

Kedar Khare
Mansi Butola
Sunaina Rajora

Fourier Optics and Computational Imaging

Second Edition



Ane Books
Pvt. Ltd.



Springer

Fourier Optics and Computational Imaging

Kedar Khare · Mansi Butola · Sunaina Rajora

Fourier Optics and Computational Imaging

Second Edition



Ane Books
Pvt. Ltd.



Kedar Khare
Optics and Photonics Centre
Indian Institute of Technology Delhi
New Delhi, India

Mansi Butola
Department of Physics
Indian Institute of Technology Delhi
New Delhi, India

Sunaina Rajora
Department of Physics
Indian Institute of Technology Delhi
New Delhi, India

ISBN 978-3-031-18352-2 ISBN 978-3-031-18353-9 (eBook)
<https://doi.org/10.1007/978-3-031-18353-9>

Jointly published with Ane Books Pvt. Ltd.
ISBN of the Co-Publisher's edition: 978-9-385-46229-0

2nd edition: © The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Preface to the Second Edition

Writing the second expanded edition of the book *Fourier Optics and Computational Imaging* has been a pleasure and has proved to be a healthy distraction during the difficult pandemic years. The aim of this book is to introduce Physics and Engineering students at Master's or advanced undergraduate levels to the exciting world of imaging research, so that they can get a good perspective of the field and get ready to begin research and development activities. The first edition of this book was published in late 2015 after which I introduced a 700-level course [PYL759: Computational Optical Imaging] at Indian Institute of Technology Delhi. In about a year or so, two enthusiastic PhD students Mansi Butola and Sunaina Rajora joined my research group. Writing the first edition of this book initiated a number of email exchanges with readers across the globe, and the book also received a good number of citations in the research literature much beyond my expectations. Teaching this material as a regular course over the last 5 years to bright students at Indian Institute of Technology Delhi also renewed my own understanding of a number of topics as the students asked a number of probing questions and pointed out some errors that were left inadvertently in the first edition. Most importantly, all of this led us to a number of novel research ideas that grew out of my hours and hours of discussions with the co-authors as their PhD work progressed. It was therefore natural for me to invite them to participate in the effort of writing the second edition of the book.

Compared to the first edition, this new expanded version presents a number of topics with much better explanations that students will find easier to follow. Several new topics have also been introduced that have attracted much attention in the imaging literature over the past 10 years. The book also comes with a few problems at the end of each chapter. The problems are sometimes left somewhat open-ended intentionally in order to get the students thinking. The problems have also been designed to introduce the readers to allied imaging concepts relevant to the material covered in the particular chapter. The list of references at the end of each chapter is not meant to be an exhaustive survey of literature but rather provides a few pointers to interested readers who may want to explore some topics with more depth. Research literature on imaging has been growing at an enormous pace, and it is not easy for anyone to keep track of every new development. The topic which is missing conspicuously in

this book is the interface of imaging with machine learning. This is an exciting area, and we debated a bit on whether we should include a new chapter on it. However, it was felt that covering this topic in an introductory manner and highlighting its implications for computational imaging deserves a much more detailed treatment. In any case, the framework discussed in this book will certainly provide useful background for anyone interested in applying machine learning principles to imaging systems.

A number of acknowledgments are due as this book writing project is getting completed. First and foremost, I would like to thank Prof. Nicholas George for getting me initiated into the wonderful world of imaging research at The Institute of Optics, University of Rochester, almost two decades ago while I worked with him for my PhD. He has been much more than an academic advisor to me and will be very happy to see this work getting published. Lively discussions with my senior colleagues Prof. Anurag Sharma, Prof. Joby Joseph, Prof. P. Senthilkumaran and Prof. D. S. Mehta at IIT Delhi on various topics relevant to the material covered in this book are also acknowledged. Even though the first edition of the book was available to us, writing the second expanded edition required much more time and energy than expected initially. I would like to thank my wife Sarita and daughters Rasika and Sanika for their patience and support while this edition of the book was being completed. Mansi and Sunaina would like to acknowledge the fellowship support from IIT Delhi and wholeheartedly thank their family and friends for their encouragement and continuous support over the years in pursuing higher education and research. This book was entirely written in the collaborative LaTeX environment Overleaf which has proved to be very helpful during the last 2 years as many times I could not personally meet with my co-authors due to lockdowns. The patience and support shown by Mr. Jai Raj Kapoor, Ane Books, New Delhi, while this project was underway are also very well appreciated.

New Delhi, India
December 2021

Kedar Khare
Mansi Butola
Sunaina Rajora

Preface to the First Edition

This book is an outgrowth of a series of lectures on Fourier Optics and Imaging that I have been teaching to the students at Indian Institute of Technology Delhi for the past few years. There are excellent books and other reference resources on various aspects of Optics, Photonics, Mathematics and Image Processing. However, when I started teaching students from various backgrounds and disciplines at IIT Delhi, I strongly felt the need of a single self-contained volume that covered the fundamental ideas required to appreciate the developments in Computational Imaging in a concise manner. The material in this book has been written at a level suitable for advanced undergraduate or post-graduate students who may wish to pursue research and development activities involving imaging systems for their future careers. A lot of people have contributed to my knowledge and understanding of imaging. First of all I thank my Ph.D. supervisor Prof. Nicholas George at Institute of Optics, University of Rochester, USA for introducing me to imaging as an important research area. I joined the Optics Ph.D. program in Rochester in 1999 with a strong background in basic Physics, thanks to some excellent mentors at Indian Institute of Technology Kharagpur. As a fresh Ph.D. student I was not at all clear about the research area where I could contribute my best. Prof. George provided me an opening into the exciting world of imaging and over the years he has been much more than just a Ph.D. advisor to me. Several years spent at the General Electric Global Research center in Niskayuna, NY further broadened my appreciation and expertise in relation to a wider range of practical imaging systems and I am thankful to my colleagues there for providing a great real-world learning opportunity. During the last few years I have been grateful to have company of several wonderful faculty colleagues at IIT Delhi who have helped me get started as an independent faculty member. Many thanks are also due to several enthusiastic students at IIT Delhi who make me think of new ways to understand the material discussed in this book and several allied topics. Finally I wholeheartedly thank my wife Sarita, my daughters Rasika and Sanika,

and all other family members in appreciation of their patience, encouragement and constant support over the last few of years while this manuscript was being prepared for publication.

IIT Delhi
June 2015

Kedar Khare

Contents

1	Introduction	1
1.1	Scope of Imaging Research	1
1.2	Computational Imaging Model	1
1.3	Organization of the Book	4
References		6
 Part I Mathematical Preliminaries		
2	Fourier Series and Transform	9
2.1	Fourier Series	9
2.2	Gibbs Phenomenon	11
2.3	Fourier Transform as a Limiting Case of Fourier Series	13
2.4	Sampling by Averaging, Distributions and Delta Function	15
2.4.1	Properties of Delta Function	17
2.5	Fourier Transform of Unit Step and Sign Functions	19
2.6	Fourier Transform of a Train of Delta Functions	21
2.7	Fourier Transform of a Gaussian Function	22
2.8	Fourier Transform of the Chirp Phase Function	23
2.9	Properties of Fourier Transform	24
2.10	Fourier Transform of the 2D Circ Function	26
2.11	Fourier Slice Theorem	27
2.12	Wigner Distribution	28
References		31
3	Sampling Theorem	33
3.1	Sampling Theorem via Poisson Summation Formula	34
3.1.1	Poisson Summation Formula for Bandlimited Signal	35
3.1.2	Additional Notes on the Sampling Formula	36
3.2	Sampling of Carrier-Frequency Signals	37
3.3	Degrees of Freedom in the Signal—Space-Bandwidth Product	38

3.4	Slepian (Prolate Spheroidal) Functions	39
3.4.1	Properties of Matrix $A^{(0)}$	41
3.4.2	Extrapolation of Bandlimited Functions	45
3.5	Band-Pass Analogues of Prolate Spheroidal Functions	46
	References	48
4	Operational Introduction to Fast Fourier Transform	51
4.1	Definition of Discrete Fourier Transform	51
4.2	Usage of 2D Fast Fourier Transform for Problems in Optics	53
	References	56
5	Linear System Formalism and Introduction to Inverse Problems in Imaging	59
5.1	Space-Invariant Impulse Response	60
5.2	Ill-Posedness of Inverse Problems	62
5.3	Inverse Filter	63
5.4	Wiener Filter	64
5.5	Generalized Wiener Filter	66
	References	70
6	Optimization Approach to Image Reconstruction	71
6.1	Image Denoising	71
6.1.1	Euler-Lagrange Problem in Variational Calculus	74
6.2	Functional Gradient for Complex-Valued Solutions	77
6.3	Image Deconvolution by Optimization	78
6.4	Compressive Imaging	81
6.4.1	Guidelines for Sub-sampled Data Measurement and Image Recovery	83
6.5	Optimization-Based Image Recovery Without a Free Parameter	86
6.6	Topics for Further Study	91
	References	93
7	Random Processes	95
7.1	Probability and Random Variables	95
7.1.1	Random Variables	97
7.1.2	Characteristic Function	99
7.1.3	Gaussian or Normal Distribution	101
7.2	Random Processes	103
7.2.1	Spectral Density: Wiener-Khintchine Theorem	106
7.2.2	Orthogonal Series Representation of Random Processes	107
7.3	Complex Representation of Random Processes	108
7.3.1	Mandel's Theorem on Complex Representation	110
	References	112

Part II Concepts in Optics

8 Geometrical Optics Essentials	115
8.1 Ray Transfer Matrix	116
8.2 Stops and Pupils	118
References	119
9 Wave Equation and Diffraction of Light	121
9.1 Review of Maxwell Equations	122
9.2 Weyl Representation of Spherical Waves	124
9.3 Angular Spectrum Method	125
9.4 Numerical Computation of Diffraction Fields Using Angular Spectrum Method	130
9.5 Fresnel and Fraunhofer Approximations	131
9.5.1 Transport of Intensity Equation (TIE)	133
9.5.2 Self-Imaging: Montgomery Conditions and Talbot Effect	136
9.5.3 Fractional Fourier Transform	137
9.6 Fraunhofer Diffraction	139
References	142
10 Coherence of Light Fields	143
10.1 Spatial and Temporal Coherence	143
10.1.1 Interference Law	144
10.2 van Cittert and Zernike Theorem	146
10.3 Space-frequency Representation of the Coherence Function	149
10.4 Intensity Interferometry: Hanbury Brown and Twiss Effect	150
10.5 Photon Counting Formula	152
10.6 Speckle Phenomenon	154
References	158
11 Polarization of Light	159
11.1 The Jones Matrix Formalism	159
11.2 The QHQ Geometric Phase Shifter	161
11.3 Degree of Polarization	162
11.4 Non-uniformly Polarized Light	163
References	164
12 Analysis of Optical Systems	165
12.1 Transmission Function for a Thin Lens	165
12.2 Fourier Transforming Property of a Thin Lens	167
12.3 Canonical Optical Processor	168
12.3.1 DC Block or Dark Field Imaging	169
12.3.2 Zernike's Phase Contrast Microscopy	170
12.3.3 Highlighting of Edges with Vortex Filter	172
12.3.4 Apodization Filters	172

12.4	Frequency Response of Optical Imaging Systems: Coherent and Incoherent Illumination	174
	References	178
13	Imaging from Information Point of View	179
13.1	Eigenmodes of a Canonical Imaging System	179
13.1.1	Eigenfunctions and Inverse Problems	182
	References	186

Part III Computational Imaging Systems

14	Digital Holography	189
14.1	Some Practical Aspects of Digital Holography Systems	191
14.1.1	In-line and Off-axis Configurations	191
14.1.2	Sampling of Digital Holograms	193
14.1.3	Numerical Aperture of the Hologram Recording System	194
14.2	Complex Object Wave Recovery in the Hologram Plane	194
14.2.1	Off-axis Digital Holography	194
14.2.2	Phase Shifting Digital Holography	198
14.2.3	Optimization Method for Complex Object Wave Recovery from Digital Holography	201
14.2.4	Noise Advantage Offered by the Optimization Method	203
14.3	Digital Holographic Microscopy	206
14.4	In-line Digital Holography of Particulates or Weak Scattering Objects	208
14.5	True 3D Image Reconstruction in Digital Holography	209
	References	215
15	Non-interferometric Phase Retrieval	217
15.1	Lensless Coherent X-ray Diffraction Imaging	218
15.2	Error Reduction Algorithm	219
15.3	Hybrid Input-Output (HIO) Algorithm	222
15.3.1	The Relaxed-Reflect-Reflect (RRR) Algorithm for Phase Retrieval	224
15.3.2	The Relaxed Average Alternating Reflection (RAAR) Algorithm for Phase Retrieval	225
15.4	Phase Retrieval Using Complexity Guidance	225
15.5	Phase Retrieval with Multiple Intensity Measurements: Fourier Ptychography	228
15.6	Phase Retrieval as an Optimization Problem	231
	References	233

Contents	xiii
16 Compact Multi-lens Imaging Systems	237
16.1 Compact Form Factor Computational Camera	237
16.2 Lightfield Cameras	240
16.2.1 The Concept of Lightfield	240
16.2.2 Recording the Lightfield Function with Microlens Array	242
References	246
17 PSF Engineering	249
17.1 Extending Depth of Focus from a Focal Stack	249
17.2 Extended Depth of Field with Cubic Phase Mask	252
17.3 Extended Depth of Focus Using the Log-Asphere Lens	255
17.4 Rotating Point Spread Functions	257
References	259
18 Structured Illumination Imaging	261
18.1 Forward Model and Image Reconstruction	262
18.2 Other Super-Resolution Microscopy Techniques	265
18.2.1 Simulated Emission Depletion Microscopy (STED)	265
18.2.2 Stochastic Optical Reconstruction Microscopy (STORM)	266
References	269
19 Image Reconstruction from Projections	271
19.1 X-ray Projection Data	272
19.2 Image Reconstruction from Projection Data	273
19.3 Holographic Tomography	276
References	279
20 Correlation Imaging	281
20.1 Single-Pixel Ghost Imaging	281
20.1.1 A Signal-Processing Viewpoint of Ghost Imaging	284
20.2 Imaging Through a Scattering Medium Using Speckle Memory Effect	286
References	289
Index	291

About the Authors

Kedar Khare is an Institute Chair Associate Professor at Optics and Photonics Centre and Department of Physics, Indian Institute of Technology Delhi. He additionally serves as an Associate Editor for the Journal of Modern Optics. Before joining IIT Delhi as a faculty member, he worked as a Lead Scientist for over five years at General Electric Global Research in Niskayuna, NY. Professor Khare's research interests span wide-ranging problems in computational imaging. In recent years he has been actively engaged in research on quantitative phase imaging, Fourier phase retrieval, computational microscopy, diagnostic MRI imaging, cryo-electron microscopy, and structured light propagation in turbulence. Prof. Khare received his MSc in Physics from the Indian Institute of Technology Kharagpur with the highest honors and his Ph.D. in Optics from The Institute of Optics, University of Rochester. Prof. Khare is a Senior Member of Optica.

Mansi Butola is currently a Ph.D. student at the Department of Physics, Indian Institute of Technology Delhi. Her research interests include computational imaging systems in general and her thesis work has focused on iterative phase retrieval algorithms for coherent diffraction imaging. She completed her B. Sc. (Honours) in Physics from Delhi University in 2014. After that in 2016, she received her M. Sc. in Physics from Garhwal University. She is a member of the Optica Student Chapter at IIT Delhi and has actively participated in organizing multiple seminars/workshops and outreach activities.

Sunaina Rajora is currently a Ph.D. student in the Department of Physics at the Indian Institute of Technology, Delhi. Her research interests include inverse problems in imaging, optimization algorithms, quantitative phase imaging, digital holography, and tomographic reconstruction. She received her B.Sc. and M.Sc. degrees in Physics from Aligarh Muslim University, India in 2013 and 2015, respectively. She is a member of the Optica Student Chapter at IIT Delhi and has actively participated in organizing seminars/workshops and outreach activities.

Chapter 1

Introduction



1.1 Scope of Imaging Research

Imaging systems form an integral part of our lives. Research and development efforts in diverse areas of science and technology over the last few decades have resulted in practical imaging devices that we often take for granted today. From cell phone cameras to advanced diagnostic imaging tools, from weather prediction based on satellite imagery to night vision devices meant for soldiers guarding our national frontiers, there are numerous examples of imaging systems that touch all aspects of our lives. Further, in current science and technological investigations dealing with sub-nano to astronomical length scales, imaging systems allow researchers to visualize natural phenomena or objects of interest and in effect, directly contribute to new discoveries. Imaging is therefore an active interdisciplinary research area that is relevant to a wide range of fundamental and applied problems.

The chart in Fig. 1.1 shows a wide variety of topics where imaging has become indispensable. It may not be too much of an exaggeration to say that every current science and technology student is very likely to encounter at least one of the topics listed in this chart during his/her career. While it is difficult to describe all the topics listed in Fig. 1.1 in detail in a single book, luckily for us the various imaging modalities share several common principles and mathematical ideas that go into their design. The goal of this book is to present these basic tools followed by a discussion of some specific imaging systems so as to provide the reader with sufficient background to enter the area of imaging research.

1.2 Computational Imaging Model

It is possible to describe a variety of imaging systems schematically by a simple model as shown in Fig. 1.2. A source of radiation produces waves (e.g. electromagnetic waves, sound waves) which then interact with the object of interest to be imaged.

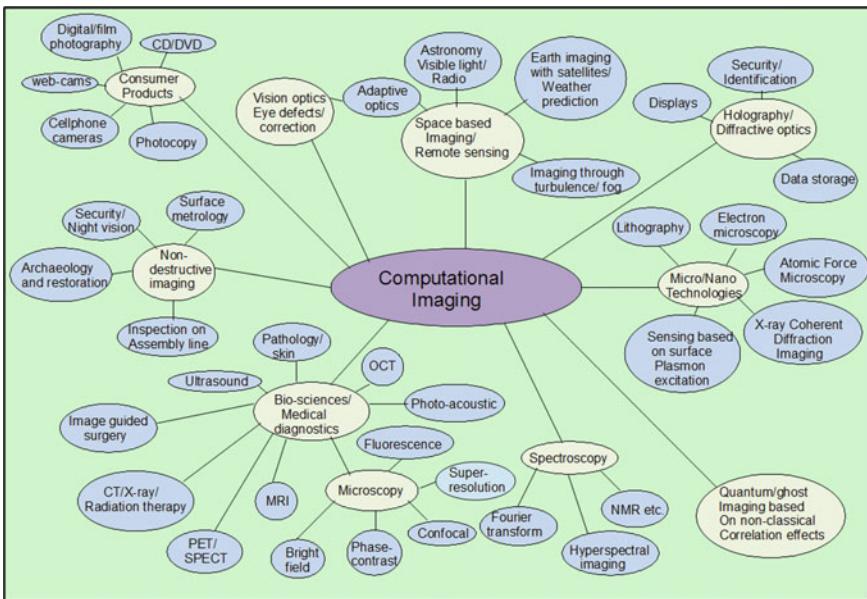


Fig. 1.1 Scope of imaging research and development

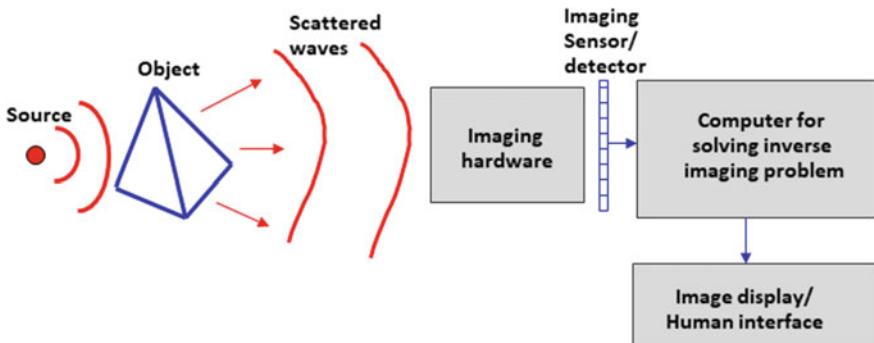


Fig. 1.2 Computational imaging model

The waves scattered from the object encode information about the object in the form of space-time variations of their intensity, phase, spectrum, polarization, coherence or other forms of correlations depending on the problem at hand. Note however that this multidimensional information in the scattered waves is not accessible to us with most traditional imaging devices as typical detectors at optical wavelengths respond only to the intensity of light waves. The central goal of an imaging system is to derive the maximum possible information that is encoded in the scattered waves so that a human observer may be able to visualize and interpret this information in the form of an image of the object. The numerical values of pixels in an image usually

represent some physical property of the object of interest. Typically, the imaging system hardware (e.g. lens assembly in a camera) is designed so as to modify or manipulate this coded information by applying some transformation to the scattered waves before they are detected by an imaging sensor. Traditional research in imaging systems involved efforts for improving the imaging hardware (e.g. lens design) so as to obtain the best possible visual images at the sensor. With the availability of computer processing power over the last couple of decades, imaging systems (e.g. those in Fig. 1.1) have increasingly become computational in nature [1–3]. In this new paradigm, the pattern recorded on the sensor may not have any visual similarity to the object of interest to be imaged. The imaging hardware is however designed such that, starting with the digitally recorded data at the imaging sensor, it is possible to solve an inverse or image reconstruction problem and to form an image of the object that is suitable for human interpretation. As imaging systems increasingly become important to scientific investigations and in our daily lives, it is often desirable to have systems with better resolution, faster imaging time, lesser sensitivity to noise, ability to capture multidimensional information (e.g. intensity and spectrum at each pixel), etc. Simple brute force scaling of traditional imaging systems for this purpose may however prove infeasible from technical as well as economic points of view. The computational imaging model has a significant role to play in this context in future. Computational imaging systems place equal emphasis on the imaging hardware part and the computational part for solving inverse imaging problems. This opens up new possibilities for obtaining unprecedented imaging performance. The current trend in various imaging system designs indicates that even the most familiar imaging tools like cameras, microscopes, telescopes, etc. can have a completely different physical form and information-gathering capabilities than what we are used to. These developments are exciting because they increase our technological capabilities to probe the nature on newer space-time scales and may lead to newer applications. Some aspects of computational imaging must therefore form a part of training for every serious science and technology student.

Distinction between image processing and computational imaging

Image processing typically involves algorithms that work on visually meaningful images. The post-processing algorithms improve the visual quality of the image or derive information of interest from the images. The processed output images can then be utilized for downstream decision-making tasks in the corresponding applications. In ***computational imaging***, the raw data recorded on a sensor is in an encoded form, and the computational algorithms are used to generate the image of interest. The raw data is not necessarily visually meaningful, but the reconstructed image is suitable for human interpretation. Additionally, computational imaging systems often estimate a whole set of new information about the object field like phase, polarization, coherence, etc. which cannot be directly measured by a sensor. Advancements in computational imaging have enabled us to explore the micro- and nano world with new imaging applica-

tions like super-resolution microscopy, cryo-electron microscopy and X-ray coherent diffraction imaging. The developments such as the reconstruction of the shadow of the M87* black hole using the Event Horizon Telescope also use the computational imaging framework.

1.3 Organization of the Book

This book is organized into three parts. The first part deals with the mathematical preliminaries that will be found useful later in the book. This material has been provided in order to make the book self-contained. The mathematical ideas presented here have not appeared in any single monograph on imaging to the best of the authors' knowledge. Discussing the mathematical ideas at the beginning will also help us in setting up a consistent notation for the rest of the book. In Chap. 2, we discuss the important tools of Fourier analysis that are used throughout the book. In Chap. 3, we present the basic ideas behind Shannon sampling theory and the concept of space-bandwidth product which is useful in quantifying information content in an image. The allied topic of the prolate spheroidal functions is also described, and an interesting connection between the sampling theorem and the prolate functions is discussed. Having introduced Fourier analysis and sampling theory, the next logical topic is the Fast Fourier Transform (FFT) presented in Chap. 4. FFT is an essential tool in almost all image-handling tasks. Instead of discussing the finer aspects of the FFT algorithm itself, we present an operational introduction to this topic which will help a reader to use FFT functions available in several standard numerical software packages effectively for simulating optics and imaging problems. Linear systems theory and associated inverse problems are essential for understanding the basic working principle of computational imaging systems. These topics are described in detail in Chap. 5. Several current computational imaging systems require image reconstruction methods that cannot be described as linear filtering operations but rather as constrained optimization problems. Some basic ideas in constrained optimization are described in Chap. 6. This chapter also includes more advanced topics such as compressive imaging that are likely to play a major role in future imaging system designs. The solution of optimization problems using a novel method of mean gradient descent that does not require empirical parameter tuning is also presented in the context of the image deconvolution problem. The last chapter of the first part presents a discussion of random processes. Conceptual treatment of imaging systems is often presented by means of deterministic equations. The generation, propagation and detection of light waves are however inherently statistical in nature. The understanding of random processes is therefore essential for taking up any imaging system design problem. The topics discussed here find direct utility in later parts of the book.

The second part of the book discusses some important concepts in Optics that are key to understanding the working of imaging systems and their capabilities or

limitations. Unlike many well-known books on Optics, the topics in this part are not covered from a historical perspective but are presented in a way that is more suitable for developing an understanding of imaging phenomena from a systems perspective. We begin this part with a description of essential concepts of ray or geometrical optics in Chap. 8 followed by details of the wave equation and diffraction phenomena in Chap. 9. Diffraction theory is essential to understand how light waves can carry information from the object of interest to the imaging system. Using the Fourier analysis tools from Chap. 2, this chapter directly introduces the Weyl representation for the spherical waves which is useful for developing the angular spectrum approach to diffraction. The exact Rayleigh-Sommerfeld-Smythe vectorial relations for the description of the diffraction problem are also presented in this chapter. Next, we discuss approximations to the exact diffraction relations, viz. Fresnel and Fraunhofer approximations, which are found to be useful in the analysis or design of several common practical system setups. The important ideas related to the coherence of light fields are described in Chap. 10. While the reader may not appreciate the reason for discussing this topic at this stage, the author believes that without the knowledge of coherence theory, one's knowledge of Optics is incomplete. This is especially so if someone is interested in designing novel futuristic imaging systems. This discussion is followed by Chap. 11 on polarization of light. Polarization is an additional degree of freedom possessed by light waves, although it has not been utilized to its fullest potential in imaging system design so far. Next in Chap. 12, we use the ideas developed in the second part of the book to show the reader how one can model simple optical systems involving free-space propagation, lens elements or other phase masks. After covering the important cases of the optical Fourier transform and the canonical or 4F imaging systems configurations, we discuss several practical systems that utilize these ideas. Concepts that help one evaluate an imaging system performance such as the optical transfer function (OTF) are also discussed. Chapter 13 which is the last chapter of the second part of the book presents a general treatment of imaging phenomena from an information point of view. A simple treatment based on the sampling theorem is used here to provide some interesting insights into the information-carrying capacity of imaging systems from object to the image space.

The third part of the book builds on the material developed in the first two parts to discuss some specific computational imaging systems. This part is intended to give the readers a feeling for the current research ideas in computational imaging. In Chaps. 14 and 15, we discuss one of the most important computational problems of interferometric and non-interferometric phase imaging. The phase of light waves scattered by an object encodes critical information about the object. The phase information is however not available to us directly and the phase can only be inferred computationally. Both chapters discuss system configurations, typical well-known algorithms and how newer ideas on image sparsity can significantly enhance the performance of phase reconstruction algorithms. Chapter 15 additionally presents a discussion on an important super-resolution configuration called Fourier ptychography. In Chap. 16, we discuss some novel compact multi-lens systems that outperform traditional cameras by combining unconventional design and computational image recovery. In particular, thin cameras and the plenoptic cameras are discussed. In

Chap. 17, we describe pupil function engineering approaches for extended depth of field imaging. The approach used here involves recording images that are blurred by nearly depth invariant impulse response which can then be used for generating all-in-focus images that are not possible to record directly with any conventional camera system. Microscopic imaging beyond traditional diffraction limit has become an essential part of basic bio-sciences research, and we discuss a powerful computational imaging modality that uses structured illumination for super-resolution imaging in Chap. 18. In Chap. 19, we present some basic ideas in image reconstruction from projections—a topic which is of considerable importance to the working of several diagnostic imaging systems as well as cryo-electron microscopy. Finally in Chap. 20, we describe a somewhat unusual imaging system setup that uses a single pixel detector for generating high-resolution images exploiting the correlations in specially designed optical beams. Another interesting correlation imaging method using the speckle memory effect is also discussed briefly.

Optical imaging is rich and arguably offers the most physical properties that provide a diversity of information. Non-optical imaging modalities often explore only a subset of them. We therefore believe that the material in this book will allow a reader to explore and appreciate the working of a wide variety of imaging modalities. The material covered in the first two parts and in some selected topics from the third part can constitute sufficient material for a semester-long course suitable for academic programs in optics and photonics.

References

1. A.F. Coskun, A. Ozcan, Computational imaging, sensing and diagnostics for global health applications. *Curr. Opin. Biotechnol.* **25**, 8–16 (2014)
2. J.N. Mait, G.W. Euliss, R.A. Athale, Computational imaging. *Adv. Opt. Photon.* **10**, 409–483 (2018)
3. Y. Altmann, S. McLaughlin, M.J. Padgett, V.K. Goyal, A.O. Hero, D. Faccio, Quantum-inspired computational imaging. *Science* **361**, 6403 (2018)

Part I

Mathematical Preliminaries

Chapter 2

Fourier Series and Transform



The ideas of Fourier series and transform play a central role in our mathematical treatment of imaging phenomena [1] and the design of image reconstruction algorithms. While these topics are extensively discussed in excellent monographs on the subject [2–4], we cover them here in detail for completeness.

2.1 Fourier Series

A periodic function $g(x)$ with period T such that

$$g(x) = g(x + T), \quad -\infty < x < \infty \quad (2.1)$$

may be represented as a Fourier series:

$$g(x) = \sum_{n=-\infty}^{\infty} G_n \exp(i2\pi nx/T). \quad (2.2)$$

This is an important idea as we shall see when studying linear systems. The question of when such an expansion exists is addressed in the Dirichlet sufficiency conditions:

1. The function $g(x)$ must be absolutely integrable over one period.
2. The function $g(x)$ must be piecewise continuous. A finite number of finite discontinuities is allowed.
3. The function $g(x)$ must have a finite number of extrema in one period. Something like $\sin(1/x)$ near $x = 0$ is not allowed.

The coefficient G_n may be determined using the following orthogonality relation:

$$\begin{aligned} \int_{-T/2}^{T/2} dx \exp[i2\pi(m-n)x/T] &= \left[\frac{\exp[i2\pi(m-n)x/T]}{i2\pi(m-n)/T} \right]_{x=-T/2}^{T/2} \\ &= T \frac{\sin[\pi(m-n)]}{\pi(m-n)} = T \delta_{m,n}. \end{aligned} \quad (2.3)$$

The coefficients G_n can therefore be obtained as

$$G_n = \frac{1}{T} \int_{-T/2}^{T/2} dx g(x) \exp(-i2\pi nx/T). \quad (2.4)$$

If $g(x)$ has a (finite) jump discontinuity at $x = x_0$, the series expansion converges to

$$g(x_0) = \left[\frac{g(x_{0-}) + g(x_{0+})}{2} \right]. \quad (2.5)$$

We illustrate the Fourier series expansion for a square wave of period T defined as

$$\begin{aligned} g(x) &= g(x+T), \\ g(x) &= 0, \quad x \in (-T/2, 0) \\ &= 1, \quad x \in (0, T/2). \end{aligned} \quad (2.6)$$

The Fourier coefficients are given by

$$\begin{aligned} G_n &= \frac{1}{T} \int_0^{T/2} dx \exp(-i2\pi nx/T) \\ &= \frac{1}{T} \left[\frac{\exp(-i2\pi nx/T)}{-i2\pi n/T} \right]_{x=0}^{x=T/2} = \frac{1 - \exp(-i\pi n)}{i2\pi n}. \end{aligned} \quad (2.7)$$

The Fourier coefficients may be listed as follows:

$$G_0 = \frac{1}{2}, \quad G_2 = G_{-2} = 0, \quad G_4 = G_{-4} = 0, \dots \text{so on.} \quad (2.8)$$

$$G_1 = \frac{1}{i\pi}, \quad G_{-1} = \frac{1}{-i\pi}, \quad G_3 = \frac{1}{3i\pi}, \quad G_{-3} = \frac{1}{-3i\pi}, \dots \text{so on.} \quad (2.9)$$

In general, all odd-order terms survive and all even-order terms except the $n = 0$ term vanish. Combining the non-vanishing terms, we can write the Fourier series expansion as

$$g(x) = \frac{1}{2} + \frac{2}{\pi} \sin\left(\frac{2\pi x}{T}\right) + \frac{2}{3\pi} \sin\left(\frac{6\pi x}{T}\right) + \dots \quad (2.10)$$

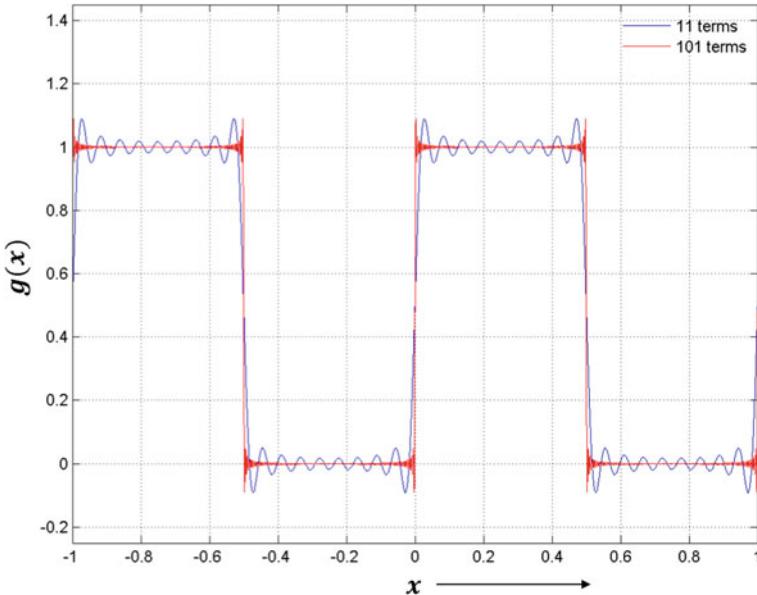


Fig. 2.1 Fourier series representation for a square wave with 11 and 101 terms as in Eq. (2.10) are shown by the blue and the red plots, respectively

Plots of the series for the first 11 and 101 terms of Eq. (2.10) are shown in Fig. 2.1. The plots show how the series converges to the square wave as the number of terms in the series representation increase. The period $T = 1$ has been used here.

