1979). *The Nature of Light: An Historical Survey*. Translated by V. Barocas. Originally published in Italian as Storia della Luce, 1939. London: Heinemann.

Sabra, A. I., trans. (1989), *The Optics of Ibn al-Haytham. Books I–II–III: On Direct Vision. English Translation and Commentary*. 2 vols, Studies of the Warburg Institute, vol. 40, London: The Warburg Institute, University of London.

Schickore, J. (2007). *The Microscope and the Eye. A History of Reflections, 1740–1878*. Chicago: The University of Chicago Press.

Smith, A. M. (1996). *Ptolemy's Theory of visual Perception: An English Translation of the Optics with Introduction and Commentary*. Transactions of the American Philosophical Society, **86**, part 2. Philadelphia: The American Philosophical Society.

Smith, A. M. (2015). *From Sight to Light, the Passage from Ancient to Modern Optics*. Chicago: The University of Chicago Press.

Wade, N. J. (1998). *A Natural History of Vision*. Cambridge: The MIT Press.

Werner, J. S., and Chalupa, L. M., Eds. (2004). *The Visual Neurosciences, Volume 1, 2*. Cambridge: The MIT Press.

Werner, J. S., and Chalupa, L. M. (Eds.). (2014). *The New Visual Neurosciences*. Cambridge: The MIT Press.

Winkelmann, A. (Ed.). (1906). *Handbuch der Physik, Zweite Auflage. Sechster Band: Optik*. Leipzig: Verlag von Johann Ambrosius Barth.

Chapter 2
Concepts and Criteria of Resolution

2.1 Introduction: What Is Resolution?

In the first chapter I defined the terms resolution, contrast, and magnification. I contrasted the terms resolution and magnification. I stressed the synergy between resolution and contrast in the microscope. In this chapter I discuss the various definitions and concepts that are used as metrics of resolution. In addition, I introduce some topics from physical optics and others from information theory that lead to a deeper understanding of the concept of resolution. I will answer the following questions about optical resolution and resolving power: What is resolution? How do we measure resolution? What limits resolution and what confounds our concepts and understanding of it? Then, I segue from a discussion of resolution concepts to the concept of superresolution, the content of Part III of this book.

Resolving power refers to the ability of an optical instrument (e.g., the human eye, a microscope, or a telescope) to separate equal intensity point sources of light as two distinct points. Resolving power is a property of an optical instrument and the properties of light. It can be expressed as the finest detail of a specimen that an optical instrument is able to resolve (Slayter and Slayter, 1992).

Resolution is the minimal distance that separates two point sources of light in the microscope to form an image in the image plane for which the two point sources of light are discernible as two separate objects. The resolution of a microscope, the minimum resolvable distance discernible in the object, is a function of several parameters: the resolving power of the instrument, the contrast and the shape of the objects, the coherence of the illumination, the optical aberrations of the optical system, and the signal-to-noise ratio in the image detector. All these parameters affect microscope resolution.

How do we measure the axial and lateral resolution of the microscope? One method to obtain an experimental measure of axial resolution is to measure the intensity of light reflected from the source as a function of the distance from the focal plane of the microscope objective and a plane mirror placed on the

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_2

microscope stage. This is performed by moving a mirror axially through the focal plane of the microscope objective and measuring the reflected light intensity as a function of the defocus distance (e.g., the distance above and below the focal plane). The axial resolution (along the optical axis) can then be defined as the width of the plot at half-maximum intensity. Optical aberrations in the microscope cause a strong asymmetry in the plot of light intensity versus distance. The lateral resolution of an optical microscope can be estimated by imaging a standard microscope test specimen—a slide containing various patterns with different spatial frequencies or a biological specimen such as a diatom. Another useful test object is an integrated microchip with known line spacings previously determined by an electron microscope. All test specimens or objects require validation with independently measured line spacing (i.e., prior use of electron microscopy to measure the line spacing of a test specimen).

2.2 Role of Diffraction in Image Formation

Light diffraction, which is predicted by Maxwell's equations, occurs when light propagates through a limiting aperture (e.g., a finite diameter lens, a slit, or other type of aperture). The diffraction of light in the optical microscope limits its resolution. The commonly used term diffraction-limited optical system is defined as an optical system that has a resolution limited by the phenomenon of light diffraction. In the 1600s Francesco Grimaldi coined the term diffraction, which corresponds to the deviation of light from rectilinear propagation. I have written the first half of an open-access publication in the European Physical Journal H that provides a comprehensive survey of the role of diffraction in image formation, a historical survey of the developments of diffraction, and a detailed survey of various concepts of resolution (Cremer and Masters, 2013).

Light diffraction is integral to Abbe's theory of image formation in the light microscope. Independent derivations of limiting resolution in the light microscope were formulated by Abbe, Helmholtz, and Rayleigh. Part III of this book describes and analyzes the development of far-field superresolution microscopes that are capable of surpassing the resolution limits of classical diffraction-limited microscopes. Therefore, a brief introduction to light diffraction is necessary.

What is diffraction? Imagine that we have a point source of light (i.e., a source without finite dimensions). For example, a star at a great distance from the earth. We use a circular lens to image the point source. What do we observe in the object plane? We do not observe a point of light; instead, we observe a pattern of concentric rings, each with different intensities. This is the diffraction pattern of the point object as imaged by a circular lens of finite dimension.

This phenomenon cannot be explained by classical geometrical optics. The goal of mathematical theories of diffraction is to predict the pattern of diffracted light from an object as a function of the shape and dimensions of the lens, the wavelength of light, and distance from the lens in the image space.

A good introduction to the various theories of diffraction is the book *Optics f2f: From Fourier to Fresnel* (Adams and Hughes, 2019). Born and Wolf's highly recommended book *Principles of Optics* includes a comprehensive discussion of the elements of diffraction theory (Born and Wolf, 1999). The fourth edition of *Optical Physics* provides an alternative description of the mathematical analysis of the diffraction theory of light (Lipson, Lipson, and Lipson, 2011). Lipson et al. state that in the focal region of a lens the geometrical model of light propagation is not valid; instead, we need to use a complete vector theory of diffraction. This vector theory, as described by Born and Wolf, yields a more complete description of the theory of light diffraction.

2.3 Development of the Two-Point Resolution Concept: Classical Criteria

2.3.1 George Biddell Airy on Resolution: Airy Diffraction Pattern

It should not be surprising that our discussion begins with the work of an astronomer. Major theoretical and practical technological advances began in astronomy and migrated to the field of microscopy. Seminal examples include criteria on resolution, theory of optical aberrations, charge-coupled devices as light detectors, and adaptive optics.

George Biddell Airy held joint positions at the University of Cambridge: he was a professor of astronomy and experimental philosophy as well as director of the Cambridge Observatory. In 1835 Airy was named to the renowned post of Astronomer Royal. In the same year Airy published his seminal paper that contained his analytical formula for the diffraction pattern. The Airy pattern is the eponymous name given to the diffraction pattern of a point light source (a very distant star) in an optical system (with a circular lens) that is devoid of aberrations (Airy, 1835).

The image of a distant star is a central bright circle (Airy disk) with a series of rings of differing light intensities. From his mathematical analysis Airy formed the following conclusions. First, the diameters of the rings depend on the aperture of the telescope and are inversely related to the aperture. Second, he gives an expression for the intensity of light in the various rings, with the intensity at the center of the bright circle being defined as the standard. Further calculation yields the relative intensities of successive bright rings and positions of the nodes, or black rings devoid of light.

I now transition from telescopes to microscopes. The diffraction image is formed in the diffraction plane by the microscope objective. The observed Airy diffraction pattern for a circular aperture is labeled the Airy pattern or the two-dimensional point spread function (PSF). The Airy disk or diffraction pattern consists of a

central peak of light intensity surrounded by weaker intensity rings that are separated by dark rings. Approximately 80% of incident intensity is in the central bright spot. The size of the central bright spot is proportional to the incident wavelength and inversely proportional to the numerical aperture (NA).

It is the Fraunhofer diffraction pattern that is formed by the exit pupil of the microscope objective. The radius of the Airy disk from the central maximum intensity peak to the first minimum is given as

$$r = 0.61\frac{\lambda}{\mathrm{NA}} \tag{2.1}$$

where λ is the vacuum wavelength of the light, and NA is the numerical aperture. The NA of a microscope objective or a microscope condenser was defined by Abbe in 1873, 1881 as $\mathrm{NA} = n\sin\theta$, where n is the refractive index of the medium measured at 587 nm, and θ is half the angular aperture (i.e., the half-angle of incident light rays at the top or front lens of the microscope objective; Abbe, 1873, 1881). Earlier, in 1870 Abbe began his investigations on numerical aperture in the light microscope. The symbol θ is the half-angle of the cone of light converging to an illuminated point or diverging from a point. This quantity can also be defined as the semiangle of the cone of rays from the axial object point that is received by the objective.

The Abbe theory of image formation in a microscope is discussed in Chapter 6. The object in Abbe's microscope was a periodic grating and the illumination was coherent. Abbe in 1873 concluded that the resolution of a microscopic imaging system is given by

$$\text{Resolution} \propto \frac{\lambda}{\mathrm{NA}} \tag{2.2}$$

Abbe investigated and then suggested techniques to enhance the optical resolution of a microscope. For example, he knew that oblique or off-axis illumination resulted in enhanced optical resolution. In addition, he knew that increasing the refractive index of the medium between the object and the objective could enhance the resolution of the microscope. Abbe also investigated how phase masks alter the resolution of an optical microscope. Such investigations with phase masks were published after Abbe died and are based on his lecture notes (Lummer and Reiche, 1910).

2.3.2 *Lord Rayleigh on Resolution*

John William Strutt, also known as Lord Rayleigh, was an English physicist noted for his many scientific achievements including his discovery of Rayleigh scattering (Howard, 1964; Masters, 2009). Rayleigh's life (1842–1919) covered two important yet very different periods in the history of physics: the classical period and

early modern physics, which is based on Einstein's relativity theory and the early development of quantum mechanics (Masters, 2009). In 1904 Rayleigh and William Ramsay shared the Nobel Prize in Physics for their joint discovery of the element argon.

What is perhaps less known is the fascinating connection between Rayleigh's book *The Theory of Sound* and the quantum mechanical perturbation methods of Erwin Schrödinger and other physicists. What is the source of this interesting connection? Rayleigh developed perturbation methods that have applications to acoustics. Later Schrödinger further developed these methods and formulated the Schrödinger–Rayleigh perturbation method.

Rayleigh published several papers on the resolution of optical instruments. He investigated the resolving or separating power of optical instruments (e.g., telescopes in terms of the wave theory of light; Rayleigh, 1880a, b, 1896, 1899). Then his interests turned to microscopes and he investigated the theory of the optical image with special reference to the microscope (Rayleigh, 1896). *The Collected Optics Papers of Lord Rayleigh* was published by the Optical Society of America who provide his corpus in optics in two volumes (Strutt, 1994).

What is the Rayleigh resolution criterion? Rayleigh based his resolution criterion on the function of the human visual system. He posits that a minimal contrast is required to distinguish two incoherent point sources of light that form a joint intensity distribution. The resolution definition that Rayleigh proposed was: "that two points emitting incoherent light of equal intensity are resolved if they are sufficiently separated in space so that the center (maximum intensity) of the Airy disk of a one-point object is situated at a point that corresponds to the first minimum of the diffraction pattern of the second point object" (Rayleigh, 1896). Two point sources of incoherent light will not be imaged by a lens as two distinct points of light; they form the diffraction pattern of the aperture of the lens and this diffraction pattern is the point spread function (PSF). In the absence of geometrical aberrations the point spread function is the Fraunhofer diffraction pattern of the lens aperture. Rayleigh thought of an object as many point sources of incoherent light for which the intensities could be added.

But what if there is no zero of intensity? In such a case Rayleigh stated that resolution is given by the distance for which intensity at the central minimum in the combined image of two equal point sources is 80% of the maximum of intensity on each adjacent side. Conversely, the two points will not be resolved if the plot of intensity shows a dip in the middle that is higher than 80% of the maximum intensity. Rayleigh stated that various aberrations reduce the resolution of an optical instrument. The Rayleigh metric of resolution is inappropriate for the case when there are no zeros of intensity or when they are too distant from the central maximum of the diffraction image. His resolution criterion must not be thought of as a physical law. His own expression sums up his analysis: "This rule [Rayleigh resolution criterion] is convenient on account of its simplicity and it is sufficiently accurate in view of the necessary uncertainty as to what exactly is meant by resolution" (Rayleigh, 1896). In summary, the Rayleigh criterion applies to two light sources of equal intensity. They are just resolved when the first zero of the Airy pattern of one

light source is located on the maximum of the Airy pattern of the second light source. Historically, the Rayleigh two-point resolution criterion was applied to astronomical telescopes. Goodman gives the following definition of the Rayleigh criterion of resolution: "two incoherent point sources are barely resolved by a diffraction-limited stem with a circular pupil when the center of the Airy intensity pattern generated by one point source fall exactly on the first zero of the Airy pattern generated by the second [point source] (Goodman, 2017, pp. 216–217).

2.3.3 Sparrow on Resolution

How does the criterion of Sparrow differ from that of Rayleigh? Sparrow's criterion of resolution does not make any assumptions about the human visual system, and it has some advantages over the Rayleigh criterion (Sparrow, 1916). The Sparrow criterion has the advantage over the Rayleigh criterion in that it is also applicable to coherent imaging (Barakat, 1962). The Sparrow resolution criterion is affected by the coherence properties of incident light.

Sparrow began his analysis on the separation of spectral lines and performed his experimental investigations in the physical laboratory of Johns Hopkins University. The Rayleigh criterion assumes that two points or lines have equal intensity, but Sparrow's experiments show that, while the shape of the intensity curves of two lines of unequal intensities are changed, the limiting resolution is nevertheless constant. I quote Sparrow: "The actual limit of the resolving-power of a perfect grating or prism has been found experimentally. It is found that this limit is given, for equal intensities of the two lines, by the 'undulation condition,' that is, by the condition that the central minimum shall just disappear. That gives a theoretical resolving-power about 26% greater that that obtained by the Rayleigh criterion. The limit given by the undulation condition has been found to hold for unsymmetrical doublets when the ratio of intensities of the two components is less than 10:3" (Sparrow, 1916).

The Sparrow criterion of two-point resolution is the smallest distance between two points at which the minimum in the intensity distribution of the combined two luminous points vanishes. It considers incoherent point sources: two point sources can be resolved if their combined intensity function has a minimum on the line between their centers. But the Sparrow criterion can be generalized to coherent light sources. Goodman gives the following definition: "The Sparrow resolution criterion states that two equally strong incoherent point sources are barely resolved when their separation is the maximum separation for which the image of the pair of points shows no dip at the midpoint" (Goodman, 2017, p. 229).

The Sparrow definition is that two points of equal brightness are imaged as two separate points if the intensity at the midpoint between them is equal to intensity at the points. The Sparrow minimal resolved distance is

$$\Delta x = \frac{0.51\lambda}{\mathrm{NA}} \tag{2.3}$$

The Sparrow criterion can be expressed in another way. Two point sources are just resolved if the second derivative of the resulting image illuminance distribution is zero at the point midway between the respective Gaussian image points. The mathematical expression of the Sparrow criterion is this: if the two points of light have equal intensities, then the Sparrow resolution criterion yields $\theta_{\mathrm{minimum}}$ when

$$\left(\frac{\mathrm{d}^2 I}{\mathrm{d}\theta^2}\right)_{\theta=\theta_{\mathrm{minimum}}/2} = 0 \tag{2.4}$$

Note that θ has units in radians. A mathematical expression of the Sparrow resolution criterion is

$$\theta_{\mathrm{minimum}} = \frac{0.95\lambda}{D} \ \text{(Sparrow)} \tag{2.5}$$

Sparrow's resolution criterion is modified in the case of coherent light to the following expression

$$\theta_{\mathrm{minimum}} = \frac{1.46\lambda}{D} \ \text{(Sparrow)} \tag{2.6}$$

For comparison with the Rayleigh criterion for the minimum angular separation of resolvable incoherent light sources, Rayleigh obtained the following result

$$\theta_{\mathrm{minimum}} = \frac{1.22\lambda}{D} \ \text{(Rayleigh)} \tag{2.7}$$

where the aperture has diameter D, and λ is the wavelength of light. It is assumed that the intensities of the two light sources are equal. An optical system is said to be diffraction limited when it can resolve two points that are separated by the angle $\theta_{\mathrm{minimum}}$. Therefore, two point objects are just resolvable if they are separated by this minimum angle. Note that θ has units in radians. These examples refer to the concept of angular resolution, which occurs in astronomy when we image light from two stars. The starlight from two stars has an angular separation θ (in radians) and the plane waves are imaged in a telescope. In summary, the Sparrow criterion states that two light sources are just resolved if the second derivative of the image intensity is zero at the midpoint between the two Airy patterns (2.4).

Application of both the Rayleigh and Sparrow criteria to resolution in the microscope is confounded by the occurrence of possible interference between two adjacent diffraction images. Other important confounding factors in the measurement of the minimal resolved distance in a light microscope or the angular separation in a telescope include the following: the signal-to-noise ratio of the detected intensity pattern, noise, the shape of the object, the type of illumination, and polarization effects.

2.4 Other Criteria of Resolution

2.4.1 Resolution Based on the Point Spread Function

Rayleigh observed that two point sources of light that are imaged by a lens can be represented by two diffraction patterns or Airy disks (Rayleigh, 1880a, b, 1896, 1899). It follows that the sharper these Airy disks (i.e., the smaller their diameter) the smaller the detectable distance between them and hence the better the two-point resolution. This idea has been generalized in the resolution criterion based on the full width at half maximum (FWHM) of the point spread function (PSF). Experimentally, the PSF is the normalized intensity image of a luminous point source (e.g., a normalized Airy disk); its FWHM is the diameter at one-half of the maximum intensity (giving a measure of the "sharpness"). Depending on the optical system the FWHM has different values in different directions in the object plane (coordinates x, y) and along the optical axis (z). Typically, two object points can be discriminated (resolved) from each other if their distance is larger than the FWHM.

2.4.2 Fourier-Based Resolution Criteria

Ernst Abbe (1873) developed his resolution criterion by assuming that the resolution limit of an object is the finest detail that can be discriminated. In Abbe's experiments that is the finest periodic grating (in terms of lines per millimeter) that can be imaged (i.e., by the smallest grid-to-grid distance that can still be detected by the optical system).

Jean Baptiste Fourier (1768–1830) is credited for his development of techniques to mathematically describe any continuous object as the superposition of harmonic functions. Not all functions can be Fourier transformed. The Dirichlet conditions place constraints on the well-behaved periodic function so that an infinite Fourier series will converge. A well-behaved periodic function has the following properties: the function is single valued over its period; it has a finite number of maximum and minimum values and a finite number of discontinuities; the function has no infinite discontinuities; and the Fourier transform and its inverse requires that a well-behaved function obeys the Dirichlet conditions.

What is the Fourier theorem? Any well-behaved periodic function $f(x)$ can be expressed as the sum of a series of sinusoidal functions (Gaskill, 1978; Goodman, 2015; Gray and Goodman, 1995)

$$f(x) = \frac{1}{2}C_0 + \sum_{n=1}^{\infty} C_n \cos(nk_0 x + \alpha_n) \tag{2.8}$$

The *ns* are called the orders of the terms and are harmonics, and

$$k_0 = \frac{2\pi}{\lambda} \tag{2.9}$$

where k_0 is the fundamental spatial frequency, which is the reciprocal of the period of the pattern in the image and is typically measured in cycles per millimeter. Spatial frequency can be defined as the number of waves per unit length; its units are inverse length. Each term in the above series has two Fourier coefficients: an amplitude C_n and a phase angle α_n. The goal of Fourier analysis is to calculate these two Fourier coefficients: the amplitude and the phase angle. The Fourier transform (FT) converts a distribution as a function of distance $f(x)$ to the same distribution as a function of frequency $F(\omega)$ (Gaskill, 1978; Goodman, 2015; Gray and Goodman, 1995)

$$F(\omega) = \int_{-\infty}^{+\infty} f(x)\exp(-i2\pi\omega x)\mathrm{d}x \tag{2.10}$$

where $F(\omega)$ is the frequency spectrum of the Fourier transform of $f(x)$. The inverse Fourier transform converts a distribution as a function of frequency to the same distribution as a function of distance (Gaskill, 1978; Goodman, 2015; Gray and Goodman, 1995)

$$f(x) = \int_{-\infty}^{+\infty} F(\omega)\exp(i2\pi\omega x)\mathrm{d}\omega \tag{2.11}$$

Convolution, correlation, and autocorrelation are frequently used in Fourier optics. The image in the object plane is convolution of the total object intensity and point spread function of the lens.

What is the mathematical operation of convolution? The convolution of two real functions f and g is defined as

$$h(x) = \int_{-\infty}^{+\infty} f(x')g(x-x')\mathrm{d}x' \tag{2.12}$$

The convolution operation is also written as

$$h(x) = f(x) \otimes g(x) \tag{2.13}$$

where $\otimes$ represents the convolution operation.

The correlation function is mathematically defined as

$$h_{\text{corr}}(x) = \int_{-\infty}^{+\infty} f(x')g^*(x' - x)\mathrm{d}x' \tag{2.14}$$

which is the convolution of $f(x)$ and $g^*(-x)$. The symbol for conjugation is the superscript asterisk. In this case g^* represents the complex conjugate of g. The complex conjugate of a complex number is a number that has an equal real part and an imaginary part that is equal in magnitude but opposite in sign.

The autocorrelation function is similar to the correlation function, but we set

$$h \equiv f, \text{then}$$

$$h_{\text{autocorr}}(x) = \int_{-\infty}^{+\infty} f(x')f^*(x' - x)\mathrm{d}x' \tag{2.15}$$

Therefore, the Fourier transform of the autocorrelation function is the square modulus of the transform of the function, which is also called the power spectrum.

McCutchen (1967) used Fourier analysis to analyze the effect of apertures on an imaging system. McCutchen could not conceive of superresolution microscopy based on transmitted light (McCutchen, 1967). He defined superresolution microscopy as that which achieves a resolution that is beyond that obtained in a diffraction-limited optical system in the absence of aberrations and noise. McCutchen made this claim because of the difficulty of placing an aperture or a hole nearer than 100 wavelengths from the object.

However, McCutchen was the first to conceive development of a reflected light microscope in which the point source of light has a diameter much smaller than the wavelength and is scanned over the object. The image is the convolution of the spatial spectrum of the object with the autocorrelation function of the illuminating pupil. This increases by $2/\lambda$, where λ is the wavelength of the incident light, the largest spatial frequency that forms the image. Such a microscope would be superresolving since there is a gain of more than a factor of 2.

2.4.3 *Nyquist Theorem-Based Resolution Criteria*

The material discussed in this section comes from the field of information theory and analysis of the transmission of signals (Shannon, 1949). In practice, a waveform that varies in time may be sampled (e.g., its voltage measured at particular

times). The original waveform must be reconstructed from a set of samples. The key question is: When is the sampling rate sufficient to exactly reconstruct the sample waveform?

The Nyquist–Shannon sampling theorem can be explained as follows (Goodman, 2015). We assume that a signal contains many frequencies. The definition of a band-limited signal depends on whether the frequencies are limited by a maximum frequency f_{max}. The Nyquist frequency is $2f_{max}$. The Nyquist–Shannon sampling theorem states that if the signal is sampled at a frequency that is greater than the Nyquist frequency, then it will be possible to exactly reconstruct the original signal from the samples. In summary, the sampling interval must be the reciprocal of twice the highest frequency present.

As far as microscope resolution is concerned we make the following assumption: the object can be thought of as composed by the superposition of many functions of different spatial frequencies. By analogy with the concept described in the previous paragraph the microscope must transmit all the spatial frequencies up to the "cutoff" frequency, which will determine the limiting resolution of the microscope.

2.5 Optical Transfer Function and Modulation Transfer Function

Optical transfer function (OTF) is a measure of the spatial frequencies present in an object that are transferred across an optical element or an optical system. The concept and derivation of OTF were first developed by the French physicist Pierre-Michel Duffieux. Duffieux published his concepts in a series of publications that first appeared in 1935, and a decade later he privately published his seminal book *L'intégrale de Fourier et ses applications à l'optique* (1946). In the mid-1980s John Wiley & Sons published an English translation of the second edition of his book with the title *The Fourier Transform and Its Applications to Optics*. He studied how best to reduce the Fourier components that are transmitted from an object to the image in an optical system, and this led him to define an optical system's OTF; he found it depended on both lens aperture and optical aberrations.

Duffieux defined OTF as a ratio: OTF = [Fourier transform of light distribution in the image] ÷ [Fourier transform of light distribution in the object] (Duffieux, 1946).

Duffieux was able to use Fourier analysis to explain Abbe's theory of image formation in the microscope (Duffieux, 1946). Furthermore, Duffieux formulated a theory of image formation that is compatible for both coherent and incoherent illumination in the microscope. His theory of image formation was compatible for all shapes of apertures and various types of aberrations. He greatly simplified calculations by first rejecting the use of the Fraunhofer diffraction integral and replacing it by the convolution theorem through which he demonstrated that the Fourier transform of the function that expressed the light intensity distribution in the

image can be approximated by the product of the Fourier transform of the light distribution in the object and the transform of a point source image. This mathematical analysis required a stringent restriction: the PSF should not vary significantly over the angle of field through which the Fourier components are transmitted to the image (Duffieux, 1946).

Perusal of the modern definition of OTF immediately indicates the seminal antecedent developments of Duffieux. In support of this claim I cite the modern definition of OTF by Williams and Becklund: "The Optical Transfer Function (OTF) is the frequency response, in terms of spatial frequency, of an optical system to sinusoidal distributions of light intensity in the object plane; the OTF is the amplitude and phase in the image plane relative to the amplitude and phase in the object plane as a function of frequency, when the system is assumed to respond linearly and to be space invariant. The OTF depends on and potentially describes the effect of diffraction by the aperture stop and the effects of the various aberrations" (Williams and Becklund, 1989). An imaging system's OTF is the Fourier transform of the PSF (Backer, 1992a, b; Williams and Becklund, 1989). I recommend two books that contain selected publications on the theory, measurement, and applications of OTF (Backer, 1992a, b).

Next, I introduce the modulation transfer function (MTF), sometimes called the contrast transfer function (CTF), which was developed by Harold H. Hopkins in 1962. Harold H. Hopkins (1918–1994) is credited with the development and use of the modulation transfer function in 1962. Hopkins was a student of Duffieux at the University of Besançon. Hopkins did most of his work on Fourier optics at the Imperial College of Science, Technology and Medicine (a.k.a. Imperial College London). Following on the prior works of his mentor Hopkins published his book *Wave Theory of Aberrations* in 1950, which defined different types of aberrations in terms of wavefront distortions (Hopkins, 1950). Hopkins developed the concept of partial coherence in optics (Hopkins, 1951). In his publications Hopkins corrected some invalid approximations made in the previous calculations of van Cittert (1934) and Zernike (1938), and he applied the revised calculations to the theory of Young's experiment, the stellar interferometer, and illumination in the microscope. This led Hopkins to a general theory of the formation of optical images. It should be noted that the earlier works of Ernst Abbe (1840–1905) and Lord Rayleigh (1842–1919) on Fourier optics preceded the works of Hopkins.

First, I define modulation or contrast in an image. Modulation is the ratio of maximum intensity minus minimum intensity divided by the sum of maximum and minimum intensities. Modulation transfer function (MTF) is defined as the ratio of image modulation to object modulation at all spatial frequencies. Alternatively, we can define MTF with incoherent illumination as the modulus as magnitude or amplitude of OTF.

To summarize, the Fourier transform of the point spread function (PSF) yields the optical transfer function (OTF). This function along with the modulation transfer function (MTF) and the phase transfer function (PTF) are related in the following manner

$$\mathrm{FT}(\text{image}) = \mathrm{OTF} \times \mathrm{FT}(\text{object}) \tag{2.16}$$

$$\mathrm{OTF} = \mathrm{MTF}\exp(\mathrm{iPTF}) \tag{2.17}$$

The point spread function (PSF) is the spread of light intensity from an ideal image point. The ratio of peak intensity to peak intensity in the limit of no aberrations is the Strehl ratio. An optical system for which the Strehl ratio is greater than 0.8, which corresponds to a wavefront error of $\lambda/14$ or less, is usually called diffraction limited.

2.6 Concept of Diffraction Limit

The concept of diffraction limit arises when we think of an optical imaging system acting as a filter for spatial frequencies. Because the aperture of an optical system is finite its optical transfer function (OTF) is band limited. As a consequence frequencies above a specific spatial frequency will have an OTF that is zero. Thus, these higher spatial frequencies are not transferred by the optical system. The spatial frequency at which OTF is zero is denoted as the diffraction limit to optical resolution.

First, I describe a linear system such as an optical system with an incoherent light source. Such a linear system is both linear and shift invariant (Gaskill, 1978; Sheppard, 2007; Van Aert et al., 2006, 2007). What are the characteristics of such a linear system? The key property of a linear system is the existence of the principle of superposition; the combined response to a linear combination of inputs is equivalent to a linear combination of responses. A shift-invariant system is one in which a shift in the position of the input results in an equal shift in the position of the output.

The point-spread function (PSF) can be used to characterize an optical imaging system that is both linear and shift invariant. A coherent optical imaging system is linear in complex amplitude. An incoherent optical imaging system is linear in intensity. As a consequence the image that is formed by a linear and shift-invariant optical imaging system will have an amplitude for an example of coherent illumination or an intensity for an example of incoherent imaging that is a convolution of either the amplitude distribution or the intensity distribution and the PSF of the optical imaging system.

2.7 Early Concepts of Superresolution

Ernst Abbe wrote in his 1973 seminal paper that under the specific experimental conditions of his light microscope resolution is limited by the wavelength of illumination and the numerical aperture of the microscope objective (Abbe, 1873). In his analysis the object was a periodic grating and the optics were assumed to be devoid of aberrations. Abbe derived an equation for the limiting resolution of a light microscope using oblique illumination

$$d = \frac{\lambda}{2n \sin \alpha} \tag{2.18}$$

where d is the minimum separation between two point sources of light that can be separated or resolved as two separate objects, λ is the wavelength of the illumination, and $n \sin \alpha$ is the numerical aperture of the objective lens of the microscope.

Abbe also suggested and demonstrated that resolution could be enhanced under his experimental conditions by two methods. First, a shorter wavelength of illumination would increase resolution. The ultraviolet microscope developed by August Köhler in 1904 is an application of the use of short-wavelength illumination to enhance resolution. Second, a higher numerical aperture would increase resolution. The 4Pi microscope based on two identical high-NA opposing microscope objectives is an application in which use of a higher numerical aperture increases resolution. But these enhancements did not yield a resolution beyond that for a diffraction-limited optical system. In modern terms this limit is related to a cutoff frequency, which is the highest spatial frequency that can be transferred by the optical imaging system.

Part III of this book describes and compares techniques that achieve superresolution in far-field optical microscopy. What do we mean by the term superresolution? It relates to reconstructing frequency components that are located beyond the cutoff frequency or the diffraction limit of the optical imaging system.

Ever since Abbe's 1973 publication people have dreamed of methods to achieve superresolution in microscopes. Toraldo di Francia is credited with introducing the concept of superresolution of images. In his 1952 publication in Italian Toraldo di Francia defined superresolution as detail finer than the Abbe resolution limit (Toraldo di Francia, 1952).

Toraldo di Francia made two important and prescient developments. First, he used pupil plane filters to achieve superresolution. He placed two concentric amplitude and/or phase filters in the pupil of an imaging system. But superresolution was only achieved in the central region of the PSF, and the peripheral regions of the field had decreased resolution. But as the resolution of the central lobe increased there was decreased intensity in that part of the PSF. Second, he showed the existence of evanescent waves. Today modern total internal reflection fluorescence (TIRF) microscopy is based on evanescent exciting light that only causes an extremely thin region of a cell surface to fluoresce because of its exponentially

decreasing intensity at a surface (Toraldo di Francia, 1952). Unfortunately, to date the method proposed by Toraldo di Francia has not been implemented in any practical device.

In 1967 Charles W. McCutchen suggested that it should be possible to build a superresolution optical microscope that can resolve details smaller than the diffraction limit (McCutchen, 1967). Sometimes innovation in instrument development results from someone posing critical questions. McCutchen posed the question: "Can superresolution really beat the ultimate Abbe resolution limit for a lens with an acceptance solid angle of 2π steradians?" (McCutchen, 1967).

Furthermore, his question likely stimulated the thinking of the inventors of 4Pi optical microscopes that use two opposing microscope objectives with the specimen between them (Hell, 1990; Hell and Stelzer, 1992a, b; Hell et al., 1994a, b; Hänninen et al., 1995). 4Pi microscopes have enhanced lateral resolution and significantly enhanced axial resolution, but they do not yield superresolution in both lateral and axial resolution.

Lukosz published two prescient papers that present a comprehensive overview of superresolving optical systems (Lukosz, 1966, 1967). He made two assumptions that were fundamental to his analysis: linearity and spatial invariance. I explained these two terms in the previous section. Lukosz modifies Abbe's definition of the diffraction limit of a coherent optical system. Instead of the usual metric—the bandwidth of the spatial frequencies that the optical system can transfer—he introduces a new metric: the number N of the degrees of freedom of the transmitted optical signal, which he believed to be a constant. He defines the number N as the product of the object area $\times$ the optical bandwidth $\times$ 2 (the number 2 comes from the number of independent states of light polarization) $\times$ the temporal degrees of freedom (Lukosz, 1966, 1967).

As far as superresolution optical imaging systems are concerned Lukosz invoked the spatial invariance theorem to show that the spatial bandwidth of a system is not a constant and that it can be enhanced above the classical limit by reducing the number of degrees of freedom of the information that the system can transmit (Lukosz, 1966). Lukosz provides several suggestions to reduce the number of degrees of freedom and therefore achieve superresolution: reduce the object area, enhance the bandwidth in the x-direction and decrease it in the y-direction keeping the two-dimensional bandwidth constant, and use only one state of polarization.

How does Lukosz implement his ideas in an optical imaging system? He modifies the optical system by inserting high-frequency gratings into the optically conjugate planes of object and image space. The first grating formed many copies that are shifted out of the object's diffraction pattern, and the role of the second grating was to demodulate or to separate these different portions of the object's diffraction pattern (Goodman, 2017). While this method increased the NA of the optical system it also resulted in ghost images that are distant from the main image. Lukosz mitigated the presence of these ghost images by reducing the field of view (Goodman, 2017). In general, his analysis is based on coherent illumination. But his superresolution optical imaging systems also work with partially coherent or even incoherent light.

2.8 Again, What Is Resolution?

A recent comprehensive book *The Limits of Resolution* provides a very good discussion of resolution and its limits (de Villiers and Pike, 2016).

While various concepts of resolution criterion are qualitatively useful, their quantitative implementation is difficult. When microscope images need to be interpreted quantitatively a new metric is required. From the field of electron microscopy a new criterion has evolved that is based on the relation between resolution and statistical measurement precision (van Aert et al., 2006, 2007). His quantitative resolution criterion can be applied to both coherent and incoherent optical imaging systems. His criterion represents resolution as a precise estimate of the separation between adjacent points in the object. Furthermore, it is based on precise measurements of physical parameters. Therefore, it eliminates arbitrary thresholds and assumptions about the human visual system that are fundamental to the Rayleigh resolution criterion.

On the other hand, there is much to be learned from perusal of the literature on optical lithography, especially when it comes to understanding resolution and the development of innovative new types of microscopes. Optical lithography is a photon-based technique in which an image is projected onto a photosensitive emulsion that is coated onto a substrate. Optical lithography is used for the manufacture of nanoelectronics by the semiconductor industry.

Modern resolution enhancement techniques enable optical lithography well below the wavelength of illumination light (Al-Amri, Liao, and Zubairy, 2012; Schellenberg, 2005). What factors underpin modern optical lithography and its enhanced resolution? The main advance is the ability to manipulate the wavefront of light that is projected onto the photosensitive coating on the substrate. The amplitude, phase, and the direction of illumination are all manipulated by using phase-shifting masks and off-axis illumination.

Historically, many discoveries in microscopy are made, lost, and then rediscovered. I give two examples in support of this. First, Abbe in his seminal paper of 1873 observed that phase-shifting masks in the optical system can improve resolution. Second, Rayleigh observed in 1896 that a phase object showed enhanced optical resolution in the microscope.

In summary, we see that resolution depends on a variety of factors: the coherence properties of illumination, the optical system, the geometric characteristics of the object, prior knowledge of the object, optical aberrations of the optical system, and noise. I highly recommend that the reader explore in depth the following selected references. First, a solid foundation of image science is invaluable to stimulating innovation in microscopy (Barrett and Myers, 2004). The foundations of digital signal processing and digital image processing are critical aspects of a deeper understanding of modern microscopy (Gonzalez and Woods, 2008; Madisetti and Williams, 1998). The field of X-ray crystallography provides many interesting insights into diffraction (Als-Nielsen and McMorrow, 2001; Giacovazzo et al., 2011). A critical study of electron microscopy will yield benefits for those who

work in optical microscopy. In this genre I would recommend beginning with the book *Light and Electron Microscopy* (Slayter and Slayter, 1992). I found the second edition of *Transmission Electron Microscopy, A Textbook for Materials Science* exceedingly useful; of particular merit are the chapters on diffraction and a chapter entitled "Thinking in Reciprocal Space" (Williams and Carter, 2009). The field of astronomy is replete with insights for the innovative inventor of new types of microscopes. Another stimulating source is the book *Astronomical Methods: A Physical Approach to Astronomical Observations*, which discusses topics such as speckle interferometry and adaptive optics as useful techniques for resolution enhancement (Bradt, 2004). The major advances in resolution enhancement techniques that have taken place in the field of optical lithography are very important sources of information for innovators in the field of microscopy (Al-Amri, Liao, and Zubairy, 2012; Schellenberg, 2004).

The path to true innovation and invention in microscopy is to approach the field from the viewpoint of disparate fields and related disciplines. This can be done by familiarizing yourself with the new innovations in quantum optics, astronomy, electron microscopy, optical lithography, and X-ray crystallography and then thinking across boundaries and disciplines. The rewards of innovation are great.

References

Adams, C. S., and Hughes, I. G. (2019). *Optics f2f: From Fourier to Fresnel.* Oxford: Oxford University Press.

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. M. Schultze's Archive für Mikroskopische Anatomie, IX, 413–468.

Abbe, E. (1881). On the Estimation of Aperture in the Microscope. Journal of the Royal Microscopical Society, **1**, 388–423.

Airy, G. B. (1835). Diffraction of an object-glass with circular aperture. Trans. Cambridge Philosophical Society, **5**, 283–291.

Al-Amri, M., Liao, Z., and Zubairy, M. S. (2012). Beyond the Rayleigh Limit in Optical Lithography. In: *Advances in Atomic, Molecular, and Optical Physics*, **61**, Amsterdam, Elsevier Inc., Chapter 8, pp. 409–466.

Als-Nielsen, J., and McMorrow, D. (2001). *Elements of Modern X-Ray Physics*. Chichester: John Wiley & Sons Ltd.

Baker, L., Editor (1992a). *Selected Papers on Optical Transfer Function: Foundation and Theory*. SPIE Milestone Series, MS **59**. Bellingham: SPIE Optical Engineering Press.

Baker, L., Editor (1992b). *Selected Papers on Optical Transfer Function: Measurement.* SPIE Milestone Series, MS **60**. Bellingham: SPIE Optical Engineering Press.

Barakat, R. (1962). Application of apodization to increase two-point resolution by the Sparrow criterion: I. Coherent illumination. J. Opt. Soc. Am., **52**, 276–283.

Barrett, H. H., and Myers, K. J. (2004). *Foundations of Image Science*. Hoboken: John Wiley & Sons, Inc.

Born, M., and Wolf, E. (1999). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Seventh edition (expanded). Cambridge: Cambridge University Press.

Bradt, H. (2004). *Astronomy Methods: A Physical Approach to Astronomical Observations*. Cambridge: Cambridge University Press.

Cremer, C., and Masters, B. R. (2013). Resolution enhancement techniques in microscopy. European Physical Journal, **38**, 281–344. Open access publication.

de Villiers, G., and Pike, E. R. (2016). *The Limits of Resolution*. Boca Raton, Florida: CRC Press.

Duffieux, P. M. (1946). *L'intégrale de Fourier et ses applications à l'optique*. Paris: Masson, Editeur. Republished in English as: Duffieux, P. M. (1970). *The Fourier Transform and Its Applications to Optics*, Second edition. New York: John Wiley and Sons.

Gaskill, J. D. (1978). *Linear Systems, Fourier Transforms, and Optics*. New York: Wiley-Interscience.

Giacovazzo, C., Monaco, H. L., Artioli, G., Viterbo, D., Milanesio, M., Ferraris, G., Gilli, G., Gilli, P., Zanotti, G., and Catti, M. (2011). *Fundamentals of Crystallography*. Third edition. Oxford: Oxford University Press.

Gonzalez, R. C., and Woods, R. E. (2008). *Digital Imaging Processing*. Third edition. Upper Saddle River: Pearson Prentice Hall.

Goodman, J. W. (2015). *Statistical Optics*. Second edition. Hoboken: John Wiles & Sons, Inc., pp. 288–289.

Goodman, J. W. (2017). *Introduction to Fourier Optics*. Fourth Edition. New York: W. H. Freeman and Company.

Gray, R. M., and Goodman, J. W. (1995). *Fourier transforms: An Introduction for Engineers*. Boston: Kluwer Academic Publishers.

Gustafsson, M. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of Microscopy, **198**, 82–87.

Hänninen, P. E., Hell, S. W., Salo, J., Soini, E., and Cremer, C. (1995). Two-photon excitation 4Pi confocal microscope: Enhanced axial resolution microscope for biological research. Applied Physics Letters, **66**, 1698–1700.

Hell, S. W., Lindek, S., Cremer, C., and Stelzer, E. H. K. (1994a). Measurement of the 4Pi-confocal point spread function proves 75 nm axial resolution. Applied Physics Letters, **64**, 1335–1337.

Hell, S. W., Stelzer, E. H. K., Lindek, S., and Cremer, C., (1994b). Confocal microscopy with an increased detection aperture: type-B 4Pi confocal microscopy. Optics Letters, **19**, 222–224.

Hell, S., and Stelzer, E. H. K. (1992a). Properties of a 4Pi confocal fluorescence microscope. Journal of the Optical Society of America A, **9**, 2159–2166.

Hell, S., and Stelzer, E. H. K. (1992b). Fundamental improvement of a 4Pi confocal fluorescence microscope using two-photon excitation. Optics Communications, **93**, 277–282.

Hell, S. (1990). Double confocal microscope. European Patent 0491289.

Hopkins, H. H. (1950). *Wave Theory of Aberrations*. Oxford: Oxford University Press.

Hopkins, H. H. (1951). The concept of partial coherence in optics. Proceedings of the Royal Society of London, Series A, **208**, 263–277.

Howard, J. N. (1964). John William Strutt, Third Baron Rayleigh. Applied Optics, **3**, 1091–1101.

Lipson, A, Lipson, S. G., and Lipson, H. (2011). *Optical Physics*. Fourth Edition: Cambridge University Press.

Lukosz, W. (1966). Optical systems with resolving powers exceeding the classical limit, Part 1. Journal of the Optical Society of America A **56**, 1463–1472.

Lukosz, W. (1967). Optical systems with resolving powers exceeding the classical limit, Part 2. Journal of the Optical Society of America A, **57**, 932–941.

Lummer, O., and Reiche, F. (1910). *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe*. Braunschweig: Friedrich View und Sohn.

Madisetti, V. K., and Williams, D. B. (1998). *The Digital Signal Processing Handbook*. Boca Raton: CRC Press.

Masters, B. R. (1996). *Selected Papers on Confocal Microscopy*. SPIE Milestone Series, MS **131**. Bellingham: SPIE Optical Engineering Press.

Masters, B. R. (2009). Lord Rayleigh: A scientific life. Optics and Photonics News, June, 32–41.

McCutchen, C. W. (1967). Superresolution in microscopy and the Abbe resolution limit. Journal of the Optical Society of America A, **57**, 1190–1192.

Pendry, J. B. (2000). Negative refraction makes perfect lens. Physical Review Letters, **85**, 3966–3969.

Rayleigh, L. (1880a). Investigations in optics, with special reference to the spectroscope. 1: Resolving or separating, power of optical instruments. London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **VIII XXXI**, 261–274.

Rayleigh, L. (1880b). On the resolving-power of telescopes. London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **X**, 116–119.

Rayleigh, L. (1896). On the theory of optical images, with special reference to the microscope. London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **42**, Part XV, 167–195.

Rayleigh, L. (1899). Investigations in optics, with special reference to the spectroscope. In Scientific Papers by John William Strutt, Baron Rayleigh, Cambridge, Cambridge University Press, Vol. **1**, pp. 415–459.

Shannon, C. E. (1949). Communications in the presence of noise. Proceedings of the Institute of Radio Engineers, **37**, 10–21.

Schellenberg, F. M. (2004). *Selected Papers on Resolution Enhancement Techniques in Optical Lithography*. SPIE Milestone Series, MS **178**. Bellingham: SPIE Press.

Schellenberg, F. M. (2005). A history of resolution enhancement technology. Optical Review, **12**, 83–89.

Sheppard, C. J. R. (2007). Fundamentals of superresolution. Micron, **38**, 165–169.

Slayter, E. M., and Slayter, H. S. (1992). *Light and Electron Microscopy*. New York, Cambridge University Press.

Sparrow, C. M. (1916). On spectroscopic resolving power. Astrophysical Journal, **44**, 76–86

Strutt, J. W. (1994). *The collected optics papers of Lord Rayleigh. Part A, 1869–1892; Part B, 1892–1919*. Washington, D.C.: Optical Society of America.

Toraldo di Francia, G. (1952). Super gain antennas and optical resolving power. Il Nuovo Cimento Suppl., **9**, 426–438.

Toraldo di Francia, G. (1955). Resolving power and information. Journal of the Optical Society of America A, **45**, 497–501.

Van den Bos, A., and den Dekker, A. J. (1996). Coherent model-based optical resolution. Journal of the Optical Society of America A, **13**, 1667–1669.

Van Aert, S., Van Dyck, D., and den Dekker, A. J. (2006). Resolution of coherent and incoherent imaging systems reconsidered – Classical criteria and a statistical alternative. Optics Express, **14**, 3830–3839.

Van Aert, S., den Dekker, A. J., Van Dyck, D., and Van den Bos, A. (2007). The Notion of Resolution, in: *Science of Microscopy, Vol. II*, edited by Peter W. Hawkes, John C. H. Spence, New York: Springer.

van Cittert, P. H. (1934). Die wahrscheinliche Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse beleuchteten Ebene. Physica. **1**: 201–210.

Williams, C. S., and Becklund, O. A. (1989). *Introduction to the Optical Transfer Function.* New York: Wiley-Interscience.

Williams, D. B., and Carter, C. B. (2009). *Transmission Electron Microscopy, A Textbook for Materials Science*, Second Edition. New York: Springer.

Zernike, F. (1938). The concept of degree of coherence and its application to optical problems. Physica, **5**, 785–795.

Further Reading

Bracewell, R. N. (2000). *The Fourier transform and its applications*. Third Edition, New York: McGraw-Hill.

De Graef, M., and McHenry, M. E. (2012). *Structure of Material: An Introduction to Crystallography, Diffraction, and Symmetry*. Second edition. Cambridge: Cambridge University Press.

den Dekker, A. J., and van den Bos, A. (1997). Resolution: A survey. Journal of the Optical Society of America A, **14**, 547–555.

Fellgett, P. B., and Linfoot, E. H. (1955). On the assessment of optical images. Philosophical Transactions of the Royal Society A, **247**, 369–407.

Gabor, D. (1946). Theory of communication. Journal of IEEE, **93**, 429–441.

Grimes, D. N., and Thompson, B. J. (1967). Two-point resolution with partially coherent light. Journal of the Optical Society of America A, **57**, 1330–1334.

Hopkins, H. H. (1953). On the diffraction theory of optical images. Proceedings of the Physical Society, Series A, **217**, 408–432.

Hopkins, H. H. (1955). The frequency response of a defocused optical system. Proceedings of the Physical Society, **231**, 91–103.

Hopkins, H. H., and Barham, P. M. (1950). The influence of the condenser on microscopic resolution. Proceedings of the Physical Society, Section B, **63**, Part 10, No. 370B, 737–744.

Linfoot, E. H. (1964). *Fourier Methods in Optical Image Evaluation*. London: Focal Press.

Lukosz, W., and Marchand, M. (1963). Optischen Abbildung Unter Überschreitung der Beugungsbedingten Auflösungsgrenze. Optica Acta, **10**, 241–255.

Mahajan, V. N. (1998). *Optical Imaging and Aberrations. Part I: Ray Geometrical Optics*. Bellingham: SPIE Press.

Mahajan, V. N. (2001). *Optical Imaging and Aberrations. Part II: Wave Diffraction Optics*. Bellingham: SPIE Press.

Mandel, L., and Wolf, E. (1990). *Selected Papers on Coherence and Fluctuations of Light* (1850–1966). SPIE Milestone Series, MS **19**, Part I, Part II. Bellingham: SPIE Optical Engineering Press.

Synge, E. H. (1928). A suggested method for extending microscopic resolution into the ultra-microscopic region. Philosophical Magazine, **6**, 356–362.

Sheppard, C. J. R. (1986). The spatial frequency cut-off in three-dimensional imaging. Optik, **72**, 131–133.

Chapter 3
Aberrations and Artifacts Confound Optical Resolution

> "Beware of determining and declaring your opinion suddenly on any object for imagination often gets the start of judgment, and makes people believe they see things; which better observations will convince them could not possibly be seen: therefore, assert nothing till after repeated experiments and examinations in all lights and in all positions. Pass no judgment upon things over-extended by force, or contracted by dryness, or in any manner out of their natural state, without making suitable allowances."
>
> —Baker, 1769, *The Microscope Made Easy*, Chap. 15

3.1 What You See May Not Be What It Is

In this chapter I encourage critical thinking, a skeptical approach to data, and continuous questioning, which are the key characteristics of a scientist. From the early history of microscopes people questioned the validity of observations and warned about artifacts (Baker, 1769). Are the extremely small objects that are observed with a microscope artifacts? There are many possible explanations. First, could the observations be artifacts due to the instruments? Second, could the observations be due to visual artifacts or defects in the human eye? Third, could the observations be due to alterations in the specimen either during its preparation or during the time of microscopic observation. The critical problem is how to interpret disparate observations. Different observers formulate different interpretations to explain what they observed in new microscopes. Similar questions about artifacts were often raised with the advent of new types of telescopes and confounded the interpretation of observations.

The problems of image artifacts and image fidelity have existed ever since the beginnings of microscopy. If we think of image fidelity as the mapping of an object

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_3

into the image, then a perfect mapping would have the highest image fidelity. If the optical system degrades this perfect mapping, then image fidelity is reduced.

The confounding problems of microscopy are not only limited to the optical system of the microscope. The preparation and the state of the specimen is integral to any consideration of artifacts in microscopy. How was the specimen prepared prior to its observation in the microscope? Was the specimen living? Was it fixed and sectioned? Was the contrast enhanced by optical means or with the use of stains, dyes, and other chemical treatments? If the specimen was living, the question arises as to whether there were changes in the specimen during the observation period.

The use of probes to enhance the contrast of cellular imaging is an active area of research prompting new questions about artifacts. What are the effects of probes on cells? What are the effects of illumination used in the microscope on cells? How can we determine the qualitative and quantitative nature of these putative effects? Sometimes authors state that they did not observe any morphological changes during microscopic imaging, and therefore they conclude that there are no functional changes in the cells. That is a false assertion.

Interpretations of the observations and images are often ambiguous (Baker, 1769). But there are techniques to mitigate artifacts, the problems of specimen preparation, the role of high-intensity illumination, and the difficult problem of image interpretation. One useful technique is to compare images of the same specimen obtained using different microscopic techniques. For example, specimens can be imaged using electron microscopy, X-ray microscopy, and optical microscopy and then compared. Alternatively, we can compare images from different microscopic techniques based on different methods to enhance contrast. For example, we can use reflected light microscopy and fluorescence microscopy to image the same specimens.

Before I proceed to the sections on artifacts and constraints in microscopy, I propose an approach that can be used when thinking about specimens, imaging techniques, and image fidelity. While others may develop different approaches to microscopy, what I want to do in this chapter is share the following approach with the reader.

Prior to any imaging it is necessary to give a lot of thought to the following questions. First, what questions regarding the specimen are we seeking to answer? Other questions may be more specific to the specimen. Is the specimen living or fixed? Is imaging to be performed over large areas (i.e., in vivo skin imaging to detect cancer)? What specimen preparation is required to acquire the necessary contrast for imaging to be performed? What are the effects of stains, dyes, and probes used to enhance the contrast of the specimen? Does treatment of the specimen affect how the specimen functions?

Further important questions concern modification of the optical properties of the specimen prior to imaging. Is the specimen highly scattering, highly absorbing, or thick? What is the thickness of the specimen that must be imaged to answer this question? If the specimen is highly scattering, then the depth of penetration of illumination in the microscope may be very limited. There are a variety of "clearing techniques" that can alter the refractive index of a specimen and result in a greater depth of penetration of illumination in the microscope. In general, a clearing

technique places the specimen in a variety of organic solvents or gels that remove scattering centers from the specimen and match the refractive index of the "cleared" specimen to the immersion fluid used in the light microscope. The goal of clearing techniques is to make the specimen transparent. But very few publications on clearing techniques have carefully validated alterations in the specimen induced during clearing techniques.

If the specimen is heterogeneous, what sampling techniques are required? What can stereology do to help answer this and other questions? For example, if we want to know more about the pathology of liver cells, what can sampling theory tell us about obtaining representative samples of the liver (biopsy samples)?

Let us now turn to finding answers to a number of important questions. To do this what microscopic techniques are best suited to answering such questions? What are the limits and difficulties of these microscopic techniques? What resolution and what contrast is required to answer these questions? Another critical consideration concerns the dynamics of the process that we wish to investigate with the microscope. Does the process occur in nanoseconds, milliseconds, seconds, or minutes? Is the data acquisition time of a given imaging technique sufficiently fast that we can study the dynamics of cellular or neuronal processes?

What exactly do we mean when we use the word quantitative with respect to microscopy? If we measure any property of a specimen (i.e., we assign a number to some physical property), then the measurement is considered to be quantitative. The use of the word quantitative in no way denotes precision or accuracy. Microscopic techniques that provide measurements contain either tacit or explicit assumptions. It is critical that the user and the person evaluating the experimental results have a detailed understanding of all the assumptions involved in producing numerical values. In addition, the effects of sampling theory and the dynamic range of detectors must be clearly understood before the validity of any aspect of quantitative microscopy can be interpreted.

3.2 Aberrations in Microscopy

What is it that confounds optical resolution in a microscope? This is a legitimate question since resolution can depend on several factors: the coherence properties of illumination; the type of illumination; the optical instrument; the properties of the object, point, line, and plane; prior knowledge of the object; aberrations of the optical system; optical properties of the source radiation field; and noise. The concept of resolution is ambiguous because different authors interpret resolution in different ways (see Chapter 2).

The resolution of an optical system that has no aberrations is altered by optical aberrations in the imaging system (Mahajan, 1998, 2001). Karl Strehl (1864–1940) showed that aberrations in the optical system could alter the Airy disk by reducing intensity at the maximum of the diffraction pattern, and this intensity is shifted into the edges of the diffraction pattern (Strehl, 1894).

Optical aberrations degrade an optical system's resolution. The list of optical aberrations includes chromatic aberrations, defocus, spherical aberration, coma, astigmatism, field curvature, and distortion (Born and Wolf, 1999; Mahajan, 2001; Sasián, 2013). I highly recommend the book *Introduction to Aberrations in Optical Imaging Systems* (Sasián, 2013).

In this chapter I want to clarify exactly what is meant by aberrations. First, recall that an ideal lens converts a plane wavefront into a spherical wavefront, which converges at a single point. But due to diffraction the resulting focal spot exists in three dimensions. This case is termed diffraction-limited imaging. I define an optical aberration as the phase difference between a wavefront from a diffraction-limited optical system and that of an aberrated optical system.

Consider the following ideal on-axis optical system that consists of a point source of illumination, an ideal lens, and a point image. By definition, light rays are normal to the geometrical wavefront. The wavefront is a surface of equal optical path length (OPL) that is measured along the rays and from the point source of illumination. The optical path length is defined by

$$\mathrm{OPL} = \int_a^b n(s)\,\mathrm{d}s \tag{3.1}$$

where n is the index of refraction, and ds is the element of arc length. In the ideal case a wavefront propagates from an ideal source, through the lens or optical system, and converges from a spherical wavefront to a point. But in actual optical systems the converging wavefront is not spherical but contains deformations.

Historically, aberrations had been observed and described at an early stage in the development of telescopes and microscopes (Sasián, 2013). Spherical aberration is defined as an aberration that occurs with lenses or mirrors that have spherical surfaces. Spherical aberration was observed by Roger Bacon in the 13th century, described by Kepler in 1611, and later by Descartes in 1637. Chromatic aberration is caused by the dispersion of light and is frequency dependent. Astigmatism was observed by both Young and Airy. In 1830 Lister described the phenomenon of coma in images. The seminal mathematical analysis of the wave theory of optical aberrations by Harold H. Hopkins are developed in his book *The Wave Theory of Aberrations* (Hopkins, 1950).

Let us now turn to the causes of optical aberrations. There are several causes including imperfections in the design and manufacture of optical elements, roughness of the surfaces of optical parts, poor alignment of optical elements in the microscope, and optical aberrations induced by the specimen itself. Specimens can induce optical aberrations as a result of either a refractive index mismatch between the specimen and the immersion medium or variations in the specimen's refractive index over the field. Optical aberrations represent the failure of an optical system to produce a perfect image. They are deviations caused by properties of the lens materials or geometric forms of refracting and reflecting surfaces.

Fortunately, modern microscope objectives are specially manufactured to minimize five categories of optical aberrations: spherical aberrations, coma, astigmatism, field curvature, and distortion (Masters, 2006). Such monochromatic optical aberrations are called Seidel aberrations in honor of Ludwig von Seidel who classified them. Aberrations must be corrected in listed order (i.e., to correct for astigmatism it is first necessary to eliminate spherical aberrations and coma).

Spherical aberration results in lack of a sharp focus point; instead, there is a zone of confusion or caustic. It is caused by a lens that has spherical surfaces whose peripheral regions refract light more than central regions. The optimal correction for spherical aberration of a microscope objective requires a defined object and image distance. This explains why the results of high-NA, oil immersion objectives used with a coverslip to image thick specimens are severely limited by spherical aberrations being generated at increasing distances below the coverslip. Other sources of spherical aberration are mismatch of tube length and objective, nonstandard thickness of coverslips, and poor-quality immersion oil.

Coma is a lens aberration that occurs when light is focused at points off the optical axis. The optical axis is perpendicular to the plane of the lens and passes through the center of a circular lens. The name derives from the Latin word for comet and is so-called because the aberrated image of a point looks like a comet.

Astigmatism is a blur of object point when light rays in the tangential plane and the sagittal plane focus in different planes that are perpendicular to the optical axis, and should only be corrected after spherical aberration and coma are corrected. The Seidel aberration of astigmatism is not equivalent to astigmatism as applied to human vision. In the human eye the nonspherical shape of the lens results in different foci for different meridional planes. In contrast, Seidel astigmatism can occur with perfectly spherical lens surfaces. It is first necessary to define two planes in the optical system. The tangential plane contains both an optical axis and an object point. The sagittal plane is perpendicular to the tangential plane and contains an object point. Because the tangential and the sagittal images are not coincident, the image of a point is two lines, and in between is an elliptical or a circular blur.

Field curvature is another aberration that persists even after spherical aberration, coma, and astigmatism are corrected. Using a lens subject to field curvature, object points that are in a plane will be imaged onto a paraboloidal surface. Field curvature makes a flat field appear curved and various regions of the image end up being blurred. When imaging with a high-aperture microscope objective subject to field curvature either the center or the periphery of the field of view is sharply focused.

Distortion is displacement of the entire image rather than blurring of individual points that form the image. Distortion occurs when the lens magnification varies from the center to the periphery. Distortion can occur as either pincushion or barrel distortion.

In addition to Seidel aberrations, corrections must be made for axial and lateral chromatic aberrations that cause the focus position to depend on the wavelength of the illumination light. Spherical and chromatic aberrations affect the entire field; in contrast, other types of aberrations are only important for off-axis image points.

Axial chromatic aberration occurs when different light wavelengths are not focused at a single point on the optical axis. Each color of light will focus at a different point on the optical axis. The image is surrounded by fringes of different colors that change with varying focus. A concave lens of glass of one refractive index can be joined to a second convex lens of a different refractive index to form an achromatic lens in which several wavelengths focus at the same point on the optical axis. A refractive index is defined as the ratio of the speed of light (phase velocity) in a vacuum to that in a given medium.

Lateral chromatic aberration occurs when different wavelengths are magnified at different ratios. This effect is greatest outside the visual field of the object where the light rays are more oblique. Each object is surrounded by a colored fringe. This effect can be compensated by eyepiece design and the microscope objective (in older microscopes) or by the objective alone (in modern microscopes).

3.3 Artifacts in Microscopy

The causes of artifacts include specimen alteration and damage, inadequate image validation, and inadequate image interpretation (Baker, 1769). Genetically expressed fluorescent proteins serve as one prime example. Seminal advances in microscopy were driven in many cases by advances in the synthesis and use of fluorescent probes. For example, the use of genetically expressed fluorescent proteins provides the researcher with a contrast mechanism of very high specificity. But these probes are typically overexpressed in the cell. How this overexpression affects cell function is a matter of continuing research.

Again, the history of microscopy is replete with warnings about image artifacts and false image interpretations. For the modern user of the microscope the same warnings still apply: so user beware!

References

Baker, H. (1769). *The Microscope Made Easy*. A facsimile edition is published in 1987, Lincolnwood, Illinois: Science Heritage Ltd.

Born, M., and Wolf, E. (1999). *Principles of Optics, Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th edition (Expanded). Cambridge: Cambridge University Press.

Hopkins, H. H. (1950). *Wave Theory of Aberrations*. New York: Oxford University Press.

Mahajan, V. N. (1998). *Optical Imaging and Aberrations, Part I, Ray Geometrical Optics.* Bellingham: SPIE Press.

Mahajan, V. N. (2001). *Optical Imaging and Aberrations, Part II, Wave Diffraction Optics*. Bellingham: SPIE Press.

Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press.

Sasián, J. (2013). *Introduction to Aberrations in Optical Imaging Systems*. Cambridge: Cambridge University Press.
Strehl, K. (1894). *Die Theorie des Fernrohrs auf Grund der Beugung des Lichts*. Leipzig: J. A. Barth.

Further Reading

Gu, M. (2000). *Advanced Optical Imaging Theory*. Berlin: Springer.

Chapter 4
Insights into the Development of Light Microscopes

4.1 Introduction

The purpose of this chapter is to highlight some critical insights that I have obtained from my decades-long perusal of English and German literature on the history of optical microscopes and my visits to American and European museums that contain extensive collections of microscopes.

I begin with sources that consist of both publications and material objects—the microscopes in collections located around the world. The published literature on the history of the microscope is vast and consists of books, journal publications, and letters. These materials are located in libraries, archives, and museums, as well as on the internet. There are journals that publish articles on microscopy and others that publish scholarly articles on the history of science. Furthermore, important historical books that have been out of print for many years are now available as reprinted facsimile editions. I am fortunate to possess the nine large volumes of Schmitz's *Handbuch zur Geschichte der Optik*, which in my opinion is the most comprehensive work on the history of optics and optical instruments—specifically microscopes and telescopes. Schmitz's volumes are written in German. They contain high-quality illustrations of the different optical instruments. Why is it important to study the joint development of microscopes and telescopes? Many technologies and theoretical developments had their origins in telescopes and were later transferred to microscopes (Schmitz, 1981a, b, 1982, 1983a, b, 1984a, b, 1989, 1990). Some of these developments include the early concepts of resolution, the description and analysis of optical aberrations, and the development of various light detectors such as photographic plates, photomultiplier tubes, and charge-coupled devices (CCDs).

But this analogy also has its limitations and can lead to erroneous conclusions and assumptions. There are important differences that must be considered. The light from stars is incoherent; the light from an object in a microscope is coherent or partially coherent and can exhibit interference. Off-axis optical aberrations are

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_4

exceedingly small in telescopes. But off-axis optical aberrations are large in light microscopes.

Ernst Abbe is credited with being the first to understand that image formation in the telescope and the microscope cannot be formulated by the same aspects of geometric optics and geometric ray tracing (Volkmann, 1966). Something new was required and that was Abbe's 1873 seminal theory of image formation in the light microscope due to diffraction (Abbe, 1873).

Of the many books on the history and development of the microscope I recommend several books that are published in English. John Quekett, who was the assistant conservator of the museum and demonstrator of minute anatomy at the Royal College of Surgeons of England in 1848, published *A practical treatise on the use of the microscope, including the different methods of preparing and examining animal, vegetable, and mineral structures* (Quekett, 1848). A good survey of the history of microscopes in the 18th century is the book *Eighteenth Century Microscopes, Synopsis of History and Workbook* (McCormick, 1987). Another comprehensive book that discusses the achromatic microscope is the *History of the Microscope* (Clay and Court, 1932). An extensive history of microscope optics during the period 1750–1850 is the doctoral thesis of Deiman (1992). For a deeper understanding of the cultural and social foundations of progress in microscopy during the 17th century in the Netherlands, as evidenced in the lives of Jan Swammerdam and Antoni van Leeuwenhoek, I refer the reader to the book *The Microscope in the Dutch Republic* (Ruestow, 1996). Turner's book *The Great Age of the Microscope* contains descriptions and illustrations of the entire collection of the Royal Microscopical Society in the United Kingdom (Turner, 1989).

I posit the existence of an enduring symbiotic relationship between the development of the microscope and advances in the fields of medicine and cell biology. In support of this linkage I have published a series of three articles that describe and analyze this proposition in three fields of microscopy. The first article is the "History of the Optical Microscope in Cell Biology and Medicine" (Masters, 2008). The second article is the "History of the Electron Microscope in Cell Biology" (Masters, 2009). And the third article is "The Development of Fluorescence Microscopy" (Masters, 2010). Another example of this strong link between the history of the microscope and advances in cell biology is the book *Three Centuries of Microscopes and Cell Biology* (Lemmerich and Spring, 1980).

In addition to the vast literature on the development of the light microscope I encourage people to visit museums that hold special collections of optical instruments—specifically those with collections of microscopes. I recommend visits to the following collections: The Billings Microscope Collection in the National Museum of Health and Medicine, Silver Spring, Maryland; Museum Boerhaave, Leiden, the Netherlands; Deutsches Optisches Museum, Jena, Germany; Deutsches Museum, Munich, Germany; Science Museum, London; Whipple Museum of the History of Science, Cambridge, United Kingdom; Museum of the History of Science, Oxford, United Kingdom; and the Museo Galileo (the former Institute and Museum of the History of Science), Florence, Italy. Smaller collections of

microscopes can be found in the Netherlands, France, and Italy. Furthermore, The Lentz Microscopy and Histology Collection from the Peabody Museum of Natural History at Yale University is available via the internet (Lentz, 2017).

4.2 Development of Light Microscopes: Case Studies

I will present selected case studies from the history of the microscope and discuss the insights I derived from them. These case studies were carefully selected to represent three disparate themes: how a popular book on microscopy generated popular interest in the microcosm, how a single-lens microscope made possible the visualization of microscopic living and inanimate specimens and generated significant knowledge of microorganisms, and an example from the application of the microscope to the structure of the nervous system that highlights ambiguities in the interpretation of microscopic observations.

In 1665 Robert Hooke, a member of the Royal Society in London, published a magisterial book *Micrographia* (Hooke, 1665). The book is written in Old English, which is readable by the modern reader. The complete title of this book is *Micrographia, or some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses with Observation and Inquiries thereupon*. In the 30-page preface Hooke explains his reasoning for writing the book. He posits that scientific inquiry begins with observation and the concomitant collection of data from the natural world. Hooke also wanted to promote the use of scientific instruments in scientific inquiry and he thought that his book would help to accomplish that aim.

Micrographia is replete with many detailed engravings of specimens that exhibit details that could not be observed with the naked eye. But when observed with his compound microscope they can be observed in all their magnificence. In *Micrographia* the reader sees images in exquisite detail. Some of the objects shown in the book include snowflakes, a needle point, the edge of a razor blade, cork, leaves, seeds, hair, wool, and the common fly. Hooke's microscopic observations of thin slices of cork led to his use of the term cells because he was reminded of the similarity, although of a vastly different scale, with cells in a monastery. He observed similar patterns in plants. A further notable accomplishment of Hooke is the use of dyes to increase the contrast of objects. Hooke made microscopic observations of both hair and wood in their natural state and after the application of dyes (Clark and Kasten, 1983).

These amazing images generated a lot of public interest about the microscope and its ability to help see the microscopic world that could not be appreciated with the naked eye. While the publication of *Micrographia* stimulated the interest of the general public as a result of the amazing objects whose details can only be seen with a light microscope, the microscopic images also stimulated questions concerning the validity of the images (McCormick, 1987). Are microscope images real or perhaps they are caused by artifacts introduced by microscopes? Questions on

the validation of images obtained with optical instruments were also raised with respect to the telescope. These questions have always been associated with microscopy and they remain valid today.

Details of Hooke's compound microscope can be found in *Micrographia*, and I briefly comment on the salient features (Bradbury, 1989; Hooke, 1665; Masters, 2008). First, I discuss the importance of the illumination system. The sun was often used as a source of illumination for microscopes (Masters, 2008). Hooke found that the intensity of illumination was too low for his compound microscope; therefore, he designed a new illumination system based on the flame from an oil lamp, a water-filled globe, and a lens that focused the illumination on the object. Hooke's compound microscope consisted of a microscope objective with a biconvex short focal length lens, an eyepiece lens, and a field lens (Bradbury, 1989; Hooke, 1665; Masters, 2008). Hooke noted limitations of his compound microscope that were due to spherical and chromatic aberrations: "[single-lens microscopes] do make the object appear much more clear and distinct, and magnify as much as double Microscopes: nay, to those whose eyes can well endure it,'tis possible with a single Microscope to make discoveries much better than with a double one, because the colors which do much disturb the clear vision in double Microscopes is clearly avoided and prevented in the single" (Hooke, 1665).

Next, I present a case study that shows a prescient example of how the light microscope was used to generate knowledge of microorganisms. I selected this important case because generating knowledge of the physical world is a useful metric of the utility of new developments in microscopy. I applaud instrument development in its own right; nevertheless, it is legitimate to ask what was learned as a result of using the new instrument.

Antoni van Leeuwenhoek (1632–1723) designed, constructed, and used his single-lens microscopes to observe the microscopic world; he observed living and nonliving specimens (Dobell, 1960; Ford, 1985, 1991; Fournier, 1991, 2003; Ruestow, 1996). Leeuwenhoek did not invent the single-lens microscope; they had been around for several decades prior to his construction of a single-lens microscope. Leeuwenhoek built approximately 400 single-lens light microscopes. Today only 12 of them exist, all of which are on view in the Museum Boerhaave (Fournier, 1991, 2003). The Zeiss Group gave me a working replica of a Leeuwenhoek single-lens microscope that I have used to observe the structure of the wing of a fly.

I start by describing the construction of his single-lens microscope (shown in Fig. 4.1). Then, I look at Leeuwenhoek's microtechnique. After which I discuss the seminal observations that Leeuwenhoek made with his single-lens microscopes (Dobell, 1960; Ford, 1985, 1991; Fournier, 1991, 2003; Ruestow, 1996).

Our knowledge of the optical specifications of Leeuwenhoek's microscopes derives from the following sources: Leeuwenhoek mailed his written communications on his microscope observations together with his specimens to the Royal Society in London (Ford, 1985, 1991). I am indebted to Brian Ford for his 1981 rediscovery of the original specimens that Leeuwenhoek sent; they were located in the Royal Society. In addition, Ford was able to derive the optical parameters

Fig. 4.1 Leeuwenhoek's microscope (replica) https://en.wikipedia.org/wiki/File:Leeuwenhoek_Microscope.png

(magnification, resolution) from the extant Leeuwenhoek single-lens microscope and to use them to form pictures of some of the original specimens.

It should be kept in mind that Leeuwenhoek was secretive about how he made the lenses for his microscopes and there were large differences in the construction materials, the methods of lens manufacture, and the optical properties of his various microscopes (Fournier, 2003). The basic design is shown in Fig. 4.1. A single lens was fixed between two plates and the specimen was attached to a sharp pin that could be moved to focus the object and to vary the field of view for the observation. Fluids were placed in a glass capillary that was fixed to the sharp pin of the microscope. According to Ford the lenses in the 12 extant microscopes were ground and had smaller chromatic aberrations than those of 16th-century compound microscopes (Ford, 1985, 1991; Van Zuylen, 1981). The measured resolving power of extentcof 1.35–4 μm and the magnification is in the range of 30 × to 200 × (Ford, 1985, 1991; Van Zuylen, 1981). Prior to Leeuwenhoek's death he bequeathed 26 single-lens microscopes to the Royal Society. Each of these instruments has a double-convex lens (Ford, 1985, 1991; Van Zuylen, 1981). Leeuwenhoek developed a technique to estimate the size of features in his specimens. He compared the size of his specimens with the diameter of a hair or a grain of sand. This early idea of a test object was further

developed in the next centuries for the relative measurement of the resolving power of microscope objectives. Usually, the wings of insects or the lines in diatoms were used as test objects. His specimens were usually mounted in olive oil or water.

Leeuwenhoek is credited with a number of firsts in the construction and application of the light microscope. He was the first to observe single-celled bacteria and other microorganisms, which he called "wee animalcules." He studied many specimens including the liver, brain, fat, and muscle tissues. Since the muscle specimens were transparent and without contrast Leeuwenhoek used saffron to stain them and generate contrast (Clark and Kasten, 1983). While he did not use stains on bacteria specimens, he made the first microscopic observations of bacteria. Furthermore, Leeuwenhoek made microscopic observations and wrote detailed reports on the following specimens: the cornea of the eye, the lens, the retina, the optic nerve, photoreceptors of the retina, dental plaque, muscle fiber striations, feathers and fish scales, hair, nails, bones, teeth, the liver, fat, milk, and unicellular organisms, as well as biological fluids such as blood (red blood cells were first discovered by Jan Swammerdam), urine, semen, sweat, and tears (Ruestow, 1996). Leeuwenhoek described microorganisms found in a drop of pond water (Ruestow, 1996). In 1661 Malpighi discovered blood capillaries in frog lungs. Unaware of Malpighi's work, Leeuwenhoek observed the flow of blood in the gills of tadpoles and he reported that the flow was temporally correlated with the tadpole's heartbeat (Ford, 1985, 1991; Ruestow, 1996). Leeuwenhoek also observed blood capillaries in the tail of tiny eels. All of Leeuwenhoek's reports on his microscopic observations together with his specimens were sent to the Royal Society; they were subsequently rediscovered by Ford (1985, 1991).

Handheld Leeuwenhoek microscopes were difficult to use since there was no microscope stand to give mechanical stability. Despite that limitation the corpus of knowledge is significant. I present this case study as an example of a new metric that can be used to judge the value of new developments in the history of the microscope: the metric is the significance of the knowledge gained with the invention of the new type of microscope.

A major advance in the design of microscope objectives occurred when Joseph Jackson Lister (1786–1869) developed his achromatic objectives that had minimal spherical aberration. In 1830 he published his method to construct microscope objectives: *On Some Properties in Achromatic Object-Glasses Applicable to the Improvement of the Microscope* (Lister, 1830). Lister's design principle permits the construction of microscope objectives with combinations of achromatic doublets or triplets, but without the compounding of spherical aberrations and coma. These microscope objectives had increased numerical aperture and therefore increased resolution. He used his new microscope objectives (constructed of achromatic doublets and triplets that did not multiply the spherical aberration and coma of each lens) to observe red blood cells and striated muscle.

Giovanni Battista Amici (1786–1863) was another innovator and developer of microscopes. Amici made three major improvements to the microscope: he developed a semispherical front lens achromatic microscope objective, he described the effect of coverslip thickness on image quality, and he introduced the use of

water immersion microscope objectives that enhanced both the resolution and the contrast of specimens. He was the first to use an immersion objective to increase the resolution (earlier suggested by Brewster). Amici found that the best images were obtained when the immersion liquid matched the refractive index of the glass of the coverslip and the front lens. Therefore, he used various immersion fluids, such as glycerin, and various oils. Amici used an ellipsoidal mirror objective to solve the problem of chromatic aberrations. This idea was first proposed by Christiaan Huygens as a method to avoid chromatic aberrations; earlier Newton made a similar proposal.

The next case study highlights ambiguities in the interpretation of microscopic observations. This is the story of two giants in the field of neuroscience who formulated opposing theories of the fine structure of the nervous system—the neuron theory (Cajal, 1906; Golgi, 1906). Both men were professors of anatomy in their universities and were highly respected histologists. Both Santiago Ramón y Cajal and Camillo Golgi made microscopic observations on similar specimens using similar staining methods that were first developed by Golgi. The silver staining technique that Golgi developed had a remarkable property. When Golgi applied his silver staining to tissues that contained an extremely high density of neurons, such as the cortex of the brain, the silver was only located on a sparse selection of neurons, with the absolute exclusion of any staining on other adjacent neurons. It was this incredible selectivity of silver staining that allowed Golgi and Cajal to trace the paths and connections of neurons.

But there were also significant differences in the specimens and staining techniques Cajal and Golgi used in their histological studies. Although Cajal at first used the staining techniques developed by Golgi, he subsequently made modifications in the method. Cajal also used methylene blue to stain neurons following the technique that Paul Ehrlich first described in 1887. In particular, Cajal studied embryos and young animals of several species. The result was that the neurons observed in the microscope by Cajal were sparser and shorter than the neuronal specimens Golgi investigated. These choices of specimen together with sparse silver staining led to the spectacular tracings of neurons in the nervous system (Cajal, 1995).

In 1885, while working at the University of Valencia, Cajal received a new Zeiss microscope and several microscope objectives that were designed for homogeneous immersion. As early as 1900 Cajal used Zeiss microscope objectives with numerical apertures of 1.3, 1.4, and 1.63. In addition, Cajal used microscopes made by Leitz in Wetzlar, Germany and by Reichert in Vienna, Austria. Cajal used a camera lucida to draw pictures of his microscopic observations. As a result of many years of microscopic studies Cajal believed that the neuronal organization was made up of discrete nerve cells (Cajal, 1906).

At the same time Golgi, working in Italy, formulated his independent theory of the structure of the nervous system. Golgi formulated a theory that the nervous system was formed from a large interconnected reticulum (Golgi, 1906). Golgi did in fact observe neurons with free nerve endings in his preparations. But he thought that the extremely small nerve endings in his preparations were beyond the

resolution of his microscope and therefore were invisible. This assumption allowed Golgi to have a consistent view of the nervous system as an interconnected reticulum and at the same time reconcile his observations of free nerve endings in his histological preparations.

How was this contentious conflict resolved? In 1906 the Nobel Committee awarded the Nobel Prize in Medicine to both Cajal and Golgi. I highly recommend the reader to peruse both Nobel Prize speeches (Cajal, 1906; Golgi, 1906). For Cajal and Golgi the resolution of their light microscopes was too low to observe synapses and synaptic vesicles. The discovery of synapses and synaptic vesicles was made after the invention and development of the electron microscope (Masters, 2009; Robertson, 1953).

References

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.

Bradbury, S. (1989). Landmarks in biological light microscopy. Journal of Microscopy, **155**(3), 281–305.

Cajal, S. R. (1906). The structure and connexions of neurons. Nobel Lecture, 12 December. https://assets.nobelprize.org/uploads/2018/06/cajal-lecture.pdf (Accessed on April 1, 2019).

Cajal, S. R. (1995). *Histology of the Nervous System*, *Santiago Ramón y Cajal*. Volume I, Volume II. Translated by Neely Swanson and Larry W. Swanson. New York: Oxford University Press.

Clark, G., and Kasten, F. H. (1983). *History of Staining*, 3rd edition, pp. 147–185. Baltimore: Williams & Wilkins.

Clay, R. S., and Court, T. H. (1932). *The History of the Microscope, compiles from original instruments and documents, up to the introduction of the achromatic microscope*. London: Charles Griffin and Company Limited.

Deiman, J. C. (1992). *Microscope Optics and J. J. Lister's Influence on the Development of the Achromatic Objective, 1750–1850*. A thesis submitted for the degree of Doctor of Philosophy of the University of London.

Dobell, C. (1960). *Anthony van Leeuwenhoek and his "Little animals"*. New York: Dover Publications.

Ford, B. J. (1985). *Single Lens. The Story of the Simple Microscope*. New York: Harper and Row, Publishers.

Ford, B. J. (1991). *The Leeuwenhoek Legacy*. Bristol: Biopress Limited.

Fournier, M. (1991). *The Fabric of Life. The Rise and Decline of Seventeenth-Century Microscopy*. Twente, The Netherlands: The University of Twente.

Fournier, M. (2003). *Early microscopes. A Descriptive Catalogue*. Leiden: Museum Boerhaave.

Golgi, C. (1906). The neuron doctrine-theory and facts. Nobel Lecture, 11 December. https://assets.nobelprize.org/uploads/2018/06/golgi-lecture.pdf (Accessed on April 1, 2019).

Hooke, R. (1665). *Micrographia: or, Some physiological descriptions of minute bodies made by magnifying glasses*. London: J. Martyn and J. Allestry, 1665. (first edition).

Lemmerich, J., and Spring, H. (1980). *Three Centuries of Microscopes and Cell Biology*. Heidelberg: European Cell Biology Organization.

Lentz, T. (2017). Lentz Microscopy and Histology Collection. Peabody Museum of Natural History at Yale University. https://medicine.yale.edu/cellbio/about/Lentz_Microscopy_Collection_5%20_31_17_304923_284_5094_v1.pdf (Accessed on February 25, 2020).

Lister, J. J. (1830). On some properties in achromatic object-glasses applicable to the improvement of the microscope. Philosophical Transactions of the Royal Society of London, **120**, 187–200.
Masters, B. R. (2008). History of the Optical Microscope in Cell Biology and Medicine. Encyclopedia of Life Sciences (ELS). Chichester: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0003082.
Masters, B. R. (2009). History of the Electron Microscope in Cell Biology. Encyclopedia of Life Sciences (ELS). Chichester: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0021539.
Masters, B. R. (2010). The Development of Fluorescence Microscopy. Encyclopedia of Life Sciences (ELS). Chichester: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0022093.
McCormick, J. B. (1987). *Eighteenth Century Microscopes, Synopsis of History and Workbook.* Lincolnwood, Illinois: Science Heritage Ltd
Quekett, J. (1848). *A practical treatise on the use of the microscope, including the different methods of preparing and examining animal, vegetable, and mineral structures*. London: H. Bailliere.
Robertson, J. D. (1953). Ultrastructure of two invertebrate synapses. Proceedings of the Society for Experimental Biology and Medicine, **82**, 219–223.
Ruestow, E. G. (1996). *The Microscope in the Dutch Republic, The Shaping of Discovery.* Cambridge: Cambridge University Press.
Schmitz, E.-H. (1981a). *Handbuch Zur Geschichte Der Optik*, Band I, Von der Antike bis Newton. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1981b). *Handbuch Zur Geschichte Der Optik*, Band 2, Von Newton bis Fraunhofer. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1982). *Handbuch Zur Geschichte Der Optik*, Ergänzungsband I, Das Fernrohr. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1983a). *Handbuch Zur Geschichte Der Optik*, Band 3, Teil A. Das XIX. Jahrhundert oder vom Pröblem zum Rechnen. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1983b). *Handbuch Zur Geschichte Der Optik*, Band 3, Teil B. Das XIX. Jahrhundert oder vom Pröblem zum Rechnen. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1984a). *Handbuch Zur Geschichte Der Optik*, Band 4, Teil A. Der Schritt in das XX. Jahrhundert. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1984b). *Handbuch Zur Geschichte Der Optik*, Band 4, Teil B. Der Schritt in das XX. Jahrhundert. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1989). *Handbuch Zur Geschichte Der Optik*, Ergänzungsband II, Teil A, Das Mikroskop. Bonn: Verlag J. P. Wayenborgh.
Schmitz, E.-H. (1990). *Handbuch Zur Geschichte Der Optik*, Ergänzungsband II, Teil B, Das Mikroskop. Bonn: Verlag J. P. Wayenborgh.
Turner, G. L'E. (1989). *The Great Age of the Microscope, The Collection of the Royal Microscopical Society through 150 Years*. Bristol: Adam Hilger.
Van Zuylen, J. (1981). The microscopes of Antoni van Leeuwenhoek. Journal of Microscopy, **121**, 309–328.
Volkmann, H. (1966). Ernst Abbe and his work. Applied Optics, **5**, 1720–1731.

Further Reading

Beck, R. (1865). *A Treatise on the Construction, Proper Use, and Capabilities of Smith, Beck, and Beck's Achromatic Microscopes*. London: John van Voorst, Paternoster Row. A Facsimile Edition is published by, Lincolnwood, Illinois: Science Heritage Ltd.
Bracegirdle, B. (1986). *A History of Microtechnique, The evolution of the microtome and the development of tissue preparation*. Second Edition. Lincolnwood, IL: Science Heritage Ltd.

Cajal, S. R. (1989). *Recollections of my life. Santiago Ramón y Cajal*. Translated by E. Horne Craigie with Juan Cano. Cambridge: The MIT Press.

Deiman, J. C. (1991). A myth revealed: The case of the 'Beeldsnyder achromatic objective'. Annals of Science, **48**, 577–581.

Gloede, W. (1986). *Vom Lesestein zum Elektronenmikroskop*. Berlin: VEB Verlag Technik.

Martin, L. C. (1966). *The Theory of the Microscope*. New York: American Elsevier Publishing Company.

McGucken, W. (1969). *Nineteenth-Century Spectroscopy, Development of the Understanding of Spectra 1802–1897*. Baltimore: The Johns Hopkins Press.

Rooseboom, M. (1940). Die holländischen Optiker Jan und Harmanus van Deijl und ihre Mikroskope. Janus, **44**, 185–197.

Chapter 5
Ernst Abbe and His Contributions to Optics

5.1 Introduction

> "Be good, be true, be just, remain true to your fellow man, and remain true to yourself."
>
> —the motto according to which Abbe lived his life
> from the funeral speech of Siegfried Czapski

In this chapter I suggest connections between the early life, education, and work experience of Ernst Abbe as well as his seminal contributions to optical metrology and his role as a social reformer. While several sources discuss what Abbe contributed during his career, I suggest links to help answer the question of what prompted him to do so? First, I posit that Abbe's difficult and financially unstable early life led to his penchant to work as a social reformer at Zeiss Werke and at the University of Jena. Second, I posit that Abbe's education provided him with the interest, knowledge, experience, and the analytical tools to devise, design, and construct metrology instruments of high accuracy that he later applied to optical components and optical instruments. I provide putative evidence in support of these two claims in the following sections (Fig. 5.1).

The variety of sources provides a trove of information on the life and works of Abbe. There is no substitute for reading his collected works, but his obituary as well as books and articles that contain biographical materials serve as secondary sources (Abbe, 1989; Auerbach, 1918, 1922; Cahan, 1996; Czapski and Eppenstein, 1924; Dippel, 1882; Feffer, 1994; Gerth, 2005; Hartinger, 1930; Masters, 2007; Nägeli and Schwendener, 1877; Rheinberg, 1905; Volkmann, 1966; von Rohr, 1940).

Perusal of these works provides some context to understand Abbe's written output. Abbe was fluent in both German and English. But he wrote most of his publications in German. It is natural to divide biographical sources on Abbe into those written while he was alive and those written after his death on January 14, 1905. Historians have been aided in their research by publication of the four volumes of Abbe's collected works in German: Volume I, theory of the microscope; Volume II, scientific treatises

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_5

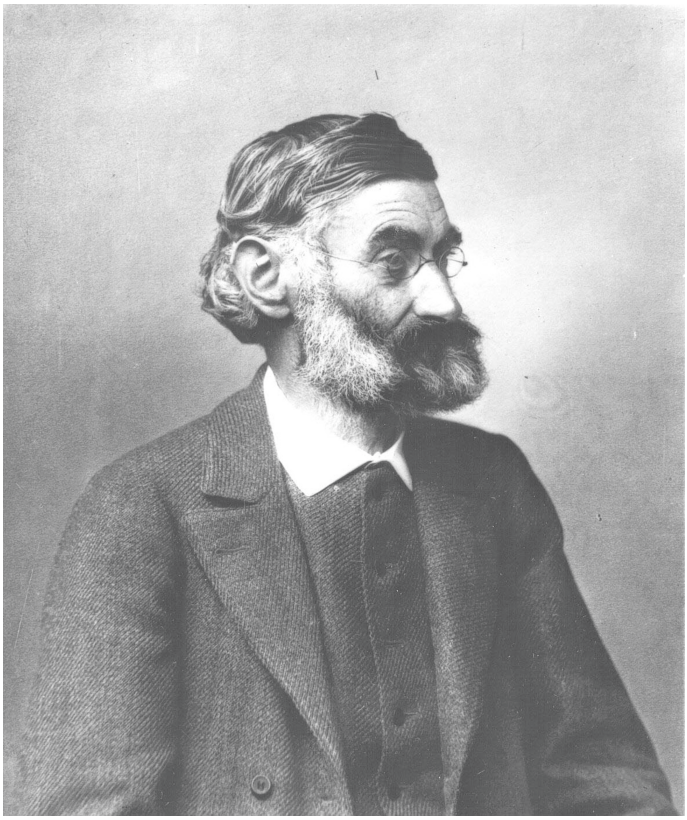


Fig. 5.1 Ernst Abbe as an older man. Reproduced with permission from the ZEISS Archives

from different fields, patents, and narratives; Volume III, sociopolitical lectures, speeches, and writings; and Volume IV, unpublished scientific and technical writings related to improving optical glass (Abbe, 1989).

My study of his collected works led me to conclude that Abbe, professor of physics at the University of Jena, physicist, and creative physicist at Zeiss Werke, published his findings related to the theory of image formation in the microscope and those related to new instrumentation not in physics journals but in journals that catered to the use and applications of microscopes (Feffer, 1994). His motivation was not widespread dissemination of his achievements in optics and instrumentation, but the promotion of new Zeiss microscopes to potential customers (Feffer, 1994). This practice is in sharp contrast with publications of other physicists, such as Helmholtz and Rayleigh, who published their work in highly respected physics journals with the aim of disseminating knowledge.

Abbe was reluctant to publish his discoveries in the scientific literature. However, Abbe's reluctance to publish his achievements changed when he learned that he had been nominated for the Nobel Prize in Physics (Cahan, 1996; Feffer, 1994). To enhance Abbe's chances of winning a Nobel Prize his colleagues began to assemble

his publications for review. This was the origin of the 1904 publication of Abbe's *Collected Works*. The first volume contained his contributions on the theory of the microscope. According to Cahan the fact that Abbe published his theory and inventions in general popular microscopy journals instead of highly respected physics journals, together with the fact that other physicists had difficulty in understanding Abbe's theory of image formation in the microscope, may have been the reason why the Nobel committee did not select Abbe for the Nobel Prize in Physics (Cahan, 1996). Feffer in his doctoral thesis describes Abbe's 1873 and 1882 publications of his theoretical work and his experiments on the microscope that were not published in physics journals, but were published in journals that catered to the practical microscope user and the potential purchaser of Zeiss microscopes (Feffer, 1994). Both Cahan and Feffer concur on this conclusion.

Abbe's failure to receive the 1904 Nobel Prize in Physics may have stimulated his desire to explain his new theory of image formation in the microscope to foreign physicists. Therefore, he devised a series of demonstrations that helped to bring his theory to a broader audience. Abbe's demonstrations probably derived from the series of laboratory experiments he used both to develop his theory of image formation in the microscope and to validate this theory with experiments. His seminal publication of 1873 on the theory of image formation in the light microscope based on diffraction together with a published eye-witness account of Abbe's experiments and demonstrations are the content of Chapter 6.

After Abbe died his colleagues published his lecture notes from the University of Jena on the geometrical theory of image formation in the 1906 *Handbuch der Physik* (Czapski, 1906). Furthermore, his theories were published posthumously in the book *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe* (Lummer and Reiche, 1910). Abbe also wrote a comprehensive chapter on geometrical and physical optics that was published in the textbook *Das Mikroskop* (Dippel, 1882). The third edition of *Grundzüge der Theorie der optischen Instrumente nach Abbe* reviewed Abbe's theoretical and experimental contributions to optical instruments (Czapski and Eppenstein, 1924).

5.2 Ernst Abbe: A Brief Biography

In this section I attempt to answer the questions posed in the introduction: not only the what but, more importantly, the why in Abbe's life. I depend on the following sources: Abbe's obituary published in the *Journal of the Royal Microscopical Society* (Rheinberg, 1905); Abbe's biography *Ernst Abbe, sein Leben, sein Wirken, seine Persönlichkeit* (Auerbach, 1918, 1922); a publication in *Die Naturwissenschaften, Zum fünfundzwanzigsten Todestage von Ernst Abbe* (Hartinger, 1930); an English publication in *Applied Optics*; *Ernst Abbe and His Work* (Volkmann, 1966); and a biography published in both German and English (Gerth, 2005) (Fig. 5.2).

I present selected key aspects of Abbe's life (1840–1905) in a chronological manner. Abbe is known as a scientist, an entrepreneur, and a prescient social

Fig. 5.2 Ernst Abbe as a young man. Reproduced with permission from the ZEISS Archives

reformer (Auerbach, 1922; Gerth, 2005; Volkmann, 1966). Ernst Carl Abbe was born on January 23, 1840 in Eisenach, Germany. I posit that Abbe's difficult childhood was permeated with long-term poverty, prolonged economic difficulties, and uncertainties that would destabilize any child.

How did Abbe's childhood experiences help to form his future progressive social views? Abbe's father worked as a spinner in a large cloth-spinning mill in Eisenach. Abbe's memories of his father's hardships as a worker had a lasting influence on his future activities in the business world. As a young boy, Abbe struggled to obtain a scholarship so that he could complete his secondary school. Following my perusal of the biographic sources cited above I conclude that Abbe's severe and difficult early childhood experiences were the foundation of his outstanding work as a social reformer in his adult life. My conclusions are confirmed by biographies of Abbe (Auerbach, 1922; Gerth, 2005; Volkmann, 1966; von Rohr, 1940).

Before I discuss Abbe's advanced studies at university, his doctoral work, and his Habilitation publication, which gave him the right to teach as a university professor, I divert the discussion to summarize his prescient sociopolitical reforms that he promulgated at the Carl Zeiss Foundation (Volkmann, 1966). The following is a partial listing of the prescient reforms that Abbe introduced in 1889 for his workers: a work day that was limited to eight hours, paid sick leave, paid vacation time, generous benefits for retired workers, security in work, and a financial

payment if a worker was terminated (Volkmann, 1966). All of these prescient advances in the way workers were treated at the Carl Zeiss Foundation had their origin in the horrible working conditions that Abbe's father endured.

Next, I continue with Abbe's education and the lasting impact that it had on his approach to the design, development, and the manufacturing of microscopes and other optical instruments at Zeiss Werke in Jena. This will address the why in relation to Abbe's style of research and his theoretical and experimental approach to optical instruments. From secondary school Abbe matriculated at the University of Jena where he developed his long-term interests in mathematics, physics, and experimental measurements. After two years at the University of Jena Abbe transferred to the University of Göttingen where he studied for a period of three years. The University of Göttingen had a world-class department of mathematics and the young Abbe was able to learn advanced mathematics from Professor Bernhard Riemann. During this period Abbe studied optics and astronomy. However, it was courses and laboratory work on the theory and use of precision measurements, the theory of experimental error, and the design, construction, and use of instruments to measure physical parameters, such as weak electric currents and magnetic fields, that prepared Abbe for his future scientific achievements in precision metrology at Zeiss Werke (Auerbach, 1922; Gerth, 2005; Volkmann, 1966; von Rohr, 1940).

In both his doctoral research and his Habilitation publication Abbe continued to concentrate on experimental and theoretical aspects of precision metrology and the theory of errors in measurements (Gerth, 2005). Abbe remained at the University of Göttingen for his doctoral studies. Professors Wilhelm Weber and Karl Snell mentored Abbe in his doctoral work, and in 1861 the young Abbe (aged 21) received a doctorate for his thesis *The Mechanical Equivalence of Heat*. Abbe then returned to the University of Jena to write his Habilitationsschrift *Ueber die Gesetzmässigkeit in der Vertheilung der Fehler bei Beobactungsreihen* (Concerning the regularity in the distribution of errors in observation series) that awarded Abbe the *Venia Docendi* and the right to teach at a university (Abbe, 1863). With these exceptional scholarly accomplishments and academic credentials Abbe became a member of the faculty at the University of Jena.

It was in Jena that Carl Zeiss, who managed a factory that made microscopes, met Abbe in 1866. Zeiss made Abbe a prescient proposal that forever changed the history of the microscope: Zeiss wanted Abbe to work on improving the design, construction, and testing of light microscopes. Zeiss tasked Abbe with developing a scientific basis for the manufacture of high-quality light microscopes.

5.3 Abbe's Contributions to Optics

Consistent with his experience and training Abbe made major contributions to the manufacture of light microscopes and other optical instruments that were produced and marketed by Zeiss Werke in Jena. Furthermore, Abbe was responsible for the

invention of precision optical instruments for testing optical components used to manufacture microscopes. Additionally, Abbe developed his theory of image formation in the optical microscope based on light diffraction. This topic is the subject of Chapter 6.

In this section I describe some of his theoretical developments, such as the Abbe sine condition and his definition of the Abbe number, as well as his invention and construction of new optical instruments for precision measurement of optical components and microscopes. Due to Abbe's excellent interdisciplinary education and his extensive experimental experience in the laboratory Abbe was well suited for his work in optics that bridged both theoretical and experimental aspects of physics. His seminal achievements in theory resulted in his brilliant theory of image formation in the light microscope based on light diffraction. His success in the design, development, and construction of high-accuracy, high-precision instruments to measure the dimensions and optical properties of glasses, liquids, lenses, microscope objectives, and complete optical systems is evidence of his experimental expertise.

5.3.1 Abbe Sine Condition

Before I describe the Abbe sine condition it is necessary to explain its significance in optical design. The Abbe sine condition provides ray-tracing constraint in the design of a lens or a complete optical system (Abbe, 1881). The sine condition, expressed as an equation, provides a bridge between the paraxial and the real aperture angles in the object and image space (Gross, 2005). If optical design is consistent with the Abbe sine condition, then optics can be corrected for spherical optical aberration. Moreover, there is a second great advantage: the capability to image points that are off the optical axis without coma. Coma is the off-axis optical aberration produced by round optical lenses. Coma is an off-axis aberration that is not symmetrical about the optical axis, and it rapidly increases as the third power of the lens aperture (Pedrotti et al., 2018). Furthermore, use of the Abbe sine condition provides a method to design microscope objectives that are devoid of spherical aberrations and coma. Such microscope objectives are called aplanic and they form sharp images both on and off the optical axis.

The Abbe sine condition is given by the following mathematical form

$$M' = \frac{n \sin \sigma}{n' \sin \sigma'} = \text{constant} \tag{5.1}$$

where M' is lateral magnification, σ and σ' are angles with respect to the axis of rays that come from an arbitrary point on the object before and after traversing the optical system, and n and n' are refractive indices of the media on both sides of the objective. The meaning of (5.1) is that all aperture angles will result in the same magnification. Lateral magnification, which results from refraction of all the zones

of a lens, must be the same for coma to be eliminated (Pedrotti, Pedrotti, Pedrotti, 2018). Feffer stated the Abbe sine condition as follows: "the Abbe sine condition requires that all rays coming from the object deliver images with precisely the same magnifications, regardless of their angle of incidence" (Feffer, 1994). An alternative statement of the Abbe sine condition is the following: in the absence of spherical aberration (spherical aberration is corrected) paraxial magnification is equal to magnification of the marginal ray (i.e., a ray with the largest angle admitted by the system aperture). If the Abbe sine condition is valid for a given optical system, then a small region of the object plane in the neighborhood of the axis is imaged sharply by a pencil of any angular divergence (Born and Wolf, 1999).

The Abbe sine condition is an important relation in geometric optics that is extremely useful in optic ray tracing for the design of lenses and optical systems. It is a good example of simultaneous yet independent invention, rediscovery, and reformulation of the original discovery. Throughout this book there are examples of invention, reinvention, simultaneous discovery, and rediscovery. According to the writings of Born and Wolf what is now termed the Abbe sine condition was derived independently and simultaneously by Abbe and Helmholtz using completely different derivations (Born and Wolf, 1999). In Chapter 7 I present my English translation of Helmholtz's publication, which contains his independent derivation of his sine condition and his independent derivation of the limiting resolution of the light microscope due to diffraction (Helmholtz, 1874). Furthermore, Born and Wolf claim that Rudolf Clausius derived the sine condition in 1864 and Helmholtz derived his sine condition from photometric analysis in 1874; it was rediscovered in 1880 by Abbe who understood its significance in optical ray tracing (Born and Wolf, 1999).

In 1879 Abbe published an English paper on the use of the Abbe sine condition in the *Journal of the Royal Microscopical Society* (Abbe, 1879). In this paper Abbe presented a new form of the sine condition. Abbe wrote that an aplanatic microscope objective will have the following constraint on light rays

$$f' = \frac{h}{\sin \sigma'} = \text{constant} \tag{5.2}$$

where h is the linear distance of a ray parallel to the optical axis from the axis, σ' is the angle that the ray forms with the axis after leaving the optical system, and f' is the focal length of the optical system. The meaning of (5.2) is that the image focal length is constant for all aperture angles and pupil heights. Abbe claimed he derived his sine condition from purely geometrical considerations. But Abbe never published such a derivation. Still, it is conceivable that Abbe made such a derivation of his sine condition.

Mansuripur begins his book *Classical Optics and Its Applications* with an excellent analysis of the Abbe sine condition (Mansuripur, 2009). Kingslake in his useful book *Lens Design Fundamentals* provides the reader with a derivation of the Abbe sine condition that is based on geometric optics (Kingslake and Johnson,

2010). Additionally, Sasián used geometric optics to derive the sine condition (Sasián, 2013). The Abbe sine condition can also be derived from Fourier optics (Simon et al., 1979).

5.3.2 Abbe Number

The optical performance of a lens depends on its design and the degree of perfection of the lens-grinding and lens-polishing processes. The quality of glass blanks from which the lens is manufactured is an important limitation. The quality of the glass depends on the purity of its components as well as the proper mixing, melting, and cooling of the glass. The goals are minimal impurities, absence of bubbles, absence of mechanical strain, and homogeneous properties of density and refractive index throughout the glass.

Glasses can be classified according to their refractive index and dispersion (variation of the refractive index with the frequency or wavelength of light). To help classify different glasses Abbe defined a number V, called the Abbe number, which he defined as follows

$$V = \frac{(n_D - 1)}{(n_F - n_C)} \tag{5.3}$$

where n_D, n_F, and n_C are refractive indices of material at the Fraunhofer D, F, and C spectral lines (589.2, 486.1, and 656.3 nm). Furthermore, Abbe designed a graph, called the Abbe diagram, which plots the refractive index at the 587.6 nm wavelength against the Abbe number.

5.3.3 Improved Manufacturing Methods and Increased Accuracy of Optical Instruments

Carl Zeiss tasked Professor Abbe, then professor of physics at the University of Jena, to develop a scientific basis for the design and manufacture of the microscope. To that end Abbe worked on several approaches: investigations into image formation in the microscope, studies of ways to improve geometrical optics so as to minimize optical aberrations in microscope objectives, the invention of optical testing methods of high accuracy and precision, and improvement of the optical qualities of the glass used to produce lenses. In Chapter 6 I describe and analyze Abbe's seminal theory of image formation in the light microscope based on diffraction. Here I describe some instruments that Abbe invented or improved for optical shop testing. In 1870 Zeiss Werke was manufacturing its own microscope objectives. Abbe used his expertise in geometrical ray tracing and the Abbe sine

condition to design and specify the precise type of glass as well as the shape and dimensions of each optical component that would form the complete microscope objective.

Abbe also designed, improved, and had several optical shop testing instruments constructed. These accurate and high-precision instruments were required to check that the manufacturing processes for the microscope objective agreed with Abbe's specifications (Volkmann, 1966).

For detailed information on these new optical shop testing instruments I used Volume II of the *Gesammelte Abhandlungen* (Abbe, 1906) for Abbe's description of the *Dickenmesser* (measures the thickness of an optical component), the *comparator* (compares the dimensions of two objects), the *Fokometer* (measures the focal length of a lens or an optical system), and the *Sphärometer* (measures the radius of curvature of a lens).

Furthermore, I based my description of some of Abbe's instruments on the book *Ernst Abbe* (Auerbach, 1922; Zeiss, 1878): Abbe's apertometer of 1870 (measures the numerical aperture of a microscope objective), Abbe's *Spektrometer* (measures the refractive index and the dispersion of glasses), Abbe's refractometer (Fig. 5.3) of 1869 (measures the refractive index of liquids or glasses), and Abbe's interferometer of 1885 (measures plane-parallel glass plates based on interference). I used the paper "Ernst Abbe and His Work" (Volkmann, 1966) to describe Abbe's instruments that precisely measure the dimensions of optical systems: the comparator (to compare dimensions of optical components), the thickness meter, and the height meter.

Figure 5.3 shows the Abbe refractometer used to measure the refractive index of liquids or glasses.

Abbe is credited with the concept of numerical aperture (Abbe, 1881). In 1870 Abbe began his investigations into numerical aperture in the light microscope (Auerbach, 1922).

5.3.4 Abbe's New Illumination System for the Microscope

Abbe realized that, in addition to the capability of Zeiss Werke to manufacture microscopes that imaged objects with high optical quality and fidelity, to be a market leader it was also important to make microscopes as user friendly as possible. In general, this goal is still applicable today. In this section I present an important microscope accessory, the Abbe condenser, which became very popular with microscopists (Volkmann, 1966).

In 1872 Abbe invented a new illumination (condenser) system for the microscope. It was published in German in 1873 in Schultze's *Archiv für mikroskopische Anatomie* and in English in 1875 in the *Monthly Microscopical Journal* (Abbe, 1873, 1875) (Fig. 5.4).

The device consists of a mirror that reflects rays of light from a source to the condenser. The mirror has no lateral movement and can only turn in a fixed point in the optical axis of the instrument. There is an adjustable iris or aperture and two nonachromatic lenses with a combined focal length of 15 mm. The upper portion of

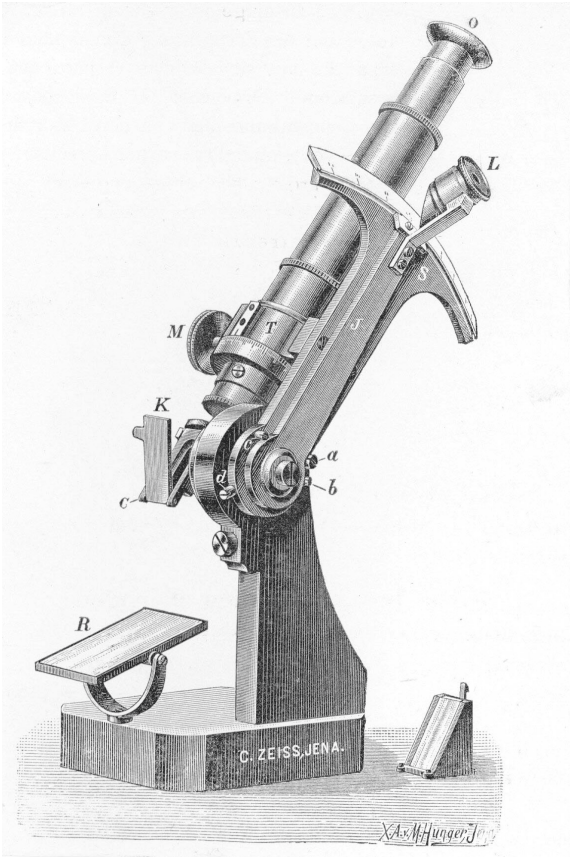


Fig. 5.3 Abbe refractometer. Reproduced with permission from the ZEISS Archives

the condenser that contained the two lenses could be readily displaced by the user in a direction perpendicular to the optical axis of the microscope. The Abbe condenser was useful because it allowed the microscope user to readily switch between direct illumination and oblique illumination. Abbe designed his condenser to have the maximum aperture for a superior focus (Abbe, 1873, 1875).

5.3.5 Abbe's Homogeneous Immersion

The Italian microscopist Giovanni Amici is credited with the use of oil and later water to fill the space between the specimen and the microscope objective. Amici made the important observation that placing liquids in the space between the objective and the specimen improved the image. Amici experimented with several kinds of liquids that he placed between the tip of the microscope objective and the specimen: water, oils, and glycerin. Amici's discovery that an improved image

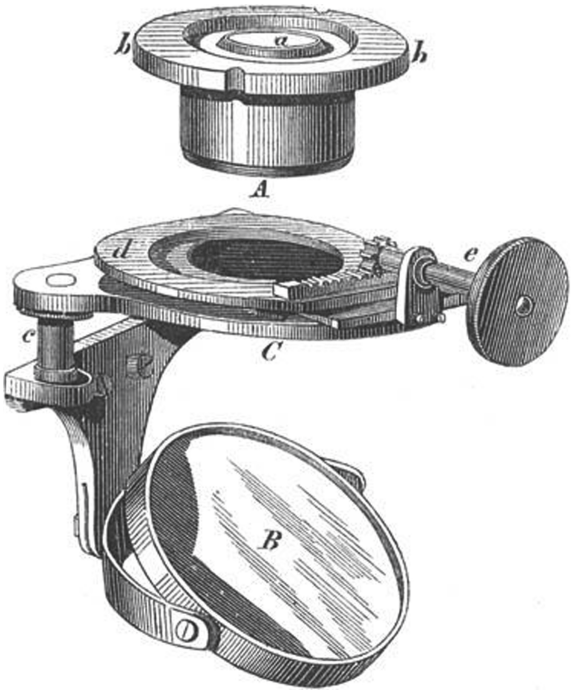


Fig. 5.4 Abbe's new illumination system for the microscope

could result from using immersion fluids to replace air was empirical and had no scientific basis in geometrical optics and ray tracing to explain such impressive results.

The space between the specimen and the tip of the microscope objective typically consisted of air. An exception was the light microscopes manufactured by Hartnack in which the tip of the microscope objective was immersed in water, which was in contact with the specimen.

Ernst Abbe corresponded with John Ware Stephenson, treasurer of the Royal Microscopical Society, over many years. It was Stephenson who proposed a solution to the problem of spherical aberrations in microscopy. Stephenson claimed that they were caused by the various thickness of the cover glasses microscopists placed over the specimen mounted on a glass slide. Stephenson was a microscopist, but was not trained as a physicist. However, Abbe was a physicist with extensive theoretical and experimental training and experience in optics. He could see that the use of immersion objectives, which Stephenson denoted as homogeneous immersion, would increase the numerical aperture (NA) of the microscope objective and therefore the resolution.

What is the origin of the word "homogeneous" in Stephenson's term homogeneous immersion? It refers to the equal or similar refractive index of all the materials from the front lens of the microscope objective to the cover glass above

the specimen. In his publications Abbe gave Stephenson credit for the idea of homogeneous immersion.

In 1879 Abbe wrote a paper in English that was published in the *Journal of the Royal Microscopical Society* titled "On Stephenson's System of Homogeneous Immersion for Microscope Objectives" (Abbe, 1879). This was followed in 1880 by Abbe's second publication titled "The Essence of Homogeneous Immersion" (Abbe, 1880). In 1889 Zeiss Werke manufactured light microscopes that finally achieved the resolution that Abbe derived in his 1873 publication: the microscope objective would be apochromatic with an NA of 1.63 when used in immersion mode with 1-bromonaphthalene (Cahan, 1996) (Fig. 5.5).

I end this brief survey with an example that clearly demonstrates the utility of innovative advances in the theory of image formation in the light microscope, the associated technological improvements, and how they combine to promote our knowledge and understanding of medicine and biology. In 1882 Professor Robert Koch, a bacteriologist working in Berlin on the agents that cause infectious diseases, was using a Zeiss microscope with a Zeiss oil immersion microscope objective and an Abbe condenser or illuminator. With this state-of-the-art microscope Koch was able to identify the microorganism that caused tuberculosis (Feffer,

Fig. 5.5 Zeiss microscope made by Carl Zeiss, Jena, 1879

1994). As described by Feffer the discovery of the tuberculosis bacillus by Koch resulted in great demand by microscopists for the now-famous homogeneous immersion microscope objectives that Abbe designed and Zeiss manufactured and advertised (Feffer, 1994).

References

Abbe, E. (1863). Ueber die Gesetzmässigkeit in der Vertheilung der Fehler bei Beobactungsreihen. [Concerning the regularity in the distribution of errors in observation series]. Dissertation for attainment of the *Venia Docendi* in the philosophy faculty in Jena in 1863. In: *Gesammelte Abhandlungen*, II. Hildesheim: Georg Olms Verlag, pp. 55–81.

Abbe, E. (1873). Ueber einen neuen Beleuchtungsapparat am Mikroskop. *Archiv für mikroskopische Anatomie*, **IX**, 469–480.

Abbe, E. (1875). V. A new illuminating apparatus for the microscope. Monthly Microscopical Journal, **XIII**, 77–82.

Abbe, E. (1879). On Stephenson's system of homogeneous immersion for microscope objectives. Journal of the Royal Microscopical Society, **2**, 256–265.

Abbe, E. (1880). The essence of homogeneous immersion. Journal of the Royal Microscopical Society, **1**, 526.

Abbe, E. (1881). On the estimation of aperture in the microscope. Journal of the Royal Microscopical Society, **1**, 388–423.

Abbe, E. (1989). *Gesammelte Abhandlungen*, I-IV. Hildesheim: Georg Olms Verlag.

Abbe, E. (1906). Messapparate [Dickenmesser, Comparator, Sphärometer] für Physiker. XI, In: Abbe, E. (1906). Gesammelte Abhandlungen, II. Hildesheim: Georg Olms Verlag, pp. 206–214.

Auerbach, F. (1918). *Ernst Abbe: sein Leben, sein Wirken, seine Persönlichkeit nach den Quellen und aus eigener Erfahrung geschildert von Felix Auerbach*. Leipzig: Akademie Verlag Gesellschaft.

Auerbach, F. (1922). *Ernst Abbe: sein Leben, sein Wirken, seine Persönlichkeit nach Quellen und aus eigener Erfahrung geschildert von Felix Auerbach*. Zweite Auflage. Leipzig: Akademische Verlagsgesellschaft.

Born, M., and Wolf, E. (1999). *Principles of Optics*, 7th (expanded) edition. Cambridge: Cambridge University Press.

Cahan D. (1996). The Zeiss Werke and the ultramicroscope: the creation of a scientific instrument in context. In: Buchwald, J. Z., editor, *Scientific Credibility and Technical Standards in 19th and early 20th Century Germany and Britain*, Dordrecht: Kluwer Academic Publishers, pp. 67–115.

Czapski, S. (1906). Geometrische Optik. In: A. Winkelmann, Ed. *Handbuch der Physik*, Zweite Auflage, Sechster Band, Optik. Leipzig: Verlag von Johann Ambrosius Barth.

Czapski, S., and Eppenstein, O. (1924). *Grundzüge der Theorie der optischen Instrumente nach Abbe*. Dritte Auflage. Leipzig: J. A. Barth Verlag.

Dippel, L. (1882). *Das Mikroskop und seine Anwendung*. Erster Theil. *Handbuch der Allegemeiner Mikroskopie*, Zweite Auflage. Braunschweig: Friedrich Vieweg und Sohn.

Feffer, S. M. (1994). *Microscopes to munitions: Ernst Abbe, Carl Zeiss, and the transformation of technical optics, 1850–1914*. Ph.D. Dissertation, University of California, Berkeley, 1994. Ann Arbor: UMI Dissertation Services.

Gerth, K. (2005). *Ernst Abbe, Scientist, Entrepreneur, Social Reformer*. Jena: Verlag Dr. Buseert & Stadeler.

Gross, H. (2005). Aberrations: The sine condition. In: *Handbook of Optical Systems, Vol. 1: Fundamentals of Technical Optics* (p. 495). Weinheim, FRG: Wiley-VCH.

Hartinger, H. (1930). Zum fünfundzwanzigsten Todestage von Ernst Abbe. *Die Naturwissenschaften*, **18**, 49–63.

Helmholtz, H. (1874). Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope. *Poggendorff's Annalen der Physik*, Jubelband 1874, 557–584, Leipzig. Translated as: "On the theoretical limits of the optical capacity of the microscope." Monthly Microscopical Journal, **16**, 15–39 (1876).

Kingslake, R., and Johnson, R. B. (2010). *Lens Design Fundamentals*. Second Edition, New York: Academic Press.

Lummer and Reiche (1910). *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe*. Braunschweig: Druck und Verlag con Friedrich Vieweg und Sohn.

Mansuripur, M. (2009). Chapter 1, Abbe's sine condition, in: *Classical Optics and its Applications*. Second Edition. Cambridge: Cambridge University Press, pp. 9–22.

Masters, B. R. (2007). Ernst Abbe and the foundation of scientific microscopes. Optics & Photonics News, February, 19–23.

Nägeli and Schwendener (1877). *Das Mikroskop*. Leipzig: Verlag von Wilhelm Engelmann.

Pedrotti, F. L., Pedrotti, L. M., and L. S. Pedrotti (2018). *Introduction to Optics*, Third Edition. Cambridge: Cambridge University Press.

Rheinberg, J. (1905). Obituary. Journal of the Royal Microscopical Society, **25**, 156–163.

Sasián, J. (2013). *Introduction to Aberrations in Optical Imaging Systems*. Cambridge: Cambridge University Press.

Simon, J. M., Ratto, J. O., and Comastri, S. A. (1979). Sine condition derivation via Fourier optics. Applied Optics, **18**, 2912–2913.

Volkmann, H. (1966). Ernst Abbe and his work. Applied Optics, **5**, 1720–1731.

von Rohr, M. (1940). *Ernst Abbe*. Jena: Gustav Fischer.

Zeiss, C. (1878). Description of Professor Abbe's apertometer, with instructions for its use. Journal of the Royal Microscopical Society, **I**, 19–22.

Further Reading

Abbe, E. (1879). On new methods for improving spherical correction, applied to the construction of wide-angled object-glasses [microscope objectives]. Journal of the Royal Microscopical Society, **II**, 812–824.

Abbe, E. (1893). Apparat zur Bestimmung der Brennweite von Linsensysten (Fokometer). In: Abbe, E. (1906) *Gesammelte Abhandlungen*, XIII. Hildesheim: Georg Olms Verlag, pp. 215–218.

Hopkins, H. H. (1950). *Wave theory of Aberrations*. New York: Oxford University Press.

Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press.

Masters, B. R. (2009). C. V. Raman and the Raman effect. Optics & Photonics News, March **20**, 40–45.

Chapter 6
Abbe's Theory of Image Formation in the Microscope

6.1 Introduction

Abbe's seminal theory of image formation in the microscope "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung" was published in *Archiv für mikroskopische Anatomie* (Abbe, 1873a). Abbe developed a theory of image formation in the optical microscope that was based on the diffraction of light by the specimen and the subsequent formation of an image. Abbe's theory introduced the concept of numerical aperture (NA) and the main factors that affect two-point resolution in the optical microscope (Abbe, 1882a, b). Earlier, in 1870 Abbe began his investigations into numerical aperture in the light microscope (Auerbach, 1918, 1922). These main factors are (1) the wavelength of illumination and (2) the numerical aperture of the microscope objective. If both these factors are known through measurements, then the two-point resolution of the optical microscope can be calculated (Abbe, 1873a).

Abbe began these experimental investigations in 1870, at which time he realized the importance of numerical aperture (NA). In 1871 Abbe designed and constructed a simple apparatus to investigate the role of diffraction on image formation in the light microscope. Various objects consisting of periodic gratings or grids and biological specimens with periodic lines and patterns were used (Auerbach, 1918, 1922). The object was illuminated and the specimen diffracted the light into several orders of diffraction. The numerical aperture of the microscope objective determined which diffraction orders could enter the microscope objective; this diffracted light formed a diffraction pattern in the back focal plane of the lens. Various masks or apertures were placed in the apparatus to select which diffraction orders could enter the aperture of the microscope objective. The resulting diffraction pattern could be observed in the back focal plane of the objective. This apparatus was originally used by Abbe to understand the role of diffraction in image formation in the light microscope. The creative genius of Abbe is demonstrated in his design and construction of these experiments, his interpretation of images in the back focal

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_6

plane of the microscope objective under various conditions, and finally his development of a comprehensive testable theory of image formation based on diffraction. Later, Abbe used this apparatus to demonstrate and explain his theory of image formation in the microscope (Abbe, 1873a).

Abbe's Experiments began when he developed experimental tools to try and understand the physical processes of image formation in the microscope. In the 1940 Zeiss catalog I found a product that was marketed as "Diffraktionsapparats nach Abbe" (Abbe's diffraction apparatus). This apparatus evolved from an earlier product that appeared in the 1878 Zeiss catalog marketed as "Diffraktionsplatte," which contained several sets of grids of different line spacings on a microscope slide and was used as a test object in the microscope. Abbe's diffraction apparatus allowed the user to perform diffraction experiments easily. In the1960s Helmut Haselmann and Dr. Kurt Michel of the Carl Zeiss Foundation developed a kit form of Abbe's Experiments (Bradbury, 1996). The demonstration equipment contained various objects and various apertures that could be inserted into the apparatus. Today the kit form of Abbe's Experiments is difficult to obtain.

John Ware Stephenson, treasurer of the Royal Microscopical Society (London), witnessed Abbe's demonstration experiments in London and subsequently described them in his 1877 publication "Observations on Professor Abbe's experiments illustrating his theory of microscopic vision" (Stephenson, 1877). Stephenson's publication helped to reduce confusion and misunderstanding among English microscopists. Because this was the only eye-witness account of Abbe in London demonstrating his theory I think it useful to share the entire publication with the reader, which is the content of Sect. 6.4 (Stephenson, 1877).

Abbe's lengthy publication had an immediate effect on the reader because it contained no equations and no figures (Abbe, 1873a). Abbe promised to write a paper in the *Jena Journal for Medical and Natural Sciences* (*Jenaische Zeitschrift für Medizin und Naturwissenschaft*) on his theory complete with mathematical analyses and derivations (Abbe, 1873a). However, Abbe died in 1905 without publishing the promised mathematical work.

On December 16, 1874 Henry E. Fripp's English translation of Abbe's paper on his theory of image formation in the microscope based on diffraction was read before the Bristol Microscopical Society. Fripp published an English translation that included both a preface and a postscript that Fripp wrote and amended in his translation. The translation titled "Contributions to the Theory of the Microscope, and the Nature of Microscopic Vision" was published in the *Proceedings of the Bristol Naturalists' Society* (Fripp, 1874a). Fripp's English translation became the standard source for people who could not read German to access Abbe's theory.

Fripp was a physician—not a physicist—and that may be part of the problem with his English translation of Abbe's publication. In Fripp's preface to his translation of Abbe's seminal publication, Fripp made several key points that I find inappropriate for a scientific translation. For example, Fripp noted that "a few paragraphs relating to illuminating apparatus, and the conditions upon which their effects depend, are omitted [Fripp deleted them], as being only supplementary to the principal subject, and also because reference is therein made to doctrines which are

somewhat at variance with English opinion and practice, and which require for their proper understanding further explanation than is afforded by the curt mention of them in Dr. Abbe's essay" (Fripp, 1874a). At the end of Fripp's English translation he added his postscript.

Following repeated perusal of Fripp's English translation, analysis of its significant failure to accurately relate the physics as it was written by Abbe, the omission of material that conflicted with the prevalent understanding of the subject in England, and finally the inaccurate preface and postscript that Fripp added to the translation, I decided to make my own English translation of Abbe's 1873 German publication. Therefore, I present my translation of Abbe's 1873 publication on his theory of image formation in the microscope in Sect. 6.2.

6.2 Barry R. Masters' Translation of Abbe's 1873 Theory of Image Formation in the Microscope

The following is my English translation of Abbe's 1873 paper "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung" that was published in German in M. Schultze's *Archiv für mikroskopische Anatomie*, **IX**, 413–468.

Contributions to the Theory of the Microscope and the Nature of Microscopic Vision

I. The Construction of the Microscope Based on Theory. II. The Dioptric Conditions for the Performance of the Microscope. III. The Physical Conditions for the Imaging of Fine Structures. IV. The Optical Power of the Microscope.

From Dr. E. Abbe,
Professor in Jena [außerordentlicher Professor]

I. The Construction of the Microscope Based on Theory

1. In our handbooks of micrography [microscopy], one occasionally finds the fact that relates to the construction of microscopes and their progressive improvement, so far have been almost exclusively empirical matters, that is the skillful and persistent trials of experienced practitioners. Now and then, the question will probably be raised: why the theory that sufficiently accounts for the mode of action of the microscope after it was constructed, does not at the same time become the basis for its construction? Why do we not construct this kind of optical instrument according to calculations that are based on theoretically developed equations as has been so successfully done since the time of Fraunhofer with the telescope and in more recent times for the optical parts of the photographic camera? The reason for the continued existence of the empirical method is generally attributed to technical difficulties, or the supposed impossibility, of maintaining the

required accuracy during the construction of microscope objectives so that the dimensions for the individual constituent lenses comply with required accuracy. This explanation appears at first sight, in fact, to be quite plausible; because the smallness of the dimensions, which is inevitable especially in the more powerful objectives, the difficulties of their production by keeping the required precision measurements can be regarded as extraordinarily large. Nonetheless, I found this concern to be incorrect after I had gained more knowledge about the facilities and the technical procedures that are applied to the construction of the microscopes in a well-conducted optical workshop. A careful consideration of the scientific and technical means available to the practical optician, and the critical comparison of various kinds of difficulties serve as a guiding thread and key to a theoretical discussion of the impacting conditions, led me to the conclusion, tested by actual successful results, that in the state of optical technology, the design of lenses and lens systems according to prescribed dimensions with all elements in a correct accuracy, can be made with an exactitude that ensures correct performance, and with greater facility than any other procedure offers for the fulfillment of the same conditions with equally good results; and that therefore, it would only be important to provide the correct calculations of the optical effect for each separate optical element into account for the success of a theoretical construction. In this viewpoint, I in conjunction with Mr. C. Zeiss in Jena have made a serious investigation to provide a secure theoretical basis to the construction of microscopes and their further perfection just as Fraunhofer has achieved for the production of astronomical telescopes. Thanks to the readiness with which Mr. Zeiss provided me supplying me through the excellent resources and proficient manpower his workshop for several years, and thanks to the zeal with which the skillful work of the leader of this workshop and his skillful agents impacted the work, this experiment after prolonged efforts achieved its goal. For some time, in the above-mentioned workshops, microscope systems which are at the pinnacle among comparable systems, from the lowest power to the highest power, have been constructed entirely by theoretical rules.

The specified constructions are regulated by strict calculation, on the basis of careful examination of the materials to be used, until the last detail for each part – every curve, every thickness of glass, each degree of aperture is determined by calculation, so that trial and error is excluded. The optical constants of each piece of glass are obtained from measurements of trial-glass prisms with a spectrometer, in order to compensate for any accidental variation of the material by a suitable variation of the construction. The individual components are ground exactly as possible to the prescribed dimensions and accurately assembled. In the highest power objectives, only a single element of construction (a lens distance) is left variable to the last, to compensate by means of an offset the inevitable small deviations of the work. It is evident here that a sufficiently thorough theory, in conjunction with an efficient technique, that uses all the tools physics can offer to practical optics, can replace the empirical methods of procedure with success in the construction of microscopes.

2. In course of the work, which led to this result, it has now been found that the current theory of the microscope in essential respects is very incomplete. In the first place, the conditions of a perfect image were pointed out and the causes of the

imperfection were discussed; this was proved to be inadequate for the real situation as it exists at the microscope. The circumstance that there is an amount of angular aperture, that is unknown in any other instrument, comes into question and makes, the terms of aberration quite useless, even for any moderately critical estimate of microscopes already constructed, to say nothing about attempts to determine beforehand the effect of combinations not yet constructed. To obtain data for an investigation of this kind, a theoretical analysis of the effect of a lens system of large angular aperture had to be carried out on a much wider mathematical basis, and in much more detail than had been previously performed; it became revealed that the correct performance of any combination of lenses, that meets the requirements of a microscope system depends on an unexpectedly large number of independent conditions; accordingly, their proper assessment is not possible without the introduction of some new factors in the general theory of the microscope.

To develop such a theory more fully in the directions that were indicated was a purely mathematical problem, which was completely done with the established principles of dioptrics. Experience and experiment concerned the inquiry only to the extent as it was a question of the contribution of the individual theoretically detectable error sources on the finished microscope; and also, to correctly estimate their very unequal significance for the practical use of the instrument. In contrast, a new deficiency in our theoretical knowledge revealed itself, which could only be met by enhanced experience. It is characterized by the uncertain, and often conflicting views concerning the significance of angular aperture of the objective, and the so-called "defining" and "resolving" powers of an objective. To remove the uncertainty on this point and to gain a clear insight into the conditions which operate here, was the condition *sine qua non* for any successful trial in the indicated direction. For upon the effect supposed to be obtained by angular aperture depends the entire direction and solution to the problem. All ratios of construction are quite different, depending as it is calculated for an objective with a 40 or 90 or 150 degrees of angular aperture. But what kind of effect was to be expected remained completely doubtful, as long as no accurate account of the real significance of these factors could be given.

3. The investigations that I have independently performed, in order to bring these matters to a conclusion, provided the results that an essential element in the optical function of the microscope has been entirely overlooked. In the explanation and interpretation of the effects of this instrument has it namely, as a matter of course been accepted as a self-understood proposition, that image formation of a microscopic object occurs in every detail, according to the same dioptric laws by which images are formed in the telescope or on the imaging plate of a camera; and one has tacitly assumed that all optical functions of the microscope are just like those of other instruments, and determined in the microscope by the geometrically traceable relations of the ray paths of refracted light. A more rigorous review of the known experiences, in which the traditional distinction between the definitions of "defining" and "resolving" powers are based, has shown that this seemingly natural assumption as inadmissible. It has been shown that, although for some special cases, it remains valid, but in general, and especially in those objects in which the

microscope should exhibit its highest performance, the generation of the microscopic images is linked with a hitherto neglected physical process, which originates in the object itself, and occurs irrespective of the construction of microscopes, although the measure of its effect is directly dependent on the construction of the objective.

The consequences of these facts directly relate to the most important problems in micrography. They led to the detection of a very specific function of angular aperture, and a connection to clear and secure notions about the so-called "defining" and "resolving" power which forms the optical capacity of the microscope, and the correct perception of the conditions that affect its performance can be accurately determined. From this follow certain practical rules for the rational construction of the microscope, as well as suggestions for good methods to test its performance. On the other hand, the expanded newly gained ground obtained by experiment and theory led to generally relevant conclusions regarding microscopic vision. Therefore, it was possible not only to fix the limits of the visible beyond which no further resolution of the microscopic structure could be expected, but also to bring to light a fact of general basis for the interpretation of microscopic perceptions that previously remained unchallenged; an error-free microscopic image, which is assumed to represent in all cases the true structure of the object, a preposition upon which all interpretation of microscopic observation has been previously based as indisputable, nevertheless for a whole class of objects and observations is unjustified.

These theoretical and experimental investigations, the main points which were previously noted, was a practical one, to obtain a safe guide for the proper determination of the conditions for the calculation of lens systems; however, they developed into a complete theory of the microscope, which involves all chapters of micrographic doctrine and adds some new chapters. This theory has a close relationship with the technical construction of the microscope and has proved useful in two ways. On the one hand, the stringent requirements imposed by the practical purpose of the work, have compelled research that no one would have thought to undertake, merely because he was writing a treatise on the microscope; but on the other hand, the actual construction of microscopes according to the principles deduced from theory, has developed the application of the most sensitive tests to which these theoretical considerations could be subjected.

The detailed communication of these studies on the theory of microscope and microscopic perception will appear in a lengthy essay in VIII volume of the *Jena Journal of Medicine and Natural Sciences*. [This never appeared as Abbe died before completing the essay.] However, I assume a concise summary of these results of the investigation will be welcome among practical microscopists, and I have taken the liberty of offering the following to the readers of this publication a brief summary of the principal results of my work. I follow here the same order and directions of investigation as in the more detailed communication, namely, first the discussion that considers aspects of the purely dioptric part of the theory and secondly proceed to consider the new factors that were previously mentioned, and their contributions to the total optical performance of the microscope; however, I

state that the following is not a reproduction of the detailed studies given elsewhere, and in no way it claims to be the full development and establishment of the facts, as will be reproduced in the more extended report.

II. The Dioptric Conditions for the Performance of the Microscope

Sections 4–12 are omitted from the translation.

III. The Physical Conditions for Imaging Fine Structures

13. The performance of the microscope does not always depend on the geometric perfection of the images alone, but in addition to this, for certain classes of objects, on the size of the angular aperture, is a long-acknowledged fact, that has greatly influenced the construction of the microscope in recent times. But what is the real meaning of this fact, nevertheless remained just as problematical as the exact nature of the peculiar quality which has been attributed to the microscope ("resolving" or discriminating power, penetration). In particular, but the question remains how much of this quality is related to the angular aperture in the ordinary scientific use of the instrument, and whether its significance extends any further than to certain particular cases, in which shadow effects were supposed to have been produced by oblique illumination. In my attempt to establish a theoretical basis for the construction of the microscope, it was of primary importance to define the exact nature of angular aperture in the normal performance of the microscope, if I did not want to run the risk of following a mere tradition that might direct my work towards aims that are of very problematical value. I think the results now communicated contain a definite conclusion of the main points in question except as new facts that become known by these results suggest further problems of another kind.

Since it had to happen first of all to determine more exactly, than had been discussed in the micrographic literature, the facts relating to the operation and the effect of the angular aperture, I strive to answer the question first by experiments: in which cases there is a distinct advantage of a large angular aperture, and in which cases no such advantage could be seen, when all other differences which may possibly influence the operation are carefully eliminated? For this purpose, a number of objectives, with very different focal lengths and very different angular aperture, were constructed with extreme accuracy, according to my calculations, and their accuracy checked for correctness; to guarantee the correctness of the observations made with them. The test objects that have been used included: all kinds of butterfly scales and diatom shells, striped muscle fibers, diamond-ruled lines on glass, groups of lines on vanishingly thin silver layers on glass, fine and coarse powder, and besides these, the small optical images of macroscopic objects (bar grating, wire net) which can be obtained by means of bubbles, or better, by objectives of a short focal length, fitted to the stage of the microscope.

14. These experiments found that:

1. As long as the angular aperture remains so large that the effect of diffraction causes no appreciable decrease of the sharpness of the image, no visual improvement arises from a difference in the outlines of the objects, i.e., the

boundaries between unequal transparent parts of the objects, if these parts are not less than 0.01 mm.

2. On the other hand, there is such a difference in favor of larger aperture for all objects showing any detail below the above limits of smallness, regardless of whether this detail is caused by unevenness of the surface or by mere differences of transparency in an infinitesimally thin layer and regardless of whether the detail is in the form of striations, grid drawings, etc. It comes out in the same sense, the mentioned optical images of macroscopic items.
3. The smaller linear dimensions of the detail in question, the larger has to be the angular aperture of the objective, if they are to be perceived with any particular type of illumination, e.g. purely central or possible oblique, and regardless of the more or less marked character of the image, and whatever the focal length and the required magnification of the objective. With a small angular aperture, the image ceases to correspond to the limit of the smallness of the corresponding magnification.
4. Where the detail of real objects in the form of striations, linear systems, and the like appears, for the same angular aperture at oblique incidence of light, and reaches a constant noticeably finer detail than with central illumination, and that is true regardless of the nature of the objects, and the possible shadow effects are completely eliminated.
5. A structure of the supposed kind, which is not revealed by a specific objective used with direct illumination is not made visible by inclining the object itself at any angle to the axis of the microscope, even if, lying perpendicularly to the microscope axis, it is perfectly resolved by oblique illumination. The resolution, follows immediately when the incident light is directed perpendicular to the plane of the objects. Thus, the increased effect under oblique illumination only depends on the inclination of the beam with respect to the microscope axis, but not by the oblique incidence of the light on the object.

The derived facts show, on the one hand the reality of a special optical quality, directly related to the angular aperture of the objective, and independent of any special perfection or magnifying power of the objective, and it is a "resolving" power or capacity, to distinguish the microscopic details, used in good agreement with the literal sense of the use of the traditional designation. On the other hand, they also contain the unambiguous indication that the imaging of very fine structural details must occur by essentially different conditions from those in which the images are formed of the contours of larger parts. In all cases where such a "resolving" power of this kind, that is a direct influence of angular aperture, be it positive or negative operates, the dioptric reunification of rays emanating from several points on the object in the focal plane of the image cannot provide an adequate explanation of the images with such details of an object, because in such a supposition the demonstrated experimental differences previously described would remain absolutely inexplicable. The result of this preliminary study, therefore, has the proper task to give the investigation the following form: namely, to determine the special causes outside of the microscope, which are involved in the formation of

images with small structural details, and then to determine separately the details of the mode and manner of their intervention in the dioptric process. Both requirements have been fulfilled theoretically and experimentally for our present requirements.

15. The wave theory of light demonstrates in the phenomena of diffraction or deflection, a characteristic change according to which material particles, in accordance with the smallness of their dimensions cause to the transmitted (possibly also on the reflected) light beams. This change consists generally in a division of an incident light ray into a group of rays of increased angular dispersion within the range which, periodic intensity maxima and minima within them. For the particular cases of regular layers, striations, point series and the like, the mathematical theory provides a complete determination of the phenomenon, which consists of this, from the rectilinear incident ray there is deflected on each opposite side a series of isolated rays at regular angular intervals from each other; these angular distances. But these angular distances are for each color proportional to its wavelength, and increase steadily from violet to red, and are also inversely proportional to the distances between the particles in the object which cause the diffraction. Therefore, when a microscopic specimen with a particular structure is illuminated by a cone of rays, such as the illuminating mirror of the microscope provides, the light does not enter the objective in the same direct line from the mirror towards the object, because the structure of the object causes a number of deflected and color dispersing rays to be separated from the rectilinear rays, and these deflected rays form larger or smaller angles with the lines of the undeflected rays depending on the higher or the lower fineness of the object's structure. Such objects therefore transmit point for point several isolated light pencils to the objective, the number of which and the arrangement within a specific angular space depends on the position of the mirror and the structure of the object.

This effect, theoretically predicted, and capable of accurate calculation, can be readily observed by examining the aperture images which accompany the images of the object as previously explained. One places an object of the type in question with a microscope so that the detail is in focus, then the ocular is removed and the image of the object is observed with the naked eye down in the open tube, or with a properly set up auxiliary microscope of very low (10–20 times) magnification which in the tube, is situated at the upper focal plane of the objective [back focal plane of the objective]. One sees then the image of the mirror (or what may otherwise be the light illuminating surface), as it is formed by the undiffracted rays, surrounded by a greater or lesser number of secondary images, in the form of impure colored spectra, whose color sequence, from the primary image always goes from blue to red.

Objects with multiple systems of intersecting lines show not only a series of diffraction images of each group in the direction of their perpendicular, the group corresponding to one row of diffraction patterns but also other such series, as to the requirements of the theory, other additional series with the angles between the perpendicular groups; butterfly scales and diatoms show this phenomenon in the greatest variety. The coarser of them permit observation with low power systems of

low angular aperture; the finer, from *Pleurosigma angulatum* upwards, require large angular apertures; in order for the diffraction images that are closest to the main image of the mirror to get into the angular aperture of the objective. For such observation, a weak immersion objective is best suited.[1]

16. The mentioned method for direct observation of the light rays emanating from microscopic objects, allows the experimental determinations of the question of what role does diffraction phenomenon play in image formation of the relevant structure. The answer has emerged readily by focusing the appropriate sample objects, and regulating of the incident light by diaphragms placed just above the objective, as near as possible to its upper focal plane [back focal plane], for the purpose of intentionally excluding one or another portion of the groups of rays that exhibit diffraction effects, the image of the preparation, as formed by the non-excluded rays could be readily observed with an ordinary ocular. The following are the immediate results of such experiments: all of the crucial tests were performed with very weak corrected objectives, 30 to 6 mm focal length, and the corresponding low magnification, but stronger objectives, particularly an immersion objective of 3 mm, were only used to control the results already obtained with coarse objects, by experiments obtained with fine diatoms. The preparations for the crucial experiments were such that their accurately known structures were previously known: various granules of powdered substances, systems of lines scratched into glass, with 0.03 to 0.002 mm line spacing, also and similar groups of lines ruled into silver deposited on glass, the silver coatings were of imperceptible thickness when imaged with the strongest microscope; groups of lines crossing each other with no difference in height were made by overlaying two glasses, their surfaces in contact were independently ruled. The facts thus obtained are:

1. When all the light separated from the incident light beam by diffraction is completely blocked out by the diaphragm, so that only the remaining non-diffracted ray pencils provides an image of the preparation, the sharpness of the contours between unequally transparent portions of the visual field was not affected, as long as the diaphragm remains wide enough as to not bring about its own diffraction, and a concomitant visible decrease of "necessary magnification." This also makes clear that the perception of separate particles is not restricted to any appreciable extent, provided that there are not more than 30 to 50 particles per millimeter. The more this number is exceeded, the more the detail disappears, so that when the fineness of the detail reaches 100 parts per mm, nothing remains visible except a homogeneous surface for any magnification, and for any type of illumination (direct or even

[1]Diatoms have been studied by Flögel (*Botan. Journal*. 1869, no. 48–45) and used to determine the minute stripe distances. The observation method used here can serve very well for the latter purpose, since it is possible, from the measured linear spacing of the diffraction spectra in aperture image, regardless of the direction of the incident light, to calculate the stripe distance when the focal length of the objective is known. It is sufficient for an ocular micrometer in the microscope to observe the aperture image. If you use very intense illumination of a small area, one can obtain the spectra that is so sharp that very regular striations such as on *Pleurosigma angulatum*, and even some Fraunhofer lines are visible.

oblique illumination). Already a couple of lines ruled on glass, for example, two diamond scratched lines or two lines in a silver layer under the designated conditions, appear indistinguishable as a broader scratch with sharp contours. With the strongest immersion objective, the figures on *Pleurosigma angulatum* are not the slightest detected and the coarse longitudinal stripes on *Hipparchia janira* remain imperceptible even at 200× magnification. In the case of granular objects and other irregularly shaped particles, the diffracted light cannot be completely separated from the undiffracted light, but accordingly occurs a possible dimming, although not an absolute disappearance of all the particles, but with such a great lack of clarity of the image follows that the finer particles of the preparation fuse into a uniform gray cloud.

2. If all the light rays are extinguished except for a single pencil of diffracted rays, as this provides a strong positive image the particles in the object in dark field, but without any detail. (Ruled lines appear to be uniformly clear flat stripes on a dark background.)

3. When at least two separate light pencils enter into the microscope, the image shows sharply defined detail, whether in the form of one or more-line systems or in the form of isolated fields. It is immaterial whether the undiffracted light is included or not with the incident cones, i.e. whether the image appears on a bright or a dark field. Other light pencils in operation result in the appearance of fresh details, but always different, according to the degree of their fineness or the nature of the markings; and this detail is not necessarily consistent with the microscopic image seen in normal illumination, nor the well-known real structures of the objects as determined in other ways. In respect to this last point the following particulars are noteworthy.

4. A simple series of lines is indeed always imaged as such, when two or more light pencils go into effect, but in doubly, or triply fine, when among those consecutive order of position, one, two, or more intermediate light pencils are jumped over. Thus, a group of only two lines in the object appears to be composed of three, four … separate sets of lines. The phantom lines thus generated cannot be distinguished with the help of smaller magnification from the normal image of actual lines of double or triple fineness, either with respect to sharpness, or in terms of their constancy of appearance, as may be shown by a conclusive experiment, in which falsely doubled image appears next to the image of an object that is actually ruled with lines of double fineness.

5. When two simple grids in the same plane intersect with any specific angle, the systems may, by appropriate regulation of the light pencil, both single line systems can be made visible simultaneously, together or separately, and by changing the form of illumination, many new line systems and various shaped fields appear with equal sharpness of delineation, which do not exist in the object. The line groups that emerge always correspond in position and distance from each other with the possible forms in which the points of intersection of the lines of the real object can arrange themselves into equidistant series.

A cross grid with perpendicular intersection shows of two secondary lines in the directions of the diagonals, whose distance from the distance of the actual lines

appear smaller as the ratio $1:\sqrt{2}$. Also, four more groups in the proportion of $1:\sqrt{5}$, each of which are inclined at an angle of 27° to a real lattice direction. When a network intersection at an angle of 60° occurs, there appears besides several smaller systems of lines, there is present a third system with lines just as strong as the real network of the object, inclined also at 60° to the others, and when one sees all three simultaneously generated, there will be seen completely demarcated hexagons, of the kind seen in *Pleurosigma angulatum* instead of the rhombic fields. I would add here, that the appearance of structure, which does not correspond to that of the known objects, is always observed with exactly the same focusing in which the image appeared well defined, and that they occurred with the same constancy with a variety of combinations of objectives and oculars, if the illumination is regulated in the same way. The influence of diffraction, which may be caused by the diaphragms above the objective, was eliminated by control experiments.

More detail about these phenomena cannot be given without an exact addressing of the theoretical laws of diffraction. For the same reason I leave here, the summary between the facts just mentioned and the under discussed optical behavior of small images. On the other hand, still had highlighted the following:

6. The partial exclusion of light pencils entering from the object (purposefully arranged in the previously described experiments) occurs unintentionally in the ordinary use of the microscope when observing very fine microscopic structures; for as soon as the detail of objects in its linear dimensions is a small multiple of the wavelength of light, even a very large angular aperture of the objective can never collect more than a small part of the many groups of diffracted light pencils at one time. This part, however, is always varying, both when the angular aperture is larger or smaller, even if the direction of the illuminating rays remains unchanged; or according as the illumination is changed, and the angular aperture remains constant. Upon this fact rests every modification in which the images of minute structures change with varying angular aperture or different directions of the illumination rays. The invariable increase of resolution by oblique illumination, namely, that both the appearance of new detail as well as the marked emergence of new detail than was discernible with central illumination, is in all cases produced by the entrance of diffracted light pencils into the larger angular aperture (with oblique illumination), which would otherwise remain outside the objective because of their greater divergence, or because of diffraction light pencils which were only imperfectly taken up when direct illumination was used, now fully enter more completely into the microscope and work with greater effect, while the direct rays are less effective. In addition to this, there frequently occurs during usual observations, accidental moments of oblique illumination which may produce the effects described in the paragraph 5. As a result, all objects with sets of similar and homogeneous sets of striations present, through a mere change of the incident light, several more sets of lines are made visible in different directions, provided that the angular aperture of the objective used has a suitable relation to the fineness of the striation, as is prominent in the case of several diatoms. Even the types of illumination which produce these effects (such as described in paragraph 4, can occur unintentionally.

In this way, for example, with high power objectives and specific positions of the mirror, the appearance of fine longitudinal lines between the coarse longitudinal stripes on *Hipparchia janira* can be visualized.

17. The facts presented here are sufficient when taken in conjunction with the unquestionable tenets of the wave theory of light, to warrant a series of the most important conclusions, which both affect the microscopic vision, as well as the construction and use of the microscope. First, the consequences in the first direction:

Any part of a microscopic preparation, which are either by their isolation (individual threads, granules and the like) or from its large dimensions relative that of the light waves, produces no appreciable diffraction effect, and is imaged in the field of the microscope according to the usual dioptric laws of concentrating of rays in a focal plane. The image purely negative, being dependent on the unequal transmission of light with partial absorption of the rays (e.g. as staining), or divergence of the rays by refraction, or diffraction of the rays by particles with internal structure. The absorption image thus formed is necessarily similar to the object itself, and can be interpreted with correct stereometric rules, and yields a completely safe conclusion on the morphological composition.

All finer structures, however, the elements of which are small and close enough to evoke an appreciable diffraction phenomenon, will not be geometrically imaged, the image will not be formed point for point, as is usually described by the reunion, in a focal point or a plane, of light pencils, emanating from the object, and which undergo various directional changes during their entrance and passage through the objective; for even when all the dioptric conditions are completely satisfied, the resulting image shows none of the fine structural details, unless at least two of the diffraction light pencils which are derived from the splitting up of the rectilinear rays are reunited.

For anyone who clearly realizes what is the assumption upon which the commonly accepted idea of similarity between an object and its optical image, the facts are sufficient to conclude that under the circumstances previously described, such an assumption is a completely arbitrary hypothesis. For everyone who makes the conditions significantly upon which the usual assumption of what is said must be sufficient to conclude. As a positive instance contrary to the arbitrary conclusion, are the previously mentioned experiments 4 and 5, on closer discussion of their results, lead by rigorous deduction, namely, that different structures always deliver the same microscopic image as soon as the difference of the diffraction effect connected to them is artificially eliminated from the action of the microscope; and that similar structures always yield different images when the diffractive effect taking place in which the microscope is artificially made dissimilar. That is to say, the images of structure that the result from the diffraction process have little constant relation to the real structure of the objects causing them, but rather are in constant relation with the image formation of the diffraction phenomena.

This is not the place to address a more detailed physical explanation of the phenomena. However, it is noted that the conclusions deduced from facts drawn from observations, find their full justification in the wave theory of light, which

shows not only why microscopic structural detail is not shown by dioptric law, but also how a different process of image formation actually occurs. It can be shown that the images of the illuminating surface, which appear in the back focal plane of the objective (the direct image and the diffraction images) must each represent, at the corresponding point, equal phases of oscillation when each single color is separately examined. These aperture images, therefore, relate to each other in the same relation as the two mirror images of a flame in Fresnel's interference experiment. The meeting of the rays issuing from them must cause, as a result of interference, a periodic alternation of light and darkness, whose relative form and dimensions depend on the number, arrangement, and the mutual distance between the interfering illuminated surfaces. The delineation of structure visualized in the field of the microscope is in all its characteristics, both in those which conform with the real nature of the objects, as well as in those which do not conform, is nothing more than the result of the process of interference occurring where all of the image forming rays encounter each other.

The mentioned sentence on the relationship between the linear distances from the axis of the microscope of elements in the aperture image, and the various inclination of the incoming rays entering the objective (explained in section 4), combined with the scholarly dioptric analysis of the microscope, as discussed in section 6, yield all the data necessary for the full implementation of the previous conclusion. It can be deduced from them, that in an achromatic objective the interference images, for all colors, coincide, and must give a general achromatic effect, which is different from all other known interference phenomena. Furthermore, that the proportionate dimensions of the thus generated images will always depend in such a way on the real structures, as the linear magnification the microscope would bring, according to the dioptric law of image formation, and (apart from possible differences of the diffraction effects) for each arrangement of the optical components or the type of illumination. All the facts of observation as they are given in section 16, are not only fully accounted for; but it is possible to calculate beforehand, and in all details, the delineation of structure image by any particular object, under any specific illumination, at a certain illumination, if only the effective diffraction phenomenon is given, namely, the number, arrangement, and the relative brightness of all the diffraction spectra.

18. The end results of these observations are: all those visible phenomena in microscopic images, that are not already accounted for by the simple absorption image, but for which the involvement of diffraction beam groups is needed, in fact all minute structural details, is as a rule, not imaged geometrically, that is, conformally with the actual structural detail of the object. However constant, strongly marked, and apparently visible, such structural elements may appear (fringe systems, drawn fields) they cannot be interpreted as morphological, but only as physical characters; not as images of material forms, but as indications of material differences of the composition of the particles forming the object. And nothing more can be safely inferred from the microscopic findings than the presence in the object that the existence of such structural peculiarities as are necessary and sufficient for the production of the diffraction phenomenon on which the images of minute details depend.

The smaller the linear dimensions of the structural elements are, the fewer in number will be the diffraction light pencils which can be effective even with the largest possible angular aperture; the less effectively can the differences of intensity in the series of these diffraction rays bring into view such structural differences as are still possible within the same relation of dimensions; and the more indefinite will be the conclusions from the image, or even from any diffraction phenomenon, with respect to the true structure of the object.

From this point of view, it is evident that all attempts to determine the structure of the finer types of diatom shells by morphological interpretation of their microscopic images, is based on invalid assumptions. Whether for example the *Pleurosigma angulatum* valve has two or three striation systems, or whether striation systems exist at all; whether the visible image is caused by isolated elevations or depressions, no microscope, however perfect, and no amount of magnification, can provide this information. What we can only say is the existence of the optical conditions necessary for the diffraction effect which is involved in the process of image formation. As far as this effect is visible in any microscope (six symmetrically located spectra, inclined by about 65° (for blue light) against the direction of the undiffracted beam, with ordinary direct illumination), it can proceed from any object which contains it within its structure, or on its surface, optically heterogeneous elements, sorted in some way into a system of equilateral triangles 0.48 μ on each side. Whatever these elements can be, organized particles or pure differences in molecular organization (condensed matter), they will always yield the familiar visible form. It does away with any reason to suppose that structural element in question can be designated as elevations or depressions, after which it is established that this has nothing at all to do with shadow effects on the visibility of the markings, or in their greater prominence with oblique illumination.[2] The distribution of light and dark on the surface of the diatom in the position of a system of hexagonal fields, appears as the mathematically necessary result of interference between the seven isolated light clusters, which is produced by diffraction, whatever may be their physical conditions of the object causing this diffraction; the position of fields of hexagons, two sides parallel to the midrib, has its sufficient reason to be visible in the diffraction spectra towards the axis of the diatom, and can be derived by calculation, without any need to know anything about the actual construction of the objects.

The existence of the same situation for a large number of pure organic structures is the work area of the histologist, in essence, we can learn from the example of the striated muscle fiber. On good preparations, thereof, the diffraction phenomena can

[2]The changes in the image of *Pleurosigma angulatum* during raising and lowering of microscope tube upon the object, proves nothing about the presence of elevations on its surface, because the same changes occur in the same way when diamond ruled lines on glass are observed. In addition, the observation of a sharply defined light source through a *Pleurosigma*, after the repeatedly mentioned method, there is no observable divergence from the refraction of the light rays which are transmitted; the diatom shell behaves exactly like a glass plate with parallel flat surfaces.

quite easily be experimentally observed and studying their effects in the microscopic image by the ways described earlier. The manifold changes in the characters of the image that were anticipated, explain to some extent the familiar discordance between the findings of different observers with respect to these structures, but also bear witness to the impossibility of interpreting in any satisfactory manner their real material composition of the tissue in terms of past attempts.

What has been argued here regarding the basics of microscopic perception, applies further, not only the morphological relations of objects, but equally the other properties which are to be deduced from the microscopic observation. That differences in transparency and colors which one perceives in the microscopic image, are not to be necessary features of the objects, but often arise from the total or partial exclusion of diffraction light clusters is sufficiently well illustrated by well-known appearances of the diatoms. But it seems important to point out that the polarization characteristics of the images of objects with microscopic structure in several views must be differently interpreted from the pure geometrical or absorption images. To conclude interpretations, concerning the substances birefringent character is at least very dangerous. For it is the open possibility that same texture conditions which cause the diffraction may simultaneously induce polarization effects, that as a mere function of the diffraction phenomenon, do not depend as in crystals, granules and the like, upon a peculiar transmission of light rays. For there remains the open possibility that same texture conditions which cause the diffraction, may at the same time cause polarization effects that, are a function of the diffraction phenomenon, do not depend, as in crystals, and in coarse granules and the like, on the transmission of the rays of light per se. That something of this kind actually occurs seems likely, from what I have seen through some perceptions of *Pleurosigma angulatum* and other diatoms, observed polarization with light, show modifications of diffraction light pencils, modifications which are likely to be difficult to explain in another way. However that may be, in any event, it is inadmissible in an object such as the striated muscle fiber, for example, whose structural detail is not dioptrically imaged, to conclude by ordinary criteria, from observations of changes in the diffraction pattern in polarized light, that the various adjacent and close birefringent layers elements show alternating character of simple and double refraction; for if a homogeneous birefringent material were present with a sufficient differentiation of structure to generate the existing diffraction effect, then the appearance of striations would arise from the polarized diffraction light pencils, and demonstrate just such modifications as the muscle fiber when seen in polarized light.

19. In connection to the previous conclusions, which are critical for the scientific use of the microscope, it appears furthermore that the limits for the resolving power are determined for both each individual objective, as well as for the microscope as a whole.

No microscopic particles are resolved (nor the characteristics of a real existing structure can be perceived) when they are located so closely together that not even the first of a series of light pencils generated by diffraction can enter the objective simultaneously with the undiffracted light rays. From this it follows that for every

degree of angular aperture there must exist a specific minimum distance of separable elements, which cannot be specified in exact figures, for the reason that this minimum differs for every color due to the unequal wavelengths for light of different colors, and also because the relative significance of each color greatly varies in the observation. Taking any specific color as the basis, the relevant minimum value is found, for pure central [direct] illumination, by dividing the wave length with the sine of half angular aperture; and half of that product when, other circumstances being equal, for the highest permissible degree of oblique illumination as the objective will admit regardless of its aperture. Therefore, even with the immersion objectives, the angular aperture can not, by any possible means, be increased beyond the degree that would correspond to 180° in air, it follows that, as the microscope might be further perfected in relation to the workable magnification power, the limit of resolving power can not be stretched beyond the figure denoting the wavelength of violet light for central illumination, nor beyond half that amount when extreme off-axis [oblique] illumination is used.

The latter limit is in fact already achieved for direct vision in the finest known diatoms and in the last line of the groups in Nobert's plate and, as far as direct vision is concerned. Only for photographic recording of microscopic images can the resolution be still noticeably richer. For because of the significantly shorter wavelength of the actinic radiation, the conditions for the photographic recording of images are much easier for every objective; namely, they present an image, in the ratio of 3 : 2 larger in detail than is seen by the eye.[3]

IV. The Optical Power of the Microscope

20. The previous research provides a foundation for accurate determination of the nature of the functions that make up the optical power of the microscope, and also for the rational definition of the claimed performance expected from the current optical combinations.

The distinction so long recognized in the micrographic literature, between "defining" and "resolving" powers receives, through the facts and evidence presented; a far wider significance than has been attributed to them by previously known facts. According to these documents, apart from two obvious exceptions, it appears that the microscopic image in general consists of a superposition of two images, equally distinct in origin and character, and can also be actually separated and examined from each other, as shown by the experiments described in the previous section. One is a negative image, in which the parts of the objects present themselves geometrically, due to the unequal emergence of light, which is caused by their mass unequally affecting the transmission of the incident light rays. One can just call this image the "absorption image," because partial absorption is the principal cause of the different amounts of emergent [transmitted] light. It is the carrier of "defining power" whose magnitude, is determined according to the terms

[3]For this reason alone, apart from all others, the performance of an objective in photography does not express the real measure of its performance in the ordinary use of the microscope.

of this kind of mapping, is determined solely by the greater or lower degree with which the direct incident light beam is brought to homofocal reunion, the conditions under which these kinds of images are formed. Thus, it is always this direct light cone, as given by the illumination source, which defines, independent of the direction in which it arrives at the objective, i.e. immaterial whether a central or a peripheral part of the free aperture receives it. Independent of this "absorption image," those parts of the object which contain internal structure, will be imaged again in a positive image, because the parts will appear to be virtually self-luminous, as a result of diffraction phenomenon that they caused. This second image, which may be called the "diffraction image" consists of, strictly speaking, as many partial images as there are isolated diffraction light pencils which enter the objective, since each produces a positive image as shown in the previously mentioned experiments, but because these partial images, taken individually, are devoid of content, the visible detail is only generated by merging two or more (their fusion into one image); their combined effect must be regarded as an independent factor. This resulting diffraction image now appears clearly as the carrier of "resolving power," or the discriminating or separating power of the microscope. Its development therefore depends first and principally on the angular aperture, in so far as this alone, according to above mentioned rules, determines the limit of its possible performance. But the actual amount will also depend on the perfection in which the partial images corresponding to the individual light pencils of diffraction that merge together; because only through the latter, that the detail which indicates existence of positive structural elements in the object is made visible. However, since the individual ray bundles, whose confocal reunion is a necessary condition for the formation of diffraction images, occupy different parts of the free aperture, and vary constantly in position depending on the character of the objects and the type of illumination, as can be assessed by looking into open tube of the microscope, it is obvious that a perfect fusion, in every case, of the several diffraction images, and then an exact superposition of the resulting "diffraction image" upon the "absorption image," is only possible, if the objective is uniformly free from spherical aberration for the area of its angular aperture.

21. Under previous ideas about the process of image formation in the microscope it could be assumed that residual aberrations would only affect the sharpness of the image definition in the objective, and that such aberrations either did not exist, or they are practically irrelevant, as long as there was no visible failure of image definition.

The verified circumstances, taken in connections with what was said of the typical form of the spherical aberration in objectives with large angular aperture, place its importance in a very different light. The individual elements of the microscopic image, both the "absorption image" as well as the various constituents of the "diffraction image" are produced by isolated light pencils of relatively small angles of divergence, almost never over 30–40°. Even with considerable residual spherical aberration can the points of such isolated light pencils, each considered by

itself, be sharp enough to leave almost no appreciable circle of confusion. Since however, with a large angular aperture these individual light pencils operate in the various parts of the free aperture simultaneously, their focal points cannot re-unite if the residual spherical aberration is considerable, but must occur behind and next to each other. The constituents of the entire image therefore do not lead to a correct fusion, but they rather are longitudinally and laterally offset from each other. Marks of structure which occur in the object at one place and level, for example various fringe systems, therefore, appear separated from each other, as well as separate from the contours of the object parts to which they belong. As a result of bias in recent times, in which the perfection of the microscope was focused on the increase of angular aperture, the conditions for such abnormal appearances, especially the presence of deceptive level differences are produced, occur in newer high power objectives in the most prolific way, as many experiences have taught me; and I do not err when I utter the opinion that the consequences of this situation play an unexpectedly large role in many disputed questions among microscopists, concerning the interpretation of confounding microscopic structural relations.

Since everyone must admit that the first and most imperative claim that can be made, which represents the consideration for the scientific use of that microscope upon the performance power of the instrument is this; that parts of the object that belong together will appear as belonging together in the microscopic image, it follows that the uniform correction of spherical aberration throughout the entire area of the aperture must be the absolute criterion and the ultimate guide for the construction of microscopes. Now it has been demonstrated, as mentioned under (7), that for a dry objective an adequate compensation of spherical aberrations will also be actually impossible when the angular aperture exceeds 110°. Therefore, this leads to the conclusion that a dry objective [in air] may be less suitable for the normal scientific use in proportion as it images fine systems of up to a distinctive boundary about 0.35 μ for oblique illumination. The greatest possible increase in resolving power can rationally only occur by means of an immersion objective, since only the immersion objective grants the possibility to increase the angular aperture arbitrarily, i.e. up to the limit of what is technically executable, without opposing the first requirement of corrected spherical aberration.

22. Following this presentation some suggestions are given with respect to the methods of testing complete microscopes.

According to the formerly known facts, as it could justify to evaluate the value of an objective by the smallness of the ultimate details, which can be visualized, and then to consider the resolution of difficult test objects of known type, as proof of the highest performance. Because it is admitted that the particular type of details in these test objects, and the specific type of illumination applied to them are not used in ordinary work, so it seemed certain that success under these peculiar circumstances of the same characteristics of the construction which gave good results will also occur in ordinary work. This now needs to be rejected for the above reasons. An examination method that ascertains the maximum limit of "resolving power,"

whether it was connected at the Nobert's test plate,[4] a diatom or the like, leads to a very exceptional direction of rays of light into the microscope, as required for this purpose by virtue of the physical conditions of the problem. But which is not needed to be repeated for any other operation; the detail only approaches the limit of resolving power when it is so fine, and causes so strong a splitting of the light by diffraction, that even under the best of circumstances, only the first deflected light pencil can enter the objective simultaneously with the direct cone of rays. When it becomes visible in the image, this is only accomplished at the outermost peripheral zone of the aperture of the objective. The most oblique incident light beam cone that the mirror provides, strikes the edge of the free aperture on one side, and the single diffracted light pencil that gains access strikes it on the other side; as you can prove by direct observation of the tracks of both light pencils in the back focal plane of the objective. Theory and experience teach us that every objective which is not a total failure, as imperfect as it may be with its correction of spherical aberration, but when its lenses are moderately well centered can always be made to work aberration free for a single zone, for example, for the outermost edge, and this permanently, if during its construction it was tested; and if it is fitted with a correction collar (adjusted during its use to emphasize the periphery of the objective), an arrangement which is often used for this purpose, rather than for its ostensible purpose, which is correction for the thickness of the cover glass.

Evidence that a lens system can resolve the very fine stripes on a diatom skeleton or on the Nobert's test plate says, strictly speaking, nothing more than that its angular aperture corresponds to the calculated diffraction angle of the relevant line distance on the test object, and that it is not so poorly constructed that a sufficient correction of the marginal zone is impossible. A test of this kind gives no means to determine what are the conditions for the correct fusion of the aperture images such an objective would present in the much less favorable case of the usual observation conditions, where one or more zones in various parts of the free aperture are simultaneously active. The same result can therefore not even make the claim, of the resolving power after its more general characteristics. There is only the limit of resolution, and thus establishes a fact, which may have a certain value in itself, because of the singular circumstances of the case, but which has no direct connection with the general performance of the objective.

Nor can the test of resolving power by direct light be estimated at a higher value. In the vicinity of the limit of resolution corresponding to this type of illumination, all of the direct light is transmitted through the central zone, and all of the diffracted through the peripheral zone of the free aperture. Apart from the circumstances that residual aberrations can be moved to the inactive middle zone of the objective, with

[4]Nobert was the first to publish his technique of ruling lines on glass for the specific purpose of testing microscopes (Nobert, 1846). Abbe used Nobert's method to make a series of test plates that consisted of sets of ruled lines of various line spacings. These sets of lines were ruled on a glass plate covered with a thin coating of silver. The light is transmitted through the lines where the silver is removed in the ruling process, and the remaining area that is covered with a very thin coating of silver is opaque and does not transmit light.

the help of a correction collar, the fact of resolution depends, even in this case, essentially on the action of the peripheral zone, because there always lie, at least two, or more, oppositely situated diffraction light pencils in the periphery, which even without the cooperation of the direct rays, make the details visible.

23. From the point of view of the theory presented here, however, another method is given, which, allows using the usual test objects, brings directly to light the particular points that mainly influence the performance quality during the normal use of the microscope. When it is desired to test, in a critical manner, the conditions for exact interaction of radiation pencils, which pass through the various parts of the free aperture, there is in fact no better tool than those natural objects as diatom skeletons and some butterfly scales; provided that the fact of resolution in itself is not made the chief consideration, but the detailed characteristics of the resulting entire image must be considered. If you select namely from a sample object of such fineness of detail that an objective to be tested, enables it to be seen even with direct illumination, and therefore it can be seen without the slightest problem with oblique illumination, it can be made then without any further aids, to enable seeing the sensitive optical path of the microscope, the production of which is effected by the testing method mentioned in (10), where an artificial test object is illuminated with two separate light pencils. The deflection of the first diffraction pencil obtains in this case such a relation to the angular aperture of the objective, that as the theory and the direct observation of the light tracks show, by setting the mirror at two specific directions, parts of all zones of the free aperture, each represented by individual lines of light are effective, and which particularly favor the emergence of any correction deficiencies. One mirror position would be when it is placed with its inner edge just outside the axis of the instrument, the whole mirror is then on one side of the axis and its surface is positioned perpendicularly to the lines in the object, so that the track of the direct illumination appears in the aperture image [back focal plane], eccentric and close to the center of the angular aperture, while the track of the diffracted ray would appear on the opposite side in the peripheral zone. The second position is that of greatest oblique illumination which allows the objective to function with no noticeable darkening of the field of view; as soon as the change of illumination is accomplished, the two tracks simply swap their previous positions. In both cases there would be, if only one set of striations is present, two separated light pencils which would occupy a part of the central and a part of peripheral zone of the free aperture, both on opposite sides of the axis at the same time. However, if the preparation contains several uniform sets of striations, although additional diffraction light pencils would pass through the objective; nothing essential is hereby amended to the conditions previously considered.

It obviously cannot be the purpose of such a test to demonstrate that each particular fault in the image forming quality of the objective in detail, as can be done with the method mentioned in (10), but to test in a general way the actual performance of an objective that would represent the normal state in the ordinary use of the microscope. The factor of chief practical importance shows itself when attention is given to the degree of correctness with which the merger of several partial images belonging to the same part of the object takes place. We have the

outline image of the objects, formed by the direct light beam, and at the same time a detailed structural image, which arises from the interference of the diffraction light pencils. In a corrected objective, each image stands out perfectly sharp, but both images should perfectly coincide without level difference and without any lateral displacement, that is both images are clearly visible, with one setting of the focus, when in the object outline and the structure are on the same level. If a lens system works satisfactorily under this kind of test, performed with a few turns of the fine adjusting screw, or at least if it performs well in the middle of the field, we can be sure that it will give correct images of any object, and with any kind of illumination. To the contrary, when the objective is focused for the contour image, the details appear to float above or below the object or to float away laterally from the contour image, a construction of the objective is indicated, which makes no guarantee that details which belong together in any preparations will be recognized as belonging together in the microscope image, however high is the resolving power of the system with the usual method of testing it. Besides, our judgment of the good qualities of a lens system, if we confine ourselves not only on the two previously designated forms of illumination, but also other mirror positions tested and their effect proved, attention being always to the characteristics of the fusion of the partial images. But in every experiment, it is hardly necessary to observe that the effective optical path must be controlled through direct observation of the aperture image.

For each objective of a large angular aperture, deviations of the kind indicated will be observed by the edge of the visual field, unless the visual angle of the eyepiece is unusually small. They mainly arise not from aberrations, but from the differences of magnification that are inevitable in the best objectives. The degree of their emergence measures the imperfection of the image formation outside of the axis.

What further belongs to our judgment about the quality of a lens system may be gathered from the examination of the colored fringes that appear on the contours of the objects, shown in the center and at the edge of the visual field. It should be noted that these types of aberrations, because they occur mainly on contour images, have a practical significance only when, with the usual use of the microscope and with direct illumination, they emerge predominately with a central position of the mirror.

For the test objects to be suitable for use in the above specified trials, they must satisfy two particular requirements. First, they must be thin and even, so that the outline borders and structure detail can be seen lying in the same level; secondly, the diffracted light pencils must have a high light intensity, so that the effect caused by them can be better appreciated in comparison to the effect of the image produced by direct illumination. For the last reason, only suitable dry objects with strong, markings, are best suited for these tests, as they always give, as can be seen in the aperture image, the brightest diffraction phenomenon, simply because the interference of intensely bright rays is required to cause strong contrasts of bright and dark parts in the microscopic image.

For the low and middle power objectives, one must use insect scales and coarser diatoms, which the micrographic manuals list, a sufficient number of suitable objects are available; for the higher power objectives however, the selection is very limited by the consideration of the first mentioned requirement [thin and even preparations]. *Pleurosigma angulatum* corresponds at best with respect to the fineness of detail as in the angular aperture of immersion systems; it is in fact quite useful even for the highest powers when delicate specimen fragments with sharply fractured edges are selected; the natural edges offer more than the lines of the midrib as a guarantee of equality of levels. To test the stronger dry systems, the coarser copies of the same objects, also in fragments, can still be used; the markings for an angular aperture of 100 degrees is a bit too fine. Furthermore, you can use broken fragments of the finer scales of *Hipparchia janira* whose crossed striae are suitable for the angular aperture of 80–90 degrees, the same also for higher amounts without disadvantage.

According to my experience in conducting an investigation of this kind, a very safe judgment of the quality of an objective can be formed after a little practice with the method recommended here. At least for higher optical powers the amount of optical capacity with regard to the functions that are independent of the amount of angular aperture, can be more accurately estimated than is possible with tests of defining power and "resolving power" analyzed separately.

What this method does not give, namely, the absolute limit of the physical discriminating power, may just as well be obtained by direct calculations from measurements of angular aperture as by direct observation of test objects. For the detailed method by which the angular aperture is obtained at any desired level of accuracy, I refer to the *Jena Journal of Medicine and Natural Sciences*. [This never appeared as Abbe died before completing the essay.]

24. In summary, some general conclusions in relation to the construction of microscopes, which draw on the facts and theories previously presented can be stated.

According to the given evidence, the optical capacity of the microscope depends on two factors which are rooted in two quite different elements of construction. The first is the geometric accuracy of the ray paths. It determines, by the magnitude of the dispersion circles on the image, the size of the smallest details, which can, on a purely geometrical basis, find expression in the image. The second factor is indicated by the capacity of the objective to meet the certain physical conditions, to which the repetition of such details is certainly connected with the integration of the light pencils split up by diffraction, without which the image remains empty of content. As on geometric principles, a detail is not imaged if its magnitude remains below that which corresponds to the diameter of the dispersion circles, reduced to the linear dimensions of the objects, so on physical grounds detail is not imaged when the angular dispersion of the diffraction pencil is so large as to make their re-union impossible (even for only two diffraction light pencils). Although the condition for both functions are alike, as has been shown, rooted solely in the objective alone; but they are rooted in quite different elements of its construction. The dioptric limit of resolution, caused by the inevitable defects of the focal union

of the light rays, can be found in the serviceable amount of magnification of the objective, and, as stated earlier, for any given level of technical perfection of the construction, and is inversely proportional to its focal length; the physical limit of resolution, however, depends solely on the angular aperture and is proportional to the sine of half the angular aperture. However, both functions are directed to the same goal, namely the visualization of particles of matter that fill infinitely small space, and both are equally indispensable for this purpose. It follows, then, that a rational construction of the microscope must aim for a balance of powers in such a way to have the limits of each approximate harmoniously. For it is evidently just as useless, to extend the physical conditions of resolving power, to an amount greatly in excess of that which can be utilized by any attainable and still serviceable magnification, as it would be, to increase the magnification beyond that which that amount of resolution requires. In the first case, where the angular aperture is in excess for the serviceable magnification corresponding to the focal length of the objective, there is a latent resolution power, which does not function in the human eye; in the second case, when the magnifying power of the objective exceeds the dioptric limits of resolution greatly beyond that which the detail accessible to the angular aperture of the objective requires, an empty magnification, that is to say, one in which the proportionate detail is absent is the consequence.

25. The consequences these considerations lead to certain rules for the proper proportion between focal length and angular aperture at the objective, which contradict in many points the habits of the past practice. These remarks provide the supplement to the statements made in (9), which seems to me to be of more general interest, because it shows the extent and the limit of the microscopic observation.

Since the theory requires a limitation of angular aperture of approximately 110° for all dry objectives, the calculation of minutest detail that remains accessible to such an objective is readily made; and suggests that at least for a rational constructed objective of this kind, in which resolution is not unfairly exulted at the expense of the real perfection of the objective, there can be no question of detail that a trained eye cannot already recognize at a good 400–500× magnification. The claims which may be made according to the present state of the optical construction technique, such a magnification can be attained in an objective even when a focal length of about 3 mm (1/8 Engl. inches). With immersion system objectives, the physical limit of resolution, even if the angular aperture is brought to the highest technically attainable value, does not extend so far that a 700–800 times magnification would not fully equal to it, and this magnification would be achieved with ease with a good construction with approximately 2 mm (1/12 inch) focal length. It may be admitted that a magnification exceeding the minimum given here, as theoretically necessary, may greatly facilitate observation and make it more certain if the additional magnification can be made as correct as possible, although it would not enrich the perception of new facts. One can scarcely attribute this value to the significance of this empty magnification beyond the stated limits, and I therefore

Fig. 6.1 A sculpture with Ernst Abbe's famous equation in Jena, Germany

come to the conclusion that the scientific value of objective, whose focal length, if dry system, is much less than 2 mm, or if for an immersion system, is much less than 1 mm, is quite problematical.

The actual capacity of the microscope, in the strict sense of correct and useful power, is in my opinion, exhausted at these limits, as long as no factors are brought forward which are completely outside the scope of the present theory. In particular, I hold the following view, there exists no microscope in which there has been seen, or will be seen, any structure which really exists in the object, and is inherent in its nature, that a normal eye cannot recognize with a sharply resolving immersion objective with 800× magnification. Recent reports especially from England of extraordinary performance, of unusually high-power objectives, 1/80 English inch focal length, are not of such a character as to induce me to change my opinion and lead me into similar error. The superiority of such objectives is said to have been proved upon objects to which the results of my observations directly apply, and which are said to appear under such magnification, as will be recognized to be entirely illusory by everyone who can understand and give an account of the optical conditions of such performances (Figs. 6.1 and 6.2).

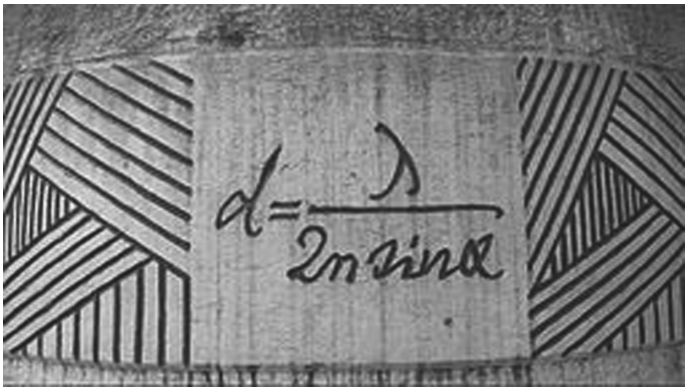


Fig. 6.2 Detail from Fig. 6.1 showing Abbe's equation for resolution in the microscope, where d is the minimum resolvable distance, λ is the wavelength of the illumination, n is the refractive index of the medium between the object and the front lens of the microscope objective, and α is half the angular aperture of the microscope objective

6.3 Commentary on Abbe's 1873 Publication

My English translation "Contributions to the Theory of the Microscope and the Nature of Microscopic Vision" covered the following parts of Abbe's "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung": Part I. The Construction of the Microscope Based on Theory; Part III. The Physical Conditions for the Imaging of Fine Structures; and Part IV. The Optical Power of the Microscope. In my translation I omitted Part II. The Dioptric Conditions for the Power of the Microscope, which comprises sections 4–12.

6.3.1 The Key Points in Abbe's "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung"

To summarize, I now present a concise listing of Abbe's main points:

1. The old empirical microscope construction methods were without a theoretical basis. They can be replaced by a new physical theory of image formation in the light microscope. Abbe demonstrated that geometrical optics alone was not capable of explaining the operation of the light microscope and its resolving power.

2. A new theoretical analysis of the microscope must include new factors and is based on the generally neglected physical process of light diffraction.
3. The resolution of the light microscope is related to the angular aperture of the microscope objective. To resolve features in the object of small linear dimensions a microscope objective with a large angular aperture is required.
4. The wavelength of illumination affects the resolution of the microscope. Shorter wavelengths of illumination in the microscope result in increased resolution.
5. The wave theory of light predicts image formation in the microscope that is due to diffraction of the light by the object. The diffraction image can be observed in the back focal plane of the microscope objective.
6. The effects of diffraction on microscope image formation can be experimentally studied using masks or apertures to stop one or other portion of the groups of diffracted rays from entering the microscope objective.
7. When at least two separate diffraction orders enter into the microscope objective, the image shows sharply defined detail.
8. The detailed structure visualized in the field of the microscope is the result of interference in which all the image-forming rays encounter each other. Abbe used the word interference at least 17 times in his publication.
9. The limits to resolving power are determined for each individual microscope objective based on the wavelength of illumination and the numerical aperture of the microscope objective.
10. The resolution of large details in an image resulted in an absorption image or "dioptric image" that carries the defining power. Diffraction resulted in the resolution of fine details. The final microscopic image consists of two superimposed images formed by two different physical processes. Abbe erred and promoted this idea, which was not supported by the physics of image formation in the light microscope. After 1901 Abbe no longer held onto these aspects of his theory, which involved the superposition of two different images formed by different physical processes (Carpenter and Dallinger, 1901).
11. The resulting diffraction image carries the resolving power, or the discriminating or separating power of the microscope. It depends on the angular aperture, which determines the limit of microscope resolution.
12. It is the obliquity of illumination light with respect to the axis of the microscope —not with respect to the plane of the object—that is critical for increased resolution with oblique light.
13. Abbe correctly claimed that the physical limit of resolution in the light microscope depends solely [at a fixed wavelength] on the angular aperture and is proportional to the sine of half the angular aperture.
14. The actual capacity [resolution] of a microscope, in the strict sense of correct and useful power, were in Abbe's opinion set at these limits, as long as no factors invoked were completely outside the scope of the present theory and the

conditions that Abbe described. This last comment on what today is called the diffraction limit of the resolution of a lens or an optical system, sometimes denoted a diffraction-limited optical system, required further explanation. The finite aperture of a lens is related to its resolution limit, the so-called diffraction limit. In Part III of this book I describe and compare various far-field super-resolution microscopes that achieve resolutions that exceed the diffraction limit. But it is important to note that these superresolution microscopes are based on optical and molecular spectroscopic techniques that are different from the scope of the theory and the conditions that Abbe carefully described in his seminal 1873 publication (Abbe, 1973a).

Abbe's 1873 theory of image formation in the microscope, as presented in his seminal paper "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," contained a significant theoretical misunderstanding that requires further discussion. In Abbe's conception of image formation in the light microscope there were two types of images involved. First, Abbe thought that the absorption image or the dioptric image was the source of the defining power of the light microscope. Second, Abbe thought that the diffraction image carried the resolution power of the light microscope. Abbe erroneously believed that the resolution of large details in an image was the result of what he called an absorption image or a dioptric image, and that the process of light diffraction was responsible for the resolution of fine details (Bradbury, 1996). Abbe thought that the fusion of these two distinct images formed the final image observed in the light microscope.

After a series of discussions with English microscopists Abbe modified his theoretical views (Bradbury, 1996). In 1901 Abbe finally publicly recanted his earlier incorrect views on his theory of image formation in the light microscope based on diffraction. I quote Abbe's words that appeared in the eighth edition of Carpenter and Dallinger's book *The Microscope and Its Revelations*: "I [Abbe] no longer maintain, in principle, the distinction between the 'absorption image' (the direct 'dioptrical image') and the 'diffraction image', nor do I hold that the microscopical image of an object consists of two superimposed images of different origin or different modes of production" (Carpenter and Dallinger, 1901). Note that Abbe used the terms absorption image and dioptrical image interchangeably in his seminal paper of 1873.

In summary, Abbe's 1873 theory of image formation in the light microscope can be expressed as follows: (1) The object diffracts incident illumination light and forms a diffraction pattern in the back focal plane of the microscope objective. (2) The diffracted light beams interfere in the image plane and form the microscopic image. (3) The light that was separated by the physical process of diffraction recombines in the image plane to form the image.

6.3.2 *Dissemination and Understanding of Abbe's Theory of Image Formation in the Microscope*

In this section I discuss the putative sources of the difficulty English microscopists had in understanding Abbe's 1973 theory as well as significant events that took place and English publications that served to mitigate the lack of understanding and therefore the acceptance of his innovative theory.

But, first, I digress with some thoughts on how science progresses with the birth, validation, dissemination, and eventual acceptance of new theories, measurements, and observations that eventually displace former theories and concepts. Progress in science depends on several factors. New experimental or theoretical knowledge must be widely disseminated in a form that can be understood by others. It should be published in peer-reviewed scientific journals that are read by the author's intended audience. It should be published in a language that can be read by a wide audience or translated into the language that is understood by specific groups in a specific country. The inventor of the new theory or the investigator who obtains new knowledge needs to explain it to disparate groups of people in various countries and through lectures, demonstrations, and precise and accurate publications. The reception of new scientific knowledge depends on all these factors. Moreover, there are the problems of bias, nationalism, religion, prejudice, and the supreme difficulty people have in departing from their traditional and long-held values and beliefs and in accepting new ideas once they have been communicated, argued, discussed, verified, and widely validated. Since science is a social phenomenon it is also subject to the foibles of human beings.

Now I return to Abbe's 1873 paper on image formation in the light microscope. To understand the context of Abbe's 1873 paper on the theory of image formation in the microscope we only have to look at the German journal where it was published: M. Schultze's *Archiv für mikroskopische Anatomie*. This was not a physics journal, but a journal whose readers were botanists, anatomists, and histologists (Feffer, 1994). Abbe's choice of journal was consistent with his aims: the improvement of light microscopes made by Zeiss Werke and getting a bigger market share based on the scientific foundation of their design and manufacture.

First, I address the putative sources of problems with accepting Abbe's theory of 1873 (Bradbury, 1996). The English translation of Fripp was the prevalent source of information on Abbe's theory in England (Fripp, 1874a). Fripp's faulty translation with his omissions and changes as well as his added preface and postscript with its negative evaluation of Abbe's theory was a primary problem. Second, microscopists in Germany and on the continent were typically scientists and physicians. But English microscopists were typically laypersons who dabbled in microscopy and worked in other fields (Bradbury, 1996). The lack of understanding of diffraction may have hindered the understanding of Abbe's theory by lay microscopists in England.

Furthermore, some of the blame for his theory of 1873 being misunderstood should be laid at the door of Abbe himself. First, Abbe propagated a false concept in his paper—the idea that the final image observed in the light microscope was composed by the merging of two images, each formed differently and each related to different aspects of resolution: the dioptric image or absorption image related to the resolution of large features and the diffraction image related to the resolution of very small details (Abbe, 1873a). It took 28 years (after his publication of 1873) for Abbe finally to admit his error (Carpenter and Dallinger, 1901). But, for almost three decades it added to the confusion and misunderstanding of his theory. Second, Abbe's publication of his theory recounted the use of masks and apertures that formed narrow cones of light. But the sole purpose of these narrow cones of illumination were for the demonstration purposes of Abbe's Experiments. Abbe never advocated the use of these narrow cones of light for the normal use of the light microscope. By way of contrast, Abbe advocated the use of high numerical objectives to enhance the resolution of the light microscope. Since Fripp in his translation of Abbe's paper stressed the use of narrow cones of light for the microscope it led to ever-more confusion and misunderstandings, with the main result being the lack of acceptance by English microscopists of Abbe's theory of image formation in the light microscope.

Next, I describe the efforts made by Abbe and members of the Royal Microscopical Society—such as John Ware Stephenson, treasurer of the Society, who often corresponded with Abbe concerning the microscope—to mitigate the problem of misunderstanding and the lack of acceptance of Abbe's theory of 1873.

In 1876 Abbe went to London with his instruments to demonstrate what was later called Abbe's Experiments or the role of diffraction in image formation in the light microscope. Abbe began developing these demonstration experiments prior to his 1873 publication, which described them in detail. Abbe made his demonstration experiments with his apparatus in front of several members of the Royal Microscopical Society. Among the observers who witnessed Abbe's experiments was Stephenson (Feffer, 1994). In 1877 Stephenson published a paper that carefully described and explained Abbe's Experiments (Bradbury, 1996). Stephenson's publication contained images of the objects, the masks used to restrict the diffraction orders that were permitted to enter the microscope objective, and an image of the diffraction pattern in the back focal plane of the microscope objective (Stephenson, 1877). Since Stephenson was a witness to the demonstrations of Abbe's Experiments his publication provides the earliest and the only eye-witness account of these demonstrations complete with images that strongly support Abbe's theory of image formation in the light microscope. Therefore, I decided to reproduce Stephenson's short publication including his images to help the reader better understand Abbe's theory.

The timeline of efforts made to improve the clarity of Abbe's 1873 publication is noteworthy. Once Abbe realized that his theory of image formation in the light microscope based on diffraction was very poorly understood and accepted in

England he traveled to England in 1876 to demonstrate his experiments in front of members of the Royal Microscopical Society. In 1877 Stephenson published a paper showing images of the objects, masks, and diffraction patterns that he witnessed during Abbe's demonstrations (Stephenson, 1877). Then, in 1878 Crisp, secretary of the Royal Microscopical Society, published another paper showing images from Abbe's demonstration experiments that provided a very clear description of the content of Abbe's 1873 paper (Crisp, 1878). Although two members of the Royal Microscopical Society, Stephenson and Crisp, understood Abbe's 1873 publication and his demonstration experiments, many other English microscopists were confused by Abbe's theory and did not accept its premises and application to the microscope (Feffer, 1994). Finally, in 1889 Abbe published a publication in English with the title: "On the effect of illumination by means of wide-angled cones of light" (Abbe, 1889). After reproducing Stephenson's 1877 publication in Sect. 6.4 I will discuss the works of Crisp and the short paper by Abbe in 1889 that led to improved acceptance of Abbe's theory of image formation in the light microscope.

6.4 Stephenson's Paper on Abbe's Experiments Illustrating Abbe's Theory of Microscopic Vision

Abbe was an honorary Fellow of the Royal Microscopical Society, London. He wrote several of his manuscripts in English and submitted them to the organization, which published them in its *Journal of the Royal Microscopical Society* (Abbe, 1879a, b). For example, in 1879 volume 2 of the *Journal of the Royal Microscopical Society* contained two papers by Abbe written in English: the first with the title "On Stephenson's system of homogeneous immersion for microscope objectives" and a second with the title "On new methods for improving spherical correction, applied to the correction of wide-angled object-glasses [microscope objectives]."

In late 1876 Abbe traveled to London to attend the International Exhibition of Scientific Instruments. He wanted to demonstrate instruments he had designed for Zeiss Werke and to evaluate the optical instruments of other instrument companies. Abbe took with him to London several instruments of his own design: spectroscopes, refractometers, and several types of Zeiss microscopes. In addition, Abbe brought with him the instrument he used to demonstrate what was later called Abbe's Experiments. These experiments were described in Abbe's 1873 paper. They were designed to demonstrate his theory of image formation in the light microscope based on diffraction.

In this section I reproduce Stephenson's (1877) entire paper because of its importance in helping explain Abbe's 1873 paper.

Stephenson, J. W. (1877). "Observations on Professor Abbe's experiments illustrating his theory of microscopic vision" (read before the Royal Microscopical Society, London, January 3, 1877). *Monthly Microscopical Journal: Transactions of the Royal Microscopical Society and Record of Histological Research at Home and Abroad*, **XVII**, 82–88.

Transactions of the Royal Microscopical Society
II. Observations on Professor Abbe's Experiments Illustrating His Theory of Microscopic Vision.
By J. W. Stephenson, F.R.A.S., Treasurer R.M.S.
(Read before the Royal Microscopical Society, January 3, 1877)

"In my opinion, the very important theory of microscopic vision which has been enunciated by Professor Abbe[5], has not received, in this country, the attention it pre-eminently deserves. The theory to which I refer, is, that the microscopic images produced by certain objects of minute detail, such as diatoms, scales of insects, and other things, are not simply direct dioptrical images, such as the mere outline of an object, but are the result, in most cases, of the combination, or fusion together, of the central pencil with certain secondary images, produced by the interference of those pencils of light, into which, by diffraction, the incident beam of light is, in passing through the object itself, resolved; in other words, that the principal or central beam of light alone is not sufficient truly to depict fine lines, small apertures, or other minute structural details, but that, as far as resolution is concerned, two or more pencils are always necessary to produce the desired effect. These pencils may, or may not, include the principle or dioptric beam, but where the latter is excluded the image necessarily appears on a dark field.

Further, his contention is, that when from any cause whatever, whether from the angles formed by the intersection of lines or the closeness of the lines themselves, whether from the aperture of the object-glass [objective], or when, by artificial means, the diffraction images, as seen within the body of the microscope, are made similar, the microscopic images themselves will be identical.

The diffraction images of a lined object, in focus on the stage of a microscope, may readily be seen by removing the eye-piece [ocular] and looking down the tube of the instrument. Here, with the light central, and the lines on the object parallel, the coloured spectra are distinctly visible, going off on either side at right angles to the direction of the striae, the most refrangible rays next to the central beam of light. The latter fact is particularly mentioned, as it has an important bearing on the limits of visibility and on the photographic reproduction of microscopic objects.

Professor Abbe has supported his views by some very striking experiments, which appear to me to be a complete practical demonstration of the truth of his mathematical deductions; and by his permission, I propose to exhibit under the

[5]A valuable translation, by Dr. H. E. Fripp (Fripp, 1874a).

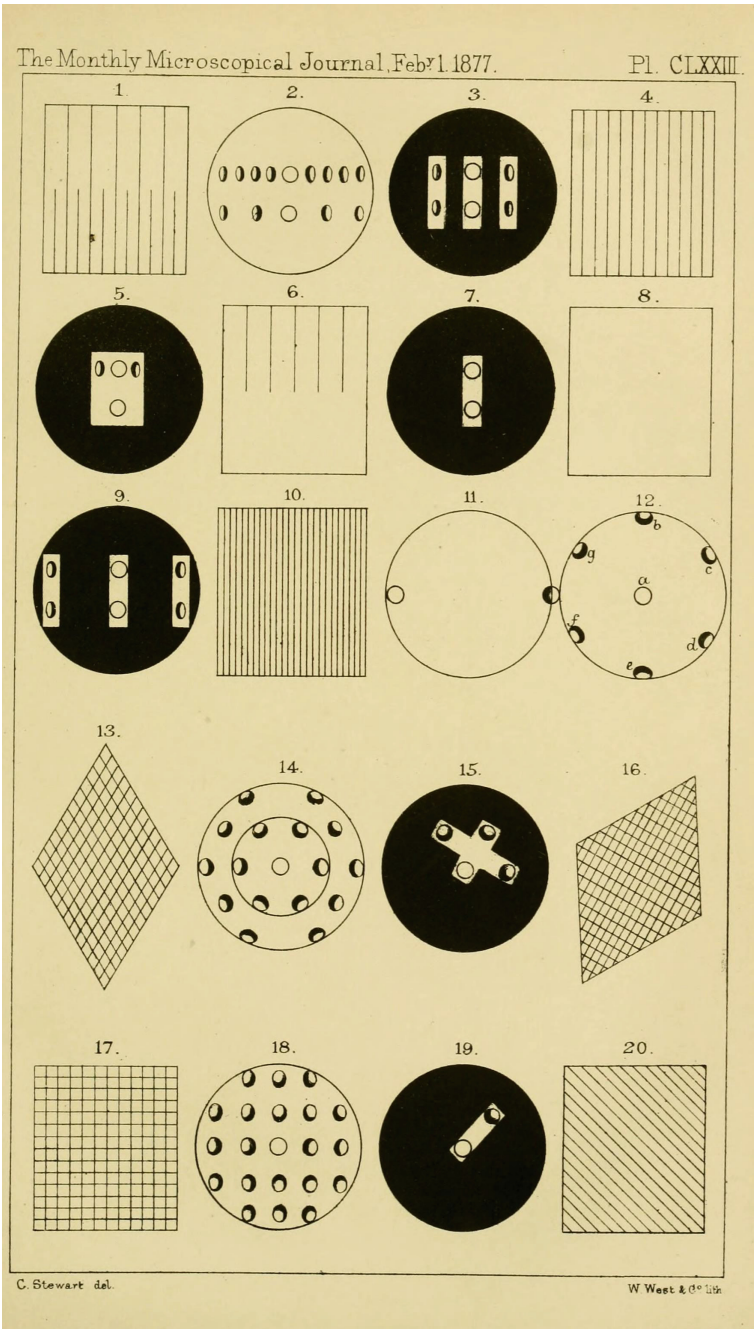


Fig. 6.3 Stephenson's plate CLXXIII on Abbe's experiments

microscope this evening four or five of those which impress me as being the most important, and therefor the most interesting (Fig. 6.3).

Explanation of Plate CLXXIII

Fig. 1. Fine grating used as the object on the stage of the microscope.

Fig. 2. Appearance presented on removing the eye-piece [ocular] and looking down tube, showing central and spectral images.

Fig. 3. Diaphragm with three slits in position, shutting out certain spectral images, and making those produced by coarse and fine lines identical.

Fig. 4. Appearance presented on examination of object (Fig. 1) when examined under the latter condition; fine lines in normal condition, coarser doubled in number.

Fig. 5. Diaphragm excluding *all* the spectral rays from finer lines, and all except the two adjacent to central beam from the coarser lines.

Fig. 6. Shows the coarse striae alone; the finer invisible in consequence of all their spectra being excluded.

Fig. 7. Diaphragm excluding *all* spectral rays.

Fig. 8. No lines visible in consequence.

Fig. 9. Diaphragm excluding all the spectra of Fig. 2, except the fourth of the coarse and second of fine grating.

Fig. 10. Appearance presented; coarse lines quadrupled and fine lines doubled in number.

Fig. 11. Effect produced by light of extreme obliquity on parallel lines of such fineness as to have nearly reached the limit of resolvability; the illuminating pencil at the edge of the field, and only the more refrangible [refracted] rays of spectral image remaining in the field on the opposite side.

Fig. 12. Appearance presented in tube by single valve of *P. angulatum*; light central. [a is the central light in the center of the figure, b-g is the six-fold arrangement of diffracted light in the periphery of the figure].

Fig. 13. Fine crossed grating (60°) used as object on stage of microscope.

Fig. 14. Spectral aspect produced by crossed grating; the spectra within smaller circle identical in form with *P. angulatum*, Fig. 12.

Fig. 15. Diaphragm in position admitting central beam and three spectral rays; the imaginary lines joining them crossing each other at right angles.

Fig. 16. Appearance of Fig. 13 under these conditions; the lines crossing each other at right angles at distances inversely to those of the spectra ($\sqrt{3}$:1).

Fig. 17. Square grating.

Fig. 18. Appearance presented in tube.

Fig. 19. Diaphragm admitting central and one spectral image.

Fig. 20. Consequent disappearance of all real lines, and substitution of diagonal lines at right angles to admitted rays.

1st Experiment. The purport of the first experiment is to illustrate the production of *identical* microscopic images by *different* structures, when, by artificial means,

the diffraction pencils arising therefrom are made similar in number and position, within the tube of the instrument, as previously mentioned.

This experiment is made on a grating formed of alternately long and short parallel lines (Fig. 1), ruled with a diamond through a film of silver, of extreme tenuity, deposited on the underside of a thin glass cover, and subsequently cemented with balsam to an ordinary glass slip, the coarser lines being about 1790 to the inch, and the finer about 3580 [this is the technique of Nobert, first published in 1846].

This grating gives rise to two sets of diffraction spectra, when placed beneath the objective, in the middle of the field, the set arising from the wider portion, being *exactly* half the distance apart of that arising from the narrower, such distances between the spectra, being inversely proportional to the distances between the lines themselves.

On removal of the eye-piece these two rows of spectra (Fig. 2) are visible, one above the other, as the eye is brought to see successively the air images at the upper end of the tube.

It is obvious from the figure, that as the wider grating gives spectra *exactly* half the distance apart, and therefore twice as numerous as those arising from the narrower, that the latter may be made to coincide with the former, in number and position (as required), by stopping out every alternate ray from the wider grating, beginning with the first.

This is readily accomplished by placing a stop close to the back combination of the objective, so constructed that a central slit will admit the central ray only, whilst another slit on each side will admit only the second spectrum of the wider and the first spectrum of the narrower grating (Fig. 3).

On replacing the eye-piece it will now be seen that the microscopic image of the narrower lines remains unaltered, but that the wider lines have doubled in number (Fig. 4), by an apparent prolongation of the shorter lines between them, making the two images identical, the upper part being distinguishable only from the lower by somewhat less brightness, which simply arises from the smaller number of real lines through which the light can pass.

Again, by stopping out all the spectra, except the fourth of the wider, and the second of the narrower (Fig. 9), the spectral aspect is again rendered similar in the two cases, and the microscopic images, though changed, will be still found to be identical, by the doubling of the narrower and quadrupling of the coarser lines (Fig. 10); but to see this distinctly, with so low a power as Zeiss' a, a, the fifth, or E, eye-piece is required.

I should state that, although in the experiment a Zeiss a, a, with a third eye-piece, has been used, any other object-glass of about 1 or 2-inch focus would do as well with a suitable stop; the stop used with the objective mentioned in this experiment has three slits, each 1/20 of an inch wide, with the same distance between them.

2nd Experiment. In this experiment, the same grating is used as before, with a diaphragm having a single central slit, so adjusted (parallel to the lines of the object) that one spectrum only will be admitted on each side, from the coarser grating, and

none whatever from the finer (Fig. 5); the object of the experiment being to show that unless one spectrum at least is admitted there is no power to resolve the lines.

An examination by the eye-piece shows that this is so; by the reduction of the aperture the finer lines (the spectra of which are excluded) have disappeared and been replaced by a plain silver band, the coarser lines appearing in their normal condition, as anticipated by theory (Fig. 6).

3rd Experiment. This experiment, like the former, illustrates the necessity of an amount of angular aperture sufficient to admit some spectral rays; in it the central slit is simply reduced to 1/30 of an inch, which is sufficient to exclude the spectra of the coarser as well as the finer lines (Fig. 7). The examination shows that, even as far as lines 1780 to the inch go, all resolving power has departed, the two gratings being replaced by a plain silver band without any trace of lines.

In all these experiments in which a slit has been employed, it will have been observed that the sides of the slit are parallel to the direction of the lines; but it will be found that if the diaphragm is turned so that the slit is at right angles to the lines, all the spectra will be readmitted and perfect definition result, proving that it is the position of the stops relatively to the striae, and not their form alone, which produces the phenomena.

The ordinary adapter used for rotating the analyzing prism of a binocular microscope is a convenient instrument for adjusting the stops, which may be placed at the end of a small tube of a size suitable for entering the objective when necessary. The same effects of duplication or obliteration of lines may be produced on such an object as *Lepisma saccharina* by using higher powers with suitable diaphragms.

The limit of visibility is a direct consequence of the demonstration that no resolution can be effected unless at least two pencils are admitted; and as the admission of a secondary or spectral image is absolutely dependent on the aperture of the objective, it follows that the resolving power is a function of such aperture, of which we know the superior limit to be 180°; when the limit of resolving power with oblique light has been reached, the illuminating ray will be seen at the extreme edge of the back lens with the spectral image on the opposite margin, as in Fig. 11.

The rule given by Professor Abbe for determining the greatest number of lines per inch which can be resolved by oblique light will be found (taking any given colour as a basis) to be equal to *twice the number of undulations in an inch multiplied by the sine of half the angle of aperture*.

As the sine of an angle can never exceed unity, the maximum will be equal to twice the number of undulations in an inch in that ray of greatest refrangibility which will afford sufficient light for the purpose. With *central* light, the maximum for any assigned colour will be equal to the number of undulations in an inch. What that colour should be is incapable of determination generally, as the capacity for appreciating light varies with different individuals.

If, for instance, we take 43μ in the spectrum as being sufficiently luminous for vision, we find the maximum, as far as *seeing* is concerned, to be 118,000 to the inch; but as the non-luminous chemical rays remain in the field after the departure

of the visible spectrum, a photographic image of lines much closer together than those named might be produced.

How little is gained in "resolving" power by such an excessive aperture is at once seen with it is considered how slowly the sines of the larger angles increase, a reduction of the angle from 180° to 128.33° causing a reduction of 10 percent only in the resolving power, with an immense increase in the general utility of the glass; or, if reduced to 106.25°, we still have a resolving power equal, on the same hypothesis, to 94,400 lines to the inch.

The next experiments are made with crossed gratings, and give equally important results.

These gratings are prepared by ruling two sets of lines through silver films as before, one set being ruled on the under side of a thin glass cover [cover slip] and the other on the slide; the two pieces of glass with the lines in contact at an angle of 60° are cemented together with Canada Balsam, and of course give rhomboid markings over the entire structure (Fig. 13).

4th Experiment. The object of the first experiment with crossed gratings is to show that with a certain arrangement of the incident light both sets of real lines disappear, and are replaced by one set of perfectly distinct spurious lines parallel to a diagonal of the rhombic figure (Fig. 14). This is effected by using a single slit stop, with the slit in the direction of one of the diagonals, when the spurious lines will appear parallel to the other diagonal, and therefore at right angles to the slit.

If a crossed slit, as in Fig. 15, is used, two sets of spurious lines will appear at right angles to each other (Fig. 16), although the real lines, from which they originate, are at an angle of 60°. The reason of this will be at once understood, from the previous experiments, to arise from the admission of two sets of spectra, the directions of which are parallel to the diagonals.

An experiment identical in principle is shown on a rectangular grating, Fig. 17, in which with a slit admitting one spectrum, as is seen in Fig. 19, both sets of lines (vertical and horizontal) disappear, and are replaced by one set (Fig. 20) intersecting the squares formed by the real lines, and therefore closer in the proportion $1:\sqrt{2}$.

5th Experiment. The object of this experiment, which is perhaps the most important of all, is to show that with only one row of spectra, the structure of such an object as that under consideration is absolutely indeterminate.

In this experiment, the slit diaphragms are entirely discarded, and the crossed grating is examined with a simple circular stop, which is used merely for the purpose of so reducing the angle of aperture that the first row of spectra only shall be admitted.

The illumination is central, and an examination of the interior of the tube shows seven pencils of light, the bright dioptric beam being in the centre of the field, with six equidistant spectral rays around the margin.

Let it now be clearly borne in mind that we are about to examine a structure which we *know to be entirely composed of distinct rhombic markings.*

On replacing the eye-piece for this purpose, we see *hexagonal* markings over the entire field, as in *Pleurosigma angulatum*, and this effect has been produced by simply so reducing the aperture relatively to the fineness of the object, that the first spectra only are admitted.

From this microscopic image, we can infer nothing as to the real structure of the object under examination; we know it to be rhombic, but it appears to be hexagonal.

But the bright central beam and six coloured spectra which have produced this result, are identical in aspect with that presented by a single valve of *P. angulatum* with central light. Compare Fig. 12 and the inner ring of Fig. 14.

This diatom with central light, under the highest powers and with the largest apertures, necessarily presents the same spectral appearance, in consequence of the fineness of the striae, or holes (whichever it may be), the dispersion being too great to admit the second row of spectra.

It has now been proved that, with the means employed, no definite inference could be drawn of the real structure of the artificial object, and it is equally certain that this demonstration will apply with equal force to the valve of *P. angulatum*, the hexagonal marking of which may, to use the words of Professor Abbe, arise from "two sets of lines, or three sets of lines or isolated apertures of any shape in the object itself."

If it were possible to admit the second row of spectra, a nearer approach to a knowledge of the true structure would be obtained, as the larger the number of diffracted rays admitted, the greater the similarity between the image and the object, the keystone of the theory being that "the interference of ALL the diffracted pencils, which come from the object, produces a copy of the real structure," as in a dioptrical image; but this, as has been abundantly shown, is rendered impossible by the great dispersive power of many fine structures.

Further illustrations of the formation of hexagonal markings may be found on the same diatom.

On bringing into focus a good specimen of *P. angulatum* flat and with distinct-looking lines, using a broad beam of central light, the six diffraction spectra before alluded to may be distinctly seen within the margin of the back lens of the objective (Fig. 12). Any two adjacent spectra combined with the central cone of light will form an equilateral triangle, and produce the well-known hexagonal markings; but as any other pencils forming an equilateral triangle will also produce hexagonal markings, a new set on a dark field may be formed by excluding the central and each alternate diffraction ray: the sides of this triangle being longer than in the common figure, in the proportion of $\sqrt{3} : 1$, the new hexagons will be three times as numerous as those usually seen, and with their sides at a different angle to the median line. The three pencils producing the interference in this case are g, c, e, or b, d, f, and the hexagons will have their sides normal to the axis of the scale, not parallel as in the common image. Not only is this so, but it follows from the theory that there must be visible three other sets of lines, bisecting the angles between the common lines, and corresponding to the combinations of the spectra, g, c, or f, d—b, f, or c, e—b, d, or g, e. All these phenomena may be observed by stopping off the pencils which are to be excluded. It is easy to get the lines bisecting the angles of the common rows, one after the other, and of these one set parallel to the axis of the

scale. For that purpose, oblique light must be used, and the central beam and one of the peripheral rays must be stopped out, leaving for instance b and f, or c and e (i.e. two spectra parallel to the median line).

In conclusion, I can only express my sincere regret that Professor Abbe's recent visit to London took place during our recess. [In 1876 Abbe traveled to London to visit an exhibition of scientific instruments. He met with several members of the Royal Microscopical Society and demonstrated his experiments in front of them]. Had it been otherwise, the Society would have been gratified by an account of his most important investigations and experiments from his own lips, very much more perfectly than I can possibly have done. But my object has been accomplished if, in bringing before the Society this wonderful contribution to microscopic science, I have induced the Fellows to appreciate the important considerations to which it necessarily gives rise."

6.5 Further Commentary and the English Reception of Abbe's "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung"

Frank Crisp, who was a lawyer, a microscopist, secretary of the Royal Microscopical Society, and editor of its journal, read a paper on the influence of diffraction in microscopic vision before the members of the Quekett Microscopical Club on June 28, 1878, and then published his paper in the club's journal (Crisp, 1878). Crisp's paper contained a very clear summary of the key points of Abbe's theory of image formation in the light microscope, and it contained one figure that illustrated Abbe's experiments (Fig. 6.4).

Crisp assumed that the reader was familiar with the interference of light. He stated that Abbe's theory can be demonstrated as follows (Crisp, 1878). The object is a periodic grating of fine lines that is illuminated with white light. The observer looks down the microscope tube after removing the eyepiece and observes a small circle of light and additional small circles. Crisp explained that the central circle is the direct image of the aperture in the diaphragm and the colored circles are diffraction images. Crisp correctly observed that at least two diffraction orders are required to form an image of the object. The object diffracts the illuminating light into several diffraction orders. Only those diffracted orders that can enter the angular aperture of the microscope objective are observed in the back focal plane of the microscope objective. Finally, Crisp stated that the interferences of all diffracted orders that entered the microscope objective form the image of the object that is seen in the image plane. The image is formed in the image plane by interference of these rays of diffracted light.

Crisp explained that Abbe in the course of his experiments was able to vary the number of diffraction orders that could enter the microscope objective and ultimately form an image of the object. Crisp explained his illustration (Fig. 6.4) to the audience of the members of the Quekett Microscopical Club and he gave a live demonstration with a microscope in order that the members could verify the points presented in Crisp's lecture/demonstration (Crisp, 1878).

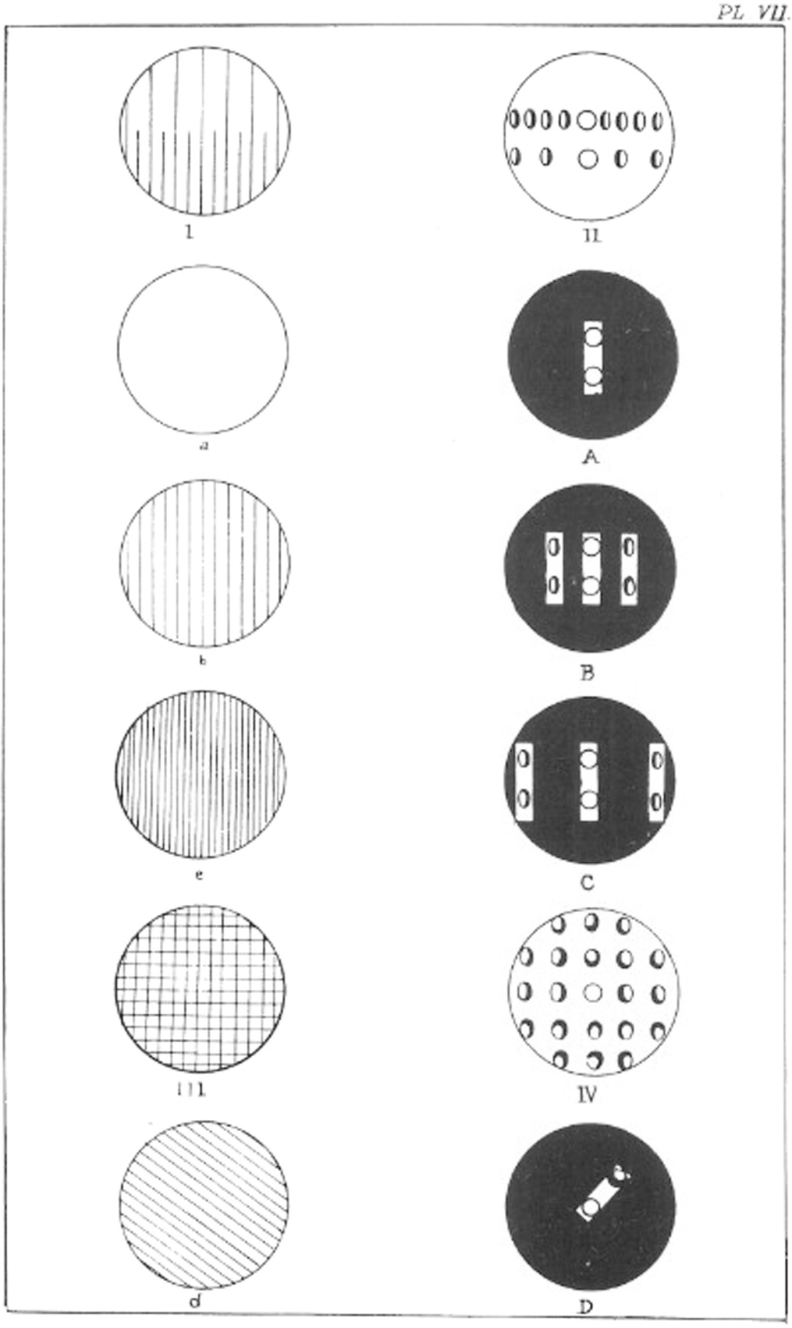


Fig. 6.4 Crisp's Plate VII illustrating Abbe's theory

Crisp described the experiments, the observations, and their interpretations as follows (Crisp, 1878). Figure 6.4 shows an object with two sets of ruled lines. The lower set of lines is twice as close as the upper set of lines. The images on the right side of Fig. 6.4 show what is observed when the eyepiece (ocular) is taken out of the microscope. The central circle is the dioptric beam, and the others are diffraction orders and their spectra. A key observation is that the distances between the spectra are inversely related to the distance between the ruled lines; the upper (widely separated lines) form spectra that are twice as close as those of the lower. The critical role of diffraction in image formation is demonstrated by excluding some orders of the diffracted light using apertures or masks and then observing their effects on the image. Crisp in his lecture and then in his publication uses different masks or apertures and observes the effect on the resolution of the resulting images in the back focal plane of the objective (Crisp, 1878). Crisp's lecture/demonstration helped to explain the foundation of Abbe's 1873 theory of image formation in the microscope based on diffraction.

Abbe realized that his 1873 paper and Fripp's English translation had contributed to misunderstandings and lack of acceptance by English microscopists. A significant misunderstanding was Abbe's use of the term "small angular apertures." Abbe only referred to their use for demonstration purposes and definitely not for general microscopy (Abbe, 1873a). What was needed was a publication that would explain this misinterpretation.

This finally occurred in 1989 when Abbe published a very brief paper "On the effect of illumination by means of wide-angled cones of light" in the *Journal of the Royal Microscopical Society* (Abbe, 1889). This publication was instrumental in getting Abbe's theory of image formation in the light microscope accepted because it helped to clear up and mitigate the confusion among English microscopists concerning the use of wide-angled cones of light for illumination in the light microscope.

6.6 Summary Remarks on Abbe's Theory and Abbe's Experiments

Abbe first published his theory of image formation in the microscope in 1873 in a German publication. In 1874 the first English translation was published by Fripp. As I previously discussed, Fripp's English translation was inaccurate and Abbe's theory ended up being misunderstood and scarcely accepted by English microscopists.

In 1889 Abbe published a paper in English in the *Journal of the Royal Microscopical Society* titled "On the effect of illumination by means of wide-angled cones of light." This publication also helped in understanding and getting Abbe's theory of image formation in the light microscope accepted because it helped to reduce the confusion concerning the use of small-angled cones of light for illumination. Abbe only intended their use for demonstration purposes—not for general use in light microscopy (for which he advocated wide-angled cones of light for illumination).

Abbe's experiments led him to two significant conclusions (Abbe, 1873a). First, geometrical optics could not explain the images of fine structure observed in the light microscope. But, the physical processes of diffraction and interference can account for the imaging of fine structure in the light microscope. Second, while the image is underdetermined by the object itself, it is entirely determined by the diffraction orders that can enter the microscope objective and form the image of the object.

Abbe had an excellent knowledge of geometrical optics, diffraction, and interference (Auerbach, 1922). In addition, he had received strong training in precision measuring instruments (Auerbach, 1922). These two factors permitted Abbe to design and construct a variety of precision measuring instruments. Because Abbe had university training in both physics and mathematics he certainly could be considered a theoretical physicist. I think that the genius of Abbe was his ability to attack a problem; in his case it was to seek a deeper understanding of the process of image formation in the light microscope to improve the design, construction, and accurate testing of microscope objectives and entire microscopes. His ability to design experiments that were critical both to the formulation of his theory and could also be used to demonstrate the validity of his theory follow from his broad interdisciplinary studies (Auerbach, 1922; Feffer, 1994). Further critical discussions of the handbooks, textbooks, monographs, book chapters, and papers relevant to Abbe's "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung" can be found in Appendix A.

References

Abbe, E. (1873a). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.

Abbe, E. (1879a). On Stephenson's system of homogeneous immersion for microscope objectives. Journal of the Royal Microscopical Society, **2**, 256–265.

Abbe, E. (1879b). On new methods for improving spherical correction, applied to the correction of wide-angled object-glasses [microscope objectives]. Journal of the Royal Microscopical Society, **2**, 812–824.

Abbe, E. (1882a). The relation of aperture and power in the microscope. Journal of the Royal Microscopical Society, Sec. **II**, 300–309, [read before the Society on May 10, 1882], [Abbe wrote the paper in English].

Abbe, E. (1882b). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Sec. **II**, 460–473. [read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Abbe, E. (1889). On the effect of illumination by means of wide-angled cones of light. Journal of the Royal Microscopical Society, Series **II**, XI, 721–724.

Auerbach, F. (1918). *Ernst Abbe: sein Leben, sein Wirken, seine Persönlichkeit nach Quellen und aus eigener Erfahrung geschildert von Felix Auerbach*. Leipzig: Akademie Verlag Gesellschaft.

Auerbach, F. (1922). *Ernst Abbe: sein Leben, sein Wirken, seine Persönlichkeit nach Quellen und aus eigener Erfahrung geschildert von Felix Auerbach*. Zweite Auflage. Leipzig: Akademie Verlag Gesellschaft.

Bradbury, S. (1996). The reception of Abbe's theory in England. Proceedings of the Royal Microscopical Society, **31**, 293–301.

Carpenter, W. B., and Dallinger, W. H. (1901). *The Microscope and its Revelations*, Eighth Edition. Philadelphia: P. Blakiston's Son & Co.

Crisp, F. (1878). On the influence of diffraction in microscopic vision. Journal of the Quekett Microscopical Club, **5**, 79–86.

Feffer, S. M. (1994). *Microscopes to munitions: Ernst Abbe, Carl Zeiss, and the transformation of technical optics, 1850–1914*. PhD Dissertation, University of California, Berkeley, 1994. Ann Arbor: UMI Dissertation Services.

Fripp, H. E. (1874a). Contributions to the theory of the microscope and of microscopic vision, after Dr. E. Abbe, Professor in Jena. Proceedings of the Bristol Naturalists' Society, **I**, 200–261.

Nobert, F. A. (1846). Ueber die Prüfung und Vollkommenheit unserer jetzigen Mikroskope. [About the testing and the perfection of our current microscopes]. Annalen Physik & Chemie, III, **143**, 173–185.

Stephenson, J. W. (1877). Observations on Professor Abbe's experiments illustrating his theory of microscopic vision. (Read before the Royal Microscopical Society, London, January 3, 1877). Monthly Microscopical Journal: Transactions of the Royal Microscopical Society and Record of Histological Research at home and Abroad, **XVII**, 82–88.

Further Reading

Abbe, E. (1873b). Ueber einen neuen Beleuchtungsapparat am Mikroskop. Archiv für mikroskopische Anatomie, **IX**, 469–480.

Abbe, E. (1874). Contributions to the theory of the microscope and the nature of microscopic vision. Translated into English by H. E. Fripp. Proceedings of the Bristol Naturalists' Society, **I**, 202–258. Read before the Bristol Microscopical Society, December 16, 1974.

Abbe, E. (1875). About a new lighting apparatus on the microscope. Translated into English by H. E. Fripp. Monthly Microscopical Journal, **XIII**, 77–82.

Abbe, E. (1880). The essence of homogeneous immersion. Journal of the Royal Microscopical Society, **1**, 526.

Abbe, E. (1881). On the estimation of aperture in the microscope. Journal of the Royal Microscopical Society, **1**, 388–423.

Abbe, E. (1883). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Sec. **II**, 790–812. [read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Abbe, E. (1989). *Gesammelte Abhandlungen*, I-IV. Hildesheim: Georg Olms Verlag. [Originally published in 1904, Jena: Verlag von Gustav Fischer].

Born, M., and Wolf, E. (1999). *Principles of Optics*, 7th (expanded) edition. Cambridge: Cambridge University Press.

Bracegirdle, B. (1987). *A History of Microtechnique, The evolution of the microtome and the development of tissue preparation*. Second Edition. Lincolnwood, IL: Science Heritage Ltd.

Bracegirdle, B. (1989). Light microscopy 1865–1985. Microscopy, **36**, 193–209.

Bradbury, S. (1989a). Landmarks in biological light microscopy. Journal of Microscopy, **155**, 281–305.

Bradbury, S. (1989b). The development of biological preparative techniques for light microscopy, 1839–1989. Journal of Microscopy, **155**, 307–318.

Brush, S. G. with Segal A. (2015). *Making 20th Century Science. How Theories Become Knowledge.* New York: Oxford University Press.

Clarke, J. A. L. (1851). Researches into the structure of the spinal cord. Philosophical Transactions Royal Society London, **141**, 601–622.

Czapski, S. (1893). *Theorie der Optischen Instrumente nach Abbe*. Breslau: Verlag von Eduard Trewendt.

Czapski, S. (1906). Geometrische Optik. In: A. Winkelmann, Ed. *Handbuch der Physik,* Zweite Auflage, Sechster Band, Optik. Leipzig: Verlag von Johann Ambrosius Barth.

Czapski, S., and Eppenstein, O. (1924). *Grundzuege der Theorie der Optischen Instrumente nach Abbe*. Third edition. Leipzig: J. A. Barth Verlag.

Dippel, L. (1869). *Das Mikroskop und seine Anwendung*. Braunschweig: Friedrich Vieweg und Sohn.
Dippel, L. (1882). *Das Mikroskop und seine Anwendung*. Braunschweig, Zweite Auflage: Friedrich Vieweg und Sohn.
Fripp, H. E. (1874b). On the limits of optical capacity of the microscope. Translated from Germany publication of Hermann von Helmholtz. Proceedings of the Bristol Naturalists' Society, **I**, 407–440.
Fripp, H. E. (1876a). On aperture and function of the microscope object glass [microscope objective]. Proceedings of the Bristol Naturalists' Society, **I**, 441–456.
Fripp, H. E. (1876b). On the physiological limits of microscopic vision. Proceedings of the Bristol Naturalists' Society, **I**, 457–475.
Gerth, K. (2005). *Ernst Abbe, Scientist, Entrepreneur, Social Reformer*. Jena: Verlag Dr. Buseert & Stadeler.
Gordon, J. W. (1901). An examination of the Abbe diffraction of the microscope. Journal of the Royal Microscopical Society, Ser II, **XXI**, 353–396.
Hammond, C. (2015). *The Basics of Crystallography and Diffraction*, Fourth Edition. Oxford: Oxford University Press.
Hartinger, H. (1930). Zum fünfundzwanzigsten Todestage von Ernst Abbe. Die Naturwissenschaften, Heft **3**, 49–63.
Helmholtz, H. (1874). Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope. Annalen der Physik Jubelband, 557–584, Leipzig. Translated as: "On the limits of the optical capacity of the microscope," Monthly Microscopical Journal, **16**, 15–39 (1876).
Kingslake, R. (1978). *Lens Design Fundamentals*. New York: Academic Press
Köhler, A. (1893). Ein neues Beleuchtungsverfahren für mikrophotographische Zwecke. Zeitschrift für wissenschaftliche Mikroskopie und für Mikroskopische Technik, **10**, 433–440.
Köhler, A. (1894). A new lighting apparatus on the photomicrographical purposes. Journal of the Royal Microscopical Society, **14**, 261–262.
Köhler, H. (1981). On Abbe's theory of image formation in the microscope. Journal of Modern Optics, **28**, 1691–1701.
Lummer, O., and Reiche, F. (1910). *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe*. Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.
Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press.
Masters, B. R. (2009). C. V. Raman and the Raman Effect. Optics & Photonics News, March 20, 40–45.
Nägeli, C., and Schwendener, S. (1867). *Das Mikroskop, Theorie und Anwendung desselben*. Erste Auflage. Leipzig: Verlag von Wilhelm Engelmann.
Nägeli, C., and Schwendener, S. (1877). *Das Mikroskop, Theorie und Anwendung desselben*. Zweite Verbesserte Auflage. Leipzig: Verlag von Wilhelm Engelmann.
Porter, A. B. (1906). On the diffraction theory of microscopic vision. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science (Series 6), **11**, 154–166.
Rayleigh, L. (1896). On the theory of optical images, with special reference to the microscope. Philosophical Magazine and Journal of Science, London, **XLII**, 167–195.
Rheinberg, J. (1901). In "Proceedings of the Society". Journal of the Royal Microscopical Society, **XXI**, 480.
Rheinberg, J. (1905). Obituary. Journal of the Royal Microscopical Society, **25**, 156–163.
Schellenberg, F. M. (2004). Selected Papers on Resolution Enhancement Techniques in Optical Lithography, SPIE Milestone Series, MS **178**, Bellingham: SPIE Press.
Stephenson, J. W. (1875). Monthly Microscopical Journal: Transactions of the Royal Microscopical Society and Record of Histological Research at home and Abroad. Vol. **XIV**, (1875). Extracts from Mr. H. E. Fripp's translation of Professor Abbe's paper on the microscope. Reprinted from the Proceedings of the Bristol Naturalists' Society, vol. I, part 2, 191–201, 245–254.
Varley, C. (1843). Varley on the method of preparing and rendering transparent objects intended to be preserved in Canada balsam. London Physiological Journal, **1**, 31.
Volkmann, H. (1966). Ernst Abbe and His Work. Applied Optics, **5**, 1720–1731.
von Rohr, M. (1940). *Ernst Abbe*. Jena: Gustav Fischer.
Zeiss, C. (1878). Description of Professor Abbe's apertometer, with instructions for its use. Journal Royal Microscopical Society, **1**, 19–22.

Chapter 7
Helmholtz's Contributions on the Theoretical Limits to the Resolution of the Microscope

"The question of if, and to what extent, the optical performance of the microscope is capable of further improvement, is a question of the greatest interest for many branches of science."

—Helmholtz, 1874

"That much of what is familiar among microscopists is almost unknown among physicists, and vice versa."

—Rayleigh, 1896

"It therefore seems that a working knowledge of the phenomena and laws of diffraction might well form a part of the equipment of everyone who uses the microscope and attempts to interpret its indications."

—Porter, 1906

7.1 Introduction

In this chapter I provide a detailed discussion of a seminal paper that independently formulated the equation for resolution in the light microscope. Although the publication of Helmholtz (1874) followed Abbe's 1873 publication, it demonstrated Helmholtz's independent discovery and derivation of the diffraction-limited resolution of the light microscope. Additionally, in the same publication (1874) Helmholtz independently derived his Sine Condition (identical with the 1873 Abbe Sine Condition) from photometric analysis (Born and Wolf, 1999). These two independent derivations by Helmholtz were based on different types of analysis than Abbe's analysis; yet they produced the same final equations. In view of the importance of Helmholtz's German publication I have translated it into English and provided my own commentary.

The main emphasis of this chapter is my translation and commentary on Helmholtz's 1874 German publication. I end this chapter with concise discussions

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_7

of other subsequent publications by Rayleigh (1896, 1903) and Porter (1906) that show the progression of mathematical analysis of physical theory and its application to the microscope.

Abbe's 1873 publication is a detailed nonmathematical treatment of image formation in the light microscope based on light diffraction. It is essentially a presentation of his phenomenological theory. Abbe never published his mathematical analysis of his 1873 theory. In the Appendix A discuss the key monographs, books, and publications that were written by Abbe's colleagues and others between 1867 and 1926. These publications include material from Abbe's lectures on the mathematical analysis of his theory of image formation that he presented at Jena University and further mathematical formulations of Abbe's theory of image formation in the light microscope (Appendix A).

7.2 Helmholtz and His 1874 Publication: On the Limits of Optical Capacity of the Microscope

I pose and then answer the following questions: Who was Herman von Helmholtz (1821–1894)? What role did Johann Benedict Listing (1808–1882) play in motivating Helmholtz to publish his German paper in 1874? Finally, what factors led me to decide to undertake a new English translation of Helmholtz's 1874 paper?

Helmholtz was a polymath, physician, inventor, experimental and theoretical physicist, educator. He is considered a paragon of 19th-century science (Cahan, 1993; Masters, 2010). After his death in 1894 a book of his lectures and speeches was published in five editions. It provides insight into his character and interests in the advancement of science, science education, and the understanding of science by the public (Helmholtz, 1903).

Helmholtz's life and work are a perfect example of the benefits of a broad interdisciplinary education. In addition to his broad studies in the humanities including philosophy, several languages, and the arts he obtained a medical degree (enabling him to gain clinical experience) and a physics graduate degree. These studies and experiences formed the foundation of his career in physiological optics, physiological acoustics, his invention of the ophthalmoscope to observe the living human retina, his work on neurophysiology, and his contributions to the promotion of science education and research (Cahan, 1993; Masters, 2010). To support his physiological research Helmholtz developed many precision physiological instruments thus demonstrating his skill in both theoretical and experimental physics. Here we see a similarity with the precision instruments that Abbe developed for his optics research.

Helmholtz's scholarly output was extraordinary—his books especially so: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (On the Sensations of Tone as a Physiological Basis for the Theory of Music) (Helmholtz, 1877), and the three volumes of *Handbuch der physiologischen Optik* (Handbook of Physiological Optics) (Helmholtz, 1909–1911).

What motivated Helmholtz to publish a paper on the resolution limits of the microscope? Helmholtz worked in the field of physiological optics as did Johann Benedict Listing who was a professor at the University of Göttingen. Abbe was a student in Listing's physiological optics course. Listing thought that he could achieve very high magnifications in the light microscope by forming two real images that were separately magnified using low-power lenses, a common optical method to form an upright image in terrestrial telescopes (Feffer, 1994). Listing posited that this technique was capable of achieving enhanced magnifications of 25,000 to 50,000× in the light microscope.

This extraordinarily high magnification should be compared with the typical magnification of 400–800× in the light microscope. Listing's speculative ideas were published in *Poggendorff's Annalen der Physik* (Listing, 1869a, b; 1871). Helmholtz read these papers and decided to initiate a theoretical investigation into the possibility of such enormous magnifications. This prompted his question: Are there theoretical limits to resolution, which Helmholtz called optical capacity, in the light microscope (Feffer, 1994)? This is another example of a research project that began with a clear formulation of a question?

Why did I decide to translate Helmholtz's 1874 publication into English? Again, I come back to the deficiencies in the translation work of Fripp, which I previously discussed in relation to his faulty translation of Abbe's 1873 publication. Fripp translated Helmholtz's paper, "Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope" into English and published it in the *Proceedings of the Bristol Naturalists' Society* (Fripp, 1874). However, I found Fripp's English translation to be inaccurate.

Both Abbe and Helmholtz arrived at similar expressions for diffraction-limited resolution in the light microscope. Furthermore, they both independently formulated the sine condition that was later called the Abbe Sine Condition.

Helmholtz independently worked on a totally different theoretical approach to arrive at his equation for diffraction-limited resolution in the light microscope. He validated his equation with his own experiments. The theory, design, and execution of the validation experiments differed between the work of Abbe and that of Helmholtz. Having completed his derivation of the limiting resolution of the light microscope Helmholtz used his experiment to show that his derivation was valid. Then and only then did he began to write his publication of 1874. It was only after Helmholtz had finished writing that he became aware of Abbe's 1873 publication. Therefore, Helmholtz added a postscript to his publication in which he acknowledges the priority of Abbe's earlier publication. Fripp did not include this postscript of Helmholtz's in his English translation (Fripp, 1874).

In 1903 J. W. Gordon published "The Helmholtz Theory of the Microscope," which was based on Fripp's 1874 translation of Helmholtz's 1874 paper (Gordon, 1903). Gordon's motivation for this commentary paper was to make the mathematical theory and its associated physical foundation accessible to members of the Royal Microscopical Society. Gordon's publication is replete with line drawings that help the reader to understand Helmholtz's theory.

Furthermore, Gordon pointed out the following errors that occur in Fripp's translation of Helmholtz's 1874 publication (Gordon, 1903). The reader of Fripp's translation is warned that the second line drawing of light rays used to illustrate eq. (6)

was added to the translation by Fripp; it was not in Helmholtz's German paper and according to Gordon it is incorrect (Gordon, 1903). Gordon also pointed out that Fripp incorrectly translated Helmholtz's use of the word *Flächenelement* into the English word "point," which makes the text connected to eq. (6) incomprehensible to the astute reader (Gordon, 1903). Fripp added a five-page preface of his own ideas and modified the original paper to such an extent that I decided to provide my own translation, which is a more accurate translation of Helmholtz's German publication.

7.2.1 *"Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope" (On the Theoretical Limits of the Optical Capacity (Resolution) of the Microscope) English translation by Barry R. Masters*

Translator's note: The equation numbering in the translation follows Helmholtz's 1874 publication.

The question of if, and to what extent, the optical performance of the microscope is capable of further improvement, is a question of great interest for many branches of science. There is some progress by the revival of Amici's suggestion of immersion lenses adopted and carried out successfully by Hartnack, but each step is slow and uncertain. We have, it is clear, now arrived at a point at which any small improvement is affected with a disproportionate effort of mental as well as mechanical labor. And yet, so far as I can see, no one has been able to give any reason why this should be, excepting the common belief that the difficulty lies in overcoming the spherical aberration of lenses so small and of such high curvature as is needed for objectives of very high magnifying power. It is a long time since Mr. Listing, one of the most eminent authorities on this subject, discussed, (*Poggendorff's Annalen* **136**, 467, 473) the means by which it might be possible to obtain magnifications ranging from 25,000 to 50,000 times, while in actual use the ordinary range of useful magnification is at the present moment limited to, from 400 to 800 times.

Moreover, the joint experience obtained by repeated efforts of practical opticians has taught us that all high magnifications combined with good definition are obtainable only through such instruments in which the objective admits a cone of light of very large angular aperture from each point of the object. We have gradually arrived at that stage of improvement in the construction of instruments in which rays of light whose direction is nearly perpendicular to the axis of the instrument are passed into and through the objective, and transmitted towards the ocular. This, it is true, happens only when a lens is used dry (i.e., the front surface is in contact with air). But the rays inclined to the axis at angles up to 87.5° actually

enter a well-built immersion lens. This angle is reduced to about 48°, however, if the instrument is used in a normal way, with a drop of water between the objective and coverslip. This angle is much larger than in the lens system of a telescope or a photographic camera obscura, because with such oblique incidence the spherical aberration, even in the carefully calculated and precisely built lenses of these instruments would be intolerably large. Why then is a broad incident beam in the microscope more advantageous than a narrow one of greater brightness, which provides the same amount of light in the instrument? The recent responses to this question appear unsatisfactory to me. For the so-called "penetration," which means the ability of the instrument to separate particles by light and shadow, whose refractive index differs very little from that of their surroundings, depends only on the ratio between the angular aperture of the illuminating light cone and the cone of points passing from the object into the lens. A sufficiently strong shading can always be formed by the narrowing of the aperture of the illuminating cone; and a comparatively large cone can be applied under the object when cones of light passing from it into the objective are also wide.

In fact, a cause now exists for the compound microscopes, which in this case causes much greater aberration of the beams of the focal plane, than from chromatic and spherical aberration, especially for narrow beams of incident light. This cause is the diffraction of light. If the same is also perhaps occasionally mentioned as a cause of degradation of the image, I have never found a methodical investigation of the nature and the magnitude of its influence. Such an investigation, however, shows that the diffraction of the radiation necessarily and inevitably increases with increasing magnification, and has an impassable limit to the further extension of microscopic vision, which has been closely approached in our newer and best instruments.

That the darkness and diffraction of the microscopic image must increase with increasing magnification, regardless of the specific construction of the optical instrument, based on a general law of optical instruments, having first established for arbitrary combinations of "infinitely thin lenses" by Lagrange[1]. The law remained almost unknown, perhaps because he has it set up in the form of equations whose coefficients are not easy to yield intuitive meanings. I myself have derived the law in a more general form (namely centered systems with refractive spherical surfaces and with arbitrary refracting media between them) in my *Physiological Optics* § 9 p. 50, I derived and tried to give it a simple and clear physical interpretation.

I recapitulate here first the theorem that is briefly described and its proof.

The same is true for any centered system of spherical refracting or reflecting surfaces through which rays pass at small angles of incidence, so fine that as to form punctiform images of punctiform objects; so that means it refracts homocentric rays, homocentrically. I designate a centered system as one in which the centers of curvature of all the refracting or reflecting spherical surfaces lie in the same straight

[1]Sur une loi générale d'Optique. *Mémoires de l'Académie de Berlin*. 1803. Cl. de Mathém. p. 3.

line, the "axis" of the system. In front of such a system and on its axis, suppose there is a luminous point belonging to some object lying in a plane, perpendicular to the axis objects $a'a''$ (Fig. 1), and the beam passes through the optical system, we call the angle between one of these rays and the axis of the system the divergence angle of that particular ray. Any plane supposed to extend through the axis and along the ray, constitutes the incidence plane of that ray at the first refraction, and will include, therefore, the same ray after its next refraction, and thus also after every subsequent refraction. This plane, which is divided by the axis into two halves, we treat one half as a positive and the other half as negative, and accordingly the divergence angle of the beam to be positive or negative depending on whether the beam proceeds towards the positive or the negative half of the plane. These postulates are stipulated, so the theorem is stated as follows:

Theorem

The product of the divergence angle of any ray, and the refractive index of the medium through which it passes at the time, and the magnitude of the image to which the rays passing through that medium belong, in a centered system of spherical refractive and reflective surfaces, remains unchanged at each refraction, if all the conditions for the formation of an accurate image are complied with. This product will therefore have the same value, after the emergence of the rays as it had before they entered into the same system.

Proof. Let $a\ b$ be a piece of the axis of the system, hh' is one of the refracting surfaces, c its center of curvature (point of convergence). Figure 1.

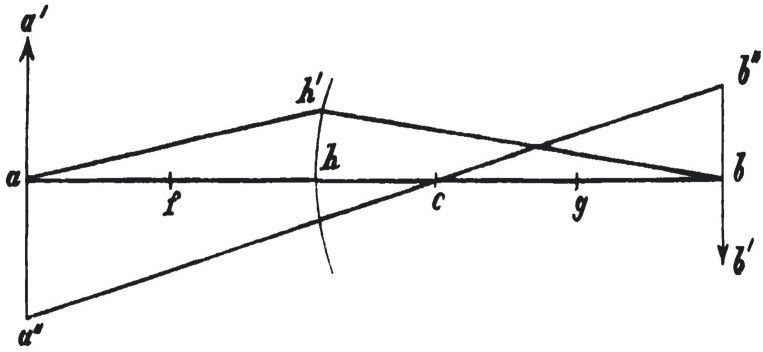


Figure 1

Let a be the point of convergence of the rays that are incident on hh'. Let b be the point of reunion of the rays that are refracted by hh'. Let point f be the front principal focus point, and point g be the back principal focus point. Let n' be the refractive index of the medium in front of hh', and let n'' represent the refractive index of the medium behind hh'. Let α' be the positive divergence angle $h'ah$ of the ray passing in the first medium through h'. Let α'' be the negative divergence angle,

in the second medium $h'bh$. Let β' be the magnitude of the image aa'' belonging to the rays in the first medium. Let β'' be the magnitude of the image bb'' belonging to the rays of the second medium. First, we have the similarity of triangles $aa''c$ and $bb''c$,

$$\frac{\beta'}{\beta''} = -\frac{ac}{cb} \tag{1}$$

Again, if we consider the short arc hh' of the refracting surface as a straight line that is perpendicular to the axis ab:

$$hh' = ah \tan \alpha' = -bh \tan \alpha''$$

or by substituting the angles for the tangents which is allowed because of the small size of the angle:

$$\frac{\alpha'}{\alpha''} = \frac{bh}{ah} \tag{2}$$

Multiplying equations (1) and (2), we get

$$\frac{\alpha' \cdot \beta'}{\alpha'' \cdot \beta''} = \frac{ac \cdot bh}{bc \cdot ah} \tag{3}$$

According to the known laws of refraction at a spherical surface, whose radius $hc = r$, the value of their principal focus is:

$$F' = hf = \frac{n'r}{n'' - n'} \text{ and } F'' = hg = \frac{n''r}{n'' - n'} \tag{4}$$

From which follow,

$$\frac{F'}{F''} = \frac{n'}{n''} \tag{4a}$$

$$F'' - F' = r \tag{4b}$$

Furthermore,

$$\frac{F'}{ah} + \frac{F''}{bh} = 1 \text{ and } \frac{F''}{ac} + \frac{F'}{bc} = 1$$

or,

$$\frac{bh}{ah} = \frac{bh - F''}{F'} \text{ and } \frac{bc}{ac} = \frac{bc - F'}{F''}$$

Division of the last two equations yields

$$\frac{bh \cdot ac}{ah \cdot bc} = \frac{F''(bh - F'')}{F'(bc - F')}.$$

But from equation (4b)

$$bh = bc + r = bc + F'' - F'$$

and also

$$bh - F'' = bc - F'$$

Hence,

$$\frac{bh \cdot ac}{ah \cdot bc} = \frac{F''}{F'} = \frac{n''}{n'} \quad \text{according to equation (4a)}$$

Therefore, equation (3)

$$\frac{\alpha' \cdot \beta'}{\alpha'' \cdot \beta''} = \frac{n''}{n'}$$

or,

$$n' \cdot \alpha' \cdot \beta' = n'' \cdot \alpha'' \cdot \beta'' \tag{5}$$

q. e. d.