2.2 Gibbs Phenomenon

We note a peculiar phenomenon which arises near the discontinuity of a periodic function that is being represented by means of the Fourier series. The Fourier series representation in Eq. (2.10) for square wave may be expressed as

$$g(x) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{1}{(2n+1)} \sin \left[\frac{2\pi(2n+1)x}{T} \right]. \quad (2.11)$$

When a finite number of terms is included in the summation above, the left-hand side has a discontinuity while the right-hand side is a sum of continuous functions. The convergence of the series sum to the periodic square wave is therefore not a point-wise convergence (near the discontinuity, one observes undershoot and overshoot) but uniform convergence. In fact, the overshoot and undershoot do not die out as the number of terms in the partial series sum increases. This interesting feature is

known by the name Gibbs phenomenon. The under and overshoot get closer to the discontinuity with an increasing number of terms such that the area under them tends to be zero. This type of convergence is termed as uniform convergence or “almost everywhere” convergence (convergence everywhere except on sets of measure zero). We may express the uniform convergence property as follows:

$$\lim_{N \rightarrow \infty} \left\| g(x) - \frac{1}{2} - \sum_{n=0}^N \frac{1}{(2n+1)} \sin \left[\frac{2\pi(2n+1)x}{T} \right] \right\|_2^2 = 0. \quad (2.12)$$

The notation $\| \dots \|_2$ above for the L_2 -norm is to be understood as

$$\|g(x)\|_2 = \left[\int_{-\infty}^{\infty} dx |g(x)|^2 \right]^{1/2}. \quad (2.13)$$

As stated above, the uniform convergence implies that $g(x)$ and its approximation need not coincide at every point. The overshoot near the point of discontinuity can be estimated using a Fourier series with a finite number of terms. Let us consider the partial sum:

$$S_N(x) = \sum_{n=0}^N \frac{1}{(2n+1)} \sin \left[\frac{2\pi(2n+1)x}{T} \right]. \quad (2.14)$$

Taking a derivative of this partial sum with respect to x gives

$$S'_N(x) = \frac{2\pi}{T} \sum_{n=0}^N \cos \left[\frac{2\pi(2n+1)x}{T} \right]. \quad (2.15)$$

Summing of this series requires a trick. We note that the consecutive terms of the series have arguments inside the $\cos(\dots)$ function that differ by $(\frac{4\pi x}{T})$ and so we multiply both sides of the equation by $\sin(\frac{4\pi x}{T})$ and use the relation $[2 \cos(A) \sin(B) = \sin(A + B) - \sin(A - B)]$. Several terms now cancel out in the summation and what remains is

$$\begin{aligned} S'_N(x) \sin \left(\frac{4\pi x}{T} \right) &= \frac{\pi}{T} \left\{ \sin \left[\frac{2\pi(2N+3)x}{T} \right] + \sin \left[\frac{2\pi(2N+1)x}{T} \right] \right\} \\ &= \frac{2\pi}{T} \sin \left[\frac{2\pi(2N+2)x}{T} \right] \cos \left(\frac{2\pi x}{T} \right). \end{aligned} \quad (2.16)$$

The derivative can thus be expressed as

$$S'_N(x) = \left(\frac{\pi}{T} \right) \frac{\sin \left[\frac{4\pi(N+1)x}{T} \right]}{\sin \left(\frac{2\pi x}{T} \right)}. \quad (2.17)$$

As we move from the discontinuity at $x = 0$ to right on the x -axis, the first maximum of the partial sum in Eq. (2.14) occurs at the place where the derivative is zero for the first time (one may verify this by going to second derivatives and checking the sign) at the point: $x_0 = T/[4(N + 1)]$. The partial sum at this point gives us an estimate of the numerical value of the overshoot:

$$\begin{aligned} S_N(x_0) &= \sum_{n=0}^N \frac{1}{(2n+1)} \sin \left[\frac{\pi(2n+1)}{2(N+1)} \right] \\ &= \frac{\pi}{2} \sum_{n=0}^N \frac{1}{(N+1)} \operatorname{sinc} \left[\frac{n+\frac{1}{2}}{N+1} \right], \end{aligned} \quad (2.18)$$

where we have used the definition:

$$\operatorname{sinc}(x) = \frac{\sin \pi x}{\pi x}. \quad (2.19)$$

For large N , the partial sum $S_N(x_0)$ may be approximated by the integral of the sinc function on interval $(0, 1)$. We therefore write

$$S_N \left(\frac{T}{4(N+1)} \right) \approx \frac{\pi}{2} \int_0^1 dx \operatorname{sinc}(x). \quad (2.20)$$

Plugging back into the Fourier series for the square in Eq. (2.11) wave, we have

$$g \left(\frac{T}{4(N+1)} \right) \approx \frac{1}{2} + \int_0^1 dx \operatorname{sinc}(x) \approx 1.0895. \quad (2.21)$$

The graph of function thus shoots above the square wave by approximately 9% even as a large number of terms are added to the series. It is to be noted that this maximum occurs closer and closer to zero (at $x_0 = \frac{T}{4(N+1)}$) as the number of terms in the partial sum increases. We shall see later in the book that the Gibbs phenomenon manifests itself in the form of ringing artifacts in images when any filtering operation is performed where the filter has a sharp edge in the frequency space.

2.3 Fourier Transform as a Limiting Case of Fourier Series

We will now consider the limiting case of the Fourier series as the period T goes to infinity which essentially is the same as saying that $g(x)$ is not periodic. We denote the discrete frequencies as

$$f_{xn} = \frac{n}{T}. \quad (2.22)$$

The difference between the consecutive discrete frequencies is given by $\Delta f_x = f_{x(n+1)} - f_{xn} = 1/T$. Further, we define

$$G(f_{xn}) = \int_{-\infty}^{\infty} dx g(x) \exp(-i2\pi f_{xn}x). \quad (2.23)$$

We may write the Fourier series expansion as

$$\begin{aligned} g(x) &= \sum_{n=-\infty}^{\infty} G_n \exp(i2\pi f_{xn}x) \\ &= \sum_{n=-\infty}^{\infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} dx' g(x') \exp(-i2\pi f_{xn}x') \right] \exp(i2\pi f_{xn}x). \end{aligned} \quad (2.24)$$

In the limit $T \rightarrow \infty$ we have

$$g(x) = \sum_{n=-\infty}^{\infty} G(f_{xn}) \exp(i2\pi f_{xn}x) \Delta f_x \rightarrow \int_{-\infty}^{\infty} df_x G(f_x) \exp(i2\pi f_x x). \quad (2.25)$$

The functions $g(x)$ and $G(f_x)$ form what is called the Fourier transform pair:

$$\begin{aligned} G(f_x) &= \int_{-\infty}^{\infty} dx g(x) \exp(-i2\pi f_x x), \\ g(x) &= \int_{-\infty}^{\infty} df_x G(f_x) \exp(i2\pi f_x x). \end{aligned} \quad (2.26)$$

The Fourier transform pair will often be denoted by the following notation: $G(f_x) = \mathcal{F}\{g(x)\}$ and $g(x) = \mathcal{F}^{-1}\{G(f_x)\}$. Once again, this definition of Fourier transform is meaningful under the Dirichlet sufficiency conditions that are similar to those for the Fourier series:

1. $g(x)$ is absolutely integrable. $\int dx |g(x)| < \infty$.
2. $g(x)$ has finite number of finite discontinuities. No infinite discontinuities are allowed.
3. $g(x)$ can have only a finite number of extrema in a finite interval on x -axis.

What do we mean by $G(f_x)$? We observe from Eq. (2.26) that $g(x)$ is a linear combination of elementary functions of the form $\exp(i2\pi f_x x)$ and $G(f_x)$ is the weight associated with a particular frequency f_x in the expansion for $g(x)$. The description in terms of $G(f_x)$ can thus be considered as an equivalent alternative description for the function $g(x)$. We will now consider a simple example of Fourier transform and introduce some notation along the way. Consider the rect function or the rectangle distribution defined as

$$g(x) = \text{rect}\left(\frac{x}{2L}\right) = \begin{cases} 1 & |x| < L \\ 1/2 & |x| = L \\ 0 & |x| > L. \end{cases} \quad (2.27)$$

The Fourier transform of the rect function may be evaluated as

$$G(f_x) = \int_{-L}^L dx \exp(-i2\pi f_x x) = 2L \text{sinc}(2Lf_x).$$

Here the sinc function is as defined in Eq. (2.19). Note that although $\text{rect}(x)$ has a discontinuity, its transform is continuous. Further, it is somewhat surprising to know that $|\text{sinc}(f_x)|$ is not absolutely integrable to have a Fourier transform or Fourier inverse in the conventional sense required by the Dirichlet sufficiency conditions. To show this, consider the intervals along f_x axis where $|\sin(\pi f_x)| \geq 0.5$. These intervals are given by $f_x \in [n + 1/6, n + 5/6]$. We therefore have

$$\int_{-\infty}^{\infty} df_x |\text{sinc}(f_x)| > 2 \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2/3}{n + 5/6} \rightarrow \infty. \quad (2.28)$$

The series on the right-hand side diverges and as a result the absolute integral of $\text{sinc}(f_x)$ does not exist. Functions such as spikes, steps and even sines and cosines do not have a Fourier transform in traditional theory. Defining Fourier transforms for such functions is however a practical necessity when representing images as 2D matrices in digital form. In fact, edges or spikes contain the most important visual information in images. The traditional Fourier transform theory must therefore be extended to take these cases into account. We will study the important case of the Dirac delta function in this context. This class of functions (spikes, steps, etc.) with no Fourier transform in conventional theory is known by the name generalized functions. We will not deal with the theory of generalized functions in detail but study some specific cases of interest starting with the Dirac delta function.

2.4 Sampling by Averaging, Distributions and Delta Function

We will now learn the important idea of the delta function that will be useful later in this book. Sampling a continuous function is a part of every measurement or digitization process. Suppose we have a signal $I(x)$ —say the intensity of ambient light along a line on a screen—that we wish to sample at a discrete set of points x_1, x_2, \dots . How do we go about it? Every detector we can possibly use to measure the intensity will have some finite width $2L$ and a sample of $I(x)$ will be an average over this detector area. So we may write the intensity at point $x = 0$, or $I(0)$ as

$$I(0) \approx \frac{1}{2L} \int_{-\infty}^{\infty} dx I(x) \operatorname{rect}\left(\frac{x}{2L}\right), \quad (2.29)$$

where the $\operatorname{rect}(\dots)$ function is defined as

$$\operatorname{rect}(x) = \begin{cases} 1, & |x| < 1/2 \\ \frac{1}{2}, & |x| = 1/2 \\ 0, & |x| > 1/2. \end{cases} \quad (2.30)$$

Now how do we improve this approximation so that we go to the ideal $I(0)$? Clearly, we have to reduce the size $2L$ over which the average is carried out. We may say that

$$I(0) = \lim_{2L \rightarrow 0} \frac{1}{2L} \int_{-L}^L dx I(x) \operatorname{rect}\left(\frac{x}{2L}\right). \quad (2.31)$$

Notice that as the length $2L \rightarrow 0$, the width of the function $\frac{1}{2L} \operatorname{rect}\left(\frac{x}{2L}\right)$ keeps reducing whereas its height keeps increasing such that the area under the curve is unity. This limiting process leads us to an impulse which is also commonly known by the name delta function. We may write

$$\delta(x) = \lim_{2L \rightarrow 0} \frac{1}{2L} \operatorname{rect}\left(\frac{x}{2L}\right). \quad (2.32)$$

Although it is commonly referred to as the “delta function” and we often call it that way, you will appreciate that it is not a function in the usual sense. When we say $f(x) = x^2$, we are associating a value for every input number x . The impulse or delta distribution is more of an idea that is the result of a limiting process. Anything that is equal to zero everywhere in the limit except at $x = 0$, where it tends to infinity, cannot be a function in the sense you may have learnt in your mathematics classes. It is not continuous nor is it differentiable in the usual way. What can it do? Given a continuous signal $g(x)$, it can sample it at any given point:

$$\int_{-\infty}^{\infty} dx g(x) \delta(x - x') = g(x'). \quad (2.33)$$

So unlike a function which takes a number as an input and produces another number as its output, the delta function takes a function as its input and produces a number $g(x')$ as its output. It is in this sense that the terminology “generalized function” for the delta function is to be understood. Another important point to note is that the delta function does make sense when used under the integral sign. For example, the sampling property above is a perfectly meaningful result. We will define $\delta(x)$ as

$$\begin{aligned} \int_{-\infty}^{\infty} dx \delta(x) &= 1 \\ \int_{-\infty}^{\infty} dx g(x) \delta(x - x') &= g(x'). \end{aligned} \quad (2.34)$$

What is the Fourier transform of $\delta(x)$? We have already seen that

$$\mathcal{F}\left\{\frac{1}{2L}\text{rect}\left(\frac{x}{2L}\right)\right\} = \text{sinc}(2Lf_x). \quad (2.35)$$

The first zero of the sinc function is at $2Lf_x = 1$. When $2L \rightarrow 0$, the first zero effectively occurs at infinity meaning that in the Fourier transform space:

$$\mathcal{F}\{\delta(x)\} = 1 \text{ for all } f_x. \quad (2.36)$$

Mathematically we may express this as

$$\delta(x) = \int_{-\infty}^{\infty} df_x \exp(i2\pi f_x x). \quad (2.37)$$

Note again that the integral is true in the sense that it represents the idea of an impulse. It should be clear now that an impulse requires equal contribution from all the frequencies. The idea of an impulse is also used often in imaging literature to denote an ideal point source. Since any object of interest can be considered as a collection of delta functions of varying strengths, the notion of the impulse will be seen to be very useful to us in later discussions in the book.

2.4.1 Properties of Delta Function

- Sampling property:** At points of continuity of a function $g(x)$ we have

$$\int_{-\infty}^{\infty} dx g(x) \delta(x - x') = g(x'). \quad (2.38)$$

If at $x = x'$ the function $g(x)$ has a finite jump discontinuity, the right-hand side of the above equation is an average value of the two limits $g(x'_+)$ and $g(x'_-)$.

- Derivatives of delta function:** All operations with delta function are to be associated with a test function under the integral sign. We evaluate the integral below by parts:

$$\begin{aligned}
& \int_{-\infty}^{\infty} dx g(x) \frac{d}{dx} \delta(x - x') \\
&= g(x) \delta(x - x')|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \frac{d}{dx} g(x) \delta(x - x') \\
&= -g'(x').
\end{aligned} \tag{2.39}$$

This property applies to multiple-order derivatives. So continuing along the lines of the above equation, we get

$$\int_{-\infty}^{\infty} dx g(x) \frac{d^n}{dx^n} \delta(x - x') = (-1)^n g^{(n)}(x'), \tag{2.40}$$

where $g^{(n)}(x)$ is the n-th order derivative of $g(x)$.

3. **Scaling property:** First of all, we note that the delta function is even: $\delta(-x) = \delta(x)$. This leads to

$$\begin{aligned}
\int_{-\infty}^{\infty} dx g(x) \delta(ax) &= \frac{1}{|a|} \int_{-\infty}^{\infty} dx g\left(\frac{x}{a}\right) \delta(x) \\
&= \frac{1}{|a|} g(0).
\end{aligned} \tag{2.41}$$

We may therefore write $\delta(ax) = \frac{1}{|a|} \delta(x)$.

4. **Delta function having another function as an argument:** What is the meaning of $\delta(g(x))$? We suppose that $g(x)$ has a zero at $x = x_1$. If $g(x)$ is a well-behaved test function, it may be expanded near the zero as $g(x_1) = (x - x_1) g'(x_1)$. (If the first derivative is zero, we may use the second or higher derivatives.) The delta function may then be expressed using the scaling property as

$$\delta(g(x)) = \frac{\delta(x - x_1)}{|g'(x_1)|}. \tag{2.42}$$

This equation may be generalized as a summation of similar terms near all the zeros of the function $g(x)$.

5. **Relation between delta function and the unit step:** Consider a function represented as

$$u(x) = \int_{-\infty}^x dt \delta(t). \tag{2.43}$$

By inspection, we observe that when $x < 0$, the integral is zero when $x > 0$ the values of the integral are equal to 1 for all x . The integral thus represents a step function defined as

$$u(x) = \begin{cases} 0 & x < 0 \\ 1/2 & x = 0 \\ 1 & x > 0. \end{cases} \tag{2.44}$$

We may now write the integral relation in reverse and observe that the derivative of the unit step is the delta function.

$$\frac{du(x)}{dx} = \delta(x). \quad (2.45)$$

In the next few sections, we will obtain Fourier transforms for some important functions of interest throughout this book.

2.5 Fourier Transform of Unit Step and Sign Functions

The Fourier transform of the unit step function is an interesting case to study. We have already seen that the derivative of a unit step is the delta function. We can rewrite this relation more generally as

$$\frac{d[u(x) + c]}{dx} = \delta(x), \quad (2.46)$$

since the derivative is not affected by the constant c . We denote the Fourier transform of $u(x)$ by $U(f_x)$ and write the derivative relation as

$$\frac{d}{dx} \int_{-\infty}^{\infty} df_x [U(f_x) + c \delta(f_x)] \exp(i2\pi f_x x) = \delta(x). \quad (2.47)$$

Taking the derivative operation inside the integral sign, we get

$$\int_{-\infty}^{\infty} df_x \{(i2\pi f_x)[U(f_x) + c \delta(f_x)] - 1\} \exp(i2\pi f_x x) = 0. \quad (2.48)$$

Since the integral is identically zero for all x , we may conclude that

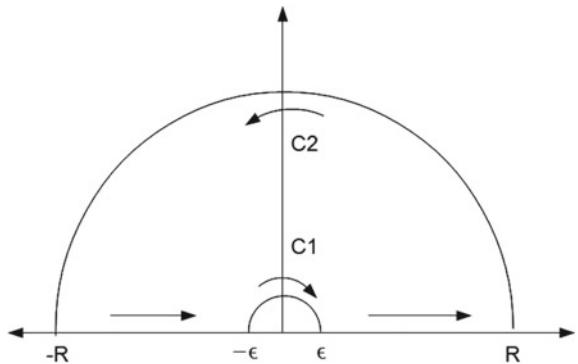
$$U(f_x) + c \delta(f_x) = \frac{1}{i2\pi f_x}. \quad (2.49)$$

Now all that remains is to determine the constant c . Taking the inverse Fourier transform of the above equation gives

$$u(x) + c = \int_{-\infty}^{\infty} df_x \frac{\exp(i2\pi f_x x)}{i2\pi f_x}. \quad (2.50)$$

The integrand on the right-hand side has a pole at $f_x = 0$ on the real line and the integral is to be understood as the Cauchy principal value. The contour used for integration is shown in Fig. 2.2. For $x > 0$, the appropriate semicircular path to be

Fig. 2.2 Contour used for complex integration in Eq. (2.50) with indent at the origin



selected for integration is in the upper half plane as per Jordan's lemma. For $x > 0$ we have

$$\int_{-R}^{-\epsilon} + \int_{C1} + \int_{\epsilon}^R + \int_{C2} = 0. \quad (2.51)$$

The integral over $C1$ gives $(-i\pi)/(i2\pi) = -1/2$ as the curve is traversed clockwise. As $R \rightarrow \infty$ the integral over $C2$ vanishes for $x > 0$. Thus, the principal value of the integral by letting $\epsilon \rightarrow 0$ is given by

$$P \int_{-\infty}^{\infty} df_x \frac{\exp(i2\pi f_x x)}{i2\pi f_x} = \frac{1}{2}, \quad x > 0. \quad (2.52)$$

In Eq. (2.50), we may now set $x > 0$ to get

$$1 + c = \frac{1}{2} \Rightarrow c = -\frac{1}{2}. \quad (2.53)$$

The Fourier transform of the unit step is thus given by

$$U(f_x) = \frac{1}{i2\pi f_x} + \frac{1}{2} \delta(f_x). \quad (2.54)$$

Once we find the transform of the unit step as above, we may evaluate the transform of the sign function $\text{sgn}(x)$ function very easily. We define

$$\text{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0. \end{cases} \quad (2.55)$$

Note that $\text{sgn}(x) = 2u(x) - 1$ so the transform:

$$\mathcal{F}\{\text{sgn}(x)\} = 2U(f_x) - \delta(f_x) = \frac{1}{i\pi f_x}.$$

Once again, the inverse transform integral is to be understood as the Cauchy principal value.

2.6 Fourier Transform of a Train of Delta Functions

We define the comb function or a periodic train of delta functions as

$$\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n). \quad (2.56)$$

The name “comb” follows from the fact that an inverted comb has a look of a train of periodically located pulses. The periodicity of the comb function allows us to represent it as a (generalized) Fourier series with period $T = 1$:

$$\text{comb}(x) = \sum_{k=-\infty}^{\infty} c_k \exp(i2\pi kx). \quad (2.57)$$

The series coefficients c_k may be evaluated as

$$c_k = \int_{-1/2}^{1/2} dx \text{comb}(x) \exp(-i2\pi kx) = 1. \quad (2.58)$$

All the Fourier series coefficients are seen to be identically equal to 1. The Fourier transform of the comb function can now be evaluated in a straightforward manner:

$$\mathcal{F}\{\text{comb}(x)\} = \sum_{k=-\infty}^{\infty} \delta(f_x - k) = \text{comb}(f_x). \quad (2.59)$$

We observe the interesting self-Fourier property that the Fourier transform of a comb function is another comb function. We will now consider another self-Fourier function.

2.7 Fourier Transform of a Gaussian Function

Before evaluating the Fourier transform of the Gaussian, we evaluate a useful integral:

$$J = \int_{-\infty}^{\infty} dx \exp(-\pi x^2). \quad (2.60)$$

We note that it is easy to evaluate J^2 in 2D polar coordinates:

$$\begin{aligned} J^2 &= \int \int dx dy \exp[-\pi(x^2 + y^2)] = \int_0^{2\pi} d\theta \int_0^{\infty} dr r \exp(-\pi r^2) \\ &= (2\pi) \frac{1}{2\pi} = 1. \end{aligned} \quad (2.61)$$

This gives an interesting result:

$$J = \int_{-\infty}^{\infty} dx \exp(-\pi x^2) = 1. \quad (2.62)$$

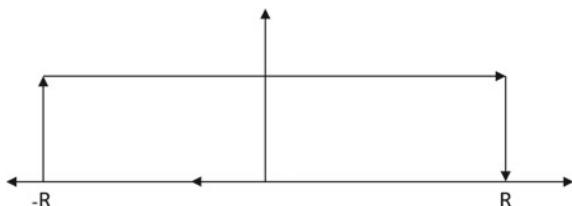
We will now evaluate the Fourier transform of a Gaussian function:

$$\begin{aligned} \int_{-\infty}^{\infty} dx \exp(-\pi x^2 - i2\pi f_x x) &= \int_{-\infty}^{\infty} dx \exp[-\pi(x + if_x)^2] \exp(-\pi f_x^2) \\ &= \exp(-\pi f_x^2) \int_{-\infty+if_x}^{\infty+if_x} du \exp(-\pi u^2). \end{aligned} \quad (2.63)$$

The contour to be used for evaluating the transform is shown in Fig. 2.3. The appropriate contour is in the upper half plane for $f_x > 0$ (as shown in Fig. 2.3) and in the lower half plane when $f_x < 0$. As $R \rightarrow \infty$, the contribution of the vertical parts of the integration paths vanishes as $\exp(-\pi(\pm R + iy)^2) \rightarrow 0$. The integral over the real line gives -1 (direction from $+$ to $-$). The integral we want to evaluate is thus equal to $+1$ since the integrand is analytic on and inside the contour, and we have

$$\mathcal{F}\{\exp(-\pi x^2)\} = \exp(-\pi f_x^2). \quad (2.64)$$

Fig. 2.3 Contour used for evaluating Fourier transform of Gaussian



2.8 Fourier Transform of the Chirp Phase Function

The Fourier transform of quadratic or chirp phase function is an important result as we shall see later in the context of Fresnel diffraction. The function of interest is $\exp(i\pi x^2)$. It is called a “chirp” since the rate of change of the instantaneous phase of the signal is linear similar to the chirping of birds. The Fourier transform of the chirp function is evaluated as follows:

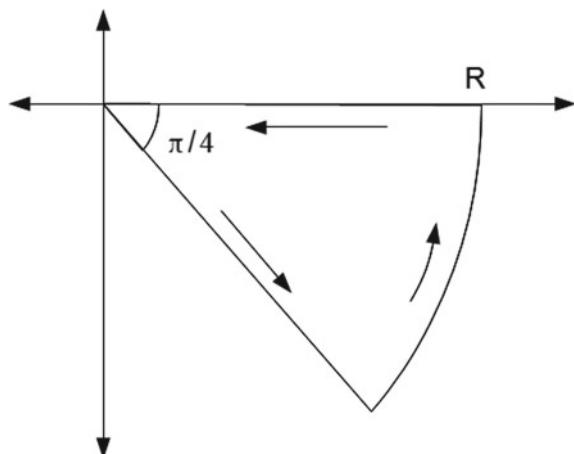
$$\begin{aligned} G(f_x) &= \int_{-\infty}^{\infty} dx \exp(i\pi x^2 - i2\pi f_x x) = \exp(-i\pi f_x^2) \int_{-\infty}^{\infty} dx \exp(i\pi x^2) \\ &= 2 \exp(-i\pi f_x^2) \int_0^{\infty} dx \exp(i\pi x^2). \end{aligned} \quad (2.65)$$

Since we know the integral of $\exp(-\pi t^2)$ on the real line, let us make a substitution:

$$-\pi z^2 = i\pi x^2 \Rightarrow z = \exp(-i\pi/4)x. \quad (2.66)$$

As x goes from 0 to ∞ , the integration path for z goes along a radial line through the origin at an angle $(-\pi/4)$ with the x-axis. We will now evaluate the integral of $\exp(-\pi z^2)$ along the 45-degree wedge contour shown in Fig. 2.4. For the integral along the slant line $z = \exp(-i\pi/4)x$, with x increasing along real line from zero to ∞ , we have $\exp(-\pi z^2) = \exp(i\pi x^2)$ as is required for the Fourier transform that we are evaluating. The integral of $\exp(-\pi z^2)$ along the curved part of the contour vanishes as $R \rightarrow \infty$. The integral along the real line is already known as we found in evaluating the Fourier transform of the Gaussian function. The integrand being analytic on and inside the wedge-shaped contour, we have the following result:

Fig. 2.4 Contour used for evaluating Fourier transform of quadratic phase function



$$\begin{aligned} \int_0^\infty dx \exp(-i\pi/4) \exp(i\pi x^2) - \int_0^\infty dx \exp(-\pi x^2) &= 0 \\ \Rightarrow \int_0^\infty dx \exp(i\pi x^2) &= \frac{1}{2} \exp(i\pi/4). \end{aligned} \quad (2.67)$$

Using this result in Eq. (2.65), the Fourier transform of interest is thus given by

$$\mathcal{F}\{\exp(i\pi x^2)\} = \exp(i\pi/4) \exp(-i\pi f_x^2). \quad (2.68)$$

The change of the sign of i in the exponential and the additional phase factor of $\pi/4$ is peculiar in this result.

2.9 Properties of Fourier Transform

In this section, we will discuss several important properties of the Fourier transform that will be used directly later in the book:

1. **Linearity** Fourier transform is a linear operation. For two functions $g_1(x)$ and $g_2(x)$ and constants c_1 and c_2 , we have

$$\mathcal{F}\{c_1 g_1(x) + c_2 g_2(x)\} = c_1 G_1(f_x) + c_2 G_2(f_x). \quad (2.69)$$

2. **Scaling property** If a function is scaled by a constant factor a , the Fourier transform is scaled too, but in the opposite way. If a function is stretched, its Fourier transform is compressed:

$$\begin{aligned} \mathcal{F}\left\{g\left(\frac{x}{a}\right)\right\} &= \int_{-\infty}^\infty dx g\left(\frac{x}{a}\right) \exp(-i2\pi f_x x) \\ &= |a| \int_{-\infty}^\infty d(x/a) g\left(\frac{x}{a}\right) \exp(-i2\pi a f_x x/a) = |a|G(af_x). \end{aligned} \quad (2.70)$$

3. **Shifting property** The Fourier transform of a shifted function gets an additional phase factor depending on the shift in coordinates:

$$\begin{aligned} \mathcal{F}\{g(x-a)\} &= \int_{-\infty}^\infty dx g(x-a) \exp(-i2\pi f_x x) \\ &= \exp(-i2\pi f_x a)G(f_x). \end{aligned} \quad (2.71)$$

4. **Energy (Parseval) theorem** The energy contained in a signal $g(x)$ is equal to the energy in its Fourier transform $G(f_x)$. Here, the energy is defined as the squared L2-norm of $g(x)$ or $G(f_x)$ evaluated with respect to x or f_x coordinates, respectively:

$$\begin{aligned} & \int_{-\infty}^{\infty} df_x |G(f_x)|^2 \\ &= \int_{-\infty}^{\infty} df_x \left\{ \int_{-\infty}^{\infty} dx g(x) e^{-i2\pi f_x x} \right\} \left\{ \int_{-\infty}^{\infty} dx' g^*(x') e^{i2\pi f_x x'} \right\}. \end{aligned} \quad (2.72)$$

The integration over f_x may now be performed by making use of the Fourier integral representation of the delta function to obtain

$$\begin{aligned} \int_{-\infty}^{\infty} df_x |G(f_x)|^2 &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' g(x) g^*(x) \delta(x - x') \\ &= \int_{-\infty}^{\infty} dx |g(x)|^2. \end{aligned} \quad (2.73)$$

5. **Convolution theorem** This is one of the most important properties of the Fourier transform and as we will see later in this book. The Fourier transform of the convolution (defined below) of two functions is given by a product of their Fourier transforms:

$$\begin{aligned} & \mathcal{F} \left\{ \int_{-\infty}^{\infty} dx' g(x') h(x - x') \right\} \\ &= \int_{-\infty}^{\infty} dx \exp(-i2\pi f_x x) \int_{-\infty}^{\infty} dx' g(x') h(x - x') = G(f_x) H(f_x). \end{aligned} \quad (2.74)$$

The convolution operation is often required to model system response in linear systems theory that is useful for modeling a number of imaging systems. A similar result—the auto-correlation theorem—may be obtained if we use $h(x) = g^*(x)$ giving

$$\mathcal{F} \left\{ \int_{-\infty}^{\infty} dx' g(x') g^*(x - x') \right\} = |G(f_x)|^2. \quad (2.75)$$

As an application of the convolution theorem, we consider the transform of a triangle function:

$$\Lambda(x) = \begin{cases} 1 - |x| & |x| \leq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (2.76)$$

We observe that the triangle function can be obtained by convolving two rectangle distributions:

$$\Lambda(x) = \text{rect}(x) * \text{rect}(x), \quad (2.77)$$

so that its Fourier transform is given by $\mathcal{F}\{\Lambda(x)\} = \text{sinc}^2(f_x)$. In fact, one may repeatedly convolve the rect function with itself and generate a family of interpolation functions known as the Box-splines.

Next, we will consider a couple of results that are important in the context of Fourier transforms of 2D functions.

2.10 Fourier Transform of the 2D Circ Function

This is the first inherently 2D case that we will consider here in the study of Fourier transforms. Several optical elements like lenses are often circular in shape, and this case is thus important. We define the circ function (or distribution) as

$$\text{circ}\left(\frac{r}{a}\right) = \begin{cases} 1, & \sqrt{x^2 + y^2} < a \\ \frac{1}{2}, & \sqrt{x^2 + y^2} = a \\ 0, & \sqrt{x^2 + y^2} > a. \end{cases} \quad (2.78)$$

Before evaluating the Fourier transform, we make transformations:

$$\begin{aligned} x &= r \cos(\theta), \quad y = r \sin(\theta), \\ f_x &= \rho \cos(\phi), \quad f_y = \rho \sin(\phi). \end{aligned} \quad (2.79)$$

We will use this notation throughout the book for 2D polar coordinates (spatial and spatial frequency domains). The Fourier integral may now be expressed as

$$\mathcal{F}\left\{\text{circ}\left(\frac{r}{a}\right)\right\} = \int_0^{2\pi} d\theta \int_0^a r dr \exp(i2\pi r\rho \cos(\theta - \phi)).$$

The integral over θ can be expressed in term of the Bessel function $J_0(\dots)$ so as to get

$$\begin{aligned} \mathcal{F}\left\{\text{circ}\left(\frac{r}{a}\right)\right\} &= 2\pi \int_0^a r dr J_0(2\pi r\rho) = 2\pi \frac{1}{(2\pi\rho)^2} \int_0^{2\pi a\rho} u du J_0(u) \\ &= (2\pi a^2) \frac{J_1(2\pi a\rho)}{(2\pi a\rho)} = (\pi a^2) \text{jinc}(2a\rho). \end{aligned} \quad (2.80)$$

Here $J_n(\dots)$ denotes the Bessel function of the first kind and of order n . In evaluating the integral, we have used the property of Bessel functions:

$$\frac{d}{du} [u^n J_n(u)] = u^n J_{n-1}(u). \quad (2.81)$$

Also in analogy with the “sinc” function, we define the “jinc” function as

$$\text{jinc}(u) = 2 \frac{J_1(\pi u)}{(\pi u)}, \quad (2.82)$$

which tends to 1 in the limit $u \rightarrow 0$. The jinc function makes its presence as a Fourier transform of circular aperture. The square of the jinc functions is also referred to as the Airy disk.

2.11 Fourier Slice Theorem

The Fourier slice theorem is an important property of Fourier transforms inherent to two or higher dimensions. The Fourier slice theorem relates the 2D Fourier transform of a function to its projection along a particular direction. We define the projection $p_\theta(t)$ of the image $g(x, y)$ as a line integral through the image $g(x, y)$ along the line L as shown in Fig. 2.5:

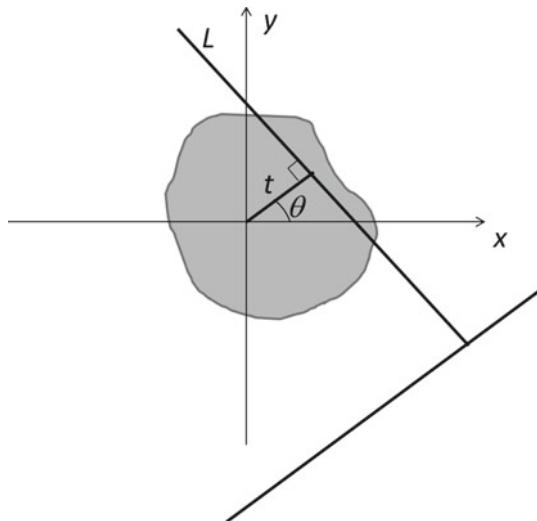
$$p_\theta(t) = \iint_{xy \text{ plane}} dx dy g(x, y) \delta(x \cos \theta + y \sin \theta - t). \quad (2.83)$$

The appearance of the delta function above is due to the fact that the line L along which the integral is to be evaluated is described by equation $x \cos \theta + y \sin \theta = t$. The line integral may be thought of as a scalar product between the function $g(x, y)$ and the line L . The Fourier transform of the projection with respect to the variable t may be related to the 2D Fourier transform $G(f_x, f_y)$ of the function $g(x, y)$ as follows:

$$\mathcal{F}_t\{p_\theta(t)\} = \int_{-\infty}^{\infty} dt \exp(-i 2\pi vt) \iint dx dy g(x, y) \delta(x \cos \theta + y \sin \theta - t). \quad (2.84)$$

The integral over t can be evaluated readily using the sampling property of the delta function. We therefore have

Fig. 2.5 Projection $p_\theta(t)$ as a line integral of object $g(x, y)$ along the line L



$$\begin{aligned}\mathcal{F}_t\{p_\theta(t)\} &= \iint dx dy g(x, y) \exp[-i 2\pi v(x \cos \theta + y \sin \theta)] \\ &= G(v \cos \theta, v \sin \theta).\end{aligned}\quad (2.85)$$

The 1D Fourier transform of the projection is thus seen to be equivalent to the 2D Fourier transform $G(f_x, f_y)$ along the line (or slice) passing through the origin which is described by $f_x = v \cos \theta, f_y = v \sin \theta$ with $v : (-\infty, \infty)$. Projections along various directions θ thus provide information about central slices through the 2D Fourier transform. This principle is central to medical imaging scanners (e.g. X-ray CT scanners) that generate images of the internal organs using projection data along multiple directions over a circle.

An interesting consequence of the Fourier slice theorem in three dimensions is the *common line theorem*. Analogous to the 2D case, the 2D Fourier transform of a projection of a 3D object (in any arbitrary direction) is equal to the planar slice of the object's 3D Fourier transform through the origin. Referring to Fig. 2.6, we note that the 2D Fourier transform of projection of a 3D object along the y direction is equal to the $x - z$ slice of the 3D Fourier transform of the object through the origin. As a consequence, if we consider any two arbitrary projections of a 3D object, they have a common intersecting line. For example, as seen in Fig. 2.6b, the 2D Fourier transforms of projections along y and z directions have a common intersecting line along the x -axis which is shown by a red line. The common line theorem allows one to find the angular relation between two arbitrary projections of a 3D object. This result finds important usage in cryo-electron microscopy [5] where the raw data effectively consists of arbitrarily oriented projections of similar virus-like particles. Once a number of projections are collected, the projection-based reconstruction cannot be performed unless the relative angular relationship between the projections is known. We will discuss more about projection-based reconstruction in Chap. 19.