It follows from this theorem, first, that when the ray B proceeding from the luminous point has a smaller absolute divergence angle than the ray A, the angle of divergence of B must be smaller after each subsequent refraction than that of A; precisely because the relevant product from our theorem for B is from the beginning smaller than that for A, and therefore must remain smaller after each refraction.

When two outgoing rays, from the same point of the axis, but with equally large divergence angles, following planes which extend in opposite directions through the axis, their divergence angles are equal after each refraction; which, incidentally, is already evident from the symmetry of the system around its axis.

If now we imagine the illuminating rays, on their way to the object, to be circumscribed by interposing a diaphragm pierced with a circular opening whose center coincides with the axial line, the plane of the diaphragm is at right angles with the optical axis, then those rays which pass through the opening close to its edge, all have equally the largest divergence angle, and retain the same relation after each refraction. These rays obviously occupy the exterior outline of cones having a circular base, and whose axis is the optical axis of the lens system, and they constitute the boundary of the cone of light proceeding from the luminous point. The divergence angle of these border rays is, in this case, throughout their entire course, the angle with the semi-aperture of the conical surface bounding the illuminating cone.

From this there follow, first, certain important results in regard to the *photometric conditions* of the microscope image.

Following from the known laws of photometry, we can equate L the quantity of light emitted from the luminous point dS upon another point ds, at distance r as follows where (r, N) and (r, n) are the angles formed between the line r and the normals N and n.

$$L = J\frac{dS \cdot ds}{r^2} \cdot \cos(r, N) \cdot \cos(r, n) \tag{6}$$

If now we understand by ds the circular aperture of the cone of rays at one of the refracting surfaces, and by dS a luminous point that is intersected by the axis so that r falls in the direction of the axis. Then,

$$\cos(r, n) = 1 \text{ and } dS \cdot \cos(r, N)$$

is the projection of dS onto a plane normal to the axis. Let α be the angle of divergence of the rays that are directed to the periphery of ds, then

$$ds = \pi \cdot r^2 \cdot \alpha^2$$

$$L = J \cdot \pi \cdot \alpha^2 \cdot dS \cdot \cos(r, N) \tag{6a}$$

The same amount of light must also be contained in the same cone of rays that continue through the following medium. We designate the corresponding quantities as: J', α', dS', N' then,

$$L = J' \cdot \pi\alpha'^2 \cdot dS' \cdot \cos(r, N') \tag{6b}$$

Now, dS' is the image of ds and its projection normal to the axis $dS' \cdot \cos(r, N')$ is the image of the corresponding projection of dS. Therefore, we have the proportion

$$dS \cdot \cos(r, N) : dS' \cdot \cos(r, N') = \beta^2 : \beta'^2$$

From which it follows:

$$J \cdot \alpha^2 \cdot \beta^2 = J' \cdot \alpha'^2 \cdot \beta'^2$$

and from equation (5):

$$J : J' = n^2 : n'^2 \tag{6c}$$

As a result, the brightness at which the surface of the image included inside the outline of the illumination cone, is independent of the direction that dS and dS' have in relation to the axis, and of their distances from the refracting surface.

From the image of dS' we may pass on to consider a second dS'', and so on. It obviously must exist between each subsequent image and dS a corresponding equation such as (6c). Thus, if $n = n'$, the object and the image lie in the same medium, *then the brightness of the optical image that is produced by rays which are inclined at a very small angle to the axis and will always be equal to the brightness of the object, except for the loss of light due to reflection and absorption.*

But this law should be valid without the restriction that the incident rays make very small angles with the axis. If that is not right, it would possible for a surface element dS, in equation (6), to emit a defined beam, an image dS' which shines towards the continuation of this beam with an increased brightness more than equation (6c) allows, then we could cause the ray pencil to pass on as parallel rays through a plane boundary surface into the air, and to enter the eye of the observer; it would be the case that you should see a object that is brighter when viewed through an optical instrument than it was before, as it is something that is contrary to all experiences with the most varied forms of transparent refractive materials. Would such possibility for light beams occur, then it could also be correctly transferred to thermal radiation, and it would contradict the law of equality between the radiation of bodies at the same temperatures. By similar considerations, as Mr. G. Kirchhoff has applied to the evidence of his law of the equality of the absorption and emission that can be carried over to our case without effort, and it would be easy to prove. One needs in the end, only the two elements dS and dS' of the same temperature and to assume the rings enclosing sheath, the transparent media is free of absorption, and a suitable diaphragm attached, also by same temperature, to delineate compressed beam. Then, if dS by this diaphragm dS' a denser ray beam passes, as this its temperature returns, must the temperature of dS' rise, that of dS fall.

A more detailed version of the law of the divergence angle. Equation (5) put forward infinitely small divergence angle as such it is indifferent if we substitute α, for $\sin\alpha$ or for $\tan\alpha$, or we substitute similar functions in which for vanishingly small α the functions are equivalent to α. Suppose a larger divergence angle of a ray pencil of circular cross section, and the illuminating surface dS is perpendicular to the axis, then:

$$L = JdS \cdot \int_0^{\alpha} 2\pi \cdot \cos\alpha \cdot \sin\alpha \cdot d\alpha' = \pi JdS \cdot \sin^2\alpha$$

If, after a series of refractions the surface dS fully and accurately imaged in dS_1, with the brightness

$$\frac{n_1^2}{n^2} J_1 \text{ and } \alpha_1$$

is the divergence angle, so the same amount of light must be:

$$L = \pi J \frac{n_1^2}{n^2} \cdot dS_1 \cdot \sin^2\alpha_1$$

because

$$dS : dS_1 = \beta^2 : \beta_1^2$$

it follows from these equations:

$$n \cdot \beta \cdot \sin\alpha = n_1 \cdot \beta_1 \cdot \sin\alpha_1 \tag{7}$$

and equation (5) is also valid for larger angles under the condition, that β and β_1 are two mutually exact images whose surfaces are perpendicular to the axis.

Brightness of the images. When the pupil of the observer is completely immersed in the ray pencil emanating from a point of the image β_1, so the observer β_1 sees image brightness as given by the Equation (6c). So if object and image both are in the air, both will be of the same brightness. This conclusion has already been drawn by Lagrange. Unfortunately, he has not investigated the second case, which in common with high magnification is usually not discussed, namely where a penetrating radiation ray pencil does not fill the entire pupil. That may have contributed a little to the oblivion into which his important paper has fallen.

If, as is often, the pupil incident light ray pencil, which always has a small divergence angle α_1 does not fill the pupil when the image β_1 in the correct viewing distance, then the brightness H of the retinal image is considered to be less than the brightness H_0 to the naked eye, whose pupil is completely filled with rays of light. If we denote by s the visual distance, and the radius p of the pupil, as is its area πp^2, the cross section of the light beam $\pi s^2 \cdot \sin^2\alpha_1$, and therefore it behaves as:

$$H : H_0 = s^2 \cdot \sin^2\alpha_1 : p^2$$

or with use of Equation (7)

$$H = H_0 \cdot \frac{s^2}{p^2} \cdot \frac{n^2}{n_1^2} \cdot \frac{\beta^2}{\beta_1^2} \sin^2 \alpha$$

The last medium in front of the eye must be necessarily air, so $n_1 = 1$, and if we denote divergence angle of the instrument with α_0, measured in air after Lister's method, then $\alpha_0 = n \cdot \sin \alpha$. We set the magnification,

$$\frac{\beta_1}{\beta} = N$$

then:

$$H = H_0 \frac{s^2 \cdot \sin^2 \alpha_0}{p^2 \cdot N^2}$$

In that magnification N_0, wherein the light cone just fills the pupil, and what we will call the normal magnification of the instrument must be $H = H_0$. This yields:

$$N_0 = \frac{s}{p} \sin \alpha_0 \tag{8}$$

and if α_0 remains unchanged:

$$H : H_0 = N_0^2 : N^2 \tag{8a}$$

if as was assumed:

$$N > N_0$$

while $H = H_0$, when $N \leq N_0$.

So this means that: *the brightness of the instrument is equal to that of the free [normal day light] eye when the magnification is equal or smaller than the normal magnification. In contrast, with the same divergence of the incident beam, the brightness of the surfaces is inversely proportional to the magnification when it is larger than the normal magnification.*

The magnification increases as equation (8) shows with the sine of the angle α_0. Meanwhile the increased value is one, if α_0 is a right angle, which is reached in the newer instruments. Then:

$$N_0 = \frac{s}{p}$$

If we set s, as usually occurs for the calculation of magnification, equal to 250 mm, and p for bright illumination equal to 1.5 mm, then:

$$N_0 = 166.7$$

and one obtains:

brightness 1/4 for magnification 333.3,

brightness 1/9 for magnification 500.0

brightness 1/16 for magnification 666.7 and so on.

This shows how quickly the brightness decreases with increasing magnification.

Would it be possible to conduct a hemispherical cone of light from an object in water with an immersion lens system and thus to produce a good picture, at these magnifications and with equal brightness, all these relations are increased in proportion: $n{:}1 = 1.3351{:}1$. But as noted above, previous instruments have been constructed only for air, and not in water with a cone of incident light that approaches hemispherical.

The *ray bundle that enters the cross section of the pupil* can be easily determined empirically. If you focus the instrument to a bright field, withdraw the eye from the ocular and extend it along the optical axis of the instrument, and look at the ocular itself. It will be seen in front of it a little bright circular area on a dark background. This is the optical image of the objective which the ocular (mainly the design of the field glass) forms. All the light which comes through the objective, and the ocular transmits, must be collected in this image. This little picture also corresponds to the cross section that several cones of light, transmitted from the brightest points of the object are collected at this point in space. In order to collect all this light and thus to obtain as large as possible and as bright as possible field of view, the pupil must be brought to the location of this little image. The relation between the surface area of this bright little image and the area of the pupil directly yields the relations to the brightness of the object to that of the small image. Only when that little image is equal to or larger than the area of the pupil, you have the full brightness.

With a telescope, the relation between the diameter of the objective and the diameter of the ocular image is equal to the magnification is already discussed by Lagrange, who suggested using this relationship to measure the magnification. In the telescope, it is not necessarily a result of such decrease in brightness with increasing magnification because the quantity of incident light can be increased indefinitely by enlarging the reflecting mirror. At the aperture of the light cone entering the microscope is, on the contrary, definitely restricted due to the limits of the angular aperture.

The course of the previous illustration shows that the relationship between brightness and magnification that is discussed here is completely independent of the particular construction of the instrument, provided only that there are sharply formed images. An increase in the magnification would therefore only be possible using much stronger light, e.g. as directed sunlight, as this has already been in view by Mr. Listing for his proposed methods to achieve very high magnification in observation.

Here, however, there are other obstacles. These are caused by *the very small divergence angle of the emerging rays*, as is required for strong magnification by equation (7).

First, the shadows of entoptic [light falling on the eye makes objects in the eye visible] objects in the eye, which densely fill this region field, ocular image of the objective, at the eye spot, becomes smaller. The retina is illuminated from this area, as if it were the light source that proceeded all light rays that enter the eye. The same is also the basis of the ray pencils from several points of the object and its retinal image, and its diameter grows, as shown before, varies and is inversely proportional to the magnification. The well known conditions which must be met in order to obtain sharp shadows of the entoptic object is very ones that require sufficient light from a relatively small surface should enter in the eye. Anyone who has ever tried to decrease the field of a microscope obtaining strong magnification by using sunlight, knows the peculiar mottled character which is then obtained in the visual field. A portion of the spots is fixed in the field of glare; others move with the eye movements. The former arises from blemishes and imperfections of the previous polishing of the ocular lens, the latter of the cornea, crystalline lens, the vitreous body of the eye. Also, this method of observation is to be seen as a method to view entoptic objects, long known and, in fact, very useful. In proportion, however, as the entoptic objects will become more apparent, there are an increasing number of fine microscopic objects that become unrecognizable.

A second unavoidable disadvantage of narrow light beam lies in the occurrence of *diffraction patterns*, whereby the outlines of viewed objects are blurred and doubled simultaneously or multiplied four times. We have to deal here mainly with the diffraction patterns as they appear when we look through a small circular opening. A bright point of light (sun reflection on a bulb), viewed through such an opening (pinprick on a map sheet), appears as a bright circular disc, which is in turn surrounded by alternate bright and dark rings. The apparent width of these rings, calculated from minimum to minimum, corresponds to very nearly to a visual angle, whose sine is equal to $\frac{\lambda}{d}$, where λ designates the wavelength of the light and d the diameter of the opening respectively. The outer rings have exactly this width, the innermost are a little wider; the radius of the innermost light ring is $1.220\frac{\lambda}{d}$. Since the smallest visual angle at which we can distinguish two fine bright lines from each other can be fixed at one minute, then the figures of the brightest yellow-green light whose wavelength is equal to 0.00055 mm is visible when the diameter of the opening $d = 1.89$ mm. The dispersion of a bright point into a circle, or a bright line into a strip must even be noticeable at larger openings.

If you look through such an aperture at any objects which shows bright points, the diffraction patterns of the individual points of light coincide, so that the circular diffraction pattern fringes of each item by itself, however, will not be recognizable. But apparently, the effect of diffraction, since it turns any bright spot into a small dispersion area, must make the imaging of objects washed out, similar to vision by the small circles of confusion in imperfect accommodation of the eye. Very fine

objects that can only be seen when the retinal image is sharply defined, will then be unrecognizable.

You can verify that this is the case by a simple experiment. The most sensitive objects are gratings with alternating light and dark parallel lines, they are composed of parallel wires or printed as white and black lines on paper. Imagine the retina is such distance from the grating that with the aid of spectacles with a perfect accommodation of the eye, the bars of the grating can just be distinguished from each other. You then move a map sheet before the eye, in which you have pierced, fine apertures of different diameters, and determine if you can see the lines of the grating through these openings just as well as without the map sheet. The lighting of the grating must be very bright, a printed paper, for example, directly illuminated by the sun so that the image seen through the opening remains sufficiently bright. In my experiments, I think, in fact, that a marked deterioration of the image is already produced by an aperture of 1.72 mm diameter. That is much more striking when you look through an even narrower opening.

Instead of the grating, there can also be printed letters used under similar conditions apply, by observing the point at a distance that they can barely be distinguished. Then they are not readable when viewed through an opening of about 1 mm in diameter. But I do not find this sample as sensitive as the grating.

Here it must of course be taken care of to maintain the best accommodation of the eye, because if this is imperfect, then insertion of the map sheet can reduce the dispersion circle on the retina and even improve the image.

The theory for the diffraction of the rays in the microscope, as will be somewhat explained below, now yields the conclusion that a single point of light in the microscopic object, when seen through the microscope, must appear exactly, as if an actual luminous point, situated in the image of the object were seen through an aperture which corresponds in terms of location and size to the ocular image of the respective smallest diaphragm aperture.

It follows, first that the diffraction must become noticeable when the ocular image has a diameter smaller than 1.89 mm, and the width of the dispersion rings produced by the diffraction must increase as inversely proportional to the diameter of this aperture, thus directly proportional to the magnification, when the incident light beam from each point in the object remains unchanged. It is therefore in these circumstances, even with a further increase in magnification, the image sharpness is unchanged due to diffraction, simply because the dispersion circles maintain the same relation to the apparent size of objects. In contrast, the deterioration which follows from the reduction in brightness and the increasing number of darker entoptical shadows increase with the magnification. It follows from this that in those magnifications which show the most detail where the smallest objects that can still be seen in the image, presented at the most suitable visual angle, that is somewhat larger than when the observer can distinguish the ever perceptible smallest objects.

Calculated according to Equation (7), corresponds to the diameter of 1.89 mm of the area of ray pencils entering the pupil, when the light entering the objective, for hemispherical propagation of the incident radiation in air, of a magnification of

264.5 times. For objectives with a narrower cone of light it would have to be set lower. Accordingly, we find in Hugo v. Mohl's *Micrographie* it indicated that magnification between 300 and 400 times permits the most detail to be appreciated, while Harting, discussing the newer microscopes with almost hemispherical formed beam, found magnifications 430–450 was most useful.

If it is a question of determining *the minimum magnitude of the detectable objects* as a measure of the accuracy of the microscopic image, so here are the same reasons as in the determination of visual acuity of the eye, reasons that I have discussed in my *Handbook of Physiological Optics,* p. 217, that you cannot use the diameter of individual bright spots or lines on a dark background, nor dark points and lines on a light background. Because then the result depends not only on the magnitudes of the images, but also on the sensitivity of the eye to weak differences of the light.

The most suitable objects are, here also, fine gratings which show alternately bright and dark bands. They have been widely used for microscopes: Nobert's glass gratings, and the line systems of diatoms and butterfly scales. Since the light of the bright stripes are certainly very much dispersed before it becomes indistinguishable, so there is only certainty on the distances between the center lines of two adjacent white stripes, and much less certainty on the original distribution of the light from the narrower or wider white stripes. So l choose as a measure for the smallest distinguishable objects, that distance of the smallest interspace between centers of two adjacent stripes, in which they can still be perceived as distinguishable.

When diffraction occurs by a square aperture it can be shown that the grating should appear as a uniformly illuminated surface with the fringe width equal to the width of grating intervals. For circular apertures, the integration for the calculation of the light distribution is very complex. When the diameter of circular aperture is equal to the side of the square aperture, the outer fringes in the spectrum of a bright spot are of the same width, but the interior fringes are wider for the circular aperture. Now for the square aperture, the separation of the bright lines of a grating with their centers around the fringe width are so broad as to erase their distinguishability, this will also have to be the case with a portion of the fringes of circular apertures that are slightly wider. I have used below, for circular apertures, the lower limit of distinguishable objects at distances equal to the width of the outer fringes. It is not impossible that by chance, there is a superposition of fringes, then objects of smaller dimensions may be partially seen and partially guessed at being seen. A certain and unambiguous perception of such will scarcely be possible.

If the size of the smallest perceptible distance is ε, the wavelength in the medium of the object is λ, the divergence angle of the incident rays is α and λ_0, α_0, are the previous named values for air. Then by the derived formulas below:

$$\varepsilon = \frac{\lambda}{2 \sin \alpha} = \frac{\lambda_0}{2 \sin \alpha_0}.$$

For white light we set the wavelength of the medium bright rays:

$$\lambda_0 = 0.00055\,\text{mm}.$$

If $\alpha_0 = 90^\circ$ then this gives:

$$\varepsilon = \frac{\lambda_0}{2} = 0.000275\,\text{mm} = \frac{1}{3636}\,\text{mm}.$$

Would it be possible to drive the rays in water at an immersion system up to hemispheric divergence, then $\alpha = 90^\circ$, λ would be equal to $\frac{3}{4}\lambda_0$, and therefore

$$\varepsilon = \frac{1}{4848}\,\text{mm}.$$

After Mr. Harting[2] measurements with a Hartnack objective system Nr. 10, the perceived smallest distances according to our calculation is:

$$\varepsilon = \frac{1}{3313}\,\text{mm}.$$

(Mr. Harting's previously mentioned number 1/5210 refers to the width of the dark interval between the lines.) In the near coincidence with the above are the measurements of Mr. L. Dippel[3] who employed the finest recognizable line systems of diatoms. He came up to 1/3500, on the finest Nobert line system and for finer lines 1/3600. Earlier details of Mr. Sollitt and Harrison[4] from the year 1853, however, go much further. Of the recognizable lines of *Navicula arcus* have been counted at 5120 to a millimeter. These results far exceed the theoretical limits for objects that are in air. However, since all recent measurements are far lower than these, I do not know whether you can consider them reliable. Also Mr. Harting, who cited this information, doubted its correctness.

Besides any possible achieved increase in beam divergence for objects lying in water, the optical performance capacity of the instruments would probably have to be increased by the application of blue light[5]. The wavelength of the line G is 0.0004282, so only about 25/32 of the above applied values of the wavelength of the strongest light. The ratio 1.28:1.00 would be in the above denominator of the values I had to multiply when applying only blue light. This yields for the actually executed immersion microscopes:

---

[2]These *Annals* CXIV.

[3]*Das Mikroskop und seine Anwendung. Braunschweig.* 1867. p. 135.

[4]*Quart. Journal of Microsc. Science* V, p. 62.

[5]Orally, I am told that Mr. Hartnack has already done this with instruments that he sent to the 1873 Vienna exhibition.

$$\varepsilon = \frac{1}{4654}\ \text{mm, instead of } \varepsilon = \frac{1}{3636}\ \text{mm.}$$

In photography, the blue light is mainly active, and the photographs seem to do a little more in fact than the eye can in white light. In a photograph of *Surinella gemma* that Dr. J. Stinde has made with a Gundlach objective at a magnification of 1000×, the lines are visible at 3800 to 4000 per millimeter.

It seems to me, beyond any doubt, that the diffraction of the radiation is the principal cause of the limited sharpness [resolution] of microscope images. In comparison with diffraction the chromatic and spherical aberrations of lenses seem to have only a negligible influence in spite of the very large angle of incidence and angle of divergence of the rays. Compared to the great effort that had to be used in the calculation and design of lenses for telescopes and photographic cameras to reduce the spherical aberration to a sufficiently small magnitude, it is surprising that lens systems of microscopes made to prescribed dimensions, and with the large angular aperture of its light cone, that spherical aberration made so little effect. Besides, I've pointed out that if there is water between the object and the cover glass as well as between the coverslip and the objective, the divergence angle is not 87°5′, as usually stated, but only to 48°30′. If, however, between object and the cover glass there is no water, a divergence angle up to 87°5′ can occur, but only for the very short distance between the object and the very close lying cover glass, so that the arising spherical aberration is not significant.

Since wide ray pencils are necessary to keep diffraction within narrow limits, of course, the illumination apparatus must be capable of providing ray pencils of the same angle, in order to clearly show the contour lines of dark objects. If there are particles in the objects that act like lenses, these can convert a narrow illuminating beam bundle into highly divergent rays and become clearly visible. Otherwise you get to see a confusion of diffractions at and in the object on the one part, and in the aperture of the microscope on the other part.

This apparently is the reason that otherwise good microscopes, if the illumination apparatus is not specifically arranged for this purpose, as with artificial illumination by a flame, provide unusable images of the contour lines of dark objects. For an immersion lens, the best illumination apparatus is constructed on the same principle, a lens of the same kind reversed. If you look at the ocular image of the objective with a loupe, it is easily known if the illumination apparatus gives a sufficiently wide beam cone.

I have to report here a *failed attempt at improvement*, the negative result is of importance. From the theory, I believed that the diffraction of a microscope can be eliminated, if one made the points of the narrow aperture, which causes the diffraction were made of mutually independent luminous points, as through the illumination lenses in the plane of this opening sharp optical image of the light source, as sunlit cloud generated.

Several years ago, in Bonn I made tests of the type on a Nobert microscope with immersion lens of excellent image sharpness. The experiment resulted in the fact

that it did not matter whether the image of the light source was in the plane of the object or the objective; the excessive diffraction fringes in the ocular remained unchanged.

I here report with experiments made with large lenses that such a procedure is not successful. If you have a good achromatic lens of about one-and-a-half-foot focal length that is situated to form a sharp image of the light source, so for example, the sky clouds, upon a glass with a lattice grating that is carved into the glass, the images of several luminous points will be projected onto this grating, and you would therefore think that the interference of the light on the adjacent transparent columns is gone. But if you look through the grating towards the objective, and position before the objective a map sheet into which you have cut fine columns, you can see with the naked eye both of these columns as exactly the same diffractions on the outer edges, as when the lens is taken away or the grid is moved out of focus.

Instead of the grid, I then used a piece of cardboard cut with two fine slits, with an interspace distances about 1 mm, and through which with the naked eye I could just see a system of very fine interference fringes in the diffraction image of another slit cut at very small acute angle. Near the top of this angle it was fine enough to show the interference fringes. These fringes did not disappear when I combined an optical image of the incident light into the plane of the double slit. In this case, there was a suspicion that chromatic or spherical aberration of the radiation should have spread across a gap of 1 mm width, but that did not occur in the least.

The explanation I can only find in the fact that the light of the objective, passing through the narrowest part of the gap, which serves as a visual object, causes so strong diffraction it afterwards reaches both openings of the double slit with a corresponding wave phase, and therefore sends interfering bundles through both slits. In order to see the interference fringes, it is necessary that their minima appear further away from each other than the width of the light strip that is imaged, and if this condition is fulfilled, the theory yields, in fact, that in the midmost bright part of the diffraction image the simple slit forms a light strip that is wider than the distance between the two columns of the double slit cardboard.

Similar relations occur but are more difficult to calculate for conditions when the object is the simple edge of a dark screen. It is known that at such an edge, ray pencils bend themselves into the dark field, which have corresponding oscillation phases, and thus when refracted by a second screen they can interfere. The result of this effect cannot be zero, appears most easily from the fact that the effect of a bright stripe can be represented as the sum of two infinite half planes that lie over each other with their edges, minus that of an equally bright full plane. Since the latter gives no interference phenomenon, as was the bright strips alone in any part of the space, such occurs if for each of the two half planes the edges slightly overlap each other. It follows that the light from a straight edge must spread itself out with the appreciable strength to the same width, as the light from a slit in the cardboard that is bounded by two other slits.

Theory of diffraction in the microscope. In conclusion, I want to specify the method of how to theoretically calculate the propagation of diffracted rays through a microscope. Instead of the simple lengths of rectilinear rays, as considered in the theory of diffraction of the light which propagates only in one optical medium, that are to be considered, you have to take into account the *optical path lengths*, i.e., the lengths which are obtained when each length of the beam is multiplied with the refractive index of the medium in which it propagates, and all products are added. The wave phases of two rays that are emitted by the same luminous spot, and have the same optical path lengths, are equal to each other at the other endpoint, because the wavelengths in medium with different refractive indices are inversely proportional to the refractive indices. Furthermore, it is known that the optical path length of all rays between two conjugate foci of the same ray pencils, in which there is a perfect union of these rays, is equally large. The proof of those referred phrases is verifiable by others in my *Handbook of Physiological Optics*, pp. 238–249.

To calculate the diffraction by the relatively narrowest aperture of the microscope, it will be necessary to treat each point c in the plane of this aperture as the beam center, its phase is determined by the optical length of the normally refracted beam, starting at the illuminated object point a until point c is reached; whose length I denote by (a c). On the other hand, the phase difference between c and the point b in the surface of the image, whose brightness is to be determined depends on the optical length (c b) from the normally refracted ray propagating from c to b. The phase of propagated movement from a through c, as a new central ray, to b thus depends on the sum of the optical path lengths $(ac) + (cb)$. The share which this beam has in the movement of point b is given by an expression of the form:

$$A \cdot \sin\left\{\frac{2\pi}{\lambda}[(ac) + (cb) - (at)] + \text{constant}\right\}$$

where λ is the wavelength in empty space, A is the velocity of propagation, t denotes time. The sum of these magnitudes taken for the collection of points c of the aperture, the factor A can be viewed as approximately independent of c, where it will ultimately determine the propagation of b.

If we consider now the rays from a and b to the point c the closest relative aperture of continuous beams extended in the direction that they have in the points c until they intersect at the points α and β, then these latter points in the medium of c are images of points a and b. Since according to the above, the optical lengths $(a\alpha)$ and $(b\beta)$ as lengths between conjugate foci are constant, so you can set:

$$(ac) = (a\alpha) - (c\alpha)$$

$$(cb) = (\beta b) - (\beta c).$$

The direction of propagation of the ray must be conceived as continuously propagating from the first to the second letter, and therefore,

$$(c\alpha) = -(\alpha c),$$

$$(\beta c) = -(c\beta)$$

Then, the expression for the effect of the individual ray on the point b is equal to:

$$A \cdot \sin\left\{\frac{2\pi}{\lambda}\left[(\alpha c) - (\beta c) - \frac{t}{a} + (a\alpha) + (\beta b)\right] + C\right\}$$

The only variables under the sine that vary with the point c are $(\alpha c) - (\beta c)$; these optical path lengths, lie completely in the medium of c, and are therefore straight lines. Thus, the diffraction phenomena of light from a at the point b, apart from the factor A, which gives the total intensity, will be the same as that of light from α for a point β. The latter can be calculated by the known method for rectilinear beams. [Note α is one-half of the angular aperture of the optical system].

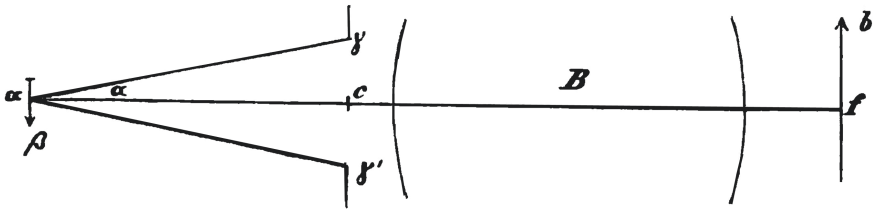


Figure 2

Let $\gamma\gamma'$ in Figure 2 be the relatively narrowest aperture, and c its center, B the part of the optical system immediately behind the aperture.

Let α be the image of axis point a of the object, further let $\alpha\beta$ its image which is lying in the medium $\gamma\gamma'$, and fb the image formed by B in the last medium. If light proceeds from α and is seen through the opening $\gamma\gamma'$ whose radius is ρ, then at α interference fringes appear in which the distance δ between each neighboring maxima (with the exception of the first two) according to known laws is equal to:

$$\delta = \frac{(\alpha c)\lambda}{2\rho} = \frac{1}{2}\frac{\lambda}{\alpha}$$

if we denote α as before, represents the divergence angle $c\alpha\gamma$ that is assumed to be very small. If N is the magnification of the image bf compared to $a\beta$, so in bf the fringe width δ' will be:

$$\delta' = N\delta = \frac{1}{2}N\frac{\lambda}{\alpha} \tag{8}$$

$$\text{or } N = \frac{n\alpha}{n'\alpha'}$$

if α' expresses the divergence angle of the emerging rays, n' is the refractive index of the last medium, n denotes refractive index of the medium at c:

$$\delta' = \frac{n}{2n'} \cdot \frac{\lambda}{\alpha'} \tag{8a}$$

If $n = n'$, the above value of the fringe width in the image bf is entirely analogous for $\alpha\beta$, and this fact shows that the fringes in the last image are just as large as you would see them through the narrowest aperture which determines the divergence angle α' of the cone of rays.

This proof assumes only that the relatively narrowest aperture is located at a point where the divergence angle of the cone of rays is very small; incidentally this may be at any point of the instrument. With an immersion microscope however, this condition is not satisfied when the lowermost boundary of the objective is the relatively narrowest aperture. This condition would be satisfied if the aperture would be situated at the upper side of the second or third lens. Thus, even if no lateral spread of propagating light on the way through the lower objective lens where the radiation cone is still strongly diverging, so it would from then on, where it has become weakly divergent or convergent, its lateral boundary, if by a diaphragm actually situated at that place or only caused by the previous course of rays, must produce diffraction.

The final result makes ultimately little difference whether the aperture is at the periphery of the ray pencils or a little in front or behind. The image of this aperture formed by the ocular lens is at a minimum dimension when it is situated at the back lens, then when it is situated at the front lens, but the difference is of no practical significance.

In equation (8) δ' is the fringe width in the last image, α is the divergence angle in the medium where the rays pass through the aperture is located, λ is the wavelength, N is the magnification of the last image. Denoting with N_1 the magnification of the last image in relation to the object, with λ_1 and n_1, the wavelength and the refractive index for the medium in which the object is located, so we can set according to equation (7)

$$\frac{n_1}{N_1} \sin \alpha_1 = \frac{n}{N} \cdot \alpha$$

since α is considered to be small; α_1 is the divergence angle in the first medium.

If we set the value of $\frac{\alpha}{N}$ in equation (8), so

$$\frac{\delta'}{N_1} = \frac{1}{2} \lambda \frac{n}{n_1} \cdot \frac{1}{\sin \alpha_1} = \varepsilon$$

or $\lambda n = \lambda_1 n_1 = \lambda_0 n_0$, which latter values relate to air, so we have

$$\frac{\delta'}{N_1} = \frac{\lambda_1}{2 \sin \alpha_1} = \frac{\lambda_0}{2 \sin \alpha_0} = \varepsilon \tag{9}$$

This ε is the true magnitude of those lengths in the object, which appears in the magnified image of width of diffraction fringes and is therefore blurred. It can thus be seen as the measure ε of the smallest distinguishable distances in the object. The ε becomes smallest when, α_0 is largest, that is equal to a right angle. So then:

$$\varepsilon = \frac{1}{2} \lambda_0 \tag{9}$$

This limiting definition is, as we see, also independent of the construction of the optical apparatus. It applies equally well for a photographic apparatus as for the connection of the microscope with the eye of the observer. These are the formulas upon which the previous calculations are based.

Postscript. The present work was finished and already prepared for publication when I looked in the April issue of the *Archives for Microscopic Anatomy* and read the work of Mr. Professor E. Abbe: "Contributions to the theory of the microscope and microscopic vison." The same is a preliminary summary of the results, partly theoretical and partly experimental investigations which for the most part are given in summary form. The theorems about divergence of the beams, the magnitude of diffraction in microscopes and their brightness, which form the basis of my conclusions are also found by Mr. Abbe, but first publicly offered without proof. In addition, but it also contains an outline of the work on the important investigations of diffraction by the microscopic objects with narrow light pencils. The special festive occasion to which this volume of the Annals is published, forbids me to withhold my work or completely withdraw it. Since it contains two theorems and some simple experiments to illustrate the theoretical considerations of Mr. Abbe, their publication may also be excused from a scientific point of view.

7.2.2 *Commentary on Helmholtz's Publication "On the Limits of the Optical Capacity [Resolution] of the Microscope"*

The publications of Abbe (1873) and Helmholtz (1874) mark a distinct paradigm shift in the understanding of image formation in the light microscope. Prior to these seminal papers geometrical optics was used for ray tracing in the light microscope and was the basis for understanding image formation. But the two independent and seminal publications of Abbe and then Helmholtz declared that geometrical optics could not explain image formation in the microscope; therefore, they introduced physical optics (in particular, the role of diffraction) to the analysis of image

formation in the light microscope. In his postscript Helmholtz acknowledged Abbe's prior work and publication. But he characterized his own work as a summary of Abbe's experimental work, which Abbe presented without any proof. Finally, Helmholtz justified his decision to publish his own paper on the grounds that it included some new, important mathematical derivations of two theorems that were significant for the design and use of the light microscope.

Having outlined the contributions of Abbe and Helmholtz I now present a summary of Helmholtz's findings as he presented them in his seminal publication (Helmholtz, 1874):

1. Helmholtz began his investigations with a question: Are there physical limits to the resolution of the light microscope?
2. What stimulated Helmholtz's question? He cites the publication of Listing, working in Göttingen, who advocated magnifications of 25,000–50,000× for the light microscope (Listing, 1869a, b).
3. Helmholtz suggested that the maximum magnification of the light microscope is limited because of decreased brightness that occurs with increased magnification.
4. Helmholtz studied the previous work of Kirchhoff on photometry to derive his equations for image brightness in the light microscope.
5. Using photometric analysis Helmholtz independently derived his own Sine Condition, which gave the relation between three factors in object space and their counterparts in image space: $n\beta \sin\alpha = n_1\beta_1 \sin\alpha_1$ (Born and Wolf, 1999). For the case of a microscope α is defined as one-half of the angular aperture. Therefore, Helmholtz derived a relation that is similar to the Abbe Sine Condition: $n'\alpha'\beta' = n''\alpha''\beta''$. But, according to Born and Wolf the origins of the sine condition preceded both Abbe and Helmholtz (Born and Wolf, 1999).
6. Helmholtz, an expert in the field of physiological optics, then used this finding to investigate the case where the light from the microscope does not fill the pupil of the observer. Helmholtz discussed the role of the observer's pupil as an important part of the optical system: illumination system, object, microscope, image, and observer. He concluded that with the very high magnifications for microscopes that Listing discussed the resulting image would have an intensity that could not be seen. Thus, the very high magnifications proposed by Listing would be unusable.
7. Helmholtz then proposed a more significant limitation in the light microscope. Light diffraction in the microscope limited the resolution of the light microscope. Helmholtz stated that diffraction fringes would obscure images and limit the resolution.
8. Helmholtz derived the limiting resolution ε, which is the smallest distance between two details in the object that can be distinguished with a nonimmersion microscope objective, as

$$\varepsilon = \frac{1}{2}\lambda_0$$

where λ_0 is the wavelength of light in air.

9. Helmholtz suggested the use of blue light, which has a shorter wavelength, to increase the resolution of the light microscope.
10. Helmholtz devised and carried out optical experiments to test his ideas on the critical role diffraction plays in image formation in the microscope. In addition, he used his knowledge of physical (wave) optics to investigate diffraction in the light microscope.

Next, I compare the works of Abbe and Helmholtz. Abbe's publication preceded that of Helmholtz. Abbe reluctantly published his paper in the *Archiv für mikroskopische Anatomie*, a journal read by anatomists, histologists, botanists, and microscopists. But, it would be unusual for this journal to be read by physicists. On the other hand, Helmholtz published his findings in *Poggendorff's Annalen der Physik*, *Jubelband*, a distinguished physics journal. Abbe's paper was devoid of figures and equations, while Helmholtz's paper contained two figures and complete mathematical derivations of his equations for both the Helmholtz Sine Condition and the diffraction-limited resolution of the light microscope. However, there were also similarities. Both publications suggested using blue light to enhance the resolution of the light microscope by decreasing the wavelength of illumination. Shorter wavelengths of illumination in the microscope result in increased resolution.

Although Abbe's publication did not contain any figures or equations, it did carefully describe the experiments that he devised and constructed to validate his theory of image formation in the light microscope based on light diffraction. My perusal of the two publications show distinct differences in the way they communicated their two independent theories. Abbe invented the concept and practical importance of numerical aperture as well as the technique of oblique illumination to enhance the resolution of the light microscope. But his 1873 German publication suffered from a lack of illustrations and mathematical analysis. Abbe's publication contained an error regarding the putative formation and merging of the absorption image and the diffraction image; and he only corrected this error decades later (Bradbury, 1996). Helmholtz's publication is remarkable in its mathematical derivations and clear presentation of the physics.

Feffer (1994) pointed out some major differences between Abbe's (1873) and Helmholtz's (1874) derivation of diffraction. According to Feffer, "In Helmholtz's treatment, the objects are self-luminous and diffraction is caused when the light pencil passes through the front opening of the objective. In Abbe's treatment, the objects are not self-luminous and the diffraction occurs in the object itself—an integral part rather than an impediment to the process of image formation" (Feffer, 1994).

I think Abbe and Helmholtz should share equal credit for their seminal work in introducing physical optics (in particular, the critical role of diffraction in their theories of image formation in the light microscope).

7.3 Rayleigh's Paper "On the Theory of Optical Images with Special Reference to the Microscope"

John William Strutt, 3rd Baron Rayleigh, a.k.a. Lord Rayleigh, is renowned for his many achievements in mathematical physics (Masters, 2009). Rayleigh was extremely competent in both theoretical and experimental physics. His seminal work in acoustics was published in his book *The Theory of Sound*. Schrödinger acknowledged the content of this book in his development of the Schrödinger–Rayleigh perturbation technique that is widely used in quantum mechanics. Similarly, Rayleigh's prior work on perturbation theory influenced the development of the Rayleigh–Ritz approximation, and Rayleigh's 1912 work influenced the development of the perturbation method of Wentzel–Kramers–Brillouin (WKB method) for quantum mechanics. Rayleigh's experimental works covered a broad spectrum of topics: theory of color, light scattering, polarization, light diffraction, resolution, and many topics in optics. The Optical Society of America published two volumes of his optics papers under the title *The Collected Optics Papers of Lord Rayleigh* (Strutt, 1994). Rayleigh validated his theory of light diffraction in the light microscope with his observations of diffraction patterns from slits of different widths (Rayleigh, 1896).

Next, I discuss Rayleigh's publications on the light microscope and their relation to the prior publications of Abbe (1873) and Helmholtz (1874). Rayleigh wrote that he read these two publications between 1894 and 1896. In a footnote to his 1896 paper Rayleigh stated that he observed Dr. Stoney demonstrate some of Abbe's experiments on image formation in the microscope.

In particular, I will discuss two of Rayleigh's papers: (1) "On the Theory of Optical Images, with Special Reference to the Microscope" (Rayleigh, 1896) and (2) "The Theory of Optical Images, with Special Reference to the Microscope" (Supplementary paper) (Rayleigh, 1903).

Rayleigh begins his 1896 publication with a review of the works of Abbe and Helmholtz. Rayleigh pointed out that Abbe and Helmholtz built on the work of Lagrange, which Abbe and Rayleigh acknowledged. Rayleigh wrote that Abbe's theory did not treat a luminous point (such as a telescope imaging light from a distant star, which can be considered a point source of light), instead Abbe's theory treated multiple plane waves incident on a grating.

Rayleigh's publication on the resolving power of optical systems introduced two new developments. First, he made use of several mathematical methods that neither Abbe nor Helmholtz applied in their works. Rayleigh applied the Fourier method and Bessel functions in his analysis of diffraction for a variety of objects in the light

microscope. Second, Rayleigh considered the phase relations of light from various regions on the object when developing his mathematical analysis. Rayleigh derived his formula for resolution in the light microscope using Fourier analysis:

$$\varepsilon = \frac{1}{2}\frac{\lambda_0}{\sin\alpha} \tag{10}$$

where ε is the smallest separation distance of two observable points, λ is the wavelength of light in vacuum, and α is the semiangular aperture of the microscope objective. Furthermore, if the wavelength of light in the medium is λ and μ is the refractive index in the medium

$$\lambda = \frac{\lambda_0}{\mu} \tag{11}$$

Then, according to Rayleigh

$$\varepsilon = \frac{1}{2}\frac{\lambda_0}{\mu\sin\alpha} \tag{12}$$

Rayleigh noted that the term $\mu\sin\alpha$ is what Abbe called the numerical aperture.

Rayleigh performed experiments to demonstrate the validity of this equation. It is significant the Rayleigh stated that his conclusions from his mathematical analysis "are in entire accordance with Abbe's theory" (Rayleigh 1896).

In Rayleigh's 1903 paper he discusses the theory of optical images with special reference to the microscope (Rayleigh, 1903). Rayleigh states that Abbe's approach was inapplicable to many problems (i.e., the case of a self-luminous object). While his 1896 publication heavily depended on use of the Fourier theorem, in his 1903 paper Rayleigh presented a more elementary analysis for the case of plane waves illuminating a wire with the plane waves propagating in a direction parallel to the axis of the microscope. Rayleigh used mathematical approximations to obtain his numerical results. He then performed experiments to validate his mathematical results (Rayleigh, 1903).

Rayleigh then attempts to reduce the confusion arising from the implications of the works of Abbe and Helmholtz. First, he points out that neither theory related to the smallness of a single object that may be made visible. He points out that the visibility of a star is only a question of brightness and has nothing to do with resolution. Resolution is only involved when one tries to recognize a double star, or to distinguish details on the surface of a planet. The astronomer can guess as to the existence of a double star, even though it may not be resolved into distinct components, since the diffraction pattern of a star is a disk and the diffraction pattern of a double star may be an oval. The same arguments about a luminous point apply to a luminous line.

7.4 Porter's 1906 Publication: "On the Diffraction Theory of Microscopic Vision"

Albert Brown Porter (1864–1909) was a physicist who did graduate work at the Johns Hopkins University. His research and publications are on image formation based on diffraction in the light microscope (Crew, 1909).

Porter's remarkable paper "On the Diffraction Theory of Microscopic Vision," replete with clear diagrams and mathematical derivations of the equations, was published in the *Philosophical Magazine* (Porter, 1906). On April 22, 1905 Porter read his paper before the American Physical Society. In only 12 pages he discussed Abbe's 1873 theory and experiments. Porter's concise treatment followed from his use of Fourier analysis to form the mathematical basis of Abbe's theory.

Porter explains the motivation for his publication as the lack of a complete mathematical development of image formation in the microscope based on diffraction. From Abbe's description of his demonstration experiments Porter realized that if the object is a transmission grating made up of a series of alternating opaque and transparent lines, then a mathematical analysis based of Fourier's theorem is suitable. Porter's conclusions, which he reached using Fourier analysis, are consistent with the experimental demonstrations that Abbe performed and described in his 1873 publication.

Briefly, illumination interacts with the periodic grating and is diffracted into several diffraction orders including the zero order or the central beam. Several outcomes may occur depending on the angular aperture of the microscope objective and the number of diffraction orders that enter the microscope objective. The back focal plane of the microscope objective contains the Fourier spectrum (diffraction pattern) of the object, and the image plane is where the Fourier components from the back focal plane of the objective combine to form the image of the object.

The angular aperture of a microscope objective determines the number of diffraction orders from the object that can enter the microscope objective. If only a zero-order diffraction beam enters the microscope objective, then there is no image of the object. If additional diffraction orders enter the microscope objective, then there is an image of the object. As the angular aperture of the microscope objective is increased more diffraction orders can enter the microscope objective and resolution of the object is increased.

It is instructive to read Porter's own description of Abbe's theory of image formation in the light microscope. I quote Porter: "If a lens is to produce a truthful image of an illuminated object, it must have an aperture sufficient to transmit the whole of the diffraction pattern produced by the object; if but part of this diffraction pattern is transmitted, the image will not truthfully represent the object, but will correspond to another (virtual) object whose whole diffraction pattern is identical with that portion which passes through the lens; if the structure of the object is so fine, or if the aperture of the lens is so narrow, that no part of the diffraction pattern due to the structure is transmitted by the lens, then the structure will be invisible no matter what magnification is used" (Porter, 1906).

Porter confirmed his mathematical analysis using simple experiments similar to those described by Abbe in his 1873 publication. Porter manipulated the spatial frequency of the image by inserting various circular apertures of different sizes, slits, and a stop in the focal plane of the microscope. Thus, he demonstrated how the image of the object could be altered by changes in spatial frequencies in the focal plane (spatial filtering). Analysis of the Fourier mathematical description and the experimental results were consistent and demonstrated the validity of the mathematical foundation (Porter, 1906).

I quote some material from Porter's paper in which he described his experiments and their interpretation. The reader is immediately reminded of the similarity with Abbe's description of his experiments. Porter wrote: "When a lens forms a real image of a grating, it does so by adding together in the focal plane the harmonic components of the diffracted light. If the illumination is central and the aperture of the lens is so narrow that it cannot pass the light represented by the second and succeeding terms of (2) [a harmonic expansion in terms of sine and cosine terms], i.e. if it passes only the non-periodic first term, or central beam, the illumination in the focal plane is uniform and no image of the lines of the grating is formed. If the first spectrum, i.e. the first periodic term in (2), is transmitted by the lens, an image is formed having a periodic structure corresponding to that of the grating, but in which the lines are much blurred. When the aperture of the lens is further widened so as to admit spectra of higher and higher orders, the definition becomes sharper and sharper and the image in general approximates more and more closely to a true representation of the object" (Porter, 1906).

Porter then described the resolution of a diffraction-limited optical system. Porter found that if the ruled lines on a diffraction grating are closer than one-half of the wavelength of the illumination light, then in the back focal plane of the microscope objective there will be no diffraction pattern and therefore no image in the image plane (Porter, 1906). This conclusion is similar to that found by both Abbe (1873) and Helmholtz (1874).

Porter, following Abbe's 1873 publication, suggested that the resolution of the light microscope could be enhanced using two methods: illumination of shorter wavelengths (i.e., illumination with ultraviolet light) and use of an immersion lens in which a liquid with a high refractive index is situated between the object and the microscope objective. Finally, Porter remarked on the real problem of artifacts or false images that are generated in the light microscope due to correlations between the aperture of the microscope objective and the fine details of the object. Abbe in his 1873 publication also pointed out this phenomenon.

Porter's publication had another important outcome. Among the critical arguments raised against Abbe's 1873 publication and Abbe's Experiments was that Abbe used atypical apertures of special shapes that were not consistent with the usual practice of microscopy (Feffer, 1994). Porter in his 1906 publication was able to replicate Abbe's experiments, even though he used circular stops similar to those used by microscopists. Therefore, Porter's publication helped to mitigate the adverse criticism of Abbe's previous work.

References

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.

Born, M., and Wolf, E. (1999). *Principles of Optics*, 7th (expanded) edition. Cambridge: Cambridge University Press.

Bradbury, S. (1996). The reception of Abbe's theory in England. Proceedings of the Royal Microscopical Society, **31**, 293–301.

Cahan, D. (1993). *Hermann von Helmholtz and the Foundations of Nineteenth-Century Science.* Berkeley: University of California Press.

Crew, H. (1909). Albert B. Porter. Science, **29**, 962–263.

Feffer, S. M. (1994). *Microscopes to Munitions: Ernst Abbe, Carl Zeiss, and the Transformation of Technical Optics, 1850–1914*. PhD dissertation, University of California, Berkeley. Ann Arbor: UMI Dissertation Services.

Fripp, H. E. (1874). On the limits of optical capacity of the microscope. Translated from Germany publication of Hermann von Helmholtz. Proceedings of the Bristol Naturalists' Society, **I**, 407–440.

Gordon, J. W. (1903). The Helmholtz theory of the microscope. Journal of the Royal Microscopical Society, **23**, 381–446.

Helmholtz, H. (1874). Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope. Poggendorff's Annalen der Physik, Jubelband 1874, 557–584, Leipzig. Translated as: "On the theoretical limits of the optical capacity of the microscope". Monthly Microscopical Journal, **16**, 15–39 (1876).

Helmholtz, H. (1877). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, English translation of the fourth edition (1877) of *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* by A.J. Ellis, N.Y., Dover Publications, 1954.

Helmholtz, H. (1903). Vorträge und Reden, Fünfte Auflage, Erster Band, Zweiter Band. Braunschweig: Friedrich Vieweg und Sohn.

Helmholtz, H. (1909–1911). *Handbuch der Physiologischen Optik*, Dritte Auflage, English translation of the third edition, three volumes (1909–1911) by J. P.C. Southall, Washington, D. C., Optical Society of America, 1924.

Listing, J. B. (1869a). Vorschlag zu ferner Vervollkommung des Mikroskops auf einem geänderten dioptrischen Wege. [Proposal for further improvement of the microscope on a modified dioptric way]. Poggendorff's Annalen der Physik, **136**, 467–472.

Listing, J. B. (1869b). Nachtrag betreffend die neue Construction des Mikroskops. [Supplement relating to the new construction of the microscope]. Poggendorff's Annalen der Physik, **136**, 473–479.

Listing, J. B. (1871). Notiz über ein neues Mikroskop von R. Winkel [Note on a new microscope by R. Winkel]. Poggendorff's Annalen der Physik, **142**, 479–480.

Masters, B. R. (2009). John William Strutt, Third Baron Rayleigh: a scientific life that bridged classical and modern physics. Optics and Photonics News, **20**, 36–41.

Masters, B. R. (2010). Hermann von Helmholtz, a 19th century renaissance man. Optics and Photonics News, **21**, 35–39.

Porter, A. B. (1906). On the diffraction theory of microscope vision. Philosophical Magazine, **11**, 154–166.

Rayleigh, L. (1896). XV. On the theory of optical images, with special reference to the microscope. Philosophical Magazine and Journal of Science, London, **XLII**, 167–195.

Rayleigh, L. (1903). VIII. On the theory of optical images, with special reference to the microscope (supplementary paper). Journal of the Royal Microscopical Society, London, 474–482.

Strutt, J. W. (1994). *The collected optics papers of Lord Rayleigh*. Part A, 1869–1892; Part B, 1892–1919. Washington, D.C.: Optical Society of America.

Further Reading

Abbe, E. (1874). Contributions to the theory of the microscope and the nature of microscopic vision. Translated into English by H. E. Fripp. Proceedings of the Bristol Naturalists Society, **I**, 202–258. Read before the Bristol Microscopical Society, December 16, 1974.

Abbe, E. (1882). The relation of aperture and power in the microscope. Journal of the Royal Microscopical Society, Section II, **2**, 300–309; 460–473. [Read before the Society on May 10, 1882], [Abbe wrote the paper in English].

Abbe, E. (1883). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Section II, **3**, 790–812. [Read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Abbe, E. (1989). *Gesammelte Abhandlungen, I-IV*. Hildesheim: Georg Olms Verlag. [Originally published in 1904, Jena: Verlag von Gustav Fischer].

Carpenter, W. B., and Dallinger, W. H. (1901). *The Microscope and its Revelations*, Eighth Edition. Philadelphia: P. Blakiston's & Son.

Crisp, F. (1878). On the influence of diffraction in microscopic vision. Journal of the Quekett Microscopical Club, **5**, 79–86.

Fripp, H. E. (1876a). On aperture and function of the microscope object glass [microscope objective]. Proceedings of the Bristol Naturalists' Society, **I**, 441–456.

Fripp, H. E. (1876b). On the physiological limits of microscopic vision. Proceedings of the Bristol Naturalists' Society, **I**, 457–475.

Fripp, H. E. (1877). Professors Helmholtz and Abbe on the optical powers of the microscope. Monthly Mathematical Journal, **17**, 103–105.

Helmholtz, H. (1851). Beschreibung eines Augen-Spiegels zur Untersuchung der Netzhaut im lebenden Auge. Berlin: A. Förstner'sche Verlagsbuchhandlung. Description of an ophthalmoscope for examining the retina in the living eye. English translation by Robert W. Hollenhorst, Chicago, Archives of Ophthalmology, **46**, 565–583 (1951).

Hopkins, H. H. (1950). *Wave theory of Aberrations*. New York: Oxford University Press.

Köhler, A. (1893). Ein neues Beleuchtungsverfahren für mikrophotographische Zwecke. Zeitschrift für wissenschaftliche Mikroskopie und für Mikroskopische Technik, **10**, 433–440.

Köhler, A. (1894). A new lighting apparatus on the photomicrographical purposes. Journal of the Royal Microscopical Society, **14**, 261–262.

Köhler, H. (1981). On Abbe's theory of image formation in the microscope. Journal of Modern Optics, **28**, 1691–1701.

Lister, J. J. (1830). On some properties in achromatic object-glasses applicable to the improvement of the microscope. Philosophical Transactions of the Royal Society of London, **120**, 187–200.

Zimmermann, A. (1895). *Das Mikroskop. Ein Leitfaden der wissenschaftlichen Mikroskopie*. [The Microscope, A Manual of Scientific Microscopy] Leipzig: Franz Deuticke.

Chapter 8
Further Insights into Abbe's Theory of Image Formation in the Microscope Based on Diffraction

8.1 Introduction

In the 1930s and the subsequent decades physicists made important advances in optics and microscopy (Duffieux, 1970; Françon, 1950, 1953, 1961; Hopkins, 1950, 1951, 1953; Van Cittert, 1934; Zernike, 1934a, b, 1938, 1942a, b, 1948, 1950, 1953a, b, 1958; Zernike and Brinkman, 1935). Their understanding of coherence and its relevance to image formation in the light microscope had enormous implications for light microscopy. These physicists introduced Fourier techniques to the problem of image formation in the light microscope. Increased understanding of the coherence properties of light led to important inventions such as the phase contrast microscope, the interference microscope, and holography. Unfortunately, books published in the 1930s and the 1940s rarely discussed these advances in physical optics. I found one book that attempts to bridge this gap between theory and application of the light microscope titled *The Theory of the Microscope* (Martin, 1966).

In this chapter I discuss the seminal work of Frits Zernike and his invention of the phase contrast microscope. The next section presents Zernike's thoughts, taken from his Nobel Lecture, on why Abbe almost succeeded in inventing the phase contrast microscope. I also discuss the prior work of others that paved the way to manipulating the phases of the various diffraction orders in the light microscope to control the contrast of transparent objects that present a difficulty in the light microscope. There are also important general lessons in Zernike's comments on innovation in science. The chapter ends with a description and discussion of Peter Evennett's five videos on YouTube that use modern devices to demonstrate Abbe's experiments. These videos also demonstrate the principle of the phase contrast microscope invented by Zernike. The Evennett videos show just how close Abbe came to inventing the phase contrast microscope.

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_8

8.2 Zernike's Insights on Abbe's Theory and the Zeiss Werke Culture

Zernike on innovation:

> ". . . I am impressed by the great limitations of the human mind. How quick are we to learn, that is, to imitate what others have done or thought before. And how slow to understand, that is, to see the deeper connections. Slowest of all, however, are we in inventing new connections or even in applying old ideas in a new field."
>
> —Frits Zernike, Nobel Lecture, December 11, 1953

In this section I introduce Frits Zernike and analyze the content of his Nobel Lecture (Zernike, 1953a). There are three main themes contained in his lecture. First, his presentation of how he discovered phase contrast and applied it to microscopy in his phase contrast microscope. Second, Abbe's role in hindering innovation in microscope development during the last decade of his life. Third, the culture at Zeiss Werke and its effect on new developments, specifically the delayed acceptance of Zernike's invention of the phase contrast microscope (Fig. 8.1).

The brief biographical sketch of Frits Zernike is based on two sources (Ferwerda, 1993; Zernike, 1953b). Frits (Frederik) Zernike (1888–1966) was a professor of mathematical and technical physics and theoretical mechanics at Groningen University, the Netherlands (Ferwerda, 1993).

I posit that Zernike's education and seminal achievements in both experimental and theoretical optics is another example of the benefits of interdisciplinary studies. At the university he majored in chemistry and had minors in physics and mathematics. Surprisingly, his graduate degrees were both in the field of chemistry; yet he developed his expertise in both advanced physics and mathematics. In support of

Fig. 8.1 Frits Zernike

this claim I quote his views on the necessary role of mathematics in understanding physical phenomena: "It is especially the general insight, however, which gains very much by the discovery of the asymptotic development. It shows that physical intuition combined with experimental ability may go far towards elucidating the main characteristics of phenomena, but that only an adequate mathematical treatment can give a satisfactory final solution" (Zernike, 1948).

According to Zernike his work on optics began in 1934 when he began his studies on the relation between lens aberrations and resulting Airy diffraction patterns (Zernike, 1934a, b; 1948). The last paragraph of his 1948 publication is insightful when it comes to understanding the scientific process, innovation, and discovery and I reprint it (Zernike, 1948): "It is especially the general insight, however, which gains very much by the discovery of the asymptotic development. It shows that physical intuition combined with experimental ability may go far towards elucidating the main characteristics of phenomena, but that only an adequate mathematical treatment can give a satisfactory final solution."

These early studies culminated with his graduate students Nijboer and Nienhuis completing their doctoral dissertations under Zernike's mentorship. Out of this work came Zernike's mathematical description of optical aberrations termed Zernike polynomials, mathematical sequences of polynomials that have the unique property of being orthogonal to the unit disk (Zernike 1934a, b; Zernike and Brinkman, 1935). Zernike published many prescient articles and books on physical optics including the study of partial coherence. A good example is his publication "The concepts of coherence and its applications to optical problems" (Zernike, 1938).

In 1953 Zernike was awarded the Nobel Prize in Physics for his demonstration of the phase contrast method, especially for his invention of the phase contrast microscope. Zernike delivered his Nobel Lecture on December 11, 1953 (Zernike, 1953a, b). A one-minute video in Swedish is available at: http://www.nobelprize.org/mediaplayer/index.php?id=357 (accessed April 4, 2019).

Zernike's Nobel Lecture "How I Discovered Phase Contrast" is the subject of this section. I strongly recommend careful study of Zernike's Nobel Lecture (Zernike, 1953a, b). In summary, first there were anomalous observations in Zernike's study of Roland ghost images from ruled gratings. Zernike distinguished amplitude gratings (amplitude objects) from phase gratings (phase objects). The former diffract light due to differences in the amplitude of light, and the latter diffract light due to unequal path lengths or phases. He cogently noted that most objects of biological or medical interest are phase objects that appear transparent in the light microscope. Stains were used to convert phase objects into amplitude objects that the human eye can observe due to differences in intensity that form contrast in the object. In Chapter 12 I describe Zernike's phase contrast microscope and the physical basis of its operation. Zernike obtained a patent for his phase contrast microscope in 1936 (Zernike, 1936). In 1950 Zernike developed a color phase microscope in which phase differences were represented as color differences (Zernike, 1950).

Zernike made the observation of the phase shift in the Roland ghosts from the principal lines. He then discovered how he could make phase objects visible by optically changing phase differences into intensity differences. That discovery was the foundation of his phase microscope. The human eye cannot detect differences of phase, but it can detect intensity differences. The insights to creativity, innovation, risk-taking in science, skepticism in science, arrogance, and expert opinion as a detriment to advances in science are all contained in his Nobel Prize lecture (Zernike, 1953a, b).

Below are Zernike's own words quoted from excerpted material of his 1953 Nobel Lecture (Zernike, 1953a, b).

"Phase contrast was not discovered while working with a microscope, but in a different part of optics. It started from my interest in diffraction gratings, about from 1920 on. Small nearly unavoidable imperfections in the location of the grooves show clearly in the optical behavior of the grating. The regularly recurring displacement of the grooves causes corresponding changes of the optical path… it, so that each strong spectral line is accompanied to right and left by a number of weak spurious lines, the so-called "Rowland ghosts."

Now it is common knowledge that in all interference phenomena differences of phase are all-important. Why then had phases never been considered before in this case, nor in the corresponding one in the microscope? Some excuse may be found in the difficulty to define them exactly. In the case of the Rowland ghosts the result was: their phases differ by 90° from the principal line.

Now I happened to know of a simple method to change this. Lord Rayleigh described in 1900 how to make very shallow etchings in glass surfaces without spoiling their optical quality, by the slow action of very dilute hydrofluoric acid. By this process I made what I called *phase-strips:* glass plates with a straight groove, a millimeter or less wide and of a uniform depth of half a wavelength. Such a phase-plate was placed in the spectrum so that a bright spectral line fell on the strip, whereas its ghosts passed through the glass beside it. In a telescope behind the phase-plate the stripes on the grating surface then stood out clearly.

For a physicist interested in optics it was not a great step to change over from this subject to the microscope. Remember that in Abbe's remarkable theory of the microscope image the transparent object under the microscope is compared with a grating. To be precise, a transmission grating is considered as the test-object and the diffraction by this grating as the primary phenomenon. At first sight this has nothing to do with the magnified image of the object formed by the microscope objective. Instead, the objective forms an image of the light source, practically in its back focal plane, consisting of a central direct image accompanied by diffracted images on both sides.

This, although on a very much smaller scale, is the analogue of the grating line with its ghosts. The light issuing from these images overlaps in the eyepiece of the microscope and by interference gives rise to stripes which, curiously enough, resemble a magnified image of the object!" Abbe's theory has been summarized in this sentence: "The microscope image is the interference effect of a diffraction phenomenon" (Zernike, 1953a, b).

I now present my commentary on Zernike's Nobel Lecture with an emphasis on innovation and creativity in science. In his lecture Zernike recounts Abbe's 1873 publication on image formation in the light microscope based on diffraction and he contrasts Abbe's early innovative and creative period with the period of significantly diminished innovation in the last three decades of his life.

In 1932 Zernike visited Zeiss Werke in Jena to demonstrate his new phase contrast microscope. But his invention was not appreciated; in fact, the people at Zeiss Werke told him that had such an invention been important the scientists at Zeiss Werke would have invented it first.

Zernike posited that Abbe's creative period became greatly diminished in 1890 when Abbe took over the administration of Zeiss Werke and concentrated on social reforms. Zernike noted that Abbe's last publication on microscopy contained the observation that transparent objects lack contrast in the light microscope (Abbe, 1889).

Zernike made the strong claim that after Abbe died in 1906 his image as an authority figure continued for 25 years and inhibited further progress in the understanding of the light microscope. Zernike commented that the lack of acceptance of Abbe's (1873) theory was caused by several disparate factors: Abbe's theory as a phenomenological theory was too abstract; it only explained the case for a point source of illumination with an object that had a periodic structure, but it did not explain the problems of imaging transparent objects.

The history of the microscope is replete with examples of simultaneous inventions that resulted in enhanced contrast, enhanced resolution, or both. As I discuss these inventions I will briefly point out instances of simultaneous independent invention.

The material that I discuss in the following paragraphs is contained in the book *Phase Microscopy, Principles and Applications* (Bennett et al., 1951, pp. 3–4).

The next paragraphs describe some of the studies that predated Zernike's invention of phase contrast microscopy. Bratuscheck stated that Abbe placed glass wedges in the back focal plane of the microscope objective. Abbe made the following observation: when he introduced a 180° phase difference between the zero-order diffraction beam and the first-order diffraction beam, then the contrast between the lines and the spaces between the lines were reversed (Bratuscheck, 1892). That observation should have led Abbe to invent the phase contrast microscope.

Bratuscheck performed the following experiment. He placed absorbing strips of black soot at the back focal plane of the microscope objective that decreased the intensity of the zero-order diffraction beam. His object on the microscope stage comprised alternate clear and weakly absorbing strips of soot. These very thin soot stripes were almost transparent (i.e., without contrast in a typical microscope). He demonstrated that when he used the previously described mask at the back focal plane of the microscope, which decreased the intensity of the zero-order diffraction beam, then the contrast of the image was increased. Furthermore, Bratuscheck noted that if he used platinum instead of soot, then this caused a slight change in the

phase of the light transmitted through the platinum stripes with respect to the phase of the light that is transmitted through the clear spaces between the stripes (Bratuscheck, 1892). It is not clear if Abbe was aware of these experimental results.

Prior to Zernike's seminal studies that led to his invention of the phase contrast microscope there were two other important publications. In 1904 Rheinberg published his study "On the influence of images of gratings of phase difference amongst their spectra" (Rheinberg, 1904). The object was a grating in which the slits were twice as wide as the bars. Rheinberg was able to completely reverse the contrast of the image by blocking all the diffraction orders except the first-order diffraction beam and the second-order diffraction beam on one side. He concluded that the contrast of the object could be altered by controlling the phase of the diffraction beams by selecting specific diffraction orders (Rheinberg, 1904). Similarly, in 1905 Conrady published his paper "An experimental proof of phase reversal in diffraction spectra" (Conrady, 1905). The well-known phenomenon of contrast reversal is demonstrated in Peter Evennett's videos that I discuss in the next section.

8.3 Abbe's Diffraction Experiments Parts 1–5, by Peter Evennett, Dresden Imaging Facility Network, 2001

Peter Evennett of the Dresden Imaging Facility Network, 2001 produced five videos on YouTube that demonstrate Abbe's experiments. I highly recommend careful viewing of each of these excellent video demonstrations narrated in English. These videos are remarkable for their clear presentation of Abbe's theory of image formation in the light microscope.

Peter Evennett's demonstration apparatus consists of a light microscope with two video cameras whose images are simultaneously shown on a split screen, the bottom portion showing the image plane of the microscope (where we observe the image of the object on the specimen stage) and the upper portion showing diffraction spectra in the back focal plane of the objective. The back focal plane of the objective and the image plane are conjugate planes. Köhler illumination is a source of white light; in special cases filters produce either red or green illumination. The image of the lamp filament is observed in the back focal plane of the objective—not in the image plane.

A small pinhole or alternatively a slit is placed in the first focal plane of the condenser to diminish illumination light and facilitate observation of diffraction spectra in the back focal plane of the microscope objective. The objects are ruled lines on a glass slide (a diffraction grating) with fine (upper portion of slide) or coarse (lower portion of slide) line separations.

Below are some of the observations that we can make while we watch the videos. The fine grating diffracts the light into widely separated diffraction spectra. The coarse grating diffracts the lights into narrowly separated diffraction spectra.

Video Part 1 illustrates the design and operation of the optical demonstration instrument, the concept of conjugate planes, and comparison of the diffraction spectra in the back focal plane and the image of the object for coarse and for fine gratings.

Video Part 2 demonstrates the effect of changing the color of illumination light and the aperture of the microscope objective on both diffraction spectra and the image of the object in the image plane. We observe that resolution is increased with light of shorter wavelengths (green light vs. red light). We also observe that resolution of the object depends on the aperture of the microscope objective. Next, the video demonstrates that the final image in the image plane is the result of interference of the diffraction orders of light from the object. One diffraction order cannot form the image; at least two different diffraction orders are required to interfere to get an image of the object. An example of the great pedagogical value of these videos is when Peter Evennett first makes a demonstration with his apparatus and then poses the question for the viewer: What could be the physical explanation for the demonstration that I just performed? That forces the viewer of the demonstrations to pause and think about possible causes.

Video Part 3 introduces the concept of coherence. Peter Evennett demonstrates that incoherent light cannot interfere and that interference requires at least two beams of coherent light. The demonstration apparatus shows that when illumination is incoherent there is no interference of the resulting diffraction orders and therefore no image is formed in the image plane of the microscope. Next, we observe a demonstration with a stop inserted into the back focal plane that obscures the zero order of the diffraction beams. The result is that the lines and the background show a reversal of intensity; when zero-order diffraction is present the background is bright and the lines of the grating are dark. But, in the absence of a zero-order diffraction beam we observe the background is dark and the lines of the grating are bright. Additionally, we conclude there are two requirements for the dark (black) lines due to destructive interference in the image of the grating (object) to be formed: the zero-order diffraction beam and any other order of a diffraction beam must be coherent, and the phase of the two beams must differ by 180° or half a wavelength. This is a dark-field image of the object. What appears most remarkable is that if another object with almost no contrast is placed on the stage, then it will still diffract the light. But its image in the image plane is barely visible. When a stop obscures the zero-order diffraction beam the object is now observed in the image plane with enhanced contrast.

Video Part 4 continues to explain that objects with very low contrast (a.k.a. phase objects such as a thin layer of cells on a microscope slide) diffract light, but the phase relation between the zero-order diffraction beam and higher order diffraction beams are one-quarter of a wavelength out of phase with each other. Such beams cannot destructively interfere. Next, a specially designed mask or stop is inserted to partially obscure the zero-order diffraction beam. This special mask causes two effects: first, it absorbs some of the intensity of the zero-order diffraction beam and,

second, it shifts the phase of the zero-order diffraction beam by one-quarter of a wavelength. Since the phase object resulted in the zero-order diffraction beam being one-quarter of a wavelength out of phase with the other diffraction beams and the special mask changes the phase of light from the zero-order diffraction beam by one-quarter of a wavelength the net result is that the zero-order diffraction beam and the other orders of diffraction beams differ in phase by half a wavelength. From the previous demonstration we know that these are the conditions necessary for destructive interference, and the almost invisible phase object is now seen as a high-contrast image in the image plane of the microscope. Insertion of the special mask results in the low-contrast phase object now being observed in the image plane as if it were a high-contrast absorption object. Again, an absorbing object will diffract light into a zero-order diffraction beam and higher order diffraction beams, and these two sets of beams differ in phase by half a wavelength. This technique is the basis behind the phase contrast microscope. The only difference is that the phase contrast microscope uses an annular ring of illumination instead of the pinhole used in the demonstration together with a phase plate or mask to shift the phase difference between the zero-order and higher order diffraction beams. Remember that the human eye is insensitive to phase differences, but is sensitive to intensity differences. The phase contrast microscope optically converts an invisible image of a completely transparent object into an image in which the details of the phase object can be observed. In this case there is no similarity between the object and its image when observed in the standard light microscope, but with the phase contrast microscope the transparent object is observed and contrast is provided by differences in light intensity. The phase contrast microscope optically converts differences in phase into differences of light intensity that the human eye can observe.

Video Part 5 poses the question: Is there a unique relation between the object and its diffraction pattern? For example, from the diffraction pattern can we deduce the shape and the structure of the object? I leave this problem for the reader to think about. This last video walks the viewer through the light microscope. Illumination light is diffracted by the object. The aperture of the microscope objective limits which diffraction orders enter the front lens of the objective. The diffraction spectra are focused in the back focal plane of the microscope objective. These diffraction beams diverge, propagate, and interfere to form an image of the object in the image plane of the microscope.

I suggest that after the reader becomes familiar with Peter Evennett's five videos demonstrating Abbe's theory of image formation the reader should return to Chap. 6 and review my English translation and commentary of Abbe's 1873 publication. Chapter 6 also contains Stephenson's 1877 publication in which he describes his eye-witness account of Abbe's experiments performed by Abbe himself. The similarity between Stephenson's publication and Evennett's five videos is striking.

Details and Links to Peter Evennett's Videos

- *Abbe's Diffraction Experiments Pt. 1 (9:35 min) by Peter Evennett, Dresden Imaging Facility Network:* https://www.youtube.com/watch?v=fAuo7NIS97U (accessed April 4, 2019).
- *Abbe's Diffraction Experiments Pt. 2 (9:38 min) by Peter Evennett, Dresden Imaging Facility Network:* https://www.youtube.com/watch?v=rPAJ2Vs9A_I (accessed April 4, 2019).
- *Abbe's Diffraction Experiments Pt. 3 (9:38 min) by Peter Evennett, Dresden Imaging Facility Network:* https://www.youtube.com/watch?v=CbPoi9bie1I (accessed April 4, 2019).
- *Abbe's Diffraction Experiments Pt. 4 (9:43 min) by Peter Evennett, Dresden Imaging Facility Network:* https://www.youtube.com/watch?v=wZ369WKiJdQ (accessed April 4, 2019).
- *Abbe's Diffraction Experiments Pt. 5 (3:48 min) by Peter Evennett, Dresden Imaging Facility Network:* https://www.youtube.com/watch?v=t6RO7fusHYk (accessed April 4, 2019).

References

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.

Abbe, E. (1889). On the effect of illumination by means of wide-angled cones of light. Journal of the Royal Microscopical Society, Series II, **IX**, 721–724.

Bennett, A. H., Jupnik, H., Osterberg, H., and Richards, O. W. (1951). *Phase Microscopy, Principles and Applications*. New York: John Wiley & Sons, Inc., pp. 3–4.

Bratuscheck, K. (1892). Die Lichtstärke-Änderungen nach verschiedenen Schwingungsrichtungen in Linsensystemen von grössen Öfnungswinkel mit Beziehung zur mikroskopischen Abbildung. [The light intensity changes according to different directions of vibration in lens systems of great angular aperture with relation to microscopic image formation] Zeitschrift für wissenschaftliche Mikroskopie, **9**, 145–160.

Conrady, A. E. (1905). An experimental proof of phase reversal in diffraction spectra. Journal of the Royal Microscopical Society, **25**, 150–152.

Duffieux, P. M. (1970). *L'intégrale de Fourier et ses applications à l'optique*. Paris: Masson, Editeur. Republished in English as: Duffieux, P. M. (1970). *The Fourier Transform and its applications to optics*, second edition. New York: John Wiley and Sons.

Ferwerda, H. A. (1993). Frits Zernike-life and achievements. Optical Engineering, **32**, 3176–3181.

Françon, M. (1950). *Le contraste de phase en optique et en microscopie*. Editions de la Rev. d'Optique theorique et Instrumental, Paris, 1–109.

Françon, M. (1953). *Le microscope à contraste de phase et le microscope interférentiel*. Paris: Centre National de la Recherche Scientifique.

Françon, M. (1961). *Progress in Microscopy*. Evanston, Illinois: Row, Peterson and Company.

Hopkins, H. H. (1950). The influence of the condenser on microscopic resolution. Proceedings of the Physical Society (B), **63**, 737–744.

Hopkins, H. H. (1951). The concept of partial coherence in optics. Proceedings of the Royal Society of London Series A, **208**, 263–277.

Hopkins, H. H. (1953). On the diffraction theory of optical images. Proceedings of the Royal Society of London Series A, **217**, 408–432.
Martin, L. C. (1966). *The Theory of the Microscope*. New York: American Elsevier Publishing Company, Inc. and London: Blackie.
Rheinberg, J. (1904). On the influence of images of gratings of phase difference amongst their spectra. Journal of Microscopy, **24**, 388–390.
Van Cittert, P. H. (1934). Die Wahrscheinliche Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse beleuchteten Ebene. [The Probable vibrational distribution in one of the one light source directly or via a lens Illuminated plane]. Physica, **1**, 201–210.
Zernike, F. (1934a). Beugungstheorie des Schneidenverfahrens und Seiner Verbesserten Form, der Phasenkontrastmethode. [Diffraction theory of the cutting process and its Improved form of the phase contrast method] Physica, **1**, 689–704.
Zernike, F. (1934b). Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. Monthly Notices of the Royal Astronomical Society, **94**, 377–384.
Zernike, F. (1936). Deutsches Reichspatent No. 636168 (September 1936).
Zernike, F. (1938). The concepts of degree of coherence and its application to optical problems. Physica, **5**, 785–795.
Zernike, F. (1942a). Phase contrast, a new method for the microscopic observation of transparent objects, Part 1. Physica, **9**, 686–698.
Zernike, F. (1942b). Phase contrast, a new method for the microscopic observation of transparent objects, Part 2. Physica, **9**, 974–986.
Zernike, F. (1948). Diffraction and optical image formation. Proceedings of the Physical Society, **61**, 158–164.
Zernike, F. (1950). Color phase-contrast microscopy: requirements and applications. Physica, **9**, 974–986.
Zernike, F. (1953a). Nobel Lecture: How I Discovered Phase Contrast. December 11, 1953. https://assets.nobelprize.org/uploads/2018/06/zernike-lecture.pdf.
Zernike, F. (1953b). Frits Zernike - Biographical. Nobelprize.org. Nobel Media AB 2014. Web. 24 Apr 2017. https://www.nobelprize.org/prizes/physics/1953/zernike/biographical/ (Accessed September 26, 2018).
Zernike, F. (1958). The wave theory of microscopic image formation. Appendix K. In: *Concepts of Classical Optics*, J. Strong, pp. 525–536. San Francisco: W. H. Freeman.
Zernike, F., and Brinkman, H. C. (1935). Hypersphärische Funktionen und die in sphärischen Bereichen orthogonalen Polynome. Verhandelingen der Koninklijke Akademie wetenschappen te Amsterdam (Proceedings of the Royal Academy, Amsterdam), **38**, 161–170.

Further Reading

Abbe, E. (1989). *Gesammelte Abhandlungen*, I-IV. Hildesheim: Georg Olms Verlag. [Originally published in 1904, Jena: Verlag von Gustav Fischer].
Airy, G. B. (1835). Diffraction of an object-glass with circular aperture. Transactions of the Cambridge Philosophical Society, **5**, 283–291.
Born, M., and Wolf, E. (1999). *Principles of Optics*, 7th (expanded) edition. Cambridge: Cambridge University Press.
Feffer, S. M. (1994). *Microscopes to munitions: Ernst Abbe, Carl Zeiss, and the transformation of technical optics, 1850–1914*. PhD Dissertation, University of California, Berkeley, 1994. Ann Arbor: UMI Dissertation Services.
Köhler, H. (1981). On Abbe's theory of image formation in the microscope. Journal of Modern Optics, **28**, 1691–1701.
Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy*: The Genesis of Live Cell Imaging. Bellingham: SPIE Press.

Michel, K. (1964). *Die Grundzüge der Theorie des Mikroskops in elementarer Darstellung*. 2. neubearbeitete Auflage. Stuttgart: Wissenschaftliche Verlagsgesellschaft M.B. H.

Nienhuis, K. (1948). *On the influence of diffraction on image formation in the presence of aberrations.* Thesis, Groningen.

Nijboer, B. R. A. (1942). *The diffraction theory of aberrations*. Thesis, Groningen.

Porter, A. B. (1906). On the diffraction theory of microscope vision. Philosophical Magazine, **6**, 154–166.

Reynolds, G. O., DeVelis, J. B., Parrent, G. B., Jr., and Thompson, B. J. (1989). *The New Physical Optics Notebook: Tutorials in Fourier Optics*. Bellingham: SPIE Optical Engineering Press.

Singer, W., Totzeck, M., and Gross, H. (2005). *Handbook of Optical Systems, Physical Image Formation* (Volume 2). Weinheim: Wiley-VCH.

Zernike, F. (1935). Das Phasenkontrastverfahren bei der mikroskopischen Beobactung. Zeitschrift für Technische Physik, **16**, 454–457.

Zernike, F. (2014). Frits Zernike Documentary. Nobelprize.org. https://www.nobelprize.org/mediaplayer/?id=357 (Accessed on September 26, 2018).

Chapter 9
Mathematical Description of Abbe's Theory of Image Formation in the Microscope Based on Diffraction

9.1 Introduction

In Abbe's 1873 seminal publication he promised another paper that would contain his mathematical analysis of image formation in the light microscope, a paper that never appeared because he died. After Abbe's death his students and colleagues developed and published details of his theory of image formation in terms of Fourier optics (see Appendix A). In the following decades physical optics based on the Fourier theory of image formation rapidly developed from the works of Duffieux (1946), Hopkins (1951, 1953), Van Cittert (1934), and Zernike (1938).

Modern views of Abbe's theory of image formation based on diffraction have been discussed at various levels in a variety of excellent textbooks and monographs. Discussions are based on Fourier's theory of image formation, which includes the mathematics of Fourier analysis, wave or physical optics, and scalar and vector approaches to diffraction theory (Goodman, 2017).

More rigorous sources present the theory in its most comprehensive and general form. Typically, the various theories invoke specific assumptions and permit approximations to be made that simplify the mathematics and provide approximate solutions. The key is to understand the assumptions and determine whether the approximate theories are valid and useful in a particular experimental situation. There is no substitute for reading the original papers, which are the primary sources, many of which are written in English, French, or German.

Comprehensive secondary sources include books that I particularly recommend: Joseph Braat and Peter Török, *Imaging Optics,* Joseph W. Goodman, *Introduction to Fourier Optics*, Fourth Edition; Min Gu, *Advanced Optical Imaging Theory*; Max Born and Emil Wolf, *Principles of Optics*, Seventh (Expanded) Edition; P. M. Duffieux, *The Fourier Transform and Its Applications to Optics*, Second Edition; Reynolds, DeVelis, Parrent Jr., and Thompson, *The New Physical Optics Notebook*: *Tutorials in Fourier Optics*; and Masud Mansuripur, *Classical Optics and Its Applications*. My understanding of optics has benefitted from reading Chapter 12

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_9

"Wave Optics" in volume 1 (Gross, 2005a), and Chap. 21 "The Abbe Theory of Imaging" in volume 2 (Singer et al., 2005) of the six-volume *Handbook of Optical Systems* (Gross, 2005b).

I recommend a second set of books that may be easier to read for readers who are less familiar with Fourier analysis and physical optics: C. S. Adams and I. G. Hughes, *Optics f2f*; Ariel Lipson, Stephen G. Lipson, and Henry Lipson, *Optical Physics*, Fourth Edition; Robert Guenther, *Modern Optics*; Geoffrey Brooker, *Modern Classical Optics*; and Ian R. Kenyon, *The Light Fantastic: A Modern Introduction to Classical and Quantum Optics*. In addition, Jerome Mertz, *Introduction to Optical Microscopy*, Second Edition (2019) is an outstanding textbook when it comes to understanding optical imaging.

9.2 Mathematical Description of Abbe's Theory of Image Formation in the Microscope Based on Diffraction

In the absence of Abbe's promised publication of his detailed mathematical analysis of his theory of image formation in the microscope we must rely on the works published by his students and colleagues in the immediate decades following his death. In Abbe's 1873 publication we find mention of Fraunhofer diffraction and multiple uses of the term interference.

Further insight into probable reconstruction of Abbe's mathematical derivation of his theory of image formation in the microscope using Fourier analysis can be found in Porter's publication "On the diffraction theory of microscopic vision" (Porter, 1906). I discussed this publication and its impact in Chapter 7.

Furthermore, I recommend that the reader study Appendix A, which is devoted to publications relevant to Abbe's "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung" (Abbe, 1873a).

In this section I discuss the mathematical description of Abbe's 1873 theory of image formation in the light microscope. The widespread implementation of Fourier transforms in optics was motivated by P. M. Duffieux's 1946 French publication of *L'intégrale de Fourier et ses applications à l'optique* (Duffieux, 1946). Duffieux formulated the image of an object light distribution as its convolution with the point spread function (PSF) or the impulse response of the optical system. The spatial frequency spectrum of the image can be described as the product of the frequency spectrum of the object distribution (amplitude for coherent image formation or intensity for incoherent image formation) and the frequency response of the optical system. The concept of a frequency-dependent optical transfer function (OTF), in which the optical system transfers different spatial frequency components of the object onto the image plane, was conceived and developed by Duffieux (Duffieux, 1946). A recommended textbook for a modern development of Fourier optics, first developed by Duffieux, is Goodman, *Introduction to Fourier Optics*, Fourth Edition (Goodman, 2017).

I now introduce a mathematical treatment that follows the analysis of Gross et al. in Chap. 21, volume 2 of *Physical Image Formation, The Abbe Theory of Imaging Optics* (Singer et al., 2005) and *The New Physical Optics Notebook: Tutorials in Fourier Optics* (Reynolds et al., 1989).

Abbe discovered that a microscope objective corrected aplanatically will form a spectrum of the object in the rear focal plane, a.k.a. the Fourier plane or the pupil plane of the objective (Abbe, 1873b). An aplanatic lens is corrected for both spherical aberrations and coma. This aplanatic correction is currently known as the Abbe Sine Condition (it was derived and described in Chapter 7). The following mathematical analysis is described for an incoherent light source in the microscope (Singer et al., 2005).

This discussion involves image formation of an object (planar) in the image plane by optical Fourier transforms made by the lens and operation of a linear filter in the pupil plane. Figure 9.1 shows the optical layout. The typical Fourier optical system is labeled a 4*f* system and is composed of two 2*f* systems: object plane to pupil plane and pupil plane to image plane. Furthermore, Fig. 9.1 shows a 6*f* optical setup due to the additional 2*f* optics of the illumination lens. The light distribution of the object space is expanded in its spatial frequency distribution by Fourier transforms. The spatial frequency distribution in image space is calculated by multiplying the frequency spectrum in the object space by a complex transmission function (usually a low-pass filter).

The model that I discuss uses Köhler illumination (Köhler, 1893, 1894). It follows that the light source is in a Fourier plane with respect to the object plane (Fig. 9.1). Köhler invented an illumination system for the light microscope that solved the following problem (Köhler, 1893, 1894): How can the intensity of field illumination that does not contain variations in the brightness distribution of the light source be achieved? In 1893 Köhler invented the subsequently named Köhler

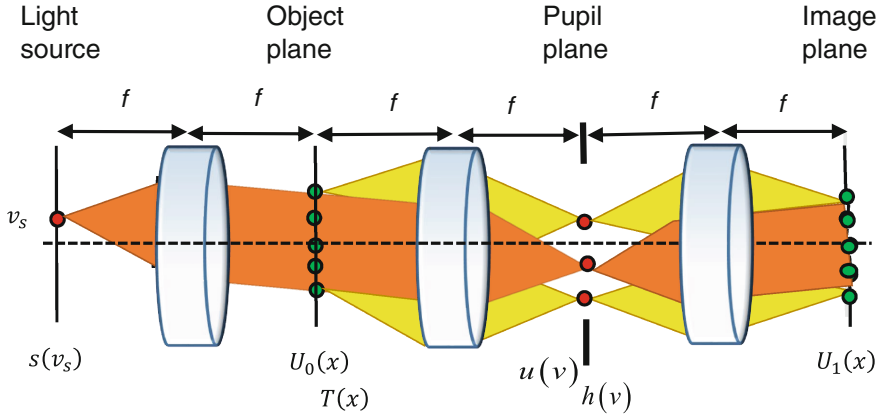


Fig. 9.1 Fourier theory of image formation in the microscope

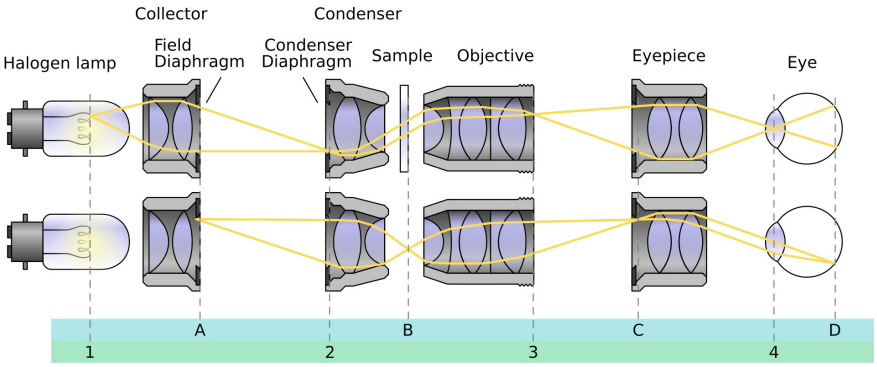


Fig. 9.2 Schematics of Köhler illumination. *Top*: Illumination beam path with conjugated planes (planes in common focus) marked on *green bar*. *Bottom*: Imaging beam path with conjugated planes marked on *blue bar*. The components lie in this order between the light source and the specimen and control illumination of the specimen. The collector/field lenses act to collect light from the light source and focus it at the plane of the condenser diaphragm. The condenser lens acts to project this light, without focusing it, through the sample. This illumination scheme creates two sets of conjugate image planes: one with the light source image and one with the specimen. Light source image planes (labeled on *green bar* in image) are found at: lamp filament (1), condenser diaphragm (2), back focal plane of the objective (3), and the eyepoint (4). Specimen image planes are found at: field diaphragm (A), specimen (B), intermediate image plane (the eyepiece diaphragm) (C), and eye retina or camera sensor (D) (https://en.wikipedia.org/wiki/Kohler_illumination, Creative Commons Attribution-ShareAlike License)

illumination system for microscopes. Köhler illumination provides uniform illumination of a specimen in the object plane (Fig. 9.2). A lens and a field stop image the light source onto the back focal plane of the condenser lens, and in the front focal plane of the condenser there is a variable aperture stop or an iris diaphragm that controls the intensity of light incident on the specimen. The light from each point on the filament of the source exits the condenser as a parallel light beam that provides uniform illumination in the object plane. The light that is incident on the specimen is from an extended region within the light source (incoherent illumination).

In summary, Köhler illumination gives the light microscope two major advantages: (1) the specimen is uniformly illuminated, and (2) the numerical aperture (NA) of the condenser and the size of the illuminated field can be independently adjusted.

I follow the description given in Chap. 21, "The Abbe Theory of Imaging," in Singer et al., (2005) and in Chapter 12, "Wave Optics," of volume 1 of *Handbook of Optical Systems: Physical Image Formation* (Gross, 2005a). The mathematical description of image formation in the light microscope that follows is shown for a one-dimensional representation.

The light field from the source denoted by $s(v_s)$ is Fourier-transformed by the lens, which is part of the illumination system, to yield a plane wave of the

illumination that is incident on the planar object, which is located on the microscope stage

$$S(x) = \int s(v_s) \cdot \exp(2\pi i v_s x) \mathrm{d}v_s \tag{9.1}$$

Next, thin-element approximation (TEA) is assumed. In TEA we assume that the z-dimension is neglected and the output wave is formed from the input wave by multiplying with a complex amplitude and phase mask. In Fig. 9.1 the x-direction is in the plane of the drawing and is perpendicular to the optic axis.

Assuming TEA the object located in the object plane is described by its complex transfer function

$$T(x) = A(x) \cdot \exp(i\phi(x)) \tag{9.2}$$

To obtain the light field distribution (object wave) that is located behind the object the incident light field distribution of the source is multiplied by the complex transfer function

$$U_0(x) = T(x) \cdot S(x) \tag{9.3}$$

The Fourier transform of $U_0(x)$ is located in the pupil plane of the optical system. The Fourier transform of the product of two functions can be mathematically expressed as the convolution of the frequency spectra of these two functions. The convolution operation is symbolized by $\otimes$, and the Fourier transform is symbolized by F. The light field distribution $u_0(v)$ of the pupil plane is then given by

$$u_0(v) = F\{U_0(x)\} = t(v) \otimes s(v) \tag{9.4}$$

Located in the pupil plane is the complex transmission function that mathematically expresses the spatial (frequency) filtering of both the aperture stop and the optical aberrations that are present in the optical system. This function is called the contrast transfer function (CTF) and is given by

$$h(v) = P(v) \cdot \exp\left(\frac{2\pi i}{\lambda} W(v)\right) \tag{9.5}$$

where $P(v)$ is the pupil function, and $W(v)$ is wavefront aberration.

Behind the aperture the light field distribution is given as a product of the field distribution $u_0(v)$ and the spatial filter function $h(v)$

$$u_1(v) = h(v) \cdot u_0(v) = h(v) \cdot [t(v) \otimes s(v)] \tag{9.6}$$

The light field in the image plane is given by another Fourier transform

$$U_1(x) = F\{h(v)\} \otimes [T(x) \cdot S(x)] = H(x) \otimes U_0(x) \tag{9.7}$$

The object wave U_0 is convoluted with the Fourier transform of the spatial filter function $h(v)$. The Fourier transform of the filter function is denoted by the amplitude distribution function $H(x)$

$$H(x) = \int h(v) \cdot \exp(2\pi i v x)\mathrm{d}v = \int P(v) \cdot \exp\left(\frac{2\pi i}{\lambda} W(v)\right) \exp(2\pi i v x)\mathrm{d}v \tag{9.8}$$

The filter $h(v)$ in frequency space is mathematically expressed in the image plane as the convolution of the object wave $U_0(x)$ and the amplitude distribution function $H(x)$

$$U_1(x) = \int U_0(x') \cdot H(x - x')dx' \tag{9.9}$$

However, light detectors measure image intensity. We can make the approximation that light intensity is given by the square of the amplitude

$$I_1(x) = |U_1(x)|^2 = |H(x) \otimes U_0(x)|^2 \tag{9.10}$$

Image formation in a light microscope with Köhler illumination can be thought of as the incoherent superposition of all the intensities to the coherent partial images of all the light source points. The light source can be considered as incoherent, coherent, or partially coherent, depending on its extension.

I now present an additional commentary on the mathematical description of Abbe's theory of image formation in the light microscope. The above simplified analysis is a scalar analysis. Several assumptions and approximations were made in scalar derivation. I mentioned the TEA where we neglect the z-dimension. In the simple model the effects of polarization and optical aberrations were also neglected. Although Fig. 9.1 shows the propagation of light waves in the microscope, the image can be considered as formed from many separate independent points of light (incoherent image formation). Further extensions of the mathematical analysis take into consideration whether the light source is incoherent, coherent, or partially coherent. Furthermore, the analysis assumed imaging with a 1:1 magnification.

The age-old question of the relation between the object and the image or the fidelity of mapping the object into its image in the light microscope is critical. The ideal image is formed by a diffraction-limited optical system. I quote Goodman's classic definition as an appropriate takeaway: "An imaging system is said to be diffraction-limited if a diverging spherical wave, emanating from a point-source object, is converted by the system into a new wave, again perfectly spherical, that converges toward an ideal point in the image plane, where the transverse location of that ideal image point is related to the transverse location of the original object point

through a simple scaling factor (the magnification), a factor that must be the same for all points in the image field if the system is to be ideal" (Goodman, 2017).

According to Abbe's 1873 theory the process of image formation in the light microscope can be understood as two Fraunhofer diffraction processes that occur sequentially. The assumptions state the object is uniformly and coherently illuminated with a plane wave. The reader should now understand that the microscope objective acts as a low-pass frequency filter that degrades the fine structures present in the object.

References

Abbe, E. (1873a). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.

Abbe, E. (1873b). Über einen neuen Beleuchtungsapparat am Mikroskop. Archiv für mikroskopische Anatomie, **IX**, 469–480.

Duffieux, P. M. (1946). *L'intégrale de Fourier et ses applications à l'optique*. Paris: Masson, Editeur. Republished in English as: Duffieux, P. M. (1970). *The Fourier Transform and its applications to optics*, second edition. New York: John Wiley and Sons.

Goodman, J. W. (2017). *Introduction to Fourier Optics*. Fourth Edition. New York: W. H. Freeman and Company.

Gross, H. (2005a). Wave Optics, in *Handbook of Optical Systems: Fundamentals of Technical Optics*, Volume 1, Chapter 12. Weinheim, FRG: Wiley-VCH Verlag GmbH & Co. KGaA.

Gross, H. (2005b). *Handbook of Optical Systems*. Volume 6, Weinheim, FRG: Wiley-VCH

Hopkins, H. H. (1951). The concept of partial coherence in optics. Proceedings of the Royal Society of London Series A, **208**, 263–277.

Hopkins, H. H. (1953). On the diffraction theory of optical images. Proceedings of the Royal Society of London Series A, **217**, 408–432.

Köhler, A. (1893). Ein neues Beleuchtungsverfahren für mikrophotographische Zwecke. Zeitschrift für wissenschaftliche Mikroskopie und für Mikroskopische Technik, **10**, 433–440.

Köhler, A. (1894). New method of illumination for photomicrographical purposes. Journal of the Royal Microscopical Society, **14**, 261–262.

Porter, A. B. (1906). On the diffraction theory of microscope vision. Philosophical Magazine, **6**, 154–166.

Reynolds, G. O., DeVelis, J. B., Parrent Jr., G. B., and Thompson, B. J. (1989). *The New Physical Optics Notebook: Tutorials in Fourier Optics*. Bellingham: SPIE Optical Engineering Press.

Singer, W., Totzeck, M., and Gross, H. (2005). The Abbe Theory of Imaging, *Handbook of Optical Systems: Physical Image Formation* Volume 2, Chapter 21. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

Van Cittert, P. H. (1934). Die Wahrscheinliche Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse beleuchteten Ebene. [The Probable vibrational distribution in one of the one light source directly or via a lens Illuminated plane]. Physica, **1**, 201–210.

Zernike, F. (1938). The concepts of degree of coherence and its application to optical problems. Physica, **5**, 785–795.

Further Reading

Adams, C. S., and Hughes, I. G. (2019). *Optics f2f.* Oxford: Oxford University Press.

Abbe, E. (1874). A contribution to the theory of the microscope and the nature of microscopic vision. Translated into English by H. E. Fripp. Proceedings of the Bristol Naturalists Society, I, 202–258. Read before the Bristol Microscopical Society, December 16, 1974.

Abbe, E. (1882a). The relation of aperture and power in the microscope. Journal of the Royal Microscopical Society, Section **II**, 300–309. [read before the Society on May 10, 1882], [Abbe wrote the paper in English].

Abbe, E. (1882b). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Section **II**, 460–473. [read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Abbe, E. (1889). On the effect of illumination by means of wide-angled cones of light. Journal of the Royal Microscopical Society, Series **II**, IX, 721–724.

Abbe, E. (1989). *Gesammelte Abhandlungen*, I–IV. Hildesheim: Georg Olms Verlag. [Originally published in 1904, Jena: Verlag von Gustav Fischer].

Abbe, E. (1883). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Section **II**, 790–812. [read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Berek, M. (1929). XXI. On the extent to which real image formation can be obtained in the microscope. Journal of the Royal Microscopical Society, **49**, 240–249.

Born, M., and Wolf, E. (1999). *Principles of Optics*, 7th (expanded) edition. Cambridge: Cambridge University Press.

Bratt, J. and Török, P. (2019). *Imaging Optics*. Cambridge: Cambridge University Press.

Feffer, S. M. (1994). *Microscopes to munitions: Ernst Abbe, Carl Zeiss, and the transformation of technical optics, 1850–1914*. PhD Dissertation, University of California, Berkeley, 1994. Ann Arbor: UMI Dissertation Services.

Gu, M. (2000). *Advanced Optical Imaging Theory*. Berlin: Springer.

Köhler, H. (1981). On Abbe's theory of image formation in the microscope. Journal of Modern Optics, **28**, 1691–1701.

Linfoot, E. H. (1964). *Fourier Methods in Optical Image Evaluation*. London: Focal Press, Ltd.

Lummer, O., and Reiche, F. (1910). *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe*. Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.

Mansuripur, M. (2009). Classical Optics and its Applications. Second Edition. Cambridge: Cambridge University Press.

Martin, L. C. (1966). *The Theory of the Microscope.* New York: American Elsevier Publishing Company, Inc. and London: Blackie.

Mertz, J. (2019). *Introduction to Optical Microscopy*, second edition. New York: Cambridge University Press.

Michel, K. (1964). *Die Grundzüge der Theorie des Mikroskops in elementarer Darstellung*. 2. neubearbeitete Auflage. Stuttgart: Wissenschaftliche Verlagsgesellschaft M.B. H.

Williams, C. S., and Beckland, O. A. (1989). *Introduction to the Optical Transfer Function*. New York: John Wiley & Sons.

Volkmann, H. (1966). Ernst Abbe and His Work. Applied Optics, **5**, 1720–1731.

von Rohr, M. (1940). *Ernst Abbe*. Jena: Gustav Fischer.

Zernike, F. (1934a). Beugungstheorie des Schneidenverfahrens und Seiner Verbesserten Form, der Phasenkontrastmethode. [Diffraction theory of the cutting process and its improved form of the phase contrast method] Physica, **1**, 689–704.

Zernike, F. (1934b). Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. Monthly Notices of the Royal Astronomical Society, **94**, 377–384.

Zernike, F. (1936). Deutsches Reichspatent No. 636168 (September 1936).

Zernike, F. (1942a). Phase contrast, a new method for the microscopic observation of transparent objects, Part 1. Physica, **9**, 686–698.

Zernike, F. (1942b). Phase contrast, a new method for the microscopic observation of transparent objects, Part 2. Physica, **9**, 974–986.

Zernike, F. (1948). Diffraction and optical image formation. Proceedings of the Physical Society, **61**, 158–164.

Zernike, F. (1950). Color phase-contrast microscopy: requirements and applications. Physica, **9**, 974–986.

Zernike, F. (1958). The wave theory of microscopic image formation. Appendix K. *In: Concepts of Classical Optics*, J. Strong, pp. 525–536. San Francisco: W. H. Freeman.

Zernike, F., and Brinkman, H. C. (1935). Hypersphärische Funktionen und die in sphärischen Bereichen orthogonalen Polynome. Verh. Akad. Wet. Amst., (Proceedings Royal Academy Amsterdam), **38**, 161–170.

Part II
Optical Techniques to Enhance Contrast in the Microscope

Introduction

In Part I we explored the microscopic world aided by the invention of the microscope. There were many questions concerning the validity of microscopic images and the complex, recurrent, and confounding problems associated with artifacts. These questions persist and require control studies to validate images.

Early concepts of resolution and resolving power were first developed for telescopes and then applied to microscopes. But experience showed that the purported analogy between the two instruments was specious. The manufacture of microscopes and their optical components was an empirical procedure and devoid of a theoretical foundation. The seminal experimental and theoretical studies of three independent investigators provided the basis for understanding resolution and image formation in the microscope. Working independently and using their own unique approaches Abbe, Helmholtz, and Rayleigh developed similar mathematical expressions for limiting resolution of the light microscope. As Abbe pointed out, limiting resolution was valid only under the physical conditions he carefully outlined in his 1873 publication. Abbe suggested that resolution of the light microscope could be enhanced in two ways: by reducing the wavelength of illumination light and by increasing the numerical aperture of the microscope objective. I pointed out in Part I that Abbe was the first to suggest reducing the wavelength to enhance resolution. This was eventually incorporated in ultraviolet, fluorescence, and electron microscopes.

I expand in Part II on inventions and the development of techniques based on Abbe's prescient suggestions as well as his decision to staff Zeiss Werke with physicists and chemists. Abbe's emphasis on a strong physical foundation for the development of new types of microscopes and other optical instruments is evident in the invention of the ultramicroscope. In Part II, in addition to discussing how contrast generation in the microscope can be enhanced, I discuss a general approach in which a problem or a limitation of existing technology occurs and then

technological solutions are found to solve or mitigate these problems. We see this approach in the development of light-sheet microscopy that evolved from prior invention of the ultramicroscope at Zeiss Werke.

Within the chapters that comprise Part II (Chapters 10–12) there are remarkable lessons to be learned. There are connections and technology transfer between technical and theoretical innovations in disparate fields and applications. This is illustrated in the discussion of the ultramicroscope that had the unique capability to locate (not image) colloidal particles in solution. Richard Zsigmondy's invention of the ultramicroscope and its subsequent use to investigate colloid chemistry ushered in the field of nanotechnology for which the inventor received the Nobel Prize in Chemistry. This invention provided the basis for the subsequent invention of the light-sheet microscope and its myriad successful applications in developmental biology and neuroscience.

In Chapter 8 (Part I) I discussed Zernike's discovery of phase microscopy. An interesting point to come out of Zernike's Nobel Lecture was how close Abbe came to discovering phase microscopy. Zernike explained how Abbe, despite his greatness as an innovator in microscopy, actually inhibited the development of phase microscopy.

With the invention and development of new microscopic techniques came new limitations. Photo-bleaching of specimen fluorescence and light damage to living specimens are limitations associated with many new and exciting microscopic techniques. There is a paucity of carefully designed and controlled experiments to understand these common problems. My objective in Part II is to discuss the limitations of each type of microscopy and come up with possible solutions. I also discuss in Part II the principles, instrumentation, applications, and the limitations of several new categories of optical microscopies such as the ultramicroscope, light sheet fluorescence microscopes, phase contrast microscopes, differential interference contrast microscopes, and modulation contrast microscopes that enhance the contrast of the specimen.

Chapter 10
Richard Zsigmondy and Henry Siedentopf's Ultramicroscope

10.1 Introduction

The concept that led to the invention of the ultramicroscope, a new type of dark-field microscope designed for the investigation of colloids, in 1903 by Richard Adolf Zsigmondy (1865–1929) is not contested (Cahan, 1996). Zsigmondy is credited with the concept of the ultramicroscope and its application to the field of colloid chemistry; he tested the ultramicroscope with various colloids and devised the technique in which the ultramicroscope is used to measure the size of ultramicroscopic colloidal particles (Cahan, 1996). Zsigmondy sometimes worked independently and other times in collaboration with the physicist Henry Friedrich Wilhelm Siedentopf (1872–1940) of Zeiss Werke. What was Siedentopf's role? He constructed the prototype ultramicroscope and subsequent versions and optimized their optical performance (Cahan, 1996).

Their collaboration is significant for several reasons. First, this is a success story involving collaboration between two researchers with disparate expertises. This is an example of the enormous benefits that can follow from interdisciplinary collaboration. Second, the microscope was conceived, designed, and constructed to solve a specific problem: how to study submicroscopic colloids. Colloids were discovered in 1861 by Thomas Graham. A colloid, or a colloidal dispersion, consists of large molecules or submicroscopic particles of one type that are dispersed in a second substance. Examples include colored glass, clouds, smoke, milk, and gelatin. Individual colloidal particles have diameters between 1 and 1000 nm. Third, the ultramicroscope does not image individual colloidal particles; it makes possible their localization and visualization. Localization is different from imaging. With the ultramicroscope colloidal particles of dimensions below the Abbe and Helmholtz limits of resolution in a light microscope can still be observed (i.e., visualized but not imaged). Siedentopf used the words "rendering visible" in his publications (Siedentopf, 1903). Fourth, the invention of the ultramicroscope is the antecedent of the modern technique of light-sheet fluorescence microscopy that I

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_10

describe in Chapter 11 (Huisken et al., 2004; Stelzer and Lindek, 1994; Voie, Burns, and Spelman, 1993).

I begin with a biographical introduction to the lives and works of Zsigmondy and Siedentopf (Cahan, 1996; Sönnichsen and Fritzsche, 2007; Zsigmondy, 1925). Zsigmondy was born in Vienna a year before Graham discovered colloids in 1865. The study of colloids was inextricably connected to Zsigmondy's interests and his professional life. His early interdisciplinary studies in chemistry and physics set the foundation for his future research in colloid chemistry. However, he majored in organic chemistry at the Technische Hochschule in Vienna and then at the Technische Hochschule in Munich. His doctorate in organic chemistry was awarded in 1889 by the University of Erlangen (Sönnichsen and Fritzsche, 2007) (Fig. 10.1).

A remarkable change in his interests occurred between 1891 and 1892 under his mentor Professor August Kundt at the University of Berlin when he became fascinated with inorganic chemistry and thus began his lifelong work on colloid chemistry. What was the connection between inorganic chemistry and colloid chemistry? The answer is glass; specifically, gold red glass. What interested Zsigmondy was the nature of the gold particles in gold red glass: were they colloidal particles? Other researchers claimed that the gold particles were in suspension. His journey from organic chemistry to inorganic chemistry and the study of gold red glass developed during his Habilitation work at the Technische

Fig. 10.1 Richard Adolf Zsigmondy

Hochschule Graz, which allowed him to become a lecturer in glass technology. His knowledge of glass and the production of colored glass led him to work at the Schott glass factory where he studied gold red glass for seven years. It was while Zsigmondy worked at the Schott glass factory that he met Siedentopf. After leaving the Schott glass factory Zsigmondy worked in Jena as an independent colloid chemist.

The then emerging field of colloid chemistry had important applications to various industries despite much controversy over the nature of colloidal particles and their roles in producing the wide variety of observable phenomena. As far as Zsigmondy was concerned chemical analysis of colloids had reached a dead-end; therefore, he reached out to physical techniques (in particular, optical techniques such as microscopy). Nevertheless, the standard bright-field light microscope could not resolve individual colloidal particles as they were below the resolution limits derived in 1873 by Abbe and in 1874 by Helmholtz.

What followed next is a lesson in the process of discovery in science. Success in research often depends on asking the right questions. This involves formulating questions in such a way that they can be experimentally addressed. Often this drives the invention and development of new instrumentation.

Zsigmondy cogently formulated several critical questions (Cahan, 1996). First, what is the size and composition of the gold particles? How many atoms compose one particle? Second, he asked a question about the scattered light that forms the Faraday–Tyndall light cone present in all colloidal solutions (van de Hulst, 1957). Zsigmondy asked whether the Faraday–Tyndall light cone is specific to gold particles in solution (van de Hulst, 1957). In that way he could experimentally answer the contentious question: are the gold particles in glass a colloidal solution or a suspension? His initial experiments using standard light microscopes failed; what was required was a microscope that could "render visible" the scattered light of the light cone from individual gold particles. From his preliminary analysis of the distance between individual gold particles he concluded that this distance was less than the wavelength of the incident light. Zsigmondy could "render visible" particles that exceeded the resolution limits given by both Abbe and Helmholtz. In the next section I describe the design, development, and applications that followed on from these inchoate investigations.

In 1905 Zsigmondy published his seminal book *Zur Erkenntnis der Kolloide. Über irreversible Hydrosole und Ultramikroskopie* (Zsigmondy, 1905). This work summarized his investigations into the nature of colloids and the ultramicroscope. Zsigmondy departed Jena in 1907 and settled in Göttingen as an associate professor at the University of Göttingen. The following year he became director of the Institute of Inorganic Chemistry.

In 1919 Zsigmondy became a full professor. The Nobel Prize in Chemistry 1925 was awarded to Zsigmondy (in 1926) "for his demonstration of the heterogeneous nature of colloid solutions and for the methods he used, which have since become fundamental in modern colloid chemistry" (Zsigmondy, 1926). Zsigmondy used his interdisciplinary knowledge of chemistry and physics to invent new instrumentation and to apply new techniques to advance colloid chemistry (Sönnichsen and Fritzsche, 2007).

Next, I present some information about Zsigmondy's collaborator Henry Friedrich Wilhelm Siedentopf (1872–1940). Siedentopf is known for his work in connection with the 1904 invention of the ultraviolet microscope by August Köhler and Moritz von Rohr at Zeiss Werke. Their motivation was to increase the resolution of the light microscope by reducing the wavelength of illumination. Shortly afterward Siedentopf made an amazing discovery: the ultraviolet light used for illumination in the ultraviolet microscope caused the object to fluoresce. But the fluorescence diminished the contrast of the image of the object! This observation led in 1908 to the development of a prototype fluorescence microscope by Köhler and Siedentopf at Zeiss Werke.

A short biographical introduction to Siedentopf is appropriate (Sönnichsen and Fritzsche, 2007). Siedentopf studied physics at Leipzig University and received his doctorate from the University of Göttingen. From 1899 to 1938 he worked at Zeiss Werke in Jena where he became director of the microscopy division in 1907. Both Siedentopf and Zsigmondy independently studied the Tyndall light cone from colloids (Sönnichsen and Fritzsche, 2007). Their collaborative research led to the development in 1902 of a new type of light microscope: the ultramicroscope. The ultramicroscope could "render visible" submicroscopic colloidal particles that were seen as points of light. Further developments led to the invention of their immersion ultramicroscope in 1913 and numerous advances and discoveries in the field of colloid science (Siedentopf, 1903; Siedentopf and Zsigmondy, 1902; Zsigmondy, 1905, 1907, 1909, 1913, 1920).

10.2 The Ultramicroscope: Design, Development, and Applications

The principle that is fundamental to the ultramicroscope is easily understood by analogy. A beam of sunlight enters a small hole in a dark room. The observer is positioned to observe the light beam from a direction that is perpendicular to the direction of propagation of the light beam. When small particles of dust in the air are transported across the light beam the observer will see very small spots of light that result from dust particles scattering the incident light.

In Zsigmondy's initial experiments the object was a cube of glass with gold colloidal particles dispersed within the glass. The source of illumination was the sun. A heliostat was used to track the relative movement of the sun and the earth. A lens was used to focus sunlight on a very small area within the glass cube. The glass cube was observed with a standard upright light microscope. The optical axis of the light microscope was perpendicular to the axis of illumination to maintain the critical constraint that no illumination light would directly enter the optical axis of the microscope. Zsigmondy was able to observe the light cone due to light scattered from individual colloidal particles (Zsigmondy, 1926).

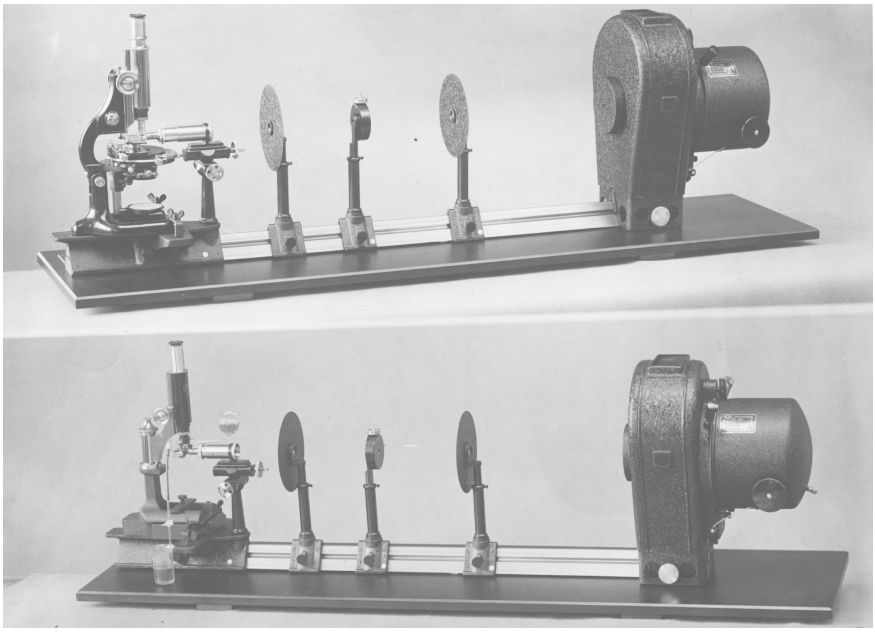


Fig. 10.2 Slit ultramicroscope after Siedentopf and Zsigmondy. Reproduced with permission from the ZEISS Archives

As Zsigmondy explained in his Nobel Lecture, it was the motivation to improve his first prototype of the ultramicroscope that led him to collaborate with Siedentopf at Zeiss Werke.

The next version of the ultramicroscope (shown in Fig. 10.2) was the slit ultramicroscope. It incorporated several changes that improved the capability of observing Tyndall light cones from individual colloidal particles. Optical components from the light source to the upright light microscope were mounted on an optical bench to provide stability. The light source was an arc lamp. A telescope objective focused the image of the light source onto an adjustable slit and a microscope objective acted as a condenser to focus the image of the slit onto the colloidal solution that was in a small dish. The optical axis of the upright microscope was perpendicular to the optical axis of the illumination system and the optical bench maintained mechanical stability.

Using the new slit ultramicroscope Zsigmondy determined the size of gold particles in both solution and in glass, the gold content of the solution was known (Zsigmondy, 1926). Furthermore, in his Nobel Lecture Zsigmondy explains that several improvements were made in the design of different versions of the ultramicroscope. He credits Siedentopf with the design of a paraboloid and a cardioid condenser, Reichert with the design of a mirror condenser, and Ignatowski and Jentzsch with the design of a ball condenser (Zsigmondy, 1926). In 1912 he invented the

immersion ultramicroscope to investigate colloids in liquids (Mappes et al., 2012). The immersion ultramicroscope was the result of collaborative work between Zsigmondy and the Göttingen microscope company Rudolf Winkel (Mappes et al., 2012; Zsigmondy, 1926).

Further insight into the instrumentation and applications of the ultramicroscope are contained in two publications by the coworkers in the development of this new type of microscope (Siedentopf, 1903; Siedentopf and Zsigmondy, 1902). In their joint paper Siedentopf and Zsigmondy introduced a new technique to visualize particles smaller than the Abbe and the Helmholtz resolution limits for diffraction-limited light microscopes (1902). To implement their technique the optical system had the following requirement: the diffraction cone of the microscope objective oriented vertically along the optical axis of the microscope and perpendicular to that axis is the horizontally oriented illumination cone of the condenser (Siedentopf and Zsigmondy, 1902). In other words, the light from the condenser light cone is unable to enter the microscope objective and propagate to the eye of the observer. But light from the particle's diffraction light cone can enter the microscope objective and thus "render visible" submicroscopic colloidal particles.

In subsequent sections of their publication Zsigmondy described his technique to determine the diameters of colloidal particles using the slit ultramicroscope. Zsigmondy then applied his new technique to investigate the question: what is the relationship between the diameter of the gold particles and the color of gold red glass? There was no correlation in his preliminary studies.

The authors made a prescient claim in their paper: they proposed that their invention of the slit ultramicroscope was ideally suited to the investigation of Brownian motion, discovered in 1828 by Robert Brown, in liquids (Siedentopf and Zsigmondy, 1902). Einstein and Smoluchowski (independently) derived theoretical models of Brownian movement based on mean square displacements of particles (Einstein, 1905, 1906; Smoluchowski, 1906). Experimental validation of the independent theories of Einstein and Smoluchowski on Brownian motion is the work of Jean Perrin who also determined an accurate value for Avogadro's number (Perrin, 1909). A major effect of Perrin's experimental work on Brownian motion was that it gave strong experimental support to the claim that atoms existed.

In 1903 Siedentopf published his seminal paper "On the rendering visible of ultra-microscopic particles and of ultra-microscopic bacteria" (Siedentopf, 1903). Two things that have great pedagogical value struck me on reading the original paper. First, the degree of caution in the writing is exceptional. Second, Siedentopf's prescient suggestion that the ultramicroscope may be useful in imaging cells and tissues is impressive. The author clearly gives words of caution on the capabilities of the ultramicroscope. He states that the ultramicroscope does not "render visible" the correct shape and size of submicroscopic colloidal particles. Although the image of an ultramicroscopic particle is always a small diffraction disk, the author proposes that a perfected version of the ultramicroscope may be useful for bacteriologists to "render visible" bacteria that are so far unknown.

The invention of the ultramicroscope led to advances in the new science of colloidal chemistry, to validation of the theories of Brownian motion that gave strong support for the existence of atoms, and to the modern development of light-sheet microscopy.

References

Cahan, D. (1996). The Zeiss Werke and the ultramicroscope: the creation of a scientific instrument in context. In: J. Z. Buchwald ed., Archimedes, volume **1**, Dordrecht: Kluwer Academic Publishers.

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik, **17**, 549–560.

Einstein, A. (1906). Zur Theorie der Brownschen Bewegung. Annals of Physics, **19**, 371–381.

Huisken, J., Swoger, J., Del Bene, F. Wittbrodt, J., and Stelzer, E. H. K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science, **305**, 1007–1009.

Mappes, T., Jahr, N., Csaki, A., Vogler, N., Popp, J., Fritzsch, W. (2012). The Invention of Immersion Ultramicroscopy in 1912-The Birth of Nanotechnology? Angewandte Chemie International Edition, **51**, 11208–11212.

Perrin, J. (1909). Mouvement brownien et réalité moléculaire [Brownian movement and molecular reality].

Siedentopf, H. and Zsigmondy, R. (1902). Über Sichtbarmachung und Größenbestimmung ultramikoskopischer Teilchen, mit besonderer Anwendung auf Goldrubingläser [About visualization and size determination of ultramicroscopic particles, with particular application to gold red glasses]. Annalen der Physik, **315**, 1–39.

Siedentopf, H. (1903). IX. On the rendering visible of ultra-microscopic particles and of ultra-microscopic bacteria. Journal of the Royal Microscopical Society, **23**, 573–578.

Smoluchowski, M. (1906). Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. Annals of Physics, **326**, 756–780.

Sönnichsen, C., and Fritzsche, W. (2007). *100 years of nanoscience with the ultramicroscope. The work of Richard Zsigmondy.* Aachen: Shaker Verlag.