2.12 Wigner Distribution

The Fourier transform $G(f_x)$ of a signal contains no information about the local behavior of the corresponding function $g(x)$. However, this does not prevent us from giving a simultaneous space-frequency description of the signal. There is a general class—Cohen class—of distribution functions that try to achieve such simultaneous description. It is clear that such a description cannot be more precise than that allowed by the uncertainty relations. One of the important space-frequency descriptions was given by Wigner in 1932. For a signal $g(x)$, the Wigner distribution is a function of both x and f_x and is given by

$$W_g(x, f_x) = \int_{-\infty}^{\infty} d\xi \ g(x + \xi/2) g^*(x - \xi/2) \exp(-i 2\pi f_x \xi). \quad (2.86)$$

For illustration let us consider two special cases:

$$g_1(x) = \exp(i2\pi f_0 x), \quad (2.87)$$

$$g_2(x) = \exp(i2\pi f_0 x^2). \quad (2.88)$$

The first function has a constant spatial frequency at all x , whereas the second signal is a chirp signal with linearly increasing local frequency. Evaluating the integral in Eq. (2.86) gives

$$W_{g_1}(x, f_x) = \delta(f_x - f_0), \quad (2.89)$$

$$W_{g_2}(x, f_x) = \delta(f_x - 2f_0 x). \quad (2.90)$$

The Wigner distribution is thus seen to provide information analogous to a musical score where the sequence of notes in time is described. The most important property of the Wigner distribution is its projection property. We first integrate the Wigner distribution with respect to the f_x variable to get

$$\begin{aligned} \int_{-\infty}^{\infty} df_x W_g(x, f_x) &= \int_{-\infty}^{\infty} d\xi g(x + \xi/2) g^*(x - \xi/2) \delta(\xi) \\ &= |g(x)|^2. \end{aligned} \quad (2.91)$$

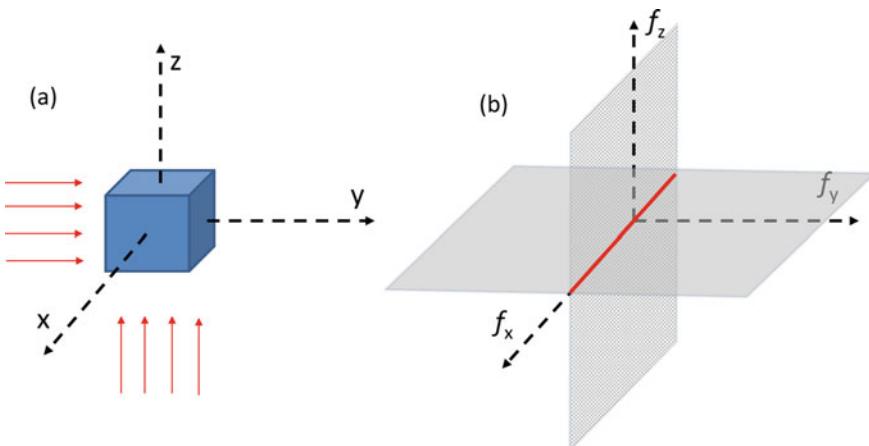


Fig. 2.6 **a** A 3D cube object with projection directions along y - and z -axes. **b** The 2D Fourier transforms of the two projections are equal to the $x - z$ and $x - y$ slices of the 3D Fourier transform of the cube object through origin. The red line shows intersection of the $x - z$ and $x - y$ planes which is the “common line” for these two projections

Further, integrating the Wigner distribution with respect to the variable x gives

$$\begin{aligned} \int_{-\infty}^{\infty} dx W_g(x, f_x) &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} d\xi g(x)g^*(x-\xi) \exp(-i2\pi f_x \xi) \\ &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} du g(x)g^*(u) \exp[-i2\pi f_x(x-u)] \\ &= |G(f_x)|^2. \end{aligned} \quad (2.92)$$

The projection of the Wigner distribution along x or f_x thus gives the complementary energy density $|G(f_x)|^2$ or $|g(x)|^2$, respectively. Wigner distribution is therefore also sometimes called a quasi-probability distribution. A closely related function called the ambiguity function is also sometimes used, particularly in radar signal processing for simultaneous space-frequency representation. It is defined as

$$A_g(x, f_x) = \int_{-\infty}^{\infty} dx' g(x' - x/2) g^*(x' + x/2) \exp(i2\pi f_x x'). \quad (2.93)$$

We will see later in this book that these functions have a close relation to the incoherent frequency response (or Optical Transfer Function) of imaging systems. We have discussed several ideas related to Fourier transforms in this chapter that we will use later in this book when discussing the propagation of light waves, analysis of imaging systems and modeling of image reconstruction algorithms.

Problems

2.1 Prove the generalized form of the energy or the Parseval theorem shown below

$$\int_{-\infty}^{\infty} dx g(x) h^*(x) = \int_{-\infty}^{\infty} df_x G(f_x) H^*(f_x).$$

In the above equation, G and H denote the Fourier transforms of g and h respectively, and the $*$ denotes the complex conjugation operation.

2.2 Evaluate the integrals below

$$\int_{-\infty}^{\infty} dx \left(\frac{\sin \pi x}{\pi x} \right)^2, \quad \int_{-\infty}^{\infty} dx \left(\frac{\sin \pi x}{\pi x} \right)^4.$$

2.3 Plot the function $q(x)$ defined below as x varies along the real line:

$$q(x) = \int_{-\infty}^x dt \delta(t - 100).$$

2.4 A linear array detector has a pixel pitch of p . The array has $(2N + 1)$ pixels. If a continuous irradiance $I(x)$ is incident on the detector array, show that the output signal at pixel k is proportional to

$$I_s(x) = \int_{-\infty}^{\infty} dx' I(x') \operatorname{rect}\left(\frac{x - x'}{p}\right),$$

where x takes discrete numerical values $x = kp$ and the index k runs from $-N$ to N .

2.5 Make a plot of the function $\operatorname{rect}(x) * \Lambda(x)$.

References

1. J. W. Goodman, *Introduction to Fourier Optics*, 2nd edn. (McGraw Hill, 1996)
2. R. Bracewell, *The Fourier Transform and Its Applications*, 2nd edn. (McGraw Hill, 1986)
3. M.J. Lighthill, *An Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, Cambridge, 1964)
4. B.G. Osgood, *The Fourier Transform and Its Applications*, Stanford Engineering Everywhere online material. (see.stanford.edu)
5. R.A. Crowther, D.J. DeRosier, A. Klug, The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. Proc. Royal Soc. London A **317**, 3193–3140 (1970)

Chapter 3

Sampling Theorem



We see the digital revolution everywhere around us which has changed human lives drastically within a span of a few decades. A lot of technologies have to get developed and come together for making a gadget like a cell phone possible. Sampling theorem is the core idea that makes it possible to faithfully represent continuous signals by means of their discrete samples [1–4]. We would like to point out that communication through electromagnetic means became practical in the early years of the twentieth century. It is then that the problem of quantifying the resources required for communicating some particular information became important. While the central result behind what we call the sampling theorem was well-known to mathematicians, the credit for developing a formalism for quantifying information in the context of communication goes to Claude Shannon. His article titled “The Mathematical Theory of Communication” (in *Bell Systems Tech. Journal* 1948) on this topic is a classic [1]. As we have already discussed before, the information of interest to us in imaging or related phenomena is encoded on light (or sound) waves scattered from the object of interest to our imaging system. While most of our models for wave phenomena are continuous, any imaging data is recorded by a discrete set of sensors. *Can we quantify the information in the scattered waves that enter our system aperture? How many samples are enough to capture the information adequately?* Intuitively it may appear at first that more samples are always better. Is that really so? Let us take an example of another kind of signal, say a speech signal which may have the highest frequency of a few kHz. It is also intuitive that it may be too much to sample the speech signal at a rate of say 1 MHz or 1 GHz when the highest audio frequencies of interest may be at most a few kHz. Similarly when we are trying to image a biological sample using a microscope whose resolution is $1 \mu\text{m}$, it may be too much if we try to sample the image field at points spaced 1 nm apart. Can we formalize this intuition in some way? Can we establish some bounds on the minimum sampling rate that is required that allows a good representation of the underlying continuous signal? We will study this important basic problem in this chapter. Sampling ideas play a

fundamental role in computational imaging systems. The discrete samples used to represent an image signal are often referred to as the pixels in an image. A central question of interest in the context of any imaging system is the number of pixels that are present in the resultant image formed by the system. In the following sections of this chapter, we will establish in detail, the main mathematical results associated with the Shannon sampling theorem and several allied concepts.

3.1 Sampling Theorem via Poisson Summation Formula

We will arrive at the Shannon sampling theorem via an interesting result on Fourier transforms—the Poisson summation formula which was known long before the sampling theorem became popular. With the knowledge of the Poisson summation formula, it is not very hard to arrive at the sampling theorem. It is just that the result gained significance due to its relevance to communication systems. We will begin by considering a signal $g(x)$ and its Fourier transform:

$$G(f_x) = \int_{-\infty}^{\infty} dx g(x) \exp(-i2\pi f_x x). \quad (3.1)$$

As the first step, we will construct a periodic function with period Δ by using the Fourier transform $G(f_x)$ that we can then represent as a Fourier series:

$$\sum_{n=-\infty}^{\infty} G(f_x + n\Delta) = \sum_{k=-\infty}^{\infty} g_k \exp(-i2\pi kf_x/\Delta). \quad (3.2)$$

We have defined the Fourier series with a negative sign in the exponential since our left-hand side is a Fourier transform. This keeps all our definitions from Chap. 2 consistent. The negative sign does not really matter since the summation k goes from $-\infty$ to ∞ . We will now determine the coefficients g_k of the Fourier series in Eq. (3.2):

$$\begin{aligned} g_k &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} df_x \sum_{n=-\infty}^{\infty} G(f_x + n\Delta) \exp(i2\pi kf_x/\Delta) \\ &= \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} \int_{(n-\frac{1}{2})\Delta}^{(n+\frac{1}{2})\Delta} du G(u) \exp(i2\pi ku/\Delta - i2\pi nk). \end{aligned} \quad (3.3)$$

Since n and k are integers, we have $\exp(-i2\pi nk) = 1$. Further, the summation over n effectively covers the entire u axis, and we can therefore write

$$g_k = \frac{1}{\Delta} \int_{-\infty}^{\infty} du G(u) \exp(i2\pi ku/\Delta) = \frac{1}{\Delta} g\left(\frac{k}{\Delta}\right). \quad (3.4)$$

The above result tells us that when a periodic function is formed by adding regularly shifted versions of Fourier transform $G(f_x)$, the corresponding Fourier series coefficients are simply the periodic samples of $g(x)$ at points $x = k/\Delta$ with integer k . This is an indication of the fact that periodic samples of $g(x)$ at a sampling interval of $(1/\Delta)$ contain information about the continuous Fourier transform $G(f_x)$. A judicious choice of Δ provides us with the sampling theorem.

3.1.1 Poisson Summation Formula for Bandlimited Signal

In order to establish the sampling theorem, we will define the class of bandlimited signals. We will call a signal $g(x)$ to be bandlimited to the frequencies $f_x : (-B, B)$ if its Fourier transform vanishes outside this frequency interval. Finite energy functions whose Fourier transform is limited to a finite interval like this are called the Paley-Wiener functions. We will consider a special case of the Poisson summation formula when the signal $g(x)$ is bandlimited as above and consider a special case $\Delta = 2B$. The shifted versions of the Fourier transform in the Poisson summation formula now do not overlap. We can then filter out the $n = 0$ term on the left-hand side of Eq. (3.2) by multiplying both sides of this equation with a rect function of width $(2B)$:

$$G(f_x) = \frac{1}{2B} \sum_{k=-\infty}^{\infty} g\left(\frac{k}{2B}\right) \exp\left(-i2\pi k \frac{f_x}{2B}\right) \text{rect}\left(\frac{f_x}{2B}\right). \quad (3.5)$$

Taking an inverse Fourier transform now yields the sampling theorem:

$$\begin{aligned} g(x) &= \frac{1}{2B} \sum_{k=-\infty}^{\infty} g\left(\frac{k}{2B}\right) \int_{-\infty}^{\infty} df_x \exp\left[i2\pi f_x \left(x - \frac{k}{2B}\right)\right] \text{rect}\left(\frac{f_x}{2B}\right) \\ &= \sum_{k=-\infty}^{\infty} g\left(\frac{k}{2B}\right) \text{sinc}\left[2B \left(x - \frac{k}{2B}\right)\right]. \end{aligned} \quad (3.6)$$

This is an exciting result. It tells us that the samples of a bandlimited signal $g(x)$ at discrete points separated $1/(2B)$ apart contain the same information as that in the continuous signal. If we know these discrete samples, the result further provides a prescription for regenerating the continuous signal via sinc-interpolation of the samples. The choice of $\Delta = 2B$ for arriving at the sampling theorem just avoided overlap of the various terms on the left-hand side of Eq. (3.2). Sampling at $2B$ samples per unit interval is the minimum rate at which the overlap is avoided which in principle allows perfect signal reconstruction from its samples. This minimum sampling rate is known by the name Nyquist rate to honor the early contributions by H. Nyquist to communication theory. As a thumb rule, a signal of the form $\cos(2\pi f_0 x)$ requires at least two samples per period. It is important to note that practical laboratory measurements do not provide “point samples” as in the sampling

formula. For example, when an image is recorded on a sensor consisting of an array of pixels, the individual measurement represents the integrated energy received over a pixel size. So in order to sample a cosine function, having just two samples that integrate the function over a finite pixel size is not enough, and typically a higher sampling rate compared to the Nyquist rate may be required.

3.1.2 Additional Notes on the Sampling Formula

An important point to note regarding the sampling theorem is that the set of shifted sinc functions forms an orthogonal and complete basis for the class of bandlimited functions. This basis is quite different from other common basis sets you may have encountered before (e.g. Hermite-Gauss functions) in that the shifted versions of the same function are here seen to be orthogonal. The origin of advanced signal representation concepts like wavelets may be traced back to this unusual kind of basis [5]. The orthogonality property may be proved as follows:

$$\int_{-\infty}^{\infty} dx \operatorname{sinc}(2Bx - m) \operatorname{sinc}(2Bx - n) = \frac{1}{2B} \operatorname{sinc}(m - n) = \frac{1}{2B} \delta_{m,n}. \quad (3.7)$$

In the calculation above we have made use of the convolution theorem. The sample $g\left(\frac{m}{2B}\right)$ of the bandlimited signal can thus be thought of as a coefficient of the series expansion in the sinc-basis:

$$g\left(\frac{m}{2B}\right) = \frac{1}{2B} \int_{-\infty}^{\infty} dx g(x) \operatorname{sinc}(2Bx - m). \quad (3.8)$$

The completeness property of the sinc-basis set is simple to prove. We observe that

$$\sum_{n=-\infty}^{\infty} \operatorname{sinc}(2Bx - m) \operatorname{sinc}(2Bx' - m) = \operatorname{sinc}[2B(x - x')]. \quad (3.9)$$

The identity follows by applying the sampling theorem to the sinc function itself. For every signal $g(x)$ bandlimited to $f_x : (-B, B)$, we have

$$\int_{-\infty}^{\infty} dx' g(x') \operatorname{sinc}[2B(x - x')] = \frac{1}{2B} g(x). \quad (3.10)$$

If we now use the completeness relation for the term $\operatorname{sinc}[2B(x - x')]$ and the relation in Eq. (3.8), we see that we arrive at the sampling theorem. The sinc-basis thus allows representation of every bandlimited function. The expression $\operatorname{sinc}[2B(x - x')]$ may therefore be thought of as the bandlimited version of $\delta(x - x')$.

3.2 Sampling of Carrier-Frequency Signals

The Shannon sampling theorem is sometimes explained simplistically by stating the thumb rule that a sinusoidal signal requires two samples per cycle in order to represent it correctly. This statement is somewhat misleading, because if one knows that the signal being sampled is a sinusoid, then it should be possible to specify the waveform only from the knowledge of its frequency, amplitude and initial phase. A similar situation arises in practice for an important class of signals that one encounters in communication, representation of narrow band optical fields and optical interferometric imaging, etc. that are of the form:

$$g(x) = a(x) \exp(i2\pi f_0 x) + a^*(x) \exp(-i2\pi f_0 x). \quad (3.11)$$

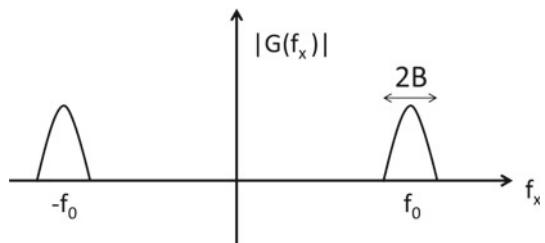
Here, $a(x)$ is called the complex envelope of the signal, and f_0 is its carrier frequency. In most cases, where such signals are studied, the effective bandwidth $2B$ of the complex envelope $a(x)$ is much less than the carrier frequency f_0 . The magnitude of the Fourier spectrum of a typical carrier-frequency signal is as shown in Fig. 3.1. It is clear that while the highest frequency in the signal is $(f_0 + B)$, it is reasonable to expect that the sampling rate required for digitizing such a signal need not be as high as $2(f_0 + B)$ samples per unit interval in x . This is because there is a lot of empty region between the two lobes in the Fourier spectrum centered on the carrier frequency $\pm f_0$. It is straightforward to verify that if we select a frequency interval $2B_0$ with

$$\frac{f_0}{2B} \geq \frac{f_0}{2B_0} = N + \frac{1}{2}, \quad (3.12)$$

for some integer N (preferably largest possible for given values of f_0 and B), then the signal $g(x)$ can be sampled at intervals $1/(4B_0)$ to get a carrier-frequency form of the sampling theorem [6]:

$$g(x) = \sum_{k=-\infty}^{\infty} g\left(\frac{k}{4B_0}\right) \operatorname{sinc}\left[2B_0\left(x - \frac{k}{4B_0}\right)\right] \cos\left[2\pi f_0\left(x - \frac{k}{4B_0}\right)\right]. \quad (3.13)$$

Fig. 3.1 Magnitude of Fourier spectrum $|G(f_x)|$ for a carrier-frequency type signal as in Eq. (3.11)



The above result can be arrived at by similar arguments as the usual low-pass sampling theorem and is sometimes referred to as the band-pass sampling theorem [7, 8]. The central idea in this result is to effectively utilize the fact that the empty region in Fourier space does not lead to aliasing even if the sampling interval is much larger than $1/[2(f_0 + B)]$. If the cosine term in the sampling relation above is written as a sum of two exponentials, the demodulated complex envelope $a(x)$ can be expressed as another sampling series:

$$a(x) = \sum_{k=-\infty}^{\infty} g\left(\frac{k}{4B_0}\right) \text{sinc}\left[2B_0\left(x - \frac{k}{4B_0}\right)\right] \exp\left(-2\pi f_0 \frac{k}{4B_0}\right). \quad (3.14)$$

The sampling rate is now comparable to the fluctuations in the complex envelope rather than the maximum signal bandwidth. The knowledge of carrier frequency f_0 is implicitly assumed in the above discussion. This result is important in the context of quantifying the degrees of freedom in a carrier-frequency signal as we will discuss in the following section.

3.3 Degrees of Freedom in the Signal—Space-Bandwidth Product

Suppose we have a signal of length $2L$ which we want to store or transmit to someone else. What resources are required? Most practical signals or images are of finite extent, and a finite length signal cannot be bandlimited at the same time in a strict sense. Its Fourier transform must in principle extend all over the frequency axis. The notion of bandwidth thus needs to be defined carefully. For a signal $g(x)$ defined over a finite interval $x : (-L, L)$, we define energy concentration ratio within some band of frequencies $f_x : (-B, B)$ as

$$\alpha = \frac{\int_{-B}^B df_x |G(f_x)|^2}{\int_{-\infty}^{\infty} df_x |G(f_x)|^2}. \quad (3.15)$$

A practical definition of bandwidth is now taken as the interval $f_x : (-B, B)$ for which the energy concentration ratio takes some value close to 1 (e.g. 0.999) as demanded by the application. This numerical value of B may be used in the sampling formula in Eq. (3.6). How many samples of the signal are needed to represent it in a fairly accurate manner? As per the sampling formula, the answer is $(2B)(2L)$. The product $N = 4BL$ is called the space-bandwidth product of a signal, and it represents the degrees of freedom in a signal. For a 2D signal or image, this number would be N^2 and it is alternately referred to as the pixels required to represent an image.

The space-bandwidth product of a signal and that for a system are two separate concepts. Suppose our object is a biological cell sample placed on a slide to be

observed under a microscope. A living cell is a very complex object and details of interest can be at a molecular or nanometer scale corresponding to spatial frequencies of the order of 10^9 m^{-1} . For a sample of size $1 \times 1 \text{ mm}$, the space-bandwidth product of the object is then of the order of $(10^{-3})^2(10^9)^2 = 10^{12}$. The system for observing the cell sample may have a much smaller space-bandwidth product. The resolution of a microscope using visible light may be $1 \mu\text{m}$. For sufficient illumination level, the space-bandwidth product for the system is thus $(10^{-3})^2(10^6)^2 = 10^6$. So if we want to use a CCD sensor as a detector at the image plane of the microscope, a sensor chip with pixel numbers of the order of 1000×1000 pixels should be sufficient. If we are operating under low-light-level conditions, the noise will make the signal from each pixel somewhat unreliable. We may have to bin the sensor chip so that an average is performed over say a 2×2 pixel block. The resulting image will thus have 500×500 pixels—resulting in a reduction in the information that can be retrieved from the system setup. The space-bandwidth product thus depends both on the system parameters and the signal-to-noise considerations.

3.4 Slepian (Prolate Spheroidal) Functions

We have defined the energy concentration ratio α in Eq. (3.15) in order to arrive at a practical definition of what we mean by bandwidth. The problem of finding the functions that achieve the maximum value of α was posed by Shannon in the 1950s. The comprehensive analysis of this problem was performed by Slepian, Pollack and Landau in a series of papers in the Bell System Tech. Journal in 1960s [9–12]. Such functions are now often referred to as the Slepian functions in honor of D. Slepian. We will study some of their interesting properties and also establish their connection to the Shannon sampling theorem. The Slepian functions will be found useful later in the book when we discuss issues such as super-resolution and information-carrying capacity of imaging systems. For a function $g(x)$ defined over $x : (-L, L)$, we will write the energy concentration ratio α explicitly as

$$\begin{aligned}\alpha &= \frac{\int_{-B}^B df_x \int_{-L}^L \int_{-L}^L dx dx' g(x)g^*(x') \exp[-i2\pi f_x(x - x')]}{\int_{-\infty}^{\infty} df_x |G(f_x)|^2} \\ &= 2B \frac{\int_{-L}^L \int_{-L}^L dx dx' g(x)g^*(x') \text{sinc}[2B(x - x')]}{\int_{-L}^L dx |g(x)|^2}.\end{aligned}\quad (3.16)$$

We observe that the above expression for α is the Rayleigh-Ritz coefficient for the sinc-kernel and the highest value of α is given by $(2B\lambda_0)$ with λ_0 being the highest eigenvalue of the following equation:

$$\lambda \phi(x) = \int_{-L}^L dx' \text{sinc}[2B(x - x')] \phi(x').\quad (3.17)$$

The prolate basis functions are numbered in decreasing order of the magnitude of the corresponding eigenvalues. The set of the highest N eigenfunctions may also be treated as the best basis for given values of L and B in the sense that they have the highest energy concentration. In their early work, Slepian and co-workers realized that the eigenfunctions of the sinc-kernel happen to be the solutions of the angular part of the prolate spheroidal differential equation

$$(L^2 - x^2) \frac{d^2\phi}{dx^2} - 2x \frac{d\phi}{dx} + \left(\chi - \frac{c^2 x^2}{L^2} \right) = 0, \quad (3.18)$$

for discrete positive values of the function $\chi(c)$. The parameter $c = 2\pi LB$ is referred to as the Shannon number and is related to the space-bandwidth product. Most of the literature on Slepian functions studies their properties based on the solution of the differential equation (3.18) and its asymptotic solutions depending on the value of c . Here, we will take a different approach based on the sampling theorem to arrive at the prolate functions. The discussion based on the solution of the differential equation is somewhat ad hoc and does not have a simple connection to the ideas such as the space-bandwidth product that are inherently related to the sampling theorem.

We note that the eigenvalue problem in Eq. (3.17) is a homogeneous Fredholm integral equation of the second kind. The kernel of the equation is Hermitian symmetric and square-integrable. Further, the kernel is positive definite:

$$\iint_{-L}^L dx dx' \text{sinc}[2B(x - x')] g(x) g^*(x') = \frac{1}{2B} \int_{-B}^B df_x |G(f_x)|^2 \geq 0,$$

for an arbitrary square-integrable function $g(x)$. The equality above holds when $g(x)$ is identically equal to zero. This property is known as non-negative definiteness of the integral kernel which ensures that the eigenvalues of the sinc-kernel are positive. We will first show that the eigenfunctions of the sinc-kernel satisfy the sampling theorem and are hence bandlimited. We note that the sinc-kernel is itself bandlimited and hence has a sampling expansion given by

$$\text{sinc}[2B(x - x')] = \sum_{m=-\infty}^{\infty} \text{sinc}(2Bx - m) \text{sinc}(2Bx' - m). \quad (3.19)$$

Using this sampling expansion in Eq. (3.17) gives

$$\begin{aligned} \lambda \phi(x) &= \int_{-L}^L dx' \left[\sum_{m=-\infty}^{\infty} \text{sinc}(2Bx - m) \text{sinc}(2Bx' - m) \right] \phi(x') \\ &= \lambda \sum_{m=-\infty}^{\infty} \phi\left(\frac{m}{2B}\right) \text{sinc}(2Bx - m). \end{aligned} \quad (3.20)$$

We observe that the eigenfunctions satisfy the sampling relation and are bandlimited as well. We now write the eigenvalue equation (3.17) as

$$\begin{aligned}\lambda \phi\left(\frac{m}{2B}\right) &= \int_{-L}^L dx' \operatorname{sinc}(2Bx' - m) \phi(x') \\ &= \int_{-L}^L dx' \operatorname{sinc}(2Bx' - m) \sum_{n=-\infty}^{\infty} \phi\left(\frac{n}{2B}\right) \operatorname{sinc}(2Bx' - n) \\ &= \sum_{n=-\infty}^{\infty} A_{mn}^{(0)} \phi\left(\frac{n}{2B}\right).\end{aligned}\quad (3.21)$$

Here, we have defined the matrix $A^{(0)}$ by the matrix elements

$$A_{mn}^{(0)} = \int_{-L}^L dx' \operatorname{sinc}(2Bx' - m) \operatorname{sinc}(2Bx' - n). \quad (3.22)$$

We make an important observation that the continuous eigenvalue problem for the sinc-kernel is thus equivalent to a discrete eigenvalue problem for the matrix $A^{(0)}$ [13]. The continuous and discrete problems share the same eigenvalues. Further, the eigenvectors of the matrix $A^{(0)}$ are the same as the Nyquist samples of the continuous prolate spheroidal functions; as a result, the continuous eigenfunctions can be obtained by sinc-interpolation of the discrete eigenvectors.

3.4.1 Properties of Matrix $A^{(0)}$

For brevity, we will denote the discrete eigenvectors of the matrix $A^{(0)}$ as

$$u_n = \left[\dots \phi_n\left(\frac{m}{2B}\right) \dots \right]^T,$$

with “T” denoting the transpose of the row vector. The matrix $A^{(0)}$ is real symmetric so that its eigenvalues and eigenvectors are real. If we normalize the continuous functions $\phi_n(x)$ over $x : (-\infty, \infty)$, it follows that

$$\int_{-\infty}^{\infty} dx \phi_n^2(x) = 1 = \frac{1}{2B} \sum_{m=-\infty}^{\infty} \phi_n^2\left(\frac{m}{2B}\right). \quad (3.23)$$

The second equality follows from the fact that $\phi_n(x)$ has a sampling expansion as per Eq. (3.20) and further using the orthogonality of the sinc functions as in Eq. (3.7). The matrix $A^{(0)}$ is centro-symmetric with respect to the element $A_{00}^{(0)}$. It is known that the eigenvectors of such matrices have definite parity—they are either even or

odd. The continuous eigenfunctions obtained by sinc-interpolation of the discrete eigenvectors thus also have a definite parity. The matrix $A^{(0)}$ being symmetric can be factored as

$$A^{(0)} = U A_d^{(0)} U^T. \quad (3.24)$$

Here, $A_d^{(0)}$ is a diagonal matrix with eigenvalues λ_n along the diagonal and the matrix U has the eigenvectors u_n as its columns. Further, the eigenvectors being orthogonal have the property that

$$U^T U = 2B\hat{\mathbf{1}} = U U^T. \quad (3.25)$$

These identities may be written explicitly to obtain the following relations that correspond to the orthogonality and completeness of the discrete basis functions:

$$\sum_{m=-\infty}^{\infty} \phi_n\left(\frac{m}{2B}\right) \phi_l\left(\frac{m}{2B}\right) = 2B\delta_{nl}, \quad (3.26)$$

$$\sum_{n=0}^{\infty} \phi_n\left(\frac{m}{2B}\right) \phi_n\left(\frac{k}{2B}\right) = 2B\delta_{mk}. \quad (3.27)$$

The structure of matrix $A^{(0)}$ is quite interesting. Since the matrix elements are defined as definite integrals over the product of shifted sinc functions, only approximately $(2B)(2L)$ matrix elements along the diagonal of the matrix centered on the matrix element $A_{00}^{(0)}$ are significant. This explains the well-known property of the sinc-kernel that only about $4BL$ highest eigenvalues for the problem are significant. This has interesting implications for signal representation using prolate functions. In Figs. 3.2 and 3.3, we show the first 8 prolate functions and their corresponding eigenvalues, respectively, computed using the sampling theorem-based approach. A 101×101 approximation of the matrix $A^{(0)}$ centered on the element $A_{00}^{(0)}$ is used for this illustration, and the matrix elements are computed using standard numerical integration methods as per the definition in Eq. (3.22). The discrete eigenvectors thus obtained are then sinc-interpolated to obtain the continuous eigenfunctions. In general, we observe that the n -th prolate function has n zero crossings in the interval $x : (-L, L)$. This appears counter-intuitive for functions of order $n > 4BL$ where the number of zero crossings go beyond the space-bandwidth product although the corresponding prolate function is still bandlimited over the infinite interval. Such functions with larger number of zero crossings in a finite interval compared to the space-bandwidth product are referred to as “super-oscillatory” in nature. We further note that as the index n goes above $4BL$, most of the energy in corresponding prolate functions is concentrated beyond the interval $x : (-L, L)$.

We will state some of the interesting properties of the prolate functions which can be easily derived from the discrete orthogonality and completeness relations in Eqs. (3.26) and (3.27):

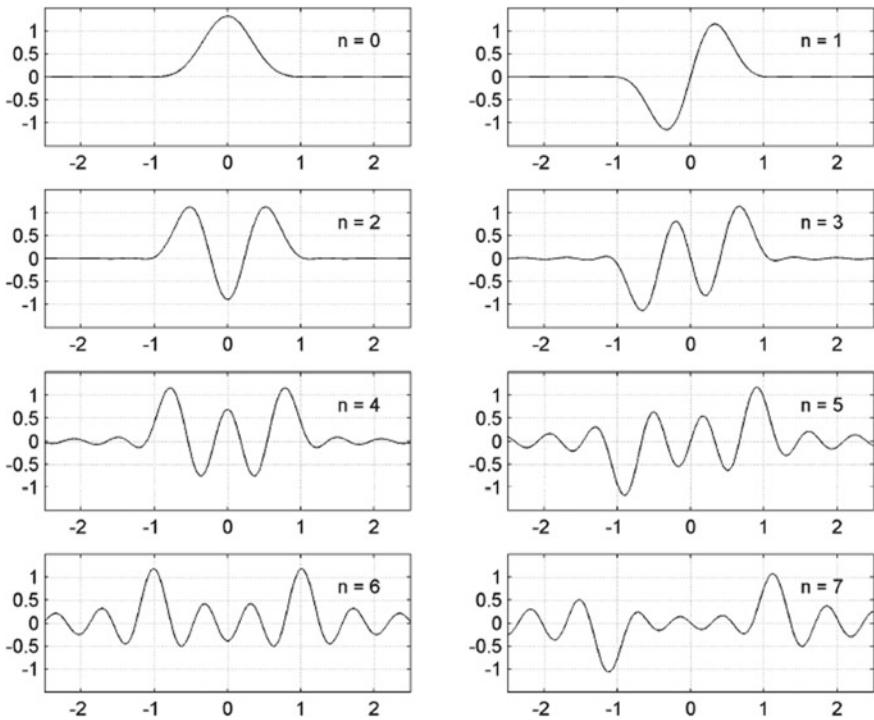
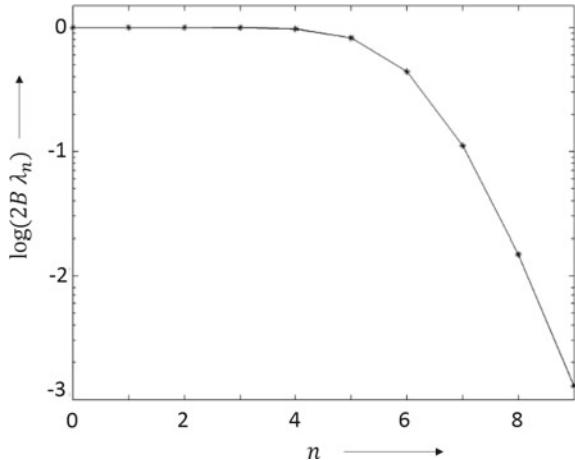


Fig. 3.2 First 8 prolate basis functions computed using sampling theorem-based approach for $L = 1$, $B = 5/\pi$

Fig. 3.3 First 10 eigenvalues of sinc-kernel computed using sampling theorem-based approach for $L = 1$, $B = 5/\pi$ shown on a log scale



- 1. Dual orthogonality** The prolate functions form an orthogonal set over $(-\infty, \infty)$ as well as over the domain $x : (-L, L)$ used for their definition:

$$\int_{-\infty}^{\infty} dx \phi_n(x) \phi_m(x) = \delta_{m,n}, \quad (3.28)$$

$$\int_{-L}^L dx \phi_n(x) \phi_m(x) = (2B\lambda_n) \delta_{m,n}. \quad (3.29)$$

The first of the identity is easy to prove if one uses sampling expansions for both $\phi_n(x)$ and $\phi_m(x)$ followed by use of the orthogonality relation in Eq. (3.26). The second orthogonality identity over finite interval can be proved as follows:

$$\begin{aligned} \int_{-L}^L dx \phi_n(x) \phi_m(x) &= \sum_{k=-\infty}^{\infty} \phi_n\left(\frac{k}{2B}\right) \int_{-L}^L dx \operatorname{sinc}(2Bx - k) \phi_m(x) \\ &= \lambda_m \sum_{k=-\infty}^{\infty} \phi_n\left(\frac{k}{2B}\right) \phi_m\left(\frac{k}{2B}\right) = 2B\lambda_m \delta_{m,n}. \end{aligned} \quad (3.30)$$

- 2. Completeness** The prolate spheroidal functions as an orthogonal basis over $(-\infty, \infty)$ for functions bandlimited to $f_x : (-B, B)$ and a complete orthogonal basis for square-integrable functions in $x : (-L, L)$. The two completeness properties may be stated as

$$\sum_{n=0}^{\infty} \phi_n(x) \phi_n(x') = 2B \operatorname{sinc}[2B(x - x')] \quad \text{for } x \in (-\infty, \infty), \quad (3.31)$$

$$\sum_{n=0}^{\infty} \frac{1}{(2B\lambda_n)} \phi_n(x) \phi_n(x') = \delta(x - x'), \quad |x|, |x'| \leq L. \quad (3.32)$$

We will not prove the completeness properties here but will refer the interested readers to [13].

We will use the orthogonality and completeness properties when considering ill-posed problems such as bandlimited extrapolation. In general, throughout the discussion in this book, we will have various occasions to refer to prolate functions when discussing information transmission capability of imaging systems. We remark that the sampling theorem-based approach to prolate functions applies equally well to eigenvalue problems associated with bandlimited kernels in general, and several interesting mathematical relations similar to the ones for prolate spheroids above can be proved for this general case.

3.4.2 Extrapolation of Bandlimited Functions

The dual orthogonality and completeness properties of the prolate spheroidal functions can be utilized for extrapolation of bandlimited functions that are known over a truncated region. Suppose a function $g(x)$ bandlimited to the frequency band $f_x : (-B, B)$ is known over a finite range $x : (-L, L)$, then one may expand it using the prolate spheroidal function set as

$$g(x) = \sum_n a_n \phi_n(x), \quad (3.33)$$

where the coefficients a_n may be determined using the orthogonality of prolate functions over the region $x : (-L, L)$:

$$a_n = \frac{1}{2B\lambda_n} \int_{-L}^L dx g(x) \phi_n(x). \quad (3.34)$$

Further, one may in principle extend the function beyond the range $x : (-L, L)$ by using the same coefficients a_n . The possibility of determining the function beyond the range $x : (-L, L)$ is however practically governed by the accuracy of our knowledge of $g(x)$ over the finite range. If, for example, the values of $g(x)$ are determined by some experimental measurement, they will contain some noise which translates to noise on the coefficients a_n . In Eq. (3.34), the evaluation of the coefficient a_n involves division by the corresponding eigenvalue λ_n . As we have already seen above, the eigenvalues are typically very small when the index $n >> 4BL$, and the corresponding prolate functions have most of their energy concentrated beyond the range $x : (-L, L)$. So although such higher order prolate functions are capable of recreating the signal beyond the finite range where the function $g(x)$ is known, their inclusion in the series representation involves division of corresponding coefficients a_n by small numbers. Since the coefficients a_n are noisy, such division by small numbers can amplify the noise making the series representation meaningless. One simple way to avoid the amplification of noise is to truncate the series representation to finite number of terms which then amounts effectively to extrapolation over a limited range beyond $(-L, L)$. The extrapolation problem is highly sensitive to the signal to noise ratio (SNR) in the measurement of $g(x)$. Numerical super-resolution imaging techniques usually can be formulated in terms of the extrapolation using the Slepian functions.

We conclude this discussion by noting that the sampling by means of the sinc-series or the Slepian function basis are equivalent to each other. The origins of the important developments in allied areas such as signal representation using wavelets can be traced to the material covered in this chapter.

3.5 Band-Pass Analogues of Prolate Spheroidal Functions

The sampling theorem-based procedure for arriving at prolate spheroidal functions has important advantages. First of all, this methodology establishes a formal connection between the sampling theorem and the minimum uncertainty signals without referring to any differential equation like that for the angular part of prolate spheroidal equation. This methodology also opens up a possibility of constructing minimum uncertainty basis sets for other band structures using the corresponding sampling theorems. We will consider here the case of band-pass analogues of prolate spheroidal functions (BPSF) [14]. A sampling theorem for the band-pass case was obtained in Sect. 3.2. This sampling relation can, for example, be utilized for constructing signals defined over length $x : (-L, L)$ that have highest energy concentration in the frequency band Ω which is a union of two intervals $f_x : (-f_0 - B, -f_0 + B)$ and $f_x : (f_0 - B, f_0 + B)$, by following a treatment similar to that in Sect. 3.4. There is no commonly known differential equation whose solutions are solutions of the minimum uncertainty functions for the frequency band Ω . With the sampling theorem-based procedure, we can once again introduce the matrix $A^{(0)}$ now defined as

$$A_{mn}^{(0)} = \int_{-L}^L dx' S[x' - m/(4B_0)] S[x' - n/(4B_0)], \quad (3.35)$$

where $S(x) = \text{sinc}(2B_0x) \cos(2\pi f_0 x)$. Further, the required continuous eigenfunctions $\{\psi_n(x)\}$ associated with the kernel $S(x)$ may be obtained by interpolating the eigenvectors of the matrix $A^{(0)}$ in Eq. (3.35) using the band-pass form of sampling formula in Eq. (3.13). The first 8 BPSFs computed using this procedure are shown in Fig. 3.4. The BPSF functions obey the band-pass form of sampling theorem in Eq. (3.13) and are therefore band-limited to the frequencies in Ω . The envelopes of the BPSF functions are observed to have a form similar to the low-pass prolate basis functions shown in Fig. 3.2. The BPSF basis functions have the highest energy concentration in the band Ω centered on the frequencies $\pm f_0$. Based on the centro-symmetric nature of the matrix $A^{(0)}$ defined in Eq. (3.35), the discretely sampled BPSF basis functions can be shown to obey orthogonality and completeness properties similar to Eqs. (3.26) and (3.27) as follows:

$$\sum_{m=-\infty}^{\infty} \psi_n \left(\frac{m}{4B_0} \right) \psi_l \left(\frac{m}{4B_0} \right) = 4B_0 \delta_{n,l}, \quad (3.36)$$

$$\sum_{n=0}^{\infty} \psi_n \left(\frac{m}{4B_0} \right) \psi_n \left(\frac{k}{4B_0} \right) = 4B_0 \delta_{m,k}. \quad (3.37)$$

Further using the above properties, it is once again possible to prove that the BPSF basis functions $\{\psi_n(x)\}$ have the curious dual orthogonality over the intervals $(-L, L)$ as well as $(-\infty, \infty)$ shown by the low-pass prolate spheroidal functions.

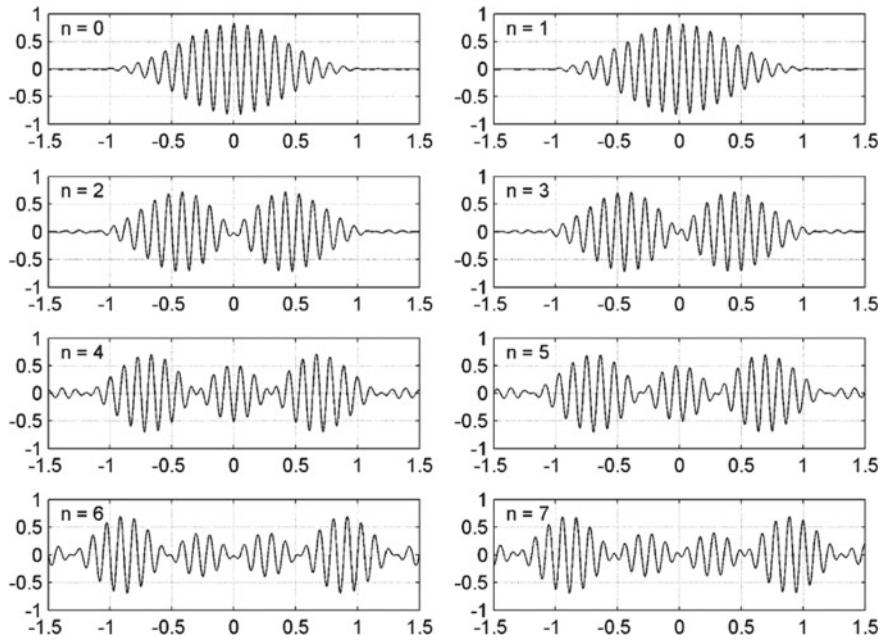


Fig. 3.4 First 8 BPSF basis functions computed using sampling theorem-based approach for numerical parameters $L = 1$, $B_0 = 1$ and $f_0 = 9B_0$. Adapted with permission from [14] ©Elsevier

Highest energy concentrated basis sets analogous to the low-pass prolate spheroidal basis can therefore be constructed using the corresponding sampling theorems and used in corresponding applications for efficient signal representation.

Problems

3.1 The angular resolution of a typical human eye is 1 arc minute. Estimate the minimum number of pixels (order of magnitude estimate is sufficient) needed for a smartphone screen of size 3×5 in.

3.2 Consider the Fourier transform pair $\psi(x) = \text{rect}(x)$ and $\tilde{\psi}(f_x) = \text{sinc}(f_x)$. Show that the uncertainty product $(\Delta x)(\Delta f_x)$ diverges, where

$$(\Delta x)^2 = \int_{-\infty}^{\infty} dx x^2 |\psi(x)|^2,$$

and

$$(\Delta f_x)^2 = \int_{-\infty}^{\infty} df_x f_x^2 |\tilde{\psi}(f_x)|^2.$$

Interpret the result and suggest a better method for defining uncertainty for long-tail functions like $\text{sinc}(\dots)$.

3.3 Show that the finite Fourier transform

$$\int_{-L}^L dx \phi_n(x) \exp(-i2\pi f_x x)$$

of the n-th prolate spheroidal function $\phi_n(x)$ defined in this chapter with respect to a length interval $2L$ and bandwidth interval $2B$ is a scaled version of the same function apart from a constant multiplier that depends on the eigenvalue λ_n as in Eq. (3.17).

3.4 Consider a function $g(x)$ which is bandlimited to the frequency range $\Omega = (f_o - B, f_o + B)$ with $f_o >> B$. Using the treatment provided in this chapter, construct a set of orthogonal functions with highest energy concentration in Ω . How will you compute the functions in this basis set?

3.5 Consider a function

$$g(x) = \text{sinc}(2Bx) \cos(2\pi f_0 x),$$

where $B = 1$ and $f_0 = 9$. Plot the function and find the number of zero crossings in this function in the range $x : (-1, 1)$. Construct a set of band-pass prolate functions $\{\psi_n\}$ numerically and show that a much smaller number of this basis set (compared to number of zero crossings) is required to represent the function $g(x)$ in the interval $x : (-1, 1)$.

References

1. C.E. Shannon, Mathematical theory of communication. Bell Syst. Tech. J. **27**, 379 (1948)
2. Abdul Jerri, The Shannon sampling theorem—its various extensions and applications: a tutorial review. Proc. IEEE **65**, 1565 (1977)
3. M. Unser, Sampling- 50 years after Shannon. Proc. IEEE **88**, 569 (2000)
4. J.R. Higgins, *Sampling Theory in Fourier and Signal Analysis* (Oxford Science Publications, 1996)
5. S. Mallat, *A Wavelet Tour of Signal Processing* (Academic Press USA, 2008)
6. K. Khare, N. George, Direct sampling and demodulation of carrier-frequency signals. Opt. Commun. **211**, 85–94 (2002)
7. A. Kohlenberg, Exact interpolation of band-limited functions. J. Appl. Phys. **24**, 1432–1436 (1953)
8. R.G. Vaughan, N.L. Scott, D.R. White, Theory of bandpass sampling. IEEE Trans. Signal Process. **39**, 1973–1989 (1991)
9. D. Slepian, H.O. Pollak, Prolate spheroidal wavefunctions, Fourier analysis and uncertainty—I. Bell Syst. Tech. J. **40**, 43–64 (1961)
10. H.J. Landau, H.O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty—II. Bell Syst. Tech. J. **40**, 65–84 (1961)

11. H.J. Landau, H.O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty—III: the dimension of the space of essentially time- and band-limited signals. *Bell Syst. Tech. J.* **41**, 1295–1336 (1962)
12. D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty—IV: extensions to many dimensions; generalized prolate spheroidal functions. *Bell Syst. Tech. J.* **43**, 3009–3057 (1964)
13. K. Khare, N. George, Sampling theory approach to prolate spheroidal wavefunctions. *J. Phys. A: Math. Gen.* **36**, 10011 (2003)
14. K. Khare, Bandpass sampling and bandpass analogues of prolate spheroidal functions. *Sig. Proc.* **86**, 1550–1558 (2006)

Chapter 4

Operational Introduction to Fast Fourier Transform



Fourier transforms will be encountered time and again in this book, and it is important for students and practitioners to have a basic operational understanding of numerical routines or functions readily available for implementing Fourier transform on discretely sampled signals and images. This chapter will not provide the details of the discrete Fourier transform and fast Fourier transform algorithms for which excellent literature is already available [1, 2]. The aim here is to provide sufficient information so that when using standard computational tools or libraries for FFT operation, a user may be able to make sense of the results. A few important aspects regarding the usage of 2D Fourier transform functions for simulating optical imaging systems are also pointed out along the way.

4.1 Definition of Discrete Fourier Transform

With an introduction to the sampling ideas as discussed in Chap. 3, we may now represent a signal over length $2L$ and an effective bandwidth $2B$ by means of samples $g(0/2B), g(1/2B), \dots, g((N-1)/(2B))$ with $N \approx 4BL$. The discrete Fourier transform (DFT) of the signal is typically defined as

$$G\left(\frac{m}{2L}\right) = \sum_{n=0}^{N-1} g\left(\frac{n}{2B}\right) \exp(-i2\pi mn/N). \quad (4.1)$$

The corresponding inverse discrete Fourier transform may be defined as

$$g\left(\frac{n}{2B}\right) = \frac{1}{N} \sum_{m=0}^{N-1} G\left(\frac{m}{2L}\right) \exp(i2\pi mn/N). \quad (4.2)$$

We notice that the discrete Fourier transform operations above are approximations to the continuous integral version of the Fourier transform when the signal is defined only over $x : (-L, L)$ by means of a discrete set of samples. The space domain samples have a periodicity of $1/(2B)$ while the frequency domain periodicity is given by $1/(2L)$. The factors $(2B)$ and $(2L)$ are usually omitted from most standard definitions, but we will retain them here explicitly. We observe that the forward and the inverse DFTs are linear transformations which may be written in a matrix form. For example, using the notation

$$\omega = \exp(i2\pi/N), \quad (4.3)$$

for N -th root of 1, the forward transformation may be written as

$$\mathbf{G} = \mathbf{F}\mathbf{g}, \quad (4.4)$$

or

$$\begin{pmatrix} G\left(\frac{0}{2L}\right) \\ G\left(\frac{1}{2L}\right) \\ \vdots \\ \vdots \\ G\left(\frac{N-1}{2L}\right) \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \cdots & \omega^{-(N-1)} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ 1 & \omega^{-(N-1)} & \cdots & \omega^{-(N-1)^2} \end{pmatrix} \begin{pmatrix} g\left(\frac{0}{2B}\right) \\ g\left(\frac{1}{2B}\right) \\ \vdots \\ \vdots \\ g\left(\frac{N-1}{2B}\right) \end{pmatrix}. \quad (4.5)$$

We note that ω satisfies the relation:

$$\omega^{(N-m)} = \omega^{-m}. \quad (4.6)$$

As a result, the frequencies $0/(2L), 1/(2L), \dots, (N-1)/(2L)$ which form the argument of the discrete Fourier transform G may be reordered as positive and negative frequencies. The typical convention used to define positive and negative frequencies is as follows:

1. **Even N case:** $v = 0/(2L)$ is the zero frequency or dc term, $v = 1/(2L)$ to $v = (N/2)/(2L)$ are treated as positive frequencies and $v = (N/2+1)/(2L)$ to $v = (N-1)/(2L)$ are treated as negative frequencies.
2. **Odd N case:** $v = 0/(2L)$ is the zero frequency or dc term, $v = 1/(2L)$ to $v = ((N+1)/2)/(2L)$ are treated as positive frequencies and $v = ((N+3)/2)/(2L)$ to $v = (N-1)/(2L)$ are treated as negative frequencies.

In defining the negative frequencies, the term $G((N-1)/(2L))$ is considered equivalent to having frequency component $v = -1/(2L)$ and so on as per the property of ω in Eq. (4.6).

We make a remark that the spatial frequency locations as obtained by the typical FFT function call are limited by the total extent $2L$ of the signal and the number of

samples N used to represent the signal. If we desire to obtain the Fourier transform at an intermediate spatial frequency location, it is preferable to use the discrete Fourier transform (DFT) as per the relation:

$$G(f_x) = \sum_{n=0}^{N-1} g\left(\frac{n}{2B}\right) \exp[-i2\pi f_x n/(2B)], \quad (4.7)$$

which is valid at arbitrary intermediate spatial frequency locations. Later when discussing numerical processing for off-axis digital holography, we will show that such DFT operation computed near a given integer pixel location $m/(2L)$ in the Fourier domain allows us to determine fractional fringe shift.

4.2 Usage of 2D Fast Fourier Transform for Problems in Optics

Without consideration of any properties of the transformation matrix \mathbf{F} , we observe that computing the discrete Fourier transformation would require N^2 multiplication operations. A great advance was made by Cooley and Tukey in 1965 [3] when they published an algorithm that utilized the symmetry properties of the \mathbf{F} matrix and reduced the number of multiplication operations to the order of $N \log N$. The corresponding algorithm for computing the DFT is known as the Fast Fourier Transform (FFT). The idea behind the FFT algorithm can be traced back in history to a publication by Gauss [4], but the significance of this received attention after the Cooley and Tukey work. FFT has now become a benchmark against which other algorithms are often compared in the context of their computational complexity. We will not provide a detailed discussion of the FFT theory but only introduce the readers to operational aspects of using FFT functionality that is now available in widely used computational tools such as MATLAB, Octave, Scilab and Python-NumPy. More advanced users may also want to consider using the standard FFTW libraries. As we shall see in later chapters of this book, Fourier transforms occur naturally in the study of the propagation of light waves, optical information processing and description of frequency response of imaging systems, to name a few topics. Digital processing methods used for modeling image formation therefore extensively require the use of the FFT algorithm. In these applications, it is important to understand the meaning of results that any standard 2D FFT tool provides and the correct usage of functions such that the results make sense from a Physics point of view.

We will provide an illustration of the computation of the 2D Fourier transform of the 2D rect function. A rect function of size equal to 11×11 pixels is defined on a 2D grid of size 255×255 pixels as shown in Fig. 4.1a. The amplitude and phase of the images obtained by a simple application of the function `fft2` from standard computational tools above are shown in Fig. 4.1b and c, respectively. The result does not look like the expected 2D sinc function. We note that the central lobe of the sinc

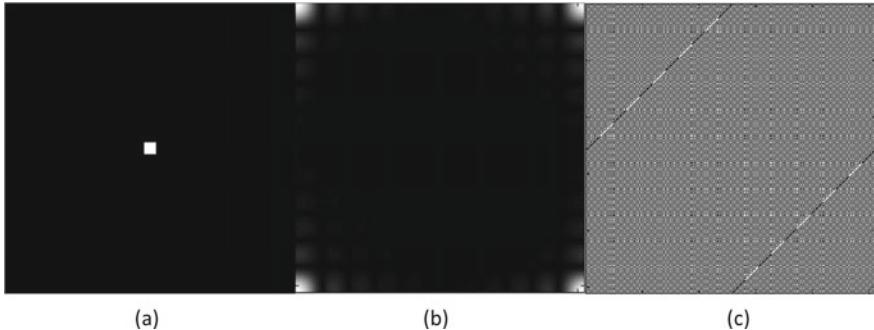


Fig. 4.1 Result of applying the function `fft2(....)` on **a** 2D rect-function object, **b** Amplitude and **c** Phase of Fourier transform. The dark and bright pixels in the phase function correspond to phase values of 0 and π , respectively

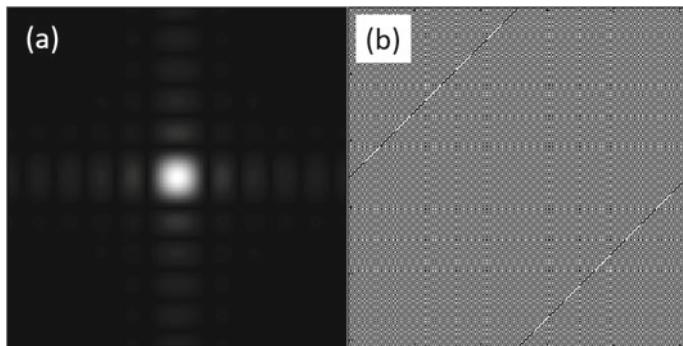


Fig. 4.2 Result of applying the sequence `fftshift(fft2(...))` on 2D rect-function object, **a** Amplitude and **b** Phase of Fourier transform. The dark and bright pixels in the phase function correspond to phase values of 0 and π , respectively

function is now distributed near the corners of the resultant image. The reason for this is that the standard FFT tools have the zero-frequency component at the beginning of the resultant vector and not at its center. A common remedy for this is to use the `fftshift` function on the result which brings the zero frequency back to the center of the image. Applying the sequence `fftshift(fft2(...))` on the rect function gives the amplitude and phase as shown in Fig. 4.2a and b, respectively. The amplitude part here looks sensible, but the phase is still not what one would expect for a sinc function. This unexpected result may be understood as follows. When calculating the Fourier transform as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \text{rect}\left(\frac{x}{a}, \frac{y}{a}\right) \exp[-i2\pi(f_x x + f_y y)], \quad (4.8)$$

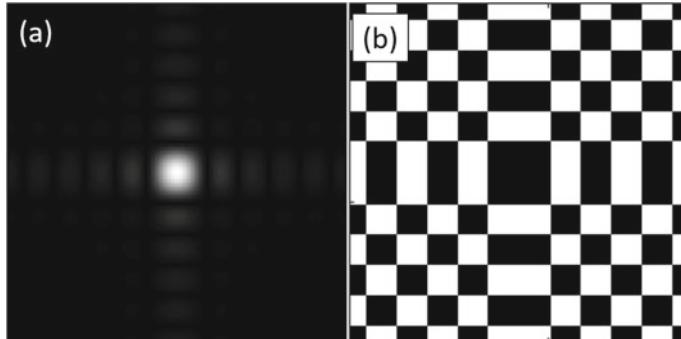


Fig. 4.3 Result of applying the sequence `fftshift(fft2(ifftshift(...)))` on 2D rect-function object, **a** Amplitude and **b** Phase of Fourier transform. The dark and bright pixels in the phase function correspond to phase values of 0 and π , respectively

the correct phase value applied to the center of the rect function corresponds to zero (since $x = 0, y = 0$). However, according to the standard FFT convention, this phase multiplier is applied at the corner of the image. The initial image to be Fourier transformed must therefore be modified appropriately using the standard `ifftshift` function that takes the center of the image to the corners by swapping the quadrants. The appropriate sequence for 2D FFT in most programming platforms for the result to be meaningful from a Physics standpoint (e.g. in describing phenomena-like diffraction) is thus given by `fftshift(fft2(ifftshift(...)))`. The amplitude and phase of the 2D FFT of the rect function computed with this sequence are shown in Fig. 4.3a, b, respectively. Both the amplitude and phase are now what one desires as per conventions in Optics. We mention that the appropriate usage of functions for the inverse FFT operation is similarly: `fftshift(ifft2(ifftshift(...)))`. The functions `fftshift` and `ifftshift` are typically identical for even-dimensional vectors but differ by one pixel shift for odd-dimensional vectors. The two sequences for forward and inverse FFT as described above may be used irrespective of whether the dimensions of vectors involved are even or odd to yield the correct phase behavior as per standard conventions in Optics.

We have provided an operational introduction to FFT routines in computational tools commonly used by researchers. The subtle points regarding the usage of `fftshift` and `ifftshift` functions along with the `fft` routines are illustrated here with examples. These topics have been discussed here since they appear to be common sources of errors when beginning researchers or students start working with FFT-based simulations related to Optical imaging and diffraction phenomena.

Problems

4.1 Interpolation using the FFT function: Based on the definition provided in Sect. 4.1, show that increasing the length of the signal in the coordinate space by zero padding allows us to interpolate the Fourier transform of the signal.

4.2 Evaluating numerical derivatives using FFT: Numerical derivative of a discretely sampled function $g(m\Delta x)$ for integer values of m and sampling interval Δx may be expressed by the central difference formula:

$$\frac{\Delta g}{\Delta x} = \frac{g(m\Delta x - \Delta x) - g(m\Delta x + \Delta x)}{2\Delta x}.$$

Using the shift theorem, show that central differentiation is equivalent to multiplying the FFT of the function $g(m\Delta x)$ by the modified wave number $i \sin(2\pi f_x \Delta x)/(\Delta x)$. Compare your result when you instead use the Fourier domain filter ($i2\pi f_x$) which is used to define the continuous derivative dg/dx . For more details, see [5].

4.3 Using the FFT function available in commonly available tools, explain how you will evaluate the discrete cosine transform (DCT) and the Hilbert transform (HT) of a real-valued image.

4.4 Consider two real-valued images $g_1(x, y)$ and $g_2(x, y)$ of similar class (e.g. face images, natural scenes and text images) and compute their 2D FFTs denoted by G_1 and G_2 . Assuming that the pixel number of the two images is the same along both row/column directions, form two images $g'_1(x, y)$ and $g'_2(x, y)$ as

$$g'_1(x, y) = |\mathcal{F}^{-1}\{|G_1| \exp[i \arg(G_2)]\}|,$$

and

$$g'_2(x, y) = |\mathcal{F}^{-1}\{|G_2| \exp[i \arg(G_1)]\}|.$$

Explain what you observe.

4.5 Verify using 2D FFT computation that (a) translation of an image does not affect its Fourier magnitude, and (b) rotating an image leads to rotation of its Fourier magnitude by the same angle. Are the above properties valid for the phase of the Fourier transform?

References

1. B.G. Osgood, *The Fourier Transform and Its Applications*. Stanford Engineering Everywhere online material. (see.stanford.edu)
2. R. Bracewell, *The Fourier Transform and Its Applications*, 2nd edn. (McGraw Hill, 1986)

3. J.W. Cooley, J.W. Tukey, An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **19**, 297 (1965)
4. M.T. Heideman, D.H. Johnson, C.S. Burrus, Gauss and the history of the Fast Fourier transform. *IEEE ASSP Mag.* **4**, 14 (1984)
5. S. Rajora, M. Butola, K. Khare, Calculating numerical derivatives using Fourier transform: some pitfalls and how to avoid them. *Eur. J. Phys.* **39**, 065806 (2018)

Chapter 5

Linear System Formalism and Introduction to Inverse Problems in Imaging

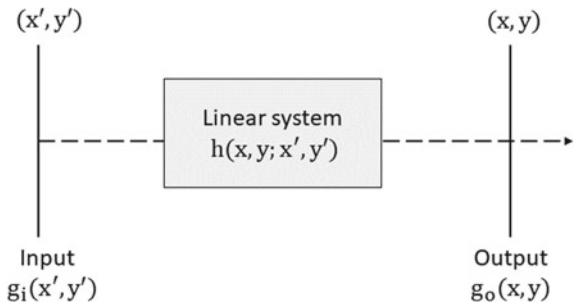


Having introduced Fourier transform theory, ideas such as the delta function (and other generalized functions) and the sampling theorem, we are now in a position to develop the linear systems formalism that will be found useful when studying diffraction, imaging systems and optical information processing, to name a few [1, 2]. A schematic diagram of a linear system is shown in Fig. 5.1. We have an input signal $g_i(x', y')$ that can represent the complex amplitude or intensity of some wavefield (mostly light waves for topics in this book). This input is acted on by some system to give an output $g_o(x, y)$ in another plane. Several topics of interest to this book can be cast in terms of this picture. For now, we will simply suppose the input and output to be complex-valued functions in general without assigning them any physical meaning. The physical meaning will be clear as we study diffraction, imaging and other related topics. The meaning of linearity here is that if we have two inputs g_{1i} and g_{2i} resulting in individual outputs g_{1o} and g_{2o} respectively, then any linear combination $(\alpha g_{1i} + \beta g_{2i})$ when provided at the input end of the system gives an output: $(\alpha g_{1o} + \beta g_{2o})$. Later in this book, we will associate the input and output functions with the light wavefields or irradiances at the input and output of an imaging system. The Maxwell equations which help describe the propagation of light waves (electromagnetic waves) are linear in nature and the importance of linear system model in imaging applications is therefore not very surprising. The linearity property suggests that instead of working with an arbitrary input function, we may want to work with some elementary set of functions that forms a suitable basis for representing the input signal. Analysis of the system then reduces to studying the response of the system to the basis functions.

An interesting basis set is the set of spikes or delta functions located at each point of the input plane. We can for example represent the input as a sum of delta functions:

$$g_i(x', y') = \iint du dv g_i(u, v) \delta(x' - u, y' - v). \quad (5.1)$$

Fig. 5.1 Linear system input-output model



In order to find the output of the system $g_o(x, y)$, all we need to do now is to find the response of the system to a delta spike located at each point (u, v) , weigh the response by the value of the input $g_i(u, v)$ and then integrate over the (u, v) plane. We will denote the response of the system to the delta spike by $h(x, y; x', y')$. In other words, if the input to the system is a delta spike at (x', y') , the response of the system will be given by the function h in the (x, y) plane. Suppose we have a microscope looking at a “point” source, you probably know that the output is some blur (e.g. Airy rings) which in this case is the function h . If we think of the object to be imaged as a collection of point sources, then the resultant image value at the point (x, y) in the output plane will be a weighted sum of blur functions $h(x, y; x', y')$ with weights given by $g_i(x', y')$. We will call $h(x, y; x', y')$ the impulse response of the linear system. The output of the system can therefore be represented in terms of the input as

$$g_o(x, y) = \iint du dv g_i(u, v) h(x, y; u, v). \quad (5.2)$$

The idea of impulse response is common in many branches of engineering and sciences and an alternate name for this concept in the physics literature is Green’s function.

5.1 Space-Invariant Impulse Response

The idea of impulse response is particularly useful when the system has an additional property of space-invariance. A system is space-invariant if the functional form of the impulse response does not depend on where the delta function is located in the input plane. The result of shifting the delta impulse in the input plane is simply to translate the response function $h(x, y; x', y')$ in the output plane—it will not change its form or shape. If the system is space invariant, we can represent the impulse response as

$$h(x, y; x', y') = h(x - x', y - y'). \quad (5.3)$$

Suppose that a point source of unit strength located at $(x', y') = (0, 0)$ produces an output $h(x, y)$. Space-invariance implies that shifting the input point source to $(x', y') = (a, b)$ will produce a shifted output $h(x - a, y - b)$. For most common optical imaging systems, this space-invariance model is a good approximation—the model is however not exact. For a typical imaging system, the space-invariance holds for point sources located near the optical axis. One may observe distortions in the form of h for points far-off from the optical axis. For all practical purposes, we will assume that the space-invariance approximation is valid as this will allow us to use the Fourier transform theory—particularly the convolution property—and develop a formalism for understanding and analyzing imaging phenomena. With the space-invariance property holding true, the input-output relation for linear systems is a convolution relation:

$$g_o(x, y) = \iint dx' dy' g_i(x', y') h(x - x', y - y'). \quad (5.4)$$

An equivalent relation in Fourier space may be written using the convolution property of the Fourier transforms:

$$G_o(f_x, f_y) = H(f_x, f_y) G_i(f_x, f_y). \quad (5.5)$$

Here, G_o , H and G_i denote the 2D Fourier transforms of the output g_o , the impulse response h and the input g_i , respectively. $H(f_x, f_y)$ describes the effect of the system on the input in the spatial frequency space and is commonly referred to as the transfer function of the system. The relation above is very important and central to the study of diffraction and imaging phenomena in general. Note that the multiplication in frequency space is much easier for conceptual understanding than convolution for the design and analysis of optical imaging systems. The transfer function gives the individual spatial frequencies a weight which has in general both an amplitude and a phase. What are the elementary basis functions in this picture? They are simply the Fourier transforms of the shifted delta-function basis—or complex exponentials $\exp[i2\pi(f_x x + f_y y)]$. We can think of the operation of the imaging system as follows—decompose the input and output in terms of the complex exponentials (Fourier transform), and relate the individual spatial frequency components of the input and output by the transfer function. The final output g_o in (x, y) domain may then be calculated by inverse Fourier transforming the resultant product. When space-invariance approximation cannot be made over the entire field of view, the image may be divided into small patches such that the space-invariance approximately holds over an image patch.

5.2 Ill-Posedness of Inverse Problems

With the computational methods forming an integral part of the imaging system, it is very important to discuss some fundamental issues that arise when solving the inverse problem of obtaining the input function g_i using full or partial measurement of the output g_o using some detector. If we associate with g_o the output wavefield, then any optical detector will typically detect the irradiance that is proportional to $|g_o|^2$. For simplicity and for the purpose of discussion here, we may assume that the complex-valued g_o is measurable in the laboratory, e.g. by means of an interferometric arrangement. It is important to note that the measurement process is inherently statistical in nature and what is usually available to us is the output function including noise $n(x, y)$, which in its simplest form may be assumed to be additive. The input-output model for a linear system may thus be stated more explicitly as

$$\tilde{g}_o(x, y) = \iint dx' dy' g_i(x', y') h(x - x', y - y') + n(x, y). \quad (5.6)$$

The noise arises out of a combination of the statistical nature of the detection process as well as the fundamental statistical fluctuations associated with light waves themselves. The noise $n(x, y)$ is never ideally zero in any practical system. Let us further consider the nominal form of impulse response $h(x, y)$ associated with a delta-function input. The transfer function $H(f_x, f_y)$ for any practical system extends over a finite range of spatial frequencies, and the impulse response function $h(x, y)$ has a spread which is inversely related to the spread of the transfer function in spatial frequency space. As a result, the forward process of going from object space to image space is typically a blurring operation. We observe an important property of the inverse problem of obtaining $g_i(x, y)$ based on the measurement $\tilde{g}_o(x, y)$. Consider a solution $g_i(x, y)$ that would ideally produce $g_o(x, y)$ at the output of the system. Interestingly, any other solution of the form $[g_i(x, y) + A \cos(\omega_x x + \omega_y y)]$ for sufficiently large values of ω_x and ω_y produces an output very close to $g_o(x, y)$ since the highly oscillating additive term averages out to zero as a result of the blurring nature of the impulse response function $h(x, y)$. It is not easy to distinguish between such widely differing solutions since they produce almost the same output within noise limits at the output end. This possibility that widely differing input functions may produce practically indistinguishable outputs as measured in a laboratory setting is commonly referred to as the ill-posedness. The ill-posedness as a general property of practical inverse problems was first discussed formally by Hadamard in the 1920s [3]. Careful formulation of the inverse problem is required in order to handle the difficulties due to ill-posedness, so that a physically meaningful solution to the problem at hand may be obtained for a given computational imaging system [4]. We will illustrate some of the important methods in this regard in the following sections.

5.3 Inverse Filter

A simplistic solution to the problem of obtaining the input $g_i(x, y)$ from a measured output \tilde{g}_o is the inverse filter solution. This method is included here as it clearly illustrates the difficulties due to ill-posedness as discussed in the previous section. Disregarding the noise in the output and using Eq. (5.5), one may nominally write the inverse solution as

$$g_i(x, y) = \mathcal{F}^{-1} \left[\frac{G_o(f_x, f_y)}{H(f_x, f_y)} \right]. \quad (5.7)$$

The function $1/H(f_x, f_y)$ is referred to here as the inverse filter. Unfortunately, we do not have access to $G_o(f_x, f_y)$ but only its noisy version $\tilde{G}_o(f_x, f_y) = \mathcal{F}[\tilde{g}_o(x, y)]$. The solution $\tilde{g}_i(x, y)$ estimated using the noisy output may be written as

$$\tilde{g}_i(x, y) = \mathcal{F}^{-1} \left[\frac{\tilde{G}_o(f_x, f_y)}{H(f_x, f_y)} \right]. \quad (5.8)$$

In numerical computation, one may replace any zero value in H with a small constant to avoid dividing by zero. The illustration in Fig. 5.2 shows the effect of a 20×20 pixel square averaging filter (with 1% additive noise) on a picture followed by an attempt at image recovery using the simple inverse filter. The recovery in Fig. 5.2c appears completely meaningless. The reason for this is that we have simply divided the Fourier transform of the blurred image (b) by the Fourier transform $H(f_x, f_y)$. The function $H(f_x, f_y)$ has zeros (or very small values) over finite regions or isolated lines/points, and dividing by these small values greatly enhances the corresponding frequency components. In particular, in the presence of noise (which typically spreads equally over all spatial frequencies) some of these components may produce completely undesirable large oscillations in the recovery as seen in Fig. 5.2c. It is important to note that if the averaging filter $h(x, y)$ is applied to this recovery, it produces a result that is almost the same as the blurred picture in Fig. 5.2b. So it is

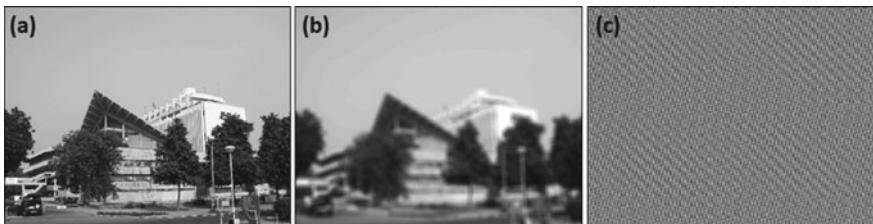


Fig. 5.2 Illustration of simple inverse filter: **a** Original image, **b** Image blurred by 20×20 pixel square averaging filter. Uniform random noise has been added to this image and **c** Recovery using simple inverse filter in Eq. (5.8)

a mathematically valid solution that is consistent with the blurred image data but from a practical standpoint, this recovery is not useful at all. This illustration shows the difficulties associated with the ill-posedness of the inverse problem. The simplistic solution provided by simple inverse filter operation is therefore not useful. A solution must therefore be constructed that gives a visually meaningful result and is reasonably consistent with the blurred image data. We observe that even for a simple-looking problem, a more involved treatment of the inverse problem is required in order to obtain a meaningful image recovery. We remark that the presence of noise in the data and the difficulty associated with de-blurring a noisy image is a problem of fundamental nature that cannot be solved simply by better hardware.