Stelzer, E. H. K., and Lindek, S. (1994). Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: confocal theta microscopy. Optics Communications, **111**, 536–547.

van de Hulst, H. C., (1957). *Light scattering by small particles*. New York: John Wiley & Sons.

Voie, A. H., Burns, D. H., and Spelman, F. A. (1993). Orthogonal-plane fluorescence optical sectioning: three-dimensional imaging of macroscopic biological specimens. Journal of Microscopy, **170**, 229–236.

Zsigmondy, R. (1905). *Zur Erkenntnis der Kolloide. Über irreversible Hydrosole and Ultramikroskopie.* Jena: Verlag von Gustav Fischer.

Zsigmondy, R. (1907). *Über Kolloid-Chemie mit besonderer Berücksichtigung der anorganischen Kolloide.* Leipzig: Verlag von Johann Ambrosius Barth.

Zsigmondy, R. (1909). *Colloids and the Ultramicroscope: A Manual of Colloid Chemistry and Ultramicroscopy,* tr. Jerome Alexander. New York: Wiley.

Zsigmondy, R. (1913). Über ein neues Ultramikroskop. Physikalische Zeitschrift, **14**, 975–979.

Zsigmondy, R. (1920). *Kolloidchemie, ein Lehrbuch*, Dritte Auflage. Leipzig: Verlag von Otto Spamer.

Zsigmondy, R. (1925). "Zsigmondy - Biographical". Nobelprize.org. Nobel Media AB 2014. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1925/zsigmondy-bio.html (Accessed on April 1, 2019).

Zsigmondy, R. (1926). "Richard Zsigmondy – Nobel Lecture: Properties of Colloids". Nobelprize.org. Nobel Media AB 2014. Nobel Lecture, December 11, 1926. https://www.nobelprize.org/uploads/2018/06/zsigmondy-lecture.pdf (Accessed April 1, 2019).

Chapter 11
Light-Sheet Fluorescence Microscopy

> "What has the light-sheet microscope brought us in 10 years? We have seen hearts beating in real time, the full development of various embryos from eggs to viable and fertile adults, bacteria colonizing sterile guts, growing microtubule asters, plant roots growing for weeks, neuronal activity in live fish brains and deep views inside tissues on scales that had seemed inaccessible to microscopy. Light-sheet microscopy makes it possible to image any organism in a near-physiological context without damage and loss of signal. These microscopes can open windows into living cells to reveal the spectacle of biological processes with high resolution in space and time."
>
> —Reynaud, Peychl, Huisken, and Tomancak (2015).

11.1 Introduction

In light-sheet fluorescence microscopy (LSFM) the optical axes of the image-forming microscope objective and the objective providing illumination of the specimen are at right angles to one another (Olarte et al., 2018). A sheet of light is incident on the specimen, and fluorescence is produced and hence detected only in that plane. To obtain images of multiple planes for three-dimensional reconstruction the light-sheet is displaced relative to the specimen either by scanning the illumination or by moving the specimen through the light-sheet (Huber et al., 2001). LSFM has the intrinsic capability of optically sectioning the specimen. An LSFM differs markedly from epi-illumination fluorescence microscopes, in which illumination is incident on the entire specimen. The key advantage of LSFM lies in the associated decreased photobleaching and phototoxicity in the specimen—up to three orders of magnitude, depending on the specimen—as compared with an epi-illumination fluorescence microscope (Reynaud et al., 2008).

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_11

The invention and development of LSFM was inspired by the work of Richard Adolf Zsigmondy and Wilhelm Siedentopf, the inventors of the ultramicroscope that would "render visible" submicroscopic colloidal particles (Chapter 10).

Innovation in instrumentation is often helped through knowledge of advances in disparate fields, as is made clear in a review of the early antecedents of LSFM (Huisken and Stainier, 2007, 2009). Prior to modern innovations in LSFM light-sheets were used in the field of photography. The problem addressed was how to extend the depth of field (Zampol, 1960). Light-sheets relieved problems of photomicrography related to deep fields of view (McLachlan, 1968; Simon, 1965). In another invention, the object is displaced through a light-sheet and the camera collects the light and is focused on the plane illuminated by the light-sheet (Huber et al., 2001).

In some of the early LSFM developments the nature of the specimen was a design consideration. Voie et al. employed what they called orthogonal-plane fluorescence optical sectioning (OPFOS) to investigate a guinea pig cochlea stained with a fluorescent dye (Voie, 1996, 2002; Voie and Spelman, 1995; Voie et al., 1993). In their instrument the light-sheet was stationary and the specimen, of the scale of a few millimeters, was rotated with respect to the light-sheet. Lateral resolution was 10 μm and axial resolution 26 μm. Concurrently, Stelzer's group, motivated by the desire to improve axial resolution, was developing a confocal microscope based on an oblique illumination system (Lindek and Stelzer, 1994; Lindek et al., 1994; Stelzer et al., 1995). In their 1995 publication they cited Voie's development of OPFOS. Stelzer's group developed theta confocal microscopy, which was further developed into a new variant of LSFM termed selective plane illumination microscopy (SPIM). Another example is the use of LSFM in the field of oceanography where the specimens are microbes in ocean water samples (Fuchs et al., 2002).

Modern innovation in LSFM has resulted in advances in both neuroscience and developmental biology. The utility of LSFM in the biomedical sciences is reflected by LSFM being selected as the Method of the Year 2014 by *Nature Photonics* (Keller et al., 2015).

In this and the following chapters on superresolution optical microscopes I discuss the invention and development of new types of microscopes that as a result of two decades of innovation have increased resolution and image contrast. I first present the limitations of existing microscopes and then explain how new instrumental developments either mitigate or overcome these limitations. To do this I attempt to answer a number of questions: What was the problem investigators sought to solve? What was the solution they came up with? How did the newly developed or invented instrument work? What was achieved with this new instrument? And, finally, what were the limitations, cautions, and trade-offs associated with this new instrument?

11.2 Review of Instrument Design, Capabilities, and Limitations of Light-Sheet Fluorescence Microscopy

In 2008 Reynaud et al. published a general introduction to the field of LSFM that provides a link between Siedentopf and Zsigmondy's 1903 invention of the ultra-microscope and this chapter on LSFM (Reynaud et al., 2008; Siedentopf and Zsigmondy, 1903).

Figure 11.1 shows how a specimen is illuminated from the side with a thin light-sheet that illuminates each plane successively. The center of the light-sheet overlaps with the focal plane of the detection system, and fluorescence is detected along an axis orthogonal to the illumination axis. The fluorophores that are located above and below the focal plane are not illuminated and therefore do not contribute to photobleaching or phototoxicity. To obtain a three-dimensional image with isotropic resolution the LSFM obtains multiple views of the same volume but in various directions.

It was also in 2008 that the Stelzer group invented digital scanned laser light-sheet fluorescence microscopy (DSLM) and demonstrated its capability for quantitative in vivo imaging of entire Japanese killifish embryos (Keller and Stelzer, 2008). The problem they addressed was how to significantly increase the imaging speed and the signal-to-noise ratio (SNR) and at the same time decrease the energy of the light-sheet. Their solution was DSLM, which achieved a 50-fold increase in imaging speed and a 10- to 100-fold increase in signal-to-noise ratio. DSLM uses a laser scanner to displace the light-sheet through the specimen. Good depth of penetration, critical in the imaging of entire embryos, is achieved by using

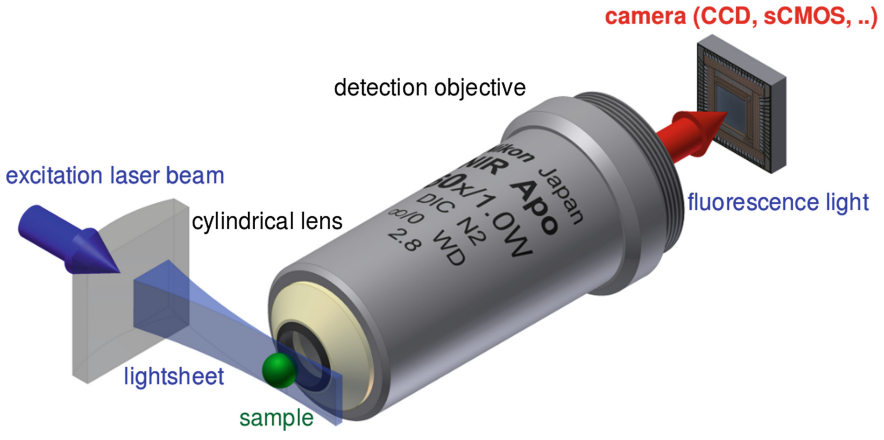


Fig. 11.1 Principle of light-sheet fluorescence microscopy (LSFM). From https://en.wikipedia.org/wiki/Light_sheet_fluorescence_microscopy. Creative Commons Attribution-ShareAlike License

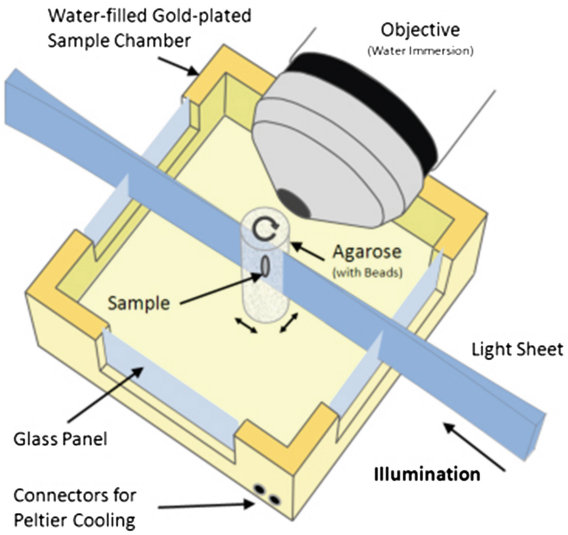


Fig. 11.2 Principle of selective-plane illumination microscopy (SPIM). SPIM technology offers fast optical sectioning and minimally invasive three-dimensional acquisition of specimen fluorescing over time. It does so by focusing a thin laser light-sheet into the specimen and taking two-dimensional images of the illuminated slice with a perpendicularly positioned detector (CCD camera). Three-dimensional stacks are obtained by moving the specimen orthogonal to the light-sheet between consecutive images. By mounting the sample in a rigid medium (e.g., agarose) and hanging it inside the sample chamber in front of the detection lens it is possible to rotate the sample and collect three-dimensional stacks from multiple angles (views). *CCD*, Charge-coupled device. From http://openspim.org/Welcome_to_the_OpenSPIM_Wiki. Content is available under Creative Commons Attribution Share Alike unless otherwise noted

an illumination microscope objective with a low NA. By imaging the same volume of the specimen along multiple directions the effects of shadows in the image can be greatly reduced. Their invention of the DSLM follows the group's previous invention of SPIM (Fig. 11.2) (Huisken et al., 2004).

The next development solved two ubiquitous problems associated with LSFM (Huisken and Stainier, 2007). Their paper "Even Fluorescence Excitation by Multidirectional Selective Plane Illumination Microscopy (mSPIM)" describes technical extensions of light-sheet-based microscopy by multidirectional illumination (Huisken and Stainier, 2007).

What were the problems the authors sought to solve? First, the problem of shadowing in the excitation path that resulted from light absorption by the specimen. Second, the problem of the spread of limits to the light-sheet due to light scattering within the specimen. To address these problems they pivoted the light-sheet and illuminated the specimen sequentially from two opposite directions, the two resulting images being fused to form a single enhanced image. The authors validated the capability of mSPIM by imaging live zebrafish embryos. Their mSPIM microscope was similar to other LSFM optical designs that were optimized

for fixed tissue such as (high-resolution) orthogonal plane fluorescence optical sectioning (HR-OPFOS) (Buytaert and Dirckx, 2007; Voie et al., 1993).

Subsequently, Huisken and Stainier authored a comprehensive and critical review of selective-plane illumination microscopy techniques in developmental biology (Huisken and Stainier, 2009).

In 2007 Buytaert and Dirckx, working at the University of Antwerp, authored a comprehensive paper titled "Design and Quantitative Resolution Measurements of an Optical Virtual Sectioning Three-Dimensional Imaging Technique for Biomedical Specimens, Featuring Two-Micrometer Slicing Resolution" (Buytaert and Dirckx, 2007). In addition to presenting an innovative LSFM their publication is exemplary in its clarity of instrument description, evaluation, and calibration. There are important lessons to be learned from the authors' review of the history of the field and the limitations of alternative microscopic techniques. Included in their paper are critiques describing the lack of citations and missing details on resolution in previous publications by other groups; the deficiencies of these groups reflect the quality of peer reviews that failed to correct these deficiencies. According to Buytaert and Dirckx, for example, the group of Huisken et al. in 2004 used a similar technique invented by Voie et al. in 1993 to image live embryos. Huisken et al. named their LSFM technique selective-plane illumination microscopy (SPIM) (Huisken et al., 2004). Huisken et al. did not cite the original work of Voie et al., (1993), but the SPIM method is essentially identical to the earlier OPFOS technique, apart from the fact that imaging is performed on very small objects so that higher resolutions can be obtained (Buytaert and Dirckx, 2007).

Buytaert and Dirckx compare mechanical sectioning with Voie et al.'s development of orthogonal plane fluorescence optical sectioning (OPFOS) microscopy (Voie et al., 1993). In the OPFOS microscope a hyperbolic light-sheet of constant thickness is formed in the focal region using a cylindrical lens. The problem to be solved was having to make a trade-off between image width and axial resolution. With a low numerical aperture (NA) the light-sheet maintains the same thickness over a comparatively large distance. With a high NA the axial resolution is greater, but this gain is only over a smaller distance. In the lateral plane the resolution is diffraction limited, but in the axial direction the resolution of the OPFOS instrument is 14 μm.

The solution to the problem of having to make a trade-off between specimen size and image resolution was the invention of a new type of LSFM, the high-resolution OPFOS (HR-OPFOS) (Buytaert and Dirckx, 2007). As explained by the authors the usual light-sheet is formed by passing a Gaussian beam through a cylindrical lens. The light-sheet has its smallest thickness at the focal point (i.e., the beam waist), the sheet thickness increasing on both sides of the waist. As illustrated in Fig. 11.3 the confocal parameter is defined as twice the Rayleigh range for the Gaussian beam (i.e., the distance for which the beam thickness is essentially constant). The persistent problem in LSFM is that the image will be in focus only over the length of

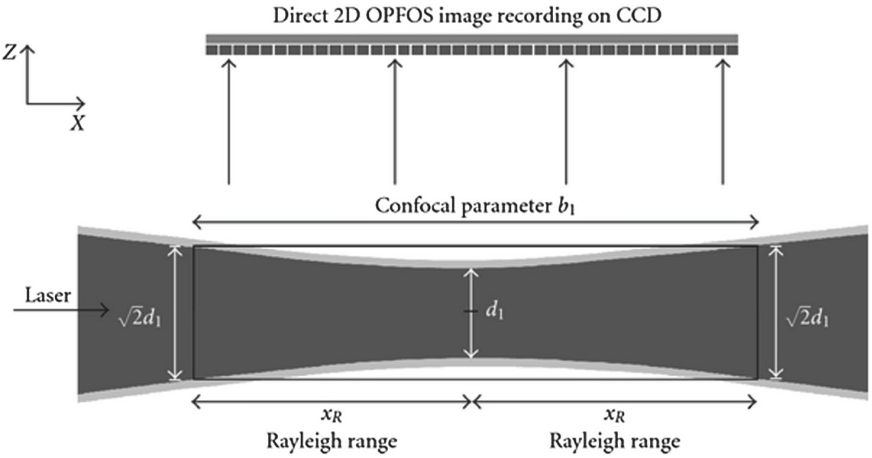


Fig. 11.3 Hyperbolic focus profile of a cylindrical lens. OPFOS records two-dimensional images in an approximated planar sheet defined by the confocal parameter zone b_1 where the thickness is considered constant at $\sqrt{2}d_1$. The *dark-gray area* in the center represents the $1/e^2$ intensity profile. *CCD* , Charge-coupled device; *OPFOS*, orthogonal-plane fluorescence optical sectioning. From Buytaert, J. A. N., Descamps, E., Adriaens, D., and Dirckx, J. J. J. (2012). Review Article. The OPFOS Microscopy Family: High-Resolution Optical Sectioning of Biomedical Specimens. Anatomy Research International, **2012**, Article ID 206238, 9 pages, http://dx.doi.org/10.1155/2012/206238

the confocal parameter. The authors invented a new type of LSFM that scanned the specimen across the beam waist and then combined the single-image columns to form a reconstructed in-focus image of the entire specimen. The invention of the HR-OPFOS, a diffraction-limited LSFM, decouples the relation between section thickness (axial resolution) and image width. Buytaert and Dirckx measured a FWHM axial resolution of 2.6 μm and a FWHM transverse resolution of 2.3 μm (Buytaert and Dirckx, 2007).

There are certain practical constraints to the use of the HR-OPFOS microscope. First, the specimen must fluoresce. Second, the light-sheet must traverse the specimen without excessive light scattering or refraction. To comply with the second requirement the specimen must be cleared, dehydrated, decalcified, have its refractive index match that of the fluid surrounding the specimen, and stained with a fluorescent probe. Many specimens in their raw state scatter a significant fraction of the incident light-sheet and as a consequence the latter does not penetrate the full thickness of the specimen. The solution to this common problem dates from 1911 when Spalteholtz published his technique for clearing specimens (Spalteholtz, 1911). The Spalteholtz clearing technique converts an opaque specimen into a transparent specimen by matching the refractive index of the entire specimen with a mixture of various oils that have refractive indices similar to that of proteins. If the specimen contains bone or calcified tissue, then the calcium (a strong scatterer of

light) must be removed (decalcified) before the index-matching technique is applied: the clearing technique cannot make bone transparent.

Buytaert et al. present a comprehensive assessment of the limitations of LSFM (Buytaert et al., 2012). They note that the complex steps required for the preparation of a specimen present a distinct disadvantage, and the specimen may suffer from significant shrinkage. The authors point out that image quality is strongly dependent on the degree of transparency of the specimen following its preparation. Inhomogeneities in refractive index, absorption, and residual scattering within the volume of the specimen result in out-of-focus illumination, decreased intensity of the light-sheet deep within the specimen, and stripes (dark and bright regions) and shadows (dark regions) in the image.

To mitigate the problem caused by varying distances between the light-sheet and the detection microscope objective at varying depths of the specimen LSFM designers have come up with different solutions. One solution is to displace the position of the specimen chamber in a direction orthogonal to the light-sheet (Buytaert and Dirckx, 2007; Huisken et al., 2004). Voie et al. came up with an alternative solution to this problem: rotating the specimen within the specimen chamber (Voie, 2002; Voie et al., 1993).

Complex specimen preparation is a precondition to the use of OPFOS and its variants when nonliving large specimens are involved. What is needed is an LSFM capable of imaging living specimens. The solution was provided by the Stelzer group working at the EMBL laboratory in Heidelberg, Germany (Stelzer et al., 1995). Citing the OPFOS technique the Stelzer group stated that SPIM followed from their previous work on oblique theta confocal microscopy (Lindek and Stelzer, 1994). Their technique is simple and elegant. They used specimens that were naturally transparent: live animal embryos such as medaka (*Oryzias latipes*) and fruit fly (*Drosophila melanogaster*) embryos that they embedded in agarose. With their SPIM LSFM and naturally transparent specimens they obtained multiple image stacks in various planes by rotating the specimens and then by means of three-dimensional computer reconstruction formed the final image.

The problem of stripes in unidirectional and multidirectional LSFM is caused by structures in the illumination light path that absorb or scatter light to a high degree. These bright and dark stripes significantly deteriorate image quality. Another problem is how to obtain micrometer resolution in macroscopic specimens. The solution to these problems was presented in a publication from Dodt et al., "Ultramicroscopy: Three-Dimensional Visualization of Neuronal Networks in the Whole Mouse Brain" (Dodt et al., 2007). The authors, who named their LSFM technique in honor of Zsigmondy, pointed out the problem with previous LSFM designs: it was not feasible to image the entire neuron network in an intact brain. Their solution was the ultramicroscope, a new LSFM used with fixed and cleared transparent specimens. The ultramicroscope incorporated two light-sheets that simultaneously illuminate the specimen from opposite sides of the specimen. This optical arrangement reduced the occurrence of stripes in the image of the specimen. The specimen was a fixed whole mouse brain that was cleared as part of specimen preparation. Resolution was at the scale of single cells. Green Fluorescent Protein

(GFP)-expressed neurons emitted fluorescence to form the high-resolution, high-contrast image of each optical section. The resulting images clearly showed single neurons, and three-dimensional reconstruction demonstrated dendritic trees and the spines of neurons (Dodt et al., 2007). The fruit fly and mouse embryo specimens allowed the authors to image cellular details based on intrinsic autofluorescence. Another important application of their ultramicroscope was the rapid phenotyping of mouse mutants (Figs. 11.4, 11.5 and 11.6).

In 2008 new LSFM variants were developed. Holekamp et al. found a way to image rapid physiological phenomena such as action potentials in neurons over wide fields of view. To do this they invented a new LSFM that had a thin light-sheet whose location was coupled to the focal plane of the detection microscope objective (Holekamp et al., 2008). The orientation of both illumination axes and the optic axis of the microscope was 45° relative to the specimen, which was in the horizontal orientation. The axial resolution of the microscope was 5 μm. Their LSFM development was called objective-coupled planar illumination (OCPI), and had applications in live brain imaging. How does OCPI compare with similar variants of LSFM (Dodt et al., 2007; Fuchs et al., 2002; Huisken et al., 2004)? The advantage of the OCPI microscope is that is can quickly scan through the thickness of the specimen while the specimen remains stationary.

Another new development was thin-sheet laser imaging microscopy (TSLIM) by Santi et al. It had many improvements over previous LSFMs: bidirectional light-sheet illumination from ultramicroscopy, image stitching from HR-OPFOS, and a combination of cylindrical lenses with aberration-corrected objectives from mSPIM. All these features were used in their TSLIM microscope (Santi et al., 2009).

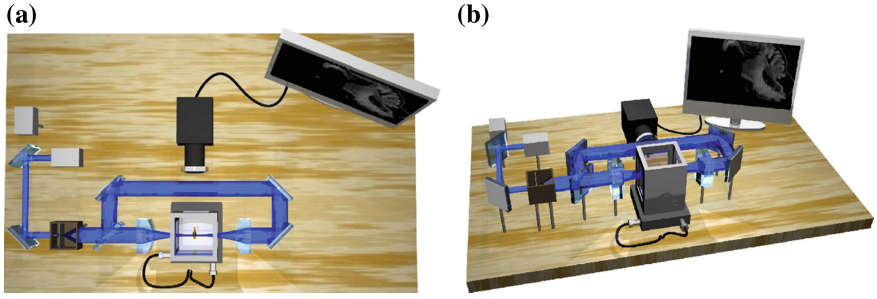


Fig. 11.4 **a** Three-dimensional representation of an HR-OPFOS setup with two-sided cylindrical lens sheet illumination and two laser wavelengths (*green* and *blue*). The *blue laser* is active here. From Buytaert, J. A. N., Descamps, E., Adriaens, D., and Dirckx, J. J. J. (2012). Review Article. "The OPFOS Microscopy Family: High-Resolution Optical Sectioning of Biomedical Specimens." Anatomy Research International, **2012**, Article ID 206238, 9 pages, http://dx.doi.org/10.1155/2012/206238. **b** Three-dimensional representation of an HR-OPFOS setup with two-sided cylindrical lens sheet illumination and two laser wavelengths (*green* and *blue*). The *blue laser* is active here. From Buytaert, J. A. N., Descamps, E., Adriaens, D., and Dirckx, J. J. J. (2012). Review Article. "The OPFOS Microscopy Family: High-Resolution Optical Sectioning of Biomedical Specimens." Anatomy Research International, **2012**, Article ID 206238, 9 pages, http://dx.doi.org/10.1155/2012/206238

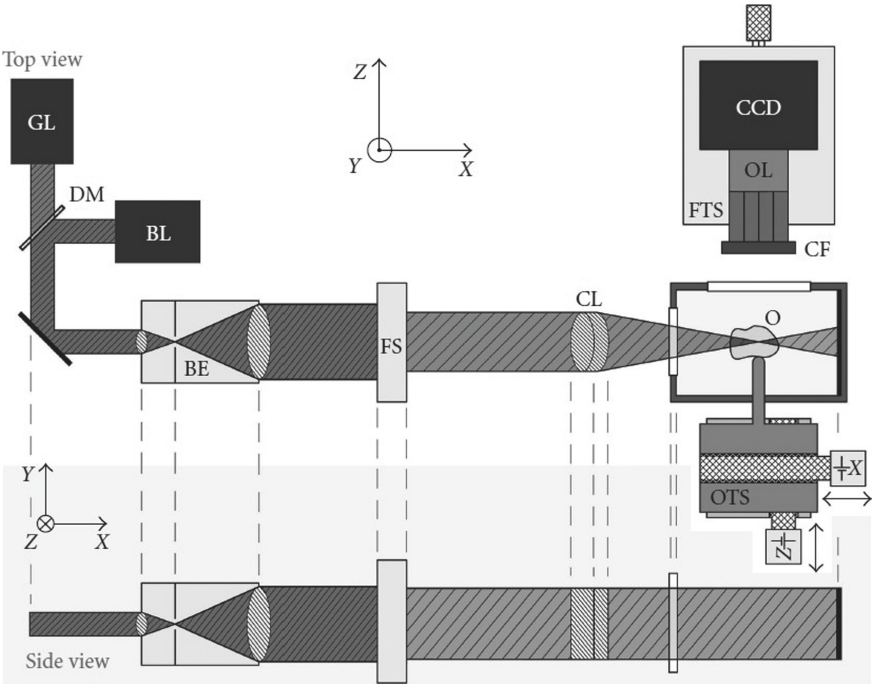


Fig. 11.5 Schematic of an HR-OPFOS setup. Light from a *green laser* (GL) or *blue laser* (BL) passes through a Keplerian beam expander (BE) with a spatial filter, a field stop (FS), and a cylindrical achromat lens (CL) that focuses the laser along one dimension within the transparent and fluorescent object (O). A two-axis motorized object translation stage (OTS) allows the specimen to be scanned and imaged at different depths. The fluorescence light emitted by the object is projected onto a charge-coupled device (CCD) camera by a microscope objective lens (OL) with a fluorescence color filter (CF) in front. The focusing translation stage (FTS) is used to make the objective lens focal plane coincide with the laser focus. From Buytaert, J. A. N., Descamps, E., Adriaens, D., and Dirckx, J. J. J. (2012). Review Article. "The OPFOS Microscopy Family: High-Resolution Optical Sectioning of Biomedical Specimens." Anatomy Research International, **2012**, Article ID 206238, 9 pages, http://dx.doi.org/10.1155/2012/206238

An excellent review of optical-sectioning light microscopy by Mertz (2011) provides another perspective. In 2008 Lim, Chu, and Kim developed HiLo microscopy, a very fast imaging method that is optimal for in vivo microscopy and yields wide fields of view (Lim et al., 2008; Lim et al., 2011; Mertz and Kim, 2010;). HiLo microscopy only requires two images: one image based on structured illumination microscopy (SIM) (for details see Chapter 13) and one widefield image. The great advantage is lack of complexity in the microscope and fast imaging speeds (Lim et al., 2008).

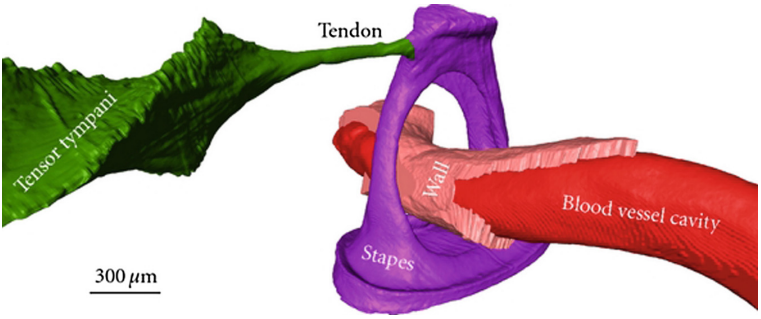


Fig. 11.6 Three-dimensional OPFOS reconstruction of a gerbil showing a surface mesh of the stapes, a blood vessel running through it, and the tensor tympani muscle attached to the head of the stapes. The blood vessel wall and inner cavity are both separately modeled (voxel size 1.5 × 1.5 × 5 μm). From Buytaert, J. A. N., Descamps, E., Adriaens, D., and Dirckx, J. J. J. (2012). Review Article. The OPFOS Microscopy Family: High-Resolution Optical Sectioning of Biomedical Specimens. Anatomy Research International, **2012**, Article ID 206238, 9 pages, Anatomy Research International (this is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited)

11.3 Optical Projection Tomography

Sharpe et al. found a way to obtain high-resolution three-dimensional images of large specimens. They understood that techniques used for three-dimensional microscopy, such as confocal microscopy, deconvolution, and optical coherence tomography (OCT), place constraints on the maximum thickness of a specimen. The optical sectioning techniques described thus far in this chapter penetrate no deeper than 1–2 mm, which is insufficient for imaging large specimens. Their solution was development of the optical analogue of X-ray computed tomography, a technique they described as optical projection tomography (OPT) (Sharpe et al., 2002; Sharpe, 2004, 2008).

OPT was developed to exploit the bright-field contrast of a sample by acquiring several light transmission images (or projections) from different directions; the three-dimensional structure of the sample is then reconstructed using a back projection algorithm based on the Radon transform (Kak and Slaney, 1988). OPT microscopy works by acquiring several light transmission images (or projections) from different directions within the specimen. The microscope acquires two-dimensional projections of the specimen. The three-dimensional structure of the specimen is reconstructed from these projections. OPT operates in two modes: fluorescence OPT and bright-field (transmission) OPT. Haisch (2012) is a comprehensive review article on optical tomography that provides detailed explanations and comparisons of many of the modern methods used for three-dimensional reconstruction.

The main advantage of OPT is its ability to achieve diffraction-limited resolution in large specimens. OPT can form diffraction-limited images of either fluorescent specimens or nonfluorescent biological specimens that have a maximum thickness of about 15 mm. Furthermore, OPT can be used to study developmental biology and gene functions by mapping the distribution of RNA in tissue and protein expression in embryos (Sharpe et al., 2002). Sharpe et al. also demonstrated time-lapse OPT and its application to the study of organogenesis in mammals (Boot et al., 2008).

The main limitation of OPT is its dependence on nonscattered light. If the specimen is not transparent, then it must be cleared. This limitation applies of course to all types of LSFM. It is advisable to use infrared light as the illumination source for absorbing specimens that are stained with colored dyes or colored precipitates because longer wavelength light is less scattered.

How does the OPT microscope operate? As described by Sharpe et al. the specimen is placed in a transparent cylinder that contains agarose gel. The cylinder is rotated in increments of 0.9° and an image is acquired at each of the 400 rotated orientations. Back projection was used to get a three-dimensional image of the specimen.

Mayer et al.'s paper "OPTiSPIM: Integrating Optical Projection Tomography in light-sheet Microscopy Extends Specimen Characterization to Nonfluorescent Contrasts" illustrates how innovative instrumentation development can incorporate the advantages of OPT and light-sheet tomography in a new type of hybrid microscope with unique capabilities (Mayer et al., 2014). The operating principle behind the OPTiSPIM hybrid involves a laser-formed light-sheet that is incident on the specimen in a direction orthogonal to the detection axis, much as is the case with LSFM. The specimen is sequentially displaced across the light-sheet to yield a three-dimensional distribution of fluorescence from the entire specimen. If the specimen is larger than the field of view of the LSFM, then the three-dimensional tiles are recorded and stitched together in a computer. The resolution and field of view are thus decoupled and diffraction-limited resolution is obtained in large specimens.

11.4 Instrumentation: Construction, Advantages, Limitations, and Applications

In the last decade LSFM has become a popular tool for cell biologists, neuroscientists, and developmental biologists. This section will discuss the basic technique and its variants, as well as the advantages and limitations of each instrument when applied to a specific specimen. The following questions are addressed for each of the publications discussed: What problem or increased capabilities were related to instrument design? What was new about the instrument? How did the instrument solve or mitigate problems with previous designs? What specimens and studies

were best suited for examination with the new instrument? What was learned about the specimen using the new instrument? These questions are important and deserve consideration by developers of new types of LSFM.

The next variant of LSFM solved an operational constraint that originated, perhaps surprisingly, with the specimen holder. Cell biologists often use specimens that are prepared on glass microscope slides and coverslips. The use of these specimen holders is not compatible with existing LSFMs. A solution was Dunsby's invention of a new variant of the LSFM called oblique plane microscopy (OPM) (Dunsby, 2008). An OPM uses one high-NA microscope objective to illuminate an oblique plane in the specimen and the same objective for the detection of fluorescence from the specimen. One-half of the objective angular aperture is used for the illumination path and the other half is used for the detection path. OPM has several advantages over SPIM: the technique can be used with both nonfluorescent and fluorescent specimens whose imaging is based on reflected and scattered light. Most importantly, an OPM can be used for specimens on glass slides and coverslips and is also compatible with biochips or cells in multiwell plates. The technique maintains the low photobleaching and low phototoxicity of other LSFM systems (Dunsby, 2008). An OPM has a number of limitations such as the full aperture of the microscope objective is not used because only one-half of the objective angular aperture is used for the detection pathway, causing a reduction in optimal resolution. Another limitation is that the OPM system is not suited to the imaging of highly scattering specimens, as is the case with all types of LSFMs.

The increasing needs and constraints of developmental biologists, cell biologists, and neurobiologists often drive the innovation of new types of LSFMs. Difficulties arise when it is necessary to image specimens that are large, multicellular, and nontransparent, which is associated with high absorption and light scattering. This can decrease image contrast. These problems have found partial solutions in trade-offs. It was found, for example, that early-stage developing embryos are more transparent and therefore better suited to LSFM than are late-stage developing embryos. However, a more general solution led to the invention and development of a new form of LSFM.

The solution was provided by Keller et al. who combined two existing microscope techniques: digital scanned laser microscopy (DSLM) invented by Keller and Stelzer (2008) and structured illumination microscopy (SIM) invented by Neil et al., (1997). The result was a new variant of LSFM that they called DSLM-SI (Keller et al., 2010). SIM is the subject of Chapter 13. The use of SIM mitigates against background light scatter and provides optical sectioning by discriminating against light that is above and below the focal plane.

Building on the prior publication of Neil et al., (1997), Keller et al. replaced the standard light-sheet with illumination that consists of a pattern of periodic stripes. Keller et al. found that using mechanical ruling to form the stripes led to problems: different mechanical rulings were required for different frequencies resulting in low image acquisition speeds. The innovation involved varying the intensity of laser illumination in step with scanning resulting in periodic intensity patterns on the specimen. These patterns could be easily and rapidly adjusted to match the

changing optical properties of the specimen over time such as during the developmental process of live embryos (Keller et al., 2010).

Validation of a new instrument involves the capability of the instrument to acquire images that have advantages over those that can be obtained with previous instruments. The DSLM-SI instrument yielded outstanding performance in imaging a live zebrafish for 58 hours of its development. The team also imaged live *D. melanogaster* during development and were able to track the movement of individual cells over time.

Krzic et al. provided the next innovation, which was motivated by the needs of developmental biologists for a microscope that could image entire living embryos during their early development over a period of several hours (Krzic et al., 2012). The microscope should allow individual cell or nuclei tracking for all the cells in the embryo as well as provide images of subcellular components and organelles. The resolution requirements to accomplish these imaging tasks were severe: temporal resolution of a few seconds over a period of several hours and spatial resolution over a range extending from submicrons to millimeters.

Large opaque specimens also presented problems. The absorption and scattering of light in a specimen significantly degrade image quality. To mitigate such undesirable effects it was standard procedure in SPIM to rotate the specimen, but specimen rotation introduced additional problems. One involved image misalignment. Another involved the combination of rotation and the requirement to image many orientations of the specimen, which resulted in increased photobleaching and phototoxicity. Furthermore, decreased image acquisition speed negated the capacity to image dynamic processes occurring in the embryo during early development.

Krzic et al.'s solution was to develop a new variant of SPIM that obviated the need to rotate the specimen and provided the required spatial and temporal resolution to satisfy the requirements of developmental biologists previously described (Krzic et al., 2012). Recall that SPIM had its origins in the invention of the ultramicroscope used to investigate submicroscopic colloids and had undergone many development stages in the form of LSFM (Dodt et al., 2007; Huisken and Stainier, 2007, 2009; Huisken et al., 2004; Siedentopf and Zsigmondy, 1903; Voie et al., 1993). The SPIM method used a light-sheet for illumination of one section of the specimen and fluorescence was collected by a detector microscope objective oriented on an axis that was perpendicular to the light-sheet.

Krzic et al. invented a new form of LSFM that they called a multiview selective plane illumination microscope (MuVi-SPIM). It used two illumination microscope objectives and two detection microscope objectives. It achieved both good temporal resolution at high speed and subcellular spatial resolution over an entire large specimen (Krzic et al., 2012). Both detection microscope objectives were focused on the same focal plane. The illumination and the detection orientations were fixed in the microscope with the advantageous results that multiview fusion of the images occurred in real time.

What were the advantages of the MuVi-SPIM? First, there was no need for specimen rotation. The specimen was displaced through the plane of the light-sheet by a piezoelectric stage translation device. Each illumination microscope objective

formed a single light-sheet, and each illumination microscope objective was used sequentially to illuminate the specimen. The two opposing illumination objectives and the two opposing detector objectives focused on the specimen from four directions. The illumination direction and the detection direction were perpendicular in their orientation. These four microscope objectives yielded four three-dimensional images. The two detector microscope objectives that operated simultaneously resulted in the MuVi-SPIM having twice the light efficiency of previous LSFM designs. The MuVi-SPIM acquired a three-dimensional stack of images from one light-sheet, and then the instrument switched to the other light-sheet and acquired a second three-dimensional image stack. The MuVi-SPIM acquired, with no need for image rotation, four three-dimensional images that were in registration and combined them into a single image of the live specimen.

The next new design of an LSFM came from Tomer et al. working at the Howard Hughes Medical Institute in Ashburn, Virginia. They developed quantitative high-speed imaging of entire developing embryos using simultaneous multiview light-sheet microscopy (Tomer et al., 2012).

Tomer et al. noted that the depth of penetration into large living specimens was severely limited with previous light microscope designs. Additionally, there was the need for temporal resolution suitable for imaging rapid physiological and developmental changes. Previous designs of LSFMs based on sequential multiview imaging did not meet the requirements of rapid temporal resolution. Tomer et al. addressed the penetration problem by inventing a new instrument that they called SiMView (Tomer et al., 2012). The SiMView design was suitable for either single-photon or two-photon simultaneous multiview image acquisition. The authors claimed that SiMView acquired four optical views of the specimen with a 20 ms time shift between each view, and that high temporal resolution was not dependent on the size of specimen. When the SiMView instrument was compared with previous designs of LSFMs that used sequential acquisition of four views of the specimen, the advantage was quite remarkable. There was a 20-fold increase in temporal resolution. The SiMView instrument was designed to acquire images of live specimens over a period of several days.

The basic components of the SiMView instrument were laser illumination sources, two light-sheet arms for the illumination system, two detector arms that used scientific complementary metal oxide semiconductor (sCMOS) cameras as the imaging detectors, a special four-view chamber to contain the live specimen under physiological condition, and a four-axis positioning system to position the specimen in the chamber. For operation in the two-photon mode a pulsed Ti:sapphire laser was used.

Tomer et al. validated the imaging capabilities of the SiMView instrument by tracking the process of embryogenesis in specimens of entire *D. melanogaster* embryos with 30 second temporal resolution. They also studied the long-term imaging of the development of neural systems by imaging neuroblast cell lineages in vivo (Tomer et al., 2012).

Wu et al.'s paper "Inverted Selective Plane Illumination Microscopy (*i*SPIM) Enables Coupled Cell Identity Lineaging and Neurodevelopmental Imaging in

Caenorhabditis elegans" presents a new LSFM design (Wu et al., 2011). The term lineaging in developmental biology refers to the tracking of cells in a developing embryo as the cells divide, differentiate, and migrate to various regions in the embryo and form specific tissues and organs. The authors wanted to investigate the neurodevelopment of *Caenorhabditis elegans* by imaging the entire embryo every 2 seconds over a 14 hour period. It was desirable, if not an absolute requirement, that there should be no signs of phototoxicity induced by the incident light-sheet during this imaging period. *C. elegans* is a useful model system for neurodevelopment because its neural connectivity data are mapped and in the public domain, and its nervous system consists of 302 neurons, making it one of the simpler nervous systems to study. Wu et al. presented further insights into advanced optical imaging techniques for neurodevelopment and pointed out that previous designs of three-dimensional microscopes failed in their temporal resolution and should be ruled out due to their light-induced phototoxicity (Wu et al., 2013). In addition, they found it desirable to prepare specimens on a coverslip for imaging in the microscope. They addressed these problems by inventing the inverted selective-plane illumination microscopy (*i*SPIM), which they described as a noninvasive (nonperceptible phototoxicity) high-speed three-dimensional imaging microscope suitable for studying living specimens over long periods (Wu et al., 2011).

The performance of the *i*SPIM was extraordinary. In their validation of the capabilities of the microscope they acquired 25,000 volumetric images of *C. elegans* embryogenesis resulting in the complete three-dimensional visualization of the identity of each nuclei and its location. Wu et al. accomplished this by combining two-color *i*SPIM with computerized lineaging methods.

How does *i*SPIM work (Fig. 11.7)? Previous implementations of LSFM systems required special sample preparation (e.g., the specimen was surrounded by an agarose gel, which prevented mounting the specimen on conventional coverslips). Wu et al. designed *i*SPIM to have the geometry of an inverted microscope and the specimen to be used with conventional coverslips (Wu et al., 2011) (Fig. 11.7).

As shown in Fig. 11.7 the inverted microscope setup consists of an objective mount that contains the excitation microscope objective and the detection microscope objective. Both share a common focus in the specimen, which is on a coverslip. To rapidly acquire images from the specimen the objective mount is attached to a *Z*-axis translation stage that can synchronously translate both microscope objectives and their common focus through the specimen. As is the case with all LSFMs the *i*SPIM design has the illumination and detection axes perpendicular in their orientation. Both water immersion microscope objectives have long working distances (3.5 mm) and high NA (0.8). Wu et al. tuned the operating characteristics of their *i*SPIM to investigate *C. elegans* embryogenesis: the beam waist of their light-sheet was 1.2 μm at the center of the embryo. The lateral spatial resolution was 0.52 μm and the axial resolution was 1.7 μm.

Wu et al. discuss the limitations of using Gaussian beam light-sheets (diffraction spreading of the light-sheet on either side of the beam waist) and point out the advantages of using Bessel beams to form light-sheets (Fahrbach et al., 2010; Planchon et al., 2011). Of course, Bessel beam light-sheets contain side lobes that

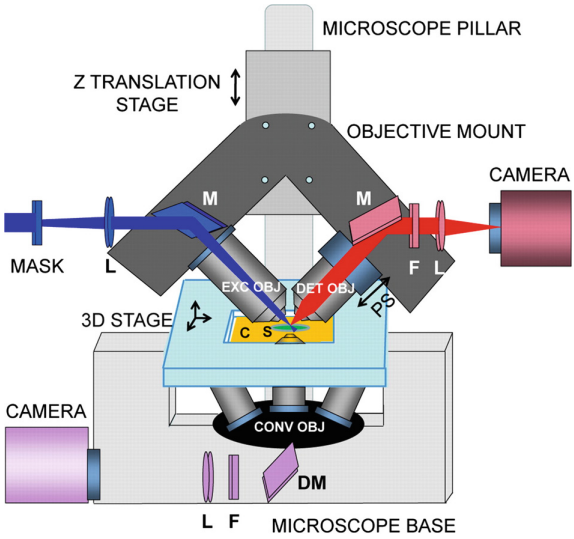


Fig. 11.7 *i*SPIM plane illumination on an inverted microscope base. Two water-immersion objectives are mounted onto a *Z*-axis translation stage that is bolted directly onto the illumination pillar of the inverted microscope. The demagnified image of the rectangular slit (MASK) is reimaged with lenses (L) and an excitation *i*SPIM objective (EXC OBJ), thus producing a light-sheet at the sample (S). For clarity, relay lens pairs between L and EXC OBJ are omitted in this schematic. Sample fluorescence is detected (DET OBJ) using appropriate mirrors (M), emission filters (F), lenses (L), and a camera. EXC OBJ is fixed in place and the light-sheet is scanned through the sample using a galvanometric mirror (not shown). A piezoelectric objective stage (PS) moves DET OBJ in sync with the light-sheet, ensuring that detection and excitation planes are coincident. The sample S is mounted on a coverslip (C) that is placed on a three-dimensional translation stage, thus ensuring correct placement of S relative to *i*SPIM objectives. S may be viewed through objectives (CONV OBJ), dichroic mirrors (DM), and optics in the conventional light path of the inverted microscope. From Wu, Y., Ghitani, A., Christensen, R., Santella, A., Du, Z., Rondeau, G., Bao, Z., Colón-Ramos, D., and Shroff, H. (2011). Inverted selective plane illumination microscopy (*i*SPIM) enables coupled cell identity lineaging and neurodevelopmental imaging in *Caenorhabditis elegans*. Proceedings of the National Academy of Sciences of the United States of America, **108**, 17708–17713

increase the background. Wu et al. have considered alternative techniques to improve the axial resolution of the *i*SPIM and suggested some promising approaches based on previous work of other groups (Dunsby, 2008; Swoger et al., 2007).

Our discussions of innovative developments of various types of LSFMs have elucidated the importance of the type of light beam(s) that forms light-sheet(s); selection and preparation of the specimen should take into consideration the transparency, refractive index inhomogeneities, absorbance, the required penetration depth, the specimen's physical dimensions, the requirements and the constraints of the specimen chamber, the use of coverslips to hold the specimen, the requirements of a prior clearing procedure, and the requirements of specific spatial and temporal resolutions. The following discussion highlights the critical role of

high-speed image acquisition as well as of real-time data processing, visualization, and analysis. A suitable basis for this discussion is provided by a publication of Schmid et al. at the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, which describes a four-lens SPIM setup for "high-speed panoramic light-sheet microscopy that reveals global endodermal cell dynamics" (Schmid et al., 2013). In the publication's acknowledgment section the authors point out that research began in the laboratory of Professor Didier Stainier at the University of California, San Francisco.

I begin with Schmid et al. (2013)'s statement of the problems with previous versions of LSFMs and then move on to a discussion of their solution to this by inventing a new type of LSFM. The authors divide the problems into two parts: data acquisition and processing, and the biological problem. First, the use of sCMOS cameras for imaging detectors imply data rates that exceed 1 GB per second presenting massive data transfer, data storage, and data analysis difficulties. These factors present a significant bottleneck for optimal use of the microscope. What is required is some technique to perform these operations and data analysis on the fly in real time (Schmid et al., 2013).

The second problem to be addressed by a new type of LSFM is specific to the specimen that researchers wish to study. In the study of early development of the zebrafish the endoderm is the germ layer of interest. Schmid et al. wanted to study the putative coordination of cells that comprise this layer, which surrounds the surface of the yolk. The importance of such knowledge is related to the fact that problems of cell movement in the gastrulation stage of development can cause malformation of organs in the zebrafish. Their solution to these different but related problems with previous LSFM designs was a new design of LSFM: a four-objective SPIM instrument with an integral real-time image processor that achieves very high-speed radial maximum intensity projection of the specimen. A critical advantage of the design is that it preserves the topology of tissue. Their LSFM can acquire high-resolution images of the entire endoderm of gastrula-stage zebrafish embryos during the image acquisition period (Schmid et al., 2013).

A second feature of Schmid et al.'s new LSFM involves the application of techniques from cartography for projective geometry mapping of three-dimensional surfaces or volumes into two-dimensional projections. For example, there are several projections that can be used to map regions of three-dimensional earth into two-dimensional maps that preserve a specific set of properties of the original three-dimensional volume. The authors' innovative data visualization techniques display the entire tissue of the spherical or ellipsoidal specimen as a single two-dimensional image (Schmid et al., 2013). The data acquisition systems employ real-time data registration, which compensates for the movement of embryos. Individual endoderm cell migration during the course of gastrula development is tracked and finally displayed on two-dimensional maps that represent the entire endoderm of the zebrafish (Schmid et al., 2013).

Schmid et al. (2013)'s new type of SPIM uses four microscope objectives and is designed and optimized for three-dimensional imaging of the entire ectoderm of the zebrafish. The detection system consists of two sCMOS cameras that image the

common focal plane of the specimen. The transparent specimen chamber is in the axial direction of the LSFM and contains several zebrafish embryos in a 1.5% agarose gel, which provides the advantage of multisample specimen mounting in the microscope. The specimen chamber is displaced along the axial direction; images of the specimen are collected in steps of 2 μm. The two illumination microscope objectives alternately form their light-sheet in the focal plane for each position of the axial scan movement. The axes of the illumination systems and their light-sheets are perpendicular to the axes of the detection cameras. There is no requirement for image registration since the two detection cameras are aligned to form identical images. The authors found that they could enhance image quality by rotating the specimen chamber by 45°.

Cartographic image projection techniques could be applied to the specimen since the zebrafish embryo in its early stages of development is spherical to a good approximation. Schmid et al. (2013) point out that their instrument and image visualization techniques can be adapted to other specimens such as fly embryos and those of worms.

The next publication I consider (Bassi et al., 2015) presents a new LSFM that is a multimode, hybrid optical design based on two known techniques: optical tomography based on bright-field contrast, and SPIM based on fluorescence to generate contrast. Bassi et al.'s innovative approach is titled "Optical Tomography Complements light-sheet Microscopy for In Toto Imaging of Zebrafish Development".

I begin with a statement of the problem as perceived by Bassi et al. They state what seems to be obvious: all types of fluorescence microscopy can only image cells or tissues that are fluorescent. Typically, genetic modification is used to overexpress a type of fluorescent protein; alternatively, fluorescent probes can be used to stain cellular components (e.g., nuclei or the cytoskeleton). These techniques can be highly specific to a single type of protein or to a single cellular component. Nonfluorescent regions of the specimen do not contribute to image contrast. Autofluorescence may be considered an alternative, but its very low fluorescence necessitates high-intensity illumination that can lead to significant phototoxicity. Bassi et al. searched for a technique that was capable of imaging nonfluorescent regions of the specimen and could complement the fluorescent images that can be obtained with SPIM techniques.

Bassi et al.'s solution to their problem was to combine two well-known techniques in a single instrument. As noted earlier, Sharpe et al. in 2002 invented optical projection tomography (OPT) for three-dimensional visualization of nonfluorescent specimens. In OPT microscopes bright-field contrast (based on light transmission) is used to acquire projections from many directions and three-dimensional reconstruction is performed from multiple projections following the back projection algorithm (Kak and Slaney, 1988). Bassi et al. combined OPT and SPIM in a standard SPIM instrument. There are three basic components to this multimodal new instrument (Bassi et al., 2015). A high frame rate sCMOS camera is used to detect light from the specimen. Four water immersion microscope objectives provide illumination light-sheets and detection. The two light-sheets

illuminate the specimen from two opposing sides. Transmitted light comes from an LED located at the back of one of the microscope objectives. The fourth microscope objective is used to collect fluorescence and transmitted light. The specimen can be rapidly rotated to yield multiviews of it from various directions. Light from an LED source illuminates the specimen from various directions to produce a set of transmission images. The specimen is translated across the focal plane of the detection microscope objective, and 20 transmission images are acquired. High-pass digital filtering is used to obtain an in-focus image. Multiple projections are obtained from 360 directions to form a data set that is used for three-dimensional reconstruction of the specimen. Optically sectioned volumes of the specimen are obtained by a filtered back projection algorithm (Fauver et al., 2005; Kikuchi and Sonobe, 1994).

Bassi et al. validated the performance of their hybrid LSFM by imaging live zebrafish embryos for periods of several hours. The movies included in the supplementary material to their publication clarify what can be achieved with this new multimodal LSFM and its utility in the study of developmental biology (Bassi et al., 2015).

There is high pedagogical value in constructing your own LSFM as an individual or as a member of a group. It is also enjoyable and can provide a great sense of accomplishment. With access to journals, such as *Review of Scientific Instruments* and *Nature Methods*, and websites that provide detailed listings of parts, suppliers, and manufacturers, as well as details on assembly, calibration, operation software, and data processing, not to mention dangers and pitfalls, the construction projects become realistic for students and researchers.

The website OpenSPIM from *Nature Methods* provides instructions, lists of parts, and blueprints (Pitrone et al., 2013): http://openspim.org/Welcome_to_the_OpenSPIM_Wiki (accessed April 21, 2019).

Another website is the open-source platform called OpenSpinMicroscopy that is also published in *Nature Methods* (Gualda et al., 2013). The link to OpenSpinMicroscopy is: https://sites.google.com/site/openspinmicroscopy/ (accessed April 3, 2019).

The OpenSpinMicroscopy website provides the user with hardware recommendations and various software plugins. Their site also lists parts, manufacturers, and assembly instructions for the user to construct various types of LSFMs: selective plane illumination microscope (SPIM), a digitally scanned light-sheet microscope, and an optical projection tomography (OPT) microscope. The software also provides a Micro-manager Java plugin for the user to have control over the full operation of these microscopes. Another useful feature is details of the design for a specimen chamber that can be constructed by three-dimensional printing.

A third recommendation is a technical paper that appeared in the journal *Review of Scientific Instruments* titled "Basic Building Units and Properties of a Fluorescence Single Plane Illumination Microscope," which contains a detailed description of the SPIM setup (Greger et al., 2007).

Another good source of information is the *Nature Protocols*' publication "Dual-View Plane Illumination Microscopy for Rapid and Spatially Isotropic Imaging," which contains detailed instructions for the construction and operation of

this LSFM (Kumar et al., 2014). This LSFM is designed for use with specimens on glass coverslips and for imaging embryos of the transparent nematode *C. elegans*. The publication provides lists of parts, manufacturers, and instructions for the assembly, optical alignment, specimen preparation, instrument control software, and tips for optimal operation of the LSFM.

Additionally, there is excellent information related to LSFMs and specimen preparation on the websites of commercial manufacturers. Below are several commercial websites that contain LSFM product information and application notes on specimen preparation and mounting. Material provided by these websites contain many practical suggestions:

- http://www.leica-microsystems.com/products/confocal-microscopes/details/product/leica-tcs-sp8-dls/ (accessed April 21, 2019)
- http://www.zeiss.com/microscopy/en_us/products/imaging-systems/lightsheet-z-1.html (accessed April 21, 2019)
- https://www.med.upenn.edu/cdbmicroscopycore/assets/user-content/documents/Zeiss_White_Paper_LightsheetZ1_Sample-Preparation.pdf (accessed April 21, 2019).

The five SPIM-specific specimen-mounting techniques—embedded, clipped, enclosed, flat, and flow-through—are described in detail in the Zeiss Lightsheet Z.1 Sample Preparation White Paper.

The Applied Scientific Information (ASI) website provides details on two types of LSFMs: oblique selective plane illumination microscope (oSPIM) and selective plane illumination microscopes (iSPIM and diSPIM):

- http://www.asiimaging.com/index.php/products/light-sheet-microscopy/ (accessed April 21, 2019)
- http://www.andor.com/learning-academy/spim-selective-plane-illumination-microscopy (accessed April 21, 2019).

11.5 Innovation on Microscope Illumination: Bessel and Airy Beams

In the previous section various illumination and detector geometries for LSFMs were discussed. In this section LSFM techniques with illumination based on Bessel beams, Gaussian beams, and Airy beams are presented, along with an introduction to lattice light-sheet microscopy. The general problem is how to improve the resolution and penetration depth of LSFMs. Solutions involve the use of Bessel beams and Airy beams to form the light-sheet. The first set of publications describes the optics of various types of beams, and the second set discusses the use of these beams in innovative designs of LSFMs and their applications.

First, what is a Bessel beam? In 1987 Durnin concluded that the Helmholtz equation describing the propagation of light waves has solutions that represent diffraction-free (i.e., nonspreading) modes (Durnin, 1987; Durnin et al., 1987). The optical properties of these diffraction-free modes, referred to as Bessel beams, are both fascinating and useful. Bessel beams have narrow-beam radii and two critical characteristics: the central spot of intensity can in theory have a radius of a few wavelengths, and the beam propagates without diffractive widening. Using simple optical elements Durnin and colleagues were able to produce a quasi-Bessel beam experimentally that had the characteristics of a Bessel beam over a finite distance. Today such quasi-Bessel beams are generally referred to simply as Bessel beams.

Durnin noted that a monochromatic wave that propagates in the z-direction and has a transverse amplitude described mathematically by a zero-order Bessel function of the first kind (Durnin, 1987) would be a solution to the Helmholtz equation. Hence the name Bessel beam. The important characteristic, a nondiffracting beam, derives from the property that a Bessel beam has the same intensity distribution in all planes perpendicular to the z-axis; in other words, the transverse intensity distribution remains constant and is independent of the propagation distance (Durnin et al., 1987). The publication "Bessel beams: Diffraction in a new light," provides a good introduction to the subject (McGloin and Dholakia, 2005).

Durnin, Miceli, and Eberly concluded from computer simulations that diffraction-free Bessel beams and Gaussian beams have similar efficiencies in the transport of power, and that a Bessel beam can have a larger depth of field than a Gaussian beam of the same spot size but with a loss of power (Durnin et al., 1987, 1988). They note that the power of Bessel beams is equally distributed between the rings of the beam, and thus as the number of rings in the beam increases the power in the central core diminishes. However, an increase in the number of rings in the Bessel beam yields a corresponding increase in the propagation distance over which the beam is effectively nonspreading (Durnin et al., 1987).

How are Bessel beams generated? Durnin et al. formed a Bessel beam by placing an annular aperture in the back focal plane of a lens and illuminating it with a plane wave (Durnin et al., 1987). An alternative method is to use an axicon, which behaves like a conical lens. The axicon converts a Gaussian beam into a Bessel beam with much greater efficiency than does an annular aperture since it uses a greater portion of the Gaussian beam (Brzobohatý et al., 2008). Bessel beams can also be formed from a Gaussian beam that is incident on a computer-generated hologram (Fahrbach et al., 2010).

11.5.1 Bessel Beams and Their Use in Light-Sheet Fluorescence Microscopy

What are the advantages, disadvantages, and limitations of Bessel beam illumination in LSFM? This question is addressed in a paper titled "Microscopy with

self-reconstructing beams" (Fahrbach et al., 2010). A self-reconstructing light beam can encounter an arbitrary obstacle in its path and, after propagating past the obstacle, effectively reform itself to its original beam profile. Zero-order Bessel beams have been shown experimentally to be self-restructuring (Bouchal et al., 1998). In their 2010 publication Fahrbach, Simon, and Rohrbach made the prescient suggestion that self-reconstructing beams may be useful in light-sheet microscopy of complex and highly scattering specimens. They demonstrated the use of Bessel beams in the formation of high-contrast images of highly scattering specimens such as human skin. They named their method microscopy with self-reconstructing beams (MISERB). A light-sheet was produced by illuminating a spatial light modulator that played the role of a computer-generated hologram and produced the Bessel beam profile. Their results show a 50% increase in penetration depth in human epidermis tissue over that achieved with conventional Gaussian beam illumination. Their publication attracted the attention of the light-sheet fluorescence microscopy community.

Planchon et al. published a seminal article "Rapid Three-Dimensional Isotropic (Equal Axial and Transverse Resolution) Imaging of Living Cells Using Bessel Beam Plane Illumination" (Planchon et al., 2011). The term plane illumination microscopy refers to a previous publication on mSPIM (Huisken and Stainier, 2009). The authors noted the fundamental trade-off between increases in the spatiotemporal resolution of the microscope that reduce the number of photons incident and emitted at each voxel in a given time period, and an associated reduction in the image signal-to-nose ratio. They warn that this aspect of Bessel beam imaging is often overlooked by researchers. When live cells are observed by microscopes based on epi-illumination (widefield, structured illumination, three-dimensional-photoactivated localization, confocal and stimulated emission depletion microscopy) the illumination is incident on the full thickness of the specimen. Although photobleaching is important, since it reduces the number of fluorescent molecules and therefore the number of photons emitted, a more critical problem is the cellular effects of phototoxicity on live specimens. LSFM minimizes these effects, but does not eliminate them.

Planchon et al. also detail a critical problem with LSFM when it comes to imaging large multicellular specimens at single-cell resolution. They then offer their solution. When Gaussian beams are used in LSFM there is trade-off between the minimum thickness of the light-sheet and the field of view over which it remains reasonably uniform. This results in less-than-optimal reduction in background rejection and photobleaching since there is significant excitation from out-of-focus planes. The thick light-sheets formed by Gaussian beams result in asymmetrical resolution in the axial and transverse directions with the axial resolution three to four times less than the transverse resolution (Planchon et al., 2011). Axial resolution and optical sectioning are related but different. Axial resolution is the smallest axial distance between two points in an object that are resolved (in a noise-free system), whereas optical sectioning is the efficiency of the system to reject out-of-focus light (Gao et al., 2014).

Planchon et al. use Bessel beams as a solution to problems with Gaussian beam light-sheets. In their Bessel beam LSFM they form Bessel beams by projecting an

annular aperture in the rear pupil of the excitation microscope objective. Bessel beams have the property, not present in Gaussian beams, that their central peak width can be decoupled from their longitudinal extent. This can be accomplished by altering the thickness of the annulus. The use of excitation and detection microscope objectives of equal NA resulted in isotropic resolution. They achieved high-speed imaging without displacing the specimen by scanning the light-sheet with a galvanometer-mounted mirror and synchronously moving the detection microscope objective (Planchon et al., 2011).

Although Bessel beam LSFM reduced the out-of-focus background compared with that produced with Gaussian light-sheets, there remained limitations to the use of Bessel beams since the sidelobes of Bessel beams contain a significant portion of the beam energy: the combined effect of beam sidelobes and the scanning technique used in Planchon et al.'s Bessel beam LSFM resulted in slowly diminishing sidelobes in the axial point-spread function (PSF). To mitigate these problems they combined the technique of structured illumination microscopy (SIM) with their Bessel beam LSFM.

Optical sectioning is achieved by projecting a periodic pattern onto the focal plane of a wide-field microscope and acquiring three images; for each image the pattern is translated in steps of one-third period (Neil et al., 1997). The image is computationally obtained and is an optical section of the object in the focal plane. The principle to obtain optical sectioning is as follows: the projected periodic pattern has the highest amplitude in the focal plane and it weakens in out-of-focus planes; the algorithm keeps the strongly modulated information at the focal plane and discards the less modulated information from planes that are above and below the focal plane.

The Bessel beam plane illumination microscope of Planchon et al. uses a galvanometer to sweep the circular Bessel beam in the x-direction across the plane of focus of the detection microscope objective, which forms a scanned light-sheet and acquires an image at a single z (axial plane) direction plane in the specimen.

But the problem of sidelobes in Bessel beams still remains. To mitigate this problem and improve the axial resolution of their microscope Planchon et al. adapted the SIM technique to fit their Bessel beam LSFM. The authors adopted the approach Neil et al. took with their Bessel beam plane illumination microscope. The Bessel beam was incident at discrete, periodic points across the total x-field of view. The authors then acquired three images whose periodic pattern was shifted for each image by one-third period, and the final image was computationally obtained (Planchon et al., 2011). They achieved diffraction-limited three-dimensional imaging of living cells. Their microscope has outstanding capabilities: three-dimensional isotropic resolution (equal in x, y, and z-directions) of 0.3 μm, image acquisition speeds of 200 image planes per second, and the capability to acquire hundreds of three-dimensional data volumes of single living cells from tens of thousands of image frames.

These preliminary investigations into the application of Bessel beams to LSFM have been further developed and extended for thickly labeled fluorescent specimens and for three-dimensional live cell imaging (Gao et al., 2012; Gao et al., 2014).

Olarte et al. investigated image formation by linear and nonlinear digital-scanned light-sheet fluorescence microscopy using Gaussian and Bessel beam profiles (Olarte et al., 2012). To characterize the effects of these four different imaging modalities they constructed a microscope that can easily be switched to any one of the four modalities while imaging the same specimen.

In the introduction to their publication these authors summarized the advantages and limitations of LSFM. The principal advantage is that photodamage to the specimen is significantly mitigated. The principal limitations include broadening of the light-sheet by scattering within the specimen, stripe artifacts along the illumination axis caused by the absorption and scattering of illumination by the specimen, and the lack of homogeneity in the light-sheet caused by diffraction from the limiting aperture. Furthermore, there is an important trade-off between the depth of field of the cylindrical lens and the thickness of the light-sheet, which leads to a reduced capacity to form optical sections.

Olarte et al. used high-NA microscope objectives with Bessel beam light-sheets, nonlinear excitation, and SIM. Using the same sample of *C. elegans* and a single ultrashort-pulsed laser their LSFM could switch between Bessel beams and Gaussian beams, and between linear and nonlinear excitation. Olarte et al.'s conclusions are again a reminder of the many trade-offs encountered in light microscopy: to deal with Bessel beam sidelobes a confocal detection system is advisable; Bessel beams combined with two-photon excitation improve image contrast and depth of penetration, but average power two orders of magnitude higher is required; and for specimens with low fluorescence, Gaussian beams are suggested since the signal-to-noise ratio is more important than contrast and optical-sectioning capacity.

A comprehensive critical publication in *Nature Protocols* by Gao et al. contains detailed information on the theory, design, construction, testing, and use of a Bessel beam plane illumination microscope (Gao et al., 2014). The authors make an important point that can be generalized: the microscope operates in several modes and each mode should be selected with full knowledge of the specimen and the parameters that should be optimized for an investigation to answer specific experimental questions. In other words, the investigator must be fully aware of the nature of the specimen and tune the hardware and software parameters of the instrument to best study specific questions regarding the specimen. The biological questions that microscopic investigation has to answer include knowing the physiological state of the live specimen and how to preserve it while imaging; knowing the trade-offs among resolution, optical sectioning, imaging speed, signal-to-noise ratio, and photo effects that alter the normal biology of the specimen. Gao et al. describe a Bessel beam LSFM that can operate in a number of modes such as Bessel beam plane illumination, Bessel structured illumination microscopy (SIM), two-photon Bessel beam plane illumination, and two-photon SIM. Each mode will be optimized for different parameters and the nature of the specific specimen will dictate selection of a particular mode of operation. Again, the issue of trade-offs occurs as many of the desired parameters cannot be independently selected without adversely affecting other parameters. For example, if we desire higher spatial resolution, then there is a negative trade-off between enhanced photodamage and

photobleaching and reduced imaging speed. Another example is if we want to optimize the optical-sectioning capability of the microscope, then the choice is between two-photon Bessel beam with plane illumination or two-photon SIM. Two-photon SIM has the limitation that it results in high photodamage to the specimen.

For readers who might want to construct their own optical instruments Gao et al. give a detailed description of the equipment as well as details of the optics, electronics, and software required to construct their version of a Bessel beam LSFM (Gao et al., 2014). Readers with the necessary skills can follow their instructions to build and operate a copy of their microscope.

I now turn to a discussion of the publication "Lattice Light-Sheet Microscopy: Imaging Molecules to Embryos at High Spatiotemporal Resolution" for its pedagogical value (Chen et al., 2014).

First, Chen et al. state the problem to be solved: for prolonged studies of live cells and living developing embryos it is necessary to have high spatial resolution to image small details of cell structure and to have high image acquisition rates to be able to follow the rapid dynamics of cellular processes. Unfortunately, high spatial resolution and high temporal resolution are mutually opposed.

Chen et al. cleverly address this problem by developing a new type of LSFM that has very thin light-sheets formed from two-dimensional optical lattices. They define an optical lattice as a two-dimensional or a three-dimensional interference pattern that is nondiffracting because the cross-section of the pattern remains constant as it propagates. The extremely thin light-sheet in the microscope achieves high axial resolution with negligible photobleaching of fluorescent molecules and background fluorescence that is out of the focal plane. Simultaneous illumination of the entire field of view allows image acquisition rates of hundreds of planes per second even with very low peak excitation.

Building upon Gao et al. (2012)'s "Noninvasive Imaging Beyond the Diffraction Limit of three-dimensional Dynamics in Thickly Fluorescent Specimens," Chen et al. used the design principles of three-dimensional, superresolution, structured illumination microscopy (SIM), which is developed in Chapter 13. To achieve the high temporal resolution required and make it simultaneous with superresolution Chen et al. used a linear array of seven parallel, noninteracting Bessel beams to compensate for the longer acquisition time required with SIM. A uniform light-sheet was formed by superresolution SIM or by dithering the Bessel beam lattice to form a uniform light-sheet.

What are the advantages and the limitations of lattice light-sheet microscopy for in vivo microscopy? The advantages include increased imaging speed (200–1000 planes imaged each second), as well as very low specimen photobleaching and phototoxicity (Chen et al., 2014). On the other hand, optical aberrations introduced by the specimen result in limitations of all forms of light microscopy including lattice light-sheet microscopy. In general, lattice light-sheet microscopy suffers from light scattering and optical aberrations at depths greater that 50 μm below the surface of the specimen. These problems can be mitigated by using adaptive optics in the

excitation and detection optical paths to compensate for these wavefront aberrations (Wang et al., 2014).

In 2018 a paper in *Science* highlighted the spectacular progress that has been made by combining lattice light-sheet microscopy (LLSM) with adaptive optics (AO) (Liu et al., 2018). LLSM yields three-dimensional images of whole cells with high spatiotemporal resolution. AO corrects optical aberrations from the specimen and the instrumentation. The combination of these two techniques (AO-LLSM) yields three-dimensional subcellular processes that can be imaged in their natural multicellular environment at diffraction-limited resolution for wide fields of view. Liu et al. (2018) have demonstrated their AO-LLSM imaging technique in zebrafish embryos, *C. elegans* nematodes, and leaves.

Another paper in *Science* demonstrated the new superresolution capabilities that result from combining expansion microscopy (ExM), a new specimen preparation technique, with lattice light-sheet microscopy (LLSM) to form a new technique that the authors called expansion/LLSM (ExLLSM), which I will now discuss (Gao et al., 2019). Details of specimen preparation, the instrument, and image processing and analysis are given in the paper's supplementary material. The authors are interested in optical-imaging studies of neural circuits in the brain. Such neural circuits are highly complex due to the thousands of protein types that comprise the brain and to the fact that the size of the components in these neural circuits vary over seven orders of magnitude.

First, Gao et al. 2019 state the utility and limitations of applying existing imaging techniques to the specific problem of imaging neural circuits in the brain. Electron microscopy has the capacity to image millimeter-sized brain tissue specimens. However, electron microscopy applied to brain tissue does not have sufficient contrast to differentiate specific proteins. Since there are biological variations between individual brains it is necessary to image a large number of specimens of brain tissue to map the average brain. Electron microscopy and superresolution fluorescence microscopy lack the speed to acquire brain-wide or cortex-wide images of multiple specimens. On the other hand, confocal fluorescence microscopy forms images that have adequate molecular contrast. However, it does not have sufficient resolution to localize specific molecules within submicrometer-sized neural structures. The lack of sufficient resolution precludes dense neural tracing (following the paths and connections of multiple neural structures). The various techniques that achieve superresolution fluorescence microscopy are limited by their rapid bleaching of fluorescent molecules, which precludes their use in large-volume imaging.

What is the technological solution that Gao and his coauthors achieved? In the preceding paragraphs I described the physics and instrumentation of lattice light-sheet microscopy. Expansion microscopy is a technique that achieves superresolution imaging with nanoscale precision of fixed cells and tissues on conventional microscopes that operate with diffraction-limited resolution (Chen et al., 2015; Chen et al., 2016).

The latest version of the technique is called protein retention ExM (proExM). It operates as follows (Tillberg et al., 2016). The tissue specimen is first fixed. This

technique permits the use of conventional fluorescently labeled antibodies and streptavidin, as well as fluorescent proteins to mark the features of interest. These marked features are then chemically anchored to a polyacrylamide/polyacrylate gel that permeates the tissue. Next, the tissue is treated with protease to digest the tissue. The digestion step yields the important result of removing lipids, proteins, and optically inhomogeneous molecules that are not bound to the gel. Finally, the gel is expanded in water isotropically. This expansion forms an enlarged phantom of the tissue that retains the tissue's original relative distribution of fluorescent tags (Tillberg et al., 2016). The effective resolution of the expanded gel is given by the original resolution of the diffraction-limited imaging microscope divided by the expansion factor (the number equal to the number of times gel dimensions are increased). The result is that the expanded gel has a refractive index that is similar to that of water. The expanded gel can be imaged without aberrations to a depth of more than 500 mm using conventional water immersion microscope objectives.

The critical question here is whether the ExM (proExM) technique preserves the structures that are present in the specimen prior to application of the expansion technique? The authors validated and demonstrated the utility of proExM for multicolor superresolution ($\sim$70 nm) imaging of cells and mammalian tissues in conventional microscopes. The validation procedure must be performed on a case-by-case basis for each type of specimen. The gold standard is correlative microscopy in which electron microscopy is used to validate the images obtained with expansion/LLSM (ExLLSM) .

Gao et al. combined the technique of proExM and LLSM and succeeded in imaging an entire *Drosophila* brain in 2–3 days (they did the same with a mouse cortex). With multiple fluorescent markers they achieved an effective resolution of $60 \times 60 \times 90$ nm giving a $4\times$ expansion. These studies resulted in molecular contrast and nanoscale resolution (Gao et al., 2019).

There are a number of limitations to the combined technique of proExM and LLSM when applied to imaging a fly brain or a mouse cortical column at the millimeter scale? The first limitation involves image acquisition speed and the researcher's requirement to complete data acquisition in a few days rather than several weeks. Gao et al. calculated that to achieve this data acquisition they would require imaging speeds of 100 million voxels per second. Such a data acquisition speed required the use of an image-processing and data storage pipeline that can operate at very high imaging speeds. The second problem involves photobleaching that quenches fluorescence from deep portions of the specimens before they are imaged. Photobleaching is worse for thicker specimens, longer imaging times, and higher illumination intensities that are used for faster imaging. The third problem involves the requirement to achieve close to isotropic resolution so that neural tracing and quantification of nanoscale neurostructures will not be limited by the axis of the poorest resolution.

11.5.2 Airy Beams and Their Use in Light-Sheet Fluorescence Microscopy

In the previous section (Sect. 11.5.1) the nonspreading nature of Bessel beams was introduced (Durnin et al., 1987). In fact, there are other types of diffraction-free beams including higher order Bessel beams and Mathieu beams (Gutiérrez-Vega et al., 2000). Berry and Balazs demonstrated theoretically that a solution to the Schrödinger equation for a free particle is a nonspreading Airy wave packet (Berry and Balazs, 1979; Siviloglou et al., 2007, 2008; Sztul and Alfano, 2008).

Vettenburg et al. published a paper in *Nature Methods* titled "Light-Sheet Microscopy Using an Airy Beam" (Vettenburg et al., 2014). This paper was followed by another by Yang et al. from the same group as Vettenburg et al. titled "A Compact Airy Beam light-sheet Microscope with a Tilted Cylindrical Lens" (Yang et al., 2014).

Vettenburg et al. begin with a discussion of the limitations of Gaussian and Bessel beams in LSFM. Although these topics were discussed in the previous section, the formulation of problems by Vettenburg et al. leads to their further understanding. First, why is the width of the illumination beam in LSFM important? The authors reply that the width of the illumination beam affects axial resolution. Specifically, the lateral or transverse resolution of the LSFM is dependent on the NA of the detection microscope objective, whereas axial resolution is dependent on two parameters: the NA of the detection microscope objective and the width of the illumination beam. In previous instrument designs employing Gaussian beams the uniformly illuminated field of view is very limited. Other LSFM designs use multiple exposures from various angles to increase the uniformly illuminated field of view. But that technique results in longer acquisition times and increased exposure of the specimen to illumination with increased risk for photobleaching and phototoxicity. What is required is the use of a thin illumination beam over a large field of view that would minimize the specimen's exposure to illumination and yield isotropic resolution. Second, what problems are associated with the use of Bessel beams? As discussed in previous sections, the sidelobes of Bessel beams increase the background fluorescence of the specimen. This problem can be mitigated by combining single-photon (not multiphoton) Bessel beams with SIM or with confocal systems.

Vettenburg et al. offer a solution to these problems in their paper "Light-Sheet Microscopy Using an Airy Beam" (Vettenburg et al., 2014). Their LSFM uses a single-photon Airy beam that propagates with an invariant intensity profile. The single-photon Airy beam is also "self-healing." Their claim includes the statement that an Airy beam yields a 10-fold increase in the field of view when compared to single-photon Gaussian beam light-sheets when a similar NA for the objectives in the illumination and detection optical paths is used. Bessel beam light-sheets also show a larger field-of-view over Gaussian beam light sheets in an LSFM. But the Airy beam with its asymmetric excitation pattern has another advantage over Bessel beams: for the Airy beam light-sheet all the fluorescence contributes to the contrast.

The authors explain this by noting that the sidelobes of Airy beams contribute positively to image formation. This is not the case for the sidelobes of Bessel beams.

I now consider Yang et al.'s paper "A Compact Airy Beam light-sheet Microscope with a Tilted Cylindrical Lens" (Yang et al., 2014). This paper is unique in that the authors present a low-cost modification to a standard LSFM that is based on the open-access light-sheet microscope OpenSPIM, an open-access platform that provides detailed instructions for the construction of a standard LSFM. An advantage of the light-sheet based on the Airy beam compared with the Bessel beam is that the Airy beam does not require scanning to form the light-sheet; instead a cylindrical lens is used to form the time-averaged distribution of intensity in the light-sheet. The result is a simpler optical system (Yang et al., 2014).

New developments in instrumentation and applications demonstrate the efficacy of these techniques. Scientists in the Shroff group recently published their latest findings in the open-access journal *Optica*. Wu et al.'s "Simultaneous Multiview Capture and Fusion Improves Spatial Resolution in Wide-Field and Light-Sheet Microscopy" is a particularly important paper showing rapid, continuous developments in the field of LSFM (Wu, Chandris, Winter, Kim, Jaumouillé, Kumar, Guo, Leung, Smith, Rey-Suarez, Liu, Waterman, Ramamurthi, La Riviere, and Shroff, 2016). This paper demonstrates an LSFM technique based on a three-lens microscope that uses digital deconvolution algorithms to form a single image with enhanced resolution.

Fluorescence microscopes collect a small part of the light emitted by the specimen at a given instant of time; this reduces spatial resolution. Wu et al. provide a solution in which their three-lens LSFM simultaneously detects and subsequently fuses and deconvolves multiple views of the specimen. Their LSFM can achieve twofold enhancement in three-dimensional resolution over single-view widefield or dual-view LSFM (Wu et al., 2016). In their triple-view LSFM, illustrated in Fig. 11.8, planar illumination is alternately formed by the upper microscope objectives (A/B), and at the same time the collection of light is from the other upper microscope objective (B/A) and the lower microscope objective (C). Microscope objectives A/B are fixed in position, while microscope objective C is displaced in the vertical direction to collect fluorescent light from the inclined illumination plane.

Another group, at the University of Cambridge in the United Kingdom, published their proposed technique in a paper in *Optics Letters* titled "triSPIM: light-sheet Microscopy with Isotropic Super-Resolution" (Manton and Rees, 2016). The triSPIM LSFM is a significant advance because experiments demonstrate an isotropic superresolution of 235 nm and computer simulations show a resolution of 120 nm. Three identical microscope objectives positioned at the corner of a cube define their respective orientations. Several raw images are required for each reconstructed image.

The following discussion highlights a recent major advance in light-sheet fluorescence microscopy outlined in a paper titled "*In Toto* Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level"

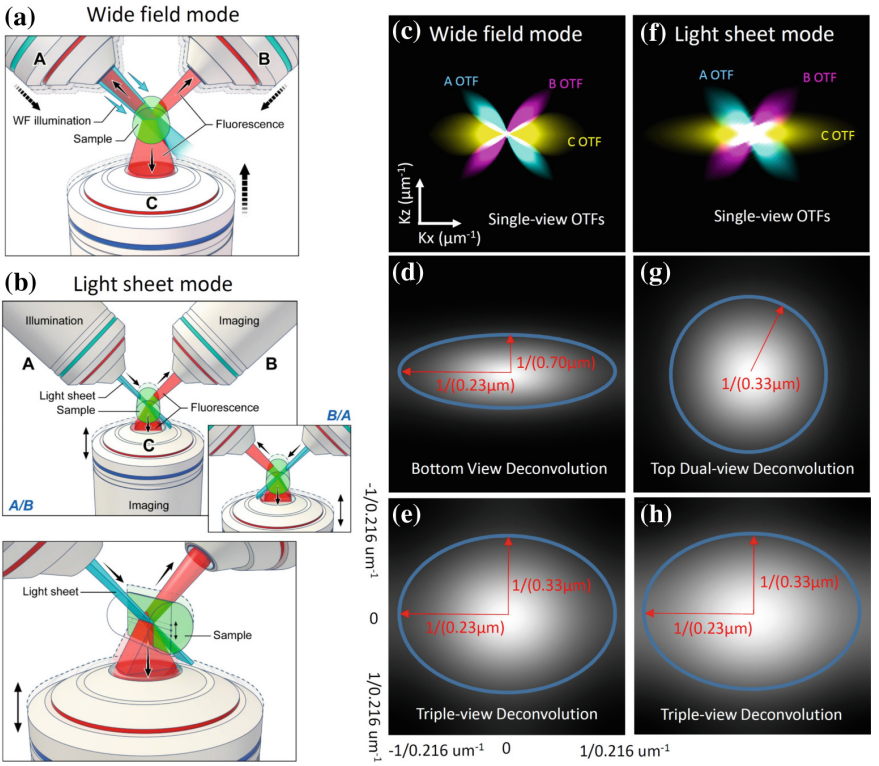


Fig. 11.8 Triple-view light-sheet fluorescence microscope. Schematic representation of triple-view widefield and light-sheet microscopy. **a** In triple-view widefield microscopy widefield illumination is introduced to the sample via one of the two upper objectives (A or B), and all three objectives (A, B, C) simultaneously collect fluorescence emissions from the sample volume. **b** In triple-view light-sheet microscopy planar illumination is alternately introduced by either of the upper objectives (A/B), with concurrent collection from the other upper objective (B/A) and lower objective (C). Note that A/B are stationary, while C is swept vertically to collect fluorescence from the inclined illuminated plane. Inset shows alternating illumination provided by B; lower panel shows perspective view. **c** Optical transfer functions (OTFs) for each objective in widefield mode assuming 0.8 NA for A/B and 1.2 NA for C. A cross section along *Kx*/*Kz* directions is shown to highlight resolution anisotropy between lateral and axial directions. Comparative OTFs are also shown after deconvolution of **d**. C alone and **e** all three views. **f** OTFs for each objective in light-sheet mode assuming the same NAs as in **c**. Comparative OTFs are also shown after joint deconvolution of **g**. A, B and **h** all three views. In both widefield and light-sheet microscopy deconvolution of all three views improves resolution. *Blue ellipses* and *red arrows* in **d**, **g**, **e**, and **h** indicate lateral and axial diffraction limits. From Wu, Y., Chandris, P., Winter, P. W., Kim, E. Y., Jaumouillé, V., Kumar, A., Guo, M., Leung, J. M., Smith, C., Rey-Suarez, I., Liu, H., Waterman, C. M., Ramamurthi, K. S., La Riviere, P., Shroff, H. (2016). Simultaneous multi-view capture and fusion improves spatial resolution in wide-field and ligh-sheet microscopy. *Optica*, **3**, 897–910. (*Optica* is an open-access, online-only journal dedicated to the rapid dissemination of high-impact peer-reviewed research across the entire spectrum of optics and photonics, published monthly by the Optical Society of America, OSA.)

(McDole et al., 2018). This paper fills the need for a microscope with the capability to image mouse development in toto with cellular resolution from the embryological stages of gastrulation to early organogenesis (formation of organs).

What is unique about this publication is that it is much more important because it is an exemplar of what science can be. First, McDole et al. clearly state the problems with current microscopic techniques. Second, they present their solution, which in this case is new hardware and software as the foundation of a new type of microscope. Third, the authors discuss the limitations of their new technology. Fourth, they use their new technological development to generate significant new knowledge. Fifth, they disseminate the knowledge with sufficient details to permit other researchers to replicate their innovative hardware and software.

The problem McDole et al. tackled concerned the development of the mouse embryo, which is important in the general study of mammalian development. They were particularly concerned with cell proliferation, differentiation, migration, and the formation of tissues and organs, which are very difficult to study with live imaging in the developing embryo. The mouse embryo is extremely sensitive to light-induced phototoxicity. Previous studies of mouse embryo development were performed with confocal microscopes. However, these studies were limited to imaging only parts of the embryo and imaging was limited to a 24 hour period due to the onset of phototoxicity. Furthermore, as a mouse embryo develops it undergoes changes in position, size, and optical properties, which affect the capabilities of optical microscopes during long-term imaging.

McDole et al.'s innovative solution to the limitations of previous microscopes built on the previous work of Tomer et al. (2012) and Royer et al. (2016) to construct an adaptive light-sheet microscope that images mouse development at the single-cell level. Their adaptive light-sheet microscope can adapt to dramatic changes in size, shape, and optical properties of a postimplantation mouse embryo, and armed with that capability it can image mouse embryo development from gastrulation to early organogenesis. Their adaptive microscope can culture and image postimplantation mouse embryos over a 48 hour period at the cellular level. The authors also developed software that controls the microscope in its adaptive functions and detects cell division.

McDole et al. described the limitations of their new adaptive light-sheet microscope. The size of the growing embryo has a 6 mm limit along the anteroposterior axis and a 1 mm limit along the mediolateral axis. The adaptive light-sheet microscope can make optical corrections once every 3 minutes. This may not be sufficient for higher resolution optical imaging. Another limitation is the effect of light scattering and higher order optical aberrations, which limits the depth of imaging at high resolution.

The significance of the knowledge generated by McDole et al.'s adaptive light-sheet microscope can be seen from the fate maps and videos that augment their groundbreaking publication. The authors reconstructed individual cell tracks, dynamic fate maps, and maps of tissue morphogenesis through the entire embryo. A fate map is a two-dimensional representation that shows cell location and fate

(tissue or organ that cells form). Fate mapping is a method used to trace cell lineages and is a basic tool of developmental biology.

These results were obtained from several embryos and McDole et al. formed an "average mouse embryo" from analysis of all the data, which allowed the authors to create a statistical, dynamic map of development from multiple embryos. Dynamic fate maps are important since they permit visualization of development at a unique level of detail. An example is the imaging of patterns of cell division and cell behavior during neural tube formation and heart development.

Besides innovation and development of unique hardware and software McDole et al. attached great value to dissemination of details of their work to serve diverse groups and users that may differ in their expertise. To this end their publication contains full documentation of the code for all their computational tools, guides for the practical use of software, complete blueprints of the adaptive light-sheet microscope, all the databases of embryonic development, and four-dimensional dynamic visualizations at cellular resolution. All this information is available in a public depository.

However, not all researchers have the inclination or technical ability to construct their own adaptive light-sheet microscope. To meet the needs of these individuals McDole et al. partnered with the Janelia Advanced Imaging Center and constructed a copy of the adaptive light-sheet microscope described in their publication. This instrument is available to the scientific community and is maintained by people with expertise in the working of the microscope and its software components. They are happy to act as mentors to users in all aspects of their work such as sample preparation, imaging, and image analysis (https://www.aicjanelia.org/ accessed April 3, 2019).

Developments that have taken place in LSFM are works in progress, and from a study of progress made in the last decade we can predict similar progress in the next. While advances in instrumentation and biological probes typically result in the generation of new biological knowledge, it is necessary to judge new instrumentation advances using the critical metric of asking questions and seeing whether the new instruments can answer them. Another way of judging them is by looking at the impact and significance of the new biological knowledge discovered by use of the new instruments.

References

Bassi, A., Schmid, B., and Huisken, J. (2015). Optical tomography complements light-sheet microscopy for in toto imaging of zebrafish development. Development, **142**, 1016–1020.

Berry, M. V., and Balazs, N. L. (1979). Nonspreading wave packets. American Journal of Physics, **47**, 264–267.

Boot, M. J., Westerberg, C. H., Sanz-Ezquerro, J., Cotterell, J., Schweitzer, R., Torres, M., and Sharpe, J. (2008). In vitro whole-organ imaging: 4D quantification of growing mouse limb buds. Nature Methods, **5**(7), 609–612. https://doi.org/10.1038/nmeth.1219. Epub 2008 May 30.

Bouchal, Z., Wagner, J., and Chlup, M. (1998). Self-reconstruction of a distorted nondiffracting beam. Optics Communications, **151**, 207–211.

Brzobohatý, O., Čižmár, T., and Zemánek, P. (2008). High quality quasi-Bessel beam generated by round-tip axicon. Optics Express, **16**, 12688–12700.

Buytaert, J. A., and Dirckx, J. J. (2007). Design and quantitative resolution measurements of an optical virtual sectioning three-dimensional imaging technique for biomedical specimens, featuring two-micrometer slicing resolution. Journal of Biomedical Optics, **12**, 014039-1—014039-113.

Buytaert, J. A. N., Descamps, E., Adriaens, D., and Dirckx, J. J. J. (2012). Review Article. The OPFOS microscopy family: High-resolution optical sectioning of biomedical specimens. Anatomy Research International, **2012**, Article ID 206238, 9 pages. http://dx.doi.org/10.1155/2012/206238 (Accessed April 3, 2019).

Chen, B.-C., Legant, W. R., Wang, K., Shao, L., Milkie, D. E., Davidson, M. W., Janetopoulos, C., Wu, X. S., Hammer, J. A., Liu, Z., English, B. P., Mimori-Kiyosue, Y., Romero, D. P., Ritter, A. T., Lippincott-Schwartz, J., Fritz-Laylin, L., Mullins, R. D., Mitchell, D. M., Bembenek, J. N., Reymann, A.-C., Böhme, R., Grill, S. W., Wang, J. T., Seydoux, G., Tulu, U. S., Kiehart, D. P., and Betzig, E. (2014). Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution, Supplementary Materials. Science, **346**, 1257998. https://doi.org/10.1126/science.1257998 Medline.

Chen, F., Tillberg, P. W., and Boyden, E. S. (2015). Optical imaging. Expansion microscopy. Science, **347**, 543–548.

Chen, F., Wassie, A. T., Cote, A. J., Sinha, A., Alon, S., Asano, S., Daugharthy, E. R., Chang, J-B., Marblestone, A., Church, G. M., Raj, A., and Boyden, E. S. (2016). Nanoscale imaging of RNA with expansion microscopy. Nature Methods, published online 4 July 2016, https://doi.org/10.1038/nmeth.3899.

Dodt, H-U., Leischner, U. Schierloh, A., Jährling, N., Mauch, C. P., Deininger, K., Deussing, J. M., Eder, M., Zieglgänsberger, W., and Becker, K. (2007). Ultramicroscopy: Three-dimensional visualization of neuronal networks in the whole mouse brain. Nature Methods, **4**, 331–336.

Dunsby, C. (2008). Optically sectioned imaging by oblique plane microscopy. Optics Express, **16**, 20306–20316.

Durnin, J. (1987). Exact solutions for nondiffracting beams. I. The scalar theory. Journal of the Optical Society of America A, **4**, 651–654.

Durnin, J., Miceli, Jr., J. J., and Eberly, J. H. (1987). Diffraction-free beams. Physical Review Letters, **58**, 1499–1501.

Durnin, J., Miceli Jr., J. J., and Eberly, J. H. (1988). Comparison of Bessel and Gaussian beams. Optics Letters, **13**, 79–80.

Fahrbach, F. O., Simon, P., and Rohrbach, A. (2010). Microscopy with self-reconstructing beams. Nature Photonics, **4**, 780–785.

Fauver, M., Seibel, E., Richard Rahn, J. R., Meyer, M. G., Patten, F. W., Neumann, T., and Nelson, A. C. (2005). Three-dimensional imaging of single isolated cell nuclei using optical projection tomography. Optics Express, **13**, 4210–4223.

Fuchs, E., Jaffe, J. S., Long, R. A., and Azam, F. (2002). Thin laser light-sheet microscope for microbial oceanography. Optics Express, **10**, 145–154.

Gao, R., Asano, S. M., Upadhyayula, S., Pisarev, I., Milkie, D. E., Liu, T-L., Singh, V., Graves, A., Huynh, G. H., Zhao, Y., Bogovic, J., Colonell, J., Ott, C. M., Zugates, C., Tappan, S., Rodriguez, A., Mosaliganti, K. R., Sheu, S-H., Pasolli, H. A., Pang, S., Xu, C. S., Megason, S. G., Hess, H., Lippincott-Schwartz, J., Hantman, A., Rubin, G. M., Kirchhausen, T., Saalfeld, S., Aso, Y., Boyden, E. S., Eric Betzig. E. (2019). Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution. Science, **363**, 245–261, January 18, 2019. https://doi.org/10.1126/science.aau8302.

Gao, L., Shao, L., Higgins, C. D., Poulton, J. S., Peifer, M., Davidson, M. W., Wu, X., Goldstein, B., and Betzig, E. (2012). Noninvasive imaging beyond the diffraction limit of three-dimensional dynamics in thickly fluorescent specimens. Cell, **151**, 1370–1385.

Gao, L., Shao, L., Chen, B.-C., and Betzig, E. (2014). three-dimensional live fluorescence imaging of cellular dynamics using Bessel beam plane illumination microscopy. Nature Protocols, **9**, 1083–1101.

Greger, K, Swoger, J., and Stelzer, E. H. (2007). Basic building units and properties of a fluorescence single plane illumination microscope. Review of Scientific Instruments, **78**, 023705.

Gualda, E. J., Vale, T., Almada, P., Feijó, J. A., Martins, G. G., and Moreno, N. (2013). OpenSpinMicroscopy: An open-source integrated microscopy platform. Nature Methods, **10**, 599–600. https://sites.google.com/site/openspinmicroscopy/ (Accessed April 3, 2019).

Gutiérrez-Vega, J. C., Iturbe-Castillo, M. D., and Chávez-Cerda, S. (2000). Alternative formulation for invariant optical fields: Mathieu beams. Optics Letters, **25**, 1493–1495.

Haisch, C. (2012). Optical tomography. Annual Review Analytical Chemistry, **5**, 57–77.

Holekamp, T. F., Turaga, D., and Holy, T. E. (2008). Fast three-dimensional fluorescence imaging of activity in neural populations by objective-coupled planar illumination microscopy. Neuron, **57**, 661–672.

Huber, D., Keller, M., and Robert, D. (2001). three-dimensional light scanning macrography. Journal of Microscopy, **203**, 208–213.

Huisken, J., Swoger, J., Del Bene, F. Wittbrodt, J., and Stelzer, E. H. K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science, **305**, 1007–1009.

Huisken, J., and Stainier, D. Y. R. (2007). Even fluorescence excitation by multidirectional selective plane illumination microscopy (mSPIM). Optics Letters, **32**, 2608–2610.

Huisken, J., and Stainier, D. Y. R. (2009). Selective plane illumination microscopy techniques in developmental biology. Development, **136**, 1963–1975.

Kak, A. C., and Slaney, M. (1988). *Principles of computerized tomographic imaging*. IEEE Press.

Keller, P. J., and Stelzer, E. H. K. (2008). Quantitative in vivo imaging of entire embryos with digital scanned laser light-sheet fluorescence microscopy. Current Opinion in Neurobiology, **18**, 624–632.

Keller, P. J., Schmidt, A. D., Santella, A., Khairy, K., Bao, Z., Wittbrodt, J., and Stelzer, E. H. K. (2010). Fast, high-contrast imaging of animal development with scanned light-sheet-based structured-illumination microscopy. Nature Methods, **7**, 637–642.

Keller, P. J., Ahrens, M. B., and Freeman J. (2015). Light-sheet imaging for systems neuroscience. Nature Methods, **12**, 17–29.

Krzic, U., Gunther, S., Saunders, T. E., Streichan, S. J., and Hufnagel, L. (2012). Multiview light-sheet microscope for rapid in toto imaging. Nature Methods, **9**, 730–733.

Kikuchi, S., and Sonobe, K. (1994). Three-dimensional computed tomography for optical microscopes. Optics Communications, **107**, 432–444.

Kumar, A., Wu, Y., Christensen, R., Chandris, P., Gandler, W., McCreedy, E., Bokinsky, A., Colón-Ramos, D. A., Bao, Z., McAuliffe, M., Rondeau, G., and Shroff, H. (2014). Dual-view plane illumination microscopy for rapid and spatially isotropic imaging. Nature Protocols, **9**, 2555–2573.

Lim, D., Chu, K. K., and Mertz, J. (2008). Wide-field fluorescence sectioning with hybrid speckle and uniform-illumination microscopy. Optics Letters, **33**, 1819–1821.

Lim, D., Ford, T. N., Chu, K. K., and Mertz, J. (2011). Optically sectioning imaging with speckle illumination HiLo microscopy. Journal of Biomedical Optics, **16**, 016014.

Lindek, S., Pick, R., and Stelzer, E. H. K. (1994). Confocal theta microscope with three objectives lenses. Review of Scientific Instruments, **65**, 3367–3372.

Lindek, S., and Stelzer, E. H. K. (1994). Confocal theta microscopy and 4Pi-confocal theta microscopy. In: C. J. Cogswell and K. Carlsson (Eds.), *Three-dimensional microscopy, image acquisition and processing*. SPIE, v. 2184 of *Proceedings of the SPIE*, p. 188.

Liu, T.-L. et al. (2018). Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms. Science, **360**, 284–297. eaaq1392 (2018). https://doi.org/10.1126/science.aaq1392. Supplementary Materials www.sciencemag.org/content/360/6386/eaaq1392/suppl/DC1 (Accessed April 3, 2019). [Science 360, eaaq1392 (2018). https://doi.org/10.1126/science.aaq1392. Observing the cell in its native state: Imaging subcellular dynamics in

multicellular organisms. Tsung-Li Liu, Srigokul Upadhyayula, Daniel E. Milkie, Ved Singh, Kai Wang, Ian A. Swinburne, Kishore R. Mosaliganti, Zach M. Collins, Tom W. Hiscock, Jamien Shea, Abraham Q. Kohrman, Taylor N. Medwig, Daphne Dambournet, Ryan Forster, Brian Cunniff, Yuan Ruan, Hanako Yashiro, Steffen Scholpp, Elliot M. Meyerowitz, Dirk Hockemeyer, David G. Drubin, Benjamin L. Martin, David Q. Matus, Minoru Koyama, Sean G. Megason, Tom Kirchhausen, Eric Betzig].

Manton, J. D., and Rees, E. J. (2016). triSPIM: light-sheet microscopy with isotropic super-resolution. Optics Letters, **41**, 4170–4173.

Mayer, J., Robert-Moreno, A., Danuser, R., Stein, J. V., Sharpe, J., and Swoger J. (2014). OPTiSPIM: Integrating optical projection tomography in light-sheet microscopy extends specimen characterization to nonfluorescent contrasts. Optics Letters, **39**, 1053–1056.

McDole, K., Guignard, L., Amat, F., Berger, A., Malandain, G., Royer, L. A., Turaga, S., C., Branson, K., and Keller, P. J. (2018). In toto imaging and reconstruction of post-implantation mouse development at the single-cell level. Cell, **175**, 1–18. https://doi.org/10.1016/j.cell.2018.09.031 (Accessed April 3, 2019).

McGloin, D., and Dholakia, K. (2005). Bessel beams: Diffraction in a new light. Contemporary Physics, **46**, 15–28.

McLachlan, D., Jr. (1968). Microscope. US Patent 3,398,634. August 27, 1968.

Mertz, J. (2011). Optical sectioning microscopy with planar or structured illumination. Nature Methods, **8**, 811–819.

Mertz J., and Kim, J. (2010). Scanning light-sheet microscopy in the whole mouse brain with HiLo background rejection. Journal of Biomedical Optics, **15**, 016027.

Neil, M.A., Juškaitis, R., and Wilson, T. (1997). Method of obtaining optical sectioning by using structured light in a conventional microscope. Optics Letters, **22**, 1905–1907.

Olarte, O. E., Andilla, J., Gualda, E. J., and Loza-Alvarez, P. (2018). Light-sheet microscopy: A tutorial. Advances in Optics and Photonics, **10**, 111–179. https://doi.org/10.1364/AOP.10.000111 (Accessed April 3, 2019).

Olarte, O. E., Licea-Rodriguez, J., Palero, J. A., Gualda, E. J., Artigas, D., Mayer, J., Swoger, J., Sharpe, J., Rocha-Mendoza, I., Rangel-Rojo, R., and Loza-Alvarez, P. (2012). Image formation by linear and nonlinear digital scanned light-sheet fluorescence microscopy with Gaussian and Bessel beam profiles. Biomedical Optics Express, **3**, 1492–1505.

Pitrone P. G., Schindelin J., Stuyvenberg L., Preibisch S., Weber M., Eliceiri K. W., Huisken J., and Tomancak, P. (2013). Nature Methods, **10**, 598–599.

Planchon, T. A., Gao, L., Milkie, D. E., Davidson, M. W., Galbraith, J. A., Galbraith, C. G., and Betzig, E. (2011). Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. Nature Methods, **8**, 417–423.

Reynaud, E. G., Kržič, U., Greger, K., and Stelzer, E. H. K. (2008). light-sheet-based fluorescence microscopy: More dimensions, more photons, and less photodamage. HFSP Journal, **2**, 266–275.

Reynaud, E. G., Peychl, J., Huisken, J., and Tomancak, P. (2015). Guide to light-sheet microscopy for adventurous biologists. Nature Methods, **12**, 30–34.

Royer, L. A., Lemon, W. C., Chhetri, R. K., Wan, Y., Coleman, M., Myers, E. W., and Keller, P. J. (2016). Adaptive light-sheet microscopy for long-term, high resolution imaging in living organisms. Nature Biotechnology, **34**, 1267–1278.

Santi, P. A., Johnson, S. B., Hillenbrand, M., GrandPre, P. Z., Glass, T. J., and Leger, J. R. (2009). Thin-sheet laser imaging microscopy for optical sectioning of thick tissues. BioTechniques, **46**, 287–294.

Schmid, B., Shah, G., Scherf, N., Weber, M., Thierbach, K., Campos, C. P., Roeder, I., Aanstad, P., and Huisken, J. (2013). High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics. Nature Communications, **4**, 1–10.

Sharpe, J. (2004). Optical Projection Tomography. Annual Review of Biomedical Engineering, **6**, 209–228.

Sharpe, J. (2008). In vitro whole-organ imaging: 4D quantification of growing mouse limb buds. Nature Methods, **5**, 609–612.

Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sørensen, J., Baldock, R., and Davidson, D. (2002). Optical projection tomography as a tool for three-dimensional microscopy and gene expression studies. Science, **296**, 541–545.

Siedentopf, H., and Zsigmondy, R. (1903). Uber Sichtbarmachung und Größenbestimmung ultramikoskopischer Teilchen, mit besonderer Anwendung auf Goldrubingläser (About visualization and size determination of ultramicroscopic particles, with particular application to gold red glasses). Annalen der Physik, **10**, 1–39.

Simon, W. (1965). Photomicrography of deep fields. Review of Scientific Instruments, **36**, 1654–1655.

Siviloglou, G. A., Broky, J., Dogariu, A., and Christodoulides, D. N. (2007). Observation of accelerating Airy Beams. Physical Review Letters, **99**, 213901-1–213901-4.

Siviloglou, G. A., Broky, J., Dogariu, A., and Christodoulides, D. N. (2008). Ballistic dynamics of Airy beams. Optics Letters, **33**, 207–209.

Spalteholtz, W. (1911). Über das Durchsichtigmachen von Menschlichenund Tierischen Präparaten (On making transparent human and animal tissue samples). Leipzig: Verlag S. Hirzel.

Stelzer, E. H. K., Lindek, S., Albrecht, S., Pick, R., Ritter, G., Salmon, N. J., and Stricker, R. (1995). A new tool for the observation of embryos and other large specimens: Confocal theta fluorescence microscopy. Journal of Microscopy, **179**(1), 1–10.

Swoger, J, Verveer, P, Greger, K, Huisken, J., and Stelzer, E. H. K. (2007). Multi-view image fusion improves resolution in three-dimensional microscopy. Optics Express, **15**, 8029–8042.

Sztul, H. I., and Alfano, R. R. (2008). The Poynting vector and angular momentum of Airy beams. Optics Express, **16**, 9411–9416.

Tillberg, P.W., Chen, F., Piatkevich, K. D., Zhao, Y., Yu, C. C., English, B. P., Gao, L., Martorell, A., Suk, H-J., Yoshida, F., DeGennaro, E. M., Roossien, D. H., Gong, G., Seneviratne, U., Tannenbaum, S. R., Desimone, R., Cai, D., and Boyden, E. S. (2016). Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. Nature Biotechnology, **34**, 987–992. https://doi.org/10.1038/nbt.3625; pmid: 27376584 (Accessed April 3, 2019).

Tomer, R., Khairy, K. Amat, F., and Keller, P. J. (2012). Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. Nature Methods, **9**, 755–763.

Vettenburg, T., Dalgarno, H. I. C., Nylk, J., Coll-Lladó, C., Ferrier, D. E. K., Čižmár, T., Gunn-Moore, F. J., and Dholakia, K. (2014). Light-sheet microscopy using an Airy beam. Nature Methods, **11**, 541–544.

Voie, A. H. (1996). Three-dimensional reconstruction and quantitative analysis of the mammalian cochlea [dissertation]. Seattle (WA), University of Washington.

Voie, A. H., Burns, D. H., and Spelman, F. A. (1993). Orthogonal-plane fluorescence optical sectioning: three-dimensional imaging of macroscopic biological specimens. Journal of Microscopy, **170**, 229–236.

Voie, A. H., and Spelman, F. A. (1995). Three-dimensional reconstruction of the cochlea from two dimensional images of optical sections. Computerized Medical Imaging and Graphics, **19**, 377–384.

Voie, A. H. (2002). Imaging the intact guinea pig tympanic bulla by orthogonal-plane fluorescence optical sectioning microscopy. Hearing Research, **171**, 119–128.

Wang, K., Milkie, D. E., Saxena, A., Engerer, P., Misgeld, T., Bronner, M. E., Mumm, J., and Eric Betzig, E. (2014). Rapid adaptive optical recovery of optimal resolution over large volumes. Nature Methods, **11**, 625–628.

Wu, Y., Ghitani, A., Christensen, R., Santella, A., Du, Z., Rondeau, G., Bao, Z., Colón-Ramos, D., and Shroff, H. (2011). Inverted selective plane illumination microscopy (*i*SPIM) enables coupled cell identity lineaging and neurodevelopmental imaging in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America, **108**, 17708–17713.

Wu, Y., Christensen, R., Colón-Ramos, D., and Shroff, H. (2013). Advanced optical imaging techniques for neurodevelopment. Current Opinion in Neurobiology, **23**, 1090–1097.

Wu, Y., Chandris, P., Winter, P. W., Kim, E. Y., Jaumouillé, V., Kumar, A., Guo, M., Leung, J. M., Smith, C., Rey-Suarez, I., Liu, H., Waterman, C. M., Ramamurthi, K. S., La Riviere, P. J., and Shroff, H. (2016). Simultaneous multiview capture and fusion improves spatial resolution in wide-field and light-sheet microscopy. Optica, **3**, 897–910.

Yang, Z., Prokopas, M., Nylk, J., Coll-Lladó, C., Gunn-Moore, F. J., Ferrier, D. E. K., Vettenburg, T., and Kishan, D. (2014). A compact Airy beam light-sheet microscope with a tilted cylindrical lens. Biomedical Optics Express, **5**, 3434–3442.

Zampol, P. (1960). Method of Photography. US Patent 2928734 A.

Further Reading

Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., and Keller, P. J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nature Methods, **10**, 413–420.

Booth, M. J. (2007). Adaptive optics in microscopy. Philosophical Transactions of the Royal Society A, **365**, 2829–2843.

Bouchard, M. B., Voleti, V., Mendes, C. S., Lacefield, C., Grueber, W. B., Mann, R. S., Bruno, R. M., and Hillman, E. M. C. (2015). Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms. Nature Photonics, **9**, 113–119.

Breuninger, T., Greger, K., and Stelzer E. H. K. (2007). Lateral modulation boosts image quality in single plane illumination fluorescence microscopy. Optics Letters, **32**, 1938–1940.

Dean, K. M., Roudot, P., Welf, E. S., Danuser, G., and Fiolka, R. (2015). Deconvolution-free subcellular imaging with axially swept light-sheet microscopy. Biophysical Journal, **108**, 2807–2815.

Denk, W., Strickler, J. H., and Webb, W. W. (1990). Two-photon laser scanning fluorescence microscopy. Science, **248**, 73–76.

Engelbrecht, C. J., and Stelzer, E. H. K. (2006). Resolution enhancement in a light-sheet-based microscope (SPIM). Optics Letters, **31**, 1477–1479.

Fahrbach, F. O., Voigt, F. F., Schmid, B., Helmchen, F., and Huisken, J. (2013). Rapid three-dimensional light-sheet microscopy with a tunable lens. Optics Express, **21**, 21010–21026.

Greenberger, D. M. (1980). Comment on "non-spreading wave packets." American Journal of Physics, **48**, 256.

Gualda, E. J., Simão, D., Pinto, C., Alves, P. M., and Brito, C. (2014). Imaging of human differentiated three-dimensional neural aggregates using light-sheet fluorescence microscopy. Frontiers in Cellular Neuroscience, **6**. http://journal.frontiersin.org/article/10.3389/fncel.2014.00221/abstract (Accessed: April 3, 2019). https://doi.org/10.3389/fncel.2014.00221.

Gustafsson, M. G. L. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of Microscopy, **198**, 82–87.

Gustafsson, M. G. L., Shao, L., Carlton, P. M., Wang, C. J. R., Golubovskaya, I. N., Cande, W. Z., Agard, D. A., and Sedat, J. W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophysical Journal, **94**, 4957–4970.

Helmchen, F., and Denk, W. (2005). Deep tissue two-photon microscopy. Nature Methods, **2**, 932–940.

Kamei, M., and Weinstein, B. M. (2005). Long-term time-lapse fluorescence imaging of developing zebrafish. Zebrafish, **2**, 113–123.

Keller, P. J., Schmidt, A. D., Wittbrodt, J., and Stelzer, E. H. K. (2008). Reconstruction of zebrafish early embryonic development by scanned light-sheet microscopy. Science, **322**, 1065–1069.

Konopka, C. A., and Bednarek, S. Y. (2008). Variable-angle epifluorescence microscopy: A new way to look at protein dynamics in the plant cell cortex. Plant Journal, **53**, 186–196.

Kumar, S., Wilding, D., Sikkel, M. B., Lyon, A. R., MacLeod, K. T., and Dunsby, C. (2011). High-speed two-dimensional and three-dimensional fluorescence microscopy of cardiac myocytes. Optics Express, **19**, 13839–13847.

Lim, J., Lee, H. K., Yu, W., and Ahmed, S. (2014). light-sheet fluorescence microscopy (LSFM): Past, present and future. Analyst, **139**, 4758–4768.

Lindek, S., Salmon, N., Cremer, C., and Stelzer, E. H. K. (1994). Theta microscopy allows phase regulation in 4Pi(A)-confocal two-photon fluorescence microscopy. Optik, **98**, 15–20.

Lindek, S., Cremer, C., and Stelzer, E. H. K. (1996). Confocal theta fluorescence microscopy with annular apertures. Applied Optics, **35**, 126–130.

Mahou, P., Zimmerley, M., Loulier, K., Matho, K. S., Labroille, G., Morin, X., Supatto, W., Livet, J., Débarre, D., and Beaurepaire, E. (2012). Multicolor two-photon tissue imaging by wavelength mixing. Nature Methods, **9**, 815–818.

Mahou, P., Vermot, J., Beaurepaire, E., and Supatto, W. (2014). Multicolor two-photon light-sheet microscopy. Nature Methods, **11**, 600–601.

Overfelt, P. L., and Kenney, C. S. (1991). Comparison of the propagation characteristics of Bessel, Bessel-Gauss, and Gaussian beams diffracted by a circular aperture. Journal of the Optical Society of America A, **8**, 732–745.

Palero, J., Santos, S. I. C. O., Artigas, D., and Loza-Alvarez, P. (2010). A simple scanless two-photon fluorescence microscope using selective plane illumination. Optics Express, **18**, 8491–8498.

Pan, C., Cai, R., Quacquarelli, F. P., Ghasemigharagoz, A., Lourbopoulos, A., Matryba, P., Plesnila, N., Dichgans, M., Hellal, F., and Ertürk A. (2016). Shrinkage-mediated imaging of entire organs and organisms using uDISCO. Nature Methods. Received 1 March; accepted 26 July; published online 22 August 2016; https://doi.org/10.1038/nmeth.3964.

Ritter, J. G., Veith, R., Siebrasse, J.-P., and Kubitscheck, U. (2008). High contrast single-particle tracking by selective focal plane illumination microscopy. Optics Express, **16**, 7142–7152.

Santi, P. A. (2011). light-sheet fluorescence microscopy: A review. Journal of Histochemistry & Cytochemistry, **59**, 129–138.

Siedentopf, H. (1903). IX. On the rendering visible of ultra-microscopic particles and of ultra-microscopic bacteria. Journal of the Royal Microscopical Society, **23**, 573–578.

Shroff, H. (2016). Simultaneous multi-view capture and fusion improves spatial resolution in wide-field and light-sheet microscopy. Optica, **3**, 897–910.

Silvestri, L., Bria, A., Sacconi, L., Iannello, G., and Pavone, F. S. (2012). Confocal light-sheet microscopy: Micron-scale neuroanatomy of the entire mouse brain. Optics Express, **20**, 20582–20598.

Siviloglou, G. A., and Christodoulides, D. N. (2007). Accelerating finite energy Airy beams. Optics Letters, **32**, 979–981.

Stelzer, E. H. K. (2015). Light-sheet fluorescence microscopy for quantitative biology. Nature Methods, **12**, 23–26.

Stelzer, E. H. K., and Lindek, S. (1994). Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: Confocal theta microscopy. Optics Communications, **111**, 536–547.

Stelzer, E. H. K., Hell, S.W., Lindek, S., Stricker, R., Pick, R., Storz, C., Ritter, G., and Salmon, N. (1994). Nonlinear absorption extends confocal fluorescence microscopy into the ultraviolet regime and confines the observation volume. Optics Communications, **104**, 223–228.

Swoger, J, Huisken, J, and Stelzer, E. H. (2003). Multiple imaging axis microscopy improves resolution for thick-sample applications. Optics Letters, **28**, 1654–1656.

Tokunaga, M., Imamoto, N., and Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. Nature Methods, **5**, 159–161.

Truong, T. V., Supatto, W., Koos, D. S., Choi, J. M., and Fraser, S. E. (2011). Deep and fast live imaging with two-photon scanned light-sheet microscopy. Nature Methods, **8**, 757–760.

Vaziri, A., and Shank, C. V. (2010). Ultrafast widefield optical sectioning microscopy by multifocal temporal focusing. Optics Express, **18**, 19645–19655.

Zanacchi, F. C., Lavagnino, Z., Donnorso, M. P., Bue, A. D. Furia, L., Faretta, M., and Diaspro, A. (2011). Live-cell three-dimensional super-resolution imaging in thick biological samples. Nature Methods, **8**, 1047–1049.

Chapter 12
Phase Microscopy to Enhance Contrast

12.1 Introduction

This chapter answers the following question: How to image transparent objects (e.g., live unstained cells) in the light microscope? Most biological specimens are transparent (i.e., in the light microscope the image is of constant intensity across the field of view and there is no contrast). Transparent specimens, such as thin layers of cells, are called phase objects. They absorb almost no incident light; however, they change the phase of the transmitted light compared with the background. The problem of imaging phase objects in the light microscope is that the human eye detects differences in light intensity but cannot detect phase differences. Similarly, photodetectors only detect light intensity.

The difficulty that transparent specimens present has long been known to microscopists. Some attempts at early solutions involved the use of highly absorbing stains or dyes that altered light absorption of the specimen compared with the background. With this in mind, in this chapter I describe three optical methods capable of enhancing contrast of the specimen in the light microscope.

I discuss three different approaches that optically manipulate the phases of light beams in the microscope to generate contrast: phase contrast microscopy (PCM), differential interference microscopy (DIC), and modulation contrast microscopy (Pluta, 1988, 1989, 1993). I will compare these techniques and discuss the theory, instrumentation, and applications of each one.

To see an image of a specimen in a microscope it is necessary to generate contrast. Resolution and contrast are equally important in the design and construction of microscopes. In Chapters 6 and 8 I introduced the connection between Abbe's theoretical and experimental investigations into the theory of image formation in the light microscope based on his diffraction theory. Zernike in his Nobel Lecture explains not only how Abbe missed the discovery of phase contrast, but how Abbe hindered the development of phase contrast. In this chapter I discuss the phase contrast microscope and the differential contrast microscope in more detail

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_12

and review some of the modern developments that often had their origins in many of the older publications.

The early development of phase microscopy led to ever-expanding developments in the theory, instrumentation, and applications of variants of phase microscopy (Pluta, 1988, 1989, 1993). The interference of light and the development of various types of interferometers is the foundation of phase microscopy and its variants (Malacara, 2007). The application of phase imaging to living cells and tissues is reviewed in a recent book (Popescu, 2011). The Schlieren method represents another way of generating contrast from a phase object (Lipson et al., 2011). The Schlieren method was first observed by Hooke (1665). Jean Bernard Léon Foucault invented the Schlieren method in 1859, and August Toepler in 1864 further developed it and applied it to the investigation of supersonic motion.

It seems incredible that Abbe did not discover phase contrast microscopy, especially since he used glass wedges in the back focal plane of the microscope objective to manipulate the phase between different diffraction orders (Bennett et al., 1951). In fact, Abbe discovered that contrast between lines and their spaces in the object grating were reversed when he caused a phase difference of 180° between zero-order and first-order diffraction spectra (Bratuscheck, 1892). Bratuscheck discovered that he could enhance the contrast of a transparent specimen by using thin strips of soot in the microscope that altered the phase of the light (Bratuscheck, 1892).

Zernike invented the phase contrast microscope as an optical solution to this problem (Saylor et al., 1950; Zernike, 1934a, b, 1935, 1936, 1938, 1942a, b, 1948, 1953a, b, c, 1958).

The phase contrast microscope is an optical device that converts differences in the phases of light transmitted through a transparent object into differences of light intensity in the image. These differences in light intensity are detectable by the human eye and other light detectors on the microscope. Zernike developed phase contrast as a means of testing the quality of telescope mirrors (Ferwerda, 1993). In his experiments he validated his technique in much the same way Abbe validated his experiments. Furthermore, Zernike (unlike Abbe) derived and published a mathematical foundation for the phase contrast technique. Then Zernike applied the phase contrast technique to light microscopy to solve the old problem of imaging transparent specimens for which he was issued a patent (Zernike, 1936). His visit to the Zeiss firm ended in failure and personal humiliation as he described in his Nobel Prize lecture (Zernike, 1953a). But in 1941 two researchers at the Zeiss firm presented extensive experimental studies that validated the utility of Zernike's phase contrast microscope (Köhler and Loos, 1941; Loos, 1941). A historical time-lapse movie 2.5 hours in duration is available on YouTube that shows the use of Zernike's phase contrast microscope to observe the prophase stage during meiosis in spermatocytes of the locust *Psophus stridulus*. This is a historic movie made in the early 1940s by Kurt Michel of the Zeiss firm in Jena using Zernike's newly invented phase contrast microscope (Michel, ca. 1943). In the following decades the phase contrast microscope became the most popular microscope in biological laboratories. There are a couple of variants of the phase contrast microscope: the reflection mode phase contrast microscope and the color phase contrast microscope

(Françon, 1950, 1953, 1961; Saylor et al., 1950). Applications of the phase contrast microscope include cytology, hematology, and mycology, as well as chemistry, mineralogy, crystallography, and the examination of surfaces (Bennett et al., 1951; Françon, 1961).

While credit for the invention of many microscopic techniques is often ambiguous and disputed, historians agree that Zernike invented the phase contrast microscope. A number of optical techniques were invented in the following years to enhance the contrast of weakly absorbing phase objects such as biological specimens (Kubitscheck, 2013; Oldenbourg and Shribak, 2010; Reynolds et al., 1989).

Georges Nomarski is usually credited with the invention of the differential interference contrast (DIC) microscope (Nomarski, 1953, 1955, 1960, 1975; Nomarski and Weil, 1955). However, according to Oldenbourg and Shribak (2010) F. H. Smith preceded him. The DIC technique was invented by Smith in 1947 (Smith, 1947, 1952, 1955, 1956). In the following sections I will elaborate on the contributions of Smith and later Nomarski to the development of the DIC microscope.