5.4 Wiener Filter

The idea of using the statistical properties of noise to design an appropriate filter for inverse recovery was first proposed by Norbert Wiener [5]. We will assume that a filter $w(x, y)$ is such that when applied to the output $\tilde{g}_o(x, y)$ an estimate of the solution $\tilde{g}_i(x, y)$ given by

$$\tilde{g}_i(x, y) = \iint dx' dy' \tilde{g}_o(x', y') w(x - x', y - y') \quad (5.9)$$

is obtained that is consistent with the noisy data. Further, considering the ambiguity in the solution as discussed in the previous section, we need a criterion for arriving at a desirable solution. One possible criterion is that the expected value of the L2-norm squared error is minimized. We can state this condition as

$$\text{minimize } E^2 = \langle \|g_i(x, y) - \tilde{g}_i(x, y)\|_2^2 \rangle. \quad (5.10)$$

The brackets $\langle \dots \rangle$ denote the ensemble average. The L2-norm of a function $g(x, y)$ denoted by $\|g\|_2$ is defined as

$$\|g\|_2 = \left[\iint dx dy |g(x, y)|^2 \right]^{1/2}. \quad (5.11)$$

In order to derive the filter, we first express the error E^2 in Fourier transform domain by making use of the Parseval theorem (see Eq. (2.73)):

$$\begin{aligned} E^2 &= \langle \|G_i(f_x, f_y) - \tilde{G}_i(f_x, f_y)\|^2 \rangle \\ &= \langle \|G_i(f_x, f_y) - W(f_x, f_y)[G_i(f_x, f_y) H + N(f_x, f_y)]\|^2 \rangle \\ &= \langle \|(1 - W H)G_i - W N\|^2 \rangle. \end{aligned} \quad (5.12)$$

Here, $W(f_x, f_y)$ and $N(f_x, f_y)$ denote the Fourier transforms of the filter $w(x, y)$ and the noise realization $n(x, y)$, respectively. We expand the terms in the above expression and equate the cross terms containing $\langle G_i(f_x, f_y)N^*(f_x, f_y) \rangle$ or its complex conjugate to zero. This is justified since the noise is not expected to be correlated with the input object function $g_i(x, y)$ that we want to recover. Further, we will denote the power spectrum associated with the object and noise as $S_{obj}(f_x, f_y) = \langle |G_i(f_x, f_y)|^2 \rangle$ and $S_{noise}(f_x, f_y) = \langle |N(f_x, f_y)|^2 \rangle$. The squared error as above may now be written more explicitly as

$$\begin{aligned} E^2 = \iint df_x df_y [& |1 - W(f_x, f_y)H(f_x, f_y)|^2 S_{obj}(f_x, f_y) \\ & + |W(f_x, f_y)|^2 S_{noise}(f_x, f_y)]. \end{aligned} \quad (5.13)$$

We note that the error E^2 is a function of both W and W^* . In order to find the appropriate $W(f_x, f_y)$ that minimizes the error, we must equate $\nabla_{W^*} E^2$ to zero, and the resulting solution for $W(f_x, f_y)$ is given by

$$W(f_x, f_y) = \frac{H^*(f_x, f_y)}{|H(f_x, f_y)|^2 + \frac{S_{noise}(f_x, f_y)}{S_{obj}(f_x, f_y)}}. \quad (5.14)$$

For more details on taking a gradient with respect to the conjugate W^* function for minimization purposes, the reader will have to wait till the next chapter on constrained optimization. The typical point spread functions are blurring in nature and as a result, the magnitude $|H(f_x, f_y)|$ fall off at higher spatial frequencies. The de-blurring process must therefore appropriately increase the weight associated with higher spatial frequencies. The filter contains the object and noise power spectra explicitly. The second term in the denominator above is the inverse of signal-to-noise ratio (SNR) at the particular spatial frequency (f_x, f_y) . For spatial frequencies where the SNR is high and $|H(f_x, f_y)|^2$ is much larger than the inverse SNR, we get an expression close to the inverse filter. On the other hand, when the SNR is low, the inverse SNR term may become comparable or larger than $|H(f_x, f_y)|^2$. As a result, the filter will not correctly recover the corresponding spatial frequencies in the original image. Approximate knowledge of the object and noise power spectra is sufficient for most practical purposes. The estimate of object power spectrum may be made based on the average Fourier spectrum associated with images of a particular class (e.g. natural scenery, faces, text), etc. that one may be interested in recovering for a particular problem. The noise spectrum may also be estimated by careful calibration of the imaging system. In Fig. 5.3a, we show image recovery using the Wiener filter for the same example used in Sect. 5.3. We notice that the recovery is now meaningful unlike that in Fig. 5.2c. The filter has recovered the details lost in the blurring process. The result also has a somewhat noisy appearance and suffers from some ringing artifacts near sharp edges as shown in Fig. 5.3b, which is a magnified view of a small portion of the recovery in Fig. 5.3a. These artifacts are typical of the Wiener filter recovery process near high-frequency features like edges.

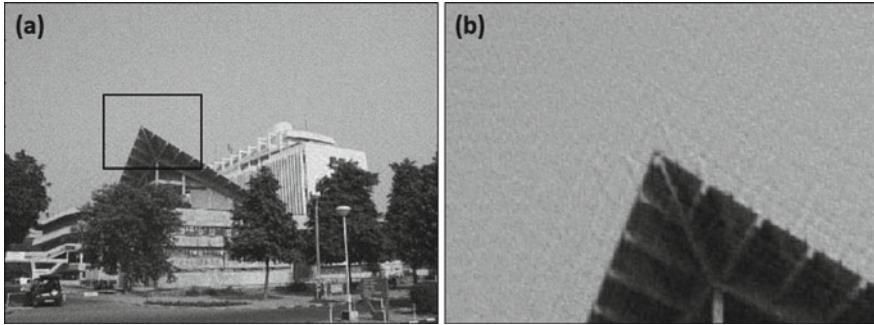


Fig. 5.3 **a** Image recovery using Wiener filter, **b** Magnified portion of the marked rectangle in recovered image

5.5 Generalized Wiener Filter

As discussed in the previous section, the Wiener filter cannot effectively recover spatial frequencies near the zeros of the transfer function $H(f_x, f_y)$. Suppose we image an object using K different imaging systems whose impulse responses may be given by $h_k(x, y)$ with $k = 1, 2, \dots, K$. We therefore have K output images:

$$\tilde{g}_{ok}(x, y) = \iint dx' dy' g_i(x', y') h_k(x - x', y - y') + n_k(x, y). \quad (5.15)$$

We will denote the corresponding transfer functions by $H_k(f_x, f_y)$. In a manner similar to the previous section, our goal now is to determine filters $w_k(x, y)$ that provide an estimate $\tilde{g}_{in}(x, y)$ of the input. The estimated solution is expressed as

$$\tilde{g}_{in}(x, y) = \sum_{k=1}^K [\tilde{g}_{ok} * w_k(x, y)]. \quad (5.16)$$

Here, $*$ represents the convolution operation. Our aim is to find the filters $w_k(x, y)$ such that the expected estimation error

$$E^2 = \langle \|g_{in} - \tilde{g}_{in}\|_2^2 \rangle \quad (5.17)$$

is minimized. Once again, we write the estimated error E^2 in the Fourier transform domain using Parseval's theorem:

$$\begin{aligned}
E^2 &= \left\langle \|G_{in} - \sum_{k=1}^K (G_{in}H_k + N_k)W_k\|_2^2 \right\rangle \\
&= \left\langle \|G_{in} \left[1 - \sum_k H_k W_k \right] + \sum_k N_k W_k\|_2^2 \right\rangle \\
&= \iint df_x df_y \left\{ S_{object} |1 - \sum_k H_k W_k|^2 + S_{noise} \sum_k |W_k|^2 \right\}. \quad (5.18)
\end{aligned}$$

Here, $W_k(f_x, f_y)$ represents 2D Fourier transform of the filter function $w_k(x, y)$. In the above processing steps, we have assumed that the different noise realizations are uncorrelated among themselves and also not correlated with the input object. For simplicity, we have assumed that the noise realizations $n_k(x, y)$ in all the measurements have the same statistical properties and a spectrum denoted by $S_{noise}(f_x, f_y)$. To minimize E^2 , we once again equate $\partial E^2 / \partial W_m^* = 0$, ($m = 1, 2, \dots, K$) leading to

$$S_{object} \left[1 - \sum_k W_k H_k \right] H_m^* = S_{noise} W_m. \quad (5.19)$$

In order to proceed further, one may multiply both sides of the above equation by H_m and sum over all m . Following straightforward rearrangement of terms gives the following expression for the k th filter:

$$W_k(f_x, f_y) = \frac{H_k^*(f_x, f_y)}{\sum_k |H_k(f_x, f_y)|^2 + \frac{S_{noise}(f_x, f_y)}{S_{object}(f_x, f_y)}}. \quad (5.20)$$

The solution $\tilde{g}_{in}(x, y)$ may then be expressed as

$$\tilde{g}_{in}(x, y) = \mathcal{F}^{-1} \left[\sum_{k=1}^K \tilde{G}_{ok}(f_x, f_y) W_k(f_x, f_y) \right]. \quad (5.21)$$

We note that the denominator of Eq. (5.20) now contains a sum $\sum_k |H_k(f_x, f_y)|^2$. If the zeros of individual system transfer functions $H_k(f_x, f_y)$ do not coincide, then the problems associated with division by small numbers do not occur with the generalized Wiener filter. The expressions in Eqs. (5.20) and (5.21) will be utilized in the discussion of structural illumination microscopy reconstruction. We will however illustrate the generalized Wiener filtering for the image shown in Fig. 5.3 where two square impulse responses with size 20×20 pixels and 24×24 pixels have been used for blurring the object. The blurred images are observed in Fig. 5.4a, b, respectively. The generalized Wiener filter recovery is shown in Fig. 5.4c. A part of the recovery is zoomed and shown in Fig. 5.4d. We observe that compared to Fig. 5.3b, the recovered image using the generalized Wiener filter now has significantly reduced

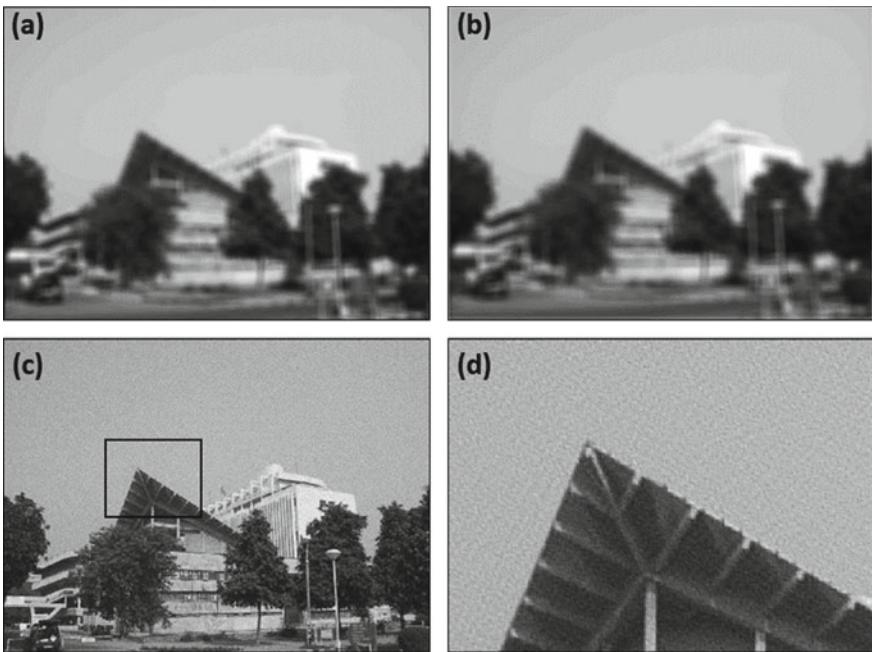


Fig. 5.4 **a, b** Blurred noisy images using two square-shaped filters of size 20×20 pixels and 24×24 pixels respectively, **c** recovery of original object using generalized Wiener filter and **d** magnified portion of the recovered image

ghosting artifacts demonstrating the advantage of having two diverse measurements of the same object over a single measurement.

While Wiener filter [5] or its generalized version [6] takes into account the noise power spectrum, the perceived quality of the recovered image is not necessarily optimal. For example, images with low local noise that still retain sharp edges give better perception of visual quality. Requirements of this kind may be incorporated into the reconstruction algorithm by selecting appropriate constraints as we will describe in the next chapter.

Problems

- 5.1** Computationally verify that the image recovery due to a simplistic inverse filter leads to meaningless recovery as shown in Fig. 5.2 for a test image of your choice. Further observe the image obtained by convolution of the apparently meaningless inverse filter recovery with the blurring function.

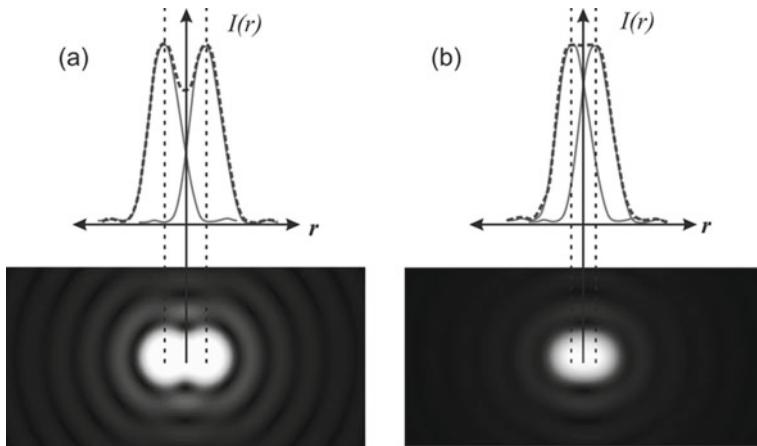


Fig. 5.5 Rayleigh resolution criterion

5.2 Computationally verify the illustrations in Figs. 5.3 and 5.4 for any image of your choice. Add pseudo-random noise to the blurred images in your simulation. Observe the effect of varying the 1/SNR factor in the denominator of the filters used.

5.3 A camera mounted sideways on a fast-moving car captures an image. Describe what the form of the point spread function of the imaging system is. Compute the Wiener filter associated with the corresponding point spread function.

5.4 A camera with a point spread function $h(x, y)$ is used to record two images of the same scene which are to be de-blurred later using a Wiener filter. The two images are recorded with average light levels of 10^4 photons/pixel and 10^2 photons/pixel. Which de-blurred image (after Wiener filtering) do you expect to have a better resolution? Explain in detail and verify computationally with a suitable image and a Gaussian impulse response with 5 pixel full width at half maximum.

5.5 Traditionally, imaging systems are assumed to be limited by the two-point Rayleigh resolution criterion. When two closely located point sources are imaged, the recorded image may look like that shown in Fig. 5.5a, b. The two points are considered to be just resolved when the peak of the pattern due to one point source is co-located with the first minimum of the pattern due to the second source. What do you think is the difference between the Rayleigh resolution and computational resolution which may be achieved after deconvolution? What do you think is the main limitation on the achievable computational resolution?

References

1. A.V. Oppenheim, *Signals and Systems* (Prentice-Hall, USA, 1982)
2. R.C. Gonzalez, R.E. Woods, *Digital Image Processing* (Prentice-Hall, USA, 2007)
3. M. Bertero, P. Boccacci, *Introduction to Inverse Problems in Imaging* (IOP, UK, 1998)
4. A.N. Tikhonov, Y.V. Arsenin, *Solutions of Ill-Posed Problems* (Halsted Press, New York, 1977)
5. C.W. Helstrom, Image restoration by the method of least squares. *J. Opt. Soc. Am.* **57**, 297–303 (1967)
6. L.P. Yaroslavsky, H.J. Caulfield, Deconvolution of multiple images of the same object. *Appl. Opt.* **33**, 2157–2162 (1994)

Chapter 6

Optimization Approach to Image Reconstruction



We have already stated in the first chapter of this book that designing algorithms for the inverse or image reconstruction problem using measured sensor data is an integral part of every computational system. The difficulties in solving them via linear filtering approaches were discussed in the last chapter. The ill-posed nature of image reconstruction problems means that they inherently have multiple solutions, and in order to select the optimal solution, they are now increasingly modeled as optimization problems with appropriate constraints [1–4]. In this chapter, we provide an introduction to formulating the inverse problem, the role of constraints and selection of free parameters. We discuss some simple approaches for solving the inverse problems and further provide pointers to more advanced techniques for the interested reader. We will start our discussion with the image denoising problem followed by image deconvolution and finally conclude with introduction to the exciting recent ideas such as compressive imaging . While a variety of constraint choices may be used for inverse imaging problems, we will consider the total variation penalty as a concrete example in this chapter. This choice of constraint has been made here since it has proved to be very powerful from image quality point of view and secondly it does not require any prior knowledge of topics such as wavelets or similar transforms that are out of the scope of this book [5].

6.1 Image Denoising

Image or video data forms a major portion of digital data streams. One of the problems often encountered even with well-designed imaging systems is the noisy appearance of recorded images that give a perception of low image quality to the end user. The noise results from various factors, e.g. electronic readout noise from the sensor, statistical fluctuations in number of photons in the illuminating light beam and the

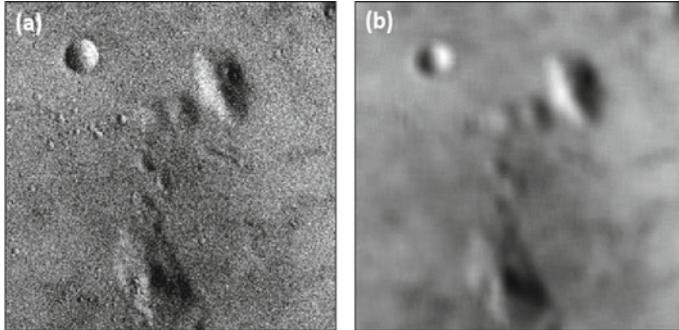


Fig. 6.1 **a** Noisy image (Moon surface) and **b** the result of convolution with a 11×11 square shaped averaging filter

statistical nature of detection process. While one may argue that the noise may be reduced by using sufficiently high light levels, there are applications such as (i) astronomy where light levels may be inherently low or (ii) medical X-ray imaging where increasing X-ray dose is not desirable. Removing noise from an image is thus an important practical problem. While denoising is not a full-fledged inverse problem, it is a good starting point for introducing the optimization framework.

Figure 6.1a, b shows a noisy image and that obtained by convolving it with a square-shaped (11×11 pixels) filter function, respectively. While the filtered image has less noisy appearance, the edges in this filtered image appear blurred, indicating that simple averaging over pixels is not sufficient for our purpose. Given a noisy image, we would like to obtain another image which has smoother appearance but no loss in sharpness or resolution. It is important to remember that edges in an image play an important role in our perception of the image quality. A scheme must therefore be realized such that the small local noise oscillations are suppressed while retaining edges at the same time. Given a noisy image $g(x, y)$, we will now attempt to find another image $\hat{g}(x, y)$ which is close to $g(x, y)$ but has the desired properties mentioned above. A visual examination of Fig. 6.1 suggests that while several structures in this image are clearly visible, the most annoying part about it is its grainy appearance. Our algorithmic procedure for removing noise must therefore impose a penalty on local fluctuations in the image. For example, denoising may be modeled as a problem of minimizing a cost function of the form:

$$C(\hat{g}) = \|g(x, y) - \hat{g}(x, y)\|_2^2 + \alpha \psi(\hat{g}). \quad (6.1)$$

Here, the notation $\|\dots\|_2$ once again denotes the L2-norm, α is a real positive number and ψ is a penalty function or constraint function. The parameter α gives an appropriate weight to the penalty term in relation to the L2-norm squared error term. A powerful choice for ψ is the total variation (TV) functional:

$$\psi(\hat{g}) = TV(\hat{g}) = \|\nabla \hat{g}\|_1 = \sum_{m=\text{all pixels}} |\nabla \hat{g}|_m. \quad (6.2)$$

The TV functional is seen to be the L1-norm of the gradient which is more explicitly defined in the last term above. The subscript m denotes the pixel index in the image. The contribution of every pixel in the TV of any given image is thus proportional to the magnitude of the gradient at that pixel. A penalty may also be designed such that this contribution is proportional to the L2-norm square of the gradient. Making the penalty in the form of L1-norm, however, gives relatively more weight to the small local oscillations due to noise. Since edges occupy relatively few pixels in a typical image, the contribution of small local oscillations to the penalty function is usually higher compared to that due to edges when using L1-norm penalty. Any numerical procedure for minimizing the cost function in Eq. (6.1) thus gives a lot of importance to small local oscillations while preserving edges. We will now describe a simple gradient descent procedure to minimize the cost function in Eq. (6.1). Given a particular guess solution \hat{g} of the problem, we wish to consider a small variation $\delta\hat{g}$ in it such that

$$C(\hat{g} + \delta\hat{g}) < C(\hat{g}). \quad (6.3)$$

This inequality may be achieved if we make $\delta\hat{g}$ proportional to the negative functional gradient of the cost function. The iteration using gradient descent thus progresses as

$$\hat{g}^{(n+1)} = \hat{g}^{(n)} - \tau [\nabla_{\hat{g}} C(\hat{g})]_{\hat{g}=\hat{g}^{(n)}}. \quad (6.4)$$

Here, τ is the step size in direction of negative gradient. The numerical value of τ is determined for each iteration to make sure that the cost function does reduce as per Eq. (6.3). For example, one may start with some nominal value $\tau = 1$ and get a test guess for $\hat{g}^{(n+1)}$. If this guess solution does not satisfy Eq. (6.3), τ is reduced to, say, $\tau/2$ and so on till this condition is satisfied. This procedure is known as simple backtracking line search [3]. A more involved strategy for selection of step size τ is known as the Armijo search where we look for a step size where the cost function reduces at least as much as a fraction of the first-order Taylor expansion for $C(\hat{g}^{(n+1)})$. The condition may thus be stated as

$$C(\hat{g}^{(n+1)}) < C(\hat{g}^{(n)}) - c_1 \tau \|[\nabla_{\hat{g}} C(\hat{g})]_{\hat{g}=\hat{g}^{(n)}}\|_2^2, \quad (6.5)$$

for a constant $c_1 \in (0, 1)$. The aim of the Armijo rule is to ensure that the selected step size τ is not too small and provides a sufficient decrease in the cost function. There is an additional stronger Wolfe condition for line search which we will not discuss here. Before proceeding with the denoising problem further, we will need to digress a bit in order to explain the meaning of the functional gradient $[\nabla_{\hat{g}} C(\hat{g})]$ and how one may evaluate it in practice when dealing with images.

6.1.1 Euler-Lagrange Problem in Variational Calculus

We present a brief discussion of a general method for optimization of real positive-valued cost functionals of the form:

$$C(\hat{g}) = \iint dx dy \mathcal{L}(\hat{g}, \nabla_x \hat{g}, \nabla_y \hat{g}). \quad (6.6)$$

The cost function in Eq. (6.1) is already in this form. We will encounter this form from time to time in this book. We start by noting that the directional derivative of $C(\hat{g})$ in the direction of another image vector \hat{h} is defined by the following relation:

$$\begin{aligned} C(\hat{g} + \epsilon \hat{h}) &= C(\hat{g}) + \epsilon (\nabla_{\hat{g}} C) \cdot \hat{h} + O(\epsilon^2) \\ &= C(\hat{g}) + \epsilon \iint dx dy [\nabla_{\hat{g}} C](x, y) \hat{h}(x, y) + O(\epsilon^2). \end{aligned} \quad (6.7)$$

Here, the notation $O(\epsilon^2)$ indicates terms of the order of ϵ^2 which are negligibly small in the limit $\epsilon \rightarrow 0$. Since we need to find how one may determine the quantity $[\nabla_{\hat{g}} C]$, we must perform a variation in Eq. (6.6) and find the required relation.

For simplicity, we will consider the 1D case of minimizing a cost function of the form:

$$C(\hat{g}) = \int dx \mathcal{L}(\hat{g}, \hat{g}_x), \quad (6.8)$$

where $\hat{g}_x = \nabla_x \hat{g}$. Performing the variation of \hat{g} gives

$$\begin{aligned} C(\hat{g} + \epsilon \hat{h}) &= \int dx \mathcal{L}(\hat{g} + \epsilon \hat{h}, \hat{g}_x + \epsilon \hat{h}_x) \\ &= \int dx \left[\mathcal{L}(\hat{g}, \hat{g}_x) + \epsilon \frac{\partial \mathcal{L}}{\partial \hat{g}} \hat{h} + \epsilon \frac{\partial \mathcal{L}}{\partial \hat{g}_x} \hat{h}_x \right]. \end{aligned} \quad (6.9)$$

The last term in the integrand above can be simplified using integration by parts:

$$\int dx \frac{\partial \mathcal{L}}{\partial \hat{g}_x} \hat{h}_x = \left[\frac{\partial \mathcal{L}}{\partial \hat{g}_x} \hat{h} \right]_{\text{boundaries}} - \int dx \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial g_x} \right) \hat{h}. \quad (6.10)$$

Assuming that the variation \hat{h} is zero at the integration limits, we have

$$C(\hat{g} + \epsilon \hat{h}) = C(\hat{g}) + \epsilon \int dx \left[\frac{\partial \mathcal{L}}{\partial \hat{g}} - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial g_x} \right) \right] \hat{h}. \quad (6.11)$$

Comparing with Eq. (6.7) yields the well-known result:

$$\nabla_{\hat{g}} C = \left[\frac{\partial \mathcal{L}}{\partial \hat{g}} - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial \hat{g}_x} \right) \right]. \quad (6.12)$$

A 2D version of the result may be stated as follows:

$$\nabla_{\hat{g}} C = \left[\frac{\partial \mathcal{L}}{\partial \hat{g}} - \frac{\partial}{\partial x} \left(\frac{\partial \mathcal{L}}{\partial \hat{g}_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial \mathcal{L}}{\partial \hat{g}_y} \right) \right] = \left[\frac{\partial \mathcal{L}}{\partial \hat{g}} - \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial \nabla \hat{g}} \right) \right]. \quad (6.13)$$

The minimum of the cost function as given in Eq. (6.1) is achieved when the functional gradient defined in Eq. (6.13) is zero, so that, we have reached a stationary point of the cost function to first order in ϵ . Practically one may stop an iteration process as in Eq. (6.4) when the relative change in the solution in successive iterations as given by $\|\hat{g}^{(n+1)} - \hat{g}^{(n)}\|_2 / \|\hat{g}^{(n)}\|_2$ is below a pre-defined threshold. More specifically, our goal is to determine the functional gradient for the cost function in Eq. (6.1). As per Eq. (6.13), the functional gradient contribution of the first term in Eq. (6.1) is simple to evaluate. For the second term involving the total variation, we may however make an approximation by introducing a small constant δ in the definition as

$$TV(\hat{g}) \approx \sum_m \sqrt{(\nabla_x \hat{g})_m^2 + (\nabla_y \hat{g})_m^2 + \delta^2}. \quad (6.14)$$

Here, δ is a constant that may be orders of magnitude smaller than the typical gradient magnitudes involved in the problem. Introduction of this constant does not affect the numerical value of $TV(\hat{g})$ significantly; however, this modification is amenable to differentiation required in the definition Eq. (6.13). With this approximation, the functional gradient corresponding to the cost function in Eq. (6.1) is given by

$$\nabla_{\hat{g}} C(\hat{g}) = -2(g - \hat{g}) - \alpha \nabla \cdot \left[\frac{\nabla \hat{g}}{\sqrt{|\nabla \hat{g}|^2 + \delta^2}} \right]. \quad (6.15)$$

Note that the required derivatives in the above expression may simply be evaluated using central differences over rows or columns of the image matrix, e.g.,

$$[\nabla_x \hat{g}]_{m,n} = \frac{g(m, n+1) - g(m, n-1)}{2}, \quad (6.16)$$

and

$$[\nabla_y \hat{g}]_{m,n} = \frac{g(m+1, n) - g(m-1, n)}{2}, \quad (6.17)$$

with one-sided differences applied at the image boundaries. Evaluating the expression for functional gradient thus involves simple algebraic steps applied to the pixels

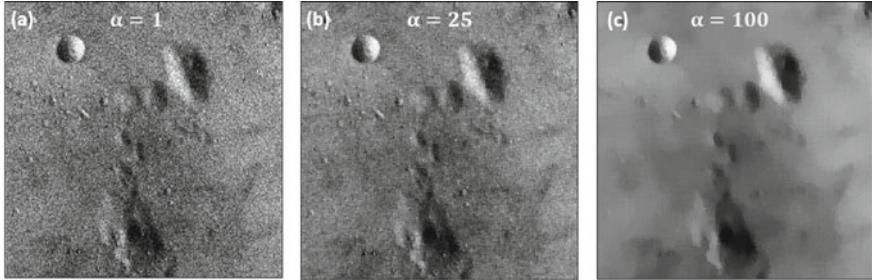


Fig. 6.2 Denoised version of image in Fig. 6.1a by iterative TV-based denoising for three values of the parameter $\alpha = 1, 25, 100$

of the image matrix. It may be noted that the expression for $\nabla_{\hat{g}} C$ contains a free parameter α . The choice of α thus determines the nature of solution. The optimization formulation thus provides a family of solutions, and an appropriate solution of this family needs to be selected based on requirements of the underlying application. In Fig. 6.2, we show result of iterations as in Eq. (6.4) for three values of the parameter $\alpha = 1, 20, 100$. For each of the three cases, 20 iterations of the iterative procedure have been used. Compared to Fig. 6.1a, we observe that the result in Fig. 6.2a for $\alpha = 1$ has not changed much in terms of noise reduction. The result for $\alpha = 100$ shown in Fig. 6.2c appears to be overly smoothed. The result for the intermediate value $\alpha = 25$ in Fig. 6.2b has noise reduction as well as retention of major features. The edge preservation is in fact superior in all the three cases in Fig. 6.2 when compared to that in Fig. 6.1b. It is important to note that the edges have been preserved but the high-frequency noise has been nearly eliminated. This may be understood if we take a look at the functional gradient for the TV penalty part (second term in Eq. (6.15)). The expression

$$\nabla_{\hat{g}} TV(\hat{g}) = -\nabla \cdot \left[\frac{\nabla \hat{g}}{\sqrt{|\nabla \hat{g}|^2 + \delta^2}} \right] \quad (6.18)$$

has a form of anisotropic diffusion. *The TV reduction procedure therefore diffuses the individual pixel values to its neighboring pixels such that the diffusion is inversely proportional to the local gradient magnitude at a particular pixel.* The edge preserving smoothing seen in Fig. 6.2 is a result of this anisotropic diffusion property.

Some discussion on selection of the α parameter is required here as this choice can be seen to affect the nature of the recovered image. While it is difficult to make a general statement regarding which of the three images in Fig. 6.2 will be suitable for a given application, generically one may evaluate the quantity $E(\alpha) = \|g - \hat{g}_\alpha\|_2$ where \hat{g}_α denotes the denoised solution for a particular choice of α . This choice of α may be considered suitable if the error $E(\alpha)$ is approximately of the same order as the L2-norm of the noise. This method of selecting the free parameter α is the essence of the Morozov discrepancy principle [1]. The main idea behind the suitability of

this choice is that, if the value of $E(\alpha)$ is much below or above the noise norm, this implies that we are either trying to match the noise (Fig. 6.2a) or over-smoothing (Fig. 6.2b) the input noisy image, respectively. In the illustration in Fig. 6.2, the quantity $E(\alpha)$ for the three values of α is observed to be 0.04, 0.18 and 0.30 times the norm $\|g\|_2$ of the original noisy image in Fig. 6.1a. In order to estimate the noise level, we may nominally take a small flat region in the original image where there is no visible structure. The mean signal value in this region is a measure of the signal level whereas the standard deviation of pixel values in this region can be related to the noise norm. The estimated noise level by this approximate methodology is observed to be $0.2\|g\|_2$ indicating that the choice $\alpha = 25$ in the illustration in Fig. 6.2 has provided a reasonable solution in accordance with the discrepancy principle. Before proceeding further, we will provide a brief description of how complex-valued images may be handled in the optimization framework. The complex-valued solutions will be seen to be found useful later in the book particularly for coherent imaging systems that aim to measure both amplitude and phase of the image.

6.2 Functional Gradient for Complex-Valued Solutions

We will encounter a number of cases where the image solution $\hat{g}(x, y)$ we are seeking is complex-valued, while the cost function is real and positive. The cost function will often in fact be a function of both \hat{g} and \hat{g}^* . Most of the discussion in typical optimization literature deals with real-valued variables and therefore it is important to understand how one needs to proceed in order to compute the steepest descent direction for updating the iterative solution. One possible way to handle this situation is to treat the problem as having twice as many variables consisting of the real and imaginary parts of the solution. However, when optimizing separately over real and imaginary parts, the expressions for the functional derivatives usually become cumbersome. It is also not clear what is the best way to combine the real and imaginary parts that have been optimized separately. An interesting solution to this problem involves what is known as complex or Wirtinger derivatives. The idea behind the complex derivatives is quite simple as we will explain below. We observe that the variation of the cost function $C(\hat{g}, \hat{g}^*)$ may be written as

$$\begin{aligned} C(\hat{g} + \epsilon h, \hat{g}^* + \epsilon h^*) - C(\hat{g}, \hat{g}^*) &= \epsilon [\nabla_{\hat{g}} C \cdot h + \nabla_{\hat{g}^*} C \cdot h^*] \\ &= 2\epsilon \operatorname{Re} [\nabla_{\hat{g}} C \cdot h] \\ &= 2\epsilon \operatorname{Re} \langle \nabla_{\hat{g}^*} C, h \rangle. \end{aligned} \quad (6.19)$$

In the last line above, we used the fact that the cost function $C(\hat{g}, \hat{g}^*)$ is real-valued so that $\nabla_{\hat{g}} C = (\nabla_{\hat{g}^*} C)^*$. In order to find the steepest descent direction, we need to select the variation h in the direction $-\nabla_{\hat{g}^*} C$, so that there will be

maximum decrease in the cost function. The gradient of $C(\hat{g}, \hat{g}^*)$ with respect to \hat{g}^* is known as the Wirtinger derivative [6]. Following Eq. (6.13), the functional gradient may be expressed as

$$\nabla_{\hat{g}^*} C = \left[\frac{\partial \mathcal{L}}{\partial \hat{g}^*} - \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial \nabla \hat{g}^*} \right) \right]. \quad (6.20)$$

With the Wirtinger derivative in hand, the complex-valued optimization problem can essentially be treated in the same way as the real-valued case, except that the derivative is now evaluated with respect to the complex conjugate image \hat{g}^ .* We will have an occasion to use this idea when discussing the phase imaging problem as well as in the case of Fourier measurements.

6.3 Image Deconvolution by Optimization

In this section, we add some complexity to the denoising problem and study the image deconvolution or de-blurring problem with TV penalty as a constraint term. The data in this case is a blurred image with noise which we will try to recover. The data model is as follows:

$$\tilde{g}_o(x, y) = g_i(x, y) * h(x, y) + n(x, y), \quad (6.21)$$

which may also be represented in a matrix operator form as

$$\tilde{g}_o = \hat{A}g_i + n, \quad (6.22)$$

with \hat{A} denoting the system or blurring linear operator. In a fashion similar to the previous section, we model the optimization problem where the cost function is defined as

$$C(g_i) = C_1 + \alpha C_2 = \|\tilde{g}_o - \hat{A}g_i\|_2^2 + \alpha TV(g_i). \quad (6.23)$$

We already know the functional gradient associated with the TV penalty denoted by C_2 in the equation above. In order to compute the functional gradient for the error term C_1 , we proceed as follows:

$$\begin{aligned} C_1(g_i + \epsilon q) &= \sum_m [\tilde{g}_o - \hat{A}(g_i + \epsilon q)]_m^2 \\ &= C_1(g_i) - 2\epsilon \langle (\tilde{g}_o - \hat{A}g_i), (\hat{A}q) \rangle + O(\epsilon^2). \end{aligned} \quad (6.24)$$

The notation $\langle \dots, \dots \rangle$ above denotes the scalar product defined as follows:

$$\langle g_1, g_2 \rangle = \iint dx dy g_1^*(x, y) g_2(x, y). \quad (6.25)$$

We now observe that

$$\begin{aligned} \frac{C_1(g_i + \epsilon q) - C_1(g_i)}{\epsilon} &= -2\langle (\tilde{g}_o - \hat{A}g_i), (\hat{A}q) \rangle + O(\epsilon^2) \\ &= -2\langle \hat{A}^\dagger(\tilde{g}_o - \hat{A}g_i), (q) \rangle + O(\epsilon^2). \end{aligned} \quad (6.26)$$

Here, the notation \hat{A}^\dagger refers to the Hermitian transpose of the operator \hat{A} . This expression has been obtained by using the well-known property of scalar products. We can now write an expression for the functional gradient of C_1 by comparison with Eq. (6.7):

$$\nabla_{g_i} C_1 = -2\hat{A}^\dagger(\tilde{g}_o - \hat{A}g_i). \quad (6.27)$$

From implementation standpoint, we need to understand the meaning of the operation \hat{A}^\dagger corresponding to convolution by the point spread function $h(x, y)$ as in Eq. (6.23). To this end, we once again use the scalar product identity:

$$\langle v_1, \hat{A}u_1 \rangle_{data} = \langle \hat{A}^\dagger v_1, u_1 \rangle_{object}. \quad (6.28)$$

Here, the subscripts on the scalar product $\langle \dots, \dots \rangle$ above imply that the scalar products have been evaluated over the data and object space coordinates, respectively. In particular,

$$\begin{aligned} \langle v_1, \hat{A}u_1 \rangle_{data} &= \iint dx dy v_1^*(x, y) \iint dx' dy' u_1(x', y') h(x - x', y - y') \\ &= \iint dx' dy' [\iint dx dy v_1(x, y) h(x - x', y - y')]^* u_1(x', y'). \end{aligned} \quad (6.29)$$

Comparing with Eq. (6.28), we have

$$\hat{A}^\dagger v_1 = \iint dx dy v_1(x, y) h^*(x - x', y - y') = v_1(x, y) * h^*(-x, -y). \quad (6.30)$$

The \hat{A}^\dagger operation thus corresponds to convolution with the complex conjugated and inverted impulse response $h^*(-x, -y)$. The overall functional gradient needed for the de-blurring cost function in Eq. (6.23) may therefore be written as

$$\nabla_{g_i} C = -2[\tilde{g}_o(x, y) - \hat{A}g_i(x, y)] * h^*(-x, -y) - \alpha \nabla \cdot \left[\frac{\nabla g_i}{\sqrt{|\nabla g_i|^2 + \delta^2}} \right]. \quad (6.31)$$

Once again all the operations in the above equation are straightforward to compute and an iterative procedure as in Eq. (6.4) may be implemented readily for obtaining

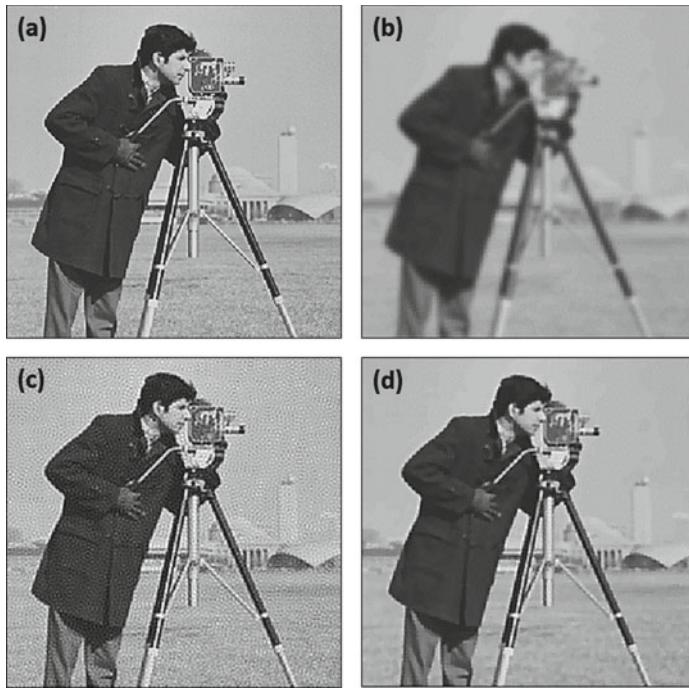


Fig. 6.3 **a** Original image, **b** blurred noisy image, de-blurring using **c** Wiener filter and **d** iterative deconvolution with total variation penalty

the de-blurred image. In Fig. 6.3, we show an illustration of image de-blurring for the standard cameraman phase image. The original and blurred images (obtained using a Gaussian function) are shown in Fig. 6.3a, b, respectively. The Wiener filter recovery and the iterative total variation based recovery are shown in Fig. 6.3c, d, respectively. The Wiener filter recovery has some grainy artifacts that are removed in the iterative recovery; otherwise, the resolution of features in the two reconstructions is similar. The numerical value of the parameter α was 0.01 in this illustration and the Wiener filter recovery in Fig. 6.3c was used as the initial guess for the iterative deconvolution. While we have discussed the simple gradient descent iteration in this example, iterative schemes that offer better convergence properties like the conjugate gradient or the Nesterov gradient method or Newton iteration may be used [2, 3]. These methods will also need the computation of functional gradient as in Eq. (6.31). The iterative optimization based solution generally provides improved image quality when the solution is constrained with the appropriate penalty functions. While image quality is a topic beyond the scope of this book, it is important to note that in several applications like diagnostic imaging, the image quality perceived by the end user plays a very important role in design of image recovery algorithms.