In 1975 Robert Hoffman invented an alternative type of microscope to overcome a problem with DIC microscopy called the Hoffman modulation contrast (HMC) microscope (Hoffman and Gross, 1975a, b). Hoffman's patent is titled: "Microscopy systems with rectangular illumination particularly adapted for viewing transparent objects." These three types of optical contrast-enhancing microscopes—phase contrast microscopy, differential contrast microscopy, and Hoffman modulation contrast microscopy—were indispensable to the important development of tissue culture to study the biology of living cells.

12.2 Phase Contrast Microscopy

First, I restate the problem and then offer Zernike's solution. Second, I describe the components of the phase contrast microscope and the propagation of light through it. Third, I describe the manipulation of phases and amplitudes of light transmitted by the background and the specimen. Fourth, I discuss the advantages and limitations of the phase contrast microscope.

The problem we are attempting to resolve is how to image a transparent object. For our discussion we assume that illumination conforms to the requirements of Köhler illumination that provides uniform illumination. We assume that the microscope slide contains a specimen that is a transparent phase object such as a living cell. It is assumed that the object has a different refractive index from the medium in which it is situated. A transparent object does not absorb transmitted light. The light transmitted through the phase object termed direct light does not deviate during propagation. However, light that is deviated by the phase object is retarded (undergoes a phase change or shift) due to the refractive index and thickness of the object. Deviated light is out of phase with direct or undeviated light. At the image plane direct light and deviated light have almost equal

intensities. Their interference in the image plane has a constant intensity. Hence a transparent object has no contrast. Although a transparent object alters the phase of transmitted light, it remains transparent since light detectors, such as the human eye, cannot detect phase changes. Zernike solved this problem by designing and constructing a light microscope with phase contrast optics that was capable of converting phase changes into intensity changes that could be detected.

The optical components of the phase contrast microscope are illustrated in Fig. 12.1. I describe the components from the illumination system at the bottom of the figure to the image plane at the top of the figure. Light from the source is incident on the condenser that contains a diaphragm with an annular ring (shown as a clear region) and a condenser lens. The diaphragm is situated in the front focal plane of the condenser. The phase contrast microscope objective consists of the objective lens and a phase plate that is placed in the back focal plane of the microscope objective. The phase plate contains a ring (shown as a clear region) in which the optical path length is altered in the ring compared with the region inside and outside the ring. The ring can be formed on a thin plate of glass by depositing dielectric material in the shape of a ring or etching dielectric material from the ring; in either case the optical path length is altered in the region of the ring. The shift in the phase of light in the phase plate depends on the wavelength of illumination. The phase plate alters direct light that passes the annular zone in two ways: first, the amplitude is reduced to match the amplitude of the diffracted light, and second, the phase difference is altered from 90 to 180°. The phase shift and amplitude reduction result in a high-contrast phase contrast image formed by interference in the image plane of the microscope.

As Zernike noted, the phase contrast microscope images thin transparent objects with increased darkness with a bright background which is similar to absorbing objects in a bright field microscope. If the phase plate retards the phase of direct light by a quarter of a wavelength relative to the diffracted light, then the thicker or higher refractive index particle will appear brighter on a dark background, which is the case with negative phase contrast microscopy. Conversely, if the phase plate advances the phase of direct light by a quarter of a wavelength relative to the diffracted light, then thicker or higher refractive index particles will appear darker on a bright background, which is the case with positive phase contrast microscopy.

There are critical constraints to the design of the illumination system of the microscope: undiffracted light (*dashed arrows* in Fig. 12.1) propagates through the region of the ring in the phase plate and diffracted light (*solid arrows* in Fig. 12.1) propagates through the phase plate in regions outside the ring. Undiffracted light (zero order) and diffracted light (nonzero order) are collected by the microscope objective and pass through the phase plate resulting in the object being detected at the image plane.

In 1935 Zernike was issued a patent for his invention of the phase contrast microscope (Zernike, 1936). In 1942 Zernike published two papers that explain the principles, instrumentation, and applications of the phase contrast microscope in

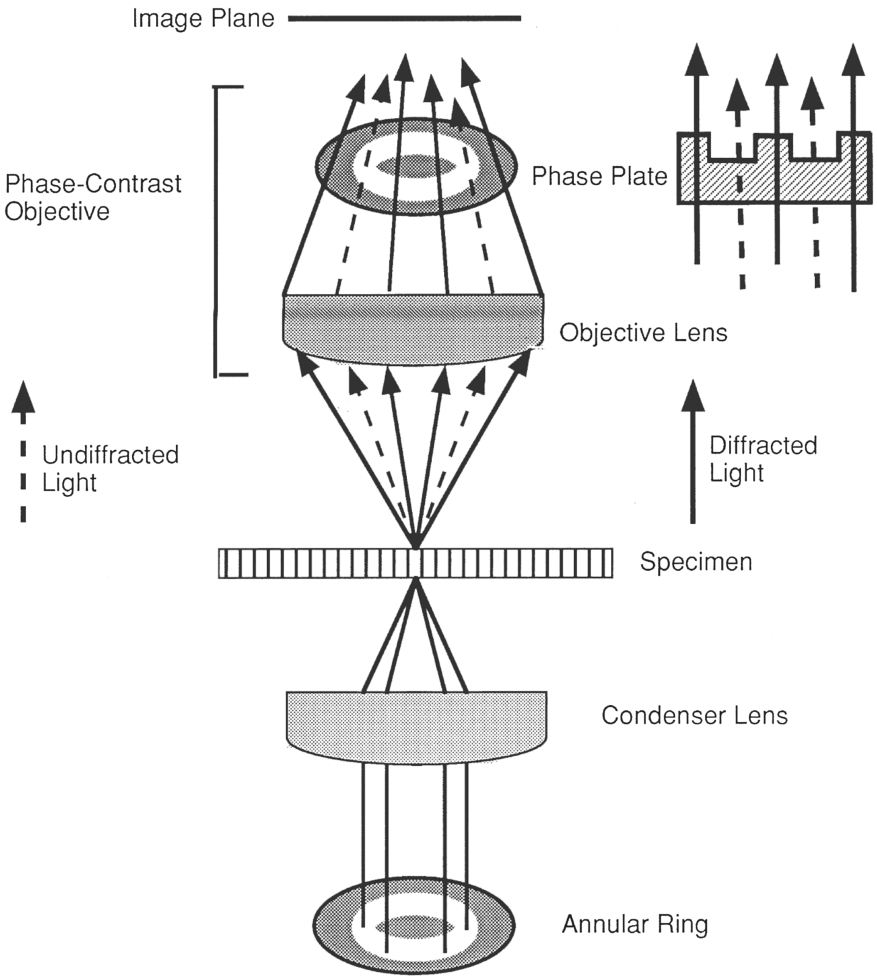


Fig. 12.1 Schematic diagram showing the optical components of a phase contrast microscope. Diffracted light from the specimen is shown as *solid arrows*; undiffracted light is shown as *dashed arrows*. From Chap. 5, Fig. 5.1, p. 59, Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press

terms of physical optics (Zernike, 1942a, b). The principle behind the phase contrast microscope is the spatial separation of direct zero-order light from diffracted light at the back focal plane of the microscope objective. Zernike understood that diffracted light is retarded in phase by 90° compared with direct, zero-order light. Zernike understood that for diffracted and direct light beams to interfere in the image plane (formation of an image with contrast) it is necessary to retard the

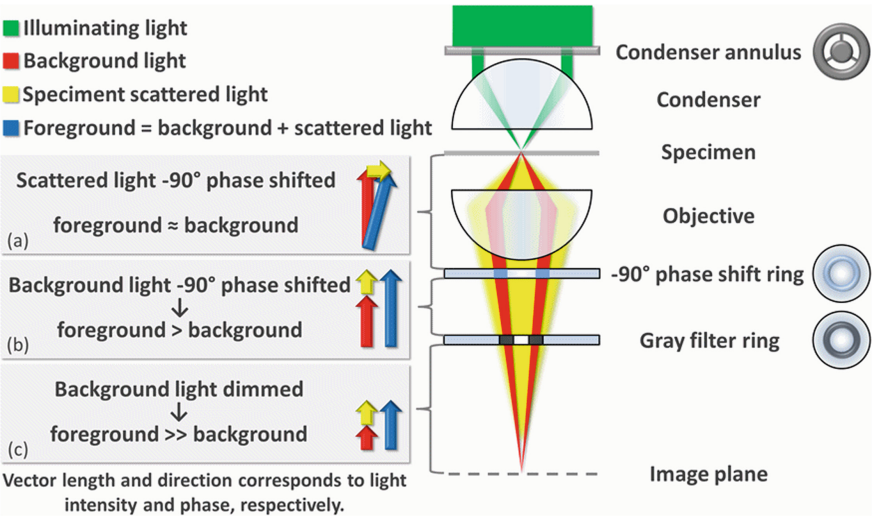


Fig. 12.2 The basic principle that makes phase changes visible in phase contrast microscopy. By Egelberg [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons. This file is licensed under the Creative Commons. Wiki Phase contrast microscopy

central region of the incident light beam in the rear focal plane of the objective by one quarter of a wavelength.

Figure 12.2 should help the reader to understand the basic principle that makes phase changes visible to the human eye and other light detectors using phase contrast microscopy. The diagram shows that image contrast in enhanced by constructive interference between diffracted light and direct or background light rays from the specimen and by reducing the amount of background light that propagates to the image plane. Background light in Fig. 12.2 is attenuated by a separate optical component

What are the advantages and the limitations of Zernike's phase-contrast microscopy. The main advantage of Zernike's phase contrast microscope is that it converts transparent objects into high-contrast images. Because the technique is insensitive to polarization and birefringence effects it is widely used in cell culture where cells grown in plastic containers can also be directly observed by the phase contrast microscope. Phase contrast microscopes are best suited for thin objects.

However, there are significant limitations to the phase contrast microscope. Image resolution is less than optimal because the full numerical aperture of the microscope is not used. In normal use the phase contrast microscope forms a bright "halo" around cell borders in positive phase contrast or a shadow in negative phase contrast. This confounds visualization of cell edges. This artifact is due to incomplete separation of direct light and diffracted light transmitted through the ring region of the phase plate. Furthermore, there is ambiguity in the interpretation of phase contrast images because intensities in the image may not directly correlate

with structures in the object. The phase contrast microscope is ideally suited for nonquantitative imaging of thin transparent objects such as thin sections of tissue or a thin layer of cells.

Many biological objects are not completely phase objects or absorption objects; there is a mixture of phase and amplitude information in the microscope (Streibl, 1985). The Zernike phase contrast microscope ideally works when there are very small phase shifts and almost no absorption. In 2002 a new type of microscope called the quantitative phase amplitude microscope was invented to circumvent the problem of mixed phase and absorption in the phase contrast microscope (Barone-Nugent et al., 2002). Their microscope separates amplitude and phase information and forms images that are linear in amplitude and phase.

12.3 Differential Interference Contrast Microscopy

The differential interference contrast (DIC) microscope was invented by Smith in 1947 (Smith, 1947, 1952, 1955, 1956; see also Oldenbourg and Shribak, 2010). Nomarski developed his own version of it called the Nomarski differential interference contrast microscope between 1953 and 1955 (Nomarski, 1953, 1955, 1960, 1975; Nomarski and Weil, 1955; see also Allen et al., 1969).

Zernike's phase contrast microscope produces images in which intensity is proportional to optical path differences. Images formed in the DIC microscope are based on the rate of change of the optical path of the object along the shear direction, and intensity is proportional to the first derivative of the difference in the optical path.

Before I describe the optical components that make up Nomarski's differential interference contrast microscope and its operation it is necessary to explain birefringence, Wollaston prism, and Nomarski's modification of the Wollaston prism. Finally, I will discuss the advantages and limitations of the DIC microscope and compare the phase contrast microscope with the DIC microscope.

A good graphical introduction to the understanding of birefringence is found in Hecht's *Optics* (Hecht, 2017). The optical properties of crystals are dependent on the symmetry properties of crystals. Crystals such as quartz or calcium carbonite (calcite) that are birefringent materials produce double images. Birefringence is the double refraction of light due to orientation-dependent differences in the refractive index of the crystal. When incident light that is linearly polarized interacts with a birefringent material the light is split (sheared or separated) into two separate components that have perpendicular planes of polarization (Hecht, 2017). The two components propagate over different paths in the crystal and exit the crystal as two rays of light that are linearly polarized. The electric field vectors of these two rays vibrate in perpendicular planes.

A Wollaston prism can be constructed from two birefringent quartz or calcite prisms that are cemented together at their base. The conjoined prisms form two right triangular prisms with their optical axes in the perpendicular direction.

A Wollaston prism will cause incident light to form two perpendicularly polarized diverging beams. The diverging beams begin to diverge at the interface of the two prisms that form the Wollaston prism.

A Nomarski prism was invented by Nomarski and is a modification of Wollaston's prism for use in the Nomarski differential contrast microscope (Nomarski, 1953, 1955, 1960; Nomarski and Weil, 1955). It consists of two birefringent quartz or calcite prisms that are cemented together at their base. But, in this case the two prisms are different. One prism is similar to the ones used to form the Wollaston prism and its optical axis is parallel to the prism surface. But the second prism is made with its optical axis at an oblique angle to the surface of the prism. Therefore, the optical axis of one prism is oriented obliquely to the optical axis of the second prism. In the Nomarski prism the incident light beam enters the prism and is transformed into two light beams each orthogonally polarized. The unique feature of the Nomarski prism is that the two emerging beams each orthogonally polarized converge at a focal point that is situated outside the prism. The focal point of the two emerging beams is called the interference plane. The Nomarski prism is an important modification because it permits the interference plane to be located at the back focal plane or the diffraction plane of the microscope objective.

The optical components of the Nomarski differential contrast microscope are illustrated in Fig. 12.3. I describe the propagation of light from the source at the bottom of the figure to the image plane at the top of the figure. The optical components of a DIC microscope include a linear polarizer located between the light source and the condenser, a modified Wollaston prism (Nomarski prism) located close to the iris in the back focal plane of the condenser, a Nomarski prism located in the back focal plane of the microscope objective, and a linear polarizer or analyzer in front of the image plane.

Light paths within the DIC microscope are shown in Fig. 12.4. I describe light paths from the left to the right side of the figure. The illumination system is configured for Köhler illumination. Light from the source is unpolarized and a polarizing filter transforms it to a beam of polarized light that is incident to the Wollaston (Nomarski) prism yielding two beams each orthogonally polarized. The distance between the two beams is called the shear. Light from the condenser lens is incident on the object on the microscope stage. Light transmitted by the object is phase-shifted relative to light transmitted by the background that is not phase-shifted. The object's different thicknesses, refractive indexes, and the rate of change of these quantities with distance (in the direction of shear) in the object modify the two beams. The two light beams each orthogonally polarized enter the aperture of the microscope objective and focus in the back focal plane of the microscope objective. The second Nomarski prism located in the back focal plane of the microscope objective recombines the incident beams that are orthogonally polarized into one beam of polarized light. The upper Nomarski beam-combining prism is movable and can be used to compensate for phase shifts in the object. A polarizing filter removes directly transmitted light. When the illumination is white light, the rate of change of optical path differences within the object is observed in the eyepiece as differences in intensity and color.

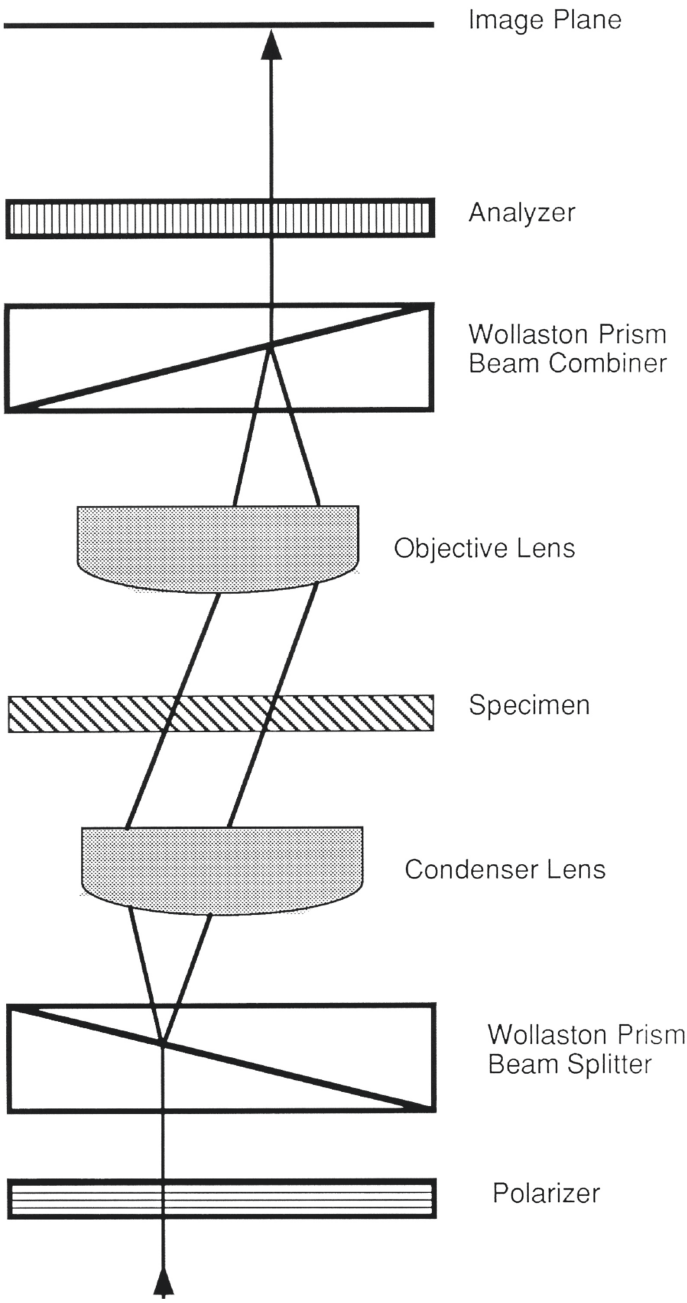


Fig. 12.3 Schematic diagram illustrating the principle behind a differential interference contrast (DIC) microscope. From Chap. 5, Fig. 5.2, p. 61, Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press

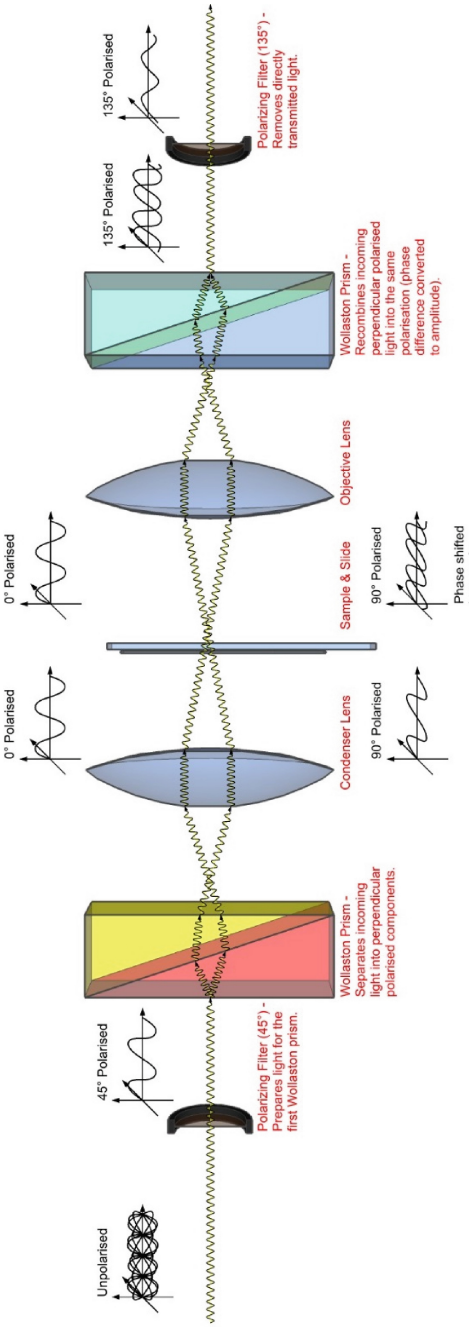


Fig. 12.4 Differential interference contrast (DIC) light paths. From Wikimedia Commons, Richard Wheeler (Zephyris)—en: Image: DIC Light Path.png

The microscopes schematically shown in Figs. 12.3 and 12.4 are transmission light microscopes. The DIC microscope can also be constructed to work as an epi-illumination DIC microscope. In that case a single Nomarski–Wollaston prism is required together with a beamsplitter (Hartman et al., 1980; Lessor et al., 1979).

How does DIC microscopy compare with Zernike's phase contrast microscopy? First, image formation is different. DIC microscopy is based on the rate of change in optical path difference between the object and the background in the direction of shear. Phase contrast microscopy is based on the absolute magnitude of optical path difference. Phase contrast microscopy is optimal for thin objects.

DIC microscopy uses the whole numerical aperture of the microscope objective and yields images with enhanced axial and lateral resolution. The DIC image is devoid of the ubiquitous bright diffraction "halo" around the edges of objects that are seen with phase contrast microscopy. The DIC image looks like a three-dimensional object illuminated with oblique light. The three-dimensional appearance enhances linear structures and edges in the object. Furthermore, the DIC microscope minimizes out-of-focus light in image formation, therefore it is optimized for imaging thick objects due to its enhanced axial resolution. The DIC microscope is an "optically sectioning" microscope that excludes light outside the focal plane. In this respect the DIC microscope is similar to the confocal microscope (Masters, 1996, 2006).

There are a number of limitations to DIC microscopes. They are more expensive than phase contrast microscopes; they cannot be used with plastic Petri dishes because plastic is birefringent; and apochromatic microscope objectives that affect polarized light cannot be used in DIC microscopy. The problem of image interpretation is similar to warnings given about phase contrast microscopy: image interpretation is difficult and sometimes misleading. The DIC microscope does not produce topographically accurate images.

12.4 Hoffman Modulation Contrast Microscopy

Robert Hoffman and Leo Gross invented the Hoffman modulation contrast microscope (HMC) to enhance contrast in unstained transparent biological specimens (Hoffman and Gross, 1970, 1975a, b). The HMC microscope converts phase gradients into changes in intensity and thus generates contrast from transparent phase objects. Light intensity in the image is proportional to the first derivative of the optical density of the object. The HMC microscope makes use of oblique illumination, which as Abbe showed results in resolution being doubled. The word "modulation" used in the Hoffman modulation contrast microscope refers to the fact that light intensities in the image are modulated above and below an average gray value.

The basic generic components of the HMC microscope are shown in Fig. 12.5. I now provide an explanation of the components and their functions as described in their papers (Hoffman and Gross, 1975a, b). Recent versions of the HMC

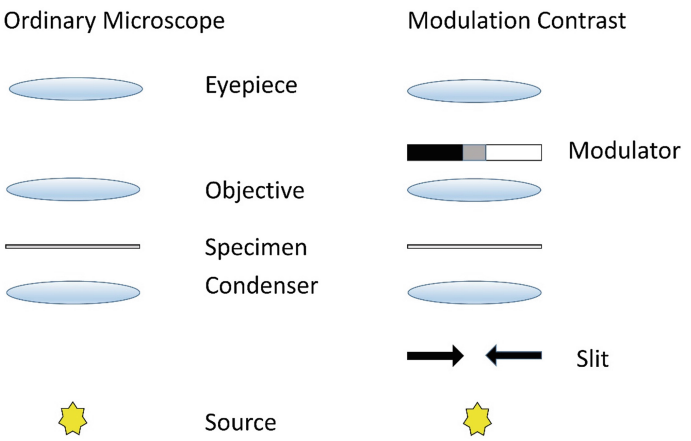


Fig. 12.5 Light paths in the Hoffman modulation contrast microscope

microscope have the slit aperture and modulator displaced from the optical axis, which permits more of the numerical aperture of the microscope objective to be utilized and results in higher resolution. The slit aperture used to form oblique illumination is located in the front focal plane of the condenser lens. The modulator is placed in the back focal plane of the microscope objective; the back focal plane is conjugated with the front focal plane of the condenser that contains the slit aperture. Therefore, both the slit aperture and the modulator are in conjugate planes. The modulator contains three regions where light transmittance differs: the region on the left in Fig. 12.5 has 1% transmittance, the middle region has 15% transmittance, and the region on the right side of the modulator has 100% transmittance. The modulator does not affect the phase of the light transmitted. Illumination light transmitted through the object and the structure of the object result in different degrees of refraction. This determines which light rays are incident on the three different regions of the modulator in the back focal plane of the microscope objective. This enhances image contrast.

During operation of an HMC microscope the image of the slit aperture is positioned in the region of the modulator that has 15% transmittance; adjacent to this region on one side is a region of low transmittance and on the other side a region of maximum transmittance. The modulator selects opposite gradients to produce opposite intensities creating an optical shadowing effect. The two opposite sides of an object imaged with an HMC microscope appear to be darker and brighter, respectively, than the background. An HMC microscope can produce "optical sections" of the object in which light from above and below the focal plane is excluded from the image of the object. The modulator only affects zero-order beams; higher order diffracted beams from the object are not affected in an HMC microscope. Transparent phase objects have enhanced contrast due to differential absorption of refracted light beams. The image observed in an HMC microscope

has an intensity that is proportional to the phase gradient; it appears as a three-dimensional structure that is very similar to images obtained with a differential contrast microscope. Image contrast can also be enhanced by using polarizing filters.

How do HMC microscopes compare with phase contrast microscopes and differential contrast microscopes. The HMC microscope differs from the phase contrast microscope in that oblique illumination forms images with enhanced resolution; there is "optical sectioning" of the object; there is a wide field of view; the image does not show the "halos" seen in phase contrast microscopes around the edges of cells; and there is a significant increase in contrast at boundaries and gradients in the object. HMC microscopes have an important advantage over differential contrast microscopes: they can be used with birefringent objects such as plastic tissue culture dishes. In addition, HMC microscopes are less expensive than DIC microscopes. HMC microscopes have wide utility in the biological laboratory and the clinic where they are indispensable to in vivo fertilization. HMC microscopes provide high-contrast images of thick, transparent, spherical objects such as oocytes. I have restricted my discussion to HMC microscopes that operate in transmission mode. However, when HMC microscopes need to image opaque objects, such as metal surfaces or integrated circuits, the reflected light HMC microscope is the only option.

The HMC microscope has a couple of limitations. One concerns the problem of deciding whether the image contains elevated or depressed structures with respect to a reference surface. The other concerns image interpretation, a limitation that persists in the phase contrast microscope too.

References

Allen, R. D., David, G. B., and Nomarski, G. (1969). The Zeiss-Nomarski differential interference equipment for transmitted-light microscopy. Zeitschrift für Wissenschaftliche Mikroskopie und Mikroskopische Technik, **69**, 193–221.

Barone-Nugent, E. D., Barty, A., and Nugent, K. A. (2002). Quantitative phase-amplitude microscopy I: optical microscopy. Journal of Microscopy, **206**, 194–203.

Bennett, A. H., Jupnik, H., Osterberg, H., Richards, O. W. (1951). *Phase Microscopy, Principles and Applications*. New York: John Wiley & Sons, Inc., pp. 3–4.

Bratuscheck, K. (1892). Die Lichtstärke-Änderungen nach verschiedenen Schwingungsrichtungen in Linsensystemen von grössen Öfnungswinkel mit Beziehung zur mikroskopischen Abbildung. [The light intensity changes according to different directions of vibration in lens systems of great angular aperture with relation to microscopic image formation] Ztschr. f. wiss. Mikr., **9**, 145–160.

Ferwerda, H. A. (1993). Frits Zernike-life and achievements. Optical Engineering, **32**, 3176–3181.

Françon, M. (1950). *Le contraste de phase en optique et en microscopie*. Editions de la Rev. d'Optique theorique et Instrumental, Paris, 1–109.

Françon, M. (1953). *Le microscope à contraste de phase et le microscope interférentiel*. Paris: Centre National de la Recherche Scientifique.

Françon, M. (1961). *Progress in Microscopy*. Evanston, Illinois: Row, Peterson and Company.

Hartman, J. S., Gordon, R. L., and Lessor, D. L. (1980). Quantitative surface topology determination by Nomarski reflection microscopy. II. Microscope modification, calibration, and planar sample experiments. Applied Optics, **19**, 2998–3009.

Hecht, E. (2017). *Optics*, fifth edition. Boston: Pearson Education, Inc.

Hoffman, R., and Gross, L. (1970). Reflected-light differential-interference microscopy: principles, use and image interpretation. Journal of Microscopy, **91**, 149–172.

Hoffman, R., and Gross, L. (1975a). The modulation contrast microscope. Nature, **254**, 586–588.

Hoffman, R., and Gross, L. (1975b). Modulation Contrast Microscope. Applied Optics, **14**, 1169–1176.

Hooke, R. (1665). Of a New Property in the Air. *Micrographia*, Observation LVIII, 217–219, London.

Kubitscheck, U. (2013). *Fluorescence Microscopy, From Principles to Biological Applications*. Weinheim, Germany: Wiley-Blackwell, 81–95.

Köhler, A., and Loos, W. (1941). Das Phasenkontrastverfahren und seine Anwendungen in der Mikroskopie. Naturwiss. **29**, 49–61.

Lessor, D. L., Hartman, J. S., and Gordon, R. L., and. (1979). Quantitative surface topology determination by Nomarski reflection microscopy. I. Theory. Journal of the Optical Society of America, **69**, 357–366.

Lipson, A., Lipson, S. G., Lipson, H. (2011). *Optical Physics*, 4th Edition. Cambridge, UK: Cambridge University Press.

Loos. W. (1941). Das Phasenkontrastverfahren nach Zernike als biologisches Forschungsmittel. Wiener klinische Wochenschrift, **20**, 849–853.

Malacara, D. (2007). *Optical Shop Testing*, Third Edition. New Jersey: John Wiley & Sons, Inc.

Masters, B. R. (1996). *Selected Papers o Confocal Microscopy*. Bellingham, WA: SPIE Optical Engineering Press.

Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press.

Michel, K. (ca. 1943). Historic time lapse movie, from 2.5 hours in real-time, demonstrating meiosis in spermatocytes of the locust *Psophus stridulus* viewed with the Zernike Phase-Contrast Microscope. Zeiss Microscopy, Jena. YouTube, June 1, 2013. https://www.youtube.com/watch?v=Ge4k3uiB3qw (Accessed April 3, 2019).

Nomarski, G. (1953). Dispositif interférentiel à polarisation pour l'étude des objects transparents ou opaques appartenant à la classe des objects de phase, French patent, No. 1.059.124.

Nomarski, G. (1955). Microinterféromètre differential à ondes polarisées. Journal de Physique et Le Radium, **16**, 9S-11S.

Nomarski, G., (1960). Interferential Polarizing Device for Study of Phase Objects. United States Patent, No. 2,924,142. Centre National de la Recherche Scientifique, Paris, France.

Nomarski, G. (1975). Optical differentiation using oblique coherent illumination and pure amplitude spatial filter. Einladung und Programm zur 76. Tagund der GDaO, 20 bis 24, May 1975, in Bad Ischl (Austria) p. 80.

Nomarski, G., and Weil, A. R. (1955). Application à la métallographie des méthodes interférentielles à deux ondes polarisées. Revue de Métallurgie, **52**, 121–134.

Oldenbourg, R., and Shribak, M. (2010). Microscopes. In: *Handbook of Optics*, Third Edition. Volume I, Geometrical and Physical Optics, Polarized Light, Components and Instruments, Bass, M., and Mahajan, V. N., editors, New York: McGraw-Hill, 28–39.

Pluta, M. (1988). *Advanced Light Microscopy*. Volume 1, Principles and Basic Properties [Image formation in the microscope within the scope of diffraction theory and Fourier optics]. Amsterdam: Elsevier.

Pluta, M. (1989). *Advanced Light Microscopy*. Volume 2, Specialized Methods [Phase contrast microscopy, Differential interference contrast]. Amsterdam: Elsevier.

Pluta, M. (1993). *Advanced Light Microscopy*. Volume 3, Measuring Techniques [Quantitative polarizing microscopy, Microinterferometry, Measurements of the refractive index and/or thickness of microobjects. Amsterdam: Elsevier.

Popescu, G. (2011). *Quantitative Phase Imaging of Cells and Tissues*. New York: McGraw-Hill Education.
Reynolds, G. O., DeVelis, J. B., Parrent Jr., G. B., and Thompson, B. J. (1989). *The New Physical Optics Notebook: Tutorials in Fourier Optics*. Bellingham: SPIE Optical Engineering Press.
Saylor, C. P., Arthur T. Brice, A. T., and Zernike, F. (1950). Color phase-contrast microscopy: requirements and applications. Journal of the Optical Society of America, **40**, 329–334
Smith, F. H. (1947). Interference microscope. U.S. patent 2,601,175 (August 5, 1947).
Smith, F. H. (1952). Interference Microscope. United States Patent, Number 2,601,175.
Smith, F. H. (1955). Microscopic Interferometry. Research (London), **8**, 835–395.
Smith, F. H. (1956). *Microscopic interferometry. Modern Methods of Microscopy*. London: Butterworths, p. 76.
Streibl, N. (1985). Three-dimensional imaging by a microscope. Journal of the Optical Society of America A, **2**, 121–127.
Zernike, F. (1934a). Beugungstheorie des Schneidenverfahrens und Seiner Verbesserten Form, der Phasenkontrastmethode. [Diffraction theory of the cutting process and its Improved form of the phase contrast method] Physica, **1**, 689–704.
Zernike, F. (1934b). Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. Monthly Notices of the Royal Astronomical Society, **94**, 377–384.
Zernike, F. (1935). Das Phasenkontrastverfahren bei der mikroskopischen Beobactung. Z. Tech. Phys., **16**, 454-457.
Zernike, F. (1936). Deutsches Reichspatent No. 636168 (September 1936).
Zernike, F. (1938). The concepts of degree of coherence and its application to optical problems. Physica, **5**, 785–795.
Zernike, F. (1942a). Phase contrast, a new method for the microscopic observation of transparent objects, Part 1. Physica, **9**, 686–698.
Zernike, F. (1942b). Phase contrast, a new method for the microscopic observation of transparent objects, Part 2. Physica, **9**, 974–986.
Zernike, F. (1948). Diffraction and optical image formation. Proceedings of the Physical Society, **61**, 158–164.
Zernike, F. (1953a). Nobel Lecture: How I Discovered Phase Contrast. December 11, 1953. https://www.nobelprize.org/uploads/2018/06/zernike-lecture.pdf (Accessed on April 2, 2019).
Zernike, F. (1953b). Frits Zernike - Biographical. Nobelprize.org. https://www.nobelprize.org/prizes/physics/1953/zernike/biographical/ (Accessed April 2, 2019).
Zernike, F. (1953c). NobelPrize.org Documentary. https://www.nobelprize.org/prizes/physics/1953/zernike/documentary/ (Accessed April 2, 2019).
Zernike, F. (1958). The wave theory of microscopic image formation. Appendix K. In: *Concepts of Classical Optics*, John Strong, pp. 525–536. San Francisco: W. H. Freeman. Zernike, F. (2014).

Further Reading

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.
Abbe, E. (1989). *Gesammelte Abhandlungen, I-IV*. Hildesheim: Georg Olms Verlag. [Originally published in 1904, Jena: Verlag von Gustav Fischer].
Berek, M. (1929). XXI. On the extent to which real image formation can be obtained in the microscope. Journal of the Royal Microscopical Society, **49**, 240–249.
Beyer, H. (1965). *Theorie und Praxis des Phasenkontrastverfahrens*. Frankfurt a.M, Leipzig: Akademische Verlagsgesellschaft Geest und Portig KG.
Born, M., and Wolf, E. (1999). *Principles of Optics*, 7th (expanded) edition. Cambridge: Cambridge University Press.

Cogswell, C. J., and Sheppard, C. J. R. (1992). Confocal differential interference contrast (DIC) microscopy including the theoretical analysis of conventional and confocal DIC imaging. Journal of Microscopy, **165**, 81–101.

Conrady, A. E. (1905). An experimental proof of phase reversal in diffraction spectra. Journal Royal Microscopical Society, **25**, 150–152.

Duffieux, P. M. (1970). *L'intégrale de Fourier et ses applications à l'optique*. Paris: Masson, Editeur. Republished in English as: Duffieux, P. M. (1970). *The Fourier Transform and its applications to optics*, second edition. New York: John Wiley and Sons.

Hoffman, R. (1977). The modulation contrast microscope: principles and performance. Journal of Microscopy, **110**, 205–222.

Hopkins, H. H. (1950). The influence of the condenser on microscopic resolution. Proceedings of the Physical Society (B), **63**, 737–744.

Hopkins, H. H. (1951). The concept of partial coherence in optics. Proceedings of the Royal Society of London Series A, **208**, 263–277.

Hopkins, H. H. (1953). On the diffraction theory of optical images. Proceedings of the Royal Society of London Series A, **217**, 408–432.

Köhler, H. (1981). On Abbe's Theory of Image Formation in the Microscope. Journal of Modern Optics, **28**, 1691–1701.

Martin, L. C. (1966). *The Theory of the Microscope*. New York: American Elsevier Publishing Company, Inc. and London: Blackie.

McMahon, P. J., Barone-Nugent, E. D., Allman, B. E., and Nugent, K. A. (2002). Quantitative phase-amplitude microscopy II: differential interference contrast imaging for biological TEM. Journal of Microscopy, **206**, 204–208.

Michel, K. (1964). *Die Grundzüge der Theorie des Mikroskops in elementarer Darstellung*. 2. neubearbeitete Auflage. Stuttgart: Wissenschaftliche Verlagsgesellschaft M.B. H.

Nienhuis, K. (1948). *On the influence of diffraction on image formation in the presence of aberrations*. Thesis, Groningen.

Nijboer, B. R. A. (1942). *The diffraction theory of aberrations*. Thesis, Groningen.

Porter, A. B. (1906). On the diffraction theory of microscope vision. Philosophical Magazine, **6**, 154–166.

Rheinberg, J. (1904). On the influence of images of gratings of phase difference amongst their spectra. Journal of Microscopy, **24**, 388–390.

Singer, W., Totzeck, M., and Gross, H. (2005). *Handbook of Optical Systems, Physical Image Formation* (Volume 2). Weinheim: Wiley-VCH.

Zernike, F., and Brinkman, H. C. (1935). Hypersphärische Funktionen und die in sphärischen Bereichen orthogonalen Polynome. Verh. Akad. Wet. Amst., (Proceedings Royal Acad. Amsterdam), **38**, 161–170.

Part III
Far-Field Superresolution Optical Microscopy

Introduction

Part III contains a comprehensive analysis of far-field superresolution optical microscopes. I divided these microscopes into three major classes and devoted a separate chapter to each class. Chapter 13, "Structured Illumination Microscopy," develops the theory of moiré interference patterns and Fourier optics as the foundations of linear and nonlinear SIM. Chapter 14, "Stimulated Emission Depletion Microscopy and Related Techniques," begins with an introduction to molecular spectroscopy. Next, Einstein's theory of stimulated emission is fully developed. STED microscopy uses a spiral waveplate to form a STED beam with an annular intensity profile that has zero intensity at the center. I present a detailed discussion of the fundamental physical theory that explains the operation of spiral waveplates. This discussion is centered on the theory of optical vortices. Chapter 15, "Localization Microscopy with Active Control," begins with a discussion of localization versus resolution. The theory and practice of single-molecule detection is then developed. I introduce the Nyquist and Shannon theorem as it applies to single-molecule detection and localization. A discussion of the key publications of far-field superresolution stochastic optical microscopy that stresses the photochemistry of photoactivatable and/or photoswitchable fluorescent probes with emphasis on the advantages and limitations of each type of microscope is provided. Part III ends with a coda, Chapter 16, that compares the various types of far-field superresolution optical microscopes that were described in Part III. The emphasis is on trade-offs, necessary cautions, validation and interpretation of images, and the effects of light on the specimen. Again, the emphasis is on posing critical questions.

When authors publish their work they carefully describe previous related works; the theoretical background; details about instrumentation, data acquisition, and analysis; specimen preparation; and conclusions. What is missing is the roadmap leading to their discovery. What were the key influences that brought about their success? Which colleagues pointed them in the correct direction? What wrong turns were

taken? Did the current funding situation encourage or discourage innovation and creative exploration? What hints or prior knowledge did they subconsciously learn in their roles as members of study sections, or in the peer review process, that perhaps influenced their pathway to discovery? This last important question brings the reader back to the topic of ethics in science and the critical topic of responsible conduct of research. These are the topics that I teach in my university courses. I have made my course presentation of "Responsible Conduct of Research" available at https://storage.googleapis.com/springer-extras/zip/2020/978-3-030-21690-0.zip.

Another topic of interest is how credit is attributed to researchers for their discoveries. Does the credit go to the first person to publish, the first person to file for a patent, or the first person to have the patent granted? Studies show that there is a trend toward increasing the number of authors contributing to single publications. Research advances in which progress is little more than incremental are sometimes published. Investigators who spend the time to perform due diligence on controls and validation of the results and their conclusions are under extreme competitive pressures to publish before their colleagues scoop them. This intense competition for priority and the subsequent fame, honor, and funding that accompanies it is the new norm. However, this situation has its disadvantages, one of which is the increasing number of publications retracted from key journals.

We also know of examples of manuscripts rejected from journals because they were ahead of their times. They were outside the mindset of reviewers, editors, and grant-funding agencies. On the other hand, there are many examples of publications that appeared in high-impact scientific journals only to be later shown to be fraudulent and, hence, subsequently retracted by the journals. The historical record also details the foibles of the Nobel Prize Committee in the singular awarding of the Nobel Prize to C. V. Raman for the Raman effect. The historical record shows that the Nobel Committee misinterpreted the letters of support for the Russian physicists Landsberg and Mandelstam who had simultaneously discovered the same effect as Raman (Masters, 2009).

We are fortunate to have Eric Betzig's Nobel Lecture available. Titled *Single Molecules, Cells, and Super-Resolution Optics* (Betzig, 2014a, b, c) the lecture material comprises a video, lecture slides, and the text. Betzig's Nobel Lecture is remarkable in several aspects. First, he clearly and honestly draws a roadmap showing all the detours, frustrations, false starts, missing links, and confounding aspects of peer review and grant-funding agencies. Second, he clearly and honestly credits colleagues, teachers, friends, mentors, and students who contributed in large or small ways to his successes. Third, he clearly and honestly points out the limitations of all superresolution far-field optical microscopy techniques when applied to biological specimens and specifically to live cell and living organisms. Betzig ends his Nobel Lecture with a plea to researcher to take risks when they do science. While that may at first seem to be at odds with the severe pressures of grant-funding agencies and the competition to be the first to publish, risk taking is at the core of curiosity-based research.

The advances in instrumentation and fluorescent probes described in Part III highlight the interdisciplinary nature of scientific research. The history of the microscope is replete with examples of invention, reinvention, and simultaneous invention. Those intrepid investigators who ventured into disparate fields of physics, chemistry, and biology led the way to seminal discoveries. A good example of interdisciplinary research is *Selected Papers on Resolution Enhancement Techniques in Optical Lithography* (Schellenberg, 2004).

Consistent with the title of this book—*Superresolution Optical Microscopy: The Quest for Enhanced Resolution and Contrast*—for each type of superresolution optical microscopy there are antecedent theoretical and experimental advances, the discovery or independent discoveries of new invention(s), and the subsequent progression of new advances, new developments, new probes, and new applications. I place subsequent achievements that follow the initial invention(s) in temporal order at the same time as I discuss them and point out the advances as well as the cautions and limitations of each new superresolution far-field optical microscope. With the publication of numerous review articles by the authors of various research groups I try to focus of what I consider the most important points and emphasize trade-offs, cautions, and limitations.

To be consistent with the plan of this book, as outlined in the Preface and in Parts I and II, I frame innovative advances by stating the problem or the question to be investigated and the technological breakthrough that provided a solution to or overcame a limitation of previous instruments. As stressed in the second part of the title of this book, *The Quest for Enhanced Resolution and Contrast* is developed in all the chapters. Connections between distant researchers tackling similar problems and key publications in disparate fields of study are discussed and highlighted.

Throughout Part III of this book I discuss publications on far-field superresolution microscopy that I believe are the most appropriate. Since my aim is not to write a review article I selected articles that I believed were the first publications of a new invention or technique in the areas of instrumentation, probe development and application, specimen preparation, image interpretation and validation, and control experiments. The pedagogical content of each publication was also an important criterion for selection. Although I have read several hundred publications, I used my judgment to select a small subset from the total corpus of published works. For a complete search of the publications on far-field superresolution microscopy I refer the reader to published review articles, many of which are cited in the references of each chapter. The quest is a work in progress as is the case with all science.

References

Betzig, E. (2014a). *Single molecules, cells, and super-resolution optics: Lecture*. https://www.nobelprize.org/mediaplayer/?id=2407. Accessed 2 April 2019.

Betzig, E. (2014b). *Single molecules, cells, and super-resolution optics: Lecture*. https://assets.nobelprize.org/uploads/2018/06/betzig-lecture.pdf. Accessed 2 April 2019.

Betzig, E. (2014c). *Single molecules, cells, and super-resolution optics: Nobel lecture*. https://assets.nobelprize.org/uploads/2018/06/betzig-lecture-slides.pdf. Accessed 2 April 2019.

Masters, B. R. (March 2009). The Raman effect, its discovery and attribution of credit. *Optics and Photonics News*, *20*, 40–45.

Schellenberg, F. M. (2004). *Selected papers on resolution enhancement techniques in optical lithography. SPIE Milestone Series Volume MS 178*. Bellingham: SPIE Press.

Further Reading

van de Linde, S., Heilemann, M., & Sauer, M. (2012). Live-cell super-resolution imaging with synthetic fluorophores. *Annual Review of Physical Chemistry*, **63**, 519–540.

Chapter 13
Structured Illumination Microscopy

"I like to think in frequency space, rather than in real space."
—Mats G. L. Gustafsson (1960–2011).

13.1 Introduction

Structured illumination microscopy (SIM) is an optical technique that has the capability of enhancing the lateral and axial resolution of a fluorescence widefield microscope. Structured illumination refers to the periodic patterns of light intensity such as dark and light regions of intensity that are formed on the specimen by projecting a grating onto the specimen or by forming a periodic pattern of light intensities by the interference of two or more laser beams.

Structured illumination is spatially modulated light that can be represented by a series of bright and dark stripes. Uniform intensity light is passed through a grating or a Ronchi ruling consisting of alternating opaque and transparent stripes of equal width, and the resulting light is spatially structured into a series of dark and bright stripes. For a given orientation on the specimen the pattern is successively displaced and images that correspond to each displacement are detected. This procedure is repeated for multiple orientations, and the final image is derived computationally. As we have seen in previous chapters the finite aperture of the microscope objective sets a limit to the highest spatial frequencies that can be collected by the microscope objective and that limits the resolution of the microscope.

The discussion in this chapter requires an understanding of Fourier optics and its associated terms. In conventional position space or real space we are familiar with Cartesian coordinate space that is composed of x, y, and z coordinates in orthogonal directions. The first term to understand is spatial frequency. Spatial frequency measures how often the periodic components of a two-dimensional image repeat per unit of distance. The unit of spatial frequency is cycles per millimeter. An image

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_13

with fine details would have high spatial frequencies, and an image with coarse details would have low spatial frequencies.

The next term to introduce is reciprocal space (a.k.a. spatial frequency space or Fourier space). In the literature on SIM many explanations contain graphical representations of spatial resolution in reciprocal space. The Fourier transform of a two-dimensional image in real space yields a series of points in reciprocal space, where each point corresponds to a spatial frequency. Instead of the Cartesian coordinates of x, y, and z of real space in frequency space we have three orthogonal directions that we label k_x, k_y, k_z, where k is called the wavenumber, which is the spatial frequency in units of cycles per millimeter or in radians per unit of distance. The angular wavenumber is given as

$$k = \frac{2\pi}{\lambda} \tag{13.1}$$

where λ is the wavelength.

In reciprocal space low-frequency information (from coarse-image details) is located near the origin and high-frequency information (from fine-image details) is located at greater distances from the origin. The Fourier transform of a two-dimensional image in real space yields the frequency components of the image as points in reciprocal space. Conversely, the inverse Fourier transform of an image in reciprocal space transforms the information into a spatial image in real, Cartesian space. I discussed several of the basics of Fourier transforms—wavenumber, frequency space, point spread function (PSF), and optical transfer function (OTF)—in Chapter 2. Note that the OTF is defined for an optical system that is both linear and space invariant. Goodman explains that "a linear optical system is one with inputs and outputs that are real-valued functions such as intensity or complex-valued functions such as field amplitude which obey the property of superposition. Superposition is the possibility of decomposing the input into elementary functions" (Goodman, 2017). Goodman defines a linear optical system as space invariant "if the image of a point source object changes only in locations as the point source moves over the entire field" (Goodman, 2017). The OTF is the capability of an optical system to transfer the amplitude and phase information from the object plane to the image plane as a function of spatial frequency. To gain insight into the OTF I recommend the book *Introduction to the Optical Transfer Function* (Williams and Becklund, 1989).

The lateral resolution of a light microscope can be graphed as a circle in reciprocal space. The origin of the circle, the point at its center, corresponds to zero spatial frequency. The radius of this circle is proportional to the numerical aperture (NA) and inversely proportional to the wavelength of illumination. Consistent with this graphical form of resolution in reciprocal space only the information within the circumference of the circle is observable in the optical system; high-frequency information that corresponds to high-resolution information located outside the circumference of the circle is not observable.

Another term key to understanding SIM is the passband (a.k.a. support of the OTF), those regions of spatial frequency for which the OTF is not zero. Support of the OTF is a set of spatial frequency components of the specimen that can be reconstructed. However, spatial frequency components that are outside support of the OTF cannot be used to reconstruct the image in real space. In the case of fluorescence light microscopy, support of the OTF is a circle with its center at the origin. There exists a cutoff radius

$$k_{\text{cutoff}} = 4\pi\text{NA}/\lambda \tag{13.2}$$

where λ is the wavelength of the fluorescence, and NA is the numerical aperture of the microscope objective (Gaskill, 1978; Goodman, 2005). The minimal resolvable separation distance of two points of the specimen is

$$d_{\text{min}} = \lambda/2\text{NA} \tag{13.3}$$

There are a number of excellent sources that provide further insights into these topics (Bracewell, 1999; Creath et al., 2007; Gaskill, 1978; Goodman, 2005; Williams and Becklund, 1989).

Since the theory of SIM is often explained in terms of moiré patterns it makes sense to delve deeper into this topic (Creath et al., 2007; Kafri and Glatt, 1990). Interestingly, the term moiré pattern is not an eponymous French name. Moiré is the name given to a type of cloth embossed with a wave-like design that has been impressed into the cloth by opposing rollers (Creath et al., 2007). We often see moiré patterns in our day-to-day surroundings. We can observe moiré patterns when two Ronchi rulings or similar periodic patterns (alternating opaque and transparent regions) of similar period are superimposed with a small angle between them. We then can observe a beat or fringe pattern that varies with the relative angle between the two patterns of approximately equal spacing. There are many common examples of moiré patterns (e.g., two layers of silk overlapping at an angle). Other common examples are when cloth curtains overlap or when a person wearing striped clothes is viewed on television; the resulting fringes are due to the superposition of stripes on the person's clothes and the striped pattern of the video camera.

In 1874 Lord Rayleigh made a prescient suggestion that is the foundation of SIM. Rayleigh used moiré patterns to measure the quality of a set of almost identical ruled gratings. The interline spacings of each of the gratings was too small to be observed in a light microscope. But Rayleigh was still able to measure the quality of the two gratings by observing the moiré pattern formed by the two gratings (Creath et al., 2007).

In modern times the moiré effect has been applied to optical metrology (Creath et al., 2007). Rowe and Welford in 1967 developed the fringe projection technique to determine surface topography. A moiré pattern is projected onto an object and the pattern is viewed from different directions thus enabling surface contours to be determined (Rowe and Welford, 1967).

Moiré patterns are related to interferometry (Oster and Nishijima, 1963). Another important prescient finding likely motivating the development of SIM is when two gratings are placed on top of one another and rotated at a small relative angle, then a moiré pattern will be produced with the important characteristic that the spacing between the lines of the new pattern is larger than the spacing between the lines of the original two gratings. Superposition of the two similar gratings will form an interference pattern with a reduced spatial frequency. The key effect to be noted is that formation of a moiré pattern is consistent with downmodulation of the spatial frequencies of each grating. This is similar to what Rayleigh previously observed. For the simple case given above (i.e., two line gratings superimposed at a small relative angle) analysis can be undertaken in terms of the superposition of two plane waves that slightly differ in their propagation directions. Bright regions of the moiré pattern represent constructive interference of the waves in phase, and dark regions of the moiré pattern are regions where the waves are out of phase and form destructive interference (Creath et al., 2007; Oster and Nishijima, 1963).

SIM illuminates the specimen with a sinusoidal light pattern. The moiré effect is an interference phenomenon that is key to SIM. Two patterns are superimposed and a beat pattern of moiré fringes is formed. Using a fluorescent specimen as an example one pattern is the periodic pattern of light intensity and the second pattern could be the spatial distribution of fluorescence in the specimen. These two patterns form the moiré pattern.

The key to understanding the basis of SIM is that high spatial frequencies are higher than the frequency cutoff of the OTF (i.e., those outside support of the OTF). Normally these high spatial frequencies cannot pass through the optical system, but they are demodulated or downshifted toward lower spatial frequencies and therefore they can pass through the microscope objective. This high spatial frequency information is encoded into the moiré pattern and can be computationally decoded to restore the high-frequency information that otherwise would be missing. The moiré fringes or beat patterns are recorded for several different phases and orientations of the structured illumination pattern, and since the pattern of structured light is known details of the high-frequency specimen pattern can be computationally derived. Therefore, the spatial resolution of a widefield fluorescence microscope is enhanced. The details of this decoding for linear SIM and nonlinear SIM will be described in subsequent sections. Lateral resolution is enhanced by a factor of 2 for linear implementation of SIM; however, superresolution requires nonlinear implementation of SIM.

In summary, SIM depends on the moiré effect in which fine periodic patterns of light are projected onto the specimen resulting in downmodulation of the high spatial frequencies necessary to observe fine details in the image of a specimen. Such downmodulation shifts high spatial frequencies into the passband (i.e., support of the OTF). Enhanced resolution is obtained by separating this information from unmodulated frequencies and computationally shifting it to its true location in Fourier space. The reason several orientations of the grid pattern are required is that SIM only increases support of the OTF in the direction of the grid pattern; therefore,

several orientations of the grid illumination pattern are required for isotropic resolution in two dimensions. SIM recovers spatial frequencies from outside the OTF support region in reciprocal space thereby increasing support of the OTF, and therefore resolution is increased.

13.2 Antecedents of Structured Illumination Microscopy

The prescient theoretical work of A. W. Lukosz is considered the foundation of SIM. Section 2.7 discussed his seminal theoretical contributions. In the 1960s Lukosz and coworkers proposed a new technique based on structured illumination combined with widefield microscopy to extend the optical bandwidth of the transmitted spatial frequencies. Suggested extension of the optical bandwidth was limited to optical systems that are both linear and space invariant. Lukosz reviewed coherent optical systems whose resolving powers exceed the classical diffraction limit (Lukosz, 1966, 1967; Lukosz and Marchand, 1963). Lukosz proposed a method to extend the optical bandwidth of the optical system: two one-dimensional gratings are placed into the conjugate planes of the object space and the image space, the second one-dimensional grating being used to decode spatial frequency information. Although calculations were performed for coherent illumination, the authors stated that the method also worked for partially coherent or incoherent illumination. They calculated the transfer of spatial frequencies in their modified optical system (Lukosz, 1966, 1967; Lukosz and Marchand, 1963). A similar scheme was proposed by Grimm and Lohmann (1966) in their paper "Superresolution image for one-dimensional specimens."

The prescient work of Lukosz and coworkers serves as a direct link to the later work of others to develop SIM (Lukosz, 1966, 1967; Lukosz and Marchand, 1963). In fact, Lukosz's optical technique has been directly extended to the more general case of imaging three-dimensional fluorescent specimens with enhanced resolution. The one-dimensional grid pattern can be formed in two ways. A one-dimensional grid pattern can be imaged onto the object plane or it can be formed by the interference of two collimated beams (Cragg and So, 2000; Frohn et al., 2000; Gustafsson, 2000; Heintzmann and Cremer, 1999, 2002).

Three-dimensional microscopic imaging requires high lateral (transverse) and axial resolution. Typically, three-dimensional data from the specimen are obtained by acquiring a series of images as the specimen is stepped through the focal plane of the microscope. The axial resolution is limited by diffraction and noise that affect the signal-to-noise ratio (SNR).

Next, I describe a series of innovative techniques based on image interference microscopy that enhance the axial or lateral resolution of widefield optical microscopes. Perusal of the papers by Cragg and So (2000), Frohn et al., (2000), Gustafsson (2000), and Heintzmann and Cremer (1999, 2002) will demonstrate their links to the ideas expressed by Lohmann (1978). Lohmann described an optical technique that enhanced the transmission of axial frequencies through the

optical system (Lohmann, 1978). Lohmann's technique involved using a laser to illuminate the specimen with focused light from two opposing microscope objectives; the two spherical wavefronts interfere and form a fringe pattern at the common focus within the specimen. Building on this early concept of Lohmann in 1995 Gustafsson and coworkers developed image interference microscopy (I^2M) in which a fluorescent specimen is mounted between two opposing high-NA microscope objectives (Gustafsson et al., 1995). The specimen is then scanned axially and the fluorescent images are combined by a beamsplitter. The superimposed images form an interference pattern in a charge-coupled device camera. The authors achieved a seven-fold enhancement in axial resolution of the light microscope.

Other investigators continued to build on the previous work of Lohmann and used interference to enhance the axial resolution of the light microscope. Lanni and colleagues developed a microscopic technique based on standing wave excitation of fluorescence, which enhanced axial resolution (Bailey, Farkas, Taylor, and Lanni, 1993; Lanni and Bailey, 1994; Lanni et al., 1993; Lanni, Taylor, and Waggoner, 1986). In standing wave fluorescence microscopy (SWFM) two coherent plane wave beams from the laser superimpose in the specimen volume where they interfere.

Light waves can interfere and form an interference pattern, but only if their polarization state is the same, otherwise there is no interference. A number of constraints must be met for SWFM: the two incident light waves are polarized normal to their common plane of incidence (s-polarization), are of equal amplitude, and cross at complementary angles $(\theta, \pi - \theta)$ relative to the axis of the microscope, then the resulting interference pattern has an electric field intensity that only varies axially in a periodic manner. When these constraints are met, there is a five-fold enhancement in the axial resolution of the light microscope. Such standing wave excitation of specimen fluorescence can be understood as a shift of the OTF in the axial direction in reciprocal space by a distance that is equal to the spatial frequency of the standing wave. Using the SWFM technique information about the specimen, not present in images from a standard widefield fluorescent microscope, can be obtained (Bailey et al., 1993; Lanni and Bailey, 1994).

There are several variants of the SWFM technique. One publication describes a fluorescence microscope in which axial resolution is increased to better than 0.05 μm by using standing wave excitation of fluorescence (Bailey, Farkas, Taylor, and Lanni, 1993). Laser excitation light results in standing waves that are due to interference; these standing waves result in high-resolution optical sectioning (enhanced axial resolution) of the specimen. A fluorescent specimen illuminated with a standing wave will have bright and dark regions in the image. Several images of the specimen are recorded as the spacing of the nodes and the phase of the standing wave are altered. From these images a three-dimensional image of the specimen can be obtained (Lanni, Taylor, and Waggoner, 1986).

SWFMs are based on axially structured illumination. Two counter-propagating laser beams that constitute the excitation light interfere. Interference of these nonfocused laser beams results in modulation of light intensity in the axial direction. The moiré effect is central to the SWFM. Spatially changing axial intensity

encodes high axial resolution information into the fluorescence intensity observed. The computer-decoding technique is based on acquiring three images of the specimen, each at a different phase of the standing wave. The advantages of SWFM techniques are enhanced axial resolution and rapid image acquisition speed in the absence of laser scanning of the specimen.

There are a number of limitations to the SWFM technique, the major one being fringe ambiguity (i.e., lack of a unique assignment of interference fringes in excitation intensity). For specimens that are very thin (e.g., less than 0.25 μm) there is no fringe ambiguity. Ideally, to avoid this problem specimens must have a thickness less than one period of the fringe excitation light pattern. However, for thicker specimens this fringe ambiguity is problematical when it comes to three-dimensional visualization of the specimen. The origin of the problem and its effect on the formation of a valid three-dimensional image of the specimen is the three-dimensional OTF, a toroid with a missing cone surrounding the k_z at the origin (Streibl, 1984, 1985). The term missing cone refers to the OTF graph in reciprocal space in which the cutoff frequency for the axial component of the OTF k_z is zero at the origin for a standard light fluorescence microscope.

The principal limitation of SWFM techniques is fringe ambiguity, which is the subject addressed in a couple of publications I want to discuss. The first publication "True optical resolution beyond the Rayleigh limit achieved by standing wave illumination" addresses the problem of fringe ambiguity and subsequent loss of information necessary for correct 3D reconstruction (Frohn et al., 2000). Harmonic excitation light microscopy (HELM) is a method that allows lateral resolution and optical sectioning to be greatly enhanced yielding two-dimensional imaging with a lateral resolution of 100 nm (Frohn et al., 2000). Resolution can be enhanced by a factor or 2 compared with conventional fluorescence microscopy and by a factor of 1.5 compared with confocal microscopy. The HELM technique illuminates the specimen by means of a two-dimensional interference pattern from a laser. In a subsequent publication entitled "Three-dimensional resolution enhancement in fluorescence microscopy by harmonic excitation" these authors used three-dimensional harmonic excitation (3D-HELM) to demonstrate that lateral and axial resolution (optical sectioning) could be much improved by raising the number of shifted copies of the OTF that are added together. Their 3D-HELM technique achieved an isotropic point spread function (PSF) with an FWHM equal to 100 nm (Frohn et al., 2001).

The next two publications I want to discuss present a method to convert a standard widefield light microscope that has no axial resolution into an optically sectioning light microscope with axial resolution (Neil et al., 1997). The new instrument stemmed from the authors' analysis of the OTF of a standard brightfield microscope and a fluorescence microscope: only the zero spatial frequency of the OTF does not vary with defocusing above and below the focal plane (Neil et al., 1997).

In Neil et al., (1997) a grid pattern is projected onto the specimen and in Neil et al., (1998a) two light beams from a single laser interfere to project an interference fringe pattern onto the specimen. These methods of illuminating the specimen with fringe patterns were used in the subsequent development of SIM.

Neil et al., (1997) wanted to find a solution to widefield light microscopes having no axial resolution. When observing thick specimens, image quality and contrast are degraded due to light from above and below the focal plane. Confocal microscopy achieves axial resolution and concomitant optical sectioning (Masters, 1996). The authors' task was to develop a less expensive and less complex light microscope capable of optical sectioning. The authors worked in the Electrical Engineering Department at Oxford University and were inspired by the mathematical description of square-law detection. This led by analogy to mathematical analysis of their invention of a "Method of obtaining optical sectioning by using structured light in a conventional microscopy" (Neil et al., 1997). In their publication the authors cite the analogy to square-law detection in communication systems (Carlson, 1986).

The way Neil et al., (1997)'s method works in practice involves an incoherent lamp projecting the image of a grating pattern onto the specimen. Three images of the specimen with the imposed grating pattern are obtained. For each image a piezoelectric translation device shifts the grating by a third of its period. A simple mathematical analysis (shown in the next paragraphs) removes the grating projection from the images. The best optical sectioning is achieved when the period of the grating is a bit larger than the diameter of the Airy disk. Optically sectioned images of the specimen appear similar to those obtained with a confocal microscope.

In their next publication Neil Juškaitis, and Wilson (1998a, b) use an argon ion laser to form two beams that interfere and illuminate the specimen with a single-frequency interference pattern. In much the same way as shown in their previous publication the interference pattern is adjusted to form three relative phases of $0°, 120°, 240°$ and at each relative phase an image I_1, I_2, I_3 is acquired

$$I_p = \left\{ (I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_2 - I_3)^2 \right\}^{1/2} \tag{13.4}$$

The conventional image I_0 is given by

$$I_0 = \frac{I_1 + I_2 + I_3}{3} \tag{13.5}$$

Their new microscope is based on spatial heterodyning and achieves real-time optical sectioning of the specimen. Advantages over their previous microscope (1997) are easier optical alignment of components and use of a variety of laser wavelengths.

In a variant of the technique described in these two publications (Neil et al., 1997, 1998a, b) decoding can be carried out by acquiring a single exposure using the chromatic channels of a color camera, and the three displacements of the grating can be replaced with a stationary color-grating pattern (six equally spaced stripes of

different colors) as described in the publication "Single-exposure optical sectioning by color structured illumination microscopy" (Krzewina and Kim, 2006). The problem the authors solved was how to increase the imaging speed of three-dimensional specimens.

In summary, I have introduced the critical terms, concepts, and instrumental techniques that provide a solid background to continued discussion of SIM. There are a number of limitations to these techniques: mechanical stability of the microscope and the specimen during image acquisition, photodamage to live specimens, photobleaching (permanent loss of the ability to fluorescence) of fluorescent molecules, degradation of the moiré fringe pattern or beat pattern with increasing thickness of the specimen, optical aberrations, and the signal-to-noise ratio (SNR).

In the following sections I explore linear SIM (Section 13.3) and nonlinear SIM (Section 13.4). Linear implementation of SIM can enhance the resolution of a standard widefield fluorescence microscope by a factor of 2. Fluorescence intensity in linear SIM is linearly related to excitation intensity. The structured light pattern can only be focused to half the wavelength of excitation light because of diffraction. To exceed this limitation to resolution it is necessary to introduce nonlinearities, the basis of nonlinear SIM. Nonlinear SIM occurs at high excitation intensities when the fluorescent emission intensity is related to the excitation intensity in a nonlinear manner. Nonlinear SIM offers enhanced resolution and is theoretically unlimited.

13.3 Linear Structured Illumination Microscopy

The term linear SIM relates to the linear relation between illumination intensity and fluorescence intensity. This is the case for low-intensity illumination when photosaturation and photobleaching are negligible.

Following the publication of Lukosz and Marchand (1963) a number of researchers suggested and experimentally demonstrated that resolution of the optical microscope could be enhanced by linear SIM. Experiments to overcome the conventional optical resolution limit in the lateral direction (object plane) using a single objective lens approach and widefield-based fluorescence detection (e.g., using CCD cameras) occurred at the end of the 1990s (Gustafsson, 1999, 2000; Heintzmann and Cremer, 1999). Other important publications added to our appreciation of SIM (Gustafsson, 1999, 2000; Gustafsson et al., 1995, 1999, 2000; Heintzmann, 2003; Heintzmann and Benedetti, 2006; Heintzmann and Cremer, 1999, 2002).

As I pointed out earlier SWFM forms an illumination pattern in the axial direction (z) (Bailey et al., 1993) that is similar to spatially modulated illumination (SMI) (Hausmann et al., 1997). However, the new technique of SIM is different in that it illuminates the specimen in the object plane of the light microscope

(x, y) with spatially modulated structured light. The problem to be addressed is how to enhance the lateral resolution of the light microscope.

The underlying principle of these approaches is the creation of a spatially modulated illumination pattern not only in the axial direction (z), as in standing wave fluorescence microscopy (SWFM) (Bailey et al., 1993) or in SME microscopy (Hausmann et al., 1997), but also in the object plane (x, y).

A solution to this problem can be achieved by implementing two techniques. The first involves positioning a diffraction grating in the illumination beam at the conjugate object plane and projecting it through the objective lens into the specimen. The second involves forming a structured illumination pattern in the object plane by two or more laser beams interfering (Best et al., 2011; Frohn et al., 2000; Gustafsson, 2005; Schermelleh et al., 2008). The specimen and the illumination pattern are then displaced relative to each other in precise steps. A widefield detection image is acquired at each position by a charge-coupled device (CCD) camera. The images obtained are used to calculate an image with enhanced resolution using an algorithm based on the structure of the Fourier space.

In real space linear SIM can enhance effective optical resolution up to a factor of 2 compared with conventional widefield microscopy. Alternatively, in Fourier space the area of accessible information (a.k.a. support of the OTF) is twice as large for linear SIM as for a standard widefield microscope.

I now turn to a discussion of how linear SIM improves the resolution of the widefield fluorescence microscope (Gustafsson et al., 1995; Heintzmann and Cremer, 1999).

To answer this question it is useful to consider the question from the viewpoint of reciprocal space or Fourier space. In previous sections I provided requisite definitions and explanations of key terms and concepts that are necessary to understanding linear SIM. The specimen's distribution of fluorescent molecules contains high spatial frequencies. The microscope can only capture those frequencies that are within the support of the OTF; frequencies that are located outside the support of the OTF are lost. A sinusoidal light pattern (structured light) is projected onto the specimen (Gustafsson, 2000) or it is formed as interference fringes on the specimen.

This light pattern has three peaks in Fourier space: a central peak for the offset of the sinusoidal pattern, and two adjacent peaks for sinusoidal modulation (Fig. 13.1c). Fluorescence from the specimen excited by the sinusoidal excitation pattern corresponds to the product of the distribution of fluorescent molecules in the specimen and the illumination pattern. In Fourier space this multiplication of the two patterns is equivalent to convolution of their two Fourier transforms. A minimum of three images are acquired for three different positions of the structured light pattern (with three different phases that vary by 120°) and the structured illumination is then rotated by 60° and 120°, which yields nine images, and then the information in these images is used to separate the three components that are superimposed in each of the Fourier images. Each image is recorded using a

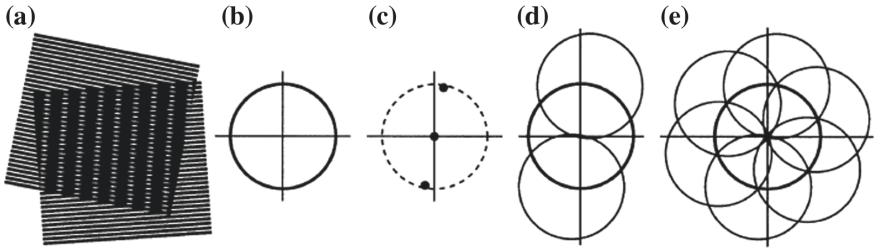


Fig. 13.1 Concept of resolution enhancement by structured illumination. (**a**) If two line patterns are superposed (multiplied), their product will contain moiré fringes (seen here as the apparent vertical stripes in the overlap region). (**b**) A conventional microscope is limited by diffraction. The set of low-resolution information that it can detect defines a circular 'observable region' of reciprocal space. (**c**) A sinusoidally striped illumination pattern has only three Fourier components. The possible positions of the two side components are limited by the same circle that defines the observable region (dashed). If the sample is illuminated with such structured light, moiré fringes will appear which represent information that has changed position in reciprocal space. The amounts of that movement correspond to the three Fourier components of the illumination. The observable region will thus contain, in addition to the normal information, moved information that originates in two offset regions (**d**). From a sequence of such images with different orientation and phase of the pattern, it is possible to recover information from an area twice the size of the normally observable region, corresponding to twice the normal resolution (**e**) From Gustafsson, M. G. L. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of Microscopy, **198**, 82–87. Used with permission

CCD camera. Finally, the separated images are shifted back to their origins and then recombined. This procedure is similar to SWFM (Bailey et al., 1994). The inverse Fourier transform will yield a new two-dimensional image with enhanced lateral resolution in real space.

Enhanced resolution is achieved because frequencies located outside support of the OTF are recovered since effective support of the OTF is increased by this procedure. This process can only increase support of the OTF in the direction of the illumination pattern. To obtain the enhancement in several directions in the specimen it is necessary to repeat the procedure with different directions of structured illumination (Fig. 13.1). Frequency mixing of the illumination sinusoidal pattern with the specimen caused by the moiré effect results in downmodulation of fine sample detail into the frequency support region of the detection optical transfer function. The factor of resolution enhancement is given by the ratio by which the size of the frequency region of the image is increased (i.e., increase in the area of support of the OTF). For linear SIM lateral resolution is enhanced by a factor of 2. To achieve two-dimensional isotropic resolution enhancement this procedure has to be repeated for at least three orientation angles and three phase shifts of the sinusoidal illumination pattern for each orientation (Fig. 13.1e). This explanation

applies to two-dimensional imaging; however, using a different type of structured illumination pattern the technique can be extended to three dimensions.

Linear SIM can also be used to enhance axial resolution of the widefield fluorescence microscope. Axial resolution in the widefield fluorescence microscope is lower than lateral resolution. In reciprocal space there is a zero near the origin. That is the origin of the "missing cone" problem and the reason widefield fluorescence microscopy lacks any capacity for optical sectioning (Streibl, 1984).

The problem to be considered is how to design of an SIM capable of enhancing axial resolution? For SIMs in which periodic structured light is formed by interference there are two possibilities. First, two-beam illumination forms an interference pattern in the object plane where the specimen is located. The resulting pattern has a spatial intensity modulation that is one-dimensional. Second, three-beam illumination forms an interference pattern in the object plane; however, this interference pattern has a spatial intensity modulation that is two-dimensional. For the three-beam case there is a spatial intensity modulation on one orientation in the focal plane and a spatial intensity modulation that is oriented parallel to the optical axis of the SIM instrument.

The three-beam SIM has a significant advantage over the two-beam SIM. A trade-off exists for the two-beam SIM between achieving maximum enhancement of lateral resolution and axial resolution. The two-beam SIM has a missing cone in its OTF and then, when lateral resolution is enhanced, there is no axial resolution or optical sectioning. This is not a problem for thin specimens. Alternatively, the two-beam SIM can be employed in combination with the total internal reflection fluorescence (TIRF) microscope, which limits the imaging depth to a very thin region adjacent to the glass slide on which the specimen is mounted.

However, a three-beam SIM has no requirement for a trade-off. Both lateral resolution and simultaneous axial resolution can be enhanced. The two-beam SIM has two diffraction orders, while the three-beam SIM is based on three diffraction orders. The origin of spatial modulation in the axial direction is the Talbot effect.

What is the Talbot effect? The Talbot effect is a diffraction effect named for Talbot (1836). A plane wave that is incident to a diffraction grating will cause an image of the grating to be repeated at a specific distance called the Talbot length from the plane of the grating. These repeated images are called Talbot images or self-images of the grating. It was Rayleigh who studied this effect and concluded that Talbot images were due to Fresnel diffraction (Rayleigh, 1881). When a beam of coherent light is incident to a transmission diffraction grating the output is a number of diffracted waves. The Talbot effect and moiré interference phenomena are related.

I now return to the discussion of how to construct an SIM that has enhanced lateral and axial resolution. To achieve this dual enhancement capability the first requirement is that the structured illumination pattern must be divisible into a lateral and an axial component. In real space an image of the specimen is considered a convolution of the structure of the specimen and the PSF with the addition of noise. In reciprocal space an image of the specimen is the Fourier transform of the

specimen multiplied by the OTF with the addition of noise. To achieve our goal of an SIM with enhanced lateral and axial resolution the following is required. In reciprocal space the usual pattern of structured light with three frequencies is required for lateral resolution enhancement, and four other frequencies in the axial plane are required to fill in the missing cone in the axial direction. This constraint means that it is necessary to acquire five images at five different phases for each orientation of the structured light pattern and at each three-dimensional focus. When these requirements are fulfilled an SIM will have the capacity to increase both lateral resolution and axial resolution by factors of 2. Next, we see how this concept is implemented in a real SIM instrument.

In 2008 Gustafsson and coworkers published their seminal paper "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination" in the *Biophysical Journal* (Gustafsson et al., 2008). This is a prime example of a three-beam SIM instrument. The authors succeeded in doubling the lateral and axial resolution in SIMs (a lateral resolution of 100 nm and an axial resolution of <300 nm) by forming a structured light pattern on the specimen that is periodic in the lateral and axial directions. The pattern is generated by a diffraction grating that forms three beams of light that are all coherent, and these beams project an interference pattern onto the specimen (Fig. 13.2).

This interference pattern encodes high-frequency information that can then be computationally manipulated to yield a three-dimensional volume image of the specimen in which lateral and axial resolution are enhanced by a factor of 2 over a standard light microscope (Fig. 13.3).

The processing steps involved in linear SIM have been elucidated (Gustafsson et al., 2008) and are summarized here:

1. Images are corrected for a flat field.
2. Separation of information mixed in the images. The mix of several images with different pattern orientations must be separated in reciprocal space (Fig. 13.1). The Fourier components that are now separated are then shifted to their correct positions in reciprocal space.
3. For each pattern the orientation and angle of the sinusoidal pattern are determined by maximizing the cross-correlation between each high-resolution component in the region of reciprocal space where they overlap.
4. Inverse filtering and reassembly. Each component of information is a low-pass filtered version of that part of the specimen structure that originated in various frequency space locations. They are assigned back to where they came from (this action increases support of the OTF) and low-pass filtering is reversed yielding the final reconstructed image with enhanced resolution (Fig. 13.3).

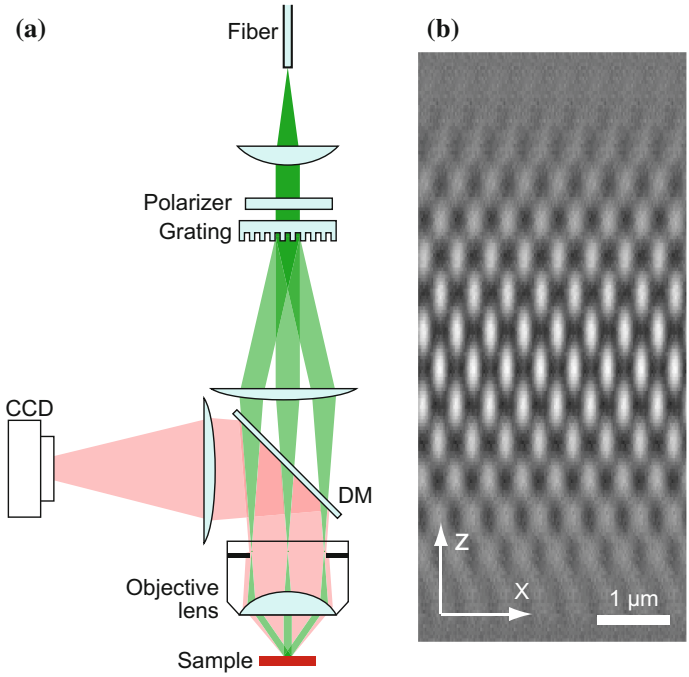


Fig. 13.2 (**a**) Simplified diagram of the structured illumination apparatus. Scrambled laser light from a multimode fiber is collimated onto a linear phase grating. Diffraction orders −l, 0 and +1 are refocused into the back focal plane of an objective lens. The beams, recollimated by the objective lens, intersect at the focal plane in the sample, where they interfere and generate an intensity pattern with both lateral and axial structure (**b**). The finite axial extent of the pattern is related to the axial broadening of its spatial frequencies. Emission light from the sample is observed by a charge-coupled device (CCD) camera via a dichroic mirror (DM). From Gustafsson, M. G. L., Shao, L., Carlton, P. M., Wang, C. J., Golubovskaya, I. N., Cande, W. Z., Agard, D. A., and Sedat, J. W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophysical Journal, **94**, 4957–4970. Used with permission

The constraints of specimen thickness and labeling density lead Gustafsson et al. to suggest the use of different types of microscopes. They suggest that SIM is preferable for thin specimens that are sparsely labeled, and that confocal microscopy is optimal for very thick and dense specimens.

Furthermore, when comparing two-dimensional SIM with three-dimensional SIM Gustafsson et al. found the latter technique required the acquisition of more images, which increased photobleaching and phototoxicity of the specimen (Gustafsson et al., 2008).

Another problem to be addressed is how to design a SIM instrument that will yield not only enhanced lateral and axial resolution, but also enhanced isometric (equal in all three orthogonal directions) resolution. The solution was found by Shao et al. (2008). Realizing that linear SIM can be further enhanced by combining

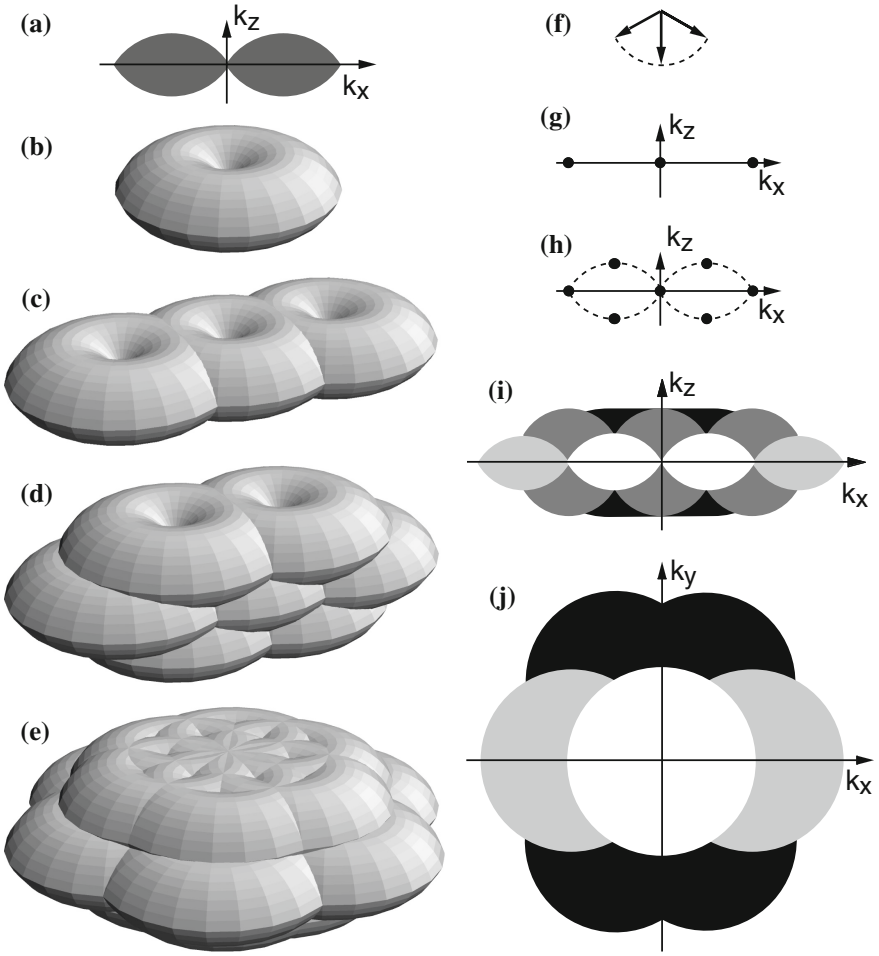


Fig. 13.3 Enlargement of the observable region of reciprocal space through structured illumination. (**a**—**e**) Observable regions for (**a** and **b**) the conventional microscope, and for structured illumination microscopy using two illumination beams (**c**), and three illumination beams in one (**d**) or three (**e**) sequential orientations. (**f**) The three amplitude wave vectors corresponding to the three illumination beam directions. All three wave vectors have the same magnitude . (**g**—**h**) The resulting spatial frequency components of the illumination intensity for the two-beam (**g**) and three-beam (**h**) case. The dotted outline in panel h indicates the set of spatial frequencies that are possible to generate by illumination through the objective lens; compare with the observable region in panel a. An intensity component occurs at each pairwise difference frequency between two of the amplitude wave vectors. (i, j): xz (**i**) and xy (**j**) sections through the OTF supports in panel b (shown in white), panel c (light shaded), panel d (dark shaded), and panel e (solid). The darker regions fully contain the lighter ones. From Gustafsson, M. G. L., Shao, L., Carlton, P. M.,Wang, C. J., Golubovskaya, I. N., Cande, W. Z., Agard, D. A., and Sedat, J. W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophysical Journal, **94**, 4957–4970. Used with permission

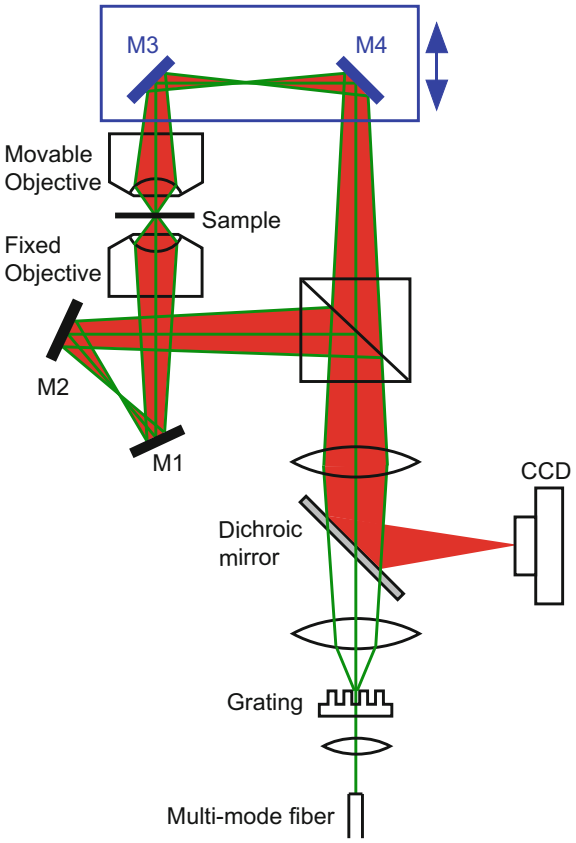


Fig. 13.4 A schematic drawing of an I^5S microscope. The illumination light passes first through a transmission grating, which diffracts it into three beams (*green lines*), and then through a beam splitter, which splits each beam and directs three beams to each of the two opposing objective lenses. The same beam splitter combines the two beams of emission light (*red*) from the sample onto the camera. The movable objective lens can be positioned in *X*, *Y*, and *Z* with respect to the stationary objective lens. Mirrors M3 and M4 can be translated together to adjust the path-length difference. The grating can be rotated and laterally translated to control the orientation and phase of the illumination pattern. From Shao, L., Isaac, B., Uzawa, S., Agard, D. A., Sedat, J. W., and Gustafsson, M. G. L. (2008). I^5S: wide-field light microscopy with 100-nm-scale resolution in three dimensions. Biophysical Journal, **94**, 4971–4983. Used with permission

it with two opposing objectives both of which are used for illumination and detection from two sides, the authors described a new type of widefield fluorescence microscopy that produces 100-nm spatial resolution in all three dimensions

(isometric) by using structured illumination in a microscope that has two opposing objective lenses (Fig. 13.4).

Previously two opposing microscope objectives have been used in 4Pi confocal fluorescence microscopy (Hell and Stelzer, 1992a, b). However, the I^5S technique involves combining structured illumination with I^5M in which two opposing microscope objectives are used to illuminate the specimen from two sides (Gustafsson et al., 1995, 1999). The two opposing microscope objectives have a combined OTF that increases lateral support of the OTF resulting in enhanced axial resolution and partially filling the missing cone. The two opposing microscope objectives permit six-beam structured illumination that has increased sinusoidal modulation in the axial direction. Lateral support of the OTF is the same as for two- or three-beam structured illumination microscopes; however, support of the OTF for the axial direction is greatly enhanced with the result that axial resolution is increased. The I^5S microscope achieves 100-nm resolution (enhanced isometric resolution of a widefield light microscope) in three directions (Gustafsson et al., 2008) (Fig. 13.5).

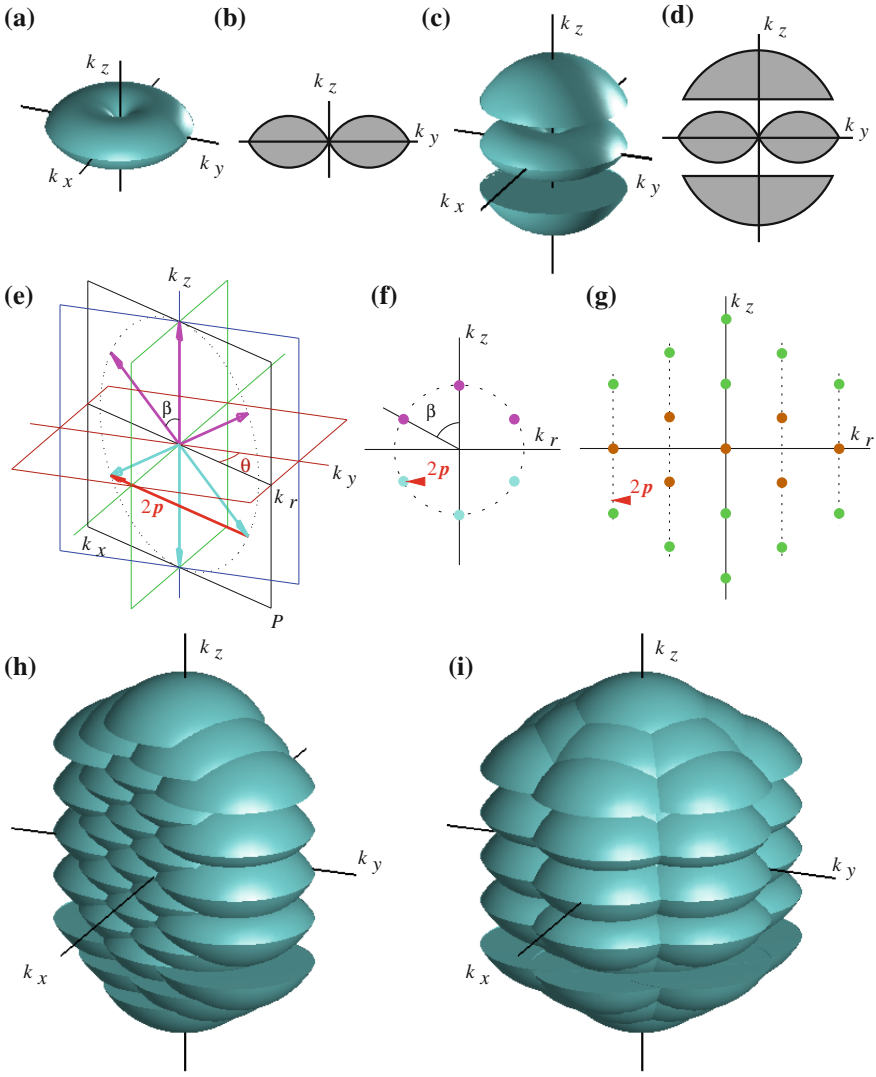
This discussion of linear SIM has demonstrated remarkable advances in the design of SIM instruments. Such advances include enhanced lateral resolution and axial resolution as well as advent of the I^5S widefield light microscope that achieves 100-nm resolution in the lateral and axial directions. A logical question posed by researchers was whether there are SIM techniques that can exceed the resolution enhancement achieved with linear SIM. The answer is yes. How this is achieved is described in the next section.

13.4 Nonlinear Structured Illumination Microscopy

In the previous section I discussed linear SIM in which excitation intensity was sufficiently low that fluorescence intensity was proportional to it. In linear SIM both the structured illumination and detection are diffraction limited. Under such conditions linear SIM was able to enhance lateral and axial resolution by factors of 2.

In this section I describe nonlinear SIM as a technique to exploit nonlinear fluorescence spectroscopy in combination with structured light of sufficient intensity to induce nonlinear effects in the photophysics of fluorescent molecules. Nonlinear optical effects include processes such as multiphoton absorption and harmonic generation (Boyd, 2008). Nonlinear SIM exploits either fluorescence saturation or photoswitching (using light to switch the photophysical properties of a fluorescent molecule between two different states) (Hirvonen, 2008).

The resolution achieved by a nonlinear SIM is theoretically without limits; in practice, resolution is limited by the noise associated with higher harmonics. Also noteworthy are the detrimental effects of the high intensities of illumination required to produce the nonlinearities exploited in nonlinear SIM. These detrimental effects include photobleaching (decreased fluorescence of the specimen due to light-induced


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
k z
k x
k y
k r
β
θ
2p
P

◂**Fig. 13.5** Increased support of the OTF by I^5S. Support of the OTF for conventional widefield microscopy (**a**, **b**) and for coherent detection through two opposing objective lenses (**c**, **d**) in 3D rendering (**a**, **c**) and in axial cross section (**b**, **d**). The supports are drawn for monochromatic emission light. Spatial frequency components of illumination (**e**–**g**). I^5S uses six mutually coherent illumination beams, three entering through one objective lens (*magenta arrows* in **e**) and three through the other objective lens (*cyan arrows* in **e**). All six beams lie within one plane *P* of frequency space (*black frame* in **e**). Corresponding frequency components of light amplitude shown within plane *P* (**f**). The six illumination beams interfere pairwise; each pair of amplitude dots in (**f**) give rise to one illumination intensity component at the position given by the difference vector between the two amplitude dots. This leads to a total of 19 distinct intensity components (**g**). *Green dots* in (**g**) represent interference between beams that passed through different objective lenses; these are not present in SIMs with a single objective lens. **h** Effective support of the OTF produced by the illumination beams shown in (**e**). This effective OTF support equals a convolution of the detection of support of the OTF in (**c**) with the illumination intensity in (**g**). In (**h**) the resolution extension lies wholly within plane *P*, but by repeating it with different pattern orientations (different values of the angle θ as defined in **e**), the overall effective support of the OTF of I^5S can be made nearly spherical (**i**). From Shao, L., Isaac, B., Uzawa, S., Agard, D. A., Sedat, J. W., and Gustafsson, M. G. L. (2008). I^5S: wide-field light microscopy with 100-nm-scale resolution in three dimensions. Biophysical Journal, **94**, 4971–4983. Used with permission

destruction of fluorescent molecules) and photodamage (even to fixed tissues). These factors limit its use for live cell and tissue imaging. Fortunately, reversible photoswitching of fluorescent proteins in the specimen is an alternative technique to achieve nonlinear SIM and can yield the nonlinearities required for nonlinear SIM; however, this is achieved at light intensities six orders of magnitude below that required to induce fluorescence saturation (Hirvonen, 2008).

Nonlinear SIM is a widefield fluorescence imaging technique that can theoretically provide unlimited resolution. Nonlinear SIM was proposed by Heintzmann et al., (2002) in their publication "Saturated patterned excitation microscopy-a concept for optical resolution improvement." Gustafsson (2005) demonstrated nonlinear SIM in his publication "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution." Any nonlinear fluorescence process can be utilized to introduce higher order harmonics into a structured illumination pattern. While nonlinear SIM can achieve arbitrarily high resolution, in practice it is limited by phototoxicity induced in the specimen and the signal-to-noise ratio.

Nonlinear SIM can be based on the nonlinear fluorescence of a molecule near saturation (Gustafsson, 2005; Heintzmann et al., 2002). Fluorescence saturation is one way of achieving nonlinear SIM (Heintzmann et al., 2002). Fluorescence is linearly related to intensity at low levels of excitation intensity. If excitation intensity is doubled, then fluorescence intensity is doubled. The standard fluorescence microscope is usually operated in the linear regime. However, fluorescence will no longer be linearly related to excitation at very high intensities of excitation. A plot of fluorescence intensity will show saturation with increasing excitation intensity. Sinusoidal structured light in this very high excitation intensity regime is no longer sinusoidal, but exhibits a flattening of intensity peaks in real space. In reciprocal space this distortion results in higher order harmonics with additional

orders. Since these higher order harmonics are located in support of the structured illumination OTF they can be used in nonlinear SIM to enhance resolution well beyond the diffraction limit. This explains why nonlinear SIM based on high-excitation induced saturation of fluorescence can result in resolution beyond the diffraction limit.

For linear SIM the low levels of excitation result in a pattern in reciprocal space of only three orders; however, with the very high excitation intensities of nonlinear SIM the distorted sinusoidal patterns of the structured illumination when mapped in reciprocal space have more than three additional orders, which can be located outside the frequency support of the structured illumination OTF.

To explain limitations to resolution achieved by nonlinear SIM it is critical to understand that the additional higher orders exhibit ever-lower intensities. This explains why the SNR of the images is the limiting factor to the degree by which resolution is enhanced with nonlinear SIM based on saturation of fluorescence at very high intensities of excitation. These high illumination intensities have other effects such as increased photobleaching of the specimen leading to reduced fluorescence emission and therefore a reduction in the SNR.

The high intensities required for nonlinear SIM lead to saturation of fluorescence because photon absorption can only occur from the ground state to the excited states. Due to the finite time of deexcitation (intrinsic fluorescence decay rate) from higher excited states to the ground state of fluorescent molecules, fluorescence saturation will occur at very high excitation intensities since fewer molecules will be in the ground state as more and more are in the excited states. These effects result in additional harmonics in SIM illumination with a concomitant increase in spatial frequencies of the illumination. It is these additional harmonics that are used to achieve the increased spatial resolution of nonlinear SIM (Gustafsson, 2005).

I will first delve deeper into the publication "Saturated patterned excitation microscopy—a concept for optical resolution improvement" (Heintzmann et al., 2002). This publication is based on using computer simulations to validate theoretical concepts. Then, I will discuss Gustafsson's experimental implementation of saturated pattern excitation microscopy (SPEM) (Heintzmann et al., 2002) in his publication "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution" (Gustafsson, 2005).

The motivation behind the publication "Saturated patterned excitation microscopy—a concept for optical resolution improvement" was to develop a new type of microscope to increase the support of the OTF by more than a factor of 2 without the requirement of stimulated emission or complex depletion techniques. Heintzmann et al. show that nonlinear effects, such as fluorescence saturation, that occur at high intensities can be used to expand the support of the OTF. When the computationally derived enhanced OTF is back-transformed with an inverse Fourier transform it is possible to detect information from the specimen at very high spatial frequencies and therefore at very high resolution—in fact, at arbitrary resolution. In practice, the highest resolution of the microscope is limited by the SNR of the data.

In addition to their working out the theory of SPEM the authors also suggested some experimental realizations based on different types of far-field epi-fluorescence microscopes. In all designs a high-intensity light source is a key requirement.

Motivated by Heintzmann et al., (2002), who developed the concept of saturation of the excited state to enhance the resolution of SIM, Gustafsson built a nonlinear SIM that demonstrated the capability of saturated structured illumination microscopy (SSIM) to achieve a two-dimensional resolution of <50 nm (Gustafsson, 2005). The ultimate resolution that SSIM can attain is limited by the SNR, and the SNR is limited by the extent of photobleaching of the specimen.

Alternatively, nonlinear SIM can be based on photoswitchable fluorophores (Hirvonen, 2008). Hirvonen found a way to build a nonlinear SIM that operates on significantly less power than the SSIM that Gustafsson (2005) constructed.

First, I define photoswitchable fluorescent probes. Specific wavelengths of light can switch a molecule between two states: the dark OFF state (nonfluorescent) and the bright ON state. The OFF state cannot be excited to fluoresce. The ON state can be excited to fluoresce. These photoswitchable fluorescent probes can be used in nonlinear SIM to form nonlinearities between the illumination and the emission. The activation of switchable fluorescent molecules is a stochastic or random process, and the activation process will exhibit a saturation curve analogous to the previously described saturation curve for fluorescence, except the shape of the activation curve is exponential. Typically, all switchable fluorescent molecules are first activated (ON state), and then structured illumination that has a sinusoidal pattern of intensity is used to deactivate switchable fluorescent molecules. Saturation occurs in the deactivation process. Saturation is a function of total light absorbed (intensity $\times$ time). The main advantage of the technique is that it works with low-intensity illumination, which produces minimal photodamage to biological specimens (Hirvonen, 2008). In particular, a nonlinear SIM based on photoswitchable probes requires six orders of magnitude less light intensity than is required for SSIM that is based on saturation of the excited state. This implementation of nonlinear SIM achieved a resolution of 40 nm.

What are the limitations of a nonlinear SIM based on photoswitchable fluorescent probes? First, the photophysics and spectroscopic properties of these photoswitchable probes severely restrict selection of suitable fluorescent molecules. Another required property of these photoswitchable probes is that they must show a complete OFF state and a complete ON state to be effective. Furthermore, these photoswitchable probes should be resistant to photobleaching and must be very intense emitters to result in a large SNR.

The publication "Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50 nm resolution" provides a seminal tutorial example of this technique (Rego et al., 2012). The authors constructed a nonlinear SIM that yields superresolution and is suitable for biological imaging. The publication notes that Mats G. L. Gustafsson died on April 17, 2011. His innovative and seminal work on the invention and development of SIM is an

outstanding achievement. Gustafsson was given credit for conceiving the idea that led to the project in Rego et al.

Rego et al. found a way to enhance the resolution of a widefield light microscope to overcome diffraction-limited resolution and achieve superresolution without the extremely high intensities required for saturation (SSIM) that induce the necessary nonlinearities. Their solution was a technique that uses the reversible photoswitching of a fluorescent protein to achieve nonlinear SIM with an illumination intensity six orders of magnitude lower than that required for saturation SIM. The authors realized that for a superresolution SIM technique to be compatible with biological specimens it must operate with illumination intensities many orders of magnitude below that required for previous saturation methods (SSIM).

They invented and developed a new SIM technique and demonstrated that it was compatible with biological specimens. To do that they achieved a 40 nm resolution on a specimen that consisted of purified microtubules labeled with the fluorescent photoswitchable protein Dronpa. They succeeded in generating the required nonlinear response with light intensities of 1–10 W/cm^2. By combining their SIM technique with a total internal reflection fluorescence (TIRF) SIM they achieved two-dimensional resolution that exceeded the diffraction limit by a factor of 4.

In addition, they demonstrated that their technique could image cellular structures by imaging the mammalian nuclear pore and actin cytoskeleton (Rego et al., 2012).

How does Rego et al.'s new technique compare with linear SIM techniques? First, the new technique based on photoswitchable probes requires the acquisition of more phases and orientations of structured light patterns on the specimen to be able to separate higher order harmonic information. This is an absolute requirement to achieve isotropic resolution.

Rego et al. address the limitations that superresolution nonlinear SIMs have when it comes to live cell imaging: increased image acquisition times due to the higher numbers of images required and the resultant motion blur of images. To help understand the increased complexity of data acquisition associated with nonlinear SIM based on photoswitchable probes I list some of the steps that take place in data acquisition. First, fluorescent molecules are switched to the ON state. Second, the structured light pattern switches fluorescent molecules to the OFF state. However, fluorescent molecules in the dark regions of the illumination pattern are still in the ON state and their fluorescence is detected. These steps are repeated for different phases of the illumination pattern and for different orientations of the illumination pattern.

There is a need for advances in faster pattern formation instrumentation and the development of new photoswitchable probes with increased photostability. Fast image acquisition speeds for SIM have been achieved using spatial light modulators, an approach that may help to form illumination patterns on the specimen at great speeds. A good example is the 2012 publication "Super-resolution video microscopy of live cells by structured illumination" (Kner et al., 2009).

13.5 Overview of Structured Illumination Microscopy

In this chapter I have discussed innovations that have enhanced lateral and axial resolution as compared to diffraction-limited widefield fluorescence microscopes. Gustafsson's publication "Extended resolution fluorescence microscopy" is of great pedagogical value (Gustafsson, 1999). He critically explains the principles and limitations of a variety of microscopic techniques that enhance resolution. SIM achieves high-resolution imaging and optical sectioning in widefield microscopes.

These SIM techniques can be divided into linear SIM that achieves resolution enhancement by a factor of 2, and nonlinear SIM that can achieve superresolution imaging. The aim is to minimize photobleaching of fluorescent probes and phototodamage to living specimens. Furthermore, the effects of aberrations and mechanical stability of the specimen and the microscope require consideration. Moreover, it is important to critically consider the effects of illumination light intensity on the biological specimen. In addition, temporal resolution of the SIM technique is an important consideration for live specimens. Further considerations involve size of the specimen. There exist severe size limitations to specimens in microscope methods that use two opposing microscope objectives. Other important considerations relate to the speed of data acquisition. For example, 4Pi microscopy is a specimen-scanning technique. In comparison, I^5M SIM is a widefield technique that can acquire images orders of magnitude faster than with 4Pi microscopy.

Specimen size, transparency, mobility, and vitality all determine the limitations and suitability for each type of microscope. A SIM instrument requires that sinusoidal patterns within the specimen are not degraded by absorption and scattering of incident light, that light intensity does not affect the structure or viability of the specimen, and that the signal-to-noise ratio of the image is sufficient to computationally decode high-frequency information. The unique requirements of fluorescent labeling vary for each type of specimen presenting another important consideration for SIM applications. The choice of specimens for each type of SIM is limited according to these considerations.

Earlier in the chapter I discussed SIM instruments that use a one-dimensional sinusoidal grid pattern. This typically involves a line pattern being formed on the specimen and phase images being taken at different positions of the line grid. In 2014 Schropp and Uhl published a new technique "Two-dimensional structured illumination microscopy" (Schropp and Uhl, 2014). This SIM technique uses rectangular and hexagonal patterns. A two-dimensional pattern leads to a more isotropic power spectral density and a slightly enhanced signal-to-noise ratio in the SIM image evaluated compared with one-dimensional SIM. The authors demonstrated theoretically and experimentally that an analogous method using hexagonal patterns yields still better isotropic sectioning (Schropp and Uhl, 2014).

Distortion of the illumination pattern due to the specimen or the illumination optics in SIM was addressed in the recent publication "Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slice blind-SIM reconstruction" (Jost et al., 2015).

In this chapter I have discussed the creative work, much of which is ongoing, of many investigators. The theoretical foundations, instrumental implementations, and limitations of specimens, illumination intensity, the signal-to-noise ratio, photobleaching and phototoxicity to the specimen, as well as optical aberrations all play a critical role in the successful use of new types of microscopes with enhanced lateral and axial resolution. One metric of success is the improvement in resolution. Another metric, perhaps equally important or more important, is the new knowledge and understanding of biological systems that only became possible as a result of newly developed optical microscopes.

For example, from the plethora of biological applications I selected the seminal *Science* publication "Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy" (Schermelleh et al., 2008). In their study of the mammalian nucleus using three-dimensional structured illumination microscopy (a subdiffraction technique that can produce multicolor, three-dimensional images of whole cells with enhanced lateral and enhanced axial resolution) they imaged chromatin, nuclear lamina, and the nuclear pore complex. They reported several features not visualized by conventional microscopy.

References

Bailey, B., Farkas, D. L., Taylor, D. L., and Lanni, F. (1993). Enhancement of axial resolution in fluorescence microscopy by standing-wave excitation. Nature, **366**, 44–48.

Bailey, B., Krishnamurthi, V., Farkas, D. L., Taylor, D. L., and Lanni, F. (1994). Three-dimensional imaging of biological specimens with standing wave fluorescence microscopy. Proceedings of SPIE, **2184**, 208–213.

Best, G., Amberger, R., Baddeley, D., Ach, T., Dithmar, S., Heintzmann, R., and Cremer, C. (2011). Structured illumination microscopy of autofluorescent aggregations in human tissue. Micron, **42**, 330–335.

Boyd, R. W. (2008). *Nonlinear Optics*, Third Edition. San Diego: Academic Press.

Bracewell, R. (1999). *The Fourier Transform and Its Applications,* Third Edition, New York: McGraw-Hill.

Carlson, A. B. (1986). *Communications Systems*. New York, McGraw-Hill.

Cragg, G. E., and So, P. T. C. (2000). Lateral resolution enhancement with standing evanescent waves. Optics Letters, **25**, 46–48.

Creath, K., Schmitt, J., and Wyant, J. C. (2007). Optical Metrology of Diffuse Surfaces, in: *Optical Shop Testing,* Third Edition, Daniel Malacara, Editor. Hoboken: John Wiley & Sons, pp. 756-807.

Frohn, J. T., Knapp, H. F., and Stemmer, A. (2000). True optical resolution beyond the Rayleigh limit achieved by standing wave illumination. Proceedings of the National Academy of Sciences USA, **97**, 7232–7236.

Frohn, J. T., Knapp, H. F., and Stemmer, A. (2001). Three-dimensional resolution enhancement in fluorescence microscopy by harmonic excitation. Optics Letters, **26**, 828–830.

Gaskill, J. D. (1978). *Linear systems, Fourier Transforms, and Optics*. New York: John Wiley and Sons.

Goodman, J. W. (2005). *Introduction to Fourier Optics, Third Edition.* Greenwood Village, Colorado: Roberts & Company.

Goodman, J. W. (2017). *Introduction to Fourier optics* (4th ed.). New York: W. H. Freeman and Company.

Grimm, M. A., and Lohmann, A. W. (1966). Super resolution image for one-dimensional objects. Journal of the Optical Society of America, **56**, 1151–1156.

Gustafsson, M. G. L., Sedat, J. W., and Agard, D. A. Method and apparatus for three-dimensional microscopy with enhanced depth resolution. US Patent RE38,307, E1. Filed 1995, reissued 11 November 2003.

Gustafsson, M. G. L. (1999). Extended resolution fluorescence microscopy. Current Opinion in Structural Biology, **9**, 627–634.

Gustafsson, M. G. L. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of Microscopy, **198**, 82–87.

Gustafsson, M. G. L., Agard, D. A., and Sedat, J. W. (1995). Sevenfold improvement of axial resolution in 3D widefield microscopy using two objective lenses. Proceedings of SPIE, **2412**, 147–156.

Gustafsson, M. G. L., Agard, D. A., and Sedat, J. W. (1999). I^5M: 3D widefield light microscopy with better than 100 nm axial resolution. Journal of Microscopy, **195**, 10–16.

Gustafsson, M. G. L., Agard, D. A., and Sedat, J. W. (2000). Doubling the lateral resolution of wide-field fluorescence microscopy using structured illumination. Proceedings of SPIE, **3919**, 141–150.

Gustafsson, M. G. L. (2005). Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. Proceedings of the National Academy of Sciences USA, **102**, 13081–13086.

Gustafsson, M. G. L. (2008). Super-resolution light microscopy goes live. Nature Methods, **5**, 385–387.

Gustafsson, M. G. L., Shao, L., Carlton, P. M., Wang, C. J., Golubovskaya, I. N., Cande, W. Z., Agard, D. A., and Sedat, J. W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophysical Journal, **94**, 4957–4970.

Hausmann, M., Schneider, B., Bradl, J., and Cremer, C. (1997). High-precision distance microscopy of 3D-nanostructures by a spatially modulated excitation fluorescence microscope. Proceedings of SPIE, **3197**, 217–222.

Heintzmann, R. (2003). Saturated patterned excitation microscopy with two-dimensional excitation patterns. Micron, **34**, 283–291.

Heintzmann, R., and Benedetti, P. A. (2006). High-resolution image reconstruction in fluorescence microscopy with patterned excitation. Applied Optics, **45**, 5037–5045.

Heintzmann, R., and Cremer, C. (1999). Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. Proceedings of SPIE, **3568**, 185–196.

Heintzmann, R., and Cremer, C. (2002). Axial tomographic confocal fluorescence microscopy. Journal of Microscopy, **206**, 7–23.

Heintzmann, R., Jovin, T. M., and Cremer, C. (2002). Saturated patterned excitation microscopy-a concept for optical resolution improvement. Journal of the Optical Society of America A, **19**, 1599–1609.

Hell, S., and Stelzer, E. H. K. (1992a). Properties of a 4Pi confocal fluorescence microscope. Journal of the Optical Society of America A, **9**, 2159–2166.

Hell, S., and Stelzer, E. H. K. (1992b). Fundamental improvement of a 4Pi confocal fluorescence microscope using two-photon excitation. Optics Communications, **93**, 277–282.

Hirvonen, L. (2008). *Structured illumination microscopy using photoswitchable fluorescent proteins*. PhD thesis, King's College London, UK.

Jost, A., Tolstik, E., Feldmann, P., Wicker, K., Sentenac, A., and Heintzmann, R. (2015). Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slice blind-SIM reconstruction. PLoS ONE, **10**(7), e0132174. https://doi.org/10.1371/journal.pone.0132174.

Kafri, O., and Glatt, I. (1990). *The Physics of Moiré Metrology.* New York: John Wiley & Sons.

Kner, P., Chhun, B. B., Griffis, E. R., Winoto, L., and Gustafsson, M. G. (2009). Super-resolution video microscopy of live cells by structured illumination. Nature Methods, **6**, 339–342.

Krzewina, L. G., and Kim, M. K. (2006). Single-exposure optical sectioning by color structured illumination microscopy. Optics Letters, **31**, 477–479.
Lanni, F., and Bailey, B. (1994). Standing-wave excitation for fluorescence microscopy. Trends in Cell Biology, **4**, 262–265.
Lanni, F., Bailey, B., Farkas, D. L., and Taylor, D. L. (1993). Excitation field synthesis as a means for obtaining enhanced axial resolution in fluorescence microscopes. Bioimaging, **1**, 187–196.
Lanni, F., Taylor, D. L., and Waggoner, A. S. (1986). Standing wave luminescence microscopy. Patent, US 4621911 A, Nov 11, 1986.
Lohmann, A. W. (1978). Three-dimensional properties of wave-fields. Optik, **51**, 105–117.
Lukosz, W. (1966). Optical systems with resolving powers exceeding the classical limit, Part 1. Journal of the Optical Society of America, **56**, 1463–1471.
Lukosz, W. (1967). Optical systems with resolving powers exceeding the classical limit. II. Journal of the Optical Society of America, **57**, 932–941.
Lukosz, W., and Marchand, M. (1963). Optischen Abbildung Unter Überschreitung der Beugungsbedingten Auflösungsgrenze. Optica Acta, **10**, 241–255.
Masters, B. R. (1996). *Selected Papers on Confocal Microscopy*. Bellingham, SPIE Press.
Neil, M. A. A., Juškaitis, R., and Wilson, T. (1997). Method of obtaining optical sectioning by using structured light in a conventional microscopy. Optics Letters, **22**, 1905–1907.
Neil, M. A. A., Juškaitis, R., and Wilson, T. (1998a). Real time 3D fluorescence microscopy by two beam interference illumination. Optics Communications, **153**, 1–4.
Neil, M. A. A., Wilson, T., and Juškaitis, R. (1998b). A light efficient optically sectioning microscope. Journal of Microscopy, **189**, 114–117.
Oster, G., and Nishijima, Y. (1963). Moiré patterns. Scientific American, **208**, 54–63.
Rayleigh, L. (1881). On copying diffraction gratings and on some phenomenon connected therewith. Philosophical Magazine, **11**, 196–205.
Rego, E. H., Shao, L., Macklin, J. J., Winoto, L., Johansson, G. A., Kamps-Hughes, N., Davidson, M. W., and Gustafsson, M. G. L. (2012). Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution. Proceedings of the National Academy of Sciences USA, **109**, E135–E143.
Rowe, S. H., and Welford, W. T. (1967). Surface topography of non-optical surfaces by projected interference fringes. Nature, **216**, 786–787.
Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M. C., Agard, D. A., Gustafsson, M. G. L., Leonhardt, H., and Sedat, J. W. (2008). Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. Science, **320**, 1332–1336.
Schropp, M., and Uhl, R. (2014). Two-dimensional structured illumination microscopy. Journal of Microscopy, **256**, 23–36.
Shao, L., Isaac, B., Uzawa, S., Agard, D. A., Sedat, J. W., and Gustafsson, M. G. L. (2008). I^5S: wide-field light microscopy with 100-nm-scale resolution in three dimensions. Biophysical Journal, **94**, 4971–4983.
Streibl, N. (1984). Fundamental restrictions for 3-D light distributions. Optik, **66**, 341–354.
Streibl, N. (1985). Three-dimensional imaging by a microscope. Journal of the Optical Society of America A, **2**, 121–127.
Talbot, H. F. (1836). Facts relating to optical science. Philosophical Magazine, **9**, 401–407.
Wicker, K., Mandula, O., Best, G., Fiolka, R., and Heintzmann, R. (2013). Phase optimization for structured illumination microscopy. Optics Express, **21**, 2032–2049.
Williams, C. S., and Becklund, O. A. (1989). *Introduction to the Optical Transfer Function.* New York: Wiley-Interscience.

Further Reading

Alexandrov, S. A., Hillman, T. R., Gutzler, T., and Sampson, D. D. (2006). Synthetic aperture fourier holographic optical microscopy. Physical Review Letters, **97**, 168102-1 to -4.

Ball, G., Demmerle, J., Kaufmann, R., Davis, I., Dobbie, I. M., and Schermelleh, L. (2015). SIMcheck: A toolbox for successful super-resolution structured illumination microscopy. Scientific Reports, **5**, 15915. https://doi.org/10.1038/srep15915.

Betzig, E. (2005). Excitation strategies for optical lattice microscopy. Optics Express, **13**, 3021–3036.

Bewersdorf, J., Schmidt, R., and Hell, S. W. (2006). Comparison of I^5M and 4Pi-microscopy. Journal of Microscopy, **222**, 105–117.

Blanca, C. M., and Hell, S. W. (2002). Axial superresolution with ultrahigh aperture lenses. Optics Express, **10**, 893–898.

Buscher, D. F. (2015). *Practical Optical Interferometry. Imaging at Visible and Infrared Wavelengths*. Cambridge, UK: Cambridge University Press.

Chung, E., Kim, D., Cui, Y., Kim, Y-H., and So, P. T. C. (2007). Two-dimensional standing wave total internal reflection fluorescence microscopy: Superresolution imaging of single molecular and biological specimens. Biophysical Journal, **93**, 1747–1757.

Cremer, C., and Masters, B. R. (2013). Resolution enhancement techniques in microscopy. The European Physical Journal H, **38**, 281–344. (Open Access article).

Débarre, D., Botcherby, E. J., Booth, M. J., and Wilson, T. (2008). Adaptive optics for structured illumination microscopy. Optics Express, **16**, 9290–9305.

Dubois, A., Vabre, L., Boccara, A.-C., and Beaurepaire, E. (2002). High-resolution full-field optical coherence tomography with a Linnik microscope. Applied Optics, **41**, 805–812.

Egner, A., and Hell, S. W. (2005). Fluorescence microscopy with super-resolved optical sections. Trends in Cell Biology, **15**, 207–215.

Fiolka, R., Shao, L., Rego, E. H., Davidson, M. W., and Gustafsson, M. G. L. (2012). Time-lapse two-color 3D imaging of live cells with doubled resolution using structured illumination. Proceedings of the National Academy of Sciences USA, **109**, 5311–5315.

Heintzmann, R., and Ficz, G. (2007). Breaking the resolution limit in light microscopy. Methods in Cell Biology, **81**, 561–580.

Jost, A., and Heintzmann, R. (2013). Superresolution multidimensional imaging with structured illumination microscopy. Annual Review of Materials Research, **43**, 261–82.

Kam, Z., Hanser, B., Gustafsson, M. G. L., Agard, D. A., and Sedat, J. W. (2001). Computational adaptive optics for live three-dimensional biological imaging. Proceedings of the National Academy of Sciences USA, **98**, 3790–3795.

Lim, D., Chu, K. K., and Mertz, J. (2008). Wide-field fluorescence sectioning with hybrid speckle and uniform-illumination microscopy. Optics Letters, **33**, 1819–1821.

Lim, D., Ford, T. N., Chu, K. K., and Mertz, J. (2011). Optically sectioned in vivo imaging with speckle illumination HiLo microscopy. Journal of Biomedical Optics, **16**, 016014-1 to 016014-8.

Lukyanov, K. A., Fradkov, A. F., Gurskaya, N. G., Matz, M. V., Labas, Y. A., Savitsky, A. P., Markelov, M. L., Zaraisky, A. G., Zhao, X., Tan, W., and Lukyanov, S. A. (2000). Natural animal coloration can be determined by a nonfluorescent green fluorescent protein homolog. Journal of Biological Chemistry, **275**, 25879–25882.

Martínez-Corral, M., and Saavedra, G. (2009). The Resolution Challenge in 3D Optical Microscopy, in: *Progress in Optics*, Emil Wolf, Editor, **53**, 1–67.

Mertz, J. (2011). Optical sectioning microscopy with planar or structured illumination. Nature Methods, **8**, 911–819.

Neil, M. A. A., Juškaitis, R., Wilson, T., Laczik, Z. J., and Sarafis, V. (2000a). Optimized pupil-plane filters for confocal microscope point-spread function engineering. Optics Letters, **25**, 245–247.

Neil, M. A. A., Squire, A., Juškaitis, R., Bastiaens, P. I., and Wilson, T. (2000b). Wide-field optically sectioning fluorescence microscopy with laser illumination. Journal of Microscopy, **197**, 1–4.

Porter, A. B. (1906). On the diffraction theory of microscope vision. Philosophical Magazine, **6**, 154–156.

Rego, E. H., and Shao, L. (2015). Practical structured illumination microscopy. In: *Advanced fluorescence Microscopy: Methods and Protocols in Molecular Biology*. Peter J. Verveer (ed.), vol. **1251**, 175–192.

Reasenberg, R. D. (Ed.) (1998). Astronomical interferometry. Proceedings of SPIE, **3350**, entire volume. Bellingham: SPIE.

Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. The Journal of Cell Biology, **190**, 165–175.

Schwentker, A., Bock, H., Hofmann, M., Jakobs, S., Bewersdorf, J., Eggeling, C., and Hell, S. W. (2007). Wide-field subdiffraction RESOLFT microscopy using fluorescent protein photoswitching. Microscopy Research and Technique, **70**, 269–280.

Shabtay, G., Mendlovic, D., Zalevsky, Z., and Lipson, L. (2001). The optimal system for sub-wavelength point source localization. Optics Communications, **198**, 311–315.

Shao, L., Kner, P., Rego, E. H., and Gustafsson, M. G. L. (2011). Super-resolution 3D microscopy of live whole cells using structured illumination. Nature Methods, **8**, 1044–1046.

Shao, L., Winoto, L., Agard, D. A., Gustafsson, M. G. L., and Sedat, J. W. (2012). Interferometer-based structured-illumination microscopy utilizing complementary phase relationship through constructive and destructive image detection by two cameras. Journal of Microscopy, **246**, 229–236.

Sheppard, C. J. R. (2007). Fundamentals of superresolution. Micron, **38**, 165–169.

Wicker, K. (2013). Non-iterative determination of pattern phase in structured illumination microscopy using auto-correlations in Fourier space. Optics Express, **21**, 24692–24701.

Wicker, K., and Heintzmann, R. (2014). Resolving a misconception about structured illumination. Nature Photonics, **8**, 341–344.

Wilson, T., Neil, M. A. A., and Juškaitis, R. (1998). Real-time three-dimensional imaging of microscopic structures. Journal of Microscopy, **191**, 116–118.

Zalevsky, Z., and Mendlovic, D. (2004). *Optical Superresolution*. Springer Series in Optical Sciences. New York: Springer.

Chapter 14
Stimulated Emission Depletion Microscopy and Related Techniques

14.1 Introduction

Stimulated Emission Depletion (STED) microscopy is a far-field, scanning fluorescence microscope technique that yields superresolution images of the specimen (Hell and Wichmann, 1994). The *Nobel Prize in Chemistry 2014* was awarded jointly to Eric Betzig, Stefan W. Hell, and William E. Moerner "for the development of super-resolved fluorescence microscopy." Hell is recognized for his seminal contribution to STED microscopy.

STED microscopy is based on two diffraction-limited, coaxial laser beams (excitation laser and STED laser) that are tightly focused by the high numerical aperture microscope objective (Braat and Török, 2019; Ganic et al., 2003; Gu, 2000). The diffraction-limited STED laser beam is of a longer wavelength than the excitation laser beam. The diffraction-limited excitation laser beam causes the transition of fluorescent molecules from the ground state to the excited state.

The STED beam has an annular intensity profile with null or zero intensity on its axis. A spiral phase plate converts a Gaussian beam into a beam with an annular intensity profile with a null or a zero intensity at its center. In the periphery of the diffraction-limited point spread function (PSF) the STED beam depletes all molecules in the excited state via stimulated emission (postulated by Albert Einstein in 1916). However, where the STED beam has null intensity on its axis the molecules in the excited state can emit photons (spontaneous fluorescence emission) and this fluorescence is detected. Spontaneous emission is only from a region (the null region of the STED beam) that is smaller than the diffraction-limited PSF and therefore superresolution is achieved with STED microscopy. Point-scanning, which involves moving the specimen with respect to the optical axis of the microscope, is used to generate a superresolution image of the specimen.

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_14

14.1.1 Introduction to Molecular Spectroscopy

In this section I review those principles of molecular spectroscopy that provide a basis for understanding light–molecule interactions. For background reading I recommend Boyd (2008); Loudon (2000); Masters (2014); Masters and So (2008); and Valeur and Berberan-Santos (2012).

There are several ways that a molecule can interact with light. These interactions can be divided into linear effects and nonlinear effects. A linear interaction is characterized by a linear relationship between light intensity and the magnitude of the effect. In linear fluorescence emitted light is proportional to the intensity of the excitation. A nonlinear interaction, on the other hand, is characterized by a nonlinear response to incident light intensity. Examples of nonlinear interaction include multiphoton excitation microscopy, second harmonic generation (SHG), third harmonic generation (THG), sum and difference frequency generation, and supercontinuum generation (Boyd, 2008; Masters and So, 2008). At high light intensities, light can alter the optical properties of the medium (e.g., its absorption coefficient, refractive index, and birefringence).

The remaining discussion of molecular photophysical processes is divided into excitation processes and de-excitation processes. Deexcitation processes are further divided into radiative processes that involve the emission of photons and nonradiative processes that do not. The discussion is framed by energy states and transitions between them (as shown in Fig. 14.1).

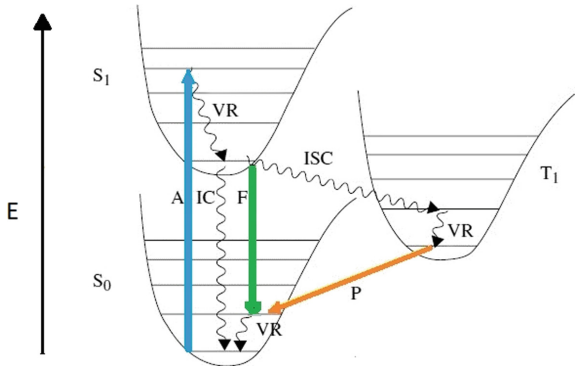


Fig. 14.1 Energy state diagram shows the electronic-vibronic energy levels and energy transitions (radiative and nonradiative transitions) of a molecule. The black upward arrow signifies increasing energy (E). The ground singlet state is S_0, the first excited singlet state is S_1, and the first excited triplet state is T_1. Higher energy singlet and triplet states can also exist; they are not shown. Transitions from absorption (A), shown as a blue upward pointing arrow; fluorescence (F) is shown as a green downward pointing arrow; and phosphorescence (P), shown by an orange arrow from T_1 to S_0. The wavy lines indicate various types of nonradiative transitions: vibrational relaxation (VR), internal conversion (thermal relaxation) (IC) and intersystem crossing (S_1 to T_1 transitions) (ISC). The horizontal lines within each electronic energy level represent the vibronic levels. The vibrational states are associated with rotational energy states which are not shown. In the classical Perrin-Jablonski diagram all of these transitions are depicted as vertical lines. Internal conversion from the T_1 to S_0 and triplet-triplet transitions (T_1 to T_2) transitions are not shown

Molecules, like atoms, have discrete (quantized) energy states. Electronic energy states are associated with a manifold of vibrational energy states, and each vibrational energy state is associated with a manifold of rotational energy states. The energy difference between vibrational states is much greater than that between rotational energy states. Vibrational energy states show decreasing energy separation between adjacent states; they merge to form a continuum in high vibrational states. Electronic states have energy separations that are larger than those of vibrational states, which are larger than those for rotational states. As a general rule, the energy of the photons required to excite these different types of energy states corresponds to the type of energy states; roughly speaking, ultraviolet and visible photons can excite electronic energy levels, infrared photons can excite vibrational states, and microwave photons can excite rotational energy levels.

In Fig. 14.1 electronic states are labeled as singlet electronic states (S_n) or triplet electronic states (T_n), where $n = 0, 1, 2, \ldots$ are the first, second, and higher electronic states (not shown in the figure). Singlet and triplet electronic states denote the spin configurations of electrons in the molecule (Masters and So, 2008; Valeur and Berberan-Santos, 2012). The term electron spin is not related to the classical mechanics of a particle rotating on its axis. Electron spin is a quantum-mechanical property that does not exist in classical mechanics. For a comprehensive introduction to this fascinating quantum-mechanical phenomenon see *The Story of Spin* (Tomonaga, 1997). The singlet state is characterized by molecular orbitals (quantum-mechanical wave functions for a molecule) for which the lowest energy state, the ground state, has a pair of electrons each of opposite spin; the excited singlet state has one electron in each higher energy orbital, and each electron has opposite spin. The net spin is zero. The triplet state is characterized by high-energy molecular orbitals that contain one electron each; however, the spins are parallel. The triplet state has a net spin of 1. The name triplet state refers to three states of equal energy. Most molecules exist in the singlet state; oxygen is a rare exception as it exists in the triplet state. The multiplicity of the energy state of a molecule is given by $2S + 1$, where S is the net spin or total spin angular momentum. Singlet states have a multiplicity of 1. Triplet states have a multiplicity of 3. As shown in Fig. 14.1 the triplet state has a lower energy than the corresponding singlet state, which is consistent with Hund's rule: for a given electron configuration the term with maximum multiplicity has the lowest energy.

Absorption of a photon is the basic process of excitation of a molecule to a higher energy state. If the photon energy equals the energy difference between the ground state energy and the first excited state, then absorption of the photon results in the first excited state of the molecule labeled S_1. This absorption process takes place in 10^{-15} s (Valeur and Berberan-Santos, 2012).

Molecules are capable of a variety of transitions between electronic, vibrational, and rotational energy states (Lewis and Kasha, 1944; Kasha, 1950, 1960). However, there are constraints from quantum mechanics—so-called selection rules—that determine which transitions are allowed (characterized by intense absorption or emission) and which are not allowed, or forbidden (those with very weak

absorption or emission). Selection rules indicate which transitions between energy states will be experimentally observed.

For the process of absorption of a photon there are two classes of selection rules. The first class consists of symmetry-forbidden transitions. Application of group theory to quantum-mechanical transition probabilities shows some transitions to be forbidden (low probability of occurring) due to the symmetry of the initial and final states. Nonetheless, even for symmetry-forbidden transitions vibrational coupling of the two states may allow these forbidden transitions to be observed. The second class of selection rules covers spin-forbidden transitions. Singlet-to-singlet and triplet-to-triplet transitions are allowed, whereas transitions between singlet and triplet multiplicities and between triplet and singlet states are forbidden. However, these selection rules are not absolute. Spin-orbit coupling, which refers to the interaction of spin and the orbital magnetic moments of the electron, can cause weak transitions between the states of different multiplicities. For example, the process of intersystem crossing (ISC in Fig. 14.1) from the first excited singlet state to the first triplet state (S_1 to T_1) can occur via spin–orbit coupling (Valeur and Berberan-Santos, 2012).

As noted above a molecule in the excited state can undergo de-excitation to the ground state via a number of processes: radiative de-excitation processes that involve the emission of photons and nonradiative de-excitation processes that proceed without the emission of photons. Depending on the energy of the photon absorbed the molecule could be in a higher electronic state and at a high vibrational level of that electronic state. For example, the absorption of a photon could induce the transition S_0 to S_1 (transition from the ground state to the first excited singlet state, and from the lowest vibrational energy level in S_0 to a higher vibrational level in S_1).

There are various kinds of radiationless or nonradiative transitions that can occur in complex molecules in an excited state (see Fig. 14.1). Nonradiative transitions occur without the emission of a photon. This is often the case if the molecule is located in a solvent. The excited state may consist of a higher vibrational energy level of the first excited singlet state or of a higher energy singlet state.

Vibrational relaxation (VR) is the transfer of a molecule's vibrational energy to solvent molecules, which results in lowering the vibrational energy to the lowest vibrational level of a given electronic state. Within 10^{-12} to 10^{-10} s the electrons are in the lowest vibrational energy state of a given electron state (e.g., the first singlet state or higher energy singlet states; Valeur and Berberan-Santos, 2012).

Another type of nonradiative transition, internal conversion (IC), is a very fast de-excitation process between two electronic states of the same spin multiplicity (e.g., S_2 to S_1 or S_1 to S_0). During IC, energy is nonradiatively dissipated from the first excited singlet state S_1 to vibrational levels of the singlet ground state S_0. Internal conversion occurs within 10^{-11} to 10^{-9} s (Valeur and Berberan-Santos, 2012).

Intersystem crossing (ISC) is a forbidden process in which energy is transferred nonradiatively from the singlet state to a lower energy triplet electronic state (S_1 to T_1) following vibrational relaxation in the first excited singlet electronic state (Fig. 14.1). Intersystem crossing depends on spin–orbit coupling (Kasha, 1950, 1960).

Intersystem crossing is forbidden but can occur with low probability via spin–orbit coupling. In general, ISC is the radiationless transition from the lowest excited singlet level to the lowest triplet level of the molecule. As shown in Fig. 14.1 the energy of the first excited triplet state (T_1) is lower than the energy of the first excited singlet state (S_1) (Valeur and Berberan-Santos, 2012).

Why are electronic transitions shown in energy-level diagrams (Fig. 14.1) as vertical transitions? In polyatomic molecules the time for electronic transition (10^{-15} s) is much shorter than the time for nuclear rearrangement (10^{-13} s); therefore, during electronic transition from the ground state to the excited state the molecules do not have time to rearrange their nuclear positions and hence remain in approximately the same nuclear configuration.

There is another important nonradiative process that occurs in excited states: resonance energy transfer (RET) (Förster, 1946, 1951; Masters, 2014). RET is the process by which energy is transferred nonradiatively from an excited donor molecule to the excited state of another molecule called the acceptor through intermolecular long-range dipole–dipole coupling. Energy transfer between the donor and the acceptor is mediated by nonradiative energy transfer.

The absorption of a photon transforms the ground state of the molecule into an excited state. This process can result in transforming the molecular structure, with significant consequences for microscopy. Photobleaching, for example, is a process whereby a fluorescent molecule is permanently changed into a nonfluorescent molecule. Photoactivation, by way of contrast, is the process by which light is used to transform a nonfluorescent molecule into a fluorescent one. A molecule can be turned into either a dark state (a nonfluorescent form of the molecule) or into a bright state (a fluorescent form of the molecule) by illumination with light of appropriate wavelength. These processes are integral to superresolution microscopy based on localization with active control (Chapter 15).

Let us now consider radiative transitions. These processes, which occur with the emission of a photon, serve to deexcite the excited state. Of particular importance to us is the process of fluorescence. Fluorescence is typically defined as the emission of photons accompanied by an S_1 to S_0 transition after absorption of a photon (see Fig. 14.1). Fluorescence is composed of photons with an energy that corresponds to the difference in energy between the lowest vibrational level of the excited singlet electronic state and the vibrational level of the ground state. In this section I discuss spontaneous fluorescence or the emission of a photon from the lowest excited singlet state. In the next section I describe the process of stimulated emission that is fundamental to STED superresolution microscopy.

While a molecule can be excited to higher electronic states, emission usually occurs from the lowest singlet state S_1 (Kasha's rule). As stated by Kasha the emitting electronic level of a given multiplicity is the lowest excited level of that multiplicity (Kasha, 1950, 1960). There is very fast relaxation from high vibrational levels to the lowest vibrational levels, and photon emission occurs from the first excited singlet state S_1. The intensity of fluorescence is related to the population of the first excited state, and the STED technique is based on depopulation of the first excited singlet state.

Fluorescence is composed of photons with an energy difference that corresponds to the lowest vibrational level of the excited singlet electronic state S_1 to the vibrational level of the ground state S_0. The mirror image rule states that the absorption and fluorescence emission spectra are similar. That is a consequence of the fact that the ground state and the excited states have similar energy separations in their vibrational energy levels.

The practical utility of fluorescence is the observation that the maximum of the first absorption band and the maximum of the fluorescence emission band are not identical. In fact, this separation of energies measured in wavenumbers is called the Stokes shift. Because of the Stokes shift we can separate excitation light from extremely weak fluorescent emission using simple colored glass filters. Because the excitation intensity is much greater than the emission intensity in fluorescence in the absence of the Stokes shift it would be extremely difficult or impossible to separate and detect the extremely weak fluorescence.

Another important parameter is fluorescence lifetime. Fluorescence lifetime is the time it takes for the intensity of photons emitted from the first excited singlet state of an extremely large number of fluorescent molecules to decay by a factor of $1/e$ following an extremely short pulse of excitation light. Still another important parameter is fluorescence quantum yield, the fraction of excited molecules that revert to the ground state S_0 with the emission of fluorescence photons. A molecule with a high fluorescence quantum yield would exhibit strong fluorescence. The fluorescence lifetime of the excited singlet state is 10^{-10} to 10^{-7} s (Valeur and Berberan-Santos, 2012). This is significantly longer than the time for absorption of the photon. Absorption occurs in about 10^{-15} s.

Of particular interest to us in connection with STED is photosaturation. When the excitation intensity is low, the intensity of fluorescence is proportional to the intensity of excitation. However, with very high intensities of excitation light the number of fluorescent molecules in the ground state are depleted and these molecules are trapped in excited singlet and triplet electronic states. This depletion of molecules from the ground state leads to decreased absorption and therefore decreased fluorescence. Photon absorption can occur if the molecule is in the ground electronic state.

Another radiative process with the emission of a photon occurs from the first excited triplet state T_1 to the vibrational energy levels of the singlet ground state S_0. This is the process of phosphorescence (Lewis and Kasha, 1944). Phosphorescence is a forbidden transition process. As shown in Fig. 14.1, the triplet states have energy levels that are lower than the corresponding set of singlet state energy levels. Therefore, phosphorescence occurs at longer wavelengths than does fluorescence. Electronic transition from the first excited triplet state to the singlet ground state is forbidden because these two electronic states have different multiplicities and de-excitation rates are very slow; indeed, they are on the order of milliseconds to seconds. The excited triplet state is the most stable state of the molecule.

14.1.2 Einstein's 1916 Concept of Stimulated Emission

In this section I introduce stimulated emission, another deexcitation process of the excited state of a fluorescent molecule. In his seminal and prescient publications of 1916 Albert Einstein predicted the necessary existence of stimulated emission (Einstein, 1916a, b). Decades later the invention of the maser in 1954 and the invention of the laser in 1960 validated Einstein's prediction of stimulated emission (Bertolotti, 1999; Loudon, 2000; Maiman, 2018; Masters, 2012; Silfvast, 2004).

Albert Einstein was fascinated by the nature of light throughout his life. To help inaugurate UNESCO's Year of Light in 2015 I wrote an essay "What is Light?" for the *International Commission for Optics* Newsletter (Masters, 2015). "What is Light?" is in English and translated into 15 languages. The article's URL is e-ico.org. Einstein's contributions to the emission and absorption of radiation can be found in Einstein (1916a, b); Masters (2012, 2015); and Masters and So (2008).

What is the relation between stimulated emission that can be initiated by a pulse of intense light with very specific characteristics and the STED technique used in scanning, far-field, superresolution light microscopy? The excited state of a fluorescent molecule can be depleted by a specific high-intensity pulse of light. The key to the superresolution STED technique is that this pulsed depletion light beam in addition to having a specific wavelength and intensity has a specific spatial structure: specifically, the beam has an annular intensity profile with a zero intensity at its center. Only in the vicinity of this null is spontaneous fluorescence detected, and the region of concern can be small compared with the diffraction limit. In the following section I elaborate on how such a spatial structure of the beam in the STED technique is achieved with a spiral waveplate and how this special type of structured light makes theoretically unlimited microscope resolution possible.

In spontaneous emission an atom or molecule in a higher energy state emits a photon with a random phase and a random direction and the molecule returns to the lower energy ground state. In stimulated emission an atom or molecule in an excited energy state interacts with an incident photon, which stimulates the emission of a photon. The emitted photon has the same frequency, phase, polarization, and propagation direction as photons in the incident field (Silfvast, 2004). We say that the emitted wave is *coherent* with the incident wave. Stimulated emission is the inverse of induced or stimulated absorption. Einstein postulated stimulated emission in his publication "Emission and Absorption of Radiation in Quantum Theory" (Einstein, 1916a). His paper initiated a probabilistic approach to quantum physics. It contained his so-called "*A* and *B* coefficients" and his prediction of the processes of stimulated emission.

Einstein made several simplifying assumptions in his publications: a two-state system for an atom made up of an upper state of higher energy (now called the excited state) and a lower state of lower energy (now called the ground state); light can cause atoms to gain or lose energy; and transitions between higher and lower energy states can occur with the absorption or emission of a photon whose energy is equal to the energy difference between the two states.

How did Einstein deduce the existence of stimulated emission? Silfvast in his book *Laser Fundamentals*, 2nd Edition, presents a modern derivation of the equations from Einstein's 1916 publications (Silfvast, 2004). First, Einstein assumed a collection of atoms (molecules) in radiative thermal equilibrium. Under this condition the same number of atoms per unit time will absorb radiation as the number that will emit radiation. The Boltzmann distribution describes the relation between the population densities of the upper or excited state N_u and the lower or ground state N_l

$$\frac{N_u}{N_l} = \frac{g_u}{g_l}\exp(-E_{ul}/kT) \tag{14.1}$$

where g is the statistical weight of a given level, E_{ul} is the energy difference between the two states, k is the Boltzmann constant, and T is the absolute temperature.

Second, Einstein realized that there are two types of transitions—not one—from the excited state to the ground state. One is the emission of radiation without any external factors (emission is independent of radiation density); this is similar to the situation of Rutherford's law of radioactive decay. Today, we call this type of process spontaneous emission, and it is characterized by *Einstein's A coefficient*. The emitted photon can propagate in any direction. The spontaneous (emission) transition probability is Einstein's coefficient A_{ul}, the rate (number per unit time) of spontaneous transitions. The number of spontaneous transitions from the upper state to the lower state per unit time and per unit volume is $N_u A_{ul}$.

The other type of transition requires an interaction between an atom in the excited state and incident radiation: nothing occurs without the presence of incident light. If incident light of the appropriate energy is absorbed and the atom makes a transition from the ground state to the excited state the process is called induced absorption. If the atom is initially in the excited state, then interaction with the photon of appropriate energy will cause the transition from the excited state to the ground state with the emission of a photon. This is the process of stimulated emission, and as suggested above the emitted photon has the same frequency (from the law of conservation of energy), phase, polarization, and propagation direction (from the law of conservation of momentum) as the photon in the incident light. The stimulated processes (i.e., stimulated absorption and stimulated emission) are characterized by *Einstein's B coefficient*. The process of stimulated emission is proportional to the photon energy density $u(\nu)$ at the frequency ν_{ul} and to the population at the specific level. Assuming the proportionality constant for stimulated or induced transitions is B (Einstein's B coefficient), the number of upward transitions (lower energy state to upper energy state) per unit volume per unit time per unit frequency is $N_l B_{lu} u(\nu)$. The number of downward transitions (upper energy state to lower energy state) per unit volume per unit time per unit frequency is $N_u B_{ul} u(\nu)$. The constants A_{ul}, B_{ul}, and B_{lu} are Einstein's A and B coefficients.

Einstein's next assumption was the principle of microscopic reversibility (a.k.a. the principle of detailed balance). It states that for radiation in equilibrium the rate

of energy transfer from lower energy states to higher energy states must equal that from upper to lower states. Since by assumption the population densities N_u and N_l are in radiative thermal equilibrium downward radiative flux must equal upward radiative flux between the two levels

$$N_u A_{ul} + N_u B_{ul} u(v) = N_l B_{lu} u(v) \tag{14.2}$$

From this equation and (14.1) it follows that

$$g_l B_{lu} = g_u B_{ul} \tag{14.3}$$

Therefore, the probability of stimulated emission is identical to the probability of stimulated absorption, and

$$B_{ul} = \frac{c^3}{8\pi h v^3} A_{ul} \tag{14.4}$$

These are the relations between stimulated emission and absorption coefficients B_{ul} and B_{lu} and their relationship to the spontaneous emission coefficient A_{ul}. The ratio of stimulated emission to spontaneous emission from level u is

$$\frac{B_{ul} u(v)}{A_{ul}} = \frac{1}{\exp(h v_{ul}/kT) - 1} \tag{14.5}$$

It follows that stimulated or induced emission is significant and dominates spontaneous emission only for temperatures at which kT is equal or greater than the photon energy $h v_{ul}$. Silfvast (2004) deduces several important relations from (14.5). First, when the ratio $h v_{ul}/kT = \ln 2$, or 0.693, both sides of (14.5) equal 1. This means that the rates of stimulated emission and spontaneous emission are equal. For electronic transitions in the green region of the spectrum this condition occurs at a temperature of 33,500 K. Thus, stimulated emission would predominate at temperatures that occur in stars. However, on earth stimulated emission in lasers predominates over spontaneous emission and the ratio in (14.5) is much greater than 1 (Silfvast, 2004).

Einstein postulated the existence of stimulated emission so that energy levels in the atom in equilibrium with the radiation field can be characterized by the Boltzmann distribution and at the same time be consistent with the Planck radiation law. Without stimulated emission Einstein's B coefficient would be zero and the Boltzmann distribution would not characterize states at thermal equilibrium. Assuming stimulated emission is at thermal equilibrium the number of atoms gaining energy must equal the number losing energy. From this reasoning Einstein deduced that the probabilities of induced absorption and stimulated emission are equal: $B_{ul} = B_{lu} = B$.

Einstein suggested that the coefficients A and B could be calculated if a new version of electrodynamics and mechanics were available that agrees with the

quantum hypothesis (Einstein, 1916a). This prediction was fulfilled in 1927 when Dirac used his version of quantum mechanics to derive Einstein's *B* coefficient and Einstein's *A* coefficient (Dirac, 1927).

14.1.3 *Spiral Phase Plate Conversion of a Gaussian TEM_{00} Laser Beam to a Helical Beam with an Annular Intensity Profile and Zero Intensity at the Center*

Central to the practical implementation of stimulated emission depletion microscopy in a STED microscope is a spiral phase plate (sometimes referred to as a vortex phase mask) that converts a Gaussian TEM_{00} laser beam into a beam with an annular intensity profile that has a null in the electromagnetic field intensity at its center. In this section and the next one, I use the dissimilar notation that is used in each author's publication and in figures from websites.

The spiral phase plate schematically illustrated in Fig. 14.2 is a transmissive plate whose thickness increases linearly with azimuthal angle over one full turn. The step demarking the beginning and end of the spiral introduces an optical path length difference of exactly one wavelength of the incident monochromatic laser beam corresponding to a phase shift of 2π radians in light (Andrews, 2015). It is this null on its axis that permits emitted photons from nondepleted excited states of on-axis molecules to be restricted to a region or volume smaller than the conventional diffraction limit. These molecules undergo spontaneous fluorescence emission. In that manner STED superresolution optical microscopy is achieved. The spiral phase plate is sufficiently large that it effectively modulates the entire Gaussian STED beam, which is centered on the axis of the phase plate.

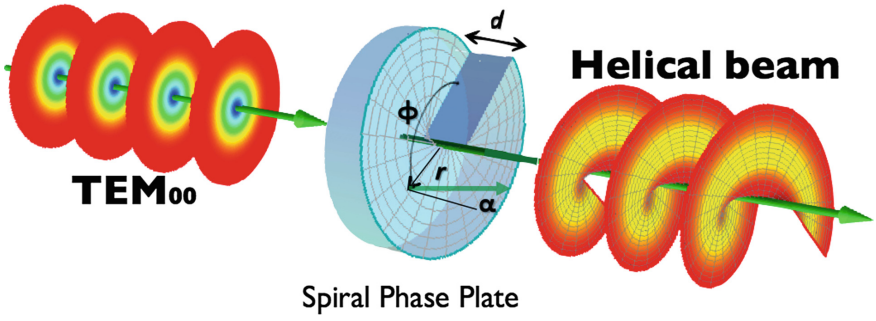


Fig. 14.2 Spiral phase plate conversion of a Gaussian TEM_{00} laser beam to a helical beam with an annular intensity profile and zero intensity at the center. On the left side is the planar wavefront of a Gaussian laser beam in TEM_{00} mode, in the center is the spiral phase plate, and on the right side is the modulated wavefront in the shape of a helical beam. The spiral phase plate step height is shown by **d** and the azimuthal angle is shown by **ϕ**. The axis of the helical beam is a null in the field intensity. This image is made by author, E-karimi, and licensed under Creative Commons Attribution-share Alike 3.0

The spiral phase plate has a specific design, its thickness increases circumferentially around the plate. Its surface resembles one turn of a spiral staircase.

The physical dimensions of the spiral phase plate depend on the wavelength of incident light. Therefore, for each wavelength to be used a specific spiral phase plate must be microfabricated (Oemrawsingh et al., 2004; Sueda et al., 2004). The optimal step height, the distance between the lowest and highest part of the spiral phase plate d of a spiral phase plate for light with wavelength λ is given by

$$d = \frac{l\lambda}{(n-1)} \tag{14.6}$$

where l is an integer (topological charge), and n is the refractive index of the medium from which the plate is formed (Andrews, 2015).

Next, I provide a heuristic explanation of why the spiral phase plate transforms a Gaussian beam plane wave into a beam with an annular intensity distribution (i.e., a null or an on-axis zero intensity). My perusal of current books and publications on STED microscopy indicates the absence of an in-depth explanation of the operation of spiral phase plates. This section is my attempt to fill this lacuna. What is a spiral phase plate and how does it work? A planewave from an STED laser is incident on a spiral waveplate, which causes a change of phase [0; 2π] across the laser beam. The high numerical aperture microscope objective focuses the phase-shifted laser beam and at the focus there is destructive interference. Destructive interference results in a null or zero intensity in the center of the STED beam.

This annular intensity profile requires that the center of the STED laser beam and the center of the spiral phase plate are correctly aligned. Three factors can result from the absence of null intensity in the center of an annular beam: (1) the absence of alignment between the center of the Gaussian STED beam and the center of the spiral waveplate, (2) optical aberrations, and (3) light scattering.

14.1.4 Vortex Beams and Singular Optics

There is growing interest in the topic of vortex beams and the branch of optics called singular optics. The spiral phase plate converts the incident Gaussian beam into an optical vortex (sometimes called a screw dislocation or phase singularity), a type of optical singularity that has a spiral phase wavefront around a singularity point where the phase is undefined. An optical vortex is a point of zero intensity in an optical field.

A plane wave that has a singularity is defined by a phase change of $2l\pi$, where l is an integer called the topological charge, around a point in a transverse plane. Because the phase at that point is undefined there is zero intensity and the beam has a doughnut or annular intensity profile in the transverse plane (Gu, 2000).

Another name for a spiral phase plate is a vortex phase plate. In the following discussion I explain the origin of the terminology and some of the physics of vortex beams and their applications to STED microscopy. First, I present a qualitative description of the operation of a spiral phase plate. A Gaussian beam with a planar wavefront is transmitted through a spiral phase plate. The resulting beam undergoes a phase change and the phase front of the resulting beam has a spiral aspect. The resulting beam carries angular momentum. The second new property of the beam emerging from the spiral waveplate is that the beam has a null of the field amplitude on its axis. This is required because the field would have to simultaneously have multiple phase values if it was not null or zero on the beam axis since the axis is at all possible circumferential angles.

Rozas (1999, p. 35) wrote in his "Generations and Propagation of Optical Vortices," a dissertation submitted to the Faculty of Worcester Polytechnic Institute for the degree of doctor of philosophy in physics: "An important attribute of the vortex is that its core remains dark as the beam propagates. This may be understood from the point of view of destructive interference between rays diffracted into the core. Let us consider a circle of infinitesimal radius centered on the vortex core. For each (arbitrary) point on this circle having a phase Φ there is a point with phase $\Phi + \pi$, symmetrically located with respect to the center of the vortex. According to Huygens' principle (M. Born and E. Wolf, *Principles of optics*, 6th ed. Cambridge U. Press, New York, 1997) all points of the circle will radiate, giving rise to destructive interference owing to the π phase difference between symmetric points."

For a rigorous and comprehensive background I recommend perusal of these books: *Structured Light and Its Applications*, 1st Edition: *An Introduction to Phase-Structured Beams and Nanoscale Optical Forces* (Andrews, 2008); *The Angular Momentum of Light* (Andrews and Babiker, 2013); *Fundamentals of Photonics and Physics*, Volume I (Andrews, 2015); *Optical Angular Momentum* (Allen et al., 2003); and *Optical Tweezers: Principles and Applications*, 1st Edition (Jones et al., 2016).

Another publication I recommend is "Helical-wavefront laser beams produced with a spiral phase plate" (Beijersbergen et al., 1994) in which the authors demonstrate experimentally that a spiral phase plate can convert a Gaussian TEM_{00} laser beam into a helical wavefront beam with a phase singularity (zero intensity) on its axis.

The following publications provide an understanding of the field of optical vortices and their use in STED microscopy. The comprehensive chapter by Wisniewski-Barker and Padgett on orbital angular momentum emphasizes the mathematical analysis, physics, and optical devices that constitute this rapidly developing field of optical vortices (Andrews, 2015). Other significant related papers and dissertations that provide deeper understanding of the relevant optics and applications to STED microscopy include the paper "Orbital angular momentum of light and the transformation of Laguerre–Gaussian laser modes" (Allen et al., 1992); the review article "Singular Optics: Optical Vortices and Polarization Singularities" (Dennis et al., 2009); Kellner's dissertation "STED microscopy with Q-switched microchip lasers" (2007); Klar's dissertation "Progress in stimulated emission

depletion microscopy" (2001), especially the sections on resolution increase by offset beams and phase retardation masks; Schönle's dissertation "Point spread function engineering in fluorescence spectroscopy" (2003); Willig's dissertation "STED microscopy in the visible range" (2006); Schoonover's dissertation "Studies in Singular Optics and Coherence Theory" (2009); Soskin and Vasnetsov's review article "Singular optics" (2001); and Sheppard and Choudhury's article "Annular pupils, radial polarization, and superresolution" (2004).

Many problems in diffraction and focusing can be solved using paraxial optics and scalar theory. A paraxial light ray has a small inclination to the optical axis. When scalar theory is insufficient, then the vectorial theory of diffraction is required. A modern comprehensive treatment of vectorial and scalar theory of diffraction and focusing is the subject of the new book *Imaging Optics* (Braat and Török, 2019). The previous standard for these topics is the book *Principles of Optics*, 7th Expanded Edition (Born and Wolf, 1999).

The in-depth study of optical vortices and singular optics began with the discovery that wavefronts can contain dislocation lines (Nye and Berry, 1974). In their publication "Dislocations in wave trains," Nye and Berry described ultrasonic pulses reflected in air from a rough surface and observed that the scattered waves contained dislocations. They concluded that dislocations are mathematically lines along which the phase is indeterminate, a condition implying that the wave amplitude is zero.

Some years later a new development advanced the field from ultrasound frequencies to the optical domain. Vaughan and Willets formed a phase singularity in a laser beam by combining high-order Hermite–Gaussian modes, which are solutions of the paraxial wave equation expressed in Cartesian coordinates (Vaughan, and Willetts, 1979).

Allen et al. (1992) is a seminal paper describing how light beams with helical phase fronts within the paraxial approximation carry a well-defined orbital angular momentum. Allen et al. formed beams that carry orbital angular momentum using two cylindrical lenses that transform a Hermite–Gaussian beam into a Laguerre–Gaussian beam. Mathematical analysis demonstrated that orbital angular momentum of light occurs when the phase increases by a multiple of 2π along a closed contour that encircles a dislocation line where phase is undetermined and the intensity is zero (Nye and Berry, 1974). Optical vortices are a prime example of singularities in optics (Dennis et al., 2009; Schoonover, 2009). Another term for an optical vortex is a phase singularity, which is a point of zero intensity.

Optical vortices can occur as special solutions (called Laguerre modes) to the wave equation expressed in cylindrical coordinates (D'Alessandro, and Oppo, 1992). (Allen et al. 1992) demonstrated that a light beam with orbital angular momentum can be described using a phase in the transverse plane of $\phi(\theta) = l\theta$, where θ is the angular coordinate, and l is an integer. Furthermore, the local momentum stream is not in the propagation direction, it is directed along the normal to the phase fronts, and has a discrete value of $l\hbar$ per photon for beams with helical phase fronts characterized by an $\exp(il\theta)$ azimuthal phase dependence, where l is the azimuthal index (Allen et al., 1992; Wisniewski-Barker and Padgette, 2015). The orbital angular

momentum of $l\hbar$ per photon is not the total orbital angular momentum; it is the component of total orbital angular momentum in the direction of the beam axis (Allen et al., 1992; Andrews, 2008). Light beams that have helical phase fronts have special properties: the momentum is not oriented parallel to the direction of beam propagation, but is tilted with respect to the beam axis (Padgett and Allen, 1955; Turnbull et al., 1996).

While the angular momentum of vortex light beams is theoretically important and has many experimental applications, I frame the discussion on the use of spiral phase plates to form a focused beam with an annular intensity profile and, most importantly, with an on-axis zero intensity at the center of the beam.

Next, I explain formation of the beam profile that has an annular or doughnut-shaped intensity pattern with zero intensity at it center (i.e., the required intensity profile for the STED beam that is central to its implementation in STED microscopy).

A Laguerre–Gaussian beam has a scalar optical singularity called an optical vortex. Its unique properties arise from its phase term, a simple function of the azimuthal angle (Goodman, 2017). The optical vortex is a point in the transverse plane where light intensity is zero and where phase is indeterminant (i.e., it is not uniquely defined). Furthermore, the topological charge of the vortex or its azimuthal index is labeled l, the multiple of 2π that phase advances over a closed loop of the vortex (Allen et al., 2003). Topological charge, which is sometimes denoted by m or l, is the number of 2π cycles in a 360° rotation of the vortex phase plate. The typical spiral phase plate that is used in STED microscopy has one 2π cycle that covers the entire 360° of the surface; therefore, topological charge is 1 ($l = 1$ or $m = 1$). If there is an increase in topological charge (e.g., from $1m$ to $2m$), then there is an increase in the area of the central dark region where intensity is null or zero (Gu, 2000). This effect is illustrated in Fig. 14.3 where topological charge is labeled with m.

Integer l is the same as the number of 2π phase cycles ($2\pi l$) around the optical vortex centered on the beam axis. This follows due to the term $\exp(il\theta)$ and is consistent with the helical form of the wavefront. Because the radiation intensity of a Laguerre–Gaussian beam must be zero at the beam axis for a spiral phase structure the intensity pattern results in an annular or doughnut-shaped pattern; thus, these beams are called doughnut beams or optical vortices. The name Laguerre–Gaussian refers to the complex functions formed from Laguerre polynomials used to mathematically describe a Laguerre–Gaussian beam (Jones et al., 2016).

The origin of the name optical vortices or vortex beams stems from helical phase fronts having a phase singularity running down their center, an intensity null or zero intensity surrounded by a phase change (Berry et al., 1979; Berry, 2004). For example, Laguerre–Gaussian beams carry orbital angular momentum in their tilted wavefront, the center position of the vortex beam where there is no energy or momentum. The mathematical similarity between helically phased beams and superfluid vortices led to these phase singularities being named optical vortices (Coullet et al., 1989).

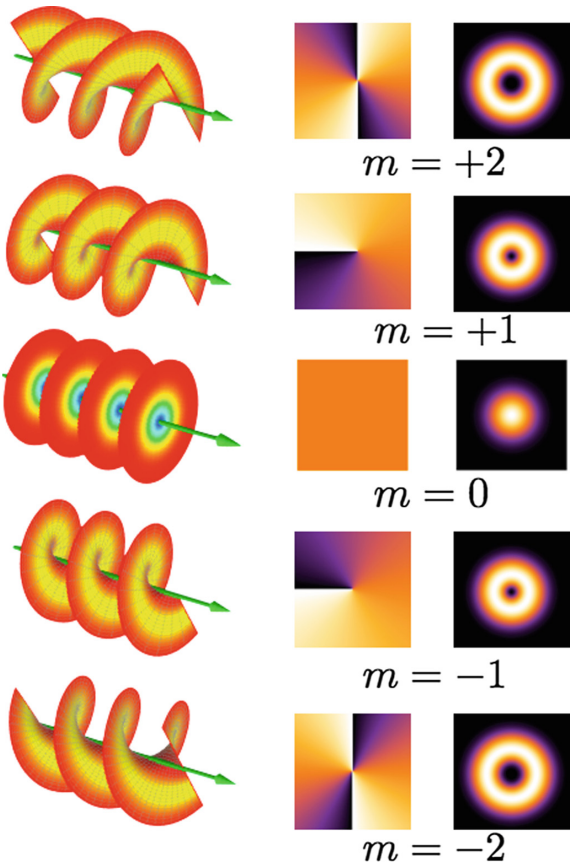


Fig. 14.3 Different modes, four of which are optical vortices. The columns show the helical structures, phase front, and intensity of beams. When the topological charge of the beam is $m = \pm 1$, $m = \pm 2$, there is a null or zero intensity on its axis. When $m = 0$, the beam has a Gaussian intensity profile. This image is made by author, E-karimi, and licensed under Creative Commons Attribution-share Alike 3.0

Keeping all this in mind I now want to discuss the early seminal publication "Helical-wavefront laser beams produced with a spiral phase plate" (Beijersbergen et al., 1994). The authors demonstrated experimentally that a spiral phase plate can convert a TEM_{00} laser beam into a helical wavefront beam with a phase singularity or zero intensity on its axis. This was the first early demonstration of a spiral phase plate that converts a Gaussian beam into a vortex beam with an annular intensity profile and null or zero intensity on its axis. It is likely that this seminal advance in the use of a spiral waveplate was not lost on the inventors of STED microscopy. However, the first decade of publications from the Hell group neither used a spiral phase plate in their STED microscope nor did they cite the prescient publication of Beijersbergen et al., (1994).

There are alternatives to using a spiral phase plate to form a beam with an annular intensity profile and zero intensity at its center (e.g., holographic elements using diffractive optical elements and spatial light modulators). Another approach involves using a superposition of modes: Hermite–Gaussian modes and Laguerre–

Gaussian modes form complete orthonormal sets, and an arbitrary amplitude distribution can be described by an appropriate complex superposition of modes from either set. By getting three or more plane wave components of similar intensity to interfere a variety of vortices in any field cross section can be obtained (O'Holleran et al., 2006; Masajada and Dubik, 2001).

14.2 Stimulated Emission Depletion Microscopy

In previous sections I described the basics of molecular spectroscopy including the definition and explanation of terms, explanation of Einstein's prescient deduction of the process of stimulated emission, and a comprehensive discussion of the physical basis of a spiral phase plate. Now we are in a position to discuss the workings of a scanning, far-field, superresolution STED microscope.

14.2.1 Historical Perspectives

The attribution of credit for an invention is a complex, sometimes contentious, and often difficult process. Who can legitimately take the title of inventor? What about the occurrence of independent, simultaneous invention? Is credit based on the first person to present the ideas at a scientific conference, to publish a paper, or to file a patent? Is the patent filing date, or the date of patent publication, or the date the patent was granted the deciding factor for the attribution of credit? Is the granting of a patent sufficient to attribute credit to an author or should construction of a working instrument be required for the attribution of credit? Patents can be contested. The existence of prior art could result in a patent being declared invalid. For readers who wish to gain a deeper understanding of the patent process I recommend the book *Consider a Spherical Patent, IP and Patenting in Technology Business* (Gortych, 2014).

Taking a hierarchical view of inventions we start with the concept of the invention. Second, there is validation and proof of concept that may be in the form of a working model or instrument. Third, there is the new knowledge obtained using the new instrument to acquire data in response to scientific questions posed. This third aspect may take a short time or many decades to occur. I am thinking particularly of the large interval between Einstein's deduction of stimulated emission and the invention of the laser. Another example is Rabi's precise measurements of the magnetic properties of nuclei and the development of clinical magnetic resonance instruments for the imaging of soft tissues in the human body.

Among the numerous examples of contested attributions are disputes over the following inventions: Raman scattering, the laser, the confocal microscope, the scanning mirror confocal microscope, the multiphoton excitation microscope, and the use of dichroic mirrors in fluorescence microscopes (Masters, 1996, 2001, 2003, 2006, 2009).

I now present three patents that relate to superresolution, far-field, scanning optical microscopy. I describe the key concepts from a variety of sources leaving it up to readers to make their own assessments of the validity of invention priority. The inventors I discuss in the following paragraphs have been granted many patents. I recommend reading their full patents and urge readers to search the European and US Patent Offices using their search engines to obtain the patents on superresolution. Perusal of these patents will yield insight into the prior work of others and to the thinking of the inventors at the time the patents were filed.

On April 10, 1986 Victor Okhonin filed a patent in the former Soviet Union (USSR), SU 1374922, "Method of investigating specimen microstructures." The Russian patent can be found at http://patents.su/4-1374922-sposob-issledovaniya-mikrostruktury-obrazca.html (accessed on April 19, 2019). The English translation of this patent undertaken by Victor Okhonin himself can be found at https://www.researchgate.net/profile/Victor_Okhonin/publication/272021175_STED_Priority_1986_Eng_Transl/links/54d8ca860cf2970e4e793c8b.pdf?origin=publication_detail English translation (accessed April 19, 2019).

Okhonin's patent was subsequently published on July 30, 1991 in *Soviet Patents Abstracts*, Section EI, Week 9218, Derwent Publications Ltd., London, GB; Class S03, p. 4. This patent was cited in the following US patents: US Pat. No. 5,394,268 A (1993); Field synthesis and optical subsectioning for standing wave microscopy invented by Frederick Lanni, D. Lansing Taylor, and Brent Bailey; and US Pat. No. RE38307 E1 (1995); Method and apparatus for three-dimensional microscopy with enhanced resolution (SIM) invented by Mats G. L. Gustafsson, John W. Sedat, and David A. Agard.

Okhonin's patent involves the formation of standing waves of light. He explains that the fluorescence of the specimen is quenched [he uses the words forced transition, which I assume means stimulated emission] everywhere, except where the intensity of the quenching beam is zero or nearly so. The patent makes the critical point that it is possible to decrease the size of the region where quenching does not occur by increasing the intensity of light that forms the standing wave pattern.

Readers can see that the patent describes something akin to stimulated emission depletion (STED) microscopy. However, something critical is missing. A key component of STED, introduction of nonlinearities via saturation, is not explicitly described in Okhonin's patent. This component in principle yields optical resolution unlimited by diffraction.

On July 15, 1994 Stephen C. Baer, of Cambridge, Massachusetts, filed a patent, "Method and apparatus for improving resolution in scanned optical system." US Pat. No. 5,866,911 was published on February 29, 1996 as International Application WO 9606369A2 and issued on February 2, 1999. Perusal of US Pat. No. 5,866,911 yields the following details. On the first page of the patent under the title *Other Publications* the following publications are listed: S. Hell, "Improvement of Lateral Resolution in Far-Field Light Microscopy by Using Two-Photon Excitation with Offset Beams," *Optics Communications*, vol. 106, Amsterdam,

The Netherlands, 1994, pp. 19–24; and S. Hell and J. Wichmann, "Breaking the Diffraction Resolution Limit by Stimulated-Emission-Depletion-Fluorescence Microscopy," *Optics Letters*, vol. 19, Washington, U.S.A., 1994, pp. 780–782. In this list of related publications Baer also cites: Arimoto and Kawata, "Laser-Scan Fluorescence Microscope with Annular Excitation," *Optik*, vol. 86, Stuttgart, Germany, 1990, pp. 7–10, No. 1. This patent does not cite Okhonin's patent SU 1374922 published on July 30, 1991.

Details of Hell and Wichmann (1994) are discussed in subsequent paragraphs of this section. In their publication there is no discussion of a quenching beam with zero intensity at its center on axis. The STED quenching beam is formed by overlapping Gaussian beams that do not possess any intensity zeros.

In contrast to Hell and Wichmann (1994), Baer's patent is based on a quenching beam composed of overlapping Airy disks that are situated around a zero point at such a distance that a section of the first minima (zero intensity) of the Airy disks intersect with the zero intensity point. In Hell and Kroug (1995), which introduced ground state depletion (GSD) microscopy, the authors similarly formed a depletion beam with zero intensity by overlapping the Airy disks that were situated in such a way that their zero rings intersect at the zero point.

There are two important points in the Baer patent that critical readers should note. First, Baer's patent refers to a quenching beam that has a central intensity minimum. This minimum is slightly above zero intensity due to light scattering, which can be reduced but not eliminated. There is no central zero intensity. Second, in the Baer patent and similarly in the Okhonin patent there is no explicit statement that fluorescence saturation forms nonlinearities that can bring about diffraction-unlimited resolution. As far as I am aware neither Okhonin nor Baer constructed a working microscope that demonstrated the instrument's capabilities as described in their separate patents. The third patent I present for discussion is US Pat. No. 5,731,588 A (filed February 11, 1995 and published March 24, 1998); Process and device for optically measuring a point on a sample with high local resolution invented by Stefan Hell and Jan Wichmann.

The authors state the principle behind their invention in their patent and provide details for implementation of their concept. They provide a method to narrow the point spread function of excitation light and thereby achieve enhanced lateral and axial spatial resolution. The microscope uses two pulsed lasers, one for the excitation of fluorophores and one for stimulated emission—pulses from both lasers are in temporal sequence. The excitation light pulse is in the range 10^{-15} to 10^{-9} s and the pulse length for stimulated emission light is 10^{-12} to 10^{-9} s.

The next point is critical since it was missing from both Okhonin's and Baer's patents. The light source for the stimulated emission beam must be sufficiently intense to induce a nonlinear relation between the intensity and the energy state occupied by the fluorescent molecule. Nonlinear effects induced by the high intensity serve to sharply limit the point spread function resulting in enhanced spatial resolution. To summarize, Hell and Wichmann's patent does not feature a stimulated emission depletion beam with zero intensity at its center [that would

come in subsequent patents]. Nevertheless, the important possibility of a nonlinear connection between intensity and the specimen occupying an excited state is suggested.

Hell and Wichmann's patent cites the following patents of Stephen C. Baer: US Pat. No. 5,952,668 (filed August 28, 1997 and published September 14, 1999); Resolution in microscopy and microlithography; US Pat. No. 6,259,104 (filed June 28, 1999 and published July 10, 2001); Superresolution in optical microscopy and microlithography; and US Pat. No. 6,903,347 B2 (filed July 9, 2001 and published June 7, 2005); Superresolution in microlithography and fluorescence microscopy. This patent of Baer's describes STED far-field, scanning fluorescence microscopes with parallelized regions of STED beams.

Having completed my discussion of the above three patents with an emphasis on their differences and omissions, I now describe some of the groundbreaking early publications from the group of Stefan Hell. These publications give readers an insight into technical problems and their solutions. The sequence of subsequent publications from the Hell group actually reflected the temporal development of their ideas on superresolution optical microscopy allowing us to reconstruct the trajectory of their achievements.

I am particularly interested in development of the link between zero intensity in the STED annular beam and its potential to achieve theoretically unlimited resolution in the far-field light microscope. A second topic of interest is introduction of the spiral phase plate, which converts a Gaussian beam into a beam with an annular intensity profile that has null or zero intensity at its center.

In 1994 Hell worked at the University of Turku in Finland. He published a paper titled "Improvement of lateral resolution in far-field light microscopy using two-photon excitation with offset beams" based on his modeling and calculations (Hell, 1994). Hell proposed exciting a fluorescent molecule with two photons, each from separate offset light beams. Simulations showed that the point spread function (PSF) was smaller than the normal two-photon PSF. This modeled enhancement only occurred in the direction of beam displacement. Modeling showed significant resolution enhancement: if offset in the lateral plane was 5.3 optical units, then the two-photon PSF was 59% smaller than the comparable value without any beam offset. The simulation showed it was possible to improve resolution by a factor of 2. Interestingly, Hell stated the potential of specimen light toxicity from the high light intensities required for two-photon excitation microscopy.

In 1994 Hell and Wichmann working at the University of Turku published their technique for "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion [later denoted STED] fluorescence microscopy" (Hell and Wichmann, 1994). This resolution enhancement of 4.5× was only in one dimension; however, it led the way for future developments of STED. The basic principle underlying STED microscopy can be explained in the following way. One beam excites a fluorescent molecule, absorption of a photon forces transition from the ground state to a higher energy electronic state, and after very rapid vibrational relaxation the molecule is in the first excited singlet state. The function of the STED beam is to induce emission from the first excited state and therefore to deplete the

excited state prior to the occurrence of spontaneous fluorescence. The authors obtained superresolution, far-field light microscopy by depleting the first excited state and thereby inhibiting spontaneous fluorescence in the peripheral regions of the excitation PSF (Hell and Wichmann, 1994).

How was their concept implemented in a STED microscope? The STED laser is a picosecond pulsed laser operating at 100 MHz that has a peak power of 1300 MW/cm^2. The STED beam is divided into two beams and pulses are temporally separated from pulses of the excitation beam, which has a Gaussian beam profile. Both offset STED beams and the excitation beam are focused on the specimen and spontaneous fluorescence is detected using a point detector. Since alignment of the STED beams and the excitation beam is critical the beams are stationary with respect to the specimen. It is by moving the specimen that we get a two-dimensional image of it. Lateral resolution of the specimen is enhanced in one lateral direction. The authors observed that as the intensity of STED beams increases so the depleted area of the focal region becomes smaller with steeper edges and the PSF decreases in size, which means resolution is enhanced beyond the diffraction limit (Hell and Wichmann, 1994).

The next publication I discuss gives a clear description of how the STED technique works to reduce the effective PSF of a diffraction-limited excitation beam and therefore achieve superresolution or a resolution that exceeds diffraction-limited resolution. In 1999 Klar and Hell working at the Max Planck Institute in Göttingen published a paper titled "Subdiffraction resolution in far-field fluorescence microscopy" (Klar and Hell, 1999). Although the experiments they described achieved an enhancement in lateral resolution, it was an improvement in only one direction. Nevertheless, their achievement is notable and serves as a useful example and explanation of the STED technique. The authors were able to inactivate fluorescence from an outer region of the focus using stimulated emission. How did the authors do this? A single mode-locked Ti:sapphire laser was the light source. Laser output (766 nm) was divided into two beams, one was frequency-doubled to form the excitation beam (383 nm) and the other was the STED beam (766 nm). A $\lambda/2$ waveplate in the STED beam ensured the two beams had the same polarization. After excitation fluorescent molecules vibrationally relax into the lowest singlet state (S_0) and the STED pulse quenches this state. Optical delay causes STED pulses to be at the focus a few picoseconds after excitation pulses.

In their publication the authors validate that the STED technique achieved a resolution enhancement in the y-direction by demonstrating that the full width half maximum (FWHM) was 106 nm ($\lambda/3.6$) while in the absence of a STED beam the FWHM was 150 nm. What is significant here is that the authors suggest that a second beam in the x-direction or in the z-direction (axial direction), or the application of a STED beam with an annular shape (doughnut shape), could enhance resolution in other directions. The authors did not explicitly explain that the words used to describe the proposed intensity profile of the STED beam (i.e., annular shape or doughnut shape) tacitly implies null or zero intensity at the center of the beam.

The next advance in the development of the STED technique and its application to living cells was the publication titled "Fluorescence microscopy with diffraction

resolution barrier broken by stimulated emission" (Klar et al., 2000). This advance was based on the use of stimulated emission (STED) to quench or deactivate fluorescent molecules in the first excited singlet state at the edge of the focal spot. In the axial or z-direction the spot size of excitation was six times less than that of diffraction-limited imaging. In the lateral plane, the x–y direction, the fluorescence spot had a diameter of 90–100 nm. The authors had successfully demonstrated that the STED technique works with living cells.

How did their STED microscope achieve this significant enhancement in spatial resolution? Since previous publications from the Hell group describe how the STED technique can be used to achieve enhanced spatial resolution I will not repeat the discussion's key points. Instead, I will focus on how the authors achieved an annular-shaped STED beam.

The design of the early STED microscope did not use a vortex phase plate (Klar et al., 2000). Instead of a vortex phase plate the authors used a STED phase plate made from a glass plate that had a thin layer of MgF_2 in a central round region. This layer of MgF_2 introduced a delay of $\lambda/2$ with the amplitude of light having a reversed sign. How was the annular intensity pattern with a central minimum formed? The authors explained that one-half of the amplitude in the entrance pupil was phase-reversed (Klar et al., 2000). The focused STED beam brought about destructive interference at the focal spot, which caused the annular STED beam to have an intensity minimum at its center. This null or zero intensity at the center of the annular beam is coaligned with the intensity maximum of the laser beam used for exciting the fluorophores. In the ideal case a STED beam has zero or null intensity at its center. Aberrations can lead to the distortion of STED beam wavefronts with the result that the center of the annular STED beam is not zero intensity.

The STED beam had a PSF with high-intensity regions above and below the focal plane. Wherever the STED beam had high intensity, it forced fluorescent molecules in the first excited singlet state to transition to the ground state—the driving process being stimulated emission. Stimulated emission depleted the excited state in the periphery of the focused (diffraction-limited) excitation beam; however, in the center of the STED beam there was zero intensity and therefore no stimulated emission occurred. Excited fluorescent molecules not in the region of the high-intensity STED beam underwent spontaneous fluorescence. To form an image of the specimen the authors translated the two beams (the excitation beam and the STED beam) over the specimen and detected the emission of fluorescence. This publication also describes the correlation between the intensity of the STED beam and enhanced resolution; in principle, as the intensity of the STED beam increases so does the possibility of unlimited spatial resolution.

The PSF of focused light can be decreased by fluorescence saturation. To achieve such a saturation there are two requirements: first, there must be an intensity null or zero in the center of the focal region and, second, fluorescent molecules must have a saturable linear transition that is reversible. At sufficiently high intensities of the STED beam saturation will occur; at even higher intensities the ratio of stimulated to spontaneous emission is increased. As a consequence the

focal volume is reduced and resolution is enhanced. Correspondingly, decreased focal volume leads to unlimited resolution. In practice, the achievable resolution is constrained by phototoxicity induced by high light intensities.

14.2.2 Stimulated Emission Depletion Foundations, Instrumentation, and Applications

Stimulated emission depletion (STED) microscopy uses stimulated emission to deplete the number of fluorescent molecules in an excited singlet state (Eggeling et al., 2015). This only occurs in a particular zone of the focal area. If the STED beam has an annular intensity profile with null or zero intensity at its center, then there will be no depletion in that region at the center of the STED beam and excited molecules in the center of the STED beam can spontaneously fluoresce.

However, the excitation beam and the STED beam are diffraction limited. That being the case how can resolution enhancement be brought about? As I previously explained in the STED technique fluorescent emission is only from the central part of the focal area. Therefore, the PSF is smaller and consequently lateral resolution is enhanced and can exceed the diffraction limit.

The intensity of the STED beam is another important parameter linked to the degree of resolution enhancement that can be achieved. Increasing the intensity of the STED beam reduces the size of the excitation PSF and therefore brings about enhanced resolution. This enhancement is theoretically unlimited. However, the resolution that can be achieved is limited by phototoxicity in live specimens and optical aberrations that distort the wavefronts of the STED beam. The STED depletion curve that relates residual spontaneous fluorescence emission with the applied intensity of the STED beam is nonlinear (Hell, 2003, 2009; Schönle, 2003; Schönle et al., 2008). At high intensities of the STED beam there is almost no spontaneous emission of fluorescent molecules. The stimulated emission process exploited in STED is a reversible saturated process.

The STED microscope uses two critically aligned pulsed laser beams that output two synchronized trains of pulses (Fig. 14.4). Correct alignment of the excitation beam and the STED beam is critical: the center of the STED beam with zero intensity must be collinear with the maximum of the excitation beam intensity. The excitation beam, which is composed of 0.2 ps pulses in the visible region, induces single-photon excitation of fluorescent molecules located in the focal volume. The STED beam is composed of pulses in the near-infrared region. Each excitation pulse is followed by a STED pulse. Stimulated emission induced by the STED beam forces electronic transition of excited molecules to an upper vibrational level of the ground state. Very rapid vibrational relaxation serves to prevent the STED beam from re-exciting fluorescent molecules.

There are temporal constraints on the STED pulse. The STED pulse is shorter than the lifetime of the fluorescent molecule, but must be longer than the time for

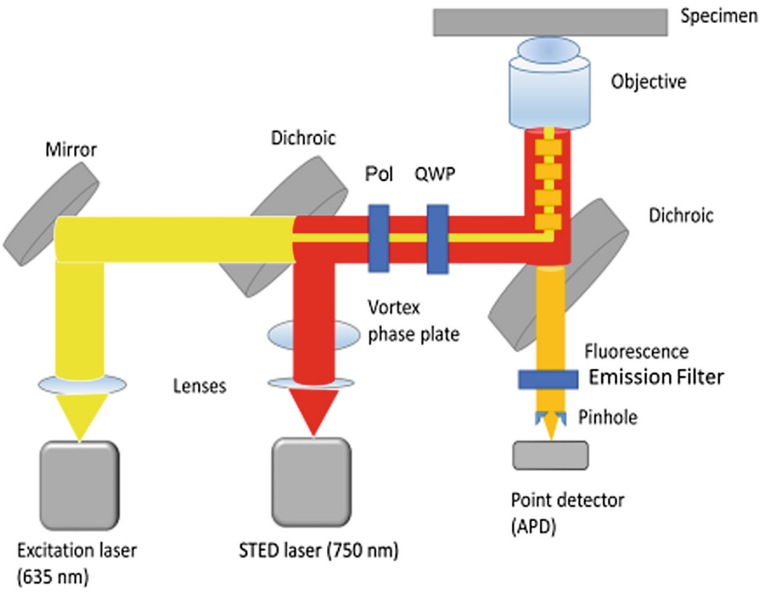


Fig. 14.4 Principle of the STED microscope. Coaxial sequences of excitation pulses (shown in *yellow*) are followed by longer wavelength stimulated emission depletion (STED) pulses (shown in *red*) for fluorescence inhibition. After passing through dichroic mirrors and emission filters specimen fluorescence (shown in *orange*) is detected through a pinhole by a point detector, here an avalanche photodiode (APD). The emission filter only allows photons emitted from fluorescent molecules to pass; photons from stimulated emission are not detected and do not contribute to the image. The vortex phase plate (a.k.a. the spiral phase plate) in the path of the STED beam converts the STED beam into a torus or a beam profile with zero intensity at its center. Between the spiral phase plate and the objective are a polarizer (Pol) and a quarter waveplate (QWP). The combination of the polarizer (half-wave plate) and the quarter-waveplate caused the STED beam and the excitation beam to have a circular polarization. Fluorescent molecules in the outer region of the excitation beam are driven from the excited state to the ground state by stimulated emission of the STED beam. Only in the center of the STED beam, where there is zero intensity, are molecules able to fluoresce via spontaneous emission. The STED beam reduces the effective excitation spot size. Stimulated emission induced by the STED beam is redshifted to longer wavelengths, which is not the case with spontaneous emission. Stimulated emission is not detected by the point detector since it is blocked by the dichroic mirror and the fluorescence emission filter. Only spontaneous emission from the decreased effective excitation spot is allowed to pass through the dichroic mirrors and detected by the point detector. The schematic is not drawn to scale

vibrational relaxation. Stimulated emission must occur before spontaneous emission can occur. In this example the STED pulses are 40 ps long. The STED beam operates at a wavelength that is on the red edge of the emission spectrum of the fluorescent molecule and photons emitted during the stimulated emission process are not detected. However, the photons emitted from the undepleted center region are detected to generate an image of the specimen.

The excitation beam and the STED depletion beam have a relative delay of a few picoseconds. This delay is important since it considers rapid vibrational relaxation at the lowest vibrational level of the first excited state. This enhances the probability

of the STED beam inducing stimulated emission and therefore depleting the number of molecules in the excited state. The excitation beam is focused to an Airy intensity profile. When the STED beam has an annular intensity profile with null or zero intensity at its center, application of the STED pulse will depopulate fluorescent molecules in the focal region of the STED beam with the exception of its geometric center where there is zero intensity. Increasing the intensity of the STED pulse brings about the nonlinear phenomenon of saturation allowing the effective PSF of the excitation beam to be made smaller than the PSF of a diffraction-limited spot that would occur in the absence of a STED pulse (Hell et al., 2003). It is theoretically possible to achieve unlimited resolution as a result of increasing the intensity of the STED beam.

The selection of fluorescent molecules and fluorescent proteins that label the specimen is an important consideration since their spectroscopic properties must be compatible with the wavelengths and the pulse characteristics of commercial lasers that serve as light sources for the STED microscope. Moreover, desirable fluorescent probes are photostable and bright (high quantum efficiency). As I previously stated the alignment of the excitation beam and the STED beam is critical. This alignment is readily disturbed by mechanical vibrations and thermal effects, which must be minimized.

The relation between the resolution of the STED microscope and the intensity of the STED beam has been expressed in an equation (Harke, 2008; Harke et al., 2008; Hell, 2004; Hell et al., 2003). In the case of an ideal system without optical aberrations and light scattering (14.7) describes the enhanced resolution of the STED microscope as a result of increased intensity of the STED beam.

The radius d that defines the lateral resolution of the STED beam is given by

$$d = \frac{\lambda}{2\mathrm{NA}\sqrt{1 + \frac{I}{I_{\mathrm{SAT}}}}} \tag{14.7}$$

where λ is the wavelength, NA is the numerical aperture of the microscope objective, I is the maximum intensity of the STED beam, and I_{SAT} is the saturation intensity of the STED beam (i.e., the intensity with which the fraction $1/e$ of the molecules are depleted or switched off by the STED beam; in some STED publications the saturation intensity is defined as the STED intensity that reduced the fluorescence by one half of its original value (Harke et al., 2008)) (Harke, 2008; Harke et al., 2008; Hell, 2003, 2004; Hell et al., 2003).

Equation (14.7) points up some interesting constraints and consequences. First, each fluorescent molecule will have a unique saturation intensity. Optimal fluorescent probes for STED microscopy will have low saturation intensities. When maximum intensity of the STED beam is equal to the saturation intensity, then resolution becomes diffraction limited. As the ratio of I/I_{SAT} increases so does the resolution of the STED microscope. In practice, optical aberrations and light scattering prevent theoretical resolution from being achieved.

An interesting case occurs when $I \gg I_{\mathrm{SAT}}$ because nonlinearities are then operating and resolution of the STED microscope can not only exceed the

diffraction limit it is also theoretically unlimited. Again, the caveat is the deleterious effects of photobleaching, which affects the signal-to-noise ratio (SNR), and phototoxicity, which affects the specimen.

A key element of the STED microscope is the spiral phase plate (discussed in Section 14.1.3). The spiral or vortex phase plate converts a laser beam with a Gaussian intensity profile into a beam with an annular intensity profile and, most importantly, with null or zero intensity at its center. In simple terms how does this occur? The spiral phase plate introduces phase information that changes with position into a plane wave before the microscope objective. A focused STED beam with spatially varying phase information will bring about destructive interference at the center of the STED beam, such interference resulting in null or zero intensity at the center of the annular beam.

Furthermore, the spiral or vortex phase plate operates with a STED beam that is circularly polarized. The use of circularly polarized light together with the spiral phase plate yields a focal spot with an annular intensity profile and with null or zero intensity at the center. This intensity profile when implemented in STED increases the resolution isotropically, thus fluorescence is independent of the lateral emission dipole distribution (Willig, 2006).

Because of its importance to the understanding of STED microscopy there is a need to explain why there is continuous reduction in the effective PSF of the excitation beam in the focal region when STED beam intensity is continuously increased above saturation intensity. Figure 14.5 illustrates this phenomenon in one dimension; however, the ideas incorporated in the figure can easily be extended into three dimensions.

When a smaller effective PSF is scanned over the specimen, an unlimited (theoretically only limited by the laser power of the STED laser and photodamage and photobleaching of the specimen) subdiffraction or superresolution image is obtained. The saturation level of the STED beam is that intensity at which one-half of the fluorophores are inhibited by stimulated emission (Harke et al. 2008). There are a number of assumptions such as the Gaussian excitation laser beam is coaxial with a STED laser beam that has an annular intensity profile with zero intensity at its center, the Gaussian excitation beam has an intensity peak that is centered on the optical axis, and the annular-shaped STED laser beam has zero intensity at its center and is centered on the optical axis.

For illustration purposes I introduce two intensities of the STED beam that I denote low intensity and high intensity. For the low-intensity case (two left panels of Fig. 14.5) the intensity of the annular-shaped STED beam is zero at the optical axis and the maxima of intensity are in the annular ring. When the maximum intensity of the STED beam is slightly higher than the saturating intensity, then intensity in the outer regions of the STED beam is greater than the saturation intensity and the effective observation volume is slightly narrowed. For the high-intensity case (two right panels of Fig. 14.5) the spatial distribution of the STED beam is unchanged from the low-intensity case. The very significant change here is that the STED beam is now greater than the saturation intensity closer to the center. Now all the molecules in the periphery of the STED beam are forced into the

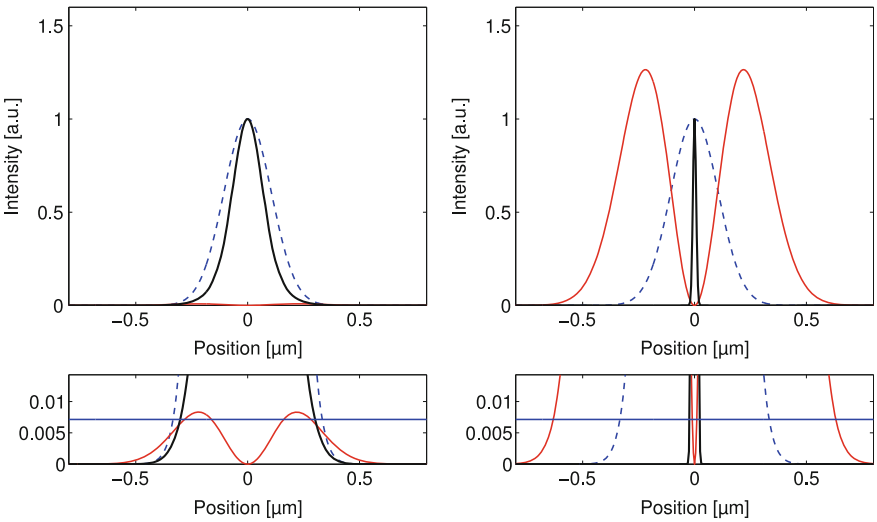


Fig. 14.5 Interaction and spatial intensity profiles of excitation and depletion foci needed to generate a small region of subdiffraction size where fluorescence emission is still allowed. Lower panels are enlargements of upper panels. The *blue horizontal line* indicates saturation intensity (the intensity at which half the fluorescence is inhibited by stimulated emission). Intensity cross sections through the overlaid foci: *blue dashed line* indicates excitation intensity distribution, *red line* STED intensity distribution, and *black line* effective observation volume (describing the image of a point object). *Left*: low STED intensity. Only in the outer part of the focus is STED intensity above the saturation intensity of the dye. The effective observation volume is slightly narrowed. *Right*: high STED intensity. Although spatial distribution of the STED intensity is unchanged, the STED intensity is above the saturation intensity even close to the center. The effective observation volume is thus strongly narrowed because all molecules in the periphery are kept dark. Note that the different intensities are not drawn to scale: the STED intensity (*red line*) is usually much higher than the excitation intensity (*blue dashed line*) Fornasiero, Rizzoli, Eds. (2014). *Super-Resolution Microscopy Techniques in the Neurosciences*. Chapter 3, M. A. Lauterbach and C. Eggeling, *Foundations of STED Microscopy*. Figure 5, p. 53. Humana Press is a part of Springer

dark state by stimulated emission. The focal volume from which there is spontaneous emission of fluorophores that form the image gets smaller and smaller with a concomitant increase in resolution as the intensity of the STED beam increases. Therefore, the effective point spread function is significantly narrowed resulting in unlimited superresolution. In principle, there is no limit to resolution enhancement as the power of the STED laser is increased, but again increased laser power will result in photobleaching and photodamage to the specimen. There are also the problems of optical aberrations and light scattering, as well as the persistent problem of the diminishing number of photons from ever-smaller excitation volumes that affect the SNR of the image.

14.3 Ground State Depletion Microscopy

In 1995 Hell and Kroug working at the University of Turku, Finland published their results of computer simulations and modeling of a new technique in a paper titled "Ground-state-depletion fluorescence microscopy: A concept for breaking the diffraction resolution limit" (Hell and Kroug, 1995).

Ground state depletion (GSD) microscopy differs from STED microscopy in that it does not use stimulated emission. Instead, it uses the basic idea of achieving enhanced resolution beyond the diffraction limit by reducing the size of the effective PSF of the excitation beam. The authors achieved this by depleting the ground state electronic energy level of fluorescent molecules in the periphery of the focus. Their simulations indicate that this method can achieve a lateral resolution of 15 nm (Hell and Kroug, 1995).

How does GSD microscopy work? As previously discussed and as shown in Fig. 14.1 transitions from the first excited singlet state to the singlet ground state are allowed. Nonradiative transitions from the first excited singlet state to the first excited triplet state are forbidden (Kasha, 1950, 1960). Although transitions are forbidden there is a low probability they may still occur via a process called intersystem crossing (ISC), which is due to spin–orbit coupling that is enhanced by heavy atoms in the molecule.

Hell and Kroug ran their computer simulations for the molecule fluorescein using a number of parameters: an assumed excitation wavelength of 488 nm; a fluorescence lifetime of 4.5 nm; an ISC lifetime of 100 ns; and a phosphorescence (radiative transition from the first excited triplet state to the singlet ground state) lifetime of 1 μs. While the STED technique is optimally run with pulsed lasers, the newly proposed GSD technique uses continuous laser light of significantly lower intensity than STED beams. This results in the significant advantage of decreased photobleaching of fluorescent molecules and decreased phototoxicity of the biological specimen.

During continuous illumination provided by the laser the fluorescein molecules cycle from the ground state to the first excited state and back to the ground state with the emission of a photon (the process of fluorescence). During each cycle a small fraction is temporally trapped in a long-lived triplet state. Over many cycles this fraction increases and eventually the ground state is depleted. If the ground state is depleted, there cannot be any induced absorption from the ground state to the first excited state and then the subsequent radiative transition back to the ground state with the emission of fluorescence.

How does depletion of the ground state of the fluorescein molecule achieve enhanced lateral resolution in the GSD microscope? In Hell and Kroug's simulation they assumed the laser beams were offset, with the minima of the laser beams overlapping at the geometrical focus and the maximum of one laser beam and the first side maximum of the second beam overlapping. The key to resolution enhancement was the high laser intensities that caused the triplet state to saturate. A second probe laser beam was used to excite ground state molecules located in the

center of the focal area that were not depleted. The effective PSF of this excitation laser beam was reduced, and therefore enhanced lateral resolution was not only achieved but was capable of exceeding the resolution given by the diffraction limit (Hell and Kroug, 1995). In summary, the fundamental new concept was the reversible fluorescence bright/dark switch to a dark state with a long lifetime which achieved superresolution imaging.

What are the limitations to GSD microscopy? The de-excitation time from the triplet state to the ground state on the order of 1–5 μs limits the pixel-scanning speed. Furthermore, fluorescent molecules in the excited triplet state can be bleached via photochemical reactions.

More than a decade after GSD was theoretically proposed by Hell and Kroug (1995) it was experimentally demonstrated, as reported in the publication titled "Breaking the diffraction barrier in fluorescence microscopy by optical shelving" (Bretschneider et al., 2007). The authors followed up on the simulations of Hell and Kroug and demonstrated a working GSD microscope with light intensities of 10^3–10^5 W/cm^2 that achieved a two-dimensional lateral resolution of 90 nm in mammalian cell microtubules and proteins on the surface membrane of a neuron (Bretschneider et al., 2007). The light intensities required to achieve fluorescence depletion as given above are significantly lower than the beam intensities required for the STED technique: 10^8–10^{10} W/cm^2. The advantages of the GSD technique are that it uses continuous wave laser excitation at much lower intensities than in the STED technique, and it works with regular fluorescent probes used in cell biology.

How does the GSD microscope work? The principle of GSD or optical shelving involves depleting the ground state of the fluorescent molecule by trapping the molecule in a metastable triplet state (as explained in the previous section). This metastable triplet state has a lifetime of 1–10^4 μs. As with the STED microscope technique the GSD microscope uses a spiral phase plate to form a beam with an annular intensity profile that has null or zero intensity at its center. A circularly polarized laser beam is required in addition to the spiral phase plate. The effective PSF of fluorescence emission is significantly smaller than the wavelength thus achieving subdiffraction imaging of the specimen.

To reduce the degree of photobleaching from a molecule in a long-lived triplet state (Bretschneider et al., 2007) used oxygen scavengers to reduce the oxygen levels surrounding the specimen. Because GSD microscopy is a far-field scanning technique there is a delay for molecules in the triplet state to de-excite and return to the singlet ground state before the beam moves to the next position on the specimen. This delay limits the speed of data acquisition.

The resolution of a GSD microscope, which has a focal spot of zero intensity at its center and a reversible saturable fluorescence depletion transition, can be mathematically expressed as (Hell, 2003)

$$d = \frac{\lambda}{2\mathrm{NA}\sqrt{1 + \frac{I_D^{\max}}{I_S}}} \tag{14.8}$$

where d is the resolution, λ is the wavelength, NA is the numerical aperture of the microscope objective, $I_D^{\max}$ is the maximum value of the peak bordering the zero, and I_S is the saturation intensity of the GSD beam, the intensity at which half the fluorescence is depleted.

In 2008 Fölling et al. published a paper on ground state depletion and single-molecule (GSDIM) return, a variant of the GSD technique, titled "Fluorescence nanoscopy by ground-state depletion and single-molecule return" (Fölling et al., 2008). The motivation behind development of this technique was to overcome a constraint fundamental to the techniques of photoactivation localization in stochastic optical reconstruction microscopy (STORM) and photoactivation localization microscopy (PALM), which are extensively discussed in the next chapter (Chapter 15). That constraint is that the photoactivatable probes used in STORM and PALM must meet the stringent requirements of specific photophysics and switching kinetics, compatibility with specific lasers, biocompatibility with the specimen, and labeling techniques. The main advantage of GSDIM is the use of ordinary fluorescent probes. Another advantage is that the technique is simple and that the recording of fluorescence is continuous.

How does GSDIM work? The fluorescent molecules used in GSDIM are common fluorescent probes. The technique operates by switching most fluorescent molecules into a long-lived, metastable triplet state. The positions of molecules that have remained in the ground state or that have spontaneously returned to the ground state are recorded. The GSDIM and GSD techniques have differences and similarities. Both use oxygen scavengers to mitigate the occurrence of photobleaching. However, GSDIM differs from GSD in that the former requires a fluorophore to recover to the singlet ground state only once. GSDIM is a stochastic process in which the positions of individual molecules are recorded. The GSD technique defines the position of molecules that have intensity zeros in the annular illumination beam. However, the GSDIM and GSD techniques use similar photophysical mechanisms to operate. The GSDIM technique is less complex to implement than the GSD technique (Fölling et al., 2008).

The next publication I discuss demonstrates the efficacy of GSD microscopy to achieve diffraction-unlimited spatial resolution. Titled "Metastable Dark States Enable Ground State Depletion Microscopy of Nitrogen Vacancy Centers in Diamond with Diffraction-Unlimited Resolution" the paper is a seminal work investigating the application of GSD microscopy techniques (Han et al., 2010). What the authors achieved was remarkable. They linked the luminescent triplet state and the dark state of diamond NV^- color centers (characterized by their long lifetimes) and obtained a spatial resolution of less than 20 nm, which is better by a factor of 10 than the diffraction limit.

Let us start by providing an explanation of negatively charged nitrogen–vacancy (NV^-) color centers in diamonds. Diamond is a three-dimensional lattice of carbon atoms. The NV^- forms when a nitrogen atom substitutes for a carbon atom that has a neighboring vacancy; these so-called color centers in diamond exhibit extraordinary photostability. Electronic transition from the triplet ground state to the triplet luminescent state occurs at 560 nm. The de-excitation process results in a luminescence from 600 to 850 nm and a luminescence lifetime of 12 ns. Nitrogen–vacancy color centers in diamonds are another important type of probe for biological imaging with superresolution optical microscopy (Balasubramanian et al., 2014).

The general problem associated with STED techniques and their variants was that they all suffer from high incident light intensity, which induces photobleaching or irreversible destruction of the fluorophore. The depletion of excited states requires a light intensity of several gigawatts per square centimeter if a spatial resolution below 10 nm is to be achieved. What was required was a low-intensity method to achieve similar resolution.

Balasubramanian et al. achieved their goal by transiently inhibiting emission in the GSD microscopy technique. This was achieved by transferring the NV^- to a dark state that was metastable and using another laser pulse to measure the emission of those NV^- that persist in the on state. The authors used three wavelengths for initialization, GSD, and a wavelength to read out persistent NV^- sites.

14.4 Reversible Saturable Optical Fluorescence Transitions Microscopy

In Chapter 13 I discussed the use of fluorescence saturation to introduce additional harmonics in the technique of nonlinear SIM, which achieved theoretically unlimited superresolution. Kawata's group in Japan introduced another technique called saturated excitation (SAX) microscopy to improve spatial resolution (Fujita et al., 2007; Yamanaka et al., 2008). The principle of SAX microscopy is to temporally modulate the intensity of excitation light and to detect harmonic modulation of fluorescence caused by saturated excitation; nonlinearity is most pronounced in the center of the laser focus. Therefore, the nonlinear fluorescence signal from the center of the beam can be used to resolve structures smaller than the focal size.

Demodulated fluorescence intensity is nonlinearly proportional to excitation intensity. Fluorescence saturation results in distorted modulation and in fluorescence containing high harmonic frequencies that can be detected. Using SAX microscopy the authors achieved superresolution (beyond the diffraction limit) in three dimensions (Yamanaka et al., 2008). The signal intensity of higher harmonic frequency components is around 10 to 1000× lower than the fundamental frequency component (Yamanaka et al., 2011). A similar decrease in the signal at higher harmonics was observed with nonlinear SIM (Chapter 13). The

signal-to-noise ratio is decreased for higher harmonics. If a fluorescent probe is very stable to light damage, then it is possible to increase the acquisition time to improve the signal-to-noise ratio. Resolution versus acquisition time is an important trade-off.

One advantage of the SAX technique is that any fluorescent molecule demonstrating a saturation effect during excitation can be used as a fluorescent probe of the specimen. The limitation of the SAX microscopy technique is that the very high intensities used for fluorescence excitation also cause strong photobleaching (destruction of the fluorescent molecule). One approach to mitigating this problem with SAX microscopy is to use nanodiamonds as fluorescent probes. These fluorescent probes with nitrogen–vacancy defect centers are extremely stable under high-intensity illumination. SAX microscopy has been applied to the superresolution imaging of macrophage cells using nanodiamond probes (Yamanaka et al., 2011).

In 2005 Hofmann et al. introduced a far-field scanning superresolution technique called reversible saturable optical fluorescence transitions (RESOLFT), a generalization of STED and GSD microscopy and its variants, in their publication "Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins" (Hofmann et al., 2005). There is a patent (US Pat. No. 7,064,824) on the RESOLFT principle.

The motivation behind expanding STED and its variants into RESOLFT microscopy related to the problems and limitations of STED microscopy. A significant disadvantage of the STED technique is the requirement for the STED beam to be of high intensity to saturate stimulated fluorescence rather than spontaneous fluorescence. In the STED technique the STED beam is composed of picosecond pulses of laser light at an intensity of 100 MW/cm^2. These high intensities result in extensive photobleaching of fluorescent molecules and phototoxicity in the specimen. These deleterious effects are significantly mitigated with the RESOLFT technique. Hofmann et al. achieved theoretically unlimited spatial resolution by saturating a linear, reversible optical transition between two states of the molecule labeled state A and state B using light. The light that brought about the transition between state A and state B induced a conformational change in the molecule: one of the states was fluorescence activated (fluorescent) and the other state was nonactivated (nonfluorescent). One of the transitions from state A to state B or from state B to state A was driven by a laser beam with an annular intensity profile and null or zero intensity at its center. With this technique the authors achieved spatial resolution beyond the diffraction limit (i.e., superresolution) in the focal plane with a saturating laser pulse of a few watts per square centimeter.

What limits the RESOLFT technique? In addition to the major limitations given by the photochemistry of the fluorescent protein in the specimen and how close to a zero intensity actually exists at the center of the beam, there are the usual limitations due to light scattering and optical aberrations (Hofmann et al., 2005).

In 2007 Schwentker et al. described widefield subdiffraction RESOLFT microscopy using fluorescent protein photoswitching (Schwentker et al., 2007). One limitation with STED microscopy is that depletion by stimulated emission has to compete with a fluorescent decay rate of $(1–5\ \text{ns})^{-1}$. This requires an intensity of 10^{11}–10^{13} W/cm^2, which is provided by picosecond pulses. The technique presented by the authors depletes the ground state of the fluorescent molecule by pumping it into its triplet state. Because of the 1000× longer lifetime of the triplet state the fluorescence signal can be depleted by intensities that are reduced by the same factor.

Image acquisition time is another problem with techniques that comprise RESOLFT. Total image acquisition time for any scanning microscopy technique depends on the number of pixels that are scanned multiplied by the pixel dwell time (i.e., the time that the laser beam is situated on each pixel). Total image acquisition time also depends on the size of the scanned area. Higher image acquisition times are achieved by reducing the scanned area of the specimen.

Since RESOLFT is a scanning far-field fluorescence microscopy technique the image acquisition time is long if the microscope uses a single beam with an annular intensity profile and null or zero intensity at the center of the beam. Is there a solution to this problem? The idea to overcome this problem is to parallelize the technique.

Chmyrov et al. presented the solution in their publication titled “Nanoscopy with more than 100,000 ‘doughnuts’” (Chmyrov et al., 2013). Instead of a single annular (“doughnut”) beam with an intensity zero at its center these authors invented a new type of microscopy they termed parallelized RESOLFT microscopy.

How does the parallelized RESOLFT microscopy technique work? The authors were able to parallelize their scanning RESOLFT microscope by forming a pattern that is scanned over the specimen. The pattern is formed by two beams that result from the incoherent summation of two orthogonal sinusoidal standing waves. These standing waves are characterized by zero-intensity points at locations where valleys and crests intercept and form intensity peaks twice the intensity of the crests of standing waves. Chmyrov et al. demonstrated that spatial resolution is completely isotropic for beam intensities that are 100 times the saturation intensity. The image acquisition time for the RESOLFT technique of Chmyrov et al. is given by the dynamics of transitions between states and the total number of pixels in the field of view imaged. Other limitations to image acquisition speed are related to light scattering, optical aberrations, and the frame rate of the camera.

Their results were exceptional. Using light intensities that were 10^{-5} the intensity required for STED the RESOLFT microscope in less than a second used 116,000 doughnuts to form a superresolution image of the specimen consisting of live cells in an area of the field of view sized 120 × 100 μm (Chmyrov et al., 2013).

In this section I discussed the advantages and limitations of RESOLFT. RESOLFT operates with reduced illumination intensities compared with those required for STED. It also makes use of photoswitchable fluorescent molecules

capable of reversibly transitioning between two states: a fluorescent state (ON state) and a nonfluorescent state (OFF state). RESOLFT requires that the switching be reversible, efficient, and repeatable over many cycles. Resolution is inversely related to the square root of the number of switching cycles.

14.5 Advances in Instrumentation, Probes, and Applications

The last decade witnessed a proliferation of publications highlighting important developments in STED instrumentation, probes, and applications.

I begin with Willig et al. whose publications significantly advanced instrumentation, probe development, and applications. While technical developments are very important to successful application of instruments to investigate important biomedical problems, validation and the significance of the new knowledge are key to the endeavor.

Willig's dissertation is well worth reading. In addition to her seminal biological studies of STED her dissertation established a novel annular intensity profile of the STED beam that had null or zero intensity at the center of the beam by utilizing a spiral phase plate (Willig, 2006). Willig et al. used STED microscopy to demonstrate an important finding in neuroscience. They found that the synaptic vesicle protein synaptotagmin remains an integral patch after its fusion with the plasma membrane (Willig et al., 2006b). Willig et al. studied the use of STED microscopy with generically encoded markers. Using STED microscopy they achieved a subdiffraction lateral resolution of 70 nm with GFP-labeled viruses and the endoplasmic reticulum (ER) of mammalian cells (Willig et al., 2006a). In 2007 Willig et al. demonstrated a STED microscope that had continuous wave laser beams (Willig et al., 2007).

The next publication I discuss introduces laser diodes as an alternative light source for STED microscopy. STED microscopy typically uses stretched femtosecond pulses formed in a mode-locked Ti:sapphire laser with a synchronously pumped optical–parametric oscillator. Laser diode STED microscopy uses two laser diodes that send pulses of 300–400 ps for STED and a third laser diode that sends pulses of 50–70 ps to excite fluorescent probes (Westphal et al., 2003). Their experiments demonstrated that laser diodes can saturate depletion. Compared with a mode-locked Ti:sapphire laser with a synchronously pumped optical–parametric oscillator, laser diodes are more easily operated as a result of electronic timing, they are compact, they have low maintenance requirements, and their cost is significantly lower.

Westphal and Hell reported a novel implementation of STED that gave a focal spot size of 16 nm (corresponding to $\lambda/50$) in a far-field fluorescent microscope. In addition, they experimentally proved that the minimal resolving distance scales inversely with the square root of the saturation level of the STED beam (Westphal and Hell, 2005).

Donnert et al. demonstrated two-color far-field fluorescence microscopy that could deliver nanoscale spatial resolution. They did so by applying STED microscopy that made use of two fluorophores with different absorption and emission spectra. Green-emission and red-emission fluorescent probes were selectively excited and quenched using different beams. STED beams delivered a lateral resolution of <30 nm and 65 nm for the green and the red channel, respectively (Donnert et al., 2007).

Early studies using STED microscopy were performed on fixed cells. The first STED application to live cells (i.e., yeast cells and *E. coli* cells) was carried out by Klar et al., (2000). Live cell imaging is of great interest and is importunate in modern biological research. STED microscopy can be used to achieve superresolution imaging of rapid physiological processes. Westphal et al. demonstrated video-rate (28 frames per second) far-field STED microscopy in their investigation of synaptic vesicle movement (Westphal et al., 2008). They demonstrated video-rate far-field optical imaging with a focal spot size of 62 nm in living cells; however, this was achieved in a very limited field of view 2.5 × 1.8 μm in size.

A rate-limiting step in the advancement of STED microscopy is the development of new fluorescent probes. Such probes have minimal biological effects on cells and organisms, they are photostable and bright, and the kinetics of their radiative and their nonradiative transitions are optimally compatible with available lasers and other light sources.

The photostability of a fluorescent marker under pulsed STED beams has been investigated. An important finding was that the photostability of specific fluorophores can be affected by altering the duration of the STED pulse (Dyba and Hell, 2003). Another finding was that several new red fluorescent probes were optimal for STED microscopy (Wurm et al., 2012). Having a model of the complex kinetic processes upon which STED microscopy is based allows specific parameters to be varied and the effects on resolution and photophysical processes to be studied. This was the basis of the publication titled "Analytical description of STED microscopy performance" (Leutenegger et al., 2010).

An interesting modification of the typical STED microscope permits superresolution and images the molecular orientation of fluorescent molecules that label the specimen. This method is denoted molecular orientation microscopy by STED (MOM-STED). Reuss et al., (2010) described a new birefringent device. This was a STED beam with an annular intensity profile and null or zero intensity at the beam center that could also be modified to reveal the orientation of fluorophores in the specimen, thereby providing subdiffraction images of molecular orientation. The authors used birefringent crystals to modify polarization across the STED beam. The same beam-shaping device can be tuned so that the image of a single fluorophore is strongly dependent on the orientation of the fluorophore's transition dipole moment (Reuss et al., 2010).

The next advance in instrumentation addressed another persistent problem of STED microscopy: the extensive photobleaching of fluorescent probes and phototoxicity of the specimen caused by the high light intensities of the STED beam. This problem was solved by developing a STED microscope with time-gated detection (Vicidomini et al., 2011; Vicidomini et al., 2013). This instrument used continuous wave STED (CW-STED). The advantage of CW-STED was that superresolution imaging could be achieved in fixed and in live cells with moderate light intensities. This advance in STED instrumentation is protected by the US patent publication, US2013256564 (A1), 2013-10-03, "STED Microscopy with Pulsed Excitation, Continuous Stimulation, and Gated Registration of Spontaneously Emitted Fluorescence Light." The application was filed May 22, 2013 and published October 3, 2013. The inventors are: Hell, Stefan W. [DE]; Engelhardt, Johann [DE]; Reuss Matthias [SE]; Westphal, Volker [DE]; Eggeling, Christian [DE]; Moneron, Gael [FE]; Han, Kyu-Young [US]; Vicidomini, Giuseppe [IT]; Willig Katrin [DE].

Figure 14.6 shows the components that make up the new CW-STED microscope described in their patent. A continuous wave laser provides the STED beam and a pulsed laser is the excitation source. The detector is time-gated, which serves to sharpen the image by reducing the detected signal that does not come from spontaneous fluorescence in the center of the STED beam. This invention solves a couple of problems. First, how to form a STED beam with a continuous wave laser capable of replacing the very expensive pulsed lasers that emit very short pulses. Second, how to mitigate blur in STED microscopes that use a continuous wave laser. The origin of this image blur is a background pedestal in the effective PSF. The authors' solution is to use a continuous wave laser together with time-gated detection (Vicidomini et al., 2011; Vicidomini et al., 2013).

Vicidomini et al. (2011) in their publication titled "Sharper low-power STED nanoscopy by time gating" showed that CW-STED achieved superresolution of labeled lipids in the cell membranes of living cells. This was accomplished with minimal image blur and with only moderate light intensities for the CW-STED beam. In a subsequent paper titled "STED nanoscopy with time-gated detection: theoretical and experimental aspects" Vicidomini et al. point out that time-gated detection has another advantage in that it mitigates the effect of local variations of the emission lifetime on the spatial resolution that is achieved. Furthermore, the authors point out a limitation associated with the time-gated STED technique in that time-gating results in a reduced signal from an effective PSF (Vicidomini et al., 2013). This is another example of the trade-offs that are required in superresolution optical microscopy.

Clausen et al. published a review on pathways to optical STED microscopy in which they highlight major problems with earlier versions of STED microscope and new technologies that mitigate some of these problems (Clausen et al., 2013).

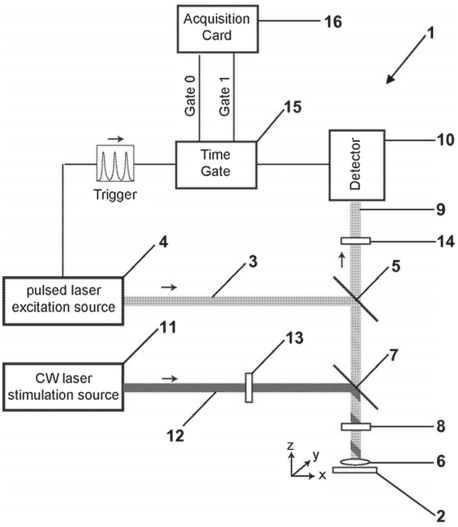


Fig. 14.6 The basic design of an embodiment of a new STED fluorescent light microscope in the US2013256564 (A1) patent. United States patent publication, US2013256564 (A1), 2013-10-03, STED Microscopy with Pulsed Excitation, Continuous Stimulation, and Gated Registration of Spontaneously Emitted Fluorescence Light. The application was filed May 22, 2013, and published October 3, 2013. The inventors are: Hell, Stefan W. [DE]; Engelhardt, Johann [DE]; Reuss, Matthias[SE]; Westphal, Volker [DE]; Eggeling, Christian [DE]; Moneron, Gael [FR]; Han, Kyu-Young [US]; Vicidomini, Giuseppe [IT]; Willig, Katrin [DE]

A recent publication titled "STED microscopy with time-gated single-photon avalanche diode" (Hernández et al., 2015) demonstrated that a fast-gated single-photon avalanche diode (SPAD) can improve the SNR of a time-gated STED image and reduce instrument complexity at the same time.

A recent 65-pp review article of lens-based fluorescence nanoscopy by Hell and his colleagues presents two decades of innovative research to achieve superresolution by a variety of techniques as interpreted by the authors (Eggeling et al., 2015). Several salient points from their review warrant discussion.

First, let us consider the concept behind saturation of a STED beam. When STED beam intensity is greater than saturation intensity, the rate of stimulated emission is greater than the rate of spontaneous emission.

Saturation intensity is given by

$$I_{\mathrm{SAT}} = (\tau \sigma_{\mathrm{STED}})^{-1} \tag{14.9}$$

where σ_{STED} is the photon cross-section of stimulated emission at the wavelength of the STED laser, and τ is the lifetime of the first excited singlet state S_1. To quote

from the review article of (Eggeling et al., 2015): "With lifetimes in the range of 1–4 ns and stimulated emission cross-sections in the range of 10−17 cm^2 (i.e. photon cross-sections $\sigma_{STED} \approx$ 25–30 cm^2/J) STED intensities of ISAT≫1–10 MW/cm^2 have to be applied to realize a sufficiently large fluorescence inhibition."

Second, how do CW-STED lasers compare with pulsed STED lasers? CW-STED lasers have been shown to be less complex and less costly than pulsed STED lasers. However, the use of CW-STED lasers is associated with two disadvantages. CW-STED lasers need to operate at five times the average power of pulsed STED lasers. The result is increased photobleaching of fluorescent molecules and increased phototoxicity of specimens. In addition, during operation of a CW-STED laser a nonnegligible portion of fluorescent molecules emit fluorescence before they are irradiated with the CW-STED beam. This fluorescence originates from regions that are peripheral to the intensity null or zero intensity at the beam center. Therefore, a pedestal exists in the effective PSF resulting in reduced contrast in images. Time-gated detection mitigates this problem (Vicidomini et al., 2013).

Third, the review article stresses the importance of parallelization in reducing image acquisition time (Chmyrov et al., 2013).

STED microscopes are limited in their actual performance by several considerations. As is the case with all types of optical microscopes theoretical resolution is affected by light scattering and optical aberrations introduced by the optical system and the specimen. There are severe restrictions on maintaining the alignment of the excitation beam and STED beams including the use of multiple STED beams.

How can STED microscopes be simplified, be made more compact, be less expensive, and be easier to maintain? Although the complexity of a STED microscope is high as reflected in the purchase price for a commercial STED microscope, Wildanger et al. designed and developed a compact STED microscope that was much less expensive and significantly less complex than previous STED instruments and achieved a superresolution of 19 nm in the focal plane (Wildanger et al., 2009). Their STED instrument used a single commercial all fiber-based supercontinuum laser source and a commercial vortex phase plate to form both an excitation beam and a STED beam with an annular intensity profile and null or zero intensity at its center (Wildanger, et al., 2008).

The new compact STED microscope achieved a lateral spatial resolution in the range of 30–50 nm in various colors. It required little alignment and afforded long-term stability. As discussed previously, the vortex phase plate required circular polarization to form a focused STED beam with an annular intensity profile and null or zero intensity at its center. However, the supercontinuum beam was not polarized resulting in one-half of the power being lost.

Wildanger et al., (2009) solved this problem by introducing a second beam path that had orthogonal circular polarization improving the performance of the supercontinuum STED microscope and yielding a resolution of 20 nm. A quarter waveplate was located in front of the microscope objective, which caused all three beams to have the required circular polarization to form a focused STED beam with an annular intensity profile and zero intensity at its center. The circular polarization of the STED beam causes all vectorial components of the light electric field to

interfere destructively at the geometric focal point, this is the source of the annular shape of the STED beam. Furthermore, the authors enhanced the lateral and axial resolution by using two phase plates (Wildanger et al., 2009).

Having come up with innovative designs of STED microscopes based on a supercontinuum laser, as described in Wildanger et al. (2008, 2009), the authors went on to develop a new STED microscope that entirely eliminated one beam path but still remained aligned by design (Wildanger, Bückers, Westphal, Hell, and Kastrup, 2009). The authors took advantage of a previous publication on the design of diffractive lenses that generate optical nulls without phase singularities (Menon et al., 2009). The new phase mask attains its wavelength-selective effect by carefully combining two optical media whose refractive indices are matched at the excitation wavelength but are different at the STED wavelength (Wildanger et al., 2009).

Another innovation involved exploiting continuous wave fiber lasers for STED microscopy such as the low-cost and turnkey continuous wave (CW) fiber lasers emitting at 592 nm. Moneron et al., (2010) described the development of a fast STED microscope based on continuous wave fiber lasers. Such a STED microscope achieved spatial resolutions of 35–65 nm in the focal plane.

Current commercial versions of the STED microscope incorporate many of the innovations in STED microscope design discussed in this chapter such as featuring a microscope objective corrected for a broad band of wavelengths. Notably, the PSF can be arbitrarily changed using two different phase plates allowing the user of a STED microscope to separately select the desired focal volume for lateral and axial resolution. Moreover, the combination of two STED beams forms an isometric focal volume.

References

Allen, L., Barnett, S. M., and Padgett, M. J. (2003). *Optical Angular Momentum*. Bristol: Institute of Physics.

Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C., and Woerdman, J. P. (1992). Orbital angular momentum of light and the transformation of Laguerre–Gaussian laser modes. Physical Review A, **45**, 8185–8189.

Andrews, D. L. (2008). *Structured Light and Its Applications, 1st Edition. An Introduction to Phase-Structured Beams and Nanoscale Optical Forces*. San Diego: Academic Press.

Andrews, D. L. (2015). *Fundamentals of Photonics and Physics. Volume I*. Hoboken: John Wiley & Sons.

Andrews, D. L., and Babiker, M. (2013). *The Angular Momentum of Light*. Cambridge: Cambridge University Press.

Baer, S. C. (1994). Method and Apparatus for improving resolution in scanned optical system. Filed: July 15, 1994, Date of Patent: February 2, 1999. U. S. Patent number: 5,866,911.

Balasubramanian, G., Lazariev, A., Arumugam, S. R., and Duan, D. W. (2014). Nitrogen-vacancy color center in diamond-emerging nanoscale applications in bioimaging and biosensing. Current Opinion in Chemical Biology, **20**, 69–77.

Beijersbergen, M. W., Coerwinkel, R. P. C., Kristensen, M., and Woerdman, J. P. (1994). Helical-wavefront laser beams produced with a spiral phase plate. Optics Communication, **112**, 321–327.

Berry, M., Nye, J., and Wright, F. (1979). The elliptic umbilic diffraction catastrophe. Philosophical Transactions of the Royal Society of London, **291**, 453–484.
Berry, M. V. (2004). Optical vortices evolving from helicoidal integer and fractional phase steps. Journal of Optics A, **6**, 259–268.
Bertolotti, M. (1999). *The History of the Laser*. Bristol: Institute of Physics Publishing.
Born, M., and Wolf, E. (1999). *Principles of Optics, 7th expanded edition*. Cambridge: Cambridge University Press.
Boyd, R. W. (2008). *Nonlinear Optics, Third Edition*. San Diego: Academic Press.
Braat, J., and Török, P. (2019). *Imaging Optics*. Cambridge: Cambridge University Press.
Bretschneider, S., Eggeling, C., and Hell, S. W. (2007). Breaking the diffraction barrier in fluorescence microscopy by optical shelving. Physical Review Letters, **98**, 218103-1–21803-4.
Chmyrov, A., Keller, J., Grotjohann, T., Ratz, M., d'Este, E., Jakobs, J., Eggeling, C., and Hell, S. W. (2013). Nanoscopy with more than 100,000 'doughnuts'. Nature Methods, **10**, 737–740.
Clausen, M. P., Galiani, S., Bernardino de la Serna, J., Fritzsche, M., Chojnacki, J., Gehmlich, K., Christoffer Lagerholm, B., and Eggeling, C. (2013). Pathways to optical STED microscopy. NanoBioImaging, 1–12. https://doi.org/10.2478/nbi-2013-0001.
Coullet, P., Gil, G., and Rocca, F. (1989). Optical vortices. Optics Communication, **73**, 403–408.
D'Alessandro, G., and Oppo, G-L., (1992). Gauss-Laguerre modes: A sensible basis for laser dynamics. Optics Communication, **88**, 130–136.
Dennis, M. R., O'Holleran, K., and Padgett, M. J. (2009). Singular Optics: Optical Vortices and Polarization Singularities. In: *Progress in Optics*, Ed. E. Wolf, **53**, 293–363.
Dirac, P. A. M. (1927). The quantum theory of emission and absorption of radiation. Proceedings of the Royal Society (London). Series A, **114**, 243–265.
Donnert, G., Keller, J., Wurm, C. A., Rizzoli, S. O., Westphal, V., Schönle, A., Jahn, R., Jakobs, S., Eggeling, C., and Hell, S. W. (2007). Two-color far-field fluorescence nanoscopy. Biophysical Journal, **92**, L67–L69.
Dyba, M., and Hell, S. W. (2003). Photostability of a fluorescent marker under pulsed excited-state depletion through stimulated emission. Applied Optics, **42**, 5123–5129.
Dyba, M., Jakobs, S., and Hell, S. W. (2003). Immunofluorescent stimulated emission depletion microscopy. Nature Biotechnology, **21**, 1303–1304.
Eggeling, C., Willig, K. I., Sahl, S. J., and Hell, S. W. (2015). Lens-based fluorescence nanoscopy. Quarterly Reviews of Biophysics, **48**, 178–243.
Einstein, A. (1916a). Strahlungs-Emission und -Absorption nach der Quantentheorie. [Emission and absorption of radiation in quantum theory] Deutsche Physikalische Gesellschaft, Verhandlungen, **18**, 318–323.
Einstein, A. (1916b). Zur Quantentheorie der Strahlung. [On the quantum theory of radiation]. Physikalische Gesellschaft Zürich, Mitteilungen, **18**, 47–62.
Fujita, K., Kobayashi, M., Kawano, S., Yamanaka, M., and Kawata, S. (2007). High-resolution confocal microscopy by saturated excitation of fluorescence. Phys. Rev. Lett. **99**, 228105.
Fölling, J., Bossi, M., Bock, H., Medda, R., Wurm, C. A., Hein, B., Jakobs, S., Eggeling, C., and Hell, S. W. (2008). Fluorescence nanoscopy by ground-state depletion and single-molecule return. Nature Methods, **5**, 943–945.
Förster, T. (1946). Energiewanderung und Fluoreszenz. Naturwissenschaften, **33**: 166–175.
Förster, T. (1951). *Fluoreszenz Organischer Verbindungen*. Göttingen: Vandenhoeck & Ruprecht.
Ganic, D., Gan, X., and Gu, M. (2003). Focusing of doughnut laser beams by a high numerical-aperture objective in free space. Optics Express, **11**, 2747–2752.
Goodman, J. W. (2017). Introduction to Fourier Optics. Fourth Edition. New York: W. H. Freeman and Company. Chapter 8. Point-Spread Function and Transfer Function Engineering, pp. 231–267.
Gortych, J. E. (2014). *Consider a Spherical Patent, IP and Patenting in Technology Business*. Boca Raton: CRC Press.
Gu, M. (2000). *Advanced Optical Imaging Theory*. Berlin: Springer, pp. 31–35.

Han, K. Y., Kim, S. K., Eggeling, C., and Hell, S. W. (2010). Metastable dark states enable ground state depletion microscopy of nitrogen vacancy centers in diamond with diffraction-unlimited resolution. Nano Letters, **10**, 3199–3203.

Harke, B. (2008). *3D STED microscopy with pulsed and continuous wave lasers*. Ph.D. thesis, George-August-University, Göttingen.

Harke, B., Keller, J., Ullal, C. K., Westphal, V., Schönle, A., and Hell, S. W. (2008). Resolution scaling in STED microscopy. Optics Express, **16**, 4154–4162.

Hell, S. W. (1994). Improvement of lateral resolution in far-field light microscopy using two-photo excitation with offset beams. Optics Communications, **106**, 19–24.

Hell, S. W. (2003). Towards fluorescence nanoscopy. Nature Biotechnology, **21**, 1347–1355.

Hell, S. W. (2004). Strategy for far-field optical imaging and writing without diffraction limit. Physics Letters A, **326**, 140–145.

Hell, S. W. (2009). Microscopy and its focal switch. Nature Methods, **6**, 24–32.

Hell, S. W., Jakobs, S., and Kastrup, L. (2003). Imaging and writing at the nanoscale with focused visible light through saturable optical transitions. Applied Physics A, **77**, 859–860.

Hell, S. W., and Kroug, M. (1995). Ground-state-depletion fluorescence microscopy: A concept for breaking the diffraction resolution limit. Applied Physics B, **60**, 495–497.

Hell, S. W., and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. Optics Letters, **19**, 780–782.

Hernández, I. C., Buttafava, M., Boso, G., Diaspro, A., Tosi, A., and Vicidomini, G. (2015). Gated STED microscopy with time-gated single-photon avalanche diode. Biomedical Optics Express, **6**, 2258–2267.

Hofmann, M., Eggeling, C., Jakobs, S., and Hell, S.W. (2005). Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. Proceedings of the National Academy of Sciences of the United States of America, **102**, 17565–17569.

Jones, P. H., Maragò, O. M., and Volpe, G. (2016). *Optical Tweezers: Principles and Applications 1st Edition*. Cambridge: Cambridge University Press.

Kasha, M. (1950). Characterization of electronic transitions in complex molecules. Discussions of the Faraday Society, **9**, 14–19.

Kasha, M. (1960). Paths of molecular excitation. Radiation Research, **2**, 243–275.

Klar, T. A., Engel, E., and Hell, S. W. (2001). Breaking Abbe's diffraction resolution limit in fluorescence microscopy with stimulated emission depletion beams of various shapes. Physical Review E, **64**, 06613–06622.

Klar, T. A., and Hell, S. W. (1999). Subdiffraction resolution in far-field fluorescence microscopy. Optics Letters, **24**, 954–956.

Klar, T. A., Jakobs, S., Dyba, M., Egner, A., and Hell. S. W. (2000). Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. Proceedings of the National Academy of Sciences of the United States of America, **97**, 8206–8210.

Leutenegger, M., Eggeling, C., and Hell, S. W. (2010) Analytical description of STED microscopy performance. Optics Express, **18**, 26417–26429.

Lewis, G. N., and Kasha, M. (1944). Phosphorescence and the Triplet State. Journal of the American Chemical Society, **66**, 2100–2116.

Loudon, R. (2000). *The Quantum Theory of Light, Third Edition*. Oxford: Oxford University Press.

Maiman, T. H. (2018). *The Laser Inventor Memoirs of Theodore H. Maiman*. New York: Springer Nature.

Masajada, J., and Dubik, B. (2001). Optical vortex generation by three plane wave interference. Optics Communications, **198**, 21–27.

Masters, B. R. (2001). *Selected Papers on Optical Low-Coherence Reflectometry & Tomography*, volume MS **165**. SPIE Milestone Series. Bellingham: SPIE Optical Engineering Press.

Masters, B. R. (2003). *Selected Papers on Multiphoton Excitation Microscopy*, SPIE Milestone Series, volume MS **175**. Bellingham: SPIE Optical Engineering Press.

Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE, Optical Engineering Press.

Masters, B. R. (2009). C. V. Raman and the Raman Effect. *Optics and Photonics News*, March 2009, pp. 41–45.

Masters, B. R. (1996). *Selected Papers on Confocal Microscopy*. SPIE Milestone Series, volume MS **131**. Bellingham: SPIE Optical Engineering Press.

Masters, B. R. (2012). Albert Einstein and the nature of light. Optics and Photonics News, **23**, 42–47.

Masters, B. R. (2014). Paths to Förster's resonance energy transfer (FRET) theory. The European Physical Journal H, **39**, 87–139.

Masters, B. R. (2015). What is light? In English and translated into 15 languages. *International Commission for Optics News Letter.* e-ico.org Accessed August 19, 2017.

Masters, B. R., and So, P. T. C. (2008). Classical and Quantum Theory of One-Photon and Multiphoton Fluorescence Spectroscopy. In: *Handbook of Biomedical Nonlinear Optical Microscopy*. Eds. Masters, B. R. and So. P. T. C., chapter 5, pp. 91–152. New York: Oxford University Press.

Menon, R., Rogge, P., and Tsai, H.-Y. (2009). Design of diffractive lenses that generate optical nulls without phase singularities. Journal of the Optical Society of America A. Optics and Image Science, **26**, 297–304.

Moneron, G., Medda, R., Hein, B., Giske, A., Westphal, V., and Hell, S. W. (2010). Fast STED microscopy with continuous wave fiber lasers. Optics Express, **18**, 1302–1309.

Nye, J. F., and Berry, M. V. (1974). Dislocations in wave trains. Proceedings of the Royal Society of London A, **336**, 165–190.

Oemrawsingh, S. S. R., van Houwelingen, J. A. W., Eliel, E. R., Woerdman, J. P., Verstegen, E. J. K., Kloosterboer, J. G., and 't Hooft, G. W., (2004). Production and characterization of spiral phase plates for optical wavelengths. Applied Optics, **43**, 688–694.

O'Holleran, K., Padgett, M. J., and Dennis, M. R. (2006). Topology of optical vortex lines formed by the interference of three, four, and five plane waves. Optics Express, **14**, 3039–3044.

Padgett, M. J., and Allen, L. (1955). The Poynting vector in Laguerre-Gaussian laser modes. Optics Communications, **121**, 36–40.

Reuss, M., Engelhardt, J., and Hell, S. W. (2010). Birefringent device converts a standard scanning microscope into a STED microscope that also maps molecular orientation. Optics Express, **18**, 1049–1058.

Rozas, D. (1999). *Generation and Propagation of Optical Vortices.* A Dissertation submitted to the Faculty of Worcester Polytechnic Institute in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Physics.

Schönle, A. (2003). *Point spread function engineering in fluorescence spectroscopy*. Doctoral Dissertation, Ruperto-Carola University of Heidelberg, Germany.

Schönle, A., Keller, J., Harke, B., and Hell, S. W. (2008). Diffraction Unlimited Far-Field Fluorescence Microscopy. In: *Handbook of Biomedical Nonlinear Optical Microscopy*, Eds. Barry R. Masters and Peter T. C. So. Oxford: Oxford University Press, Chapter 24.

Schoonover, R. W. (2009). *Studies in Singular Optics and Coherence Theory*. Doctoral Thesis, Technische Universiteit Delft, The Netherlands.

Schwentker, A., Bock, H., Hofmann, M., Jakobs, S., Bewersdorf, J., Eggeling, C., and Hell, S. W. (2007). Wide-field subdiffraction RESOLFT microscopy using fluorescent protein photoswitching. Microscopy Research and Technique, **70**, 269–280.

Sheppard, C. J. R., and Choudhury, A. (2004). Annular pupils, radial polarization, and superresolution. Applied Optics, **43**, 4322–4327.

Silfvast, W. T. (2004). *Laser Fundamentals, second Edition*. Cambridge: Cambridge University Press.

Soskin, M. S., and Vasnetsov, M. V. (2001). Singular optics. in *Progress in Optics*, edited by E. Wolf, **42**, 219–276. Amsterdam: Elsevier.

Sueda, K., Miyaji, G., Miyanaga, N., and Nakatsuka, M. (2004). Laguerre-Gaussian beam generated with a multilevel spiral phase plate for high intensity laser pulses. Optics Express, **12**, 3548–3553.

Tomonaga, S.-I. (1997). *The Story of Spin*. Chicago: The University of Chicago Press.

Turnbull, G. A., Robertson, D. A., Smith, G. M., Allen, L., and Padgett, M. J. (1996). Generation of free-space Laguerre-Gaussian modes at millimeter-wave frequencies by use of a spiral phaseplate. Optics Communication, **127**, 183–188.

Valeur, D., and Berberan-Santos, M. N. (2012). *Molecular Fluorescence, Principles and Applications, Second Edition*. Weinheim: Wiley-VCH.

Vaughan, J. M., and Willetts, D. (1979). Interference properties of a light-beam having a helicalwave surface. Optics Communication, **30**, 263–267.

Vicidomini, G., Moneron, G., Han, K. Y., Westphal, V., Ta, H., Reuss, M., Engelhardt, J., Eggeling, C., and Hell, S. W. (2011). Sharper low-power STED nanoscopy by time gating. Nature Methods, **8**, 571–573.

Vicidomini, G., Schönle, A., Ta, H., Han, K. Y., Moneron, G., Eggeling, C., and Hell, S. W. (2013). STED nanoscopy with time-gated detection: Theoretical and experimental aspects. PLOS ONE, **8**(1–12), e54421.

Westphal, V., Blanca, C.M., Dyba, M., Kastrup, L., and Hell, S. W. (2003). Laser-diode-stimulated emission depletion microscopy. Applied Physics Letters, **82**, 3125–3127.

Westphal, V., and Hell, S. W. (2005). Nanoscale resolution in the focal plane of an optical microscope. Physical Review Letters, **94**, 143903–143907.

Westphal, V., Kastrup, L., and Hell, S. W. (2003). Lateral resolution of 28 nm ($\lambda/25$) in far-field fluorescence microscopy. Applied Physics B Lasers and Optics, **77**, 377–380.

Westphal, V., Rizzoli, S. O., Lauterbach, M. A., Kamin, D., Jahn, R., and Hell, S. W. (2008). Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. Science, **320**, 246–249.

Wildanger, D., Bückers, J., Westphal, V., Hell, S. W., and Kastrup, L. (2009). A STED microscope aligned by design. Optics Express, **17**, 16100–16110.

Wildanger, D., Medda, R., Kastrup, L., and Hell, S. W. (2009). A compact STED microscope providing 3D nanoscale resolution. Journal of Microscopy, **236**, 35–43.

Wildanger, D., Rittweger, E., Kastrup, L., and Hell, S.W. (2008). STED microscopy with a supercontinuum laser source. Optics Express, **16**, 9614–9621.

Willig, K. I., (2006). *STED microscopy in the visible range*. Doctoral Dissertation, Ruperto-Carola University of Heidelberg, Germany.

Willig, K. I., Harke, B., Medda, R., and Hell, S. W. (2007). STED microscopy with continuous wave beams. Nature Methods, **4**, 915–918.

Willig, K. I., Kellner, R. R., Medda, R., Hein, B., Jakobs, S., and Hell S. W. (2006a). Nanoscale resolution in GFP-based microscopy. Nature Methods, **3**, 721–723.

Willig, K. I., Rizzoli, S. O., Westphal, V., Jahn, R., and Hell, S. W. (2006b). STED-microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis. Nature, **440**, 935–939.

Wisniewski-Barker, E., and Padgett, M. J. (2015). Orbital angular momentum. In: *Photonics: Scientific Foundations, Technology and Applications, Volume I, First Edition. Edited by David L. Andrews. pp. 321–340.*. New York: John Wiley & Sons, Inc.

Wurm, C. A., Kolmakov, K., Göttfert, F., Ta, H., Bossi, M., Schill, H., Berning, S., Jakobs, S., Donnert, G., Belov, V. N., and Hell, S. W. (2012). Novel red fluorophores with superior performance in STED microscopy. Optical Nanoscopy, **1**, 1–7.

Yamanaka, M., Kawano, S., Fujita, K., Smith, N. I., Kawata, S. (2008). Beyond the diffraction-limit biological imaging by saturated excitation microscopy. J. Biomed. Opt. **13**, 050507.

Yamanaka, M., Tzeng, Y.-K., Kawano, S. Smith, N. I., Kawata, S., Chang, H-C., Fujita, K. (2011). SAX microscopy with fluorescent nanodiamond probes for high-resolution fluorescence imaging. Biomed. Opt. Express **2**, 1946–1954

Further Reading

Allen, L., Courtial, J., and Padgett, M. J. (1999). Matrix formulation for the propagation of light beams with orbital and spin angular momenta. Physical Review E, **60**, 7497–7503.

Allen, L., and Padgett, M. (2007). Equivalent geometric transformations for spin and orbital angular momentum of light. Journal of Modern Optics, **54**, 487–491.

Allen, L., and Padgett, M. J. (2000). The Poynting vector in Laguerre-Gaussian beams and the interpretation of their angular momentum density. Optics Communication, **184**, 67–71.

Andresen, M., Stiel, A. C., Fölling, J., Wenzel, D., Schönle, A., Egner, A., Eggeling, C., Hell, S. W., and Jakobs, S. (2008). Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. Nature Biotechnology, **26**, 1035–1040.

Aquino, D., Schönle, A., Geisler, C., Middendorf, C. v., Wurm, C. A., Okamura, Y., Lang, T., Hell, S. W., and Egner, A. (2011). Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores. Nature Methods, **8**, 353–359.

Beijersbergen, M. W., Allen, L., van der Veen, H. E. L. O., and Woerdman, J. P. (1993). Astigmatic laser mode converters and transfer of orbital angular momentum. Optics Communication, **96**, 123–132.

Berning, S., Willig, K. I., Steffens, H., Dibaj, P., and Hell, S. W. (2012). Nanoscopy in a living mouse brain. Science, **335**, 551.

Bierwagen, J., Testa, I., Fölling, J., Wenzel, D., Jakobs, S., Eggeling, C., and Hell, S. W. (2010). Far-field autofluorescence nanoscopy. Nano Letters, **10**, 4249–4252.

Brown, T. G. (2011). Unconventional polarization states: Beam propagation, focusing, and imaging. Progress in Optics, **56**, 81–129.

Chen, Z., Hua, L., and Pu, J. (2012). Tight focusing of light beams: Effect of polarization, phase, and coherence. Progress in Optics, **57**, 219–260.

Cheng, W. (2013). Optical Vortex Beams: Generation, Propagation and Applications. Dissertation Submitted to the School of Engineering of the University of Dayton in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Electro-Optics.

Cremer, C., and Masters, B. R. (2013). Resolution enhancement techniques in microscopy. European Physical Journal H, **38**, 281–344.

Dyba, M., and Hell, S. W. (2002). Focal spots of size $\lambda/23$ open up far-field fluorescence microscopy at 33 nm axial resolution. Physical Review Letters, **88**, 163901-1–163901-4.

Egner, A., Jakobs, S., and Hell, S. W. (2002). Fast 100-nm resolution three-dimensional microscope reveals structural plasticity of mitochondria in live yeast. Proceedings of the National Academy of Sciences of the United States of America, **99**, 3370–3375.

Götte, J.B. (2006). *Integral and fractional orbital angular momentum of light*. Ph.D. Dissertation, Glasgow, Scotland: University of Strathclyde.

Hein, B., Willig, K. I., and Hell, S. W. (2008a). Stimulated emission depletion (STED) nanoscopy of a fluorescent protein-labeled organelle inside a living cell. Proceedings of the National Academy of Sciences of the United States of America, **105**, 14271–14276.