6.4 Compressive Imaging

In this section, we turn to an interesting application—compressive imaging—of the constrained optimization framework. From Fig. 6.3c, d, we observe that the Wiener filter solution has grainy features and thus contains more local fluctuations compared to the iterative solution that uses a total variation penalty. The iterative solution is more pleasing to eyes because of its flatter and smooth nature. Mathematically, if we want to represent such a solution in any basis, it is clear that it will have less degrees of freedom and will require less number of coefficients compared to the Wiener filter solution with grainy artifacts. In fact, research work on image compression standards over the decades has shown that natural images are compressible [7]. In other words, the number of degrees of freedom in a typical image is often much less than the number of “visual” pixels used to represent it. For example, if we inspect the image in Fig. 6.3a, we observe that there are several flat or smooth regions where the numerical values are locally almost constant or slowly varying from pixel to pixel. These correlations in pixel values may be used efficiently for storage of images as is now done in widely used image compression standards such as JPEG2000. If most images to be recorded in some application are expected to have some structure and hence such compressibility, a natural question to ask is as follows: *How many measurements are required to generate an N -pixel image?* For the imaging model in Eq. (6.22), the simple and conventional answer to this question (disregarding any compressibility in the image g_i) is that we will need approximately N measurements to generate an N pixel image. The argument here is that Eq. (6.22) is a linear system of equations, and in order to find N unknown pixel values, we need to have as many equations as the number of unknowns. However, if we suppose that the image to be recovered is compressible due to its structure, then it is conceivable intuitively that number of measurements much less than N should be sufficient. We provide a concrete example of Fourier domain measurements to illustrate this point.

Fourier domain measurements occur commonly in optics, medical imaging modalities such as MRI and X-ray CT and radio astronomy to name a few applications. Consider an image $g(x, y)$ and its Fourier transform $G(f_x, f_y)$ as shown in Fig. 6.4a, b, respectively. If the full data $G(f_x, f_y)$ (both amplitude and phase) is measured at N pixels in the Fourier domain, it is straightforward to invert it to get the image $g(x, y)$. However, we note that the image $g(x, y)$ is sparse in the gradient domain. If we inspect the gradient magnitude image as in Fig. 6.4c, we observe that the number of pixels with significant gradient magnitude is much less than the total number of pixels in the image. Can we utilize this redundancy in the image expressed via gradient domain sparsity in some way? For example, we may measure only $N/4$ Fourier domain samples randomly at locations shown as bright pixels in a binary $(0, 1)$ mask $M(f_x, f_y)$ shown in Fig. 6.5a. The corresponding masked Fourier data is shown in Fig. 6.5b. The $N/4$ data points in Fourier domain are clearly not sufficient in the conventional sense of solving for the N unknown pixel values in the image $g(x, y)$. Fourier transform being a linear transform, we now need to solve for N unknown pixels from $N/4$ linear equations. In principle, such an incomplete

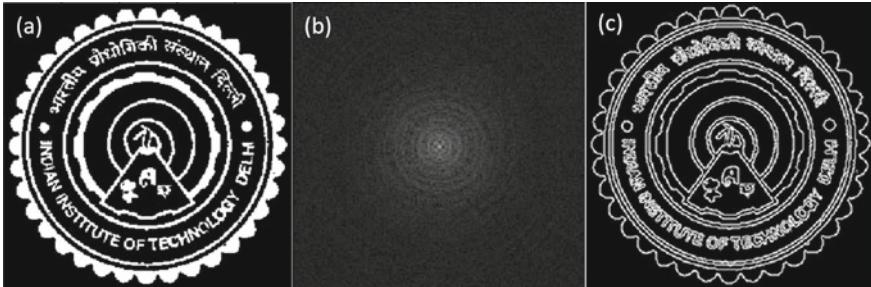


Fig. 6.4 **a** Logo image $g(x, y)$, **b** Fourier transform magnitude $|G(f_x, f_y)|$ shown on a log-scale, **c** Gradient magnitude image $|\nabla g(x, y)|$

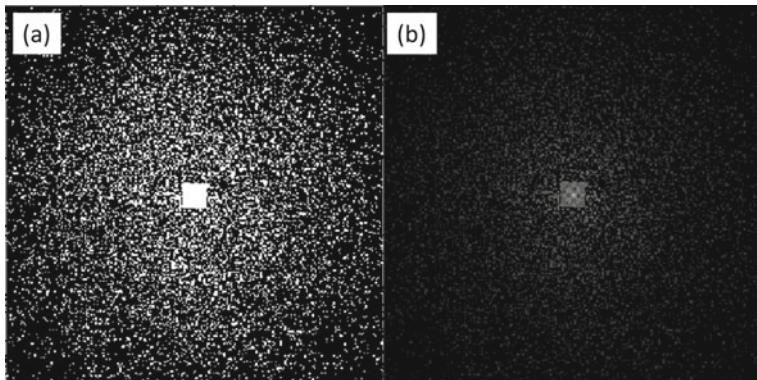


Fig. 6.5 **a** Binary mask $M(f_x, f_y)$ showing locations where Fourier data is measured, **b** Fourier transform magnitude $|M(f_x, f_y)G(f_x, f_y)|$ shown on a log-scale

system of linear equations has infinite number of solutions. To find an appropriate solution, we need some additional criteria to decide on which solution to select. As a first attempt, we consider the minimum L2-norm squared solution by minimizing the cost function:

$$C(g_i, g_i^*) = \|M(f_x, f_y)[G(f_x, f_y) - \mathcal{F}\{g_i(x, y)\}]\|^2. \quad (6.32)$$

Note that here we are looking for the best solution in the least-square sense that has a Fourier transform close to $G(f_x, f_y)$ at the measurement locations in the Fourier domain as per the mask definition. The solution for this minimization problem is shown in Fig. 6.6. We observe that the least-square solution is clearly of poor quality and is not useful in any practical application. However, since we expect the solution to be sparse in gradient domain, we may incorporate this information in the cost function by introducing the TV penalty in the cost function:

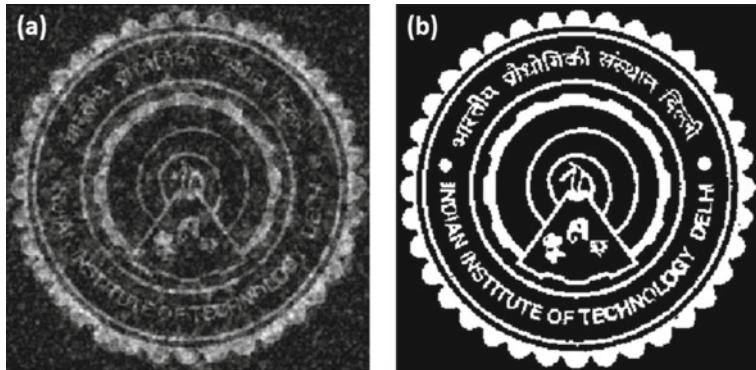


Fig. 6.6 Image recovery for **a** minimum squared L2-norm problem as in Eq. (6.32), **b** minimum squared L2-norm problem with TV penalty as in Eq. (6.33)

$$C(g_i, g_i^*) = \|M(f_x, f_y)[G(f_x, f_y) - \mathcal{F}\{g_i(x, y)\}]\|^2 + \alpha TV(g_i, g_i^*). \quad (6.33)$$

The solution to the corresponding problem with TV penalty incorporated is shown in Fig. 6.6b where we observe near-perfect image recovery with 25% measurements. Both the optimization problems above are solved using the iteration scheme described in Eq. (6.4) and Eq. (6.27). The value of the free parameter $\alpha = 0.01$ was used in this illustration. The adjoint operator A^\dagger in this illustration corresponds to the inverse Fourier transform of the Fourier data restricted by the mask $M(f_x, f_y)$.

6.4.1 Guidelines for Sub-sampled Data Measurement and Image Recovery

While we have seen excellent image recovery with data that may be considered incomplete in traditional sense, the choice of transform and the measurement scheme need some further discussion. For simplicity, we will assume that the solution $g(x, y)$ can be expanded in terms of a basis set of functions $\{\phi_n(x, y)\}$:

$$g(x, y) = \sum_n \beta_n \phi_n(x, y). \quad (6.34)$$

Further, we assume that the choice of basis is such that the vector $\boldsymbol{\beta}$ made up of the coefficients β_n is sparse in nature. In other words, a small number of coefficients β_n compared to N are significant. The choice of basis can be typically a scale-based basis such as some wavelet family that is known to be efficient for representation of natural images. The imaging model used is as in Eq. (6.22):

$$g_o = \hat{A} \sum_n \beta_n \phi_n(x, y) + n(x, y). \quad (6.35)$$

The compressive sensing problem refers to recovery of the coefficients β_n under a prior constraint that $\boldsymbol{\beta}$ is a sparse vector. As shown by Candes et al. in [7], this problem may be modeled as a minimization problem for a cost function given by

$$C(\boldsymbol{\beta}) = \|g_o - \hat{A} \sum_n \beta_n \phi_n(x, y)\|^2 + \alpha \|\boldsymbol{\beta}\|_1. \quad (6.36)$$

Here, the sparsity in components of $\boldsymbol{\beta}$ is achieved by imposing an L1-norm penalty. More explicitly, $\|\boldsymbol{\beta}\|_1 = \sum_n |\beta_n|$. The TV penalty used for the illustration in Fig. 6.6 is an example of L1-norm penalty as well, except that the image representation is not in terms of a basis set of functions.

The two popular choices of regularization are L1-norm and L2-norm. While we continue to use TV in our discussion which is an L1-norm type penalty, here we show pictorially why L2-norm fails to give a sparse solution on the other hand L1-norm provides a sparse solution. Generally, Eq. (6.34) is to be solved in N-dimension for an image represented over N-pixels grid. However, for the ease of understanding, we take a case of a 2D space as a pictorial representation in higher dimensions will be very complicated and hard to imagine. Figure 6.7a, b shows the L2-norm and L1-norm spheres, respectively. The line “L” is the set of solutions satisfying the data constraint. The optimal solution satisfying the data constraint along with L1/L2 penalty is an intersection point of the line L with L1-norm(or L2-norm) sphere in the 2D space whose axes are the coefficients β_1 and β_2 , respectively. From Fig. 6.7a, we see that the minimum L2-norm error solution on the line L is away from the axes with high probability and hence is not a sparse solution. Interestingly, we note from Fig. 6.7b that L1-norm is “pointy” in nature which suggests that the intersection point representing the solution is highly likely to be along the axis and hence the solution will be a sparse solution (as one of the coefficients has 0 value).

The geometrical picture above is not easy to imagine; however, analogy can be drawn to the fact that the TV penalty favors solutions with sparsely located edges along with the majority of nearly flat (small gradient) regions in the recovered image. Finally, we would like to mention for completeness that as per the rigorous compressive imaging (or sensing) theory, if an N pixel image is likely to have a representation with $K (<< N)$ significant coefficients in some sparsity basis, then a number of measurements of the order of $K \log N$ are required. Image recovery is then possible with high probability using an optimization algorithm employing L1-norm penalty. Another important aspect about the compressive imaging methodology is the relation between the measurement scheme and the basis on which the image is likely to be

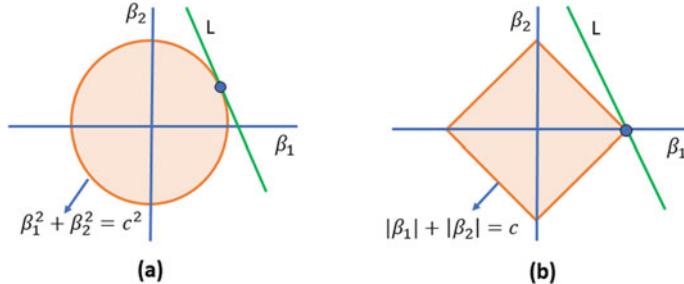


Fig. 6.7 **a** Solution with minimum L2-norm is not sparse while **b** solution with minimum L1-norm is sparse due to pointy nature of L1-norm ball

Fig. 6.8 A sparse image in pixel basis: Few pixels shown by white dots are non-zero while all other pixels are zero



sparse. We will illustrate this with a simple case involving sparsity in pixel basis. Consider a sparse image in pixel basis as shown in Fig. 6.8. Suppose we wish to measure an output g_o corresponding to this sparse object as per the model in Eq. (6.35). We notice that each measurement point g_{oj} corresponds to a linear combination

$$g_{oj} = \sum_n \beta_n (\hat{A}_j, \phi_n) + \text{noise}. \quad (6.37)$$

Here, (\hat{A}_j, ϕ_n) denotes the scalar product between the j th row of the system matrix \hat{A} (which is the j th element of the measurement basis) and the sparsity basis function ϕ_n . First, let us suppose that we select the pixel basis as a measurement basis. The basis function ϕ_n in the case of pixel basis is an image whose n th pixel has a value equal to 1 while all other pixels are zero. If we desire to recover the image $g = \sum_n \beta_n \phi_n$

from a number of measurements that are much smaller than the number of total pixels N in the image, we certainly cannot use the pixel basis as the measurement basis. Since the image as shown in Fig. 6.8 is sparse in the pixel basis, choosing the pixel basis for measurement will mean that most of the scalar products (\hat{A}_j, ϕ_n) will be zero and the recovery problem will be meaningless. On the other hand, if we choose a Fourier basis for measurement \hat{A} , most scalar products are likely to survive making it possible to recover the sparse image with the use of appropriately designed algorithm. *Thus, an important guideline to follow is that the image to be recovered cannot in general be sparse in both the ϕ_n basis and the measurement basis.* This property of the two basis sets is represented more formally in terms of what is known as the restricted isometry property(RIP). We will however not discuss this topic with mathematical details. Apart from rigorous theorems, it is important to remember that the philosophy of explicitly using the expected sparsity of the solution in design of measurements as well as image (or signal) reconstruction tasks is very important and is found useful in breaking several traditional limits on imaging performance. The possibility of solving such “incomplete data” inverse problems has important implications for how imaging systems of tomorrow may be designed. The philosophy behind ideas like compressive imaging is significant to the following classes of problems:

1. **Imaging systems beating conventionally perceived limits:** We now have a new framework for re-examining our notions about imaging system performance metrics such as resolution and sensitivity to noise and for evolving novel imaging concepts that may outperform traditional imagers in these aspects.
2. **Simpler/leaner imaging systems:** High-quality imaging performance with less measurements implies simplifying imaging system design with correspondingly reduced cost for imaging technologies in the area of healthcare, security, education, etc. with high social impact.

Traditionally image processing typically involved enhancing and modifying images and drawing conclusions out of them for applications involving some decision-making task. The new ideas in image reconstruction mean that the algorithms play important role in designing the measurements as well and this aspect offers a number of interesting future possibilities.

6.5 Optimization-Based Image Recovery Without a Free Parameter

The reader may have noticed that the choice of weight parameter α is important in almost all the optimization problems studied in this chapter. The choice of α ultimately affects the perceived image quality. In applications where the user may not be aware of the underlying reconstruction algorithms, the requirement of empirically tuning the free parameter α poses several practical difficulties. The user with a different area of expertise now has to decide which α provides a “suitable” image

and then use this image for the downstream decision-making tasks. Empirical choice of α may work when one is simply trying to prove proof-of-principle of a system concept. In practice, however, the nature of data recorded by a system may change due to noise variations or simply due to different classes of objects being imaged by the same system. A robust methodology that will not require such empirical tuning is therefore desired in practice. In this section, we explore two such approaches. The ideas discussed here may not be possible to prove easily, but they seem to work in a wide variety of inverse problems. For the purpose of illustration, we will once again use the image de-blurring problem in this section.

In order to have an algorithm that does not involve an empirical choice of weight parameter α , we need to change the philosophy of what we wish to achieve via an optimization algorithm. While typically we have framed the reconstruction problem as that of minimizing a cost function, what is more important is to reach a solution point where the two terms of the cost function balance each other and cannot be decreased simultaneously. The first algorithm in this vein is the ASD-POCS (adaptive steepest descent-projection onto convex sets) proposed by Pan et al. in the context of CT scan iterative image reconstruction [8]. The second algorithm is the Mean Gradient Descent idea proposed by the authors of this book that is inspired by the ASD-POCS algorithm. The main idea behind such approaches is to re-frame the reconstruction problem with the requirement that

$$C_1 \leq \epsilon^2, \quad (6.38)$$

such that among the solutions satisfying this condition, we wish to find the solution for which C_2 is minimal. When data are recorded in an experiment, with noise in the measurement process and/or incompleteness of the measurements, it is usually understood that the data consistency term C_1 cannot identically go to zero. The numerical value of the quantity ϵ^2 will then be decided by the specific system considerations. The ill-posedness of the inverse problems also suggests that there will be multiple solutions that satisfy the condition in Eq. (6.38). Starting with an arbitrary initial guess for the solution, we thus need to follow a path that reduces C_1 , but at the same time take the solution away from points where Eq. (6.38) is satisfied but the numerical value of C_2 is high. In particular, if we simply start minimizing C_1 from the starting guess, it may be observed that at some point the algorithm starts fitting the noise in the data thereby increasing the penalty term C_2 which is typically a function of gradient magnitudes of the solution image. The solution strategy suggested by ASD-POCS algorithm involves two distinct steps in each iteration of the algorithm. The $(n + 1)$ th iteration starts with the previous guess $g^{(n)}$ and may be described as follows:

- 1. Reduction of C_1** : The first term C_1 of the cost function (typically representing the data consistency) is reduced by following a strategy like steepest descent to obtain an intermediate solution \tilde{g} .

$$\tilde{g} = g^{(n)} - \tau [\nabla_g C_1]_{g^{(n)}}. \quad (6.39)$$

Here, the functional gradient of C_1 is evaluated with respect to the previous guess $g^{(n)}$ and the step size τ may be selected by following backtracking line search. The norm distance between the previous guess $g^{(n)}$ and the intermediate solution \tilde{g} is then computed for use in the second step of the iteration.

$$d_1 = \|\tilde{g} - g^{(n)}\|_2. \quad (6.40)$$

- 2. Adaptive reduction of C_2 :** The intermediate solution \tilde{g} is clearly better than $g^{(n)}$ with respect to C_1 , but it may need modification as it may not be optimal with respect to C_2 . The penalty term C_2 associated with the intermediate solution \tilde{g} is now reduced in N_{sub} number of sub-iterations to obtain the next guess $g^{(n+1)}$ such that the norm distance

$$d_2 = \|\tilde{g} - g^{(n+1)}\|_2 \quad (6.41)$$

is approximately equal to d_1 . In particular, C_2 reduction is performed by following sub-iterations of the form:

$$\tilde{g} \leftarrow \tilde{g} - \tau' \left[\frac{\nabla_{\tilde{g}} C_2}{\|\nabla_{\tilde{g}} C_2\|_2} \right]_{\tilde{g}}, \quad (6.42)$$

where the step size τ' is made equal to a small fraction βd_1 of the distance d_1 . The procedure of adaptive reduction of C_2 may specifically be carried out as follows. Every outer iteration involves a fixed number N_{sub} (e.g. 20) of sub-iterations in the steepest descent direction corresponding to C_2 starting with some initial value of fraction β (e.g. 0.2) at the start of the algorithm. After the N_{sub} sub-iterations, if it is determined that the distance d_2 by which the intermediate solution has changed is comparable or larger compared to d_1 (e.g. $d_2 > 0.95d_1$), then for the next outer iteration the value of β is reduced by a fixed fraction (e.g. 0.95).

Typically as one starts with an arbitrary initial guess, the C_1 reduction causes large change d_1 in the guess solution. The sub-iteration process involving N_{sub} steps is also allowed to change the solution by an equally large (or comparable) amount d_2 . As the outer iterations progress, both d_1 and d_2 start reducing in magnitude. An interesting aspect of this method is that after a number of iterations, it is typically observed that the negative functional gradients $-\nabla_g C_1$ and $-\nabla_g C_2$ computed at the successive guess solutions make a large obtuse angle with each other indicating that the steepest descent directions corresponding to C_1 and C_2 are opposing each other leading to an equilibrium point. At this point, the solution does not change much as reduction directions for C_1 and C_2 now nearly oppose each other and this brings the solution back to nearly the same place. The iterative process is stopped if C_1 has reduced to a reasonably low value (depending on system specifics) and at the same time the angle between the two steepest descent directions with respect to C_1 and C_2 is sufficiently large (practically greater than 120-degrees or more). We note that since the aim of the algorithm is to achieve balance between C_1 and C_2 rather than minimizing a

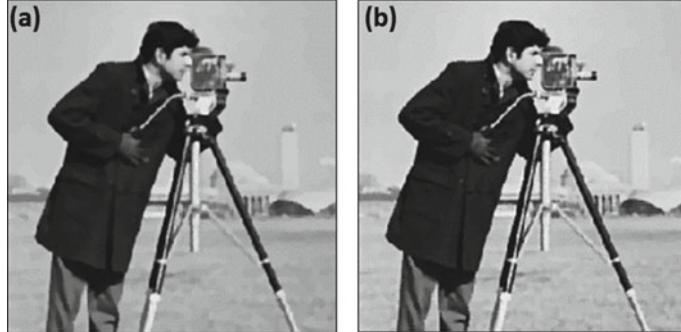


Fig. 6.9 Image recovery using **a** adaptive tuning of distances d_1, d_2 as in ASD-POCS, **b** mean gradient descent

cost function of the form $(C_1 + \alpha C_2)$, the overall procedure does not require any empirical parameter α and the ultimate balance point location in solution space adjusts itself as per the nature of the recorded data. While the original ASD-POCS article describes this process in terms of the variational inequality formalism, the nature of the solution set $C_1 \leq \epsilon^2$ may not be convex for certain problems involving complex-valued solutions. The overall strategy however is typically observed to work even in such cases. In the illustration in Fig. 6.9a, we show the image de-blurring result obtained by following the above strategy where the same C_1 and C_2 as in the de-blurring problem Eq. (6.23) are used. It is observed that the image recovery in Fig. 6.9a is comparable to that in Fig. 6.3d obtained with a suitable α parameter.

Inspired by the success of ASD-POCS type algorithm for image reconstruction, the present authors have recently proposed another idea for achieving balance between C_1 and C_2 terms of the cost function, which is referred to as the Mean Gradient Descent (MGD) algorithm [9]. The key idea behind MGD is to recognize that ASD-POCS requires the balance of the two distances d_1 and d_2 corresponding to the L2-norm change in the guess solution due to reduction of C_1 and C_2 , respectively. Taking a cue from this important aspect, MGD updates the guess solution in the direction that bisects the two directions $-\nabla_g C_1$ and $-\nabla_g C_2$. In particular, we define two normalized descent directions:

$$\hat{u}_1 = \frac{\nabla_g C_1}{\|\nabla_g C_1\|_2}, \quad (6.43)$$

and

$$\hat{u}_2 = \frac{\nabla_g C_2}{\|\nabla_g C_2\|_2}. \quad (6.44)$$

A new vector \hat{u} is introduced which is along the mean of \hat{u}_1 and \hat{u}_2 as

$$\hat{u} = \frac{\hat{u}_1 + \hat{u}_2}{2}, \quad (6.45)$$

and the guess solution is progressed by the following iterative scheme:

$$g^{(n+1)} = g^{(n)} - \tau \|g^{(n)}\|_2 [\hat{u}]_{g^{(n)}}. \quad (6.46)$$

Since \hat{u} is along the mean of the two unit vectors \hat{u}_1 and \hat{u}_2 , the update step ensures that the change in the solution $g^{(n)}$ due to reduction of C_1 and C_2 is always kept equal. The selection of the step size τ needs some attention. In typical optimization problems as explained in this chapter earlier, the step size τ is selected by line search methods which check whether the cost function is reduced sufficiently for a given step size. In MGD approach since there is no free parameter α , there is no fixed cost function to reduce. Our goal is not to reduce a cost function but to progress such that the angle between \hat{u}_1 and \hat{u}_2 is increased to a large obtuse value, so that the two cost function terms C_1 and C_2 are difficult to reduce simultaneously. In practice, numerical experiments with the scheme suggest that starting with a uniform random number as initial guess, if the step size is kept fixed at a moderate level $\tau = 0.01 - 0.05$, then after at most a few hundred iterations, the angle between \hat{u}_1 and \hat{u}_2 defined as

$$\theta_{12} = \arccos(\hat{u}_1 \cdot \hat{u}_2), \quad (6.47)$$

is observed to stagnate. At this point, the solution has several features of the desired solution but may have additional artifacts that cannot go away with the fixed step size. From this point onward, the step size may be selected by a backtracking line search strategy which is based on increase in angle θ_{12} . This process is seen to quickly reduce the step size and improve the solution further. The angle θ_{12} quickly rises to a numerical value close to 180 degrees. The iterative process may be stopped at this point as there is not much change in solution on further iterations. Since the descent directions with respect to C_1 and C_2 are opposing at this point, an equilibrium solution point is reached. When the same blurred and noisy image data is used, the resultant image recovery with MGD approach is shown in Fig. 6.9b which is similar in nature to the recovery in Fig. 6.9a.

An interesting aspect of the ASD-POCS as well as the MGD methodologies is that if the noise level in the blurred image is now changed, there is no empirical tuning of free parameters required and the same procedure provides a solution that achieves balance between C_1 and C_2 for the higher noise level. In traditional optimization approaches involving minimization of cost function, the free parameter α needs to be tuned again when the noise level in the data changes.

6.6 Topics for Further Study

In this chapter, we have provided the reader with operational introduction to the important ideas in constrained optimization that are highly valuable for image recovery problems without necessarily providing the most rigorous mathematical treatment to optimization theory for which excellent sources are already available. Starting with the simplest image denoising problem, we added next level of difficulty to the problem when discussing image de-blurring and further considered the compressive sensing ideas that suggest possibility of image recovery with data that would be considered incomplete in traditional sense. Discussion of algorithmic ideas on reconstruction that does not involve any free tuning parameter was also provided. Two major aspects that are left out here are the choice of algorithms and modeling of the cost function. In particular, while we have stuck to the simplest gradient descent scheme, many other choices such as conjugate gradient, Nesterov optimal gradient and Newton iteration have been studied in the literature. Each of these methods has its own advantages or disadvantages. The choice of final algorithm depends on several considerations such as speed, computational load, simplicity of implementation and aspects specific to the system or application under consideration. In terms of cost function modeling, there are several statistical approaches such as maximum likelihood, maximum entropy and weighted least squares that account for the knowledge of statistical nature of signal detection and noise in measurements [10]. Also while TV penalty was used in all the illustrations, a general class of Gibbs prior penalty functions also exists and the choice of penalty function again depends on the application under consideration. We believe that the reader has acquired a sense of the powerful constrained optimization techniques and their importance to computational imaging systems through this chapter. The state-of-the-art imaging systems are likely to be increasingly dependent on such optimization algorithms for image recovery from measured data.

Problems

6.1 Consider a cost function of the form:

$$C(g) = \|g_{out} - \hat{A}g\|_2^2 + \alpha \sum_m [(\nabla_x g)_m^2 + (\nabla_y g)_m^2],$$

where the first term represents the data consistency measure and the penalty term is known as the total square variation (TSV). Find an expression for the functional gradient $\nabla_g C$. Based on the functional gradient, explain what you think is the nature of the image solution favored by the TSV penalty in comparison to a similar optimization problem which uses a TV penalty.

6.2 Consider the modified Huber penalty function:

$$H(g) = \sum_m [\sqrt{1 + \frac{|(\nabla g)_m|^2}{\delta^2}} - 1].$$

Show that for pixels where $|\nabla g|^2 << \delta^2$ and $|\nabla g|^2 >> \delta^2$, the Huber penalty acts like a TSV penalty and TV penalty, respectively.

6.3 Consider an error function corresponding to the model for an interference record:

$$C(O, O^*) = \|H - |R + O|^2\|_2^2.$$

Here, $H(x, y)$ represents a recorded interference pattern (or hologram), $R(x, y)$ represents a complex-valued reference wave and $O(x, y)$ represents the unknown object wave. Evaluate the Wirtinger derivative: $\nabla_{O^*} C$.

6.4 Maximum likelihood: In this chapter, we mostly modeled the data fidelity term in the optimization problems as a least-square fit. This approach gives equal importance to all the measured data points. In case of highly noisy data, some data points may be more noisy than others (and hence less reliable). We will take a concrete example where a low-light level blurred image $d(x, y)$ is modeled as

$$d(x, y) = g(x, y) * h(x, y) + n(x, y).$$

Here, all the functions in the above equation are 2D and are denoted in terms of photon counts registered on a pixelated detector. The desired solution, the point spread function and the noise are denoted by $g(x, y)$, $h(x, y)$ and $n(x, y)$, respectively. We will assume that the counts in different pixels are independent and obey Poisson statistics (see Chap. 7). Given a solution estimate $\hat{g}(x, y)$, the probability of recording the data $d(x, y)$ is given by

$$P(d(x, y)|\hat{g}(x, y)) = \prod_{all(x,y)} \frac{[\hat{d}(x, y)]^{d(x,y)} \exp[-\hat{d}(x, y)]}{d(x, y)!},$$

where $\hat{d}(x, y) = \hat{g}(x, y) * h(x, y)$. The most likely solution $\hat{g}(x, y)$ is the one for which this conditional probability is maximized. Since the probability involves products, it is easier to maximize the logarithm instead:

$$\log[P(d(x, y)|\hat{g}(x, y))] = \sum_{(x,y)} \{d(x, y) \log[\hat{d}(x, y)] - \hat{d}(x, y) - \log[d(x, y)!]\}.$$

Write and implement an iterative algorithm that will maximize the log probability (or likelihood) [11, 12].

6.5 Geman prior: Consider a penalty function:

$$L(\nabla g) = \frac{|\nabla g|^2}{\delta^2 + |\nabla g|^2}.$$

Compare the behavior of this function relative to the total variation penalty in the context of its edge preservation property.

References

1. M. Bertero, P. Boccacci, *Introduction to Inverse Problems in Imaging* (IOP, UK, 1998)
2. S. Boyd, L. Vandenberghe, *Convex Optimization* (Cambridge University Press, UK, 2004)
3. J. Nocedal, S. Wright, *Numerical Optimization* (Springer Science and Business Media, 2006)
4. H.H. Barrett, K.J. Myers, *Foundations of Imaging Science* (Wiley, USA, 2003)
5. Mallat, *A Wavelet Tour of Signal Processing* (Academic, USA, 2000)
6. D.H. Brandwood, A complex gradient operator and its applications in adaptive array theory. Proc. IRE **30**, 11–16 (1983)
7. E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal recovery from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**, 489–509 (2006). A comprehensive list of resources/articles in Compressive Sensing is available at <http://dsp.rice.edu/cs>
8. E.Y. Sidky, X. Pan, Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Phys. Med. Bio. **53**(17), 4777 (2008)
9. S. Rajgora, M. Butola, K. Khare, Mean gradient descent: an optimization approach for single-shot interferogram analysis. J. Opt. Soc. Am. A **36**(12), D7–D13 (2019)
10. S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. **6**, 721–741 (1984)
11. W.H. Richardson, Bayesian-based iterative method of image restoration. J. Opt. Soc. Am. A **62**, 55–59 (1972)
12. L.B. Lucy, An iterative technique for the rectification of observed distributions. Astron. J. **79**, 745 (1974)

Chapter 7

Random Processes



In this chapter, we will introduce several important concepts from the theory of random processes that will be useful later when discussing coherence properties of light and other phenomena such as speckles that are important for imaging systems. When physical phenomena occur as a result of a large number of unknown (almost random) factors, the observations and predictions of experimental results can only be done satisfactorily in statistical terms [1–5]. The generation of light, light-matter interaction, propagation of light through realistic fluctuating media and detection of light cover most of the topics that are relevant to Optics and Photonics. Imaging involves all the above Physical processes in addition to reconstruction algorithms that incorporate the physical model for image formation. Therefore, the knowledge of concepts in random process theory is essential for imaging researchers.

7.1 Probability and Random Variables

The concept of probability is associated with observable phenomena or events when we do not have sufficient information to predict the precise behavior of the underlying system. If a particular event has n successes out of N trials, then as the number of trials becomes very large, we may associate an average rate of success or probability:

$$p = \lim_{N \rightarrow \infty} \frac{n}{N}. \quad (7.1)$$

In set-theoretic terms, we may denote the total event space by Ω . Then $p(\Omega) = 1$ and probability of the null set O is zero. For any subset $A \in \Omega$, $0 \leq p(A) \leq 1$. For two mutually exclusive events A_1 and A_2 in Ω ,

$$p(A_1 + A_2) = p(A_1) + p(A_2). \quad (7.2)$$

When an event is obtained by compounding two events, we have to consider joint probability of events A and B which is denoted by $p(A, B)$ or which is the same as $p(A \cap B)$. It is clear that the joint probability is less than or equal to the probability of one event alone. If M events B_1, B_2, \dots, B_M are mutually exclusive and span the whole of Ω , then

$$\sum_{j=1}^M p(A, B_j) = p(A, \Omega) = p(A). \quad (7.3)$$

For events A and B that are not necessarily mutually exclusive, the probability of their union is given by

$$p(A \bigcup B) = p(A) + p(B) - p(A \cap B) \leq p(A) + p(B). \quad (7.4)$$

Conditional probability $p(A|B)$ or A given that B has occurred is defined as

$$p(A|B) = \frac{p(A \cap B)}{p(B)}. \quad (7.5)$$

If the two events are such that $p(A|B) = p(A)$, then as far as A is concerned, it does not matter if B has occurred or not and the two events are considered to be statistically independent. From the definition of conditional probability, we observe that if events A and B are independent,

$$p(A \cap B) = p(A)p(B). \quad (7.6)$$

The above relation may be considered as a necessary and sufficient condition for statistically independent events. If two events are mutually exclusive, they cannot be statistically independent as their joint probability is equal to zero. We will next state the Bayes rule on inverse probabilities. This is a relation between the conditional probabilities $p(A|B)$ and $p(B|A)$.

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)} = \frac{p(A|B)p(B)}{\sum_{all\ B} p(A|B)p(B)}. \quad (7.7)$$

In many practical situations where the inverse probability is to be predicted, the individual probabilities $p(B)$ for all possible B 's may not be known and are often assigned an equal value for lack of additional knowledge.

7.1.1 Random Variables

The outcome or measurement from a complicated experiment is typically some numerical variable x whose value is not deterministic. We will refer to x as a random variable, which may take discrete values or any of the (infinite number of) values in some interval. Accordingly the random variable is termed as discrete or continuous random variable. The set of possible outcomes of x is called the ensemble of x . For a discrete random variable, we may associate a probability p_i with each of the possible outcomes x_i such that

$$\sum_i p_i = 1. \quad (7.8)$$

When the random variable is continuous, we talk of probability density $p(x)$ such that $p(x)dx$ equals the probability of finding an outcome in the interval $(x - dx/2, x + dx/2)$. The probability density integrated over the x -coordinate must satisfy the following relation:

$$\int dx p(x) = 1. \quad (7.9)$$

The cumulative probability density associated with $p(x)$ is defined as

$$P(x \leq X) = \int_{-\infty}^X dx p(x). \quad (7.10)$$

Based on this definition, the derivative of $P(x \leq X)$ has the following property:

$$\frac{dP(x \leq X)}{dX} = p(X). \quad (7.11)$$

The statistical behavior of the random variable can be characterized by means of expectation values and moments as we will describe next. For any function $f(x)$ of a random variable with probability density given by $p(x)$, its expectation value is defined as

$$\langle f(x) \rangle = \int dx p(x) f(x). \quad (7.12)$$

We observe that the expectation value is the weighted average of $f(x)$ depending on the probability density $p(x)$. Among the functions of x , its moments or the expectation values of the powers of x have a special place. The r th moment of x is defined as

$$\nu_r = \langle x^r \rangle = \int dx p(x) x^r. \quad (7.13)$$

The mean of the random variable may be denoted as ν_1 . For discrete random variable that takes values 0, 1, 2, ... one may define the r th moment as

$$\nu_r = \sum_{x=0}^{\infty} p(x) x^r. \quad (7.14)$$

Another useful concept is the factorial moment of a random variable defined as

$$\langle x^{(r)} \rangle = \langle x(x-1)(x-2)\dots(x-r+1) \rangle. \quad (7.15)$$

The factorial moments are sometimes easier to evaluate analytically for some commonly used discrete distributions (e.g. Poisson and binomial distributions). The moments defined so far have been defined with respect to $\langle x \rangle = 0$. If the mean $\langle x \rangle$ is a finite value, then one may talk of central moments by converting to a new variable: $\mu = (x - \langle x \rangle)$. The first moment μ_1 in this case will be zero. The second central moment μ_2 is known as the variance of the variable and it represents the spread of the distribution around the mean. It must be noted that the higher order moments like variance may not always exist (they may diverge) if the probability distribution is slowly decaying with increasing value of x . An example of this is the sinc² probability distribution which describes the probability of arrival of a photon as a result of the single slit diffraction. It is clear that the second moment of this distribution diverges. There are alternative ways to define what we mean by “spread” of such distributions as we have already discussed in Chap. 3 in relation to the energy concentration problem and the Slepian functions. The square root of the variance is often denoted by σ and is called the standard deviation of the distribution. The standard deviation may be used to normalize higher order moments. For example, the skewness of a distribution is defined by $\alpha_3 = \mu_3/\sigma^3$ and the kurtosis $\alpha_4 = \mu_4/\sigma^4$. Random variables are often expressed in what is called the standard form: $y = (x - \langle x \rangle)/\sigma$. It is easy to show that y has zero mean and unit standard deviation.

Moments of multiple random variables may be evaluated using joint probability distributions. For N real random variables x_j with $j = 1, 2, \dots, N$, we define the covariance matrix μ_{ij} as

$$\mu_{ij} = \langle \Delta x_i \Delta x_j \rangle, \quad (7.16)$$

where $\Delta x_i = (x_i - \langle x_i \rangle)$. The covariance matrix is symmetric with the diagonal elements given by the variances of the individual random variables. We will assume that all random variables x_i have zero mean for simplicity of notation in the following discussion. By using the Cauchy-Schwartz inequality, we can show the following important result:

$$\begin{aligned} & \left[\iint dx_i dx_j p(x_i, x_j) x_i x_j \right]^2 \\ & \leq \left[\iint dx_i dx_j p(x_i, x_j) x_i^2 \right] \left[\iint dx_i dx_j p(x_i, x_j) x_j^2 \right]. \end{aligned} \quad (7.17)$$

In the notation for covariance matrix μ_{ij} , the above inequality may be stated as

$$\mu_{ij}^2 \leq \sigma_i^2 \sigma_j^2. \quad (7.18)$$

The correlation coefficient defined as $\rho_{ij} = \mu_{ij}/\sigma_i \sigma_j$ can therefore take values in the range $[-1, 1]$. If two random variables are statistically independent, the correlation coefficient is equal to zero. This can be shown easily using $p(x_1, x_2) = p(x_1)p(x_2)$ in the definition of μ_{12} . If the random variables are complex-valued, the covariance matrix is made Hermitian symmetric by the following definition:

$$\mu_{ij} = \langle \Delta x_i^* \Delta x_j \rangle. \quad (7.19)$$

The absolute value $|\rho_{ij}|$ of the correlation coefficient in this case is bounded by 1.

7.1.2 Characteristic Function

The characteristic function corresponding to a probability density $p(x)$ is the expected value of $\exp(i\omega x)$ or in other words it is the Fourier transform of the probability density function.

$$C(\omega) = \langle \exp(i\omega x) \rangle = \int_{-\infty}^{\infty} dx \ p(x) \ \exp(i\omega x). \quad (7.20)$$

The probability density $p(x)$ is absolutely integrable and hence its Fourier transform always exists. The probability density may be obtained from $C(\omega)$ by inverse Fourier transformation. The characteristic function is a generating function in the sense that its derivatives are related to the moments of the distribution as follows:

$$(-i)^r \left[\frac{\partial^r C(\omega)}{\partial \omega^r} \right]_{\omega=0} = \int dx \ p(x) \ x^r. \quad (7.21)$$

Some of the properties of the characteristic function are listed as follows:

1. Since the probability density $p(x)$ integrates to 1, we have $C(0) = 1$. Additionally, we have the inequality:

$$|C(\omega)| \leq \int_{-\infty}^{\infty} dx |p(x) \exp(i\omega x)| = C(0). \quad (7.22)$$

2. Since $p(x)$ is real, $C(-\omega) = C^*(\omega)$.
3. $C(\omega)$ has a curious property that it is non-negative definite. To prove this property, one may start with

$$\langle \left| \sum_{j=1}^N a_j \exp(ix\omega_j) \right|^2 \rangle \geq 0. \quad (7.23)$$

Writing the average $\langle \dots \rangle$ in terms of $p(x)$ leads to

$$\sum_{j=1}^N \sum_{k=1}^N a_j^* a_k C(\omega_k - \omega_j) \geq 0. \quad (7.24)$$

An important result (Bochner's theorem) states that the Fourier transform of a non-negative definite function is non-negative and vice versa. As we shall see in the next section, the auto-correlation function of a stationary random process has properties similar to the characteristic function and the non-negative quantity that corresponds to its Fourier transform is the spectral density (or spectrum) of the random process.

The characteristic function is useful in determining the transformations of random variables. For example, when two statistically independent random variables x_1 and x_2 are added, the probability distribution of the resultant random variable $y = x_1 + x_2$ may be obtained using the characteristic function of y as follows:

$$\langle \exp(i\omega y) \rangle_y = \langle \exp[i\omega(x_1 + x_2)] \rangle_{x_1, x_2} = \langle \exp(i\omega x_1) \rangle_{x_1} \langle \exp(i\omega x_2) \rangle_{x_2}. \quad (7.25)$$

In other words, the characteristic function $C_y(\omega)$ is a product of the characteristic functions $c_{x_1}(\omega)$ and $c_{x_2}(\omega)$. Since characteristic function is a Fourier transform of the probability distribution, the product relation implies that the probability density of y is a convolution of probability densities of x_1 and x_2 . It is sometimes required to find the probability density associated with a transformed variable y which is a function of the random variable x , i.e. $y = f(x)$. We can once again obtain the probability density of y using its characteristic function:

$$\begin{aligned} C_y(\omega) &= \langle \exp(iy\omega) \rangle = \langle \exp[if(x)\omega] \rangle \\ &= \int dx p(x) \exp[if(x)\omega]. \end{aligned} \quad (7.26)$$

The inverse Fourier transform of $C_y(\omega)$ gives the probability density $p(y)$:

$$\begin{aligned} p(y) &= \frac{1}{2\pi} \int d\omega \int dx p(x) \exp[-i(y - f(x))\omega] \\ &= \int dx p(x) \delta[y - f(x)]. \end{aligned} \quad (7.27)$$

The probability density $p(y)$ is thus obtained by integrating $p(x)$ over all the values of x satisfying the constraint $y = f(x)$. The above relation may be made more explicit in terms of the roots $x = x_i$ where $y - f(x_i) = 0$.

$$p(y) = \int dx \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|} p(x) = \sum_i \frac{p(x_i)}{|f'(x_i)|}.$$

As a simple example, consider the transformation $y = (x - \langle x \rangle)/\sigma$. The probability density of y denoted by $P(y)$ is given by

$$P(y) = \int dx p(x) \delta[y - (x - \langle x \rangle)/\sigma] = \sigma p(y\sigma + \langle x \rangle). \quad (7.28)$$

7.1.3 Gaussian or Normal Distribution

The Gaussian random distribution is defined with a probability distribution function:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-(x - \langle x \rangle)^2/2\sigma^2]. \quad (7.29)$$

The characteristic function for the Gaussian distribution is given by

$$\begin{aligned} C(\omega) &= \frac{1}{\sqrt{2\pi}\sigma} \int dx \exp[-(x - \langle x \rangle)^2/2\sigma^2 + i\omega x] \\ &= \exp(i\omega\langle x \rangle + \sigma^2\omega^2/2). \end{aligned} \quad (7.30)$$

It is clear that all the moments of the Gaussian distribution are expressible in terms of the mean $\langle x \rangle$ and the standard deviation σ since these are the only parameters present in its characteristic function. The moments for Gaussian distribution can be calculated from derivatives of the characteristic function. For a zero mean Gaussian random distribution, the moments of arbitrary order r are given by

$$\begin{aligned} \mu_r &= 0, && \text{for odd order } r, \\ &= \frac{(\sigma^2/2)^{r/2} r!}{(r/2)!}, && \text{for even order } r. \end{aligned} \quad (7.31)$$

The Gaussian random distribution is very important in many applications. An important result known as the central limit theorem states that when a large number of independent random variables belonging to arbitrary probability density functions are added together, the resultant distribution is a Gaussian distribution. In case of thermal light sources or the speckle phenomena, when resultant light field is a superposition of contributions from a number of independent atoms or scattering centers, respectively, the field distribution is usually Gaussian in nature. We have already seen that when random variables are added, their probability density function is a convolution of the respective probability distributions. The central limit theorem states that when a large number of independent random variables are added, the probability distribution of the sum tends to be a Gaussian distribution. As an illustration of this, we can consider multiple random variables each having a rectangle function distribution. The variables are thus equally likely to take values in the range $x : (-1/2, 1/2)$.

Convolution of rect with itself produces a triangle function. Convolution of the triangle with another rect for a couple of more times gives the higher order box splines which start looking like a Gaussian distribution. The central limit theorem formalizes this result. Suppose that x_1, x_2, \dots are statistically independent variables in a normal form ($\langle x_i \rangle = 0$ and $\sigma_i = 1$ for all x_i). We construct a variable y as

$$y = \frac{1}{\sqrt{N}}(x_1 + x_2 + \dots + x_N). \quad (7.32)$$

The mean $\langle y \rangle = 0$ and the variance $\langle y^2 \rangle = 1$. For a random variable x_i in the standard form, we can write the characteristic function $C(\omega)$ as a Taylor series around point $\omega = 0$ using the relation in Eq. (7.21) as follows:

$$C(\omega) = 1 - \frac{\omega^2}{2} + O(\omega^3). \quad (7.33)$$

We note that this form is the same for all random variables x_i in normal form irrespective of the probability distribution. The characteristic function $C_y(\omega)$ is therefore given by

$$C_y(\omega) = \left[1 - \frac{\omega^2}{2N} + O(\omega^3) \right]^N. \quad (7.34)$$

For large N , we may neglect the terms of order ω^3 and higher to obtain

$$C_y(\omega) = \lim_{N \rightarrow \infty} \left[1 - \frac{\omega^2}{2N} \right]^N = \exp(-\omega^2/2), \quad (7.35)$$

which is a Gaussian. The probability distribution $p(y)$ which is obtained by inverse Fourier transforming $C_y(\omega)$ is therefore once again a Gaussian distribution. The composition of N random variables in standard form for large N thus tends to a Gaussian distribution irrespective of the nature of individual probability distributions.

Zero mean random variables x_1, x_2, \dots, x_n are called jointly Gaussian when their probability distribution is given by

$$p(\underline{x}) = \frac{1}{(2\pi)^{n/2} \det(\underline{\mu})} \exp \left[-\frac{1}{2} \underline{x}^T (\underline{\mu})^{-1} \underline{x} \right]. \quad (7.36)$$

Here $\underline{\mu}$ is the covariance matrix of the variables. Note that the joint distribution is a generalization of a single Gaussian random distribution. For independent variables, the covariance matrix is diagonal and the joint distribution simplifies to a product of individual Gaussian distributions. The joint characteristic function is given by

$$C(\underline{\omega}) = \exp \left[-\frac{1}{2} \underline{\omega}^T \underline{\mu} \underline{\omega} \right]. \quad (7.37)$$

Here, $\underline{x} = [x_1, x_2, \dots, x_n]^T$ and $\underline{\omega} = [\omega_1, \omega_2, \dots, \omega_n]^T$. In general, $\underline{\mu}$ is a symmetric matrix and therefore can be diagonalized. The transformation $(\underline{\mu})^{-1} = U \underline{\Sigma} U^T$ leads to uncorrelated variables $y = Ux$. A joint expectation of Gaussian random variables is expressible in terms of product of second-order correlations.

$$\langle x_1 x_2 \dots x_n \rangle = \sum_{\text{all combinations}} \langle x_1 x_2 \rangle \langle x_3 x_4 \dots \rangle \quad \text{for even } n \quad (7.38)$$

$$= 0, \quad \text{for odd } n. \quad (7.39)$$

This result is known as the Gaussian moment theorem and can be obtained by differentiating the characteristic function associated with the jointly Gaussian random variables. For example, a joint expectation of the form $\langle x_1^2 x_2^2 \rangle$ may be evaluated as

$$\langle x_1^2 x_2^2 \rangle = \langle x_1^2 \rangle \langle x_2^2 \rangle + 2\langle x_1 x_2 \rangle^2. \quad (7.40)$$

A random process is called complex Gaussian process if its real and imaginary parts are independent and derived from a real-valued Gaussian process. A set of n complex random variables z_1, z_2, \dots, z_n now correspond to $2n$ variables and the characteristic function is then introduced in the same way but now over $2n$ variables. The random variables z_m and z_n may however be correlated. The complex form of the Gaussian moment theorem may be stated as

$$\langle z_1^* z_2^* \dots z_k^* z_{k+1} z_{k+2} \dots z_{2k} \rangle = \sum_{\text{all } k! \text{ pairs}} \langle z_1^* z_p \rangle \langle z_2^* z_q \rangle \dots \langle z_k^* z_r \rangle. \quad (7.41)$$

A special case of this result is useful in applications involving intensity-intensity correlation and may be stated as follows:

$$\langle z_1^* z_2^* z_2 z_1 \rangle = \langle z_1^* z_1 \rangle \langle z_2^* z_2 \rangle + \langle z_1^* z_2 \rangle \langle z_2^* z_1 \rangle. \quad (7.42)$$

The proof is more involved but the procedure essentially is same and involves differentiation of the corresponding characteristic function appropriate number of times [6].

7.2 Random Processes

Random process is a concept that builds on that of the random variables. We have been talking of countable variables x_1, x_2, \dots , etc. Now suppose that we are thinking of a scalar component of electric field generated by a typical light source. The field is a non-deterministic function of the continuous parameter time and may be denoted as $x(t)$. The samples of $x(t)$ (or the numerical values $x(t)$ takes at time t) may form

a particular probability density function $p(x, t)$. Note that if we integrate over x at any given time t ,

$$\int dx p(x, t) = 1. \quad (7.43)$$

The expectation of x at time t may be defined as

$$\langle x(t) \rangle = \int dx p(x, t)x. \quad (7.44)$$

We may alternatively consider a number of realizations of the process under consideration: $x^{(1)}(t), x^{(2)}(t), \dots, x^{(N)}(t), \dots$. The values of the various realizations at time t make the probability distribution $p(x, t)$. The mean or expected value may then be thought of as an average over the possible realizations.

$$\langle x(t) \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{r=1}^N x^{(r)}(t). \quad (7.45)$$

The different realizations make what is known as an ensemble and the average above is an ensemble average. The probability density $p(x, t)$ contains information about the possible values taken by x at any time t but has no information about the possible correlations between values of x at different times. That information is contained in joint probability $p(x_1, t_1, x_2, t_2)$ which is the probability density that x takes a value x_1 at t_1 and x_2 at t_2 . Such joint probability density has more information than $p(x, t)$. In fact by integrating the joint density over x_2 , we can get $p(x_1, t_1)$. The joint density may be used in particular to define the auto-correlation function or the two-time correlation function which we will define in terms of the complex-valued random process:

$$\Gamma(t_1, t_2) = \langle z^*(t_1)z(t_2) \rangle. \quad (7.46)$$

The average $\langle \dots \rangle$ above is with respect to the joint density $p(z_1, t_1, z_2, t_2)$. In general the joint densities can be defined up to any order with n -order density having more information than $(n - 1)$ -order density. The exception to this rule is when we have Gaussian density function in which case only second-order correlations are sufficient as per the Gaussian moment theorem in Eq. (7.38).

A random process is called stationary if there is no preferred origin for t implying that the statistical nature of the process remains same at all times. Any joint density thus remains same if T is added to all times. For example,

$$p(z_1, t_1, z_2, t_2) = p(z_1, t_1 + T, z_2, t_2 + T), \quad (7.47)$$

and similarly for higher order joint densities. Clearly if we choose $T = -t_1$, the correlation function is seen to be only a function of the time difference $(t_2 - t_1)$.

$$\Gamma(t_1, t_2) \equiv \Gamma(t_2 - t_1). \quad (7.48)$$

We observe that Γ is Hermitian symmetric with respect to $\tau = (t_2 - t_1)$:

$$\Gamma(-\tau) = \Gamma^*(\tau). \quad (7.49)$$

A random process is called stationary in strict sense if joint probability densities of all orders have the stationarity property. A random process is called wide sense stationary if the mean is independent of t and the correlation $\Gamma(t_1, t_2)$ is a function of $(t_2 - t_1)$.

A random process is called ergodic if its ensemble average coincides with a long-time average of one realization.

$$\langle z \rangle = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} dt z^{(r)}(t). \quad (7.50)$$

For this to be possible, it is generally required that the correlation $\Gamma(\tau)$ dies out to zero fast for increasing τ . A single realization of the process may then be divided into multiple time intervals that are uncorrelated and may be considered as separate realizations of the process. The understanding of when time average can be treated as equivalent to the ensemble average is often important in laboratory experiments, where only one realization of the random process (e.g. corresponding to some light source) may be available. Performing time averages is then inevitable for comparing experimental results with theoretical predictions using ensemble averages. Some properties of the auto-correlation function are stated below and are straightforward to prove:

1. $|\Gamma(0)| \geq 0$.
2. $\Gamma(-\tau) = \Gamma^*(\tau)$.
3. $\Gamma(\tau)$ is non-negative definite. For n complex numbers a_1, a_2, \dots, a_n and n real numbers t_1, \dots, t_n , we have

$$\langle \left| \sum_{j=1}^n a_j z(t_j) \right|^2 \rangle \geq 0, \quad (7.51)$$

which for a stationary random process may be further simplified to

$$\sum_j \sum_k a_j^* a_k \Gamma(t_j - t_k) \geq 0. \quad (7.52)$$

We make an interesting observation that the properties of the auto-correlation function are essentially same as that of the characteristic function.

7.2.1 Spectral Density: Wiener-Khintchine Theorem

We have encountered the non-negative definiteness property earlier in case of the characteristic function associated with a probability density function. The Fourier transform of such functions is known to be non-negative (Bochner theorem). Since the correlation function $\Gamma(\tau)$ satisfies the non-negative definiteness property, its Fourier transform is non-negative and we will see next that this non-negative quantity is the spectrum of the process—or the spectral density representing the energy density in the process at a given frequency. *The Wiener-Khintchine theorem relates the auto-correlation function of a random process with its spectrum.* For a random process $z(t)$, we will denote its truncated Fourier transform by

$$Z(v, T) = \int_{-T/2}^{T/2} dt z(t) \exp(-i2\pi vt). \quad (7.53)$$

The periodogram or the power density of a single realization may be defined as

$$S_T(v) = \frac{|Z(v, T)|^2}{T}. \quad (7.54)$$

The expected value of the power spectral density of the random process is obtained by ensemble averaging the above quantity in the limit $T \rightarrow \infty$.

$$S(v) = \lim_{T \rightarrow \infty} \langle \frac{|Z(v, T)|^2}{T} \rangle. \quad (7.55)$$

This definition may be further expanded explicitly as

$$\begin{aligned} S(v) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt dt' \text{rect}\left(\frac{t}{T}\right) \text{rect}\left(\frac{t'}{T}\right) \Gamma(t - t') \exp[i2\pi v(t - t')] \\ &= \lim_{T \rightarrow \infty} \int_{-\infty}^{\infty} d\tau \Lambda\left(\frac{\tau}{T}\right) \Gamma(\tau) \exp[i2\pi v\tau] \\ &= \int_{-\infty}^{\infty} d\tau \Gamma(\tau) \exp[i2\pi v\tau]. \end{aligned} \quad (7.56)$$

In the above calculation, $\Lambda(u)$ is the triangle function defined earlier in Eq. (2.76). The triangle function appears here as a correlation of two rect functions. We observe the important result that the spectral density $S(v)$ and the auto-correlation function $\Gamma(\tau)$ are related by a Fourier transform. This important result is referred to as the Wiener-Khintchine theorem.

7.2.2 Orthogonal Series Representation of Random Processes

For a zero mean wide sense stationary random process $z(t)$, we may consider an expansion over $t : (-T/2, T/2)$ in terms of an orthonormal set $\{\psi_n(t)\}$.

$$z(t) = \sum_n c_n \psi_n(t), \quad (7.57)$$

with the expansion coefficients following a Kronecker-delta correlation:

$$\langle c_m^* c_n \rangle = \lambda_m \delta_{m,n}. \quad (7.58)$$

When such an expansion exists, we note that

$$\Gamma(t_1 - t_2) = \sum_n \lambda_n \psi_n^*(t_1) \psi_n(t_2). \quad (7.59)$$

Further due to the orthonormality of the basis functions, we have

$$\int_{-T/2}^{T/2} dt_1 \Gamma(t_1 - t_2) \psi_n(t_1) = \lambda_n \psi_n(t_2). \quad (7.60)$$

We therefore see that the orthonormal basis functions are the eigenfunctions of the integral kernel given by the correlation function. This orthonormal expansion is referred to as the Kosambi-Karhunen-Loeve expansion. We will next show that the eigenfunctions of the auto-correlation functions make the most efficient basis for representing the random process. To show this, let us assume that $z(t)$ is represented as a linear combination of any arbitrary orthogonal basis $\{\phi_n(t)\}$ in $(-T/2, T/2)$.

$$z(t) = \sum_n c_n \phi_n(t). \quad (7.61)$$

Then using the orthogonality of the basis functions, it is easy to show that

$$\int_{-T/2}^{T/2} dt |z(t)|^2 = \sum_n |c_n|^2. \quad (7.62)$$

For an approximation of $z(t)$ with finite number M of basis functions, the ratio of energy captured by the approximation to total energy $E_0 = \sum_n |c_n|^2$ is given by

$$\beta = \sum_{n=0}^M |c_n|^2 / E_0. \quad (7.63)$$

We want to select a basis set that will maximize the expectation value of β . The expectation value of $|c_n|^2$ may be evaluated as

$$\begin{aligned}\langle |c_n|^2 \rangle &= \iint_{-T/2}^{T/2} dt dt' \langle z^*(t) z(t') \rangle \phi_n^*(t) \phi_n(t') \\ &= \iint_{-T/2}^{T/2} dt dt' \Gamma(t, t') \phi_n^*(t) \phi_n(t').\end{aligned}\quad (7.64)$$

The expectation value is essentially what is known as the Rayleigh-Ritz coefficient corresponding to the Hermitian symmetric integral kernel $\Gamma(t, t')$. The highest value of $\langle |c_n|^2 \rangle$ is known to be the eigenvalue corresponding to the n -th eigenfunction of $\Gamma(t, t')$ (the eigenfunctions being orthogonal to each other). The expectation value of the energy ratio β is thus maximized if the basis set is chosen to be the eigenfunctions of the auto-correlation function. This special set of eigenfunctions (and their discrete version which is known as the principal component basis or the PCA basis) is therefore used in multiple applications like speech coding, face recognition, unsupervised classification of data, etc. and provides the most suitable basis for representing a given class of signals. As an aside, we mention here that for a bandlimited random process with a flat rect-function spectrum, the auto-correlation function is the sinc function and the corresponding orthogonal basis set is the Slepian or the prolate spheroidal functions studied in Sect. 3.4.

7.3 Complex Representation of Random Processes

The quantities such as charges and currents that produce electromagnetic fields are real quantities but it is often required to represent the electromagnetic fields as complex numbers. For a real random process $x(t)$, we denote the Fourier transform by

$$X(\nu) = \int_{-\infty}^{\infty} dt x(t) \exp(i 2\pi \nu t).\quad (7.65)$$

Since $x(t)$ is real, the transform satisfies the relation $X(-\nu) = X^*(\nu)$. The negative frequencies do not contain any additional information and we may suppress them to define a complex signal associated with the real process $x(t)$ as follows:

$$z(t) = \int_0^{\infty} d\nu X(\nu) \exp(-i 2\pi \nu t).\quad (7.66)$$

We will express the above relation as

$$\begin{aligned} z(t) &= \frac{1}{2} \int_{-\infty}^{\infty} d\nu X(\nu) \exp(-i2\pi\nu t) [1 + \text{sgn}(\nu)] \\ &= \frac{1}{2}[x(t) + iy(t)], \end{aligned} \quad (7.67)$$

where $y(t)$ is the Hilbert transform:

$$y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} dt' \frac{x(t')}{t - t'}, \quad (7.68)$$

and P denotes the Cauchy principal value. The signal $x(t)$ has a similar representation in terms of $y(t)$.

$$x(t) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} dt' \frac{y(t')}{t - t'}. \quad (7.69)$$

Such relations (or Hilbert transform pair) are often found in Physics in frequency domain and are known as dispersion relations. For example, the real and imaginary parts of refractive index obey such relationship (Kramers-Kronig relation) due to causality of material response. In the above discussion, this relation occurs in time domain since we have suppressed the negative frequencies in the signal. Finally, we observe that

$$\int dt x^2(t) = \int dt y^2(t) = 2 \int dt |z(t)|^2 \quad (7.70)$$

and

$$\int dt x(t)y(t) = 0. \quad (7.71)$$

The concept of complex signal representation plays an important role in many branches of physics and electrical engineering. The analytic signal representation was first introduced in this form by Dennis Gabor [7]. In optics and particularly in this book, we will often encounter narrowband fields such that the Fourier transform of the signal is effectively concentrated in a small frequency band $\Delta\nu$ centered on some carrier frequency ν_0 . We may typically express such signal as

$$x(t) = A(t) \cos[2\pi\nu_0 t + \Phi(t)]. \quad (7.72)$$

The Hilbert transform of the above equation for slowly varying envelope gives

$$y(t) = A(t) \sin[2\pi\nu_0 t + \Phi(t)]. \quad (7.73)$$

If $z(t)$ is the corresponding complex signal, then denoting the Fourier transform of $z(t)$ by $Z(\nu)$, we have the following expression for the complex envelope:

$$z_0(t) = A(t) \exp[i\Phi(t)] = \int d\nu Z(\nu) \exp[i2\pi(\nu - \nu_0)t]. \quad (7.74)$$

We see that in the definition of the complex envelope, we are using two functions $A(t)$ and $\Phi(t)$ to represent one function and as a result there could be infinite number of choices for how we define the complex envelope. Is there any unique choice that is in some sense optimal? There is a theorem in this regard by Mandel [8].

7.3.1 Mandel's Theorem on Complex Representation

Suppose that corresponding to a narrowband real stationary random process $x(t)$, we generate another process $y(t)$ by convolution of $x(t)$ with a filter function $k(t)$

$$y(t) = \int dt' k(t - t') x(t'), \quad (7.75)$$

and represent the corresponding complex signal as

$$z(t) = z_0(t) \exp(i2\pi\nu_0 t). \quad (7.76)$$

Here, ν_0 may be considered as the carrier frequency of the narrowband process. A measure of fluctuation in the complex envelope may then be defined as follows:

$$\rho = \langle \left| \frac{d z_0(t)}{dt} \right|^2 \rangle. \quad (7.77)$$

It is the ensemble average of the absolute squared derivative of the complex envelope. The problem posed by Mandel is to find an appropriate filter $k(t)$ that minimizes this fluctuation [8]. The solution to this problem is formulated as an optimization problem and the answer is that the filter $k(t)$ corresponds to the Hilbert transform. The sine and cosine representation is thus optimal in the sense that it produces signal with least fluctuation in the envelope or an envelope with least redundancy. Finally, we note that the spectrum of the complex signal is given by

$$S_z(\nu) = \frac{1}{4} |1 + \text{sgn}(\nu)|^2 S_r(\nu), \quad (7.78)$$

which vanishes for negative frequencies.

A 2D analogue to this theorem exists and has been studied in detail by the present author [9]. The solution of the 2D problem is that the analogue of Hilbert transform in 1D is provided by the spiral phase transform [10]. For an image signal $g_1(x, y)$, the quadrature signal as per this result is given by

$$g_2(x, y) = \mathcal{F}^{-1}[\exp(i\phi) G_1(f_x, f_y)],$$

where ϕ is the polar angle in the 2D Fourier frequency plane and $G_1(f_x, f_y)$ denotes the 2D Fourier transform of the image $g_1(x, y)$. We will not discuss this topic here any further.

Problems

7.1 Use factorial moment to compute the standard deviation of the Poisson distribution. Use the characteristic function to obtain the same result. The Poisson distribution is given by

$$p(n) = \frac{\bar{n}^n \exp(-\bar{n})}{n!},$$

and the factorial moment of the order of r is defined as the expectation value $\langle n(n - 1)\dots(n - r + 1) \rangle$.

7.2 Consider a random variable $y = x_1 + x_2$, where x_1 and x_2 are independent random variables with uniform probability distribution in the range $[-0.5, 0.5]$. What is the probability distribution of y ? Verify your result with computer simulation by generating a number of samples of x_1 and x_2 , summing them and plotting a histogram.

7.3 Simulating random numbers with a given distribution: Given a distribution $p(x)$, how can we generate sequence of x 's following this distribution? This may be achieved using the idea of cumulative distribution.

$$F(x) = \int_{-\infty}^x dx' p(x'), \quad \frac{dF(x)}{dx} = p(x).$$

For uniform random distribution $u(x)$, $x \in [-0.5, 0.5]$, the cumulative distribution goes from 0 to 1 linearly from -0.5 to 0.5 . We then have to map a practical range of interest for the cumulative distribution of $p(x)$ to $[-0.5, 0.5]$. For example, suppose we have to find a Poisson random number with mean $\bar{m} = 5$. Then we can do the following. (a) Generate a uniform random number in $[-0.5, 0.5]$ (say = 0.1) and find the cumulative distribution of uniform distribution up to this point (cumulative distribution = 0.6). (b) Start observing the cumulative distribution: $\sum_{m=0}^M \frac{\exp(-\bar{m})\bar{m}^m}{m!}$. The M for which cumulative distribution is closest to 0.6 is the random number we want to generate. Repeat the process for a new uniform random number and so on. Implement this methodology with a computer code.

7.4 An image recorded with a camera in ambient light may be considered as a probability distribution. Simulate a realization of an image of your choice at low-light level such that the whole image is made up of just 10000 detected photons.

7.5 Find the analytic signal representation associated with the signal $y = \cos^2(\omega_0 t)$.

References

1. W.B. Davenport, W.I. Root, *An Introduction to the Theory of Random Signals and Noise* (Wiley-IEEE Press, 1987)
2. L. Mandel, E. Wolf, *Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995)
3. J.W. Goodman, *Statistical Optics* (Wiley-Interscience, 2000)
4. A.M. Yaglom, *Correlation Theory of Stationary and Related Random Functions* (Springer, 1987)
5. B.R. Frieden, *Probability, Statistical Optics, and Data Testing: A Problem Solving Approach* (Springer, 1991)
6. I.S. Reed, On a moment theorem for complex Gaussian processes. *IRE Trans. Inf. Theory* **8**, 194–195 (1962)
7. D. Gabor, Theory of communications. *J. Inst. Electr. Eng.* **93**, 429–457 (1946)
8. L. Mandel, Complex representation of optical fields in coherence theory. *J. Opt. Soc. Am.* **57**, 613–617 (1967)
9. K. Khare, Complex signal representation, Mandel's theorem and spiral phase quadrature transform. *Appl. Opt.* **47**, E8–E12 (2008). (Special issue on phase space representations in Optics)
10. K.G. Larkin, D.J. Bone, M.A. Oldfield, Natural demodulation of two-dimensional fringe patterns I: general background of the spiral phase quadrature transform. *J. Opt. Soc. Am. A* **18**, 1862–1870 (2001)

Part II

Concepts in Optics

Chapter 8

Geometrical Optics Essentials



Geometrical optics is a ray-based model for describing propagation of light in small wavelength limit. We will introduce a few basic notions that will serve as a useful reference when studying Fourier optics description of imaging systems later in this book. This chapter is by no means a detailed description of geometrical optics tools that have become important for designing optical assemblies and assessing their performance.

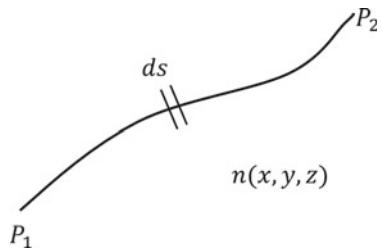
The basic guideline for tracing rays through an optical system is Fermat's principle which states that the optical path length along a ray between two fixed points is an extremum (minimum or maximum) when compared to the neighboring paths. The optical path length (Fig. 8.1) along a ray joining the two points P_1 and P_2 in space is given by

$$S = \int_{P_1}^{P_2} ds n(x, y, z), \quad (8.1)$$

where $n(x, y, z)$ is the refractive index of the medium at the location (x, y, z) along the elemental path length ds . For a given distribution of the refractive index, the ray path is derived such that the first variation δS of the integral above vanishes. The basic laws such as reflection and refraction governing ray propagation may be obtained using this principle. The ray path in this formalism is obtained by solving the corresponding Euler-Lagrange equation.

Optical system assemblies consisting of lenses or other elements made up of reflecting/refracting surfaces and segments of free space are now commonly modeled by careful ray tracing techniques that are now available in the form of a number of commercially available software tools. While ray tracing is considered an approximate model, these tools are very useful in analysis and design of complicated optical assemblies. For example, high-magnification microscope objectives may easily involve more than 20 optical elements and ray tracing is often the only tool available to predict performance of such assemblies. The availability of computational tools

Fig. 8.1 The path taken by a light ray through an inhomogeneous medium as described by a position-dependent refractive index. The optical path length is an integral of nds



for ray tracing has also helped emergence of unconventional optics in the form of aspheric and freeform optical elements.