Hein, B., Willig, K. I., Westphal, V., Jacobs, S., and Hell, S. W. (2008b). Stimulated emission depletion nanoscopy of living cells using SNAP-tag fusion proteins. Biophysical Journal, **98**, 158–163.

Heintzmann, R. (2003). Saturated patterned excitation microscopy with two-dimensional excitation patterns. Micron, **34**, 283–291.

Heintzmann, R., Jovin, T., and Cremer, C. (2002). Saturated patterned excitation microscopy—A concept for optical resolution improvement. Journal of the Optical Society of America A. Optics and Image Science, **19**, 1599–1609.

Hell, S. W. (2014). *Nanoscopy with focused light: Lecture slides*. Stefan W. Hell—Nobel Lecture. http://www.rki-i.com/cell_reg2003/hell-lecture-slides.pdf. Accessed April 20, 2019.

Hell, S. W. (1997). Increasing the Resolution of Far-Field Fluorescence Microscopy by Point-Spread-Function Engineering. In: *Topics in Fluorescence Spectroscopy; **5**: Nonlinear*

and Two-Photon-Induced Fluorescence, edited by J. Lakowicz. New York: Plenum Press, pp. 361–426.

Hell, S. W. (2007). Far-field optical nanoscopy. Science, **316**, 1153–1158.

Hell, S. W., Schrader, M., and van der Voort, H. T. M. (1997). Far-field fluorescence microscopy with three-dimensional resolution in the 100-nm range. Journal of Microscopy, **187**, 1–7.

Hell, S. W., Schmidt, R., and Egner, A. (2009). Diffraction-unlimited three-dimensional optical nanoscopy with opposing lenses. Nature Photonics, **3**, 381–387.

Kastrup, L. (2004). *Fluorescence depletion by stimulated emission in single-molecule spectroscopy*. Doctoral Dissertation, Ruperto-Carola University of Heidelberg, Germany.

Keller, J., Schönle, A., and Hell, S. W. (2007). Efficient fluorescence inhibition patterns for RESOLFT microscopy. Optics Express, **15**, 3361–3371.

Kellner, R. R. (2007). *STED microscopy with Q-switched microchip lasers*. Dissertation submitted to the Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Germany for the degree of Doctor of Natural Sciences.

Klar, T. A. (2001). *Progress in stimulated emission depletion microscopy*. Doctoral Dissertation. Ruprecht-Karls-Universität Heidelberg.

Kuśba, J., Bogdanov, V., Gryczynski, I., and Lakowicz, J. R. (1994). Theory of light quenching: Effects on fluorescence polarization, intensity, and anisotropy decays. Biophysical Journal, **67**, 2024–2040.

Lalkens, B., Testa, I., Willig, K. I., and Hell, S. W. (2011). MRT letter: Nanoscopy of protein colocalization in living cells by STED and GSDIM. Microscopy Research and Technique, **75**, 1–6.

Lauterbach, M. A., Keller, J., Schonle, A., Kamin, D., Westphal, V., Rizzoli, S. O., and Hell, S. W. (2010). Comparing video-rate STED nanoscopy and confocal microscopy of living neurons. Journal of Biophotonics, **3**, 417–424.

Leutenegger, M., Ringemann, C., Lasser, T., Hell, S. W., and Eggeling, C. (2012). Fluorescence correlation spectroscopy with a total internal reflection fluorescence STED microscope (TIRF-STED-FCS). Optics Express, **20**, 5243–5263.

Liu, Y, Ding, Y., Alonas, E., Zhao, W., Santangelo, P. J., Jin, D., Piper, J. A., Teng, J., Ren, Q., and Xi, P. (2012). Achieving $\lambda/10$ resolution CW STED nanoscopy with a Ti:Sapphire oscillator. PLoS One, **7**, e40003. https://doi.org/10.1371/journal.pone.0040003.

Liu, Y., Kuang, C., and Liu, X. (2015). The use of azimuthally polarized sinh-Gauss beam in STED microscopy. Journal of Optics, **17**(4), 045609.

Maleev, I. D., and Swartzlander, G. A., Jr. (2003). Composite optical vortices. Journal of Optical Society of America B, **20**, 1169–1176.

Meinecke, F. (1996). *Stimulierte Emission im Fluoreszenzmikroskop: Das STED-Konzept zur Überwindugn der Abbeschen Beugungsgrenze*. Diploma thesis, Ruperto-Carola University of Heidelberg, Germany.

Moneron, G., and Hell, S. W. (2009). Two-photon excitation STED microscopy. Optics Express, **17**, 14567–14573.

Okhonin, V. A. (1991). Method of investigating specimen microstructure, Patent SU 1374922, (See also in the USSR patents database SU 1374922) priority date April 10, 1986, Published on July 30, 1991, Soviet Patents Abstracts, Section EI, Week 9218, Derwent Publications Ltd., London, GB; Class S03, p. 4. Cited by patents US 5394268 A (1993) and US RE38307 E1 (1995). From the [https://www.researchgate.net/profile/Victor_Okhonin/publication/272021175_STED_Priority_1986_Eng_Transl/links/54d8ca860cf2970e4e793c8b.pdf?origin=publication_detailEnglishtranslation]. Accessed August 20, 2017.

Pezzagna, S., Rogalla, D., Wildanger, D., Meijer, J., and Zaitsev, A. (2011). Creation and nature of optical centres in diamond for single-photon emission—Overview and critical remarks. New Journal of Physics, **13,** 1–27.

Punge, A., Rizzoli, S. O., Jahn, R., Wildanger, J. D., Meyer, L., Schönle, A., Kastrup, L., and Hell, S. W. (2008). 3D reconstruction of high-resolution STED microscope images. Microscopy Research and Technique, **71**, 644–650.

Rankin, B. R., Kellner, R. R., and Hell, S. W. (2008). Stimulated-emission-depletion microscopy with a multicolor stimulated-Raman scattering light source. Optics Letters, **33**, 2491–2493.

Richards, B., and Wolf, E. (1959). Electromagnetic diffraction in optical systems II. Structure of the image field in an aplanatic system. Proceedings of the Royal Society of London A, **253**, 358–379.

Rittweger, E., Han, K. Y., Irvine, S. E., Eggeling, C., and Hell, S. W. (2009). STED microscopy reveals crystal colour centres with nanometric resolution. Nature Photonics, **3**, 144–147.

Schmidt, R., Wurm, C. A., Jakobs, S., Engelhardt, J., Egner, A., and Hell, S. W. (2008). Spherical nanosized focal spot unravels the interior of cells. Nature Methods, **5**, 539–544.

Tinnefeld, P., Eggeling, C., and Hell, S. W. (2015). *Far-Field Optical Nanoscopy*. Berlin: Springer.

Török, P., and Munro, P. (2004). The use of Gauss-Laguerre vector beams in STED microscopy. Optics Express, **12**, 3605–3617.

Vaughan, J. M., and Willetts, D. V. (1983). Temporal and interference fringe analysis of TEM_{01}^{*} laser modes. Journal of Optical Society of America, **73**, 1018–1021.

Warren, W. S., Rabitz, H., and Dahleh, M. (1993). Coherent control of quantum dynamics: The dream is alive. Science, **259**, 1581–1589.

Weiss, S. (2000). Shattering the diffraction limit of light: A revolution in fluorescence microscopy? Proceedings of the National Academy of Sciences of the United States of America, **97**, 8747–8749.

Wildanger, D., Maze, J. R., and Hell, S. W. (2011). Diffraction unlimited all-optical recording of electron spin resonances. Physical Review Letters, **107**, 017601-1–017601-4.

Xue, Y., Kuang, C., Li, S., Gu, Z., and Liu, X. (2012). Sharper fluorescent super-resolution spot generated by azimuthally polarized beam in STED microscopy. Optics Express, **20**, 17653–17666.

Zhan, Q. (2009). Cylindrical vector beams: From mathematical concepts to applications. Advances in Optics and Photonics, **1**, 1–57.

Chapter 15
Localization Microscopy with Active Control

> "Superresolution alone does not provide more information than any other imaging modality unless the experiments are well designed. The question we have to pose is whether these images are biologically meaningful, that is, can we ensure that the results are physiologically relevant? To do so, we must perform careful live-cell control and correlative experiments with alternative superresolution imaging methods, vary the irradiation doses and labeling densities, and exchange the fluorescent probes."
>
> —van de Linde, Heilemann, and Sauer (2012)

15.1 Introduction

How does localization microscopy achieve superresolution? What do we mean by the word "localization"? The answer to these questions is the content of this chapter. First, I define some terms. Localization refers to the technique of locating the centroid or geometric center of the point spread function (PSF) of a fluorescent molecule. The localization of a molecule can be obtained with higher precision than resolution; precision increases with larger numbers of detected photons and with decreasing background fluorescence. Resolution (Chapter 2) refers to the smallest distance that separates two objects for which the two objects are imaged as two distinct objects. Photoswitching fluorescent molecules is a reversible or irreversible process in which a photon can alter the emission wavelengths of the molecule (e.g., from green fluorescent to red fluorescent). Photoactivation is the process in which molecules, which are nonfluorescent (in the OFF state) before they are activated with blue or ultraviolet light, exhibit normal absorption and emission spectra (in the ON state). Reversible photoactivated fluorescent molecules can cycle between the nonfluorescent OFF state and the fluorescent ON state many times; irreversible photoactivated fluorescent molecules can only be activated to the fluorescent state

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_15

once. Superresolution microscopy achieves imaging of object details with a resolution that exceeds diffraction-limited resolution.

In a volume that has a Cartesian coordinate system the x–y coordinates define the lateral location and the z coordinate defines the axial location. If a molecule is on a dark background and is induced to emit a large number of photons a diffraction-limited spot can be readily detected. This spot of fluorescence can be fitted with a two-dimensional Gaussian function and the x, y coordinates of the centroid can be calculated (Thompson et al., 2002). In the absence of background noise, or when the background noise is insignificant compared with the fluorescent signal, the precision with which the centroid of the fluorescent spot can be located is a function of the number of emitted photons that are detected, and precision scales inversely with the square root of the number of detected photons and the standard deviation of the PSF (Mortensen et al., 2010). With the detection of a large number of photons from the fluorescent molecule the original PSF greatly exceeds the error in locating the position of the fluorescent molecule.

But there is a problem. When several thousand fluorescent molecules (typical in high-density labeling of cells) are located close together (i.e., the separation of the individual fluorescent molecule is less than the resolution of the microscope), these molecules are not able to be distinguished because the PSF of each molecule overlaps with those of adjacent molecules. The key to resolving these closely spaced fluorescent molecules is to somehow distinguish which photons are emitted by which of the fluorescent molecules. If we can switch the state of each of the molecules from the nonfluorescent OFF state to the fluorescent ON state, then we can determine which emitted photons come from which individual fluorescent molecules. If a sufficiently large number of photons are detected, then we can locate the coordinates of each fluorescent molecule in the focal plane. In summary, we can calculate the centroid of the PSF of a fluorescent molecule to a precision that exceeds the resolution of the microscope (Thompson et al., 2002).

That in a nutshell is how localization microscopy with active control (to distinguish which photons are emitted from which molecules) can achieve superresolution microscopy (Moerner, 2012). Separate fluorescent molecules that are temporally separated from their neighbors are randomly or stochastically activated, and the coordinates of each emitter is recorded. The final image is sequentially built up from the many coordinates of fluorescent molecules. Resolution of the final image is no longer limited by light diffraction; the precision of each localization and their number limits resolution. Furthermore, the density of the fluorescent label in the specimen and the size of the fluorescent probes are important factors in achieving the ultimate spatial resolution.

This chapter discusses pathways to achieving superresolution optical microscopy based on localization with active control. The detailed content presents the antecedents, the theoretical background, and the implementation of the theory into new types of superresolution microscopes, as well as their practical operation, limitations, and applications.

In 1989 Moerner and Kador published their seminal study titled “Optical detection and spectroscopy of single molecules in a solid” (Moerner and Kador, 1989). This publication motivated others to work on single-molecule imaging.

Localization microscopy is a microscopy technique based on single-molecule photocontrol (Moerner, 2014a, b). I highly recommend a study of the slides from Moerner's Nobel Prize lecture titled "Single-Molecule Spectroscopy, Imaging, and Photocontrol: Foundations for Super-Resolution Microscopy"—lecture slides (Moerner, 2014a, b). They are a useful source of interdisciplinary strategies and experimental details that provide an excellent introduction to the requisite spectroscopy of single molecules in condensed matter (Basché et al., 1997).

Comprehensive background material for this chapter can be found in the following recommended review articles: "Superresolution imaging using single-molecule localization" (Patterson et al., 2010); "Live-cell super-resolution imaging with synthetic fluorophores" (van de Linde et al., 2012); "Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging" (Ha and Tinnefeld, 2012); "Extending microscopic resolution with single-molecule imaging and active control" (Thompson et al., 2012); "Photocontrollable fluorescent proteins for superresolution imaging" (Shcherbakova et al., 2014); "Superresolution localization methods" (Small and Parthasarathy, 2014); "Fluorescence imaging for bacterial cell biology: from localization to dynamics, from ensembles to single molecules" (Yao and Carballido-López, 2014); "Modern statistical challenges in high-resolution fluorescence microscopy" (Aspelmeier et al., 2015); and "Lens-based fluorescence nanoscopy" (Eggeling et al., 2015).

This genesis had its origins in advances in single-molecule detection, an active field of research spanning several decades (Moerner, 2007, 2012). A typical organic fluorescent molecule is 1 nm in size; however, its image in a diffraction-limited fluorescent microscope is a few hundred nanometers. A point source of fluorescence is imaged in the light microscope as the point spread function (PSF).

Superresolution microscopy based on localization with active control has a long genesis. In 1976 Hirschfeld labeled a single protein with 100 fluorescent molecules and was able to image single protein molecules with his fluorescence microscope (Hirschfield, 1976). There were seminal advances during the following decades: "The optical detection of single-molecules in a solid via their absorption spectra at 1.6 K" (Moerner and Kador, 1989); "The detection of single fluorescent molecules at room temperature in solution" (Shera et al., 1990); "Illuminating single molecules in condensed matter" (Moerner and Orrit, 1999); "Single pentacene molecules detected by fluorescence excitation in a paraterphenyl crystal at 1.8 K" (Orrit and Bernard, 1990); and "Fluorescence spectroscopy of single biomolecules" (Weiss, 1999).

Two critical new ideas were required for success in achieving superresolution microscopy with single molecules (Moerner, 2014a, b). The first new idea was superlocalization (i.e., the detection and localization of fluorescent molecules separated in space). A single point source of fluorescence will not appear as a point but as a PSF whose width is defined by the diffraction limit. The foundation of superlocalization was to fit the PSF with a two-dimensional Gaussian function and

then calculate the centroid of the PSF with a precision exceeding the diffraction limit of the microscope.

A severe constraint of the technique was the requirement that there be no overlap of adjacent PSFs that would occur if two fluorescent molecules are closer than 200 nm from each other. A very low density of fluorescent molecules that are emitting photons in the specimen is mandatory to ensure compliance with the requirement of no overlap of adjacent PSFs. But, for optimal superresolution imaging the specimen needs to have a high labeling density. The fluorescent labeling density is defined as the number of fluorescent molecules per square micrometer. These opposing constraints required a new solution and in 2006 it appeared in three independent yet similar versions. While I present these inventions in separate sections, the similarities and differences among these superresolution localization microscopy techniques will become evident.

The innovative solutions were the inventions of photoactivated localization microscopy (PALM), stochastic optical reconstruction microscopy (STORM), and fluorescence photoactivated localization microscopy (FPALM) discussed in the following sections (Betzig et al., 2006; Hess et al., 2006; Rust et al., 2006). These inventions as well as their variants overcame severe constraints by localizing extremely sparse groups of photoswitchable or photoactivatable molecules.

Further insight into the putative attribution of priority for these three inventions can be obtained by a more detailed study of their publication histories and a presentation at a National Institutes of Health (NIH) conference one month prior to submission of the PALM paper to *Science*.

The three inventions of the three independent groups were published in 2006; however, their submission dates showed considerable variation. I order these papers from the earliest submission date to the latest submission date. Photoactivated localization microscopy (PALM) by Betzig, Patterson, Sougrat, Lindwasser, Olenych, Bonifacino, Davidson, Lippincott-Schwartz, and Hess [Harald F. Hess] (2006) was submitted to *Science* on March 13, 2006, accepted on August 2, 2006, and had an online publication date of August 10, 2006; fluorescence photoactivated localization microscopy (FPALM) by Hess [Samuel T. Hess], Girirajan, and Mason (2006) was submitted to the *Biophysical Journal* on June 12, 2006 and published in print in December 2006; and, lastly, stochastic optical reconstruction microscopy (STORM) by Rust, Bates, and Zhuang (2006) was submitted to *Nature Methods* on July 7, 2006 and had an online publication date of August 9, 2006.

Furthermore, it is extremely instructive to look at the NIH conference titled "2006 Frontiers in Live Cell Imaging" that was held in Bethesda, Maryland. On April 21, 2006 Eric Betzig presented a talk titled "New Approaches to Intracellular Imaging at High Spatial and Temporal Resolution." The webcast can be found at https://videocast.nih.gov/summary.asp?Live=4926&bhcp=1. Betzig's talk (day 3) begins at 1 h and 38 min into the video (accessed April 6, 2019).

Betzig clearly described the essentials of the PALM technique (i.e., the use of photoactivated fluorescent proteins expressed in cells), the important requirement of a very high labeling density, and the requirement that cells be fixed. With random

or stochastic illumination individual molecules could be made to fluoresce, and from the centroid of the PSF they could be localized at very high precision. After a molecule has emitted photons it is irreversibly bleached. The location of fluorescent molecules from thousands of images are mapped into the final superresolution image. Betzig showed data on fixed cells that achieved 10 nm transverse spatial resolution. He demonstrated that the PALM technique works when several thousand fluorescent molecules are located in a diffraction-limited volume. Betzig described the trade-offs between the density of fluorescent molecules on the specimen, the limitation of overexpression of genetically modified proteins, the importance of collecting a large number of photons, and the use of electron microscopy to validate superresolution optical images. The data presented took 12 h to acquire and Betzig discussed an optical lattice microscope that could significantly reduce the data acquisition time (Betzig, 2005). The problem of specimen drift was solved by the use of gold nanoparticles as fiduciary markers. He discussed the importance of comparative (correlative) transmission electron microscopy to validate PALM imaging. Additionally, he proposed new approaches to get spatial resolution to go beyond 10 nm.

In response to a question (the last question that followed Betzig's talk) by Xiaowei Zhuang on the accuracy of the 10 nm resolution that PALM achieved, Betzig stated that the PALM technique yields molecular spatial resolution with high precision, but the accuracy of the location is more difficult to determine and he thinks correlative transmission electron microscopy is useful in this task. Betzig's theoretical and experimental results presented in his April 21, 2006 lecture, including the images, became the basis of the PALM manuscript submitted to *Science* on March 13, 2006 and published online on August 10, 2006. That was several months prior to Xiaowei Zhuang's submission of the STORM manuscript to *Science* on July 7, 2006.

Photoactivatable fluorescent proteins (PA-FPs) can be changed from a nonfluorescent OFF state to a fluorescent ON state (photoactivation) by ultraviolet or blue light, or they can undergo light-induced changes in their spectral properties (photoswitching or photoconversion). Light of different wavelengths can switch these reversible photoswitchable fluorescent molecules between their OFF states and their ON states. Photobleaching is a different process as it is an irreversible transformation induced by light. Photoactivatable molecules include photoactivatable fluorescent proteins such as photoactivatable Green Fluorescent Protein (PA-GFP).

The second new idea was active control of emitting molecules so that there is a very sparse concentration of single fluorescent molecules that are in the ON or bright state; this satisfied the requirement that there be no overlap of PSFs during a single imaging frame (Moerner, 2012). The final image is formed from a large (tens of thousands) series of time-sequential imaging and superlocalization of single emitting fluorescent molecules that labeled the specimen. The final image consists of the combined locations of all the single fluorescent molecules that result from imaging the separate randomly (stochastically) formed sparse assembly of molecules. The combination of these two new ideas led to the success of superresolution microscopy based on localization with active control.

In 1995 Betzig proposed that superresolution could be achieved through spectral tunability (Betzig, 1995). Three-dimensional superresolution was achieved in the laboratory at low temperatures (van Oijen et al., 1999), and later at room temperature using semiconductor quantum dots that blinked (Lidke et al., 2005).

It is important to understand the difference between a deterministic process and a stochastic or random process. STED is an example of a deterministic superresolution imaging technique that uses stimulated emission to restrict fluorescence emission to specific regions that are smaller than the limit imposed by diffraction-limited resolution. Spatial resolution can be increased in the focal region by reducing the effective PSF. Scanning the specimen by precisely translating it through the optical axis of the STED microscope results in formation of a superresolution image.

Such a process is very different from stochastic or random processes that are the basis of the localization superresolution microscope techniques of PALM, FPALM, STORM, and dSTORM (discussed in subsequent sections). These stochastic superresolution imaging techniques use random or stochastic photoswitching or photoactivation, localization, and sequential reconstruction (i.e., the sum of all frames acquired from all localizations). Each frame is a superlocalization of individual, single fluorescent molecules.

How does single-molecule imaging relate to superresolution microscopy? The abovementioned superresolution techniques have a common foundation involving the use of single-molecule switching that locates single molecules, the use of photoswitching or photoactivation of fluorescent molecules to ensure that at a specific time there is one fluorescent molecule emitting photons in a specific diffraction-limited focal volume of the specimen. The final image of the specimen is reconstructed from many such cycles that are repeated to detect the photons (and therefore the positions of all single molecules in the specimen) of all the fluorophores in the specimen. These techniques serve to separate temporally and spatially the images of single fluorescent molecules, thereby forming a superresolution image of the specimen. Fluorescent molecules are switched between a nonfluorescent OFF state and a fluorescent ON state.

Although the localization of single molecules is a necessary condition, it is not sufficient to achieve superresolution microscopic imaging. A means of photoswitching fluorescent molecules is needed to provide the capability to separate the many fluorescent molecules that are in the diffraction-limited focal volume.

Note that superresolution microscope techniques are based on fluorescent labels. The image is not of the microscopic and nanoscopic structures themselves, but of the photoswitchable fluorophores that are attached to these structures. As superresolution techniques move to ever-higher spatial resolutions a point is reached where resolution is similar to the size of fluorescent proteins, which becomes the limiting resolution.

Throughout Part III of this book I discuss publications on superresolution microscopy. Since my aim is not to write a review article I selected those articles that were the first to deal with a new invention or technique in the areas of

instrumentation, probe development and application, specimen preparation, image interpretation and validation, and control experiments. The pedagogical content of each publication was also an important criterion for selection. Although I read several hundred publications, I used my judgment to select a small subset from the total corpus of published works. For a complete search of the publications on superresolution microscopy I refer the reader to review articles, such as Shtengel et al., (2014), many of which are cited in the references to each chapter.

15.2 Antecedent Publications

I now introduce three prescient works published in 1985, 1995, and 1999 by independent groups of researchers that predated the first publications of PALM, FPALM, and STORM, which occurred in 2006. These antecedent publications pointed the way for these inventions. Other pre-2006 publications that I will subsequently discuss give credence to the postulate that invention sometimes builds on the prior creative works of others. This postulate is supported by perusal of the literature on these techniques and their variants.

The first publication titled "Strategies for attaining superresolution using spectroscopic data as constraints" used computer simulations to demonstrate that superresolution can be attained for simulated point sources of light within 1/30th of the Rayleigh distance irrespective of whether the transfer function is wavelength dependent or wavelength independent (Burns et al., 1985). The authors' problem involved finding a way to achieve superresolution via the restoration of bandwidth-limited images? The specific case they simulated involved two point sources of light located within the Rayleigh distance of each other; each with different spectral properties. The authors assumed the image was a set of intensity values, a mathematical function of both positions and wavelengths. They then assumed that linear superposition was valid and therefore used the methods of rank and eigenanalysis to obtain the number of objects that were spectrally different. They then calculated the objects' location in space and the spectrum of each object. This early approach missed the critical concept of photoswitching that is key to localizing and then separating the large numbers of fluorescent molecules situated in the diffraction-limited volume of the specimen.

The second publication titled "Proposed method for molecular optical imaging" (Betzig, 1995) appeared a decade later. Betzig was experienced in the design and use of near-field scanning optical microscopy (NSOM), which uses an illuminated subwavelength-sized aperture to form a superresolution image. His motivation was to extend optical imaging to the superresolution regime (i.e., to reach beyond diffraction-limited imaging).

The problem Betzig strived to solve was the separation of fluorescent molecules located in the same focal volume since the PSFs of single fluorescent molecules can overlap. Fluorescent molecules exist in a space of m-spatial dimensions. There are

n-optical parameters that characterize fluorescent molecules. The space has $m + n$ dimensions. Betzig proposed a two-step solution to this dilemma. First, one or more unique optical characteristics would be used to identify and isolate each feature. Second, the spatial coordinates of each feature would be determined (i.e., the centroid of the PSF for each feature). Finally, the total set of coordinates for all the features would form the final image. Betzig also proposed an alternative technique: the application of an intense spatial gradient that could affect one or more of the *n*-optical parameters of fluorescent molecules. A practical way of achieving this is via the Stark shift that can be used to change the spectral parameters of adjacent molecules. Betzig proposed two general ideas that are key to achieving superresolution imaging: the localization of individual molecules and the separation of adjacent molecules within the focal volume by using a technique to modify the optical properties of these adjacent molecules.

The third publication I discuss is titled "Far-field fluorescence microscopy beyond the diffraction limit" (van Oijen et al., 1999). The problem facing the authors involved developing a far-field optical microscope capable of forming three-dimensional images with a length scale much below the Rayleigh limit? Perusal of the first two publications shows the connection and the progress made toward finding a solution to their common problem: the development of a far-field superresolution light microscope. van Oijen et al. (1999) specifically credits the prior publication of Burns et al. (1985) for suggesting the original idea.

van Oijen et al. achieved their goal by coming up with a new technique they called spectrally selective imaging (SSI). SSI can locate and separate fluorescent molecules located within the focal volume where the PSFs of adjacent molecules overlap. High-resolution laser spectroscopy is used to spectrally designate a single fluorescent molecule in the focal volume. This is accomplished by tuning the excitation laser into resonance with another fluorescent molecule; the position of individual molecules can be determined for all the molecules located within the diffraction-limited volume. A CCD camera detects the three-dimensional spatial distribution of photons from the fluorescent molecule allowing its position to be determined with very high accuracy. SSI is applied over and over to determine the precise locations of all the fluorescent molecules in the diffraction-limited volume.

van Oijen et al. demonstrated the SSI technique by imaging pentacene in p-terphenyl at a temperature of 1.2 K. The SSI technique achieved localization of fluorescent molecules with a lateral resolution of 40 nm and an axial resolution of 100 nm (van Oijen et al., 1999).

The SSI technique has its limitations (van Oijen et al., 1999). The authors state that it would be a challenge to apply their technique to biological specimens. What is the foundation for the limiting separation of two adjacent molecules that the SSI technique can resolve? The SSI technique will fail when the interaction energy of the transition–dipole moments of each molecule is similar to the energy of the induced transition. The excitation energy will not be specific to a single fluorescent molecule and the two adjacent molecules will fluoresce.

While unsuitable as a general technique for biological specimens at room temperature, it can be viewed as a close precursor to PALM, FPALM, and STORM. Again, the missing factor inherent to these techniques is the use of stochastic photoactivation/photoswitching to separate the thousands of molecules located in the same diffraction-limited volume. Such high-labeling densities are required for accurate superresolution imaging of biological specimens.

In this section I explore prescient studies on single-molecule localization and on the necessary condition of photoswitching of fluorophores between the nonfluorescent OFF state and the fluorescent ON state. Perusal of the publications related to these techniques reveals the synergistic interaction between molecular spectroscopy, photoswitchable probe development, and single-molecule localization and imaging. Integration of all these areas of research resulted in the development of superresolution microscopic imaging based on stochastic switching of molecules and their localization.

Localization of a fluorophore with nanometer precision preceded the invention of superresolution microscopy imaging techniques (described later). Of the many examples in support of this statement I choose the 1988 publication titled "Tracking kinesin-driven movements with nanometer-scale precision" (Gelles et al., 1988) and the 2003 publication titled "Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization" (Yildiz et al., 2003).

Progress in this field depended on advances in the spectroscopy of fluorescent molecules (Fernández-Suárez and Ting, 2008; Lukyanov et al., 2005). The 2005 *Nature Reviews Molecular Cell Biology* publication of Lukyanov et al. titled "Photoactivatable fluorescent proteins" evaluates the properties of available photoactivatable fluorescent proteins and their potential applications. In 2008 Fernández-Suárez and Ting published a review article titled "Fluorescent probes for super-resolution imaging in living cells," which describes the contributions of fluorescent probes to far-field superresolution imaging, focusing on fluorescent proteins and organic small-molecule fluorophores. This comprehensive review gives an in-depth comparison of the various biological imaging techniques including superresolution methods in terms of spatial and temporal resolution. Moreover, it provides a good review of various genetically encoded fluorescent proteins, quantum dots, and methods for site-specific targeting of small-molecule probes to cellular proteins.

Seminal developments in the directed mutagenesis of photoactivated fluorescent probes highlight the role probe development plays in the field of superresolution microscopy (McKinney et al., 2009; Patterson et al., 2010; Patterson and Lippincott-Schwartz, 2002; Shcherbakova et al., 2014). The linear dimensions of many organic fluorescent molecules are 1 nm, fluorescent antibodies are 10-15 nm, and fluorescent proteins, such as Green Fluorescent Protein (GFP), are 3 nm.

Another group of investigators published a seminal paper on in vivo photoactivation of GFP titled "Spatial dynamics of GFP-tagged proteins investigated by

local fluorescence enhancement" (Yokoe and Meyer, 1996), in which the authors describe a method of locally enhancing the blue excited fluorescence of GFP using a spatially focused ultraviolet laser pulse.

An early prescient approach was the publication titled "High-resolution colocalization of single dye molecules by fluorescence lifetime imaging microscopy" (Heilemann et al., 2002). Their technique was based on fluorescence lifetime imaging. Their method could distinguish and measure the distance between two dye molecules that were adjacent and separated by less than 30 nm. To demonstrate their method the authors acquired lifetime images of a mixture of Cy5 and JF9 (rhodamine derivative) molecules randomly adsorbed on a glass surface. Since these two different molecules have two different fluorescence lifetimes (Cy5 of 2.0 ns and JF9 of 4.0 ns) it is possible to assign the contribution of fluorescence made by the two fluorescent molecules to each image pixel using a pattern recognition technique. Both fluorescent molecules can be excited using the same laser wavelength. The authors achieved the first high-precision distance measurements between single conventional fluorescent molecules using only the difference in the fluorescence lifetime of the molecules.

Again, I stress that probe development is key to advancng the field. For example, the publication titled "Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer" provided the results researchers needed to produce mutations that could modify the absorption spectrum and enhance the photostability of GFP (Chattoraj et al., 1996).

Photocontrol or photoswitching of fluorescent molecules is critical to superresolution as it permits single fluorescent molecules to be detected from a multitude of adjacent molecules. Moerner and his colleagues' publication titled "On/off blinking and switching behavior of single molecules of green fluorescent protein" was an early example of this technique (Dickson et al., 1997). What was most significant was the authors' prescient proposal, which followed from their experiments, that GFP mutants could be useful as molecular photonic switches capable of being altered at the single-molecule stage (Dickson et al., 1997). The authors first immobilized GFP mutants in aqueous polymer gels that they aerated. When GFP mutants located in the gels were irradiated with 488-nm light they underwent many cycles of on/off fluorescence that lasted several seconds. This on/off emission over time is termed blinking. After several seconds of this blinking the GFP fluorescent molecules spontaneously entered a nonfluorescent OFF state. Subsequent excitation of these molecules in the OFF state with light of 405 nm photoswitched these molecules from the OFF state to a fluorescent ON state. These experimental results and the authors' prescient proposal of controlling photoswitchable fluorescent molecules at the single-molecule level was not lost on the inventors of PALM, FPALM, and STORM.

Next, I introduce the Nyquist–Shannon sampling theorem (a.k.a. the Whittaker–Shannon sampling theorem, among other names) and its application (Gonzalez and Woods, 2008; Goodman, 2017). This is another example of multiple inventions occurring sometimes without giving proper attribution to original works. Working at

Bell Laboratories (USA) Harry Nyquist published two key papers in 1928 (Nyquist, 1928a, b). In 1949 Claude E. Shannon, also working at Bell Laboratories, proved the Nyquist theorem (Shannon, 1949). The theorem is now referred to as the Nyquist–Shannon sampling theorem, the foundation of digital signal processing theory. The Nyquist–Shannon sampling theorem states that a continuous, band-limited function can be reconstructed from samples if the samples are detected at a rate that exceeds the highest frequency content of the function. Another important definition is given by the Nyquist rate, which is a sampling rate equal to twice the highest frequency. Application of the Nyquist–Shannon sampling theorem to digital image processing is best described in *Digital Image Processing*, third edition, which analyzed sampling methods as well as the Fourier transform of sampled function (Gonzalez and Woods, 2008). The one-dimensional sampling theorem can be readily extended to the imaging case of two-dimensional sampling (Goodman, 2017). An important caveat applied to both cases: if the function is undersampled (i.e., sampled at a rate that is less than twice its highest frequency), then frequency aliasing occurs which modifies the reconstructed function or the digital image.

The Nyquist–Shannon sampling theorem can be used to evaluate image resolution in densely labeled specimens (van de Linde et al., 2012). Labeling density in superresolution microscopy affects the final resolution, and the required density of the fluorescent molecules in the specimen must be sufficiently high to conform with the constraints of the Nyquist–Shannon sampling theory.

What does this specifically mean in the context of superresolution light microscopy? The Nyquist–Shannon sampling theorem provides guidance on the labeling density of the specimen with fluorescent molecules required such that the desired spatial resolution can be achieved. The distance between adjacent fluorescent molecules must be two times smaller than the spatial resolution that has to be achieved. This critical point can be illustrated in the following way: to achieve a spatial resolution of 20 nm the fluorescent molecules must be attached to the specimen with a separation of 10 nm. Another way to visualize this constraint and comply with the previous desired spatial resolution is to have a density of fluorescent molecules of 10^4 per square micrometer, which corresponds to 600 fluorescent molecules situated within the diffraction-limited focal volume of the microscope (van de Linde et al., 2012).

There are various ways to achieve three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit. The next publication I present supports the premise that creative innovation in superresolution microscopy is a work in progress and that there are many opportunities for further achievements. Pavani et al. in their publication titled “Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function” present a unique microscopy technique that achieves three-dimensional superresolution (Pavani et al., 2009). Our understanding of a rotating PSF dates from 1996 (Schechner et al., 1996).

The key concept behind their microscope is a unique type of PSF given to a single fluorescent molecule: this PSF has two lobes in the image plane of the microscope and is called a double-helix PSF (DH-PSF) because on the z-axis its

shape is a double helix, and the angle of the line between the two lobes depends on the axial position of the fluorescent molecule. The two lobes of the DH-PSF are rotated with the axial (z)-position of the fluorescent molecule. From the rotation of the PSF the axial position of the point source of light (the fluorescent molecule) can be determined. However, to prevent uncertainty as to the location of the fluorescent molecule the rotation of the PSF must be less than a complete circle or 360° (Goodman, 2017). Goodman detailed a complete four-page mathematical derivation of a rotating PSF for depth resolution (Goodman, 2017, pp. 237–240).

The basic idea is that the PSF rotates as the point source of light is located at increasing distances above or below the focal plane. This type of PSF can be used to locate the point source of light in three dimensions. The PSF is formed from superpositions of Gaussian–Laguerre modes.

The authors demonstrated that single molecules were localized in three-dimensional with 10 to 20 nm precision in single 500 ms acquisitions. The technique uses a photoactivatable fluorescent molecule, 2-dicyanomethylene-3-cyano-2,5-dihydrofuran, to achieve three-dimensional superresolution (Pavani et al., 2009).

This section will now segue into a discussion of the 2006 inventions of PALM, FPALM, and STORM, and the later inventions of their variants. These independent and seminal inventions of new types of superresolution microscopy solved the problem of imaging molecules spaced closer than the diffraction limit. The key to their success in imaging densely labeled specimens is the following sequence and its variations: repeated stochastic photoactivation, image acquisition, superlocalization, and photobleaching of nonoverlapping subsets of fluorescent molecules.

In summary, the final superresolution image is formed from the sum of many separate images. The key is to alter transitions between the ON state and the OFF state of individual fluorescent molecules. When the molecule is in the ON state the photons emitted (the more the better the resolution) are detected and the position of individual molecules can be determined (localization). The fluorescent molecules are then switched to the OFF state. Next, another random set of fluorescent molecules are turned on and the previous processes are repeated tens of thousands of times (Eggeling et al., 2015).

15.3 Photoactivated Localization Microscopy

15.3.1 Steps Toward Photoactivated Localization Microscopy

Photoactivation is the subject of this section, which introduces some of the salient principles of photoactivated localization microscopy (PALM), fluorescence photoactivated localization microscopy (FPALM), and interferometric photoactivated

localization microscopy (iPALM). I also discuss some of the major limitations, problems, cautions, caveats, and trade-offs relevant to the superresolution microscopy of biological specimens. I highly recommend the reader study the slides from Betzig's Nobel Prize lecture as well as the lecture itself titled *Single Molecules, Cells, and Super-Resolution Optics* (Betzig, 2014a, b).

Betzig and his colleagues were responsible for advances that paved the way toward superresolution imaging. This took a number of years with progress first being made in near-field imaging and subsequently in far-field superresolution microscopy when the seminal advances of others were incorporated. I will discuss some of the key influential publications.

The starting point is of course arbitrary and I begin with a paper in *Science* titled "Single molecules observed by near-field scanning optical microscopy" (Betzig and Chichester, 1993). This paper contained a number of firsts. The first room temperature near-field imaging of single molecules, the first superresolution image of single molecules, and the first localization of single molecules. The paper was further significant in that the research it contained became the inspiration behind ideas that led to the invention of PALM.

A second seminal paper in *Science* titled "Near-field spectroscopy of the quantum constituents of a luminescent system" described the separation of high-density sites in a single quantum well (Hess et al., 1994). This major achievement came about as a result of using near-field microscopy at cryogenic temperature. However, this led the way to separating fluorescent molecules in densely labeled specimens destined for imaging by far-field microscopy. The requirement was that fluorescent molecules in the ON state be separated from adjacent identical molecules by distances significantly greater than the width of the PSF (Hess et al., 1994).

In the next decade prescient publications paved the way to achieving the first and most difficult step toward Betzig's 1995 proposal for superresolution optical imaging: how to identify and isolate each fluorescent molecule from specific optical properties (Betzig, 1995). If there is only one fluorescent molecule in the diffraction-limited PSF, then we can obtain the coordinates of the centroid of the PSF with enhanced precision. In fact, if a sufficient number of emitted photons are detected from a single fluorescent molecule, we can obtain nanometer precision as to its localization. But, what happens if more than one fluorescent molecule is located within the PSF? That is a difficult problem to overcome.

In the discussion of the next two publications I address the authors' solution to bringing about superresolution microscopy when there are many single fluorescent molecules located within the diffraction-limited focal volume of the specimen.

A paper titled "Nanometer-localized multiple single-molecule (NALMS) fluorescence microscopy," based on both centroid localization and photobleaching of single fluorophores, represented a milestone toward the invention of PALM (Qu et al., 2004). If two different fluorescent molecules are located within the same diffraction-limited volume, then they can be localized and separated based on their different spectral properties. However, if the two fluorescent molecules possess

identical spectral characteristics, then they cannot be separated based on spectral properties.

The NALMS technique solves this pervasive problem and achieves nanometer localization of multiple identical fluorescent molecules situated within the same diffraction-limited PSF. How does NALMS work to make this possible?

The NALMS technique utilizes centroid localization together with photobleaching of individual fluorescent molecules to achieve superresolution. Remember that photobleaching is the irreversible switching from a fluorescent ON state to a nonfluorescent OFF state. The fluorescent molecules undergo many cycles of excitation and emission, and these cycles end with the rapid photobleaching of individual molecules. The number of fluorescent molecules in a single diffraction-limited volume can be determined by observing the number of jumps in fluorescent intensity; each jump or step represents the photobleaching of a single fluorescent molecule, and the number of jumps or steps is related to the number of fluorescent molecules within the diffraction-limited PSF.

The success of the NALMS technique depends on the ability to detect the photobleaching of individual fluorescent molecules. Other requirements are that the fluorescent molecules do not change their location over the time needed to acquire the data, and that there are a few fluorescent molecules within the diffraction-limited PSF; this criterion is only valid for a very dilute density of fluorescent labels.

Generally, however, that is not the case with fluorescent-labeled biological specimens. Superresolution microscopy requires the specimen to be densely labeled. It is important to note that there are many hundreds of adjacent fluorescent molecules within the volume of the PSF of a typical fluorescent-labeled biological specimen. In summary, NALMS is optimal for a small number of brightly emitting fluorescent molecules. The technique works by photobleaching individual fluorescent molecules one by one, with photobleaching occurring in a stochastical or random order.

Qu et al. demonstrated the capability of the NALMS technique on two identical fluorescent molecules (Cy3) separated by 8 nm. It gave a precision of 2.5 nm. The NALMS technique was also applied to DNA mapping when a resolution less than 30 nm was achieved (Qu et al., 2004).

A second paper in 2004 demonstrated the utility of photobleaching, or the irreversible transfer of a fluorescent molecule from the ON state to the OFF state, and advanced the quest for superresolution microscopy. This paper titled "Single-molecule high-resolution imaging with photobleaching" (a.k.a. SHRImP) independently highlighted the role photobleaching played in the leadup to the invention of PALM (Gordon et al., 2004). The authors' motivation was to develop a technique using a single-color fluorescent dye to measure two point separation distances of 10 nm and less. A single fluorescent dye was found to be advantageous over alternative techniques that used two colors (two different fluorescent molecules) because there were no registration problems with a single fluorescent molecule.

Gordon et al.'s technique, designed to measure distances between 10 and 100 nm, located the positions of two identical fluorescent molecules and measured

their separation with a precision of 5 nm. Quantal photobleaching of single fluorescent molecules led to this achievement (Gordon et al., 2004).

Since Gordon et al. were interested in measuring the distances of DNA subunits they validated their technique by determining the distance between single fluorescent molecules separated by 10–20 nm. Separation was brought about by attaching the fluorescent molecules to the ends of double-stranded DNA molecules stabilized on a surface (Gordon et al., 2004). This advance also highlighted the use of test objects, in this case biological objects where the physical distances are independently known, to validate the accuracy of superresolution images.

15.3.2 The Invention and Development of Photoactivated Localization Microscopy

On August 10, 2006 *Science* published an online paper titled "Imaging intracellular fluorescent proteins at nanometer resolution" (Betzig et al., 2006). The publication stated that Eric Betzig and Harald F. Hess contributed equally to this work. The late Michael W. Davidson is appropriately and graciously credited with pointing out to Betzig and Hess the wide variety of photoactivatable fluorescent proteins.

Betzig et al. (2006) represented two milestones: it was the first PALM publication, and the first publication to use correlative transmission electron microscopy (TEM) and PALM to validate the images. The latter is an extremely important validation procedure that should be used more often to determine whether the images obtained by superresolution microscopy are real and are physiologically meaningful (for live cell imaging). The use of alternative imaging techniques on the same specimen is an important validation technique (Masters, 2009a).

In a nutshell, the PALM technique and its capability can be described as a method to isolate individual photoactivatable fluorescent molecules at very high label (or labeling) densities on the specimen (on the order of $10^5/\mu m^2$) through sequential photoactivation and photobleaching of many sparse (less dense) subsets of photoactivatable fluorescent proteins (PA-FPs) of the specimen. The authors demonstrated the utility of PALM with a total internal reflection fluorescence (TIRF) microscope to image proteins located in lysosomes and mitochondria, and in studies with fixed whole cells they imaged proteins bound to the plasma membrane. The PALM technique can separate fluorescent molecules located a few nanometers from each other.

The PALM technique works by following a sequence of steps—photoactivation, locating individual fluorescent molecules, and photobleaching—which is repeated for many cycles for the acquisition of 10^4 to more than 10^5 image frames. The number of image frames that need to be acquired depends on cellular expression and the resulting spatial distribution of PA-FPs. With an average frame rate and a

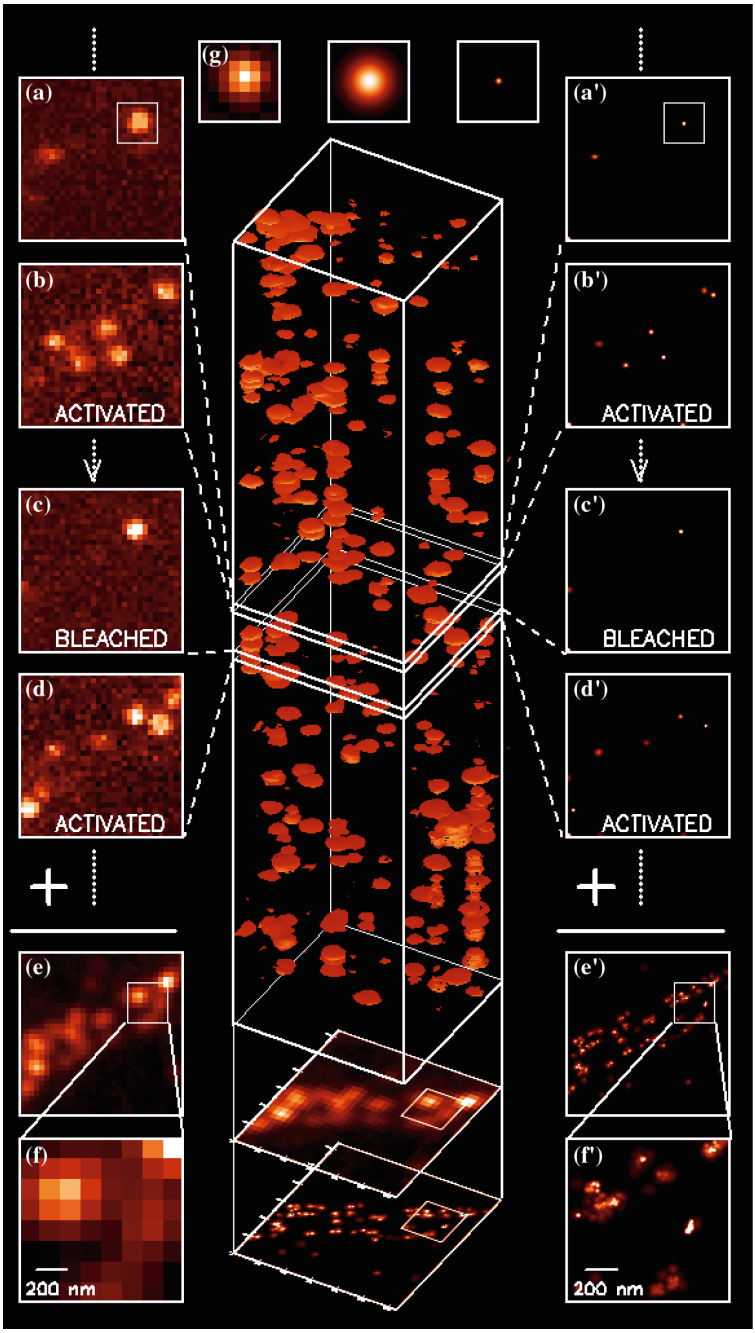
frame acquisition time of half a second to one second between 2 and 12 h are necessary to obtain a complete stack of images. This resultant image stack is then reconstructed to the final superresolution image, which contains 10^6 spatially localized molecules. Figure 15.1 shows the method and typical data subsets. Each fluorophore in PALM and FPALM is activated only once (Betzig et al., 2006).

Betzig et al. (2006) stands out as a result of its unique capability to effect separation. Previous studies have managed to separate fluorescent molecules located within a single focal area that is diffraction limited. However, in these cases the number of fluorescent molecules is very small (in the range of two to five molecules). The seminal achievement of Betzig and his colleagues was to effect separation of a very high labeling density of fluorescent molecules that reside within the diffraction-limited focal area. The fundamental key idea behind PALM is to use photoactivatable proteins and serially convert the original very dense, highly labeled set of fluorescent molecules into a series of sparse subsets, each composed of separated fluorescent molecules whose PSFs do not overlap. The selection of which individual fluorescent molecules comprise a specific subset is an entirely random or stochastic process.

The PALM instrument described in Betzig et al. (2006) has a major advantage over previously described STED superresolution microscopes: simplicity. The PALM/TIRF (total internal reflection fluorescence) superresolution microscope requires lasers, filters, and software for instrument control, image acquisition, and analysis. The light detector is an electron-multiplying charge-coupled device (EMCCD) camera with single-photon sensitivity. The authors propose their instrument is well suited for in vitro specimens and for the imaging of fixed cells (Betzig et al., 2006).

As is the case with all superresolution microscopy techniques there are trade-offs to be considered. If the number of fluorescent molecules localized is decreased during formation of the final superresolution image, then there may be improved localization and sharper images. However, in this case there will be reduced information on the distribution of labeled proteins in the specimen. The photophysics and spectroscopic characteristics of photoactivatable probes are critical to PALM effectiveness. Those PA-FPs with longer lifetimes of photobleaching will permit the detection of more photons, but data acquisition times are also increased. These considerations point to the critical importance of designing new PA-FPs with photophysical properties that enhance the effectiveness of the PALM technique. Again, probe development is crucial to the advancement and wide utility of the PALM technique. Desirable features are the development of new PA-FPs with enhanced excitation cross sections, quantum efficiency, brightness, photostability, and minimal blinking. A bright fluorescent molecule is associated with a large molar absorption coefficient and a large quantum yield. Brighter fluorescent molecules result in increased contrast compared with an autofluorescence background. Another important factor is the contrast between the ON state and the OFF state of PA-FPs.

In the years since the invention of PALM Shroff and his colleagues further developed the capabilities of PALM and demonstrated its utility for live cell


(a)
(b)
ACTIVATED
(c)
BLEACHED
(d)
ACTIVATED
(e)
(f)
200 nm
(g)
(a')
(b')
ACTIVATED
(c')
BLEACHED
(d')
ACTIVATED
(e')
(f')
200 nm

◂**Fig. 15.1** The principle behind PALM. A sparse subset of PA-FP molecules attached to proteins of interest and then fixed within a cell are activated (**a** and **b**) with a brief laser pulse at $\lambda_{act} =$ 405 nm and then imaged at $\lambda_{exc} = 561$ nm until most are bleached (**c**). This process is repeated many times (**c** and **d**) until the population of inactivated, unbleached molecules is depleted. Summing the molecular images across all frames results in a diffraction-limited image (**e** and **f**). However, if the location of each molecule is first determined by fitting the expected molecular image given by the PSF of the microscope (**g**, center) to the actual molecular image (**g**, left), then the molecule can be plotted (**g**, right) as a Gaussian that has a standard deviation equal to the uncertainty $\sigma_{x,y}$ in the fitted position. Repeating with all molecules across all frames (**a′** through **d′**) and summing the results yields a superresolution image (**e′** and **f′**) in which resolution is dictated by the uncertainties $\sigma_{x,y}$ and by the density of localized molecules. Scale: 1×1 μm in (**f**) and (**f′**), 4×4 μm elsewhere. From Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science, **313**, 1642–1645.

superresolution imaging. In the following two publications the authors significantly decreased the total time required to form the final composite superresolution image; they applied the PALM technique to live cell imaging and discussed the problems associated with the same; and they highlighted the need for very careful control studies to validate and correctly interpret such live cell superresolution studies (Shroff et al., 2007; Shroff et al., 2008).

The first publication titled "Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes" was a response to the following question (Shroff et al., 2007). What are cellular adhesion complexes (ACs)? Adhesion complexes are defined as the points where the cytoskeleton is attached to the surface to which cells migrate. They are composed of more than 90 different proteins. These adhesion complexes contain different proteins previously imaged as colocalized in the standard fluorescent microscope. First, the authors needed to obtain superresolution images of pairs of proteins to determine the actual spatial arrangement of proteins in the adhesion complex. Second, the authors wanted a technique to label the proteins in adhesion complexes that did not perturb the spatial arrangement of proteins. And, finally, the authors wanted to achieve superresolution microscopic imaging at high resolution (10–30 nm) and short acquisition time (5–30 min).

The main biological question that Shroff et al. wished to investigate concerned the structural interaction and spatial arrangement of pairs of two different proteins fused to different PA-FPs, each of which has a distinct spectrum? This publication achieved advances in both specimen preparation and PALM instrumentation. The specimen comprised whole fixed cells. In previous works specimens were labeled with fluorescent antibodies exogenously introduced to the specimen. However, this new investigation used two-color PALM with PA-FPs endogenously expressed in the cells.

First, some comments are needed on specimen preparation. There is little point in obtaining a superresolution image of the specimen if the techniques required for specimen preparation significantly alter the specimen's structure. Special precautions were required to have a specimen preparation that minimized alteration to molecules and their complexes in the specimen. Shroff et al. used a mild fixation

technique for the specimen, after which the specimen was returned to its physiological media. Their method of specimen preparation was devoid of detergents that can alter the specimen's structure, there was no need for oxygen-depleting agents, and no chemical treatments were required to alter the photophysics of the molecules (these are requirements for the STORM technique). The authors' specimen preparation had other advantages. Genetically expressed PA-FPs managed to attach to their targets. This is a great advantage over the alternative use of exogenous labels associated with problems of specificity and background fluorescence. The authors caution the readers to beware and prevent overexpression of the target protein.

Second, Shroff et al. corrected the ever-present persistent problem in all superresolution microscopic techniques of specimen drift due to thermal fluctuation and/or vibrations. They cleverly corrected this by tracking the movement of gold fiducial beads 40 and 100 nm in size. Photobleaching and phototoxicity of the specimen were minimized by ensuring photoactivation was continuously applied to the specimen with intensities in the range of 0.5–2.0 kW/cm^2. Their instrument's single-molecule frame times of 2–50 ms resulted in the desired short acquisition time of less than 30 min (Shroff et al., 2007). A logical question the astute reader might ask is: What about live cell superresolution PALM?

The following year Shroff et al. introduced the next step in their quest for enhanced resolution and contrast in live cells titled "Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics" (Shroff et al., 2008). First, I summarize their achievement and then discuss the problems and the authors' innovative solutions. Remarkably, the authors were able to study nanoscale dynamical processes in separate adhesion complexes of living cells under physiological conditions for a period of 25 min. These live cell PALM imaging studies showed the entry and exit of single paxillin molecules during the evolution of a separate AC. Live cell PALM imaging showed the molecular formation of a separate AC during individual phases of initiation, maturation, and dissolution of cellular events (Shroff et al., 2008).

To achieve their desired aim of live cell PALM Shroff et al. had to circumvent several difficult problems. First, I describe the problems with live cell PALM and, then, I discuss the authors' solutions. Finally, I describe the extremely important control studies performed by the authors. Again, I repeat a general dictum applicable to the field of superresolution microscopy: the validity of superresolution images and therefore their interpretation and analysis is only as good as the quality and design of the control experiments.

What are the difficulties and problems in superresolution live cell imaging that must be solved? Shroff et al. state five critical concerns: (1) The image sampling interval should be less than one-half of the required resolution according to the Nyquist criterion. For example, if the goal is an N-fold increase in spatial resolution in D dimensions, then the requirement is to acquire N^D-fold additional pixels. (2) So that the imaging speed or the SNR do not change the rate at which photons are detected per second must be increased by N^D. This results in the specimen receiving an N^D-fold larger dose of excitation radiation during the acquisition of each image.

(3) When the imaging technique is based on point-scanning (e.g., using a confocal microscope) the required intensity increase is even larger (i.e., a factor of N^{2D}). (4) Since spatial resolution is enhanced in the process of live cell imaging there is a requirement for faster frame rates. This is required to mitigate blurring due to the movement of subcellular components. (5) The last concern is crucial to live cell imaging. The investigation must validate that these requirements do not alter the physiological function of the cells studied. If this is not the case, then the interpretations and conclusions of live cell imaging studies are meaningless (Shroff et al., 2008).

Shroff et al. were able to provide solutions to these concerns. In their paper they described the combination of factors that eventually led to their achievement. First, they used an imaging speed that was compatible with the Nyquist criterion. Second, they developed a means of robust labeling of the target protein using a unique probe (described below). Third, the authors selected cells for their investigations that show minimal photo-induced perturbations (Shroff et al., 2008).

Probe development is key to superresolution microscope advances. In fact, it can be considered the limiting factor to advances in the field of superresolution optical microscopy. This is clearly demonstrated in the live cell PALM paper of Shroff et al. (2008). The authors used the photoactivatable probe EosFP because it has sufficient photostability for it to be located with a precision of less than 20 nm and because it is not too stable, so new subsets can be acquired in a timely manner. Furthermore, the probe EosFP features a high-contrast ratio between its fluorescent ON state and its nonfluorescent OFF state. This feature is critical to precise localization of single molecules. In their experiments the authors used live NIH 3T3 cells expressing the tandem–dimer form of EosFP fused to the AC protein paxillin (tdEosFP–paxillin). The activation intensity of 405-nm light is 0.05 kW/cm^2 to activate single tdEosFP molecules and the activation intensity of 561-nm light is 1 kW/cm^2 to excite fluorescence and eventually bleach the activated molecules (Shroff et al., 2008). It is important to note that activation and excitation intensities vary with the specific type of fluorescent molecule used in superresolution microscope investigation.

The contrast ratio between the ON state and the OFF state is defined as the fluorescence intensity ratio between these two states. Even in the OFF state photons can still be emitted from the fluorescent molecule. With the required high labeling density with fluorescent molecules in order to achieve superresolution imaging that correctly shows the structures of interest in the specimen the Nyquist criterion is a critical consideration. Another important parameter to be considered in the selection of a fluorescent molecule for superresolution microscopy is its ON/OFF contrast ratio (Thompson et al., 2002). A fluorescent molecule with a high ON/OFF contrast ratio (e.g., a contrast ratio in the range 1000–2000) will preclude the severe problem of background fluorescence.

Shroff et al.'s paper on live cell PALM is an exemplar of the control studies required to validate resulting images and their physiological meaning. Since photoeffects on cells are important concerns in live cell PALM imaging the authors used cell lines from Chinese hamster ovary (CHO) and NIH 3T3 fibroblasts found

to be photon tolerant. Exposures of 1 kW/cm^2 for a duration of 30 min showed no alterations to the pertinent cell processes the authors were investigating. To further minimize photoeffects on cells during live cell PALM imaging the authors combined the PALM instrument with a total internal reflection fluorescence (TIRF) microscope. ACs are only found at the interface between the substrate and the cells. Combining the PALM technique with a TIRF microscope allowed the illumination to be restricted to the region of the ACs (in the region of 50–100 nm) from the substrate. The TIRF microscopy technique reduces background fluorescence because it limits the focal depth of the excitation of fluorescent molecules to a thin evanescent field region. This spatially restricted illumination has the important effect of significantly lowering the cell's exposure to the photoeffects of activation and excitation light during the PALM procedure (Shroff et al., 2008).

Shroff et al. demonstrated that these solutions actually resulted in the nonperturbative conditions of the cells being imaged. This is critical to validating the superresolution technique, and hence validating the interpretation and conclusions that are based on superresolution images. The authors completed the validation process by acquiring differential contrast images of cells at the rate of one every 2 s during the PALM procedure. The internal components of the cells could be imaged in real time during the PALM procedure and the results could be compared with a set of similar cells that were not imaged with the PALM procedure (Shroff et al., 2008).

Shroff et al. (2008) contains an important lesson for all investigators who use superresolution microscopy. The lesson is live cell superresolution microscopy requires careful control studies to validate and correctly interpret superresolution images. The use of correlative microscopy, in which the specimen is imaged with superresolution light microscopy and electron microscopy, was demonstrated in the first PALM publication (Betzig et al., 2006).

15.4 Interferometric Photoactivated Localization Microscopy

A persistent goal of developers of light microscopy is the invention of microscopes that cover a wide range of scales of spatial and temporal resolution. An important example and a significant contribution to this goal is a superresolution microscopic technique with a resolution between that of electron tomography and light microscopy.

I now discuss Shtengel et al.'s paper titled "Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure" (Shtengel et al., 2009). This publication cleverly combines the two optical techniques of interferometry and PALM to come up with a new superresolution light microscope that achieves three-dimensional imaging of the ultrastructure of cells. The authors called

their new invention interferometric PALM (iPALM). They were able to simultaneously combine the technique of photoactivated location microscopy (PALM) with the technique of single-photon multiphase interferometry.

The limitations and problems besetting then-existing instruments motivated the researchers to pursue the invention and development of a new type of microscope. The existing imaging techniques of light microscopy and electron microscopy were limited by two factors: spatial resolution and molecular specificity. An understanding of the limits of previous superresolution microscopes motivated investigators to invent a new microscope to circumvent the limits of previous instruments. The discussion of the previous publication showed that combining a PALM instrument with a TIRF microscope can achieve superresolution images of the cellular adhesion complexes that connect cells to the substrate (Shroff et al., 2008). The TIRF microscope limits the focal depth from the substrate to between 50 and 100 nm.

Shtengel et al. set out to develop a new microscope capable of bridging the previous gap in the resolution of the light microscope and the resolution of electron tomography (Shtengel et al., 2009) because previous imaging techniques were limited in resolution and molecular specificity. While the technique of electron microscopy together with immunostaining offers the appropriate nanoscale spatial resolution to study subcellular nanostructures the degree of molecular specificity is impaired by the ability of large antibodies to work their way into complex structures and by their diminished specificity due to cross-reactivity. The authors noted that while progress is being made (as of 2009), the axial resolution of less than 100 nm is still not sufficient to investigate subcellular nanostructures (Shtengel et al., 2009). Their invention of iPALM is a solution to this problem.

How does iPALM work? The new iPALM consists in combining a lateral photoactivated localization microscope (PALM) with a single-photon multiphase interferometer. Since I have already discussed the PALM technique in previous sections I will limit my discussion to the single-photon multiphase interferometer. For those not familiar with interferometry and its various techniques the works by Bhushan, Wyant, and Koliopoulos (1985), Delaunay (1953), Hariharan (2003), and Malacara (2007) are all worth reading. The authors cogently point out that applying interferometry to the microscopy of biological specimens faces the problems of sharply defined surfaces with steep changes in refractive index and poor molecular specificity (Shtengel et al., 2009).

Shtengel et al. faced the daunting problem of constructing an interferometer-based microscope capable of coping with interference from two coherent light beams. The critical advantage of enormous specificity that accrues using fluorescent molecules and the good understanding of bleaching and the interaction between fluorescent molecules and light enabled the authors to solve a number of difficulties on their pathway to inventing iPALM.

The first step involved a literature search that provided insights from previous publications (Braun and Fromhertz, 1998; Dogan et al., 2008). These publications demonstrated an important principle applied in fluorescence microscopy

(Braun and Fromhertz, 1998) and in spectral self-interference fluorescence microscopy (Dogan et al., 2008).

The basic idea behind Shtengel et al.'s invention of the iPALM superresolution optical microscope is single-photon interferometric fluorescence imaging with multiphase simultaneous detection (illustrated in Fig. 15.2).

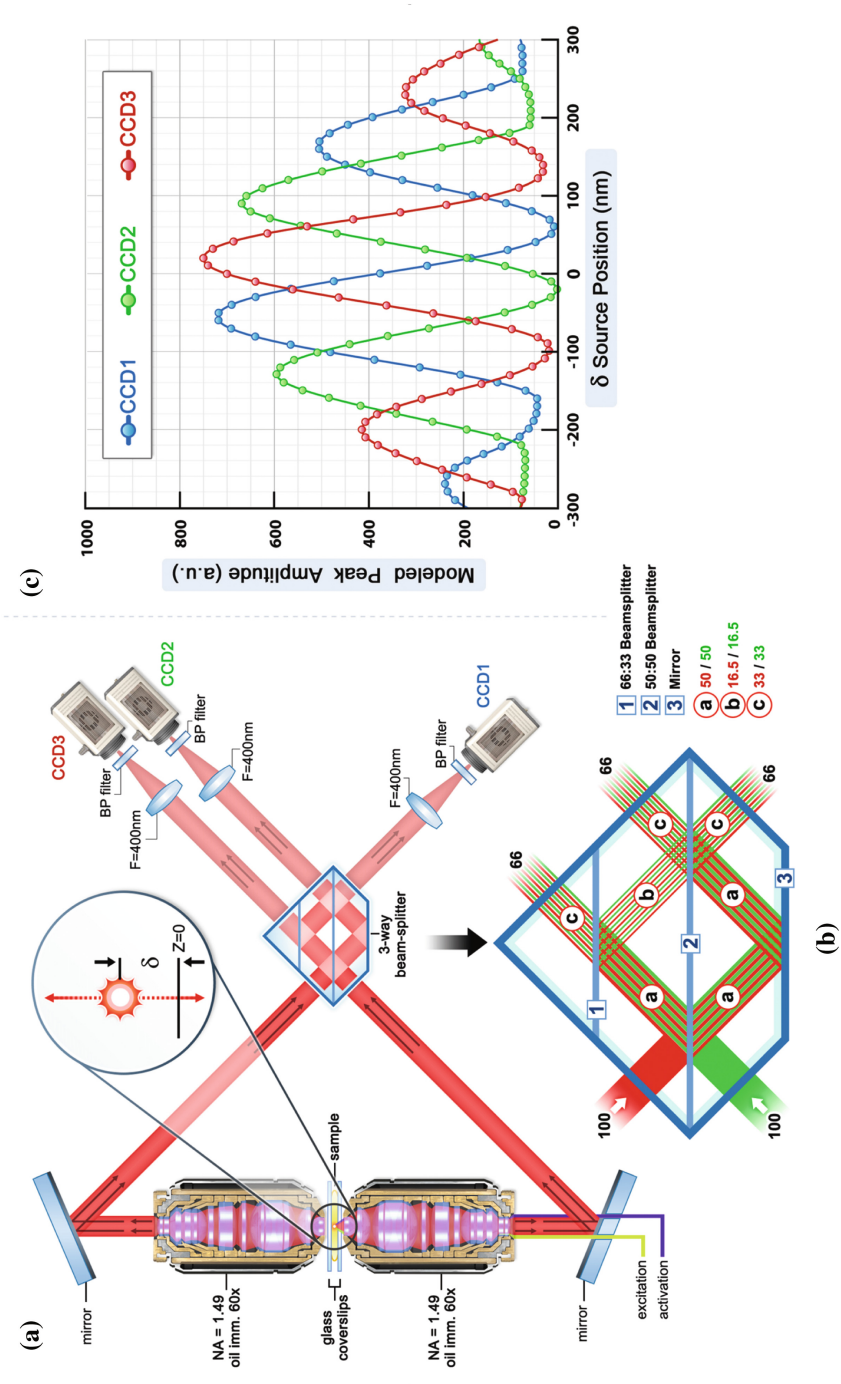
The concept is that a photon emitted from a fluorescent molecule in the process of de-excitation from the excited state to the ground state can simultaneously travel along two different optical paths (upward and downward in Fig. 15.2) and be recombined in a unique three-way beamsplitter; the photon self-interferes. The spatial axial location of the fluorescent molecule that emitted the photon is the parameter that forms the difference in optical path lengths and this yields the phase difference between the two beams. More specifically, the axial position where a given fluorescent molecule is located and at the same time is the source of the photon can be determined by measuring the relative amplitudes of images of the source with three separate electron-multiplying charge-coupled cameras (EMCCDs; iXon DU-897, Andor Technology).

Shtengel et al. used the concept of spectral self-interference fluorescence microscopy and designed a microscope with the inherent capability to have this self-interference occur from a range of lateral positions. Interference images are detected on three CCD cameras in parallel. A critical feature of the authors' design and invention of iPALM is that every single photon undergoes simultaneous multiphase detection (Shtengel et al., 2009). A multiphase interferometer yields the axial position of the fluorescent molecule, and simultaneously the PALM technique yields the lateral position in the focal plane of the specimen.

The results were a major achievement: a 10× enhancement in axial resolution and a 100× enhancement in photon efficiency relative to defocus superresolution microscopy such as 3D STORM (Huang et al., 2008). The iPALM technique achieved three-dimensional localization that was sub-20 nm (Shtengel et al., 2009). The authors used fixed cells that expressed photoactivatable or photoswitchable FP fusion proteins. The authors demonstrated a three-dimensional spatial resolution less than 20 nm (nearly isotropic imaging) using photoactivatable FPs. The specimens imaged include measurement of the 25-nm diameter of the microtubule and the arrangement of integrin receptors in adhesion complexes and in the endoplasmic reticulum (Shtengel et al., 2009).

Enhancing the range of imaging or the imaging depth within a specimen in the axial direction is what is needed. The first publication on iPALM was in 2009. At that time the range in imaging depth in the axial direction over which iPALM could be used was 250 nm. The cause of this limit in axial imaging depth stems from the nature of the calibration curve, which is periodic. This periodic structure makes it difficult to separate adjacent interference fringes and limits the imaging depth to 250 nm.

The solution came from a publication that showed how mirrors could be used to add a phase curvature in the pupil plane. This phase shift caused the PSF to take on an elliptical form (Brown et al., 2011). An elliptical PSF and astigmatic defocusing enhanced the imaging depth for iPALM to 750 nm (Shtengel et al., 2014).


(a)
mirror
NA = 1.49
oil imm. 60x
glass
coverslips
sample
NA = 1.49
oil imm. 60x
mirror
excitation
activation
δ
Z=0
3-way
beam-splitter
CCD3
CCD2
CCD1
BP filter
F=400nm
(b)
1 66:33 Beamsplitter
2 50:50 Beamsplitter
3 Mirror
a 50 / 50
b 16.5 / 16.5
c 33 / 33
66
100
(c)
Modeled Peak Amplitude (a.u.)
δ Source Position (nm)
CCD1
CCD2
CCD3

◂**Fig. 15.2** Schematics and operating principle of the multiphase interferometric microscope illustrating how the *z*-position is resolved. **a** and **b** Schematic of the single-photon multiphase fluorescence interferometer. A point source with *z*-position δ emits a single photon both upward and downward. These two beams (color-coded as *red* and *green* in **b**) interfere in a special three-way beamsplitter. **c** The self-interfered photon propagates to the three color-coded CCD cameras with amplitudes that oscillate 120° out of phase (as indicated). From Shtengel, G., Galbraith, J. A., Galbraith, C. G., Lippincott-Schwartz, J., Gillette, J. M., Manley, S., Sougrat, R., Waterman, C. M., Kanchanawong, P., Davidson, M. W., Fetter, R. D., and Hess, H. F. (2009). Interferometric fluorescent super-resolution microscopy resolves three-dimensional cellular ultrastructure. Proceedings of the National Academy of Sciences of the United States of America, **106**, 3125–3130. In journal Figure # 1

Shtengel et al. (2014) is a review article highlighting two other advances in PALM and iPALM. First, it demonstrated that fluorescent-labeling strategies are a work in progress and are developing fluorescent molecules with higher photostability, higher brightness, smaller size, and higher specificity. Another requirement for a subset of fluorescent molecules is that they are compatible with live cell imaging. Second, it demonstrated the growing importance of correlative microscopy to validate superresolution images, their analysis, and their interpretation. The first PALM publications validated superresolution images using electron microscopy. Since PALM and iPALM are still undergoing development the use of electron microscopy continues to play an important role (Shtengel et al., 2014).

I posit that correlative microscopy should be mandatory for any new superresolution microscopy technique, specimen labeling, and specimen preparation (Masters, 2009a).

15.5 Fluorescence Photoactivated Localization Microscopy

Three independent groups were responsible for the independent inventions of PALM (photoactivated localization microscopy), FPALM (fluorescence photoactivated localization microscopy), and STORM (stochastic optical reconstruction microscopy) in 2006. On June 12, 2006 Hess and his coworkers submitted their manuscript to the *Biophysical Journal*. Titled "Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy" it was accepted on August 28, 2006 and published in December (Hess et al., 2006).

This publication is noteworthy for its emphasis on validation and controls, as well as a detailed description of the theory, instrumentation, and computer simulations. To demonstrate the capability of the FPALM technique the investigators used photoactivatable green fluorescent protein (PA-GFP) molecules deposited by the evaporation of dilute solutions on a glass coverslip. Hess et al. validated the FPALM technique by comparing FPALM images of the molecule PA-GFP deposited on a terraced sapphire crystal with images from atomic force microscopy (Hess et al., 2006).

Furthermore, Hess et al. (2006) makes a couple of remarks relating to the social process of science and its foundational ethics (Masters, 2016). First, in the acknowledgment section of their publication the authors thanked Dr. George

Paterson for giving them PA-GPF samples as well as providing them with unpublished data. This is an example of how generosity and trust should be used to promote the advancement of science. Second, the authors added a note at the end of their publication stating that a publication by Betzig et al. on a related superresolution technique was being published in *Science* at the same time as Hess et al., (2006). This is an example of how citations should be made.

What is fluorescence photoactivated localization microscopy (FPALM) and how does it work to achieve superresolution optical microscopy? FPALM is a new type of superresolution optical microscopy. Photoactivatable molecules in an inactive or nonfluorescent state are activated or transferred to a state that can emit fluorescence by illumination with a 405-nm laser. The intensity of the activation laser controls the rate of photoactivation. Single molecules activated in this way are then excited using a laser of lower frequency and fluorescence is detected using a single-molecule detection technique such as a CCD imaging camera. Next, after a random number of emitted photons are detected, the molecules that emitted the fluorescence are removed from the detected field of view by either of two processes: (1) they are reversibly inactivated, or (2) they are irreversibly photobleached. These sequential steps are repeated many times allowing the precise location of each fluorescent molecule to be determined. The final superresolution two-dimensional image is formed by superposition of the locations of all fluorescent molecules. The key to success of the FPALM technique is that only an extremely small subset of all photoactivable molecules are stochastically (randomly) activated at a given instant, and therefore fluorescence is detected from only an extremely small subset of photoactivatable molecules. The FPALM technique is significant in that photoactivatable molecules can be located with enhanced precision at a low number of detected photons per fluorescent molecule—in this case only 100 photons from each fluorescent molecule (Hess et al., 2006).

Hess et al. pointed out that FPALM is unique in that it is possible to control, increase, or decrease the number of active fluorescent molecules by modifying the density-labeling rates of the photoactivation process and the photobleaching process. FPALM differs from other single-molecule imaging methods limited to a labeling density of one fluorescent molecule per square micrometer by having a labeling density that can be orders of magnitude larger (Hess et al., 2006).

Figure 15.3 illustrates the experimental geometry of an FPALM. This figure and the detailed figure legend specify the details of the FPALM technique and the instrumentation that comprises the microscope. Two lasers are used in an FPALM: a photoactivation laser consisting of a 405-nm diode laser and an excitation laser consisting of an Ar^+ (argon ion) laser.

The FPALM technique is illustrated in Fig. 15.4, which shows the sequence of events that take place to form the final superresolution image. The FPALM technique has significant advantages over the RESOLFT techniques that include STED (Hess et al., 2006). First, FPALM is a widefield superresolution optical microscopy technique that does not require scanning. This contrasts with the RESOLFT techniques that either scan the specimen or the light beam. Second, FPALM illuminates the specimen with low-intensity light (less than 100 W/cm^2). This low-intensity illumination reduces light-induced photobleaching and phototoxicity of cells. Third,

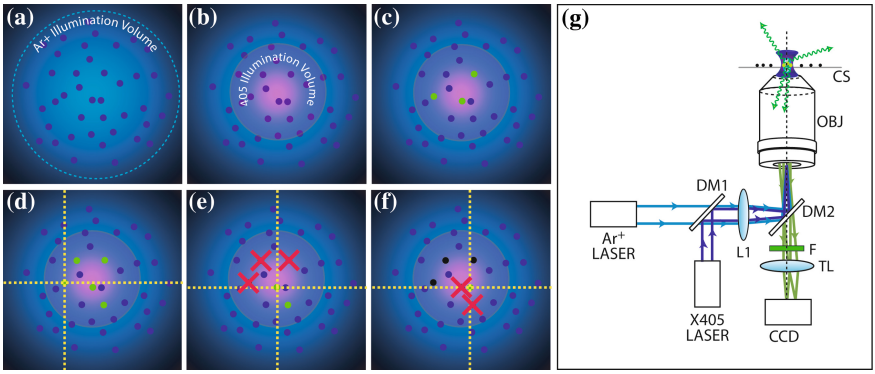


Fig. 15.3 Fluorescence photoactivated localization microscopy (FPALM). An area containing photoactivatable molecules (here PA-GFPs) is illuminated simultaneously with two frequencies of light, one for readout (here an Ar^+ ion laser, its spatial illumination profile shown in (**a**), and a second for activation (here a 405-nm diode laser, its profile superimposed in (**b**). Within the region illuminated by the activation beam inactive PA-GFPs (*small dark blue circles*) are activated (**c**), (*small green circles*), and then localized (**d**). After some time the active PA-GFPs **e** photobleach (*red* $\times s$) and **f** become irreversibly dark (*black circles*). Additional molecules are then activated, localized, and bleached until a sufficient number of molecules have been analyzed to construct an image. **g** The experimental geometry shows the 405-nm activation laser (X405), which is reflected by a dichroic mirror (DM1) to make it collinear with the Ar^+ readout laser. A lens (L1) in the back port of an inverted fluorescence microscope is used to focus the lasers, which are reflected upward by a second dichroic mirror (DM2), onto the back aperture of the objective lens (OBJ). The sample, supported by a coverslip (CS), emits fluorescence that is collected by the objective, transmitted through DM2, filtered (F), and focused by the tube lens (TL) to form an image in a camera (CCD). From Hess, S. T., Girirajan, T. P. K., and Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophysical Journal, **91**, 4258–4272

the FPALM technique is devoid of the complex optics needed to carefully align the two beams in STED microscopy. The less complex optical system of the FPALM is less expensive and relatively easy to adjust, operate, and maintain compared with the STED microscope. On the other hand, the STED microscope has higher axial resolution compared with the early versions of the FPALM (Hess, et al., 2006).

One year after the initial publication on the invention of FPALM a subsequent publication demonstrated the utility of FPALM to work with both fixed and live cells and to investigate significant biological questions. In 2007 Hess et al. published a paper titled “Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories” (Hess et al., 2007). The authors thanked Dr. George Paterson for providing PA-GFP construct and purified protein. This publication is a significant validation of the use of the FPALM method on fixed and on live cells to solve important biological problems. FPALM widefield superresolution optical microscopy permitted the investigators to use imaging to distinguish between various raft theories (Hess et al., 2007). Furthermore, their publication uses electron microscopy to validate the results of imaging with the FPALM technique. Review of a decade of publications on

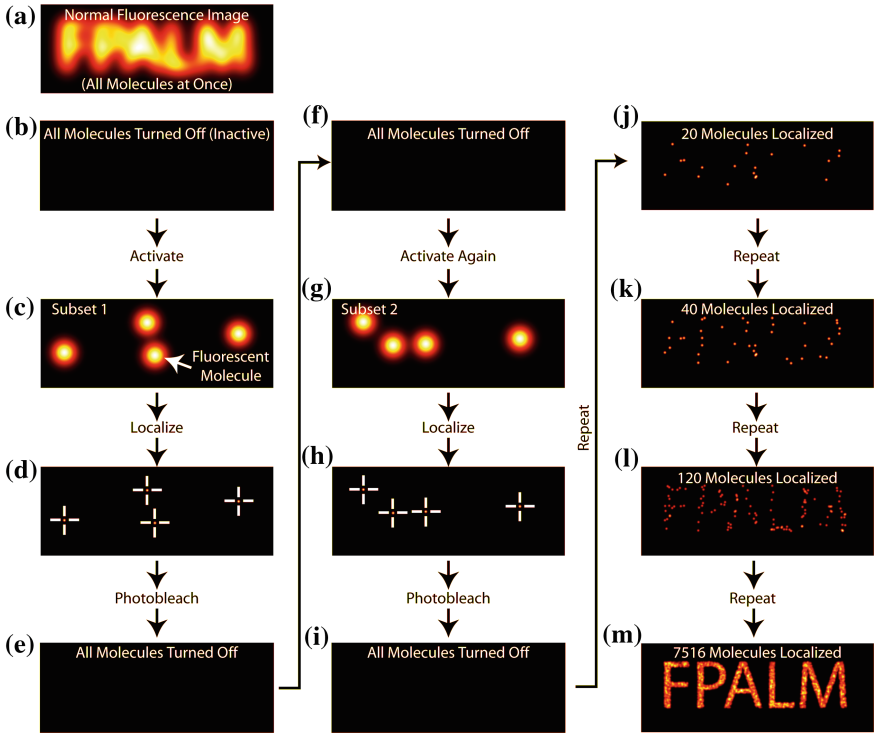


Fig. 15.4 Concept of FPALM. **a** In normal fluorescence microscopy large numbers of fluorescent molecules are visible at once, but diffraction blurs objects smaller than 200–250 nm obscuring fine details. In FPALM light is used to limit the number of visible (fluorescent) molecules. In contrast to normal fluorescent molecules FPALM uses photoactivatable fluorescent probes that are initially nonfluorescent (inactive). **b** Even under normal illumination inactive molecules are invisible. **c** A low-intensity 405-nm activation laser converts a small subset of inactive molecules into active ones (*large spots*). **d** Active molecules are imaged and localized to precisely determine their positions (*small spots*). **e** Photobleaching turns active molecules permanently off. **f–i** Starting with the remaining inactive molecules the process of activation, imaging, localization, and photobleaching is repeated, this time yielding the coordinates of a new subset of molecules. The process is repeated until enough molecules have been localized to reveal the structure of the sample. The plotted positions of localized molecules form the FPALM image. **j–m** Simulated FPALM images of a sample with increasing numbers of molecules illustrate how data are built up iteratively. FPALM can be used to image a variety of samples in two or three dimensions. With permission of author, Samuel T. Hess

superresolution microscopy shows a paucity of validation techniques comparing disparate types of microscopes (Masters, 2009a).

In 2008 details of a new version of FPALM called Biplane-FPALM (BP-FPALM) were published in "Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples" (Juette et al., 2008). I frame my discussion of this innovative new type of superresolution optical microscopy in four parts. First, what is the problem to be solved? Second, how does BP-FPALM work and provide a solution to the problems and limitations of previous superresolution optical microscopes? Third, what three-dimensional resolution is achievable with

BP-FPALM? Fourth, what are the advantages and benefits of using BP-FPALM over alternative types of superresolution microscopes?

The problem the investigators tackled was how to construct a nonscanning superresolution optical microscope based on the FPALM technique capable of achieving three-dimensional spatial resolution of thick specimens less than 100 nm in all directions? To solve this problem and reach their stated goal the authors decided to combine the technique of FPALM (Hess et al., 2006) with a previously published independently invented technique called multifocal plane microscopy (a.k.a. biplane detection microscopy) that was successfully used to achieve three-dimensional images in single-particle tracking (Prabhat et al., 2004; Ram et al., 2012).

How does BP-FPALM work and provide a solution to the problems and limitations of previous superresolution optical microscopes? The optical setup for BP-FPALM is similar to that for FPALM with only a slight modification of the optics in the detection pathway in front of the electron-multiplying charge-coupled device (EMCCD) imaging detector. The light detected from photoactivated, excited, fluorescent molecules is passed through a 50:50 beamsplitter. This beamsplitter divides light into two separate beams, each with different optical paths. One beam is transmitted by the beamsplitter and focused on one region of the EMCCD, and the second beam is reflected by the beamsplitter to a mirror and then to a second region of the EMCCD. Between the mirror and the EMCCD there is a focal plane for the reflected beam. Both images are simultaneously formed on two distinct regions of the EMCCD; each image of detected photons generated in the FPALM technique corresponds to photons traveling along different path lengths. These two images formed on different regions of the EMCCD correspond to object planes that are located either closer to or more distant from the objective by 350 nm with respect to the original object plane. The signals from the two distinct regions of the EMCCD are combined to form a three-dimensional data set. The PSF is measured from a small fluorescent bead. The shape of the PSF is determined as a function of the varying positions of the fluorescent bead relative to the location of the objective (Juette et al., 2008).

What three-dimensional resolution can be achieved with BP-FPALM? The authors achieved a lateral spatial resolution of 30 nm and an axial resolution of 75 nm over a range of depths (several micrometers) within the specimen (Juette et al., 2008).

What are the advantages and benefits of using BP-FPALM over alternative types of superresolution microscopes? First, the BP-FPALM technique that combines FPALM with a simple modification of the optical layout can be implemented on all existing superresolution optical microscopes that use PALM, FPALM, or STORM. BP-FPALM is suitable for superresolution optical imaging with live cells. BP-FPALM does not require the reducing (low oxygen) conditions that limit the STORM technique.

BP-FPALM has other inherent advantages over alternative instruments. It uses only one EMCCD camera, thus avoiding the problem of synchronization that occurs with two separate cameras. Since it is a nonscanning superresolution three-dimensional imaging technique it is significantly faster than alternative scanning techniques. BP-FPALM has another important advantage over alternative techniques in that it mitigates against the artifacts of localization in scanned

instruments that occur in single-molecule studies due to blinking and bleaching (Juette et al., 2008).

FPALM and PALM superresolution techniques use identical or similar photoactivatable fluorescent proteins; the instrumentation is nearly identical too. When these techniques are used with living cells, similar intensities are used. For example, FPALM imaging of PA-GFPs in living cells uses $\sim 10^3$ W/cm^2 of 488-nm excitation light and $\sim 10^2$ W/cm^2 of 405-nm activation light (Gould et al., 2009).

15.6 Photoactivated Localization Microscopy with Independently Running Acquisition

Another variant of the PALM technique is called photoactivated localization microscopy with independently running acquisition (PALMIRA). The name refers to the fact that this technique detects the spontaneous cycles of photoactivatable probes in the OFF state and the ON state. This process occurs in the absence of detector synchronization or independently running acquisition. In 2007 several publications introduced PALMIRA (Bock et al., 2007; Egner et al., 2007; Geisler et al., 2007). I frame my discussion in the typical format: the problem, the solution, the experimental or simulation achievement, and the limitations, advantages, and cautions with respect to alternative techniques.

What is the problem that the investigators set out to solve? The investigators sought to develop a variant of the superresolution optical microscope technique of photoactivated localization microscopy (PALM) (Betzig et al., 2006) that would significantly reduce acquisition times to 2.5 min. The PALM, FPALM, and STORM techniques have several limitations (Geisler et al., 2007). They are sensitive to the background since they operate by detecting single molecules. Another problem is the low contrast between the nonfluorescent OFF state and the fluorescent ON state of these fluorescent probes that mandates the use of low labeling densities. Furthermore, the first PALM publication had an acquisition time of 2–12 h (Betzig et al., 2006).

The experimental solution they came up with was a new technique called photoactivated localization microscopy with independently running acquisition (PALMIRA) (Geisler et al., 2007). The name PALMIRA originates from the capability of the technique to detect the emission of nontriggered spontaneous OFF–ON–OFF cycles of a single switchable fluorescent molecule; however, it operates in the absence of any detection synchronization (Geisler et al., 2007).

How does PALMIRA work? Geisler et al. report the use of a unique reversible fluorescent switchable protein called rsFastLime (RSFP). The rsFastLime fluorescent probe, a variant of the reversible switchable fluorescent protein Dronpa, has an important unique property in that the same wavelength is used to turn the molecule ON (activated state) and to excite the molecule. Because of this property a single laser can be used as the light source. An argon ion laser supplied the required 488-nm laser line. The equipment for the PALMIRA technique is a basic epi-fluorescence optical microscope, an argon ion laser, and an electron-multiplying charge-coupled device (EMCCD) camera. All superresolution optical microscopy techniques are

subject to drift and vibrations. The tracking of fluorescent microspheres is one method used to correct specimen drift. The major achievement of PALMIRA is the significant reduction in acquisition time. The authors demonstrated a 50-nm lateral resolution with a 2.5-min acquisition time for their PALMIRA microscope. Using the rsFastLime fluorescent probe results in data acquisition that is 100-fold faster than that of previous techniques (Geisler et al., 2007).

A more detailed publication on the PALMIRA technique demonstrated super-resolution microscopy of entire cells by the asynchronous localization of photo-switchable fluorescent molecules (Egner et al., 2007).

15.7 Superresolution Optical Fluctuation Imaging

In 2009 Dertinger et al. developed a new superresolution microscopy technique—not based on localization—for superresolution in three dimensions (Dertinger et al., 2009). This new three-dimensional superresolution optical microscopy technique was called superresolution optical fluctuation imaging (SOFI) microscopy (Dertinger et al., 2009; Dertinger et al., 2010a; Dertinger et al., 2010b). The SOFI microscopy technique is not a localization technique, and therefore differs from other superresolution methods that use single-molecule switching. In the SOFI method superresolution can be achieved even when many fluorescent molecules located in a diffraction-limited volume are in the ON or emitting state. The SOFI technique works even in the presence of high densities of fluorescent molecules that are simultaneously emitting. The SOFI superresolution technique is based on a sequence of images collected using an EMCCD camera together with statistical analysis of temporal fluctuations due to fluorescence blinking.

What is the problem the investigators tasked themselves with solving? They attempted to develop a new three-dimensional superresolution optical microscopy technique with two important capabilities: (1) the new technique must be as free of background intensity as possible, and (2) the acquisition time should be on the order of a few seconds. The first aim of a background-free technique is important because deconvolution image-processing techniques are very sensitive to the presence of background intensity. A background-free image is more compatible with deconvolution methods. The second aim is important because very short acquisition times reduce the effects of specimen drift. With very short acquisition times it is not necessary to use fiduciary markers to computationally correct for specimen drift caused by vibration and thermal gradients.

How does the SOFI technique differ from the superresolution optical microscopy techniques that I have described and discussed in previous sections of this chapter? SOFI differs in not being a localization technique; it achieves a narrower PSF in the microscope. SOFI differs from the other superresolution techniques discussed in this chapter in that it does not require controlled or synchronized photoactivation; it is based on independent stochastic fluctuations of photons emitted by fluorescent molecules. SOFI instrumentation is less complex than that required for alternative techniques used to achieve superresolution.

How does the SOFI superresolution technique work? The basic idea is to make a movie, a temporal sequence of images. The independent stochastic or random fluctuations of photons emitted by the fluorescent molecules are subjected to statistical analysis (Dertinger et al., 2009; Dertinger et al., 2010a; Dertinger et al., 2010b). The temporal fluorescence fluctuations of emitting molecules (a.k.a. fluorescence intermittency) is subjected to statistical analysis. However, the SOFI technique has a specific set of requirements or constraints (Dertinger et al., 2009). First, the fluorescent molecules must exist in two optically distinguishable emission states. Second, the different fluorescent molecules must repeatedly switch from one state to the second state and this switching must be random. Third, the acquired image must be formed by pixels whose dimensions are smaller than those of the diffraction limit. Good candidates for SOFI include quantum dots and the fluorescent dyes used in dSTORM (discussed in the next sections); specific buffers are used to bring about the random and independent intensity fluctuations required.

The SOFI technique posits that the spatial positions of emitting fluorescent molecules are stable during image acquisition. The temporal changes are the result of blinking (i.e., changes in the fluorescent state of separate fluorescent molecules). The blinking effect is caused by a transition to the triplet state, an OFF or dark triplet state, via the process of intersystem crossing (Dertinger et al., 2010b). Furthermore, there is no correlation in time between separate fluorescent emitters. The sequence of SOFI-acquired images are measures of the brightness of emitters and the extent of the correlation. The following result follows from detailed higher order statistical analysis (nth-order cumulant) of the acquired sequence of images: the effective point spread function is now the nth power of the original point spread function and resolution is improved by $\sqrt{n}$ (Dertinger et al., 2009). The SOFI technique can be implemented to higher nth orders, which in principle can result in ever-higher spatial resolutions. However, two factors place practical limits on the nth order of statistical analysis: (1) higher order analysis requires longer measurement times, and (2) the problem of the brightness of emitters in which the higher the order of the analysis, the larger the contribution of very intense emitters and the smaller that of less intense emitters.

Instrumentation for the SOFI technique is readily available. A typical widefield fluorescence microscope is required together with a camera that can rapidly acquire images with high sensitivity (e.g., an EMCCD camera). The SOFI technique can be used with a variety of optical microscopes: widefield, confocal, and TIRF. TIRF systems are preferred since they significantly reduce background fluorescence. A wide variety of fluorescent probes can be used with SOFI: photoswitchable fluorescent proteins, organic fluorescent probes that reveal the blinking or spontaneous fluctuations of the fluorescence, and quantum dots. The first publications demonstrated the SOFI technique using blinking quantum dots as fluorescent probes.

Finally, what was achieved? The SOFI technique produced three-dimensional superresolution images that were background free and showed greatly enhanced image contrast. The acquisition time was limited to a few seconds. The investigators achieved imaged microtubules (an α-tubulin network) of 3T3 fibroblast cells that were previously immunolabeled with quantum dots (Dertinger et al., 2009). The authors demonstrated a fivefold enhancement in spatial resolution compared with a standard widefield fluorescence microscope. Furthermore, there was a major reduction in background and enhanced contrast in images acquired using the SOFI technique (Dertinger et al., 2009). The SOFI technique is applicable to a variety of fluorescent labels: fluorescent proteins, quantum dots, and organic dyes that can be made to blink. There is another advantage to using the SOFI technique compared with alternative superresolution techniques based on localization: the very short acquisition times required in the SOFI technique mitigate the problems of specimen drift that are so prevalent in other techniques that have significantly longer acquisition times (Dertinger et al., 2010b). The main limitation of the SOFI superresolution technique is photobleaching, which becomes problematic when higher order statistical correlations are used.

15.8 Stochastic Optical Reconstruction Microscopy

15.8.1 Introduction

Three independent groups published the original key papers on new widefield superresolution microscopy based on photoswitching of fluorescent molecules located within a diffraction volume. A single fundamental idea is common to these inventions: the random or stochastic switching of fluorescent molecules between the ON and OFF states. This process converts large numbers (tens of thousands) of fluorescent molecules located in a diffraction-limited volume to a sparse subset of fluorescent molecules in the ON state. Under this condition it is possible to measure the centroid of the PSF, and therefore the position of a fluorescent molecule with a precision that is significantly below the diffraction limit. This is the basis of the superresolution capability of the PALM, FPALM, and STORM superresolution single-molecule techniques.

It is critical to stress that the measured position of the fluorescent label may not be its real position. Again, the precision of a localization measurement may not be the same as the accuracy of a localization measurement. Several factors can affect the accuracy of localization measurements: physical motion of the label, optical aberrations, optical resolution, the number of photons detected, the lifetime of the fluorescent label, and the physical size of the link between the fluorescent label and the object to which it is attached.

It is important to stress that it was these three publications that led the way to the first superresolution techniques based on single-molecule imaging and photoswitching capable of separating a large number of fluorescent molecules located in the same diffraction-limited volume. It was in this aspect that these three innovative publications differed from many previous publications that did not have the capability to separate large numbers of fluorescent molecules located in a diffraction-limited volume. The STORM technique enabled the fluorescence image of the specimen to be reconstructed by localizing single fluorescent molecules switched on and off by light of different frequencies.

Prior to the first publication of the STORM technique the Zhuang group published a paper that contained several of the technical prerequisites for STORM. It was published in its final edited form in the March 18, 2005 edition of *Physical Review Letters* (Bates et al., 2005). Their publication highlighted the importance of an optical switch to switch fluorescent molecules between the ON and OFF state. Using an optical switch to transfer fluorescent molecules between two states (ON and OFF) when implemented in a random or stochastic manner is the basis of the STORM technique. The authors demonstrated a single-molecule optical switch that could function as a short-range spectroscopic ruler (FRET) to probe distances down to 1 nm; distance measurements in this range by energy transfer have previously never been demonstrated. Distance dependence was significantly enhanced compared with Förster's resonance energy transfer (FRET) (Bates et al., 2005). The authors determined distance dependence by separating the Cy5 and Cy3 by different known distances (Bates et al., 2005). Typically, Förster's resonance energy transfer uses donor–acceptor pairs that have a Förster radius R_0 of 4–6 nm (Masters, 2014).

Bates et al. (2005) described development of the single-molecule switch, the switching mechanism, and specific buffer requirements that together formed the foundation of their invention of STORM. I emphasize the aspects of the authors' achievement that relate to the switching behavior of fluorescent probes, the conditions of the buffer that promotes the switching, and the authors' suggested mechanism. This discussion is pertinent to understanding the first STORM publication (Rust et al., 2006).

How does a molecular switch operate? First, what is the nature of the two fluorescent molecules that comprise the bipartite molecular switch? The first molecule Cy5 is the primary switch molecule called the reporter molecule, and the second molecule Cy3 that permits the primary molecule Cy3 to switch is called the activator molecule. The two indocarbocyanine synthetic molecules are members of the polymethine group. The digit that follows the letters Cy refers to the number of carbon atoms that separate the two indolenine groups. The Cy3 molecule is required to switch the Cy5 molecule from the OFF state to the ON state. The mechanism behind this is unclear.

What are the spectral properties of Cy3 and Cy5? These molecules are characterized by narrow excitation and emission spectra. The peak absorption of Cy3 is

550 nm and the peak emission is 570 nm. The peak absorption of Cy5 is 650 nm and the peak emission is 670 nm.

Operation of the molecular switch requires a buffer that contains the reducing agent β-mercaptoethanol (Nikon Corporation, 2013, 2015). In addition, a system composed of glucose, glucose oxidase, and catalase is present in the buffer to reduce the concentration of oxygen to very low levels, which serves to reduce the rates of photobleaching of fluorescent molecules (Nikon Corporation, 2013, 2015).

How did Rust, Bates, and Zhuang (2006) test the molecular switch? They covalently attached the Cy5 and Cy3 to different strands (on one end) of a double-stranded DNA molecule. The other end of the double-stranded DNA molecule was immobilized on a quartz surface.

Rust, Bates, and Zhuang (2006) used alternating laser pulses of red (638 nm) light and green (532 nm) light on the Cy5/DNA/Cy3 constructs. Another low-light intensity laser at (638 nm) continuously illuminated the construct and monitored the fluorescent state of the Cy5 molecule. The authors discovered that single molecules of Cy5 were readily switched between two states: a pulse of red laser light (638 nm) switched the Cy5 molecule to the OFF state, and a pulse of green laser light switched the molecule back to the fluorescent state or the ON state. The reversible nature of this process showed that the Cy5 molecule could be reversibly switched to an OFF state and not be permanently photobleached. The red emission of the fluorescent Cy5 molecule was controlled by the green (532 nm) laser light (Rust et al., 2006). Reversible switching was recorded over multiple cycles (Rust et al., 2006). The authors observed that whenever the green laser was turned off the red laser light converted the molecules to the OFF state; whenever the green (532 nm) laser was operating with sufficient intensities so that it could override the red laser the Cy5 molecules were converted back into an ON or fluorescent state.

A feasible mechanism underlying switching was proposed by Bates, Blosser, and Zhuang (2005). It was previously known that cyanine fluorescent dyes can undergo photoinduced transitions from their excited singlet electronic state into a triplet state that is nonfluorescent or dark. It was also known that the rate of intersystem crossing from the excited singlet state to the excited triplet state is enhanced in the presence of heavy atoms such as iodide. These heavy atoms induce spin–orbit coupling and result in increased rates of intersystem crossing.

Some four years later (2009) the actual photoswitching mechanism of cyanine fluorescent molecules was elucidated (Dempsey et al., 2009). The authors found that the concentration of the thiol and the pH of the buffer control the rate at which cyanine molecules are switched to the OFF state. Thiol–cyanine molecules in the OFF state were investigated with mass spectroscopy. It was found that the thiol group was attached to the polymethine bridge. Cyanine molecules exposed to red light switch to the OFF state, and this switching is enhanced in the presence of primary thiols in the buffer solution. This provides the mechanism behind transition to the state; the thiol group affects the conjugated pi–electron system of the cyanine molecule (Dempsey et al., 2009). There are a variety of imaging buffers affecting

the rate constants of photochemical transitions between the two states (Nikon Corporation, 2013, 2015).

Dempsey, Bates, Kowtoniuk, Liu, Tsien, and Zhuang (2009) observed the single-molecule switching mechanism operating only in the presence of β-mercaptoethanol and those components of the system removing oxygen from the buffer system. Dempsey, Bates, Kowtoniuk, Liu, Tsien, and Zhuang (2009) came to the conclusion that the switching mechanism was not from the excited singlet state to the excited triplet state, but to another OFF state; the fluorescent molecule in the excited singlet state passes through the intermediate excited triplet state and then to another OFF state (Bates et al., 2005).

15.8.2 The First Stochastic Optical Reconstruction Microscopy Publication

The first STORM publication, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)" was received on July 7, 2006; accepted on July 31, 2006; and published online in *Nature Methods* on August 9, 2006 (Rust et al., 2006). The first STORM publication cited a publication titled "Carbocyanine dyes as efficient reversible single-molecule optical switch" (Heilemann et al., 2005), but its connection to the first STORM publication is ambiguous. Heilemann et al. (2005) is fully discussed in Section 15.9.

The first STORM publication built on previous publications demonstrating the utility of carbocyanine fluorescent probes as reversible single-molecule optical switches (Bates et al., 2005; Heilemann et al., 2005). Rust et al., (2006) went on to develop the switch into a superresolution fluorescence microscopy technique built on very high-precision localization of photoswitchable fluorescent probes (STORM). In this first STORM publication no mention was made of imaging fixed or living cells.

What was the problem that STORM solved? STORM, PALM, and FPALM were all developed to solve a problem that hindered the development of superresolution microscopy: the very large numbers of fluorescent molecules located within the diffraction volume. A means of converting this very large number into a sparse subset was needed. While previous techniques could use blinking of quantum dots to precisely localize between two and five single molecules in the same diffraction-limited volume, the methods could not work with hundreds of single fluorescent molecules situated within the same diffraction-limited volume. Something new was required and in 2006 three superresolution microscopy techniques were invented that were capable of overcoming the problem.

The solution was a key development common to STORM, PALM, and FPALM. It involved random or stochastic switching of fluorescent molecules between the ON and OFF states. Eric Betzig presented his data showing how such superresolution imaging could be accomplished with his PALM earlier on April 21, 2006 at

the 2006 Frontiers in Live Cell Imaging conference held at the National Institutes of Health, Bethesda, Maryland. This process converts the large numbers of fluorescent molecules located in a diffraction-limited volume into a sparse subset of fluorescent molecules in the ON state. In each cycle only a very small fraction of fluorescent molecules are in the ON state. By forming a sparse subset of molecules it is possible to measure the position of the fluorescent molecule(s) with nanometer precision significantly below the diffraction limit. The set of all the cycles and the positions of all the fluorescent molecules is used to form the final superresolution image of the specimen labeled with reversible, switchable fluorescent probes. In a nutshell, that was the solution to the above problem.

I frame the following discussion by posing some questions. How does STORM function as a superresolution technique? How was the STORM technique validated experimentally? What resolution was achieved with STORM? And, finally, what were the concerns, problems, and limitations of STORM as presented in Rust et al., (2006)?

How does the STORM technique work? As shown in precursor publications the technique is based on the switchable fluorescent molecules Cy5 and Cy3 (located close to each other) and two lasers (one red and one green). The bipartite optical switch was illuminated with red (638 nm) and green (532 nm) laser pulses turned on and off alternately. A buffer system comprising glucose, glucose oxidase, and catalase was used to mitigate against photobleaching by depleting oxygen in the buffer. The compound β-mercaptoethanol was also present in the buffer because it is necessary for the photoswitching of cyanine dyes.

I now list the sequence of steps used in the STORM process. For the moment I ignore presenting details about specimen preparation and labeling and the technique to correct for specimen drift. The sequence of steps I list is repeated many times. First, the bipartite optical switch molecules Cy5 and the nearby Cy3 are chemically attached to the specimen using dense labeling. This is not a simple labeling process. According to the 2015 Nikon booklet titled *Super Resolution Microscope N-SIM/N-STORM* the Cy5 and Cy3 fluorescent molecules are chemically attached to the same secondary antibody. This secondary antibody containing the Cy5 and Cy3 molecules is then attached to the primary antibody that binds the target protein of the specimen. A red laser and a green laser are used to initiate the optical switching process. The process begins with all the fluorescent molecules being driven to the nonfluorescent or OFF state by applying an intense pulse of red laser light. This is followed by a green laser pulse that switches a very small fraction of the fluorescent molecules to the ON state. The red laser excites fluorescent molecules in the ON state and the photons emitted from these single molecules are collected, resulting in the positions of these molecules being located to a precision greater than the diffraction limit. The centroid of the PSF gives the molecular location. The final superresolution image is formed from the precise location of all the fluorescent molecules after correction for specimen drift. In the next step the sparse subset of molecules turned to the ON state will fluoresce until they are switched to the OFF state or until they are photobleached irreversibly.

The Cy5 optical switch is turned on and off many hundreds of times prior to photobleaching destroying the optical switch. Note that photobleaching is a permanent or irreversible process from which molecules never recover. The Cy5 fluorescent molecules in the ON state are excited with a red laser line (633 nm and an intensity of 30 W/cm^2) and the same laser turns the Cy5 molecules to the dark or OFF state. The Cy5 molecules in the OFF state are switched to the ON or fluorescent state with a green laser line (532 nm and an intensity of 1 W/cm^2). Illumination of the specimen continues for tens of thousands of cycles with radiation alternating between the red and green lasers (Rust et al., 2006).

I now summarize the steps involved in the STORM technique and include additional details (Rust et al., 2006). First, optical switch molecules are used to label the specimen in a uniform and dense manner. If labeling is not sufficiently dense, then the final superresolution image will poorly reflect the actual structure of the specimen. Second, in each time period of the repetitive cycles constituting the STORM technique light is used to alter only a very sparse subset of photoswitchable fluorescent molecules. The selection of molecules transformed into the ON state is a random or stochastic process. The key idea is that molecules in the ON state have PSFs that do not overlap with the PSF of another photoswitchable molecule. This precludes the possibility of having large numbers of fluorescent molecules simultaneously turned to the ON state while they are located in the same diffraction-limited volume. Third, the molecules in the ON state are excited and fluoresce emitting photons that are collected. The precision of localization depends on the number of detected photons emitted for the duration that the molecule is in the ON state, as well as fluorescence from the background. The localization method is based on finding the centroid of the PSF of each fluorescent molecule. Fourth, the sequence is repeated tens of thousands of times (cycles). In each cycle a different subset of fluorescent molecules is localized, and the STORM sequence continues until the entire labeled specimen is sampled. From the localized positions of all photoswitchable fluorescent molecules the final superresolution image of the specimen is formed (Rust et al., 2006).

The STORM technique can be implemented in a standard epi-fluorescence microscope or in a total internal reflection fluorescence (TIRF) microscope. The TIRF implementation reduces background fluorescence.

How was the STORM technique validated in the first STORM publication? Now that the principle of the STORM technique has been explained a new question arises in the mind of the skeptical reader. How did Rust, Bates, and Zhuang (2006) of the first STORM publication validate their new technique? All validation experiments were performed when the Cy5 and Cy3 attached to biological structures were in close proximity. For the studies with the optical switch Cy5 with Cy3 in close proximity in each cycle of the STORM technique 3000 photons were detected. This is equivalent to a precision in localization of 4 nm. However, Rust, Bates, and Zhuang (2006) measured precision in the localization of fluorescent molecules of 8 nm and attributed the difference to uncorrected specimen drift and

optical aberrations. This corresponds to a 20-nm resolution. In other studies fluorescent switches separated by 40 nm have been resolved. Rust, Bates, and Zhuang (2006) clearly resolved two optical switch fluorescent molecules separated by 46 nm on an immobilized piece of double-stranded DNA (dsDNA). In similar studies when bipartite optical switches (Cy5 with Cy3 in close proximity) were attached to RecA-coated circular plasmids of DNA the STORM technique yielded superresolution images (Rust et al., 2006). However, it is important to point out that photoswitchable molecules are attached to the specimen through a secondary antibody whose linker (fluorescent molecule to specimen) has a total length of 10 nm. Therefore, the length of the linker must be considered when calculating localization precision based on the number of detected photons. In the PALM technique the linker for photoactivatable proteins is only 2–4 nm and the additional uncertainty is much smaller.

What was the initial achievement of the STORM technique? The authors demonstrated an imaging resolution of 20 nm (Rust et al., 2006). The authors of the first STORM publication included two undemonstrated proposals in their 2006 paper: (1) they posit that the STORM technique could be used in principle with other photoswitchable fluorescent probes including fluorescent proteins (PALM and FPALM techniques), and (2) they further posit that the STORM technique can potentially be used for live cell superresolution microscopy (Rust et al., 2006).

What problems with the STORM technique were discussed or not discussed in the first STORM publication (Rust et al., 2006)? First, the optical switch is large and complex. It consists of two different molecules Cy5 and Cy3 that must be in close proximity for the switch to function. The authors refer to the Cy5–Cy3 pair of fluorescent molecules as the switch, but it consists of two distinct fluorescent molecules. Second, while the authors state that the STORM method can work with other (besides the Cy5–Cy3 pair) photoswitchable fluorescent molecules and with photoswitchable fluorescent proteins (PALM and FPALM), they do not demonstrate that proposal. Third, the authors allude to the use of STORM in superresolution live cell microscopy. However, the STORM technique requires a special buffer system to deplete oxygen from the buffer system and contains the molecule β-mercaptoethanol (Nikon Corporation, 2013, 2015). This special buffer system has not been shown to be compatible with live cell imaging. Furthermore, the cyanine dye molecules (Cy5, Cy3) do not penetrate cell membranes. For these dyes to penetrate cells detergents such as Triton are required to permeabilize the cell membranes (Nikon Corporation, 2013, 2015). Again, the need to permeabilize cell membranes questions the use of STORM for live cell imaging. Fourth, the authors mention the problems of stage drift and specimen drift. While such problems can in principle be corrected using fiduciary markers, the correction is not always complete. The problem of specimen drift is a general problem of all types of superresolution localization microscopy.

15.8.3 Developments of Stochastic Optical Reconstruction Microscopy

In this section I discuss some of the important advances and achievements that further developed the STORM technique after the initial publication on the technique by Rust et al., (2006). I exclude the technique of direct STORM (dSTORM), which I discuss in Section 15.9. In this section I discuss four innovative publications by Zhuang and her coworkers, each of which achieved significant advances in the STORM technique: multicolor STORM could be applied to fixed mammalian cells, the use of optical astigmatism to achieve axial resolution with STORM, rapid 3D STORM for live cells, and a dual-objective STORM instrument with optical astigmatism for three-dimensional imaging. I frame the discussion in two parts: statement of the problem and its solution.

Bates et al. in their *Science* paper introduced the technique of multicolor stochastic optical reconstruction microscopy (STORM) based on the development of bright, photoswitchable fluorescent probes of various colors (Bates et al., 2007). The problem involved developing different sets of optical switches that emit photons of different frequencies (colors) for the STORM technique.

How does multicolor STORM work? First, I review the nature of optical photoswitchable molecules upon which STORM is based. Each switch is composed of two different molecules that must be in close proximity; labeling the specimen consists of attaching many pairs of molecules on the specimen. The authors labeled one part of the switch as the "reporter" fluorescent molecule and the second part of the switch, which must be in close proximity, as the "activator" molecule. Without the activator molecule adjacent to the reporter fluorescent molecule the photoswitchable probe is not functional. Reviews have shown that it is the reporter fluorescent molecule that is optically switched between the fluorescent ON state and nonfluorescent OFF state. The multicolor STORM technique is based on pairing various reporter molecules with various activator molecules, thereby forming photoswitchable probes of various colors. The authors used different pairs of molecules to form photoswitchable probes with different colors. Cy3 was the activator. The reporter molecules in the switch consisted of Cy5, Cy5.5, and Cy7 (Bates et al., 2007).

What did Bates et al. achieve with their multicolor STORM technique? They labeled various DNA constructs with different optical switches: Cy5–Cy3, Cy5.5–Cy3, and Cy7–Cy3. They tested other activator and reporter molecules constituting optical switches. They observed that the various optical switches could be distinguished by their different emission wavelengths and that different pairs constituting the switches responded optimally to different laser lines used to activate the switch (Bates et al., 2007).

Similarly, they used fixed mammalian cells with different optical switches and they applied the STORM technique. They achieved a lateral resolution between 20 and 30 nm using the multicolor STORM technique (Bates et al., 2007).

The next publication I discuss addresses the problem of how to modify the instrumentation of the STORM technique such that it was capable of three-dimensional imaging (i.e., the introduction of axial resolution). Huang et al. in their paper in *Science* titled "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy" introduced optical astigmatism to modify the widefield, nonscanning STORM technique so that the axial and lateral positions of single fluorophores could be measured. The authors achieved an image resolution of 20–30 nm in the lateral dimensions and 50–60 nm in the axial dimension (Huang et al., 2008a). Inventing a nonscanning STORM technique in which resolution in the axial and lateral directions was enhanced tenfold compared with standard diffraction-limited fluorescence microscopy was notable.

How did 3D STORM work according to Huang et al. (2008a)? It had previously been demonstrated that if defocus is used to modify the image then superresolution is achieved in the axial direction with little effect on lateral localization accuracy (van Oijen et al., 1999). Motivated by the innovative 1999 publication of van Oijen et al. and other similar publications Huang et al. extended the concept of defocus in the image and applied it to the STORM technique (Huang et al., 2008a).

Instead of defocus Huang et al. used the previously published technique to introduce astigmatism into the image. They placed a weak-power cylindrical lens into the imaging path between the microscope objective and the imaging lens element that was in front of the EMCCD camera. The use of a cylindrical lens to encode axial position was first proposed in 1994 (Kao and Verkman, 1994). This resulted in two different focal planes for the x- and y-directions. Therefore, the orientation of the fluorescent molecule and its ellipticity changed as a function of its location in the axial direction (z-direction). If the fluorescent molecule is located exactly between these two focal planes, then the PSF is round; above or below this point the image of the fluorescent molecule is ellipsoidal and the long axis of the image is in the x- or y-direction (Huang et al., 2008a).

The next step is to fit the image of the fluorescent molecule with a two-dimensional elliptical Gaussian function. This yielded the peak position in terms of x and y and peak widths in x and y; all these parameters finally yielded the z-coordinate of the fluorescent molecule. The above analysis demonstrated that an experimental calibration curve could be formed and used in the STORM analysis to achieve three-dimensional STORM with superresolution in three dimensions (Huang et al., 2008a).

The next publication I discuss addresses how relatively fast, three-dimensional superresolution images of live cells can be achieved with the STORM technique. The solution was first proposed in a paper in *Nature Methods* titled "Fast, three-dimensional super-resolution imaging of live cells" (Jones et al., 2011). The manuscript was received on September 1, 2010; accepted on April 2011, 15; and published online on May 8, 2011.

How did they achieve this? As I and others have previously stated there are two factors applicable to all superresolution microscopy techniques that limit the advancement of such innovative techniques. Probe development is a major limitation to progress. Specimen preparation, which includes chemical, immunological, and genetic techniques to label the specimen with fluorescent probes, is also a critical factor in the advancement of superresolution microscopy. These impediments to progress stand in contrast to the rapid advances in detector technology that produced electron-multiplying charge-coupled device (EMCCD) cameras (iXon 860; Andor Technology) that have low noise and high-acquisition rates. An alternative to the EMCCD camera is the sCMOS camera (Andor Technology; Hamamatsu Photonics). Selection of the appropriate camera for the problem to be investigated using superresolution microscopy will depend on the questions to be addressed, the signal and noise levels, and temporal resolution requirements (the trade-offs are discussed on the Andor Technology website and the Hamamatsu Photonics website).

In the first STORM publication the authors posited but did not demonstrate that the STORM superresolution technique should be applicable to a variety of photoswitchable probes including photoswitchable proteins and that the STORM technique was applicable to live cell imaging (Rust et al., 2006). Jones et al. (2011) finally demonstrated that a variety of photoswitchable fluorescent molecules could be used in the STORM technique to image live cells (Jones et al., 2011).

How did the investigators accomplish this? Jones, Shim, He, and Zhuang (2011) labeled specific proteins in live cells with photoswitchable fluorescent molecules. The performance of different molecules making up the set of photoswitchable fluorescent molecules varies in the STORM technique. Such parameters as the photophysics of the fluorescent molecules, their brightness, trade-offs in the spatial and temporal resolution they achieve, their stability to photobleaching, their stability during storage, and the labeling density they achieve are critical factors for the user to consider. In general, it is highly desirable to have photoswitchable fluorescent molecules that are bright (i.e., they emit a large number of photons per unit time and their excitation wavelengths are positioned in the red, far-red, or infrared regions of the spectrum compared with the blue and ultraviolet wavelengths). While photodamage to cells is dependent on light intensity, oxygen concentration, and length of time exposed to light, excitation wavelengths in the red to infrared are less damaging to cells.

Jones, Shim, He, and Zhuang (2011) compared six different photoswitchable fluorescent molecules for use on live cells. Another important aspect of their publication is that they used a variety of photoswitchable fluorescent molecules on the same live cells to validate images obtained using the STORM technique. The photoswitchable fluorescent molecules they investigated consisted of the cyanine probe Alexa Fluor 647, three different molecules that can permeate cell membranes (Atto655, tetramethylrhodamine, and Oregon Green), and two Eos protein derivatives (mEos2 and tdEos). While all these photoswitchable fluorescent molecules yielded three-dimensional superresolution images, the quality of the final superresolution images varied (Jones et al., 2011). The authors corrected for specimen drift by attaching fluorescent beads to the coverglass; these beads were fiducial markers and their displacement was used to correct for specimen movement.

What was achieved? The STORM technique was used on live cells to obtain three-dimensional superresolution of 30 nm in the lateral direction and 50 nm in the axial direction (Jones et al., 2011). Furthermore, the authors demonstrated the use of two-color STORM on live cells. The question arises as to the physiological state of cells that are used in STORM live cell imaging. The authors state that the cyanine fluorescent molecules that constitute part of the optical switch required for STORM to operate switch to the OFF state by combining with thiol present in the buffer. While the authors explain that β-mercaptoethanol is required to be present in the buffer for live cell imaging, there is no discussion of the effect on living cells. The effects of β-mercaptoethanol not only on cell viability but also cell function are dependent on length of time exposed, the specific cell type, and concentration of the thiol molecule in the buffer solution.

Furthermore, Jones et al. stated that STORM required a system in which oxygen can be depleted from the buffer (Nikon Corporation, 2013, 2015). Oxygen depletion was required to mitigate against and reduce the rate of photobleaching of fluorescent molecules. In their STORM studies of live cells the authors stated that the oxygen depletion system (glucose, glucose oxidase and catalase) was added to the buffer just before the start of image acquisition. Nevertheless, critical questions remain: What do we mean by live cell imaging? What is the effect of oxygen depletion on the buffer containing the cells. What physiological and functional alterations does oxygen depletion cause? These questions are addressed in a book titled *Live Cell Imaging, A Laboratory Manual* (Goldman and Spector, 2005).

The next publication I discuss addresses the achievable resolution of superresolution microscopy techniques such as STORM and its dependence on high-precision localization and collection of a large number of photons while the molecule is emitting. When a single fluorescent molecule is illuminated with the appropriate wavelength of excitation light it will undergo a transition from the ground state to a higher energy electronic state. While there are many pathways of de-excitation, the emission of a photon from the first excited state is termed fluorescence. This process of spontaneous emission differs from the process of stimulated emission. In stimulated emission the incident photon's propagation direction is exactly the same as the propagation direction of the emitted photon. However, in spontaneous emission (fluorescence) the emitted photon can propagate in any random direction. Therefore, the collection efficiency of emitted photons in the standard epi-illumination microscope that uses a single objective to collect the photons is limited by the numerical aperture of the objective. While a fraction of emitted photons will enter the microscope objective, a much larger fraction is not detected because their direction of propagation precludes them from entering the aperture of the microscope objective. This explains why a dual-objective STORM system has enhanced photon collection efficiency and therefore enhanced precision of localization. As a result of the sufficiently high label density of fluorescent molecules on the specimen the enhanced number of collected photons will improve the lateral and the axial resolution of the STORM technique.

The solution to this problem was given in a paper titled "Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton"

(Xu et al., 2012). The investigators achieved 3D STORM using the previously described method of astigmatism imaging. They combined astigmatism imaging with two opposing microscope objectives between which the specimen was located. The dual-objective STORM microscope achieved a lateral resolution of less than 10 nm and an axial resolution of less than 20 nm. The specimen was made up of fixed cells. The authors obtained superresolution, three-dimensional images of the ultrastructure of the actin cytoskeleton.

How does dual-objective STORM work? 3D STORM is based on the use of a cylindrical lens placed in the detection pathway in front of the EMCCD camera. Experimental details are the same as those described in Huang et al., (2008a), which I previously discussed. Two opposing microscope objectives collect photons emitted from the fluorescent molecules in the specimen. The use of two opposing microscope objectives to collect photons emitted from the fluorescent probes in the STORM technique should double the number of photons detected and yield a resolution enhancement of the square root of 2 or a 1.4-fold greater resolution. There is a trade-off between the number of acquired image frames and the image quality of the final superresolution image. If more frames are acquired, then the image quality of the final superresolution image is improved, but at the cost of increased acquisition time.

Next, I discuss an innovative paper highlighting small-membrane fluorescent molecules suitable for use in STORM microscopy (Shim et al., 2012). Typically, STORM microscopy uses photoswitchable fluorescent molecules. The bipartite photoswitch actually consists of two distinct molecules that must be in close proximity for the switch to operate. As pointed out earlier, one of the photoswitchable fluorescent molecules is called the "reporter" (i.e., the fluorescent molecule that undergoes photoswitching), and the second required molecule is called the "activator." Usually these bipartite photoswitches are chemically attached to antibodies, either primary antibodies or in other cases the primary antibody with the reporter is bound to a secondary antibody. The reporter combined with the two bound antibodies constitute the probe label. While the antibodies bind with high specificity to their targets, they are difficult to introduce inside live cells and their large size limits the ultimate resolution of superresolution localization techniques. Furthermore, the STORM technique requires an oxygen depletion system and a buffer system that contains a member of the thiol group such as β-mercaptoethylamine (BME) or mercaptoethylamine (MEA) (Nikon Corporation, 2013, 2015). Using that background I am now in a position to present the problem the authors of the next publication tasked themselves with solving.

This problem involves identifying and testing a number of photoswitchable fluorescent molecules with certain characteristics. Photoswitchable fluorescent molecules should be small molecules that readily permeate live cells and attach to membranes with high specificity in and on the surface of the cells. These membrane probes should operate in the STORM technique on live cells and achieve superresolution and high temporal resolution such that the dynamic processes of membrane transformation can be studied. In addition, these small fluorescent molecules

should achieve high labeling density in live cells. These membrane probes should operate in the absence of activator molecules. Furthermore, the buffer systems used with STORM should not require the presence of β-mercaptoethylamine (BME) or mercaptoethylamine (MEA).

The solution to this problem was clearly demonstrated and explained in a paper titled "Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes" (Shim et al., 2012). What the authors discovered was remarkable. The authors began by investigating the photoswitching characteristics of eight small-molecule membrane probes. I will discuss the key achievements that stemmed from studies of DiI, a long alkyl group cyanine dye that is very lipophilic and therefore strongly binds to the plasma membrane. The authors' key unexpected finding was that the DiI densely labeled the plasma membrane within a few minutes of incubation with live cells. DiI-bound fluorescent molecules showed photoswitching characteristics in the absence of the activator molecule usually required for reporter fluorescent molecules to operate as a bipartite photoswitch. When DiI-bound fluorescent molecules were illuminated with light at 561 nm it was fluorescent and then spontaneously switched to the dark or OFF state. When the DiI-bound molecules were subsequently illuminated with light at 405 nm they switched from the OFF state to the ON or fluorescent state. While I am describing only one member of the lipophilic cyanine membrane fluorescent probe family similar characteristic switching behavior was also observed for DiD and DiR. These three cyanine fluorescent molecules are excited by different wavelengths. The STORM technique using DiI-bound molecules acquired a superresolution image in 15 s (Shim et al., 2012).

Next, Shim et al. used cationic rosamine (MitoTracker Orange and MitoTracker Red) and carbocyanine (MitoTracker Deep Red) fluorescent molecules with the STORM technique to image the dynamics of mitochondria. These fluorescent molecules readily label mitochondria via incubation with fluorescent molecules. This is a very simple labeling technique compared with the typical labeling procedure used in STORM microscopy. A good example of this is given in a booklet titled *Super Resolution Microscope N-STORM Protocol-Sample Preparation* (Nikon Corporation, 2013).

When MitoTracker Orange and MitoTracker Red were illuminated with 561-nm light they emitted photons (fluorescence) and then switched to the OFF state. Similarly, the molecule MitoTracker Deep Red was excited with 657-nm light and showed fluorescence before the molecules switched to the OFF state. Light of 405 nm photoswitched these molecules from the OFF state to the ON state. These molecules were found to be useful labels for mitochondria in live cells using the STORM technique. The authors achieved a lateral resolution of 30 nm within a few seconds.

It is important to point out that these membrane probes functioned without the requirement for an activator molecule (as was the case in previous STORM studies). Furthermore, these membrane probes did not require the addition of thiol chemicals such as β-mercaptoethylamine (BME) or mercaptoethylamine (MEA) for

the photoswitching to operate. The studies reported in Shim et al. (2012) used a standard oxygen-depleting system in the buffer (Nikon Corporation, 2013, 2015). The authors noted that in the absence of an oxygen-depleting system in the buffer the membrane probes still operate as photoswitchable fluorescent molecules, although the rate of photobleaching was increased.

Increased rates of photobleaching of fluorescent molecules have important consequences in the STORM technique. First, photobleaching (an irreversible photophysical process) a fluorescent molecule that has a reporter with a given brightness (number of photons emitted per second), will shorten the time the fluorescent molecule emits photons; therefore, the number of collected photons will be reduced. This effect will serve to reduce the precision of localization of fluorescent molecules, and therefore reduce the resolution that can be achieved with the STORM technique.

In the usual operation of the STORM technique, as is the case in other super-resolution localization techniques, the problems of specimen drift or motion must be corrected for. When the STORM technique is applied to biological processes the image acquisition rates must be sufficiently rapid to acquire accurate images of moving molecules. Shim et al. detected images of single fluorescent molecules using an EMCCD camera (iXon 860; Andor technology) with frame rates in the range of 500–900 Hz. However, with the small, membrane fluorescent molecules discussed by Shim et al. there is the additional problem that they can diffuse on the membranes where they are located.

I now switch the discussion to the topic of buffers and reagents used in the STORM technique. Although detector technology has rapidly improved with the availability of EMCCD and sCMOS cameras that have a high frame rate and significant progress has been made in the development and testing of photo-switchable fluorescent molecules, research into the development of buffers required for the STORM technique has shown less progress. The buffers used with photo-switchable fluorescent pairs of molecules typically contain β-mercaptoethylamine (BME) or mercaptoethylamine (MEA). In addition, the buffers contain glucose, glucose oxidase, and catalase, components of the oxygen-depleting system (Nikon Corporation, 2013, 2015). Different bipartite photoswitchable fluorescent pairs of molecules have different spectral properties in the presence of various buffers and different buffer concentrations (Dempsey et al., 2011). Vaughan et al., (2013) represented another important advance. The authors showed that the molecule phosphine tris(2-carboxyethyl)phosphine (TCEP) was capable of reversibly quenching the molecule Cy5. They showed that illuminating the Cy5 molecule with ultraviolet light switched it from the quenched, nonfluorescent OFF state to the fluorescent ON state; this photoswitchable switch in the presence of TCEP can be used in a superresolution STORM microscope (Vaughan et al., 2013).

Next, I return to problems associated with buffers, critical components of the STORM technique. While the research on STORM has concentrated on the development of new types of fluorescent molecules that can operate as photo-switchable molecules and on using the STORM technique to investigate a variety of

significant biological problems, significantly less attention has been given to the search for new types of buffers. I discuss two publications that appeared in 2013 from the same group at the Swiss Federal Institute of Technology (EPFL). These publications address problems relating to buffers in STORM microscopy.

The problem addressed in a paper titled "Simple buffers for 3D STORM microscopy" (Olivier et al., 2013a) can be found in a booklet titled *Super Resolution Microscope N-STORM Protocol-Sample Preparation* (Nikon Corporation, 2013). The booklet discusses complex specimen preparation associated with labeling and complex buffer systems (one of the problems associated with the STORM technique). Olivier et al. investigated three-dimensional STORM microscopy in which a simple and reproducible buffer was used. The authors sought a reproducible buffer system that would enable the performance of various STORM instruments to be compared as well as the algorithms used to process the images.

The solution the authors came up with involved Vectashield, a commercially available mounting medium. They found it was highly compatible with the fluorescent molecules Alexa-647 used in the STORM technique. Moreover, the three-dimensional superresolution STORM images obtained with this buffer were equivalent or superior to those obtained with the more complex traditional buffers used in the STORM protocol (Olivier, Keller, Rajan, Gönczy, and Manley, 2013a).

The standard buffer system used in the STORM technique contains a variety of chemicals. The oxygen-depleting system alone consists of glucose, glucose oxidase, and catalase and in addition there are thiol chemicals such as β-mercaptoethylamine (BME) or mercaptoethylamine (MEA). BME and MEA are highly toxic and special caution should be used with them in the laboratory. The reason an oxygen-depleting system is present in the buffer is to reduce the rate of photobleaching of fluorescent molecules. The disadvantage of a standard oxygen-depleting system is that over time it acidifies the buffer thus affecting the photophysical properties of the fluorescent molecules used to label the specimen in the STORM technique. To prevent this alteration in buffer acidity it is usually required to prepare the buffers just before each imaging session.

The alternative that Olivier et al. suggested was Vectashield, an anti-bleaching, mounting medium. They demonstrated its use with the STORM technique. The interaction of Vectashield with fluorescent molecules varies from fluorophore to fluorophore. Vectashield is not compatible with the fluorescent molecule Cy2, but is compatible with the common fluorescent molecule Alexa-647 when used in the STORM technique. The enhanced image quality obtained with Alexa-647 in the STORM technique is due to the medium's index-matching qualities (Olivier et al., 2013a).

Next, I discuss a paper in *PLOS One* titled "Resolution doubling in 3D-STORM imaging through improved buffers" (Olivier et al., 2013b). These investigators confronted the following problem. The STORM technique is based on photoswitching of fluorescent molecules between the fluorescent or ON state and the

nonfluorescent or OFF state. The buffer used in the STORM technique contains various chemicals that facilitate photoswitching between these states. The authors posited that the composition of buffers used with the STORM technique limits the resolution achievable. The specific problem the authors investigated concerned finding a way to optimize the chemical composition of the STORM buffer system used with the photoswitchable fluorescent molecule Alexa 647 for the number of photons emitted per switching cycle.

The solution Olivier et al. came up with involved improving the buffer systems. They labeled the specimen with the photoswitchable fluorescent molecule Alexa 647 and optimized the number of photons emitted per switching cycle with improved buffer systems. As a consequence they doubled the resolution of 3D STORM. They managed to achieve superresolution images with a lateral resolution of 10 nm and an axial resolution of 30 nm (Olivier et al., 2013b).

Olivier et al. achieved these impressive results because they were not only familiar with the technique of dual opposing microscope objectives that double the amount of photons collected, but also with an alternative method to increase localization precision and therefore resolution of the STORM technique. If the number of photons emitted from a fluorescent molecule could be increased that would also increase localization precision and therefore resolution of the STORM technique. It was the latter method the authors chose to investigate. Standard buffers used with the STORM technique include an oxygen-depleting system together with a thiol-reducing chemical such as mercaptoethylamine (MEA), mercaptoethanol (BME), or tris(2-carboxyethyl)phosphine (TCEP). The authors changed the chemical composition of the STORM buffer to optimize the photoswitching of the fluorescent molecule and therefore the resolution that can be achieved. They were able to increase the resolution of the STORM technique without modifying the optical system of the STORM microscope.

Olivier et al. used cyclooctatetraene (COT), a polyunsaturated hydrocarbon, as the new STORM buffer together with the photoswitchable fluorescent molecule Alexa 647 (Olivier et al., 2013b). The IUPAC name of the molecule is cycloocta-1,3,5,7-tetraene. Common names for this molecule are COT or annulene. An earlier publication demonstrated that the fluorescent molecule Alexa 647 shows minimal photobleaching in the presence of COT (Dave et al., 2009). Olivier et al. investigated this effect using the STORM technique with the goal of maximizing the achievable resolution.

The photoswitchable fluorescent molecule Alexa 647 can operate in blinking mode (Zhuang, 2009). Olivier et al. found that there was indeed minimal photobleaching of Alexa 647 in the presence of the new COT buffer. This blinking or spontaneous photoswitching together with minimal photobleaching achieved high localization of Alexa 647 and therefore enhanced two-dimensional and three-dimensional resolution of immunolabeled microtubules using the STORM technique (Olivier et al., 2013b).

15.9 Direct Stochastic Optical Reconstruction Microscopy

This section will introduce and evaluate the benefits and limitations of a new variant of the STORM technique dubbed direct stochastic optical reconstruction microscopy (direct STORM or dSTORM) by Heilemann et al., (2008). As usual, I frame the discussion in terms of the problem, the solution, the achievements, and the limitations and trade-offs.

First, I present the problem. A well-thought-out and clearly expressed statement of a research problem is the first step of an investigation and a prerequisite for a research grant. The problem concerns forming a photoswitchable fluorescent molecule from a single fluorescent molecule that is smaller than the typical bipartite or "tandem" photoswitches commonly used in the STORM technique.

Let us place the posed problem into context by reviewing the typical bipartite or tandem pairs used in the STORM technique (Nikon Corporation, 2013, 2015). The optical switch is formed from a bipartite pair of fluorescent molecules. As pointed out earlier, one fluorescent molecule is called the "activator" and the second fluorescent molecule is called the "reporter." For the optical switch to operate the activator molecule and reporter molecules must be in close proximity. This spatial constraint is achieved by binding one reporter molecule and several activator molecules to a secondary antibody. The secondary antibody is bound to a primary antibody. The combination of the two antibodies and the activator molecule and reporter molecules is finally bound to the target molecule of the specimen. Typical fluorescent molecules that work as the activator are Alexa Fluor 405, Cy2, and Cy3. The typical fluorescent molecule that works as the reporter is Alexa Fluor 647. A structural description of the typical bipartite photoswitch used in the STORM technique shows that the photoswitch is a large structure that has difficulty permeating cells and whose large spatial extent presents a limitation to the resolution that can be achieved with the STORM technique.

Next, I discuss the solution. What motivated the invention of dSTORM was a 2005 publication that demonstrated the capability of a reversible optical switch based on a single fluorescent molecule: carbocyanine (Heilemann et al., 2005). It is also of historical interest that this publication was cited in the first STORM publication (Rust et al., 2006). Heilemann et al. (2005) focused on single-molecule studies. They demonstrated that Cy5 fluorescent molecules can be reversibly optically switched from the fluorescent ON state to the nonfluorescent OFF state by alternate illumination of 633-nm light and 488-nm light. Operation of this Cy5 single-molecule switch is dependent on the aqueous solution containing an oxygen-depleting system and the triplet quencher β-mercaptoethylamine (BME) (Heilemann et al., 2005).

Using Heilemann et al. (2005) as a foundation, Heilemann et al., (2008) went on to develop the technique of direct stochastic optical reconstruction microscopy. They called their technique direct STORM (dSTORM) because it uses single

fluorescent molecules (Cy5 or Alexa 647) to form the optical switch. The dSTORM technique is different from the STORM technique which uses two different fluorescent molecules (the activator and the reporter) both of which are attached to an antibody.

Heilemann et al. (2008) demonstrated that the dSTORM technique could bring about superresolution microscopy with an optical switch composed of a single fluorescent molecule. The authors used the dSTORM technique to image microtubules and actin filaments in fixed mammalian cell and achieved a lateral spatial resolution of 21 nm (Heilemann et al., 2008).

One year after Heilemann et al. (2008), Zhuang published a focus/commentary paper in *Nature Photonics* stating that STORM also works with a single fluorescent molecule in the absence of an activator (Zhuang, 2009). Zhuang showed that microtubules in a cell can be labeled immunologically with the fluorescent molecule Alexa 647 and imaged with the red laser linewidth at 657 nm. Alexa 647 was used in blinking mode with STORM. Red laser light of 657 nm excited Alexa 647 to emit photons, caused the fluorescent molecule to switch from the ON state to the OFF state, and switched it back again (from the OFF state to the ON state) (Zhuang, 2009).

How do the dSTORM technique and the standard STORM technique compare? What are the advantages and limitations of each technique? The dSTORM and STORM techniques share a number of basic concepts. First, they are based on photoswitchable fluorescent molecules that can switch between the fluorescent ON state and the nonfluorescent OFF state. Second, the fluorescent molecules are turned to the ON state in a random or stochastic manner. The fluorescent molecules are separated in time so that the PSF of a molecule in the ON state does not overlap in space with the PSF of another molecule. Third, single molecules are localized with high precision. Fourth, the final superresolution image is composed of the sum of all single-molecule localizations.

Despite the similarities there are noteworthy differences between these two techniques. The dSTORM technique has several advantages over the STORM technique. For example, there is no requirement for a second fluorescent molecule (the activator in the dSTORM technique). An activator is typically used in the STORM technique together with a second, different fluorescent molecule called the reporter.

Another significant advantage of the dSTORM technique is its use of commercially available standard photoswitchable fluorescent molecules. These simple and small molecules can be used to label the intracellular structures of cells. Examples are Cy5 or Alexa 647, antigen-binding fragments (Fab), proteins, and peptides (Heilemann et al., 2008). Furthermore, when attached to small peptide tags the smaller optical switches comprising photoswitchable single fluorescent molecules are readily applied in dSTORM to achieve superresolution images of the intracellular architecture. For multispectral or multicolor dSTORM various cyanine dyes with differing spectral properties can be used. Each fluorescent molecule used

in the dSTORM technique has optimal requirements for the composition and pH of the buffer system, as well as the concentration of the thiol molecule MEA. For example, Alexa 647 operates optimally in the dSTORM technique in the presence of an oxygen-depleting system; other fluorescent molecules are less sensitive to photobleaching and work subject to more relaxed requirements.

The STORM technique has a significant advantage over the dSTORM technique, one that must be noted and taken into consideration when selecting a superresolution microscopy technique. When the bipartite or tandem reversible photoswitch composed of the reporter molecule and the activator molecule is used in the STORM technique the power of the laser line at 514 nm (used to switch the fluorescent molecule from the OFF state to the ON state) is 200× lower in the presence of an activator fluorescent molecule (STORM technique) than in the dSTORM case with the absence of an activator; however, it still remains in the microwatt to milliwatt intensity range (Heilemann et al., 2008). It is important to note that while the original STORM publication used the lower power line of the 514-nm laser, the authors were not interested in the superresolution imaging of cells (Rust et al., 2006). Subsequent publications on STORM used intensities that are similar to those used in dSTORM.

van de Linde et al., (2011) is a highly recommended paper that discusses using the dSTORM technique with fixed and live cells. Additionally, I recommend reading van de Linde et al., (2012), an important review article that elucidates the photochemistry and spectroscopy of single-molecule fluorescent labels compatible with the dSTORM technique.

15.10 General Comments on Localization Microscopy with Active Control

The superresolution localization techniques discussed in this chapter can operate with live cells and fixed cells, but different localization techniques have different trade-offs. When imaging fixed cells the investigator must carefully understand the fixation process to mitigate artifacts and interpret the final superresolution image correctly. When a superresolution technique is applied to live cells with the aim of studying dynamic processes, minimizing the effects of nonphysiological conditions is critical (e.g., spatial and temporal resolution are important considerations) (van de Linde et al., 2012).

I have discussed the problem of global specimen movement during data acquisition that is applicable to all superresolution microscopy techniques. Global movement can be due to thermal gradients and vibrations. The use of gold nanoparticles as fiduciary markers can be analyzed to correct for global movement. An important consideration in superresolution microscopy of live cells is movement

of the cells themselves and movement of the surface and internal structures of these live cells. Such movement can occur during data acquisition and will affect the accuracy of the superresolution image.

Another very important research area is the development of new and improved photoswitchable fluorescent molecules and photoactivatable fluorescent proteins. Such fluorescent molecules constitute the photoswitches and should have the following properties: they should be small molecules that can readily permeate cells and bind to intracellular organelles and structures; they should be photoswitched by light in the red, far-red, or the infrared region of the spectrum since such light is less nontoxic to live cells than blue or ultraviolet light; they should be bright (i.e., they should emit a large number of photons before they either switch to the OFF state or are photobleached); they should be resistant to spontaneous photobleaching in oxygen-containing buffers; they should be able to operate in the absence of another, different activator fluorescent molecule; and they should operate effectively in a physiological buffer.

Next, I point out some of the significant problems and limitations common to the various types of microscopies. Superresolution instrument development and probe development are works in progress.

First, some general questions for the researcher. What are the aims of the microscopic investigation? In the early phases of the invention of a new type or variation of a superresolution microscopy technique the initial publications stressed the proof of principle using a specimen with a structure already known from electron microscopy.

The eventual success of a new type of microscopy technique depends on multiple factors: compatibility with fixed cells or live cell imaging; complexity or simplicity of instrumentation; availability of appropriate probes, fluorescent dyes, or fluorescent proteins; lateral and axial spatial resolution; temporal resolution; acquisition time for each sequential frame; requirements for special buffers that are nontoxic to live cells; effects of photobleaching; effects of phototoxicity; levels of background fluorescence; precision of localization; labeling density; density of localization in the final superresolution image; effects of thick specimens; possibility of deep tissue imaging; validation and interpretation of the final image; severity of stage drift; number of photons detected from a single molecule that is emitting; and limitations of noise and optical aberrations.

There are many other factors for the researcher to consider. How do we know how to validate or interpret the superresolution image? Other investigators have suggested using a variety of light intensities, a variety of fluorescent probes, and correlative microscopy. The effects of light on the specimen during acquisition must be experimentally investigated; specific cell types can vary in their sensitivities to light (e.g., neurons are highly sensitive to phototoxicity). Tests for phototoxicity on one cell type may not be transferable to other cell types.

The quality of the final image(s) depends not only on the spatial and temporal resolution achieved but also on the contrast between the background and the labeled specimen. Therefore, great care is required to minimize background fluorescence

that can come from autofluorescence, cell media and buffers, objective immersion fluid, and microscope slides and coverslips.

Another important factor to keep in mind is that these superresolution techniques image fluorescent molecules that are bound or attached to cellular and intracellular organelles and structures. Photoswitchable fluorescent molecules are usually bound to primary and secondary antibodies (STORM technique). This presents a potential problem. As the distance between the fluorescent molecule and the two antibodies increases so does ambiguity in the interpretation of superresolution images.

An alternative labeling technique to that of exogenous fluorescent molecules involves the use of fluorescent proteins that can be genetically fused to a specific protein. However, this technique is limited in that fusion proteins are usually overexpressed. The effects of fluorescent proteins on normal cell functions must be investigated. Caution and experimental controls are the marks of a careful researcher.

In this chapter I have discussed the use of single-molecule localization together with active controls, such as photoactivation and/or photoswitching, to achieve superresolution optical microscopy.

The same year that the first publications on PALM, FPALM, and STORM appeared (2006) a new approach to superresolution was published titled "Wide-field subdiffraction imaging by accumulated binding of diffusing probes" (Sharonov and Hochstrasser, 2006). I discuss their publication for two reasons: (1) it is an exemplar of how to write a scientific research paper that is logically developed, inclusive of the background to the invention, the limitations of the method, the advantages of the method over alternative techniques, an analysis of the theory, error analysis, and experimental validation of the technique; (2) the new technique used several aspects common to alternative methods and incorporates a new technique that is unique.

The problem the investigators sought to solve was how to achieve superresolution optical microscopy without the requirement of labeling the specimen with fluorescent molecules and without the need for photoswitchable fluorescent molecules? These two strict requirements are foundational to the techniques of PALM, FPALM, STORM, and dSTORM that I have discussed in this chapter.

The solution Sharonov and Hochstrasser came up with was a new technique for single-molecule localization that could be used to achieve widefield superresolution optical microscopy. Sharonov and Hochstrasser called their new technique points accumulation for imaging in nanoscale topography (PAINT).

How does PAINT work and achieve superresolution optical microscopy? Furthermore, how did the authors achieve single-molecule localization and solve the problem of not having to locate multiple fluorescent molecules within the same diffraction volume without the use of reversible photoswitching, the common solution of alternative superresolution techniques? First, it is important to state that Sharonov and Hochstrasser (2006) demonstrated their PAINT technique not with fixed or live cells, but with lipid bilayers and vesicles formed from a single layer of lipids. Why they used lipids will become immediately clear. The authors used the fluorescent molecule Nile Red that has useful properties. Nile Red is a bright

fluorescent organic molecule, it emits a large number of photons (i.e., it is a bright fluorophore), and it is resistant to photobleaching. When Nile Red is dissolved in aqueous solution it is nonfluorescent. When it is in a lipophobic environment, such as when it is bound to lipids either in the form of a lipid bilayer or large single lipid layer vesicles, Nile Red is fluorescent and with excitation of the correct wavelength it will fluoresce with the emission of photons.

The key to the PAINT technique involves the diffusion of Nile Red molecules from an aqueous solution to a lipid bilayer or single-layer lipid vesicle. The rate at which Nile Red molecules collide with lipid structures is dependent on the diffusion coefficient and the gradient of Nile Red molecules near the lipid surface. When a Nile Red molecule is on the surface of a lipid it binds to the lipid and fluoresces. Fluorescent emission ceases when the molecule is photobleached or becomes unbound from the lipid surface. By controlling the concentration of Nile Red in the aqueous solution the collision rate can be controlled.

Photoswitches are used in alternative superresolution localization techniques to prevent the overlapping of multiple PSFs from fluorescent molecules that are simultaneously emitting. However, instead of the photoswitching operation a very low level of emitting fluorescent molecules is achieved in the PAINT technique by having a very low fraction of fluorescent molecules bound to the specimen at any given time period.

To get only one Nile Red molecule emitting photons at any one observation time within a diffraction-limited volume it is necessary to restrict the number of bound Nile Red molecules on the lipid surface to less than one molecule per square micrometer (Sharonov and Hochstrasser, 2006). This condition is fulfilled with a concentration of Nile Red of 10^{-9} M and a laser power of 1 kW/cm^2 when the probability of detecting two fluorescent Nile Red molecules on a single large single-layer lipid vesicle has been calculated as very low (Thompson et al., 2012). In addition, the frame rate at which photons are detected from a single molecule of Nile Red is less than the period between collisions. If these restrictions are valid under experimental conditions, then single-molecule localization follows in exactly the same way as in alternative superresolution techniques. If each image is a PSF of a single molecule of Nile Red, then highly precise localization beyond the diffraction limit is achievable.

What resolution did the authors achieve with the PAINT technique? With a Nile Red molecule and a lipid bilayer or a single-layer large lipid vesicle as the specimen the PAINT technique achieved a spatial resolution of 25 nm (Sharonov and Hochstrasser, 2006).

What are the advantages and limitations of the PAINT technique (Thompson et al., 2012)? One advantage of PAINT is that the organic molecule Nile Red is both resistant to photobleaching and is a very bright fluorophore. The PAINT method differs from alternative methods by not being limited by photobleaching. To increase the total number of localized molecules of Nile Red all that is required is to have a longer time to permit more Nile Red molecules to collide with and bind to lipid structures. PAINT does not label the specimen with antibodies that can disturb and distort the physiological structures of the cell because of their large size.

Typically, the resolution achieved with alternative localization techniques is limited by the label density on the specimen, but this is not an issue with the PAINT technique.

There are several different types of limitations associated with the PAINT technique. First, there is very little specificity in the binding of Nile Red and lipid structures. Second, there is a background from the Nile Red molecules in an aqueous solution. Third, Nile Red molecules bound to the surface of lipid structures can undergo diffusion, which can confound the analysis and interpretation of superresolution images.

In 2010 the PAINT technique was further developed into a new technique that the authors called universal PAINT (uPAINT) (Giannone et al., 2010). The authors built upon the PAINT technique and achieved superresolution microscopy of endogenous proteins on live cells.

Finally, I discuss a very different, yet innovative technique that can be used to achieve superresolution microscopy. I present this technique as a challenge to students who are exploring the possibility of research in the field of optics for their undergraduate or graduate programs. All the techniques that I have discussed are indeed works in progress. I am certain that new optical techniques will continue to be invented and then further developed into new instruments, new probes, and new algorithms. All these tools will help researchers advance our knowledge and understanding of living systems, their structure, function, and dysfunction.

The first requirement for success is to obtain a broad interdisciplinary background: mathematics, physics, chemistry, biology, medicine, computer science, signal processing, biological engineering, quantum optics, molecular spectroscopy, and instrument design and construction.

The second requirement for success in these endeavors is to have comprehensive knowledge of the field of optics, as well as a broad and deep understanding of optics. Many of the advances in the field of optics had their origins in disparate fields of engineering. A classic example is the laser, which was based on Einstein's theoretical work positing the existence of stimulated emission. It was incorporated into the invention of the maser, which operated in the microwave region of the spectrum. Finally, it was developed in the optical region as the laser. Remember the early lasers were incorrectly called optical masers (Maiman, 2018).

Keeping these guidelines in mind I segue to an approach to superresolution based on the optical and statistical properties of quantum emitters, one that highlights the potential of exploiting the nonclassical or quantum properties of light for the development of new types of superresolution optical microscopes.

In 1995 a publication appeared that proposed a new technique to modify the detection of PSFs and therefore enhance resolution in an optical microscope (Hell et al., 1995). What was different in their proposal? The usual approach to superresolution involved modifying light on the illumination side of the optical microscope. What the authors proposed was totally different. They used the classical properties of light for the illumination and took advantage of the nonclassical or quantum properties of light emitted from the fluorescent molecules labeling the

specimen. The authors proposed that the detection of two photons originating at the same point of the focal plane would sharpen the PSF and result in enhanced spatial resolution.

In Schwartz et al., (2013) the authors were motivated in their quest for a new type of superresolution optical microscopy by the prescient publication of Hell et al., (1995), in which a technique to achieve superresolution optical microscopy using quantum emitters was proposed.

How does the superresolution microscope of Schwartz et al. work? The technique developed by Schwartz et al. achieves superresolution imaging by using a property of light that is not understood classically. Photon antibunching can be explained and understood by considering a quantized electromagnetic field. The quantum phenomenon is called photon antibunching (i.e., the quantum property of inhibiting the emission of multiple photons simultaneously). Photon antibunching was first observed in resonance fluorescence (Kimble et al., 1977). Photon antibunching can be observed in organic fluorescent molecules and in quantum dots. The authors use a digital photon-counting camera and measure second- and third-order correlations of the photons emitted; these correlations derive from quantum antibunching.

There are similarities and differences between the work of Schwartz et al. on photon antibunching and the 2009 work of Dertinger, Colyer, Iyer, Weiss, and Enderlein on superresolution optical fluctuation imaging (SOFI). Both techniques use fluorescent molecules and operate by detecting correlations in the photons emitted from the image plane of the optical microscope. Antibunching microscopy is characterized by sub-Poissonian statistics, whereas the SOFI technique is characterized by super-Poissonian statistics (Schwartz et al., 2013).

How can a technique based on antibunching of photons become an operational superresolution microscope? This became reality with the development of a fiber bundle camera (Israel et al., 2017). The microscope was a standard epi-fluorescence microscope equipped with a single-photon detector made from a fiber bundle. Each fiber, from a bundle of 15 fibers, represented a single pixel of the camera and photons were detected at the end of each fiber by a single-photon avalanche diode (SPAD). This camera is unique in being able to detect single photons, in exceeding the temporal resolution of 1 ms of standard cameras, and in having a fill factor greater than 80%.

What is the main advantage a superresolution microscopy technique based on antibunching of emitted photons has over alternative localization techniques?

With techniques such as PALM, STORM, and variants of STORM the problem of multiple fluorescent molecules being located within a given diffraction volume has the following solution. Temporal resolution using these localization techniques is given by the requirement for very sparse photoswitching between the ON and OFF states of fluorescent molecules. This yields one fluorescent molecule in a given diffraction-limited volume and therefore very precise single-molecule localization. However, the technique based on antibunching of emitted photons yields the number of excited emitters, making it possible to avoid calculations with

multifluorescent molecules such that single-molecule localization can be preserved. Hence, the main advantage of this superresolution optical microscopy technique is enhanced temporal resolution (Israel et al., 2017).

The question often arises as to which software package is best suited to analyzing the data acquired as a result of using the various localization microscopy techniques (Deschout et al., 2014; Sage et al., 2015). Each of the software packages described in these references is based on different algorithms and contains different assumptions. Commercial superresolution microscopes come with commercial software packages. The researchers of purpose-built superresolution microscopes may choose to code their own software. In either case it is crucial to know the assumptions made, the validity of, and limitations of the algorithms that are the foundation of single-molecule localization software.

References

Aspelmeier, T., Egner, A., and Munk, A. (2015). Modern statistical challenges in high-resolution fluorescence microscopy. Annual Review of Statistics and Its Application, **2**, 163–202.

Basché, T., Moerner, W. E., Orrit, M., and Wild, U. P., Editors, (1997). *Single-Molecule Optical Detection, Imaging and Spectroscopy*. Weinheim: VCH.

Bates, M., Blosser, T. R., and Zhuang, X. (2005). Short-range spectroscopic ruler based on a single-molecule optical switch. Physical Review Letters, **94**, 108101–108104.

Bates, M., Huang, B., Dempsey, G. T., and Zhuang, X. (2007). Multicolor super-resolution imaging with photoswitchable fluorescent probes. Science, **317**, 1749–1753.

Betzig, E. (1995). Proposed method for molecular optical imaging. Optics Letters, **20**, 237–239.

Betzig, E. (2005). Excitation strategies for optical lattice microscopy. Optics Express, **13**(8), 3021–3036.

Betzig, E. (2014a). *Single Molecules, Cells, and Super-Resolution Optics*: Lecture Slides. https://www.nobelprize.org/uploads/2018/06/betzig-lecture-slides.pdf (Accessed April 6, 2019).

Betzig, E. (2014b). *Single Molecules, Cells, and Super-Resolution Optics*: Nobel Lecture. https://www.nobelprize.org/prizes/chemistry/2014/betzig/lecture/ (Accessed April 6, 2019).

Betzig, E., and Chichester, R. J. (1993). Single molecules observed by near-field scanning optical microscopy. Science, **262**, 1422–1425.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science, **313**, 1642–1645.

Bhushan, B., Wyant, J. C., and Koliopoulos, C. L. (1985). Measurement of surface topography of magnetic tapes by Mirau interferometry. Applied Optics, **24**, 1489–1497.

Bock, H., Geisler, C., Wurm, C. A., Von Middendorff, C., Jakobs, S., Schönle, A., Egner, A., Hell, S. W., and Eggeling, C. (2007). Two-color far-field fluorescence nanoscopy based on photoswitchable emitters. Applied Physics B, **88**, 161–165.

Braun, D., and Fromherz, P. (1998). Fluorescence interferometry of neuronal cell adhesion on microstructured silicon. Physical Review Letters, **81**, 5241–5244.

Brown, T. A., Tkachuk, A. N., Shtengel, G., Kopek, B. G., Bogenhagen, D. F., Hess, H. F., and Clayton, D. A. (2011). Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction. Molecular and Cellular Biology, **31**, 4994–5010.

Burns, D. H., Callis, J. B., Christian, G. D., and Davidson, E. R. (1985). Strategies for attaining superresolution using spectroscopic data as constraints. Applied Optics, **24**, 154–161.

Chattoraj, M., King, B. A., Bublitz, G. U., and Boxer, S. G. (1996). Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. Proceedings of the National Academy of Sciences of the United States of America, **93**, 8362–8367.

Dave, R., Terry, D. S., Munro, J. B., and Blanchard, S. C. (2009). Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging. Biophysical Journal, **96**, 2371–2381.

Delaunay, G. (1953). Microscope interférentiel A. Mirau pour la mesure du fini des surfaces. Rev. Opt. Theor. Instrum. **32**, 610–614.

Dempsey, G. T., Bates, M., Kowtoniuk, W. E., Liu, D. R., Tsien, R. Y., and Zhuang, X. (2009). Photoswitching mechanism of cyanine dyes. Journal of the American Chemical Society, **131**, 18192–18193.

Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M., and Zhuang, X. (2011). Evaluation of fluorophores for optimal localization-based super-resolution imaging. Nature Methods, **8**, 1027–1036.

Dertinger, T., Colyer, R., Iyer, G., Weiss, S., and Enderlein, J. (2009). Fast, background-free, three-dimensional super-resolution optical fluctuation imaging (SOFI). Proceedings of the National Academy of Sciences of the United States of America, **106**, 22287–22292.

Dertinger, T., Colyer, R., Vogel, R., Enderlein, J., and Weiss, S. (2010a). Achieving increased resolution and more pixels with superresolution optical fluctuation imaging (SOFI). Optics Express, **18**, 18875–18885.

Dertinger, T., Heilemann, M., Vogel, R., Sauer, M., and Weiss, S. (2010b). Superresolution optical fluctuation imaging with organic dyes. Angewandte Chemie International Edition, **49**, 9441–9443.

Deschout, H., Zanacchi, F. C., Mlodzianoski, M., Diaspro, A., Bewersdorf, J., Hess, S. T., and Braeckmans, K. (2014). Precisely and accurately localizing single emitters in fluorescence microscopy. Nature Methods **11**, 253–266.

Dickson, R. M., Cubitt, A. B., Tsien, R. Y., and Moerner, W. E. (1997). On/off blinking and switching behavior of single molecules of green fluorescent protein. Nature, **388**, 355–358.

Dogan, M., Yalçin, A., Jain, S., Goldberg, M. B., Swan, A. K., Selim Ünlü, M., and Goldberg, B. B. (2008). Spectral self-interference fluorescence microscopy for subcellular imaging. IEEE Journal of Selected Topics in Quantum Electronics, **14**, 217–225.

Eggeling, C., Willig, K. I., Sahl, S. J., and Hell, S. W. (2015). Lens-based fluorescence nanoscopy. Quarterly Reviews of Biophysics, **48**, 178–243.

Egner, A., Geisler, C., von Middendorff, C., Bock, H., Wenzel, D., Medda, R., Andresen, M., Stiel, A. C., Jakobs, S., Eggeling, C., Schönle, A., and Hell, S. W. (2007). Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. Biophysical Journal, **93**, 3285–3290.

Fernández-Suárez, M., and Ting, A. Y. (2008). Fluorescent probes for super-resolution imaging in living cells. Nature Reviews Molecular Cell Biology, **9**, 929–943.

Geisler, C., Schonle, A., von Middendorff, C., Bock, H., Eggeling, C., Egner, A. and Hell, S. W. (2007). Resolution of $\lambda/10$ in fluorescence microscopy using fast single molecule photo-switching. Applied Physics A, **88**, 223–226.

Gelles, J., Schnapp, B. J., and Sheetz, M. P. (1988). Tracking kinesin-driven movements with nanometer-scale precision. Nature, **331**, 450–453.

Giannone, G., Hosy, E., Levet, F., Constals, A., Schulze, K., Sobolevsky, A. I., Rosconi, M. P., Gouaux, E., Tampé, R., Choquet, D., and Cognet, L. (2010). Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. Biophysical Journal, **99**, 1303–1310.

Goldman, R. D., and Spector, D. L., Eds. (2005). *Live Cell Imaging, A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Gonzalez, R. C., and Woods, R. E. (2008). *Digital Image Processing,* Third Edition. Upper Saddle River, NJ: Pearson, Prentice Hall.

Goodman, J. W. (2017). *Introduction to Fourier Optics*. Fourth Edition. New York: W.H. Freeman and Company. Chapter 8. Point-Spread Function and Transfer Function Engineering, pp. 231–267.

Gordon, M. P., Ha, T., and Selvin, P. R. (2004). Single-molecule high-resolution imaging with photobleaching. Proceedings of the National Academy of Sciences of the United States of America, **101**, 6462–6465.

Gould, T. J., Verkhusha, V. V., and Hess, S. T. (2009). Imaging biological structures with fluorescence photoactivation localization microscopy. Nature Protocols, **4**(3), 291–308.

Ha, T., and Tinnefeld, P. (2012). Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. Annual Review of Physical Chemistry, **63**, 595–617.

Hariharan, P. (2003). *Optical Interferometry, Second Edition*. London: Academic Press.

Heilemann, M., Herten, D. P., Heintzmann, R., Cremer, C., Müller, C., Tinnefeld, P., Weston, K. D., Wolfrum, J., and Sauer, M. (2002). High-resolution colocalization of single dye molecules by fluorescence lifetime imaging microscopy. Analytical Chemistry, **74**, 3511–25177.

Heilemann, M., Margeat, E., Kasper, R., Sauer, M., and Tinnefeld, P. (2005). Carbocyanine dyes as efficient reversible single-molecule optical switch. Journal of the American Chemical Society, **127**, 3801–3806.

Heilemann, M., van de Linde, S., Schüttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., and Markus Sauer, M. (2008). Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes. Angewandte Chemie International Edition in English, **47**, 6172–6176.

Hell, S. W., Soukka, J., and Hänninen, P. (1995). Two- and multiphoton detection as an imaging mode and means of increasing the resolution in far-field light microscopy: A study based on photon-optics. Bioimaging, **3**, 64–69.

Hess, H. F., Betzig, E., Harris, T. D., Pfeiffer, L. N., and West, K. W. (1994). Near-field spectroscopy of the quantum constituents of a luminescent system. Science, **264**, 1740–1745.

Hess, S. T., Girirajan, T. P. K., and Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophysical Journal, **91**, 4258–4272.

Hess, S. T., Gould, T. J., Gudheti, M. V., Maas, S. A., Mills, K. D., and Zimmerberg, J. (2007). Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. Proceedings of the National Academy of Sciences of the United States of America, **104**, 17370–17375.

Hirschfield, T. (1976). Optical microscopic observation of single small molecules. Applied Optics, **15**, 2965–2966.

Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008a). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. Science, **319**, 810–813.

Israel, Y., Tenne, R. Oron, D., and Silberberg, Y. (2017). Quantum correlation enhanced super-resolution localization microscopy enabled by a fibre bundle camera. Nature Communications, **8**, 14786. https://doi.org/10.1038/ncomms14786.

Jones, S. A., Shim, S.-H., He, J., and Zhuang, X. (2011). Fast, three-dimensional super-resolution imaging of live cells. Nature Methods, **8**, 499–505.

Juette, M. F., Gould, T. J., Lessard, M. D., Mlodzianoski, M. J., Nagpure, B. S., Bennett, B. T., Hess, S. T., and Bewersdorf, J. (2008). Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. Nature Methods, **5**, 527–529.

Kao, H. P., and Verkman, A. S. (1994). Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position. Biophysical Journal, **67**, 1291–1300.

Kimble, H. J., Dagenais, M., and Mandel, L. (1977). Photon antibunching in resonance fluorescence. Physical Review Letters, **39**, 691–695.

Lidke, K. A., Rieger, B., Jovin, T. M., and Heintzmann, R. (2005). Superresolution by localization of quantum dots using blinking statistics. Optics Express, **13**, 7052–7062.

Lukyanov, K. A., Chudakov, D. M., Lukyanov, S., and Verkhusha, V. V. (2005). Photoactivatable fluorescent proteins. Nature Reviews Molecular Cell Biology, **6**, 885–890.

Maiman, T. H. (2018). *The Laser Inventor, Memoirs of Theodore H. Maiman.* New York: Springer Nature.

Malacara, D. (2007). *Optical Shop Testing,* Third Edition. Hoboken: John Wiley & Sons.

Masters, B. R. (2010). The Development of Fluorescence Microscopy, in: *Encyclopedia of Life Sciences (ELS)*, John Wiley and Sons, Ltd: Chichester, UK, https://doi.org/10.1002/9780470015902.a0022093.

Masters, B. R. (2014). Paths to Förster's resonance energy transfer (FRET) theory. European Physical Journal H, **39**, 87–139.

Masters, B. R. (2016). The Importance of Responsible Conduct of Research (RCR), Responsible Conduct of Research, pdf. MIT CISB, Center for Integrative Synthetic Biology, SBC e-Newsletter 2015 (1), (2). URL: https://cisb.mit.edu/wp-content/uploads/2014/08/2016MIT-RCR-byMasters.pdf. Accessed April 6, 2019.

McKinney, S. A., Murphy, C. S., Hazelwood, K. L., Davidson, M. W., and Looger, L. L. (2009). A bright and photostable photoconvertible fluorescent protein. Nature Methods, **6**, 131–133.

Moerner, W. E. (2007). New directions in single-molecule imaging and analysis. Proceedings of the National Academy of Sciences of the United States of America, **104**, 12596–12602.

Moerner, W. E. (2012). Microscopy beyond the diffraction limit using actively controlled single molecules. Journal of Microscopy, **246**, 213–220.

Moerner, W. E. (2014a). *Single-Molecule Spectroscopy, Imaging, and Photocontrol: Foundations for Super-Resolution Microscopy—Lecture slides.* https://www.nobelprize.org/uploads/2018/06/moerner-lecture-slides.pdf. Accessed April 6, 2019.

Moerner, W. E. (2014b). *Single-Molecule Spectroscopy, Imaging, and Photocontrol: Foundations for Super-Resolution Microscopy—Nobel Lecture.* https://www.nobelprize.org/prizes/chemistry/2014/moerner/lecture/. Accessed April 6, 2019.

Moerner, W. E., and Kador, L. (1989). Optical detection and spectroscopy of single molecules in a solid. Physical Review Letters, **62**, 2535–2538.

Moerner, W. E., and Orrit, M. (1999). Illuminating single molecules in condensed matter. Science, **283**, 1670–1676.

Mortensen, K. I., Churchman, L. S., Spudich, J. A., and Flyvbjerg H. (2010). Optimized localization analysis for single-molecule tracking and super-resolution microscopy. Nature Methods, **7**, 377–381.

Nikon Corporation. (2013). *Super Resolution Microscope N-STORM Protocol-Sample Preparation.*

Nikon Corporation. (2015). *Super Resolution Microscope N-SIM / N-STORM.*

Nyquist, H. (1928a). Certain topics in telegraph transmission theory. Proceedings of the IEEE, **90**, 280–305.

Nyquist, H. (1928b). Thermal agitation of electric charge in conductors. Physical Review, **32**, 110–113.

Olivier, N. Keller, D., Gönczy, P., and Manley, S. (2013b). Resolution doubling in three-dimensional-STORM imaging through improved buffers. PLoS ONE, **8**, e69004.

Olivier, N., Keller, D., Rajan, V. S., Gönczy, P., and Manley, S. (2013a). Simple buffers for three-dimensional STORM microscopy. Biomedical Optics Express, **4**, 885–889.

Orrit, M., and Bernard, J. (1990). Single pentacene molecules detected by fluorescence excitation in a paraterphenyl crystal. Physical Review Letters, **65**, 2716–2719.

Pavani, S. R. P., Thompson, M. A., Biteen, J. S., Lord, S. J., Liu, N., Twieg, R. J., Piestun, R., and Moerner, W. E. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proceedings of the National Academy of Sciences of the United States of America, **106**, 2995–2999.

Patterson, G., Davidson, M., Manley, S., and Lippincott-Schwartz, J. (2010). Superresolution imaging using single-molecule localization. Annual Review of Physical Chemistry, **61**, 345–367.

Patterson, G., and Lippincott-Schwartz, J. (2002). A photoactivatable GFP for selective photolabeling of proteins and cells. Science, **297**, 1873–1877.

Prabhat, P. Ram, S., Ward, E. S., and Ober, R. J. (2004). Simultaneous imaging of different focal planes in fluorescence microscopy for the study of cellular dynamics in three dimensions. IEEE Transactions on Nanobioscience, **3**, 237–242 (2004).

Qu, X., Wu, D., Mets, L., and Scherer, N. F. (2004). Nanometer-localized multiple single-molecule fluorescence microscopy. Proceedings of the National Academy of Sciences of the United States of America, **101**, 11298–11303.

Ram, S., Kim, D., Ward, E. S., and Ober, R. J. (2012). three-dimensional single molecule tracking with multifocal plane microscopy reveals rapid intercellular transferrin transport at epithelial cell barriers. Biophysical Journal, **103**, 1594–1603.

Rust, M. J., Bates, M., and Zhuang, X. W. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature Methods, **3**, 793–795. Published online August 9, 2006.

Sage, D., Kirshner, H., Pengo, T., Stuurman, N., Min, J., Manley, S., and Unser, M. (2015). Quantitative evaluation of software packages for single-molecule localization microscopy. Nature Methods, **12**, 717–724.

Schechner, Y. Y., Piestun, R., and Shamir, J. (1996). Wave propagation and rotating intensity distributions. Physical Review E, **54**, R50–R53.

Schwartz, O., Levitt, J. M., Tenne, R., Itzhakov, S., Deutsch, Z., and Oron, D. (2013). Superresolution microscopy with quantum emitters. Nano Letters, **13**, 5832–5836. https://doi.org/10.1021/nl402552m

Shannon, C. E. (1949). Communication in the presence of noise. Proceedings of the IEEE, **37**, 10–21.

Sharonov, A., and Hochstrasser, R. M. (2006). Wide-field subdiffraction imaging by accumulated binding of diffusing probes. Proceedings of the National Academy of Sciences of the United States of America, **103**, 18911–18916.

Shcherbakova, D. M., Sengupta, P., Lippincott-Schwartz, J., and Verkhusha, V. V. (2014). Photocontrollable fluorescent proteins for superresolution imaging. Annual Review of Biophysics, **43**, 303–329.

Shera, E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A., and Soper, S. A. (1990). Detection of single fluorescent molecules. Chemical Physics Letters, **174**, 553–557.

Shim, S.-H., Xia, C., Zhong, G., Babcock, H. P., Vaughan, J. C., Huang, B., Wang, X., Xu, C., Bi, G.-Q., and Zhuang, X. (2012). Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes. Proceedings of the National Academy of Sciences of the United States of America, **109**, 13978–13983.

Shroff, H., Galbraith, C. G., Galbraith, J. A., and Betzig, E. (2008). Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nature Methods **5**, 417–423.

Shroff, H., Galbraith, C. G., Galbraith, J. A., White, H., Gillette, J., Olenych, S., Davidson, M. W., and Betzig, E. (2007). Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. Proceedings of the National Academy of Sciences of the United States of America, **104**, 20308–20313.

Shtengel, G., Galbraith, J. A., Galbraith, C. G., Lippincott-Schwartz, J., Gillette, J. M., Manley, S., Sougrat, R., Waterman, C. M., Kanchanawong, P., Davidson, M. W., Fetter, R. D., and Hess, H. F. (2009). Interferometric fluorescent super-resolution microscopy resolves three-dimensional cellular ultrastructure. Proceedings of the National Academy of Sciences of the United States of America, **106**, 3125–3130.

Shtengel, G., Wang, Y., Zhang, Z., Goh, W. I., Hess, H. F., and Kanchanawong, P. (2014). Imaging cellular ultrastructure by PALM, iPALM, and correlative iPALM-EM. Methods in Cell Biology, **123**, 273–294.

Small, A. R., and Parthasarathy, R. (2014). Superresolution localization methods. Annual Review of Physical Chemistry, **65**, 107–125.

Thompson, M. A., Lew, M. D., and Moerner, W. E. (2012). Extending microscopic resolution with single-molecule imaging and active control. Annual Review of Biophysics, **41**, 321–342.

Thompson, R. E., Larson, D. R., and Webb, W. W. (2002). Precise nanometer localization analysis for individual fluorescent probes. Biophysical Journal, **82**, 2775–2783.

van de Linde, S., Heilemann, M., and Sauer, M. (2012). Live-cell super-resolution imaging with synthetic fluorophores. Annual Review of Physical Chemistry, **63**, 519–540.
van de Linde, S., Löschberger, A., Klein, T., Heidbreder, M., Wolter, S., Heilemann, M., and Sauer, M. (2011). Direct stochastic optical reconstruction microscopy with standard fluorescent probes. Nature Protocols, **6**, 991–1009.
van Oijen, A. M., Köhler, J., Schmidt, J., Müller, M., and Brakenhoff, G. J. (1999). Far-field fluorescence microscopy beyond the diffraction limit. JOSA A, **16**, 909–915.
Vaughan, J. C., Dempsey, G. T., Sun, E., and Zhuang, X. (2013). Phosphine quenching of cyanine dyes as a versatile tool for fluorescence microscopy. Journal of the American Chemical Society, **134**, 1197–1200.
Weiss, S. (1999). Fluorescence spectroscopy of single biomolecules. Science, **283**, 1676–1683.
Xu, K., Babcock, H. P., and Zhuang, X. (2012). Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. Nature Methods, **9**, 185–188.
Yao, Z., and Carballido-López, R. (2014). Fluorescence imaging for bacterial cell biology: from localization to dynamics, from ensembles to single molecules. Annual Review of Microbiology, **68**, 459–476.
Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. Science, **300**, 2061–2065.
Yokoe, H., and Meyer, T. (1996). Spatial dynamics of GFP-tagged proteins investigated by local fluorescence enhancement. Nature Biotechnology, **14**, 1252–1256.
Zhuang, X., (2009). Nano-imaging with STORM. Nature Photonics, **3**, 365–367.

Further Reading

Andresen, M., Stiel, A. C., Fölling, J., Wenzel, D., Schönle, A., Egner, A., Eggeling, C., Hell, S. W., and Jakobs, S. (2008). Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. Nature Biotechnology, **26**, 1035–1040.
Bachl, A., and Lukosz, W. (1967). Experiments on superresolution imaging of a reduced object field. Journal of the Optical Society of America, **57**, 163–169.
Bates, M., Huang, B., and Zhuang, X. (2008). Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes. Current Opinion in Chemical Biology, **12**, 505–514.
Betzig, E. (2014c). Biographical sketch. https://www.nobelprize.org/prizes/chemistry/2014/betzig/biographical/ (Accessed April 6, 2019).
Betzig, E., Trautman, J. K., Harris, T. D., Weiner, J. S., and Kostelak, R. L. (1991). Breaking the diffraction barrier: optical microscopy on a nanometric scale. Science, **251**, 1468–1470.
Biteen, J. S., Thompson, M. A., Tselentis, N. K., Bowman, G. R., Shapiro, L., and Moerner, W. E. (2008). Super-resolution imaging in live *Caulobacter crescentus* cells using photoswitchable EYFP. Nature Methods, **5**, 947–949.
Cremer, C., and Masters, B. R. (2013). Resolution enhancement techniques in microscopy. European Physical Journal H, **38**, 281–344. Open Access.
Dani, A., Huang, B., Bergan, J., Dulac, C., and Zhuang, X. (2010). Superresolution imaging of chemical synapses in the brain. Neuron, **68**, 843–856.
Endesfelder, U., and Heilemann, M. (2014). Direct Stochastic Optical Recnstruction Microscopy (dSTORM). In: *Advanced Fluorescence Microscopy: Methods and Protocols*, Peter J. Verveer (ed.), volume **1251**, 263–276. New York: Springer.
Fölling, J., Belov, V., Kunetsky, R., Medda, R., Schönle, A., Egner, A., Eggeling, C., Bossi, M., and Hell, S. W. (2007). Photochromic rhodamines provide nanoscopy with optical sectioning. Angewandte Chemie International Edition, **46**, 6266–6270.
Huang, B., Babcock, H., and Zhuang, X. (2010). Breaking the diffraction barrier: Super-resolution imaging of cells. Cell, **143**, 1047–1058.

Huang, B., Bates, M., and Zhuang, X. (2009). Super-resolution fluorescence microscopy. Annual Review of Biochemistry, **78**, 993–1016.

Huang, B., Jones, S. A., Brandenburg, B., and Zhuang, X. (2008b). Whole-cell three-dimensional STORM reveals interactions between cellular structures with nanometer-scale resolution. Nature Methods, **5**, 1047–1052.

Kanchanawong, P., Shtengel, G., Pasapera, A. M., Ramko, E. B., Davidson, M. W., Hess, H. F., and Waterman, C. M. (2010). Nanoscale architecture of integrin-based cell adhesions. Nature, **468**, 580–584.

Kopek, B. G., Shtengel, G., Xu, C. S., Clayton, D. A., and Hess, H. F. (2012). Correlative three-dimensional superresolution fluorescence and electron microscopy reveal the relationship of mitochondrial nucleoids to membranes. Proceedings of the National Academy of Sciences of the United States of America, **109**, 6136–6141.

Lippincott-Schwartz, J., and Patterson, G. H. (2009). Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging. Trends in Cell Biology, **19**, 555–565.

Masters, B. R. (2008). History of the Optical Microscope in Cell Biology and Medicine, in: *Encyclopedia of Life Sciences (ELS)*, John Wiley and Sons, Ltd: Chichester, UK, https://doi.org/10.1002/9780470015902.a0003082.

Masters, B. R. (2009a). Correlation of histology and linear and nonlinear microscopy of the living human cornea. Journal of Biophotonics, **2**, 127–139.

Masters, B. R. (2009b). History of the Electron Microscope in Cell Biology, in: *Encyclopedia of Life Sciences (ELS)*, Chichester, UK. John Wiley and Sons, Ltd: September 2009, https://doi.org/10.1002/9780470015902.a0021539.

Masters, B. R., and So, P. T. C. (2008). *Handbook of Biomedical Nonlinear Optical Microscopy*. New York: Oxford University Press.

Nieuwenhuizen, R. P. J., Lidke, K. A., Bates, M., Puig, D. L., Grünwald, D., Stallinga, S., and Rieger, B. (2013). Measuring image resolution in optical nanoscopy. Nature Methods, **10**, 557–562.

Ober, R. J., Ram, S., and Ward, E. S. (2004). Localization accuracy in single-molecule microscopy. Biophysical Journal, **86**, 1185–1200.

Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. Journal of Cell Biology, **190**, 165–175.

Sigal, Y. M., Speer, C. M., Babcock, H. P., and Zhuang, X. (2015). Mapping synaptic input fields of neurons with super-resolution imaging. Cell, **163**, 493–505.

Subach, F. V., Patterson, G. H., Manley, S., Gillette, J. M., Lippincott-Schwartz, J., and Verkhusha, V. V. (2009). Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. Nature Methods, **6**, 153–159.

Subach, F. V., Patterson, G. H., Renz, M., Lippincott-Schwartz, J., and Verkhusha, V. V. (2010). Bright monomeric photoactivatable red fluorescent protein for two-color super-resolution sptPALM of live cells. Journal of the American Chemical Society, **132**, 6481–6491.

Szymborska, A., de Marco, A., Daigle, N., Cordes, V. C., Briggs, J. A., and Ellenberg, J. (2013). Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. Science, **341**, 655–658.

van de Linde, S., Kasper, R., Heilemann, M., and Sauer, M. (2008). Photoswitching microscopy with standard fluorophores. Applied Physics B, **93**, 725–731.

van Engelenburg, S. B., Shtengel, G., Sengupta, P., Waki, K., Jarnik, M., Ablan, S. D., Freed, E. O., Hess, H. F., and Lippincott-Schwartz, J. (2014). Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. Science, **343**, 653–656.

Vaziri, A., Tang, J., Shroff, H., and Shank, C. V. (2008). Multilayer three-dimensional super resolution imaging of thick biological samples. Proceedings of the National Academy of Sciences of the United States of America, **105**, 20221–20226.

Vogelsang, J., Kasper, R., Steinhauer, C., Person, B., Heilemann, M., Sauer, M., and Tinnefeld, P. (2008). A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. Angewandte Chemie International Edition in English, **47**, 5465–5469.

Chapter 16
Coda: Trade-Offs, Cautions, and Limitations of Superresolution Optical Microscopes

16.1 Introduction

My aim in this chapter is to encourage critical skeptical thinking about superresolution microscopy. It is important to note that the precision and accuracy of single-molecule localization are different concepts. Typically, authors report the precision of localization of a single molecule. That is not equivalent to the accuracy of localization. The accuracy of localization is a measure of the fidelity with which single-molecule localization is identical to the actual location within the specimen. To assess the accuracy of superresolution single-molecule locations and therefore the final superresolution image we require calibration specimens whose structures are independently determined. An alternative is to use correlative microscopy (e.g., transmission electron microscopy) to validate the accuracy of information obtained from the superresolution image. Localization techniques achieve high precision in localization of a fluorescent molecule. The size of a fluorescent molecule affects the accuracy of the position of the structure to which the label is attached. This is of increasing importance as the size of the label, including its associated linker molecules, approaches the required resolution of the superresolution image.

Innovation in research and development is inhibited by aversion to risk taking. The origin of this can be found in our educational systems as well as societal norms, social pressure, and the cultures of corporations, scientific societies, publishers of journals, and especially research-funding agencies. This situation must change.

What are the trade-offs, cautions, and limitations in superresolution fluorescence microscopy? Today there are commercial superresolution microscopes based on the techniques I described in Part III. Each commercial vendor has a website—loaded with information, protocols, amazing images, fantastic resolution benchmarks, lists of publications in high-impact journals—stressing the advantages of their particular instrument. A very practical resource here is *Super-Resolution Microscopy: A Practical Guide* (Birk, 2017). The advantages of a commercial microscope include the vendor's maintenance and support for the microscope, the possibility of

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7_16

software upgrades and extensions, and in-house training that the vendor may supply. But buyer beware! Some of the benchmarks are achieved by individuals with expert skill sets and the typical researcher may not be able to replicate these published benchmarks. It is also important to point out that some of the published images that inspire awe are made with specimens that are essentially one-dimensional filaments (e.g., the cytoskeleton) or point-like objects such as vesicles. Commercial microscopes that are based on photoactivation or photoswitching and localization of fluorescent molecules have an imaging capability that is related to the shape and the extent or volume of the structure imaged (Schermelleh et al., 2010).

For those intrepid researchers with the requisite expertise in instrument design, development, and construction there is the attractive alternative option of building your own instrument. Journals, such as *Nature Protocols* and *Review of Scientific Instruments*, provide details of component sources, instrument construction, calibration, specimen preparation, and troubleshooting. There are several advantages to building your own instrument, especially when the superresolution technique can readily be adapted to a fluorescence microscope in your laboratory. Besides the learning process of constructing an instrument, which is a wonderful learning experience for students who are part of your research group, there is also the possibility to carefully select microscope components, such as the illumination system and the detection system, to optimize specific parameters.

Let us now turn to developing some of the trade-offs, cautions, and limitations of superresolution microscopes. First, some thoughts on the specimens or the objects of our investigations. Deep thought should be given to formulating questions relating to specimen investigation. It is necessary to know what spatial and temporal resolutions are required to answer the questions posed. Ask yourselves whether you truly require a superresolution microscope or whether a standard fluorescence microscope will suffice for the investigation.

The next question concerns the state of the specimen. Can we conduct our microscopic investigation on a fixed specimen? If that is the case, then how do we minimize the effects of tissue shrinkage and other artifacts due to the fixation process? What staining procedures, labeling with fluorescent probes, or the use of immunocytochemical staining is optimal to develop the contrast required in the fixed specimen. How can cells and tissue be made permeable to enhance the penetration of fluorescent tagged antibodies? For cases that require thin sections there is the option to mechanically section thick specimens.

If fixation of the specimen is not an option, then we are in the domain of live cell, live tissue, or live organism imaging. Of course, all specimens are different. They differ in size, thickness, absorption and scattering coefficients at different wavelengths, their photosensitivity, and their response to various components of the buffer solution. For live cell, tissue, and organ imaging we need to specify the temporal resolution of the microscope if we are to investigate dynamic processes. Are we imaging deep neurons in the brain of an animal over various time scales: minutes, hours, weeks, or months? Are we imaging human tissue such as in vivo cornea, lens, or retina? Are we imaging in vivo human skin? Each specimen will

have different levels of autofluorescence, absorbance and scattering coefficients, and photosensitivities. Further questions to be answered include the following: What is the required field of view? Are we in the domain of deep-tissue imaging (e.g., imaging neurons in the animal brain that are located 1 mm below the surface of the brain)? Or are we restricting our investigation to the imaging of structures that are located extremely close to the surface of cells (e.g., the use of TIRF microscopy to investigate the focal adhesion plaques of cells)?

Having answered these questions we come to the topics of specimen labeling and specimen preparation. While a confocal microscope generates images in the commonly used fluorescent mode, it also works in reflectance mode. In vivo clinical confocal microscopy of the living human eye is based on reflectance microscopy, and contrast is due to differences in the absorption and scattering coefficients of the various components of the ocular system (Böhnke and Masters, 1999). Similarly, with in vivo clinical confocal microscopy of living human skin contrast is generated from reflectance mode (Masters et al., 1997). On the other hand, the types of superresolution optical microscope that we discussed in Part III depend on a wide variety of photoactivatable or photoswitchable fluorescent proteins, or labeling with small organic molecules that may be attached to primary or to secondary antibodies for labeling specificity.

Each labeling protocol has inherent advantages and disadvantages. The resolution of the final image acquired by the localization microscopy technique depends on the labeling density of the specimen, the precision of localization, and the number of localizations. The number of photons detected from each fluorescent molecule is key to high-precision localizations. Superresolution techniques that are based on localization require a labeling density that is very high.

Genetically expressed fluorescent proteins are prone to protein overexpression that may confound correct interpretation of superresolution images. Additionally, they are much larger than small-molecule fluorescent probes. They are also typically less bright than small-molecule fluorescent probes. Many small-molecule fluorescent probes are bound to secondary antibodies attached to the primary antibody. This increases the distance from the fluorescent molecule to the binding site on the epitope that can confound correct interpretation of superresolution images. For live cell imaging a couple of questions arise. What are the effects of labeling procedures? How can we validate the physiological function of the specimen that was subjected to the labeling procedure, the specimen preparation, and the phototoxicity from the light used for the superresolution technique? The general topic of validation of superresolution images framed from a statistical point of view is another thought-provoking topic for detailed evaluation (Aspelmeier, Egner, and Munk, 2015).

Another consideration is the important question of whether our investigation can be carried out using two-dimensional superresolution microscopy? Or is three-dimensional superresolution microscopy required to answer the questions posed by our investigation? To obtain three-dimensional superresolution images large increases are required in the number of detected photons, the intensity of illumination, and the acquisition time. These requirements are huge for the superresolution techniques of STED/RESOLFT and

the localization techniques; however, they are minimal for SIM. Furthermore, SIM can be used with many fluorescent labels if they are sufficiently photostable.

The degree of complexity of the instrumentation in superresolution microscopy is also a significant concern when selecting a superresolution microscope. STED microscopes are complex and require appropriate skills to keep the STED beam and excitation beam aligned. Some superresolution microscope techniques based on localization techniques are optimal when used with TIRF microscopes. All superresolution optical microscopes are poorly suited to deep-tissue imaging (Small and Parthasarathy, 2014).

After taking all these considerations into account and the researcher is successfully operating a superresolution optical microscope on the specimen of interest there are still other problems. Vibration and temperature shifts can cause the microscope stage to move, and these shifts must be corrected. There is the ever-present background fluorescence that must be minimized. There are optical aberrations that degrade the image. Adaptive optics is one approach to minimizing optical aberrations. However, the field of probe development and the development and improvement of superresolution optical microscopes are works in progress and we can expect to see continued innovation.

What needs to be taken into consideration when selecting a particular type of superresolution optical microscope is important. Many factors need to be considered by investigators: whether they are imaging fixed cells or live cells, the required spatial and temporal resolution required to answer the research problems posed, the associated photo burden on phototoxicity and photobleaching of the specimen, the complexity of the labeling process, the achievement of high-density labeling, the complexity of the buffers required and their effects on cellular physiology and function, the requirements of specific lasers and their compatibility with available fluorescent molecules, the specificity of labels to cellular structures and components, ease of operation of the instrument, stability of the microscope to drift caused by thermal gradients and vibration, and the effects of labels on the cells under investigation. There is also the question as to whether three-dimensional imaging is required. In some cases two-dimensional imaging is sufficient to study the research questions posed. Another factor that requires careful consideration is the requirement for multicolor imaging.

The question of trade-offs was stressed in Schermelleh, Heintzmann, and Leonhardt (2010). Of course, the resolution required should be sufficient to solve the biological question posed regarding the susceptibility of the specimen to superresolution microscopy. Since all these considerations cannot be met by a single superresolution technique the investigator is faced with opposing capabilities and must make knowledgeable decisions based on trade-offs. Critical factors are the signal-to-noise ratio, the background intensity, and optical aberrations. The end goal is not the highest resolution achievable, but a solution to the biological problem posed that can only be obtained using a superresolution microscope.

Furthermore, there are the persistent dangers and problems of image artifacts that confound correct interpretation of superresolution images. This is the reason validating the superresolution is so critical. One approach is to study the same type of specimen using alternative types of microscopy (e.g., electron microscopy). That is

the approach Betzig et al. used in the first PALM publication (Betzig et al., 2006). Another approach is to study the same type of specimen using a variety of fluorescent molecules to label the specimen.

While there is great excitement about the capability of new types of superresolution optical microscopes to achieve spatial resolutions approaching molecule dimensions, it is just one metric of achievement. Another metric of success is the significance of the knowledge obtained using the superresolution optical microscope. There are a couple of questions that arise here: What new knowledge was gained using superresolution microscopes? Is the new knowledge significant to our understanding of biological structures and functions?

16.2 Highly Desirable Future Developments

The ultimate goal of microscopy is to have a variety of instruments covering a wide range of spatial and temporal scales of resolution that are ideally suitable for various specimens including biological and nonbiological specimens, fixed biological specimens, live cell specimens, and living and developing organisms. Such multiscale microscopes provide researchers with new tools to investigate biological processes bringing about a gain in knowledge that has important impacts on advances in cell biology, neuroscience, and medicine (van Engelenburg et al., 2014).

These microscopes should offer a variety of optical and nonoptical means of generating contrast in the specimens under observation. The confounding effects of radiation, electrons, or photons must be carefully controlled and understood to minimize alterations in the specimen under observation. Specifically, the phototoxic effects of light on living biological cells, tissues, and organisms must be controlled and understood by the use of careful control experiments and correlative microscopy (i.e., the use of various techniques and different labeling probes). This was illustrated in the first PALM publication (Betzig et al., 2006). A recent publication in *Science* was an example of three-dimensional superresolution microscopy and correlative electron microscopy to investigate a biological problem (van Engelenburg et al., 2014).

The theoretically unlimited resolution of nonlinear SIM and nonlinear fluorescence saturation in STED microscopy pose limits in their applicability to live cell imaging due to the effects that high intensities have on the function and structural integrity of live cells. Therefore, more work is required to further develop linear fluorescence superresolution techniques.

I want to make several comments here that are more specific to the content of Part III on superresolution optical microscopy (Thompson et al., 2012). Techniques based on single-molecule imaging and active control are works in progress. We expect to see further development yielding improved spatial and temporal resolution. Photocontrollable fluorescent protein development is today a limiting factor in the progress of superresolution optical microscopy (Shcherbakova et al., 2014).

A consequence is that many investigators are involved in probe development with the goals of finding brighter, more photostable, less phototoxic to live cells, and redshifted probes to minimize autofluorescence from cells. Furthermore, if superresolution microscopy is to be used to study dynamic cellular processes, then temporal resolution must be increased. Temporal resolution is a function of the detection system and depends on the kinetics of photoswitching photophysics. Progress in probe development should yield photoswitching probes with the photophysical switching kinetics desired. Another important problem to be worked on is the confounding effects of optical aberrations and high background fluorescence.

Many of the superresolution techniques discussed in Part III are not well suited to deep-tissue imaging in living organisms. Until new and innovative technical solutions to these problems are developed the use of light-sheet fluorescence microscopy (Chapter 11) together with adaptive optics may be the optimal microscopy technique for highly scattering specimens, developing organisms, and deep-tissue imaging. Deep-tissue imaging in high-scattering specimens is still problematical and challenging.

References

Aspelmeier, T., Egner, A., and Munk, A. (2015). Modern statistical challenges in high-resolution fluorescence microscopy. Annual Review of Statistics and Its Application, **2**, 163–202.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science, **313**, 1642–1645.

Birk, U. J. (2017). *Super-Resolution Microscopy: A Practical Guide*. Weinheim, Germany: Wiley-VCH.

Böhnke, M., and Masters, B. R. (1999). Confocal microscopy of the cornea. Progress in Retina & Eye Research, **18**(5), 553–628.

Masters, B. R., Gonnord, G., and Corcuff, P. (1997). Three-dimensional microscopic biopsy of in vivo human skin: A new technique based on a flexible confocal microscope. Journal of Microscopy, **185**, 329–338.

Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. Journal of Cell Biology, **190**, 165–175.

Shcherbakova, D. M., Sengupta, P., Lippincott-Schwartz, J., and Verkhusha, V. V. (2014). Photocontrollable fluorescent proteins for superresolution imaging. Annual Review of Biophysics, **43**, 303–329.

Small, A. R., and Parthasarathy, R. (2014). Superresolution localization methods. Annual Review of Physical Chemistry, **65**, 107–125.

Thompson, M. A., Lew, M. D., and Moerner, W. E. (2012). Extending microscopic resolution with single-molecule imaging and active control. Annual Review of Biophysics, **41**, 321–342.

van Engelenburg, S. B., Shtengel, G., Sengupta, P., Waki, K., Jarnik, M., Ablan, S. D., Freed, E. O., Hess, H. F., and Lippincott-Schwartz, J. (2014). Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. Science, **343**, 653–656.

Further Reading

Betzig, E. (2014a). *Single Molecules, Cells, and Super-Resolution Optics*: Lecture Slides. https://www.nobelprize.org/uploads/2018/06/betzig-lecture-slides.pdf (Accessed April 2, 2019).

Betzig, E. (2014b). *Single Molecules, Cells, and Super-Resolution Optics*: Nobel Lecture. https://www.nobelprize.org/prizes/chemistry/2014/betzig/lecture/ (Accessed April 2, 2019).

Betzig, E. (2014c). Biographical sketch. https://www.nobelprize.org/prizes/chemistry/2014/betzig/biographical/ (Accessed April 2, 2019).

Shtengel, G., Wang, Y., Zhang, Z., Goh, W. I., Hess, H. F., and Kanchanawong, P. (2014). Imaging cellular ultrastructure by PALM, iPALM, and correlative iPALM-EM. Methods in Cell Biology, **123**, 273–294.

van de Linde, S., Heilemann, M., and Sauer, M. (2012). Live-cell super-resolution imaging with synthetic fluorophores. Annual Review of Physical Chemistry, **63**, 519–540.

Yao, Z., and Carballido-López, R. (2014). Fluorescence imaging for bacterial cell biology: From localization to dynamics, from ensembles to single molecules. Annual Review of Microbiology, **68**, 459–476.

Appendix A
Annotated Biography of Key Publications Relevant to Abbe's *Beiträge* 1873

A.1 Introduction

This appendix supplements the material contained in Chapter 6. It contains further explanations and insights into Abbe's seminal theory with contributions from his coworkers, colleagues, and people who attended his university lectures in Jena.

In Chapter 7 I discussed the independent theoretical and experimental investigations of Helmholtz, Rayleigh, and Porter. Each investigator published detailed theoretical analyses of image formation in the light microscope and independently derived an equation for the limiting resolution of a light microscope due to diffraction. They independently confirmed Abbe's seminal paper of 1873 titled "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung." Abbe's paper was long, difficult to understand, and completely devoid of figures and mathematical equations. Abbe promised a second paper, complete with mathematical derivations; however, he died before completing this task.

This appendix reviews the published works that followed Abbe's seminal publication of 1873. Many of these articles were written by Abbe's colleagues at Zeiss Werke, some were based on Abbe's notes from his mathematical lectures at Jena University. Other contributions appeared in handbooks and textbooks of optics and microscopy. Some of these were very popular and were published in several editions. My selection includes contributions that were published after 1873 and continues with others published after Abbe's death in 1905. This selection includes eight publications in German and one in English.

George Biddell Airy was Astronomer Royal and Professor of Astronomy and Experimental Philosophy at the University of Cambridge. Abbe held a deep interest in astronomy and from 1877 to 1900 was Director of the Jena Observatory. Furthermore, Abbe was also fluent in English. Thus, it is likely that Abbe was familiar with Airy's publications.

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7

I summarize two of these publications here; one a paper on diffraction and the other a textbook (Airy, 1835, 1866). In 1835 Airy developed his formula for the diffraction pattern, called the Airy disk, which is the image of a point source of light in an aberration-free optical system (Airy, 1835). What did Airy achieve in his publication? Airy derived the mathematical equation for the diffraction pattern from a distant star, or another point source of light that is imaged by a circular lens. The result of his analysis showed that the point source of light was not imaged as a point but as a bright disk with a series of bright rings around it. The special case of the Fraunhofer diffraction of a circular aperture is given an eponymous name: the Airy pattern (Airy, 1835). Perusal of his publication demonstrates his lucid writing style.

What conclusions did Airy draw from his analysis? Beside the seminal fact that a point source of light is imaged as a bright disk with surrounding bright rings Airy calculated the positions of the rings. Furthermore, Airy noted that his calculations were in good agreement with observations (Airy, 1835).

How is Airy's seminal work related to image formation in the light microscope? In the light microscope the aperture of the microscope's objective is what causes the diffraction effects. A point source of light in the microscope results in a diffraction pattern in the diffraction plane of the microscope's objective.

Next, I discuss Airy's book titled "*Treatise on Undulatory Optics, Designed for Students in the University.*" The first edition was published in 1831 and the second edition in 1866. Airy presents his ideas as a series of propositions together with their mathematical formulation, derivations, and proofs. Then he presents specific cases that can be tested against experiments. The results of such experiments can then be compared with theory.

A.2 Contribution of Nägeli and Schwendener, *Das Mikroskop, Theorie und Anwendung desselben* (*The Microscope, Theory and Applications*) (1867, 1877)

Carl Nägeli and Simon Schwendener published the first edition of their classic book in 1867, publishing a second edition in 1877 (Nägeli and Schwendener, 1867, 1877). Carl Nägeli was a student of the botanist Matthias Schleiden. Schleiden motivated Carl Zeiss to start production of optical microscopes in his Jena shop. Simon Schwendener was a student of Nägeli. Working together in Jena they wrote a book titled *Das Mikroskop, Theorie und Anwendung desselben.*

The complete book was published in 1867 and was a milestone as it contained a unique and very comprehensive discussion of microscope optics. In support of this claim I list some of the topics from the first edition's table of contents: Theory of Microscopes; The Cardinal Points of Microscopes; Chromatic and Spherical Aberrations; Illumination with Transmitted (durchfallendem Licht) and Reflected Light (auffallendem Licht); Oblique Illumination; Testing of Microscopes for

Aberrations, Angular Aperture, Magnification, and Focal Length; Performance of Microscopes; The Relation between Focal Length and Aperture Angle; Interference in the Microscope; and Definition of "Defining Power" and "Penetrating Power" (Nägeli and Schwendener, 1867). Many of these terms are discussed in Abbe's seminal publication (Abbe, 1873).

I found it very instructive to read the first edition of *Das Mikroskop* (1867), then Abbe's seminal paper (1873), and finally the second edition of *Das Mikroskop* (1877). The first edition gives the reader a comprehensive overview of state-of-the-art light microscopes up to 1867; the second edition augments the comprehensive content with a section (without any mathematical proofs or derivations) on Abbe's theory of image formation in the light microscope based on diffraction. Nägeli and Schwendener's second edition of *Das Mikroskop* contained a 10-page section that discussed Abbe's 1873 paper on the theory of image formation in the optical microscope (Nägeli and Schwendener, 1877). In great contrast to Abbe's 1873 publication the second edition of their book had both figures and mathematical equations, making Abbe's theory of image formation in the microscope comprehensible to a wide audience. Unfortunately, the authors did not discuss the important experiments that Abbe carefully described in his 1873 publication.

A.3 Contribution of Leopold Dippel, *Das Mikroskop und Seine Anwendung*, Second Edition (*The Microscope and Its Applications*) (1882)

The next book that I discuss is Dippel's *Das Mikroskop und seine Anwendung,* which was part of the *Handbuch der Allegemeiner Mikroskopie*, Zweite Auflage and published in 1882. Leopold Dippel was a friend of Abbe. He was writing a new textbook on the theory of the microscope and contacted Abbe to determine whether Abbe could help him with his book. Subsequently, Abbe sent Dippel several chapters on geometric optics and physical optics; the text was almost completely written by Abbe (Feffer, 1994). Most of Abbe's manuscript discussed various types of geometric and chromatic optical aberrations. In addition, Abbe wrote about his theory of what was subsequently called the Abbe Sine Condition.

More importantly, Abbe included unpublished material from both theory and experiments that supported his theory of image formation in the light microscope based on light diffraction. For those who cannot read German I explain some terms here. The German words *Diffraktion* and *Beugung* both translate into the English word diffraction. The German word for refraction is *Brechung*.

How did Abbe's manuscript, sent to Dippel for incorporation into Dippel's *Das Mikroskop*, compare with Abbe's 1873 publication? There were significant differences. Abbe's 1873 publication was only text—there were no figures or equations in the publication. However, in his submission to Dippel Abbe included a detailed description of his experiments; he provided details of the structure of the

Diffraktionsplatte (test slide) that had various ruled lines with different interline spaces, and a description of the resulting diffraction pattern that he observed in the back focal plane of the microscope's objective.

A.4 Contribution of Siegfried Czapski, *Theorie der Optischen Instrumente nach Abbe* (*Theory of Optical Instruments after Abbe*) (1893)

Siegfried Czapski (1861–1907) studied physics in Berlin with Hermann von Helmholtz and Gustav Robert Kirchhoff. After completing his studies Czapski began to work in Jena at the Carl Zeiss optic works. Many years later he became the director of the Zeiss factory, a position previously held by Abbe.

In 1893 Czapski published the first edition of his major work that became a classic in the field of optical design titled *Theorie der Optischen Instrumente nach Abbe*. Although Czapski frequently cited Abbe throughout the book there is still one outstanding omission. While Czapski offered a concise, yet qualitative description of Abbe's theory of image formation in the light microscope based on light diffraction, there was no mathematical derivation of Abbe's theory of microscopic image formation.

In 1894 Czapski published "Die künstliche Erweiterung der Abbildungsgrenzen" (Theory of Spherical Aberration) in the *Handbuch der Physik*, **2**, 96. This publication has historical significance because it contained the mathematical derivation and proof of the Abbe Sine Condition. This was unique as it was the only proof and derivation of the Abbe Sine Condition that was published while Abbe was still alive (Czapski, 1894).

A.5 Contribution of Dr. Albrecht Zimmermann, *Das Mikroskop. Ein Leitfaden der wissenschaftlichen Mikroskopie* (*The Microscope, A Manual of Scientific Microscopy*) (1895)

Progress in the science and technology of microscopy are promoted by publications in physics journals and through the dissemination of practical knowledge that is based on these theoretical and practical advances. This emphasizes the importance of practical books on microscopy, read by microscopists without backgrounds in physics—something usually required to understand papers in physics journals. *Das Mikroskop* is an example of such a practical book (Zimmermann, 1895). This book contains numerous figures that are perfectly suited to its intended audience of botanists, zoologist, anatomists, and histologists.

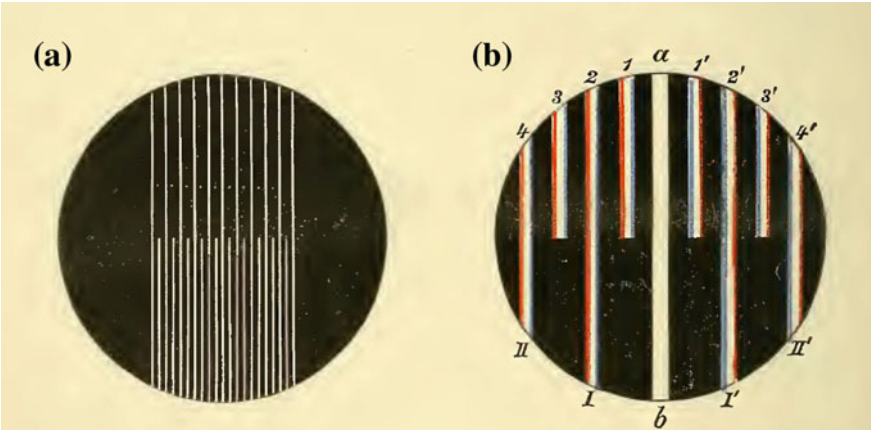


Fig. A.1 The use of the Abbe diffraction plate in a microscope

The importance of books, publications, and demonstrations that aim to educate the public should not be underestimated. However, at the time Zimmermann's book was published it would have been unusual for microscopists and the public to have access to the simple apparatus needed to reproduce the optical experiments that supported the new theories of light. The Zeiss firm developed a kit called the Abbe Diffraction Apparatus that was used by many microscopists to demonstrate Abbe's theory of image formation in the light microscope by diffraction.

In particular, Zimmermann's practical book presented the details of Abbe's experiments and then discussed the conclusions that could be derived from them. Zimmermann carefully described and illustrated the Abbe test plate (a.k.a. the Abbe diffraction plate). Colored illustrations were provided of the diffraction image observed in the back focal plane of the microscope's objective, both in the presence and absence of sets of diaphragms that altered the number of diffraction orders that could be collected by the numerical aperture of the microscope's objective. In Fig. A.1 the image on the left-hand side is the Abbe diffraction plate (the object) and the image on the right is the diffraction image observed in the back focal plane of the objective.

Zimmermann provides a qualitative analysis of Abbe's theory of the limits of resolution obtainable with the light microscope due to diffraction effects. He discusses the two important determiners of resolution in the light microscope: the wavelength of the illumination and the numerical aperture (NA). Furthermore, he explains that resolution can be enhanced by using illumination of a shorter wavelength and by increasing the NA of the microscope's objective. He suggests that with ultraviolet light as the source of illumination, with oblique illumination, and with the use of a monobromonaphthalin immersion system the maximum resolution that can be obtained is 0.12 μm. However, he then writes that this limit of resolution cannot be exceeded.

A.6 Contribution of William B. Carpenter and W. H. Dallinger, *The Microscope and Its Revelations*, Eighth Edition (1901)

In 1901 the eighth edition of *The Microscope and Its Revelations* was published simultaneously in both London and Philadelphia. The authors were W. B. Carpenter and W. H. Dallinger. Although Abbe was in poor health at the time the seventh edition was being written, he agreed to edit and comment on sections comprising Chapter 2, written by Carpenter and Dallinger, that described Abbe's work on image formation in the light microscope. Abbe was content with how the two authors represented his work. The common goal of the book's authors and Abbe was to disseminate Abbe's theories via an English book. This book, along with its figures and legends, are now in the public domain. Several important definitions and concepts were invented or developed by Abbe.

First, I list Abbe's Sine Condition, which Helmholtz derived independently of Abbe and published shortly after Abbe's publication. This condition is required of a lens or a complete optical system for the formation of sharp images from objects that are on axis and also off axis. Abbe's Sine Condition can be stated in several ways. Here is the statement from Carpenter and Dallinger: "the sines of the angles of the conjugate rays on both sides of an aplanatic optical system have the same quotient" (Carpenter and Dallinger, 1901).

Second, I state Abbe's definition of numerical aperture (NA): the sine of half the angular aperture that is multiplied by the refractive index of the medium.

Next, I relate a view that was held by Abbe (1873) and consequently taught and propagated by his followers. This view held that the image observed in the light microscope is actually composed of two images that are superimposed. One image is called the "absorption image," an image of the larger features of the object that define its gross shape and structure. The second image, which Abbe called the "diffraction image," is formed from the extremely small parts of the object and forms the small details of the image. These small detailed parts of the object are incapable of being imaged geometrically, and they result in the diffraction phenomenon that is the foundation of Abbe's theory of image formation in the light microscope.

However, 28 years after Abbe first published his seminal publication on image formation in the light microscope (Abbe, 1873a) he admitted that an error of his existed in Carpenter and Dallinger's eighth edition, published in 1901. This admission is an example of the ethical character of Abbe. Abbe was able, albeit after almost three decades from the time that he elaborated this false view as part of his theory, to change his views based on experimental observations.

I think that this initial error and Abbe's eventual correction deserves to be read in Abbe's own words. Therefore, here I directly quote Abbe as written in Carpenter and Dallinger's eighth edition: "I no longer maintain in principle the distinction between the absorption image (or direct dioptrical image) and the 'diffraction image,' nor do I hold that the microscopical image of an object consists of two

superimposed images of *different origin* or different mode of production" (Carpenter and Dallinger, 1901). Abbe goes on, "This distinction, which, in fact, I made in my first paper of 1873, arose from the limited experimental character of my first researches and the want [*sic*] of a more exhaustive theoretical consideration at that period. I was not then able to observe in the microscope the diffraction effect produced by relatively coarse objects because my experiments were not made with objectives of sufficiently long focus; hence it appeared that coarse objects (or the *outlines* of objects containing fine structural details) were depicted by the directly transmitted beam of light solely, without the co-operation of diffracted light" (Carpenter and Dallinger, 1901).

Next, Carpenter and Dallinger present a series of images that replicate what observers would see if they performed these parts of Abbe's experiments. The experiments served to show audiences of microscopists the experimental foundation of Abbe's theory of image formation in the light microscope based on diffraction theory. This part of the book served those microscopists who could not perform and witness such experiments themselves. In Abbe's seminal publication of 1973 he carefully described and interpreted these experiments. However, the publication did not contain any figures to help the reader comprehend these clever experiments and the conclusions Abbe derived from them (Abbe, 1873a).

In summary, *The Microscope and Its Revelations*, Eighth Edition is an exceptionally clear resource that provides deeper understanding of Abbe's theory of image formation in the light microscope based on diffraction (Carpenter and Dallinger, 1901). The book's production involved Abbe reading the sections on his theory and strongly approving their wording and content, the provision of significant pedagogical value via the use of figures, and the inclusion of an outstanding section on the history of the light microscope that contained high-quality engraved images of many of the early types of light microscopes.

A.7 Contribution of Siegfried Czapski in A. Winkelmann (Ed.), *Handbuch der Physik, Zweite Auflage, Sechster Band, Optik* (1906)

Winkelmann's *Handbuch der Physik, Zweite Auflage, Sechster Band, Optik*, published in 1906, is a comprehensive volume of contributed chapters on optics. It is the standard reference work and represents a single volume of the then-current (1906) knowledge on optics (Czapski, 1906).

I continue here with a discussion of some of the content of the *Handbuch* by Czapski and other contributors that related to the works of Abbe. Noteworthy is Chapter II titled "Die geometrische Theorie der optischen Abbildung" (The Geometric Theory of Optical Imaging) by S. Czapski. The subtitle of the chapter is revealing: "Nach den Universitätsvorlesungen von E. Abbe" (After the University lectures of E. Abbe). Of course, these were the lectures on geometric optics that Abbe gave at the University of Jena. Czapski presents, following Abbe, knowledge

on the subject of geometric optics, ray tracing, and aberrations for a variety of optical systems; derives all the key equations and laws; and presents specific approximations that simplify calculations. In a series of subsequent chapters Czapski defines Abbe's concept of *numerical aperture* as the product of the refractive index and the sine of one-half of the angular aperture of the lens. The concepts of entrance pupil and exit pupil and conjugate planes are all carefully defined.

It is also interesting that Czapski contributed a chapter (VIII) on the eye in which he presents the optical system of the eye; the optical properties of the models of eyes from both Helmholtz and Marius Tscherning (a Danish ophthalmologist); a discussion of the process of accommodation, which is the process by which the eye can change its focus; and a review of the aberrations of the human eye.

With reference to Abbe, Czapski' s contribution to chapter XII on the compound microscope includes a brief discussion of aperture in connection with Abbe's sine equation and his equation for the limits of resolution in the light microscope. Czapski cited Abbe's 1873 paper and Helmholtz's 1874 paper on resolution in the light microscope and the role of diffraction in image formation in the light microscope.

Furthermore, in chapter XVI Czapski presents a detailed description of several of Abbe precision instruments for the measurements of focal lengths (Abbe's Fokometer) and the aperture (Abbe's Apertometer) of a lens.

The chapter on interference of light (153 pages) by W. Feussner and the chapter on diffraction of light (87 pages) by F. Pockels represent comprehensive discussions on these two important topics that are relevant to the theory of image formation in the light microscope. These chapters form a detailed review of wave optics or physics optics up to 1906. Pockels' chapter begins with a historical introduction and then proceeds to discuss the Huygens principle and the independent contributions of Fresnel, Fraunhofer, and Sommerfeld to the theory of diffraction. Differential equations that describe the diffraction of light are then solved for a number of cases.

Pockels' chapter on the diffraction of light contains a section titled "The role of diffraction in image formation in the microscope." The basic concepts of Abbe's 1873 paper are reviewed in this section and there is a figure that illustrates the diffraction of light by the object, the entrance of some of these diffraction orders into the angular aperture of the microscope's objective to form the diffraction image in the back focal plane of the microscope's objective, and the interference of these diffractions to form an image in the image plane. Pockels also discusses Abbe's equation for the limiting resolution of the light microscope in terms of the wavelength of illumination, and the numerical aperture given symbol A (the modern symbol is NA) is stated

$$d = \frac{\lambda}{2\mathrm{A}} \tag{A.1}$$

A.8 Contributions of Otto Lummer and Fritz Reiche, *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe* (*The Theory of Image Formation in the Microscope by Ernst Abbe*) (1910)

Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe is unique and significant since the authors had permission from Mrs. Abbe to incorporate Abbe's 1888 lecture notes. This helped Lummer and Reiche in their task to formulate a consistent mathematical description of Abbe's theory of image formation in the light microscope based on diffraction (Lummer and Reiche, 1910). Furthermore, this book is probably the closest that we can get to reading the mathematical derivations based on Fourier analysis for Abbe's theory of image formation in the light microscope. After Abbe's seminal publication of 1873 was published without any figures or equations he was pressed to publish the mathematical derivations of his theory. While Abbe made several promises to write such a paper he died before they could be fulfilled. Therefore, the 1910 book by Lummer and Reiche, based on Abbe's lecture notes, indirectly gives us access to how Abbe's promised mathematical derivations would have appeared if he had lived to publish them.

Previously, in Dippel's 1882 book titled *Das Mikroskop und seine Anwendung,* which was prepared in close association with Abbe, Dippel wrote that the use of Fourier integrals is required for Abbe's theory of image formation in the microscope (Dippel, 1882). Dippel's book was published 6 years before Lummer stated that he had listened to Abbe's lecture in Jena on the theory of image formation in the light microscope based on diffraction.

I now continue with a summary of Lummer and Reiche's 1910 book. The publisher conceived of this textbook, with its physical and mathematical details, as a valuable teaching resource. The facing page of Lummer and Reiche's book contains a portrait of Abbe and his signature.

The book's foreword was written by Lummer. In it he explains how he got the opportunity to listen to Abbe's lectures in Jena. Lummer worked on precision optical measurements at Berlin's Physikalisch-Technische Reichsanstalt (PTR) (the German Physical and Technical Institute). In connection with his work he traveled to Jena to listen to Abbe's lectures on theoretical optics. In the foreword Lummer thanks Frau Professor Abbe (as he called Mrs. Abbe) for her permission to use Abbe's lecture notes and figures from his 1888 textbook.

This book, with its unique insight into how Abbe used Fourier analysis to provide a mathematical derivation of his diffraction theory, is highly recommended to readers. It is short and comprehensive. The first chapter contains the mathematical derivations for the laws of refraction as well as derivation of the Abbe Sine Condition. In the second chapter the reader is introduced to image formation for self-luminous objects (incoherent light) in terms of physical optics. Here diffraction theory is developed with derivations of Kirchhoff's principle and the Huygens–Fresnel principle. The third chapter discusses image formation of nonluminous (coherent light) objects. Lummer and Reiche treat the diffraction of light for various

cases: an illuminated slit, two parallel adjacent slits, and infinitely narrow slits for both coherent and incoherent light. Lummer and Reiche, closely following Abbe's lecture notes, use Fourier integral equations to evaluate diffraction integrals. The Abbe limit of resolution is then derived for image formation in the microscope.

In summary, Lummer and Reiche's 1910 book, although only 108 pages in length, is a very concise summary of the mathematical derivations related to Abbe's theory of image formation in the microscope based on diffraction. The derivations are given stepwise and the textbook is replete with numerous figures that make the theory clear and comprehensible.

A.9 Contributions of Siegfried Czapski and Otto Eppenstein, *Grundzüge der Theorie der optischen Instrumente nach Abbe*, Dritte Auflage (*Fundamentals of the Theory of Optical Instruments after Abbe*, Third Edition) (1924)

Czapski's *Grundzüge der Theorie der optischen Instrumente nach Abbe* was published in several editions: the first edition was published in 1893 and edited by Siegfried Czapski; the second edition was published in 1904 and edited by Siegfried Czapski and Moritz Rohr; and the third edition was published in 1924 and edited by Siegfried Czapski and Otto Eppenstein. With the exception of a few changes to the contributing authors all these editions were very similar and their chapters almost identical to those that comprised the volume *Optics* in Winkelmann's series *Handbuch der Physik, Sechster Band, Optik*, published in several editions.

Czapski and Eppenstein's 1924 edition of their *Grundzüge der Theorie der optischen Instrumente nach Abbe* is very similar in content to Winkelmann's 1906 *Handbuch der Physik, Zweite Auflage*, *Sechster Band, Optik.* These two contributed books differ in that the latter was published 18 years after the former and that each book has different contributors. However, there is some overlap of content between the two.

Chapter XVI titled "Das zusammengesetzte Mikroskop" (The Compound Microscope) was contributed by H. Boegehold (Czapski and Eppenstein, 1924). I will discuss the section titled "Die Apertur und die Grenzen der mikroskopischen Wahrnehmung" (The aperture and the limits of microscopic perception). Boegehold begins by explaining the beam paths for a compound microscope with a simple ocular and similarly a complex ocular. Abbe's Sine Condition is then derived and discussed.

After the introduction the author follows with a section titled "The aperture and the limits of microscopic perception." Abbe's limiting resolution is given in terms of the wavelength of illumination divided by the numerical aperture. For oblique illumination the value of the limiting resolution is given by the wavelength divided

by twice the numerical aperture. These limiting resolution equations are followed by some figures that show gratings with various linewidths and the resulting diffraction image observed in the back focal plane of the microscope's objective, as well as the diffraction pattern from an object that consists of crossed grating. Other figures illustrate cases when diaphragms with slits in different orientations that can remove some of the diffraction orders alter the diffraction image. The section ends with a discussion of Abbe's equation for the limiting resolution in a light microscope with oblique illumination

$$d = \frac{\lambda}{2n \sin u} \tag{A.2}$$

The limiting resolution according to Abbe's theory is the wavelength of illumination divided by twice the numerical aperture (Czapski and Eppenstein, 1924).

A.10 Contributions of the late O. Lummer, *Müller-Pouillet's Lehrbuch der Physik, Zweiter Band, 11. Auflage. Die Lehre von der strahlenden Energie (Optik)*, Erste Hälfte (*Müller-Pouillet's Physics Textbook, Second Volume, 11th Edition. The Teaching of Radiant Energy (Optics)*, First Half) (1926)

This section illustrates the role of textbooks in summarizing the state of knowledge in a specific field of physics. Multiple authors contributed chapters to such textbooks. These textbooks, as well as handbooks of physics, were to be found on the bookshelves of physicists in their academic or laboratory settings.

Müller-Pouillet's Lehrbuch der Physik (*Müller-Pouillet's Physics Textbook*) was a multivolume textbook containing chapters contributed by various authors. It was a classic set of volumes and by 1926 had been published in several editions. Its second volume titled *Die Lehre von der strahlenden Energie (Optik)* (*The teaching of radiant energy (Optics)*) was edited by Lummer who died in 1925, a year before its final publication. The contributors to this volume included H. Erggelet (from Jena), F. Jüttner (from Breslau), A. König (from Jena), M. v. Rohr (from Jena), and E. Schrödinger (from Zürich).

Quantum mechanics was the monumental physical theory suggested and developed over a 3-year interval between the 1926 first part of *Die Lehre von der strahlenden Energie (Optik)* and the 1929s part of the volume. Quantum mechanics was developed in several forms that eventually proved to be mathematically identical. Wave mechanics was developed by Erwin Schrödinger. Werner Heisenberg developed matrix mechanics. Paul Dirac developed a version of quantum mechanics based on noncommuting dynamical variables. The 1929 edition of *Die Lehre von der strahlenden Energie (Optik), Zweite Hälfte-Erster Teil.*

In: Müller-Pouillet's Lehrbuch der Physik, 11. Auflage, Zweiter Band, was edited by Karl Wilhelm Meissner from Frankfurt, with contributions from E. Buchwald (from Danzig), M. Czerny (from Berlin), E. Gehrcke (from Charlottenburg), G. Hettner (from Berlin), H. Kohn (from Breslau), R. Minkowski (from Hamburg), and W. Pauli (from Zürich).

In the 1929 edition the volume is divided into two parts. Here, I describe the first part (Meissner, 1929a). In *Die Lehre von der strahlenden Energie (Optik)*, Zweite Hälfte-Erster Teil we see the nexus of optics and atomic physics. A major portion of the volume is devoted to spectroscopy and the theoretical and experimental study of blackbody radiation. In this edition we see that many types of optical phenomena, previously explained in terms of phenomenological theories, are now explained in terms of atomic theory. The advances made after the development of quantum theory were phenomenal; our understanding of spectroscopy was greatly advanced and atomic theory was applied to the understanding of the interaction between light and matter. During the same period came advances in the wave theory of light (i.e., physical optics) that led to an improved understanding of double refraction, light–matter interactions, diffraction, interference, and the study of linear and circular polarization. Radiation theory, from Planck's to Einstein's to Compton's contributions, are discussed in great detail from both the theoretical and experimental viewpoints. There is also an excellent discussion on instrumentation used for optical measurements and for the investigation of optical phenomena. The theories of refraction, diffraction, and dispersion—including anomalous dispersion—are given very detailed consideration. The first part ends with discussion of the quantum theory of dispersion, absorption and emission in quantum theory, and the width of spectral lines.

Next, I describe the second part of *Die Lehre von der strahlenden Energie (Optik)*, Zweite Hälfte-Zweiter Teil (Meissner, 1929b). The editor of this second part was Karl Wilhelm Meissner from Frankfurt, and it had contributions from E. Back (from Hohenheim-Stuttgart), D. Coster (from Groningen), B. Gudden (from Erlangen), G. Hertz (from Berlin-Charlottenburg), A. Kratzer (from Münster), R. Ladenburg (from Berlin-Dahlem), L. Meitner (from Berlin-Dahlem), F. Paschen (from Berlin-Charlottenburg), W. Pauli (from Zürich), and R. W. Pohl (from Göttingen).

This second part of *Die Lehre von der strahlenden Energie (Optik)*, Zweite Hälfte-Zweiter Teil begins with Pauli's contribution on the general foundation of quantum theory (Meissner, 1929b). In subsequent chapters the topics of advanced spectroscopy, light–matter interaction, and the effects of electric and magnetic fields on atomic spectra are fully explained with equal weight being given to mathematical derivations and the physical principles of such phenomena.

I now return to the discussion related to Abbe's theory of image formation in the light microscope based on diffraction as discussed in *Die Lehre von der strahlenden Energie (Optik)*, Erste Hälfte (Lummer, 1926). On pages 570–571 the Abbe Sine Condition and Abbe's definition of numerical aperture are stated and discussed, but not derived. Chapter 15 is a comprehensive review of the diffraction of light. The first case considered is diffraction from a single slit; the second case being analysis

of light diffraction from a grating; the next being analysis of the diffraction of a grating with many columns.

Chapter 16 describes the theory of image formation for nonluminous objects. This chapter is about image formation with coherent illumination. In this chapter Abbe's theory is explained as follows: the coherent light is transmitted through the object which causes it to become diffracted into several diffraction orders. The numerical aperture of the objective collects several of the diffracted orders and the diffraction image can be observed in the back focal plane of the objective. Finally, the diffracted orders interfere with each other in the image place and form the image.

The experimental proof of Abbe's theory of image formation in the light microscope based on diffraction is given as a set of Abbe's experiments. In the color plates section various objects are shown as test places with engraved lines, and the diffraction images observed in the back focal plane of the microscope's objective are shown. Then, when specific apertures that select the specific orders of diffracted light allowed to enter the microscope are inserted into the optical system, the results on diffraction images are shown in the color figures.

Next, equations that show the limiting resolution of the light microscope as derived by Abbe and then independently by Helmholtz are given. Numerical examples are provided for the limiting resolution in the light microscope: for violet light with a Fraunhofer Line H the wavelength is about 0.4 μm and the numerical aperture is 1.66 for bromonaphthalene. The calculated limit of resolution is approximately 0.00025 or 1/4000 mm.

The only practical method to achieve a higher resolution is to reduce the wavelength of illumination. This is exactly the idea that A. Köhler, working at the Zeiss firm, developed in 1904—inventing a photographic microscope using ultraviolet light with a wavelength of 275 μm and replacing glass lenses with quartz lenses capable of transmitting this short wavelength of light. This ultraviolet microscope, using a photographic detector instead of the human eye (insensitive to ultraviolet light), quartz lenses, and an ultraviolet light source, achieved a limiting resolution that was approximately double that achieved for a light microscope illuminated with violet light.

The remaining sections of this chapter represent a detailed mathematical analysis of image formation in the light microscope in terms of diffraction for cases of coherent and incoherent illumination. The detailed mathematical analysis is based on the dissertation work of M. Wolfke who was awarded his doctorate in Breslau in 1910. The final section is an extensive analysis of the ultramicroscope invented by Siedentopf and Zsigmondy in 1903.

References

Abbe, E. (1873a). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, **IX**, 413–468.

Abbe, E. (1873b). Ueber einen neuen Beleuchtungsapparat am Mikroskop. Archiv für mikroskopische Anatomie. **IX**, 469–480.

Abbe, E. (1874). A contribution to the theory of the microscope and the nature of microscopic vision. Translated into English by H. E. Fripp. Proceedings of the Bristol Naturalists Society, **I**, 202–258. Read before the Bristol Microscopical Society, December 16, 1874.

Abbe, E. (1875). About a new lighting apparatus on the microscope. Translated into English by H. E. Fripp. Monthly Microscopical Journal, April.

Abbe, E. (1879a). On Stephenson's system of homogeneous immersion for microscope objectives. Journal of the Royal Microscopical Society, **2**, 256–265.

Abbe, E. (1879b). On new methods for improving spherical correction, applied to the correction of wide-angled object-glasses [microscope objectives]. Journal of the Royal Microscopical Society, **2**, 812–824.

Abbe, E. (1880a). The Essence of Homogeneous Immersion. Journal of the Royal Microscopical Society, **1**, 526.

Abbe, E. (1880b). Ueber die Grenzen der geometrischen Optik. Mit Vorbemerkungen über die Abhandlung "Zur Theorie der Bilderzeugung" von Dr. R. Altmann. *Jenaischen Gesellschaft für Medicine und Naturwissenschaft*, 71–109.

Abbe, E. (1882a). The relation of aperture and power in the microscope. Journal of the Royal Microscopical Society, Sec. **II**, 300–309, [read before the Society on May 10, 1882], [Abbe wrote the paper in English].

Abbe, E. (1882b). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Sec. **II**, 460–473. [read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Abbe, E. (1883). The relation of aperture and power in the microscope condenser. Journal of the Royal Microscopical Society, Sec. **II**, 790–812. [read before the Society on June 14, 1882], [Abbe wrote the paper in English].

Abbe, E. (1889a). On the effect of illumination by means of wide-angled cones of light. Journal of the Royal Microscopical Society, Series II, **IX**, 721–724.

Abbe, E. (1989b). *Gesammelte Abhandlungen*, I-IV. Hildesheim: Georg Olms Verlag. [Originally published in 1904, Jena: Verlag von Gustav Fischer].

Airy, G. B. (1835). Diffraction of an object-glass with circular aperture. Transactions of the Cambridge Philosophical Society, **5**, 283–291.

Airy, G. B. (1866). *Treatise on the Undulatory Theory of Optics, Designed for Students in the University.* London: Macmillan and Co.

Altmann, R. (1880a). Zur thorie der Bilderzeugung. *Archive für Anatomie und Physiologie* (Anatomische Abtheilung), 111–184.

Altmann, R. (1880b, 1882). Ueber die Vorberkungen des Hrn. Prof. Abbe zu seinen 'Grenzen der geometrischen Optik.' *Archiv für Anatomie und Physiologie* (Anatomische Abtheilung), (1880b), 354–363; (1882), 52–59.

Auerbach, F. (1918). *Ernst Abbe: sein Leben, sein Wirken, seine Persönlichkeit nach Quellen und aus eigener Erfahrung geschildert von Felix Auerbach.* Leipzig: Akademie Verlag Gesellschaft.

Auerbach, F. (1922). *Ernst Abbe: sein Leben, sein Wirken, seine Persönlichkeit nach Quellen und aus eigener Erfahrung geschildert von Felix Auerbach.* Zweite Auflage. Leipzig: Akademie Verlag Gesellschaft.

Berek, M. (1926a). Über Kohärenz und Konsonanz des Lichtes. I [On coherence and consonance of light]. Zeitschrift für Physik, **36**, 675–688.

Berek, M. (1926b). Über Kohärenz und Konsonanz des Lichtes. II. Theorie der Interferenzen bei hohen Gangunterschieden unter Berücksichtigung der Kohärenzverhältnisse [Theory of interference taking into account the coherence conditions]. Zeitschrift für Physik, **36**, 824–838.

Berek, M. (1926c). Über Kohärenz und Konsonanz des Lichtes. III. Interferenzen, welche durch Beugung entstehen [Interference which results from diffraction]. Zeitschrift für Physik, **37**, 387–394.

Berek, M. (1926d). Über Kohärenz und Konsonanz des Lichtes, IV. Die optische Abbildung nichtselbstleuchtender Objekte [Optical imaging is non self-luminous objects]. Zeitschrift für Physik, **37**, 420–450 (*1926*).

Berek, M. (1929). XXI. On the extent to which real image formation can be obtained in the microscope. Journal of the Royal Microscopical Society, **49**, 240–249.

Born, M. and Wolf, E. (1999). *Principles of Optics, 7th (expanded) edition*. Cambridge: Cambridge University Press.

Bradbury, S. (1996). The reception of Abbe's theory in England. Proceedings of the Royal Microscopical Society, **31**, 293–301.

Carpenter, W. B. and Dallinger, W. H. (1901). *The Microscope and its Revelations, Eighth Edition.* Philadelphia: P. Blakiston's Son & Co.

Conrady, A. E. (1904). XII. Theories of Microscopical Vision: A Vindication of the Abbe Theory. Journal of the Royal Microscopical Society, **24**, issue 6, 610–633.

Conrady, A. E. (1905). VI. Theories of Microscopical Vision (Second Paper). Journal of the Royal Microscopical Society, **25**, issue 5, 541–553.

Crisp, F. (1878). On the influence of diffraction in microscopic vision. Journal of the Quekett Microscopical Club, **5**, 79–86.

Czapski, S. (1893). *Theorie der Optischen Instrumente nach Abbe*. Breslau: Verlag von Eduard Trewendt. [special printing of Adolf Winkelmann's *Handbuch der Physik*].

Czapski, S. (1894). Die künstliche Erweiterung der Abbildungsgrenzen. [Theory of Spherical Aberration; proof of Abbe'sine condition]. Handbuch der Physik, **2**, 96.

Czapski, S. (1906). Geometrische Optik. In: A. Winkelmann, Ed. *Handbuch der Physik, Zweite Auflage, Sechster Band, Optik.* Leipzig: Verlag von Johann Ambrosius Barth.

Czapski, S. and Eppenstein, O. (1924). *Grundzüge der Theorie der optischen Instrumente nach Abbe. Dritte Auflage.* [Fundamentals of the theory of optical instruments of Abbe. Third Edition]. Leipzig: J. A. Barth Verlag.

Dippel, L. (1882). *Das Mikroskop und seine Anwendung. Erster Theil. Handbuch der Allegemeiner Mikroskopie Zweite Auflage*. Braunschweig: Friedrich Vieweg und Sohn.

Everett, J. D. (1904). *IX.* A Direct Proof of Abbe's Theorems on the microscopic Resolution of Gratings. Journal of the Royal Microscopical Society, **24**, issue 4, 385–387.

Feffer, S. M. (1994). *Microscopes to munitions: Ernst Abbe, Carl Zeiss, and the transformation of technical optics, 1850–1914*. PhD Dissertation, University of California, Berkeley, 1994. Ann Arbor: UMI Dissertation Services.

Fripp, H. E. (1874a). A contribution to the theory of the microscope and of microscopic vision, after Dr. E. Abbe, Professor in Jena. Proceedings of the Bristol Naturalists' Society, **I**, 200–261.

Fripp, H. E. (1874b). On the limits of optical capacity of the microscope. Translated from Germany publication of Hermann von Helmholtz. Proceedings of the Bristol Naturalists' Society, 407–440.

Fripp, H. E. (1876a). On aperture and function of the microscope object glass [microscope objective]. Proceedings of the Bristol Naturalists' Society, **I**, 441–456.

Fripp, H. E. (1876b). On the physiological limits of microscopic vision. Proceedings of the Bristol Naturalists' Society, **I**, 457–475.

Gerth, K. (2005). *Ernst Abbe, Scientist, Entrepreneur, Social Reformer*. Jena: Verlag Dr. Buseert & Stadeler.

Gordon, J. W. (1901). An examination of the Abbe diffraction of the microscope. Journal of the Royal Microscopical Society, **21**, 353–396.

Hartinger, H. (1930). Zum fünfundzwanzigsten Todestage von Ernst Abbe. Die Naturwissenschaften, Heft **3**, 49–63.

Helmholtz, H. (1851). *Beschreibung eines Augen-Spiegels zur Untersuchung der Netzhaut im lebenden Auge*. Berlin: A. Förstner'sche Verlagsbuchhandlung. Description of an ophthalmoscope for examining the retina in the living eye. English translation by Robert W. Hollenhorst, Chicago, *Arch. Ophthalmol*., **46**, 565–83 (1951).

Helmholtz, H. (1874). Die theoretische Grenze für die Leistungsfähigkeit der Mikroskope. *Poggendorff's Annalen der Physik Jubelband*, 557–584, Leipzig. Translated as: "On the theoretical limits of the optical capacity of the microscope". Monthly Microscopical Journal, **16**, 15–39 (1876).

Helmholtz, H. (1877). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, English translation of the fourth edition (1877) by A.J. Ellis, N.Y., Dover Publications, 1954.

Helmholtz, H. (1909–1911). *Handbuch der Physiologischen Optik*, Dritte Auflage, English translation of the third edition, three volumes (1909–1911) by J. P.C. Southall, Washington, D. C., Optical Society of America, 1924.

Kingslake, R. (1978). *Lens Design Fundamentals*. New York: Academic Press.

Lister, J. J. (1830). On some properties in achromatic object-glasses applicable to the improvement of the microscope. Philosophical Transactions of the Royal Society of London, **120**, 187–200.

Listing, J. B. (1869a). Vorschlag zu ferner Vervollkommung des Mikroskops auf einem geänderten dioptrischen Wege [Proposal for further improvement of the microscope on a modified dioptric way]. Poggendorff's Annalen der Physik, **136**, 467–472.

Listing, J. B. (1869b). Nachtrag betreffend die neue Construction des Mikroskops [Supplement relating to the new construction of the microscope]. Poggendorff's Annalen der Physik, **136**, 473–479.

Listing, J. B. (1869c). Notiz über ein neues Mikroskop von R. Winkel [Note on a new microscope by R. Winkel]. Poggendorff's Annalen der Physik, **142**, 479–480.

Lord Rayleigh (1896). On the Theory of Optical Images, with Special Reference to the Microscope. Philosophical Magazine and Journal of Science, London, **XLII**, 167–95.

Lummer, O. (1909). *Die Lehre von der strahlenden Energie (Optik).* In: *Müller-Pouillet's Lehrbuch der Physik, 10 Auflage*, *Zweiter Band,* Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.

Lummer, O. (1926). *Die Lehre von der strahlenden Energie (Optik),* Erste Hälfte. *In: Müller-Pouillet's Lehrbuch der Physik, 11. Auflage, Zweiter Band,* Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.

Lummer, O. and Reiche, F. (1910). *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe.* Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.

Masters, B. R. (2006). *Confocal Microscopy and Multiphoton Excitation Microscopy: The Genesis of Live Cell Imaging*. Bellingham: SPIE Press.

Masters, B. R. (2009a). John William Strutt, Third Baron Rayleigh: A Scientific Life that Bridged Classical and Modern Physics. Optics and Photonics News, **20**, no. 6, pp. 36–41.

Masters, B. R. (2009b). C. V. Raman and the Raman Effect. Optics & Photonics News, March 20, 40–45.

Masters, B. R. (2010). Hermann von Helmholtz, A 19th Century Renaissance Man. Optics and Photonics News, **21**, no. 3, pp. 35–39.

Meissner, K. W. (1929a). *Die Lehre von der strahlenden Energie (Optik), Zweite Hälfte-Erster Teil. In: Müller-Pouillet's Lehrbuch der Physik, 11. Auflage, Zweiter Band,* Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.

Meissner, K. W. (1929b). *Die Lehre von der strahlenden Energie (Optik), Zweite Hälfte-Zweiter Teil. In: Müller-Pouillet's Lehrbuch der Physik, 11. Auflage, Zweiter Band,* Braunschweig: Druck und Verlag von Friedrich Vieweg und Sohn.

Moore, H. (1940). XI. Theories of image formation in the microscope. Journal of the Royal Microscopical Society, **60**, issue 3, 140–151.

Nägeli, C. and Schwendener, S. (1867). *Das Mikroskop, Theorie und Anwendung desselben. Erste Auflage*. Leipzig: Verlag von Wilhelm Engelmann.

Nägeli, C. and Schwendener, S. (1877). *Das Mikroskop, Theorie und Anwendung desselben. Zweite Verbesserte Auflage*. Leipzig: Verlag von Wilhelm Engelmann.

Porter, A. B. (1906). On the diffraction theory of microscope vision. Philosophical Magazine, **6**, 154–166.

Rheinberg, J. (19b01). In "Proceedings of the Society". Journal of the Royal Microscopical Society, **XXI**, 480.

Schellenberg, F. M. (2004). *Selected Papers on Resolution Enhancement Techniques in Optical Lithography*, SPIE Milestone Series, MS **178**, Bellingham: SPIE Press.

Stephenson, J. W. (1875) *Monthly Microscopical Journal*: *Transactions of the Royal Microscopical Society and Record of Histological Research at home and Abroad*. Vol. XIV, (1875). Extracts from Mr. H. E. Fripp's translation of Professor Abbe's paper on the microscope. Reprinted from the *Proceedings of the Bristol Naturalists' Society*, vol. **I**, part 2, 191–201, 245–254.

Stephenson, J. W. (1877). Observations on Professor Abbe's experiments illustrating his theory of microscopic vision. (Read before the Royal Microscopical Society, London, January 3, 1877). Monthly Microscopical Journal: Transactions of the Royal Microscopical Society and Record of Histological Research at home and Abroad, **XVII**, 82–88.

Volkmann, H. (1966). Ernst Abbe and His Work. Applied Optics, **5**(11), 1720–1731.

von Rohr, M. (1940). *Ernst Abbe*. Jena: Gustav Fischer.

Zimmermann, A. (1895). *Das Mikroskop. Ein Leitfaden der wissenschaftlichen Mikroskopie*. [*The Microscope, A Manual of Scientific Microscopy*]. Leipzig and Wien: Franz Deuticke.

Appendix B
Responsible Conduct of Research

What is science? Science is a social process that generates knowledge about the physical universe and is built on trust and ethical conduct. Such generation of knowledge is always subject to experimental outcomes that can be objectively compared with predictions derived from theory. A lack of verification of theory typically results in either rejection or modification. Think of science as a set of algorithms that (if followed) can be used to generate knowledge about the physical universe. Furthermore, the process of science can be taught to others and passed on from generation to generation. As scientists, we strive to promote mentorship (critical to the teaching and training of a new generation of scientists) and cooperation, advance research, and work in a context of social responsibility.

What is pseudoscience? Pseudoscience is composed of spurious theories that are promulgated as science but are actually bogus (i.e., without objective validation). Today the boundaries between truth and falsehoods and lies, between science and pseudoscience, are becoming distorted and subsequently weakened. Facts and opinions are becoming less distinct.

In fact, science today is coming under increasingly severe attacks. We see this phenomenon in the public debates between political leaders and government officials with respect to climate change and the putative harmful effects of vaccines. Pseudoscience and nonsense can cause great harm when they affect public policy. Think of the human effects of climate change. Think of the global outbreaks of infectious diseases.

The aims of scientists include the promotion of science, the protection of science, and the advancement of science. Mentoring, risk taking, and innovation all serve to promote science. A deep understanding of scientific misconduct is necessary to enhance the protection of science. Good practices in terms of research and scientific communication all serve to advance science.

What are the impacts of scientific misconduct on the progress of science? Impacts may include scientists being misled by false publications, funding being

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7

wasted, and the time and effort of editors and peer-reviewers being wasted too. Another concern is loss of trust by the public in the process of science and the subsequent diminishing of government funding.

The set of best practices for science falls under the rubric of Responsible Conduct of Research. These principles are taught as interactive courses at universities and research institutes worldwide. Topics typically include good research practices; fraud, fabrication, and falsification; authorship; plagiarism; peer review; research misconduct; the use of animals in research; research involving human subjects; manipulation of digital images; proper citation; and intellectual property. Another important topic is the dissemination and sharing of data.

Several years ago I developed a set of lecture notes for a series of short courses I taught at the Massachusetts Institute of Technology. My "Responsible Conduct of Research" presentation has been taught at a number of universities in Europe, the Middle East, and Asia. Feedback from students and faculty members has been incorporated in the most recent version of my lecture notes. My "Responsible Conduct of Research" course presentation is available as a pdf at https://storage.googleapis.com/springer-extras/zip/2020/978-3-030-21690-0.zip.

Index

B. R. Masters, *Superresolution Optical Microscopy*, Springer Series in Optical Sciences 227, https://doi.org/10.1007/978-3-030-21691-7