8.1 Ray Transfer Matrix

The first-order transfer matrix (or ABCD matrix) formalism aims at connecting the position \mathbf{r} and direction θ of a light ray between two planes connected by free space or by an optical system, e.g. an imaging system (see Fig. 8.2). The relation between the output and input parameters of the ray is typically expressed as the ABCD matrix. This representation is an approximation that is typically valid for paraxial rays.

$$\begin{pmatrix} \mathbf{r}' \\ n'\mathbf{u}' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ n\mathbf{u} \end{pmatrix}. \quad (8.2)$$

The vectors, denoting the position of intersection of a ray \mathbf{r} at a plane and its direction, have two components and thus the ABCD matrix must contain 16 numbers in general. However, for optical systems that are symmetric about the yz -plane, it can be shown that only two numbers—the coordinate along y -axis and the corresponding direction cosine are relevant. The ABCD matrix for a 2D rotationally symmetric system thus consists of only four numbers. This representation of the ray transfer matrix is often referred to as the first-order approximation as the matrix elements A, B, C, D may be considered as resulting from first-order Taylor series expansion about the axial ray as follows:

$$\begin{aligned} A &= \frac{\partial y'}{\partial y} \Big|_{y=0, n u_y=0}, & B &= \frac{\partial y'}{\partial (n u_y)} \Big|_{y=0, n u_y=0}, \\ C &= \frac{\partial (n' u_y)}{\partial y} \Big|_{y=0, n u_y=0}, & D &= \frac{\partial (n' u_y)}{\partial (n u_y)} \Big|_{y=0, n u_y=0}. \end{aligned} \quad (8.3)$$

Using the definitions of the matrix elements A, B, C, D as in Eq. (8.3), we observe that the transfer matrices for simple cases of propagation through homogeneous medium and bending of rays at a spherical surface of radius R may be described by the following matrices. For homogeneous medium, the transfer matrix is given by

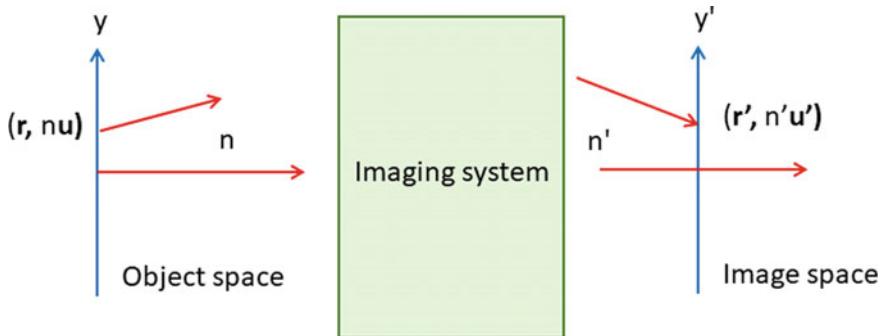


Fig. 8.2 The transfer matrix formalism connects the position and direction vectors corresponding to a ray between two planes connected by free space or by an optical system

$$T = \begin{pmatrix} 1 & \frac{d}{n} \\ 0 & 1 \end{pmatrix}. \quad (8.4)$$

Here d is the distance of propagation along z -axis and n is the index of the medium. For a curved spherical surface,

$$T = \begin{pmatrix} 1 & 0 \\ -\frac{\Delta n}{R} & 1 \end{pmatrix}. \quad (8.5)$$

Here, $\Delta n = (n' - n)$ is the difference between the refractive indices of the media on the two sides of the spherical surface. In the convention where propagation direction is from left to right, the radius R of curvature is assigned a positive sign if the center of curvature is to the right of the surface. For a thin lens in free space, with curvatures of two surfaces equal to R_1 and R_2 , we have a transfer matrix that is a product of two matrices corresponding to the curved surfaces:

$$T = \begin{pmatrix} 1 & 0 \\ -(n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) & 1 \end{pmatrix}. \quad (8.6)$$

The transfer matrix formalism is suitable to describe the behavior of a light ray as it passes through a cascade of systems. For example, if the transfer matrices \mathbf{T}_1 and \mathbf{T}_2 represent the ABCD matrices between planes I-II and planes II-III, respectively, then the transfer matrix between planes I-III is given by the matrix product $\mathbf{T}_1 \mathbf{T}_2$. This simple idea can be used to describe ray propagation through complex optical systems. In principle, the transfer matrix formalism has a close link to the Fourier optics description where a plane wave Fourier component is associated in the ray direction ($n\mathbf{u}$) with an appropriate weight.

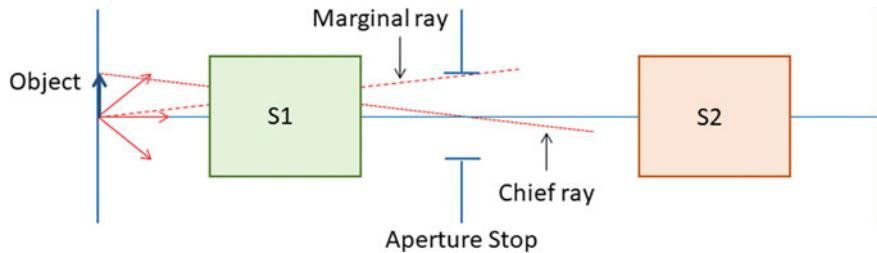


Fig. 8.3 Schematic diagram illustrating the idea of aperture stop. Also shown are the chief ray and the marginal ray

8.2 Stops and Pupils

In an optical system, the aperture stop is a physical aperture that limits the size of the cone of rays originating from an axial object point. We will consider two subsystems S_1 and S_2 as shown in Fig. 8.3. The image of an aperture stop formed in subsystem S_1 in object space is referred to as the entrance pupil while the image of the aperture stop formed by the subsystem S_2 in the image space is known as the exit pupil. The distance between the exit pupil and the image plane is an important parameter as it decides the spatial frequency scaling as we will see later when studying Fourier optics analysis of optical systems. The angular range of the cone of rays allowed by the aperture stop also decides the diffraction-limited resolution of the system. In tracing rays for understanding image formation, typically two rays are considered to be of utmost importance. The *marginal ray* originates from the axial object point and passes through the edge of the aperture stop, while the *chief ray* originates from the edge of the object and passes through the axial point of the aperture stop. The two-parameter column vectors $\begin{pmatrix} y \\ n u_y \end{pmatrix}$ corresponding to the marginal and chief rays are linearly independent and any other ray passing through the system can be thought as a linear combination of these two rays up to the first-order approximation that has been used in this description. The treatment of geometrical optics as presented here is very brief and is just supposed to provide some reference material to common terms that will be used later. The interested reader may refer to standard detailed texts on this subject [1–3].

Problems

- 8.1** Find the transfer matrix for the achromatic doublet lens with n_1 and n_2 as the refractive indices of convex and concave lenses, respectively.

8.2 Use the ABCD matrix approach to compute the focal length of a lens with on-axis thickness T and radii of curvature R_1, R_2 . The index of the lens material may be assumed to be n .

8.3 If an object at a distance z_1 from the front focus of the lens is imaged at a distance z_2 behind the back focus of a lens, find the focal length of the lens.

8.4 Find the transfer matrix for the configuration $F\text{-Lens-}2F\text{-Lens-}F$. Here, F denotes the focal length of the two lenses.

8.5 Think about how you can include a mirror (flat or curved) in the transfer matrix formulation.

References

1. M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, (Ed. 7) (Cambridge University Press, 1999)
2. W. Smith, *Modern Optical Engineering* (SPIE Press, 2007)
3. R.K. Luneberg, *Mathematical Theory of Optics* (University of California Press, 2021). [originally published in 1964]

Chapter 9

Wave Equation and Diffraction of Light



The first part of the book introduced a number of important mathematical ideas on image or signal representation and reconstruction. We are now in a position to apply these ideas to concepts in optics that are important for understanding imaging systems [1, 2]. In the previous chapter, we briefly introduced the geometrical optics or ray picture of imaging systems mainly for introducing some notation. For example, we know that a parallel set of rays are focused at the back-focal plane of a lens. Is the physical focus truly a point? How well does geometrical optics work? For answering this kind of questions, we need to treat light as a wave phenomenon. Diffraction of light is the fundamental problem of interest here. What is diffraction? Diffraction refers to any bending or spreading of light waves that cannot be attributed to reflection or refraction at a boundary between two materials. Why do waves diffract? They do so due to confinement in transverse direction and the spread of the waves may therefore be associated with a form of the uncertainty relation. As mentioned in the introductory chapter, when light waves interact with obstacles or objects of interest, the scattered waves carry information about the object with them. When designing any imaging system, we need a model for describing the waves that reach our system and the information carried by them. Diffraction theory provides a rigorous way to understand how information about features of different length scales (relative to illumination wavelength) in the object of interest is propagated to the imaging system. In the context of computational imaging systems, diffraction formalism provides a guideline for design of appropriate imaging hardware and image reconstruction algorithms.

Historically the earliest known record of observation of diffraction phenomenon is found in Grimaldi's experiments published in 1665. He observed that transition of intensity from bright to dark at the edge of a shadow region is not sharp but gradual. This observation was contradictory to the corpuscular theory of light which suggested propagation of light in a straight line. Huygens in 1698 proposed a picture for propagation of light waves. He put forth the idea of what we now call the wavefront. In this picture, every point on the wavefront is a secondary source of spherical waves. The envelope of a collection of spherical wavefronts gives the corresponding

wavefront at a future time. This qualitative intuitive picture is useful but it does not allow us to actually calculate how waves diffract when they encounter obstacles. This wave picture of light was not worked on for almost 100 years when in 1804 Thomas Young performed the two-slit experiment to put forth the idea of interference of light waves. The credit of developing a first quantitative picture for diffraction goes to Fresnel. In his 1818 paper presented to French academy of sciences, he combined Young's idea of interference and Huygens' picture of secondary waves to present a theory for diffraction. The wave theory of light was still suspected at that point of time. Poisson, who was heading the committee that examined the submitted papers, ridiculed Fresnel's idea. He showed that if Fresnel's theory was true then an opaque circular disk obstacle should have a bright point on axis. F. Arago who was a committee member at the academy went to the laboratory and performed this experiment to observe the bright spot! The experimental observation confirmed Fresnel's theory and also to a large extent the wave nature of light. Without going further into interesting historical details [3] we will begin our analysis of the diffraction phenomenon with a review of Maxwell equations that form the well-accepted basis for description of wave phenomena. In the treatment presented here we will avoid the usage of traditional ideas resulting out of contributions of Huygens, Fresnel and Kirchhoff [4]. Instead we will present the angular spectrum framework that is consistent with the Maxwell equations. *The angular spectrum method as discussed here clearly distinguishes between features of the object smaller or larger compared to wavelength of illumination.* As we will explain later in this chapter, this method is equivalent to the exact Rayleigh-Sommerfeld solution to the diffraction problem. Further the angular spectrum approach is amenable to a straightforward computation of diffraction from planar apertures that is key to several practical settings in coherent optical imaging (e.g. digital holographic imaging).

9.1 Review of Maxwell Equations

We will start by writing the Maxwell equations that govern the behavior of \mathbf{E} and \mathbf{B} fields in terms of total charges and currents [5]. The four equations in the MKS system take the form:

$$\begin{aligned}\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \\ \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0}, \\ \nabla \cdot \mathbf{B} &= 0.\end{aligned}\tag{9.1}$$

Along with the continuity relation:

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0, \quad (9.2)$$

and suitable boundary conditions one may, at least in principle, solve any problem involving electromagnetic fields. We will consider the case of EM fields in region of space where there are no charges and currents thus reducing the Maxwell equations to a simpler form:

$$\begin{aligned}\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \times \mathbf{B} &= \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \\ \nabla \cdot \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{B} &= 0.\end{aligned}\quad (9.3)$$

Since we are interested in understanding how \mathbf{E} or \mathbf{B} fields behave as they propagate in free space, we will now proceed to obtain the wave equation. Taking the curl of the first equation gives

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{\partial(\nabla \times \mathbf{B})}{\partial t}, \quad (9.4)$$

and further substitution of the curl \mathbf{B} equation gives the wave equation:

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (9.5)$$

Here we have used the relation $c = 1/\sqrt{\epsilon_0 \mu_0}$ to denote the speed of light in vacuum. The magnetic field obeys the same wave equation as the electric field. Any time dependence of the \mathbf{E} field may be modeled by Fourier decomposition, and it is therefore customary to analyze the wave equation by a monochromatic or steady state representation. In this approach, the time dependence of the solution at each frequency is purely periodic. We will represent the fields with the $\exp(-i2\pi\nu t)$ representation as

$$\mathbf{E}(\mathbf{r}, t) = \text{Re}\{\mathbf{E}_v(\mathbf{r}) \exp(-i2\pi\nu t)\}. \quad (9.6)$$

In this representation, the wave equation for the space-part \mathbf{E}_v at the frequency v takes the form:

$$(\nabla^2 + k^2)\mathbf{E}_v(\mathbf{r}) = 0, \quad (9.7)$$

which is known as the homogeneous (source-free) Helmholtz equation and k stands for the magnitude of the wave vector $k = (2\pi v)/c$. The scalar diffraction theory

that as we will study here deals with the single component of the \mathbf{E}_v field, in effect, disregarding the coupling of other field components through Maxwell equations.

9.2 Weyl Representation of Spherical Waves

In this section, we will obtain an important result regarding the 2D Fourier representation of a spherical wave which will be found useful in the next section on angular spectrum representation. The starting point of this exercise is the fact that a spherical wave $q = \exp(ikr)/(4\pi r)$ is a solution of the wave equation for a delta-function excitation source [6].

$$(\nabla^2 + k^2) q(x, y, z) = -\delta(x)\delta(y)\delta(z). \quad (9.8)$$

Here $r = \sqrt{x^2 + y^2 + z^2}$ is the radial coordinate. Writing $q(x, y, z)$ as a 3D Fourier expansion and applying the $(\nabla^2 + k^2)$ we observe that

$$\begin{aligned} & \iiint df_x df_y df_z [k^2 - 4\pi^2(f_x^2 + f_y^2 + f_z^2)] Q(f_x, f_y, f_z) \exp[i2\pi(f_x x + f_y y + f_z z)] \\ &= - \iiint df_x df_y df_z \exp[i2\pi(f_x x + f_y y + f_z z)], \end{aligned} \quad (9.9)$$

leading to

$$Q(f_x, f_y, f_z) = \frac{-1}{[k^2 - 4\pi^2(f_x^2 + f_y^2 + f_z^2)]}. \quad (9.10)$$

In order to obtain a 2D Fourier representation $\tilde{Q}(f_x, f_y; z)$ for the spherical wave $q(x, y, z)$, we have to integrate the above expression for $Q(f_x, f_y, f_z)$ over the variable f_z .

$$\begin{aligned} \tilde{Q}(f_x, f_y; z) &= - \int df_z \frac{\exp(i2\pi f_z z)}{[k^2 - 4\pi^2(f_x^2 + f_y^2) - 4\pi^2 f_z^2]} \\ &= \int df_z \exp(i2\pi f_z z) \frac{1}{4\pi\alpha} \left[\frac{1}{f_z - \alpha/(2\pi)} - \frac{1}{f_z + \alpha/(2\pi)} \right]. \end{aligned} \quad (9.11)$$

with $\alpha = +\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}$. The integral above must be understood as its principal value since the integrand has poles on the real line at $f_z = \pm(1/2\pi)\alpha$. For a realistic medium if we assume a small imaginary component in k , the two poles in the real line shift up or down as shown in Fig. 9.1 making the integral meaningful. For spherical wave propagating in the right half space we have positive z and therefore the contour includes a semicircle with radius tending to infinity in the upper half

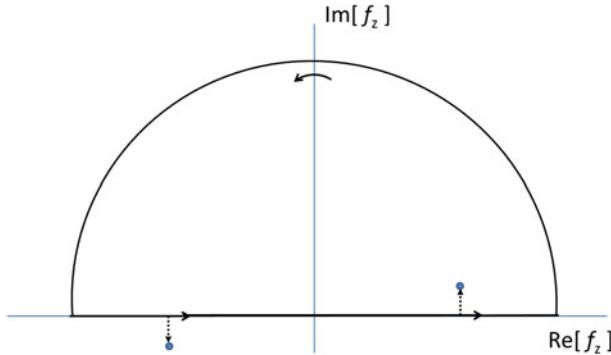


Fig. 9.1 Contour used for integration in Eq. (9.11). The two poles on the real line are shifted as shown if k has a small positive imaginary part

space. Using the Cauchy residue theorem for the simple pole inside the contour, we can obtain the value of $\tilde{Q}(f_x, f_y; z)$ as

$$\tilde{Q}(f_x, f_y; z) = \frac{\exp(ikr)}{4\pi r} = i \iint df_x df_y \frac{\exp[i2\pi(f_x x + f_y y) + i\alpha z]}{2\alpha}. \quad (9.12)$$

9.3 Angular Spectrum Method

We will now introduce an interesting approach to the diffraction problem that is consistent with the linear systems theory developed in the first part of this book and is called as the angular spectrum method [7]. We will start with the Helmholtz equation for a transverse component of the E-field, say, E_x :

$$(\nabla^2 + k^2) E_x(x, y, z) = 0. \quad (9.13)$$

Our aim here is to determine the field $E_x(x, y, z)$ as a function of the field in the $z = 0$ plane as denoted by $E_x(x, y, 0)$. The evolution of the field between these two planes is guided by means of the Helmholtz equation, and we will obtain the solution to this problem that is consistent with the boundary condition at $z \rightarrow \infty$. We begin by expressing E_x as a 2D Fourier transform in the x and y coordinates. The z coordinate is treated as special as this is the direction in which we want to find the propagated fields. The 2D Fourier representation of the field component E_x is given by

$$E_x(x, y, z) = \iint df_x df_y \tilde{E}_x(f_x, f_y; z) \exp[i2\pi(f_x x + f_y y)]. \quad (9.14)$$

Applying the Helmholtz operator and equating the integrand in the resultant Fourier expansion to zero (since the integral is identically zero for all x and y) gives us a differential equation for \tilde{E}_x in variable z :

$$\frac{d^2}{dz^2} \tilde{E}_x(f_x, f_y; z) = -[k^2 - 4\pi^2(f_x^2 + f_y^2)] \tilde{E}_x(f_x, f_y; z). \quad (9.15)$$

Denoting $\alpha = +\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}$ as in the previous section, we can write the solution of this equation as

$$\tilde{E}_x(f_x, f_y; z) = A(f_x, f_y) \exp(i\alpha z) + B(f_x, f_y) \exp(-i\alpha z). \quad (9.16)$$

We now need to make a careful choice of the functions $A(f_x, f_y)$ and $B(f_x, f_y)$ to satisfy the boundary conditions of the problem at hand. First of all we note that the Fourier spectrum of the field in $z = 0$ plane is given by

$$\tilde{E}_x(f_x, f_y; 0) = A(f_x, f_y) + B(f_x, f_y). \quad (9.17)$$

It is important to understand the behavior of α as we move away from origin in the (f_x, f_y) plane. We note that α is a real number for $k^2 \geq 4\pi^2(f_x^2 + f_y^2)$ and we will choose to take the positive value of the square root. However, α becomes imaginary when $k^2 < 4\pi^2(f_x^2 + f_y^2)$. We will use the following convention:

$$\begin{aligned} \alpha &= +\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}, & k^2 \geq 4\pi^2(f_x^2 + f_y^2) \\ &= i\sqrt{4\pi^2(f_x^2 + f_y^2) - k^2}, & k^2 < 4\pi^2(f_x^2 + f_y^2) \end{aligned} \quad (9.18)$$

Note that we now have an explicit connection between the wavelength (since $k = 2\pi/\lambda$) and the spatial frequencies f_x and f_y . We notice that when $k^2 < 4\pi^2(f_x^2 + f_y^2)$ or for spatial frequencies greater than $1/\lambda$, the A term in Eq. (9.16) decays exponentially whereas the B term increases exponentially as we go in the positive z direction. To have a physically meaningful solution as $z \rightarrow \infty$, we must set $B(f_x, f_y) = 0$. Notice that the B term corresponds to an incoming wave solution. Since we are interested in the outward propagating solutions, this condition on $B(f_x, f_y)$ is in some sense equivalent to Sommerfeld's radiation condition. This choice of $B(f_x, f_y)$ along with Eq. (9.17) gives $A(f_x, f_y) = \tilde{E}_x(f_x, f_y; 0)$ and we have the following solution for the propagation problem:

$$\tilde{E}_x(f_x, f_y; z > 0) = \tilde{E}_x(f_x, f_y; 0) \exp(i\alpha z). \quad (9.19)$$

The inverse Fourier transform of the term $\exp(i\alpha z)$ is now needed for obtaining the field $E_x(x, y, z)$ in terms of $E_x(x, y, 0)$. In order to evaluate the inverse transform we observe that

$$\begin{aligned} & \iint_{-\infty}^{\infty} df_x df_y \exp(i\alpha z) \exp[i2\pi(f_x x + f_y y)] \\ &= -\frac{\partial}{\partial z} \left\{ \iint_{-\infty}^{\infty} df_x df_y \frac{i}{\alpha} \exp(i\alpha z) \exp[i2\pi(f_x x + f_y y)] \right\}. \end{aligned} \quad (9.20)$$

We note that the quantity included in the curly brackets above is related to the Weyl representation for a spherical wave as per Eq. (9.12). We may therefore get

$$\mathcal{F}^{-1}\{\exp(i\alpha z)\} = -\frac{\partial}{\partial z} \left[\frac{\exp(ikr)}{2\pi r} \right] = \frac{\exp(ikr)}{2\pi r} \left(-ik + \frac{1}{r} \right) \frac{z}{r}, \quad (9.21)$$

where $r = \sqrt{x^2 + y^2 + z^2}$. The inversion of Eq. (9.19) may now be represented as a convolution integral given by

$$E_x(x, y, z) = \iint dx' dy' E_x(x', y', 0) \frac{\exp(ikR)}{2\pi R} \left(-ik + \frac{1}{R} \right) \frac{z}{R}, \quad (9.22)$$

where the distance R is defined as

$$R = \sqrt{(x - x')^2 + (y - y')^2 + (z - 0)^2}. \quad (9.23)$$

This relation is known as the first Rayleigh-Sommerfeld solution to the diffraction problem. It is important to note that the response of a point source located in the $z = 0$ plane is not a spherical wave (as in Huygen's original idea) but a modified form:

$$\frac{\exp(ikR)}{2\pi R} \left(-ik + \frac{1}{R} \right) \frac{z}{R}.$$

The Rayleigh-Sommerfeld solution in Eq. (9.22) may also be expressed in terms of a partial derivative in z :

$$E_x(x, y, z) = -\frac{\partial}{\partial z} \iint dx' dy' E_x(x', y', 0) \frac{\exp(ikR)}{2\pi R}. \quad (9.24)$$

We could have solved the identical problem as above in a coordinate system rotated 90 degrees with respect to the z axis and so we get a similar relation for the y -component of the diffraction field.

$$E_y(x, y, z) = -\frac{\partial}{\partial z} \iint dx' dy' E_y(x', y', 0) \frac{\exp(ikR)}{2\pi R}. \quad (9.25)$$

The diffraction fields also have a z -component that may be obtained by observing that both E_x and E_y in Eqs. (9.24) and (9.25) are represented in terms of partial derivatives with respect to z . Further using the coupling of E_z component to the x and y components through the relation $\nabla \cdot \mathbf{E} = 0$ in the right half space, we have

$$\frac{\partial}{\partial z} E_z(x, y, z) = -\frac{\partial}{\partial x} E_x - \frac{\partial}{\partial y} E_y. \quad (9.26)$$

Finally noting that E_z cannot have a non-zero z -independent additive term (in order to make the field vanish at infinity), the $E_z(x, y, z)$ component may be represented as

$$E_z(x, y, z) = \iint dx' dy' \left\{ E_x(x', y', 0) \frac{(x - x')}{R} + E_y(x', y', 0) \frac{(y - y')}{R} \right\} \frac{\exp(ikR)}{2\pi R} \left(-ik + \frac{1}{R} \right). \quad (9.27)$$

The results in Eqs. (9.24), (9.25), (9.27) may be written in a compact form as first suggested by W. R. Smythe [8]:

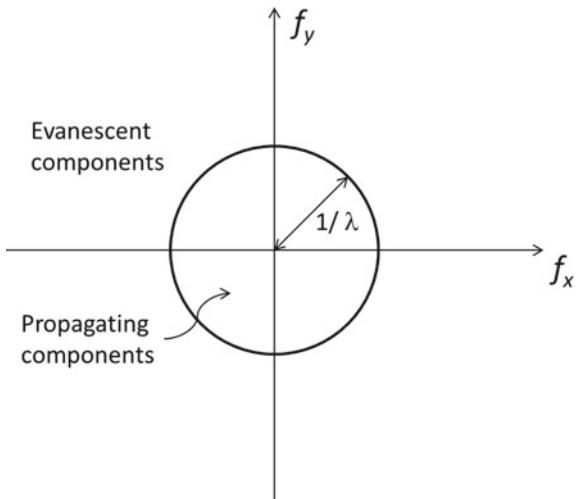
$$\mathbf{E}(x, y, z) = \nabla \times \iint dx' dy' [\hat{z} \times \mathbf{E}(x', y', 0)] \frac{\exp(ikR)}{2\pi R}, \quad (9.28)$$

and are referred to as the Rayleigh-Sommerfeld-Smythe relations. These relations are exact solutions of Maxwell equations for the electromagnetic diffraction problem. We observe that all the three components of the \mathbf{E} field at any point $(x, y, z > 0)$ in the right half space are therefore completely determined by specifying the tangential components $E_x(x, y, 0)$ and $E_y(x, y, 0)$ in the $z = 0$ plane. This may be considered as a starting point for holography. As long as we can specify or generate the required tangential fields in the hologram plane, the hologram on replay will give us an impression as if the fields entering our eye actually originated from a 3D object placed appropriately. The z -component of the field is usually ignored in many problems in para-axial regime where $(x - x')/R$ or $(y - y')/R$ are much smaller than z/R . However this component may become significant particularly in case of fields for high numerical aperture systems near the focus.

We note that the treatment of the diffraction problem provided above is exact and fully consistent with Maxwell equations. The Rayleigh-Sommerfeld-Smythe relations tell us that when the tangential fields in the $z = 0$ plane are known, all components of the electric field in the right half space ($z > 0$) are uniquely determined. It is however important to observe that the Rayleigh-Sommerfeld-Smythe relations by themselves are not sufficient to determine the tangential fields in the aperture. Determination of aperture fields that may be used for calculating diffraction fields in $z > 0$ space is a separate problem in itself and needs a careful treatment particularly when the aperture size is comparable or smaller than the wavelength λ of illumination used. This interesting topic is not within the scope of this book and will not be discussed here. Some consequences that follow readily in the angular spectrum description are as follows (refer to Fig. 9.2).

- 1. Propagating waves:** We note that Fourier components of the input field $E_x(x, y, 0)$ (or $E_y(x, y, 0)$) corresponding to spatial frequencies with $(f_x^2 + f_y^2) < 1/\lambda^2$, get assigned a phase factor $\exp(i\alpha z)$ due to propagation in $+z$ direction.

Fig. 9.2 2D Fourier space showing the propagating and evanescent spatial frequency components



tion. Their Fourier magnitude stays the same. These components are therefore called the propagating components. They carry detectable information about the obstacle/aperture/object even to large distances ($R \gg \lambda$) from the aperture. If we write out the inverse Fourier transform relation:

$$E_x(x, y, z) = \iint df_x df_y \tilde{E}_x(f_x, f_y; 0) \exp[i2\pi(f_x x + f_y y) + i\alpha z]. \quad (9.29)$$

The component corresponding to spatial frequencies (f_x, f_y) can be considered as a plane wave of weight $\tilde{E}_x(f_x, f_y; 0)$ that makes angles

$$\theta_x = \cos^{-1}(\lambda f_x), \quad \theta_y = \cos^{-1}(\lambda f_y) \text{ and } \theta_z = \cos^{-1}(\sqrt{1 - \lambda^2(f_x^2 + f_y^2)})$$

with the $+x$, $+y$, $+z$ axes respectively and this is the justification for the name “angular spectrum”. Note that zero spatial frequency $(f_x, f_y) = (0, 0)$ corresponds to a plane wave propagating in the $+z$ direction. The spatial frequencies $\lambda^2(f_x^2 + f_y^2) = 1$ correspond to plane waves propagating at 90 degrees to the $+z$ axis. As the details in the aperture plane become finer, their diffraction corresponds to plane waves making larger angles with the $+z$ axis.

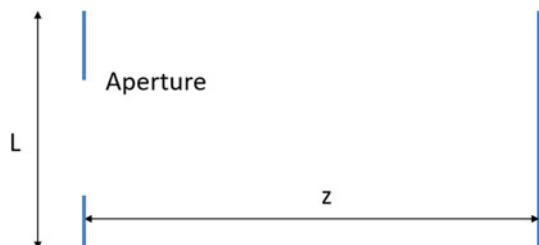
2. **Evanescent waves:** The Fourier components of the input field $E_x(x, y, 0)$ (or $E_y(x, y, 0)$) corresponding to spatial frequencies with $(f_x^2 + f_y^2) > 1/\lambda^2$, get assigned a decaying factor $\exp(-|\alpha|z)$. These wave components decay exponentially as we move in $+z$ direction from the aperture. The decaying components are typically negligible in magnitude as soon as we are several wavelengths away from the aperture when considering the problem of diffraction in the right half space.

The importance of the angular spectrum approach is that the wavelength λ appears explicitly in relation to the scale of features in the aperture. The behavior of the diffraction field for feature sizes larger and smaller than the wavelengths is seen to be distinct. The information about the sub-wavelength features is not available (for all practical purposes) unless we are in a region that is few wavelengths from the aperture plane. Probing structures such as sharp metal-coated optical fiber tips can however pick up this information when placed very close to the aperture plane as is done in near field scanning optical microscopes. Sub-wavelength structures placed very close to the aperture can also convert some of the evanescent components into propagating components resulting in possibility of detecting information about sub-wavelength features in the far-field.

9.4 Numerical Computation of Diffraction Fields Using Angular Spectrum Method

The numerical computation of diffraction fields is often required in practice. We note that the transfer function for angular spectrum given by $H(f_x, f_y) = \exp(i\alpha z)$ has highly oscillatory behavior and careful attention must be paid to sample this factor correctly. Let us suppose that we have a computational window of size $L \times L$ at the two planes of interest as shown in Fig. 9.3. Our goal is to understand how many samples N over length L should be used so that the transfer function is appropriately represented on a discrete grid. As per notations introduced in Chap. 4, the sampling interval in (x, y) domain is $\Delta x = \Delta y = L/N$. When using the FFT algorithm to evaluate the 2D Fourier transform, the corresponding interval in the spatial frequency space is given by $\Delta f_x = \Delta f_y = 1/(N \Delta x) = 1/L$. In practical settings, the sampling interval Δx is often decided by experimental considerations such as camera pixel size. We will first consider the change in the phase of the transfer function $\exp(i\alpha z)$ over one pixel in the spatial frequency space. The change in phase may for example be evaluated along the f_x -axis as

Fig. 9.3 Computational setting for angular spectrum propagation



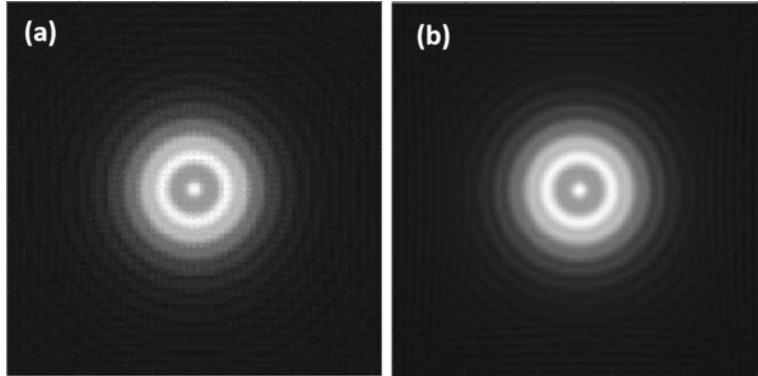


Fig. 9.4 Illustration of angular spectrum propagation of a circular aperture of radius $a = 250 \mu\text{m}$ performed on a square computational grid with $\Delta x = 5 \mu\text{m}$, $N = 256$ and propagation distance of $z = 5 \text{ cm}$. **a** and **b** show diffracted field amplitudes obtained with or without applying the low-pass filter of with $2f_0 \times 2f_0$ with f_0 defined in Eq. 9.32

$$\Delta\phi = \left[\frac{\partial(\alpha z)}{\partial f_x} \right] \Delta f_x = \frac{-4\pi^2 z f_x}{\sqrt{k^2 - 4\pi^2 f_x^2}} \cdot \frac{1}{N \Delta x}. \quad (9.30)$$

As per Nyquist criterion, the change in phase per pixel should be less than or equal to π which leads to the condition:

$$\frac{4\pi z f_x}{\sqrt{k^2 - 4\pi^2 f_x^2}} \cdot \frac{1}{N \Delta x} \leq 1, \quad (9.31)$$

or equivalently,

$$f_x \leq f_0 = \frac{1}{\lambda \sqrt{1 + (2z/N \Delta x)^2}}. \quad (9.32)$$

In order to avoid aliasing problems in angular spectrum computation, a filter of size $f_x \leq f_0$ therefore needs to be applied in the Fourier domain in addition to the transfer function [9]. In Fig. 9.4, we show illustration of diffraction field amplitude for a circular aperture of radius $250 \mu\text{m}$ over a distance of 5 cm . The computational window has 256×256 pixels and the sampling interval is $5 \mu\text{m}$. Note that when the low-pass filter as defined in Eq. (9.32) is applied, the aliasing artifacts in Fig. 9.4a are seen to vanish as shown in Fig. 9.4b.

9.5 Fresnel and Fraunhofer Approximations

While the angular spectrum solution corresponds to the exact Rayleigh-Sommerfeld diffraction relation, laboratory experiments are often performed such that the area of observation (as decided by detector size) is in the paraxial region relative to the

aperture. In such cases, a smaller extent of spatial frequencies near $f_x = f_y = 0$ are sufficient to describe the observations in a diffraction experiment. In this context, we will explain how the Rayleigh-Sommerfeld diffraction solution may be expressed in an analytically simpler form. We will start with the exact Rayleigh-Sommerfeld solution in Eq. (9.22) for one scalar component of the diffraction field and perform simplifying approximation to arrive at the Fresnel diffraction formula. The x -component E_x of the scalar field may be given by

$$E_x(x, y, z) = \iint_{S_A} dx' dy' E_x(x', y', 0) \frac{\exp(ikR)}{2\pi R} \left(-ik + \frac{1}{R}\right) \frac{z}{R}. \quad (9.33)$$

We note that for a finite aperture S_A , all the terms in the integral need not be included at distances much larger than the wavelength λ of illumination. For example, we note that for optical wavelengths $|k| >> 1/R$ as in typical tabletop experimental settings. The wide-field form of Fresnel diffraction formula may then be stated as

$$E_x(x, y, z) = \frac{1}{i\lambda} \iint_{S_A} dx' dy' E_x(x', y', 0) \frac{\exp(ikR)}{2\pi R} \frac{z}{R}, \quad (9.34)$$

Further in the paraxial (or near-axis) region, we may assume $z/R \approx 1$ and also the approximation $1/R \approx 1/z$ may be used. The term $\exp(ikR)$ needs to be handled carefully. In the paraxial region we use a binomial expansion for the distance R .

$$\begin{aligned} R &= \sqrt{(x - x')^2 + (y - y')^2 + z^2} \\ &\approx z \left[1 + \frac{(x - x')^2 + (y - y')^2}{2z^2} - \frac{1}{8} \left(\frac{(x - x')^2 + (y - y')^2}{z^2} \right)^2 + \dots \right]. \end{aligned} \quad (9.35)$$

We may neglect the third term in the binomial expansion above if it is much less than 1 radian. This condition may be stated more explicitly in the form:

$$z^3 >> \frac{k}{8} [(x - x')^2 + (y - y')^2]_{\max}^2, \quad (9.36)$$

where the subscript “max” denotes the maximum value of the quantity inside the braces, the maximum value being decided by the size of the region representing the aperture S_A . We will refer to the region of right half space satisfying this condition as the Fresnel zone. The paraxial form of the Fresnel zone diffraction formula may now be written as

$$\begin{aligned} E_x(x, y, z) &= \frac{\exp(ikz)}{i\lambda z} \iint_{S_A} dx' dy' E_x(x', y', 0) \exp \left\{ \frac{i\pi}{\lambda z} [(x - x')^2 + (y - y')^2] \right\}. \end{aligned} \quad (9.37)$$

We note that the above formula is a 2D convolution of the field distribution in the $z = 0$ plane with the Fresnel impulse response $h_F(x, y; z)$ given by

$$h_F(x, y; z) = \frac{\exp(ikz)}{i\lambda z} \exp\left[\frac{i\pi}{\lambda z}(x^2 + y^2)\right]. \quad (9.38)$$

The corresponding transfer function has already been evaluated in Sect. 2.8. Using this result, we may write the transfer function for Fresnel diffraction as

$$H_F(f_x, f_y; z) = \exp(ikz) \exp[-i\pi\lambda z(f_x^2 + f_y^2)], \quad (9.39)$$

so that, the Fresnel diffraction formula may be written in the transfer function form as

$$\tilde{E}_x(f_x, f_y; z) = \tilde{E}_x(f_x, f_y; 0) H_F(f_x, f_y; z). \quad (9.40)$$

We note that this result may also be obtained by approximating the angular spectrum transfer function for the paraxial region of right half space. Particularly we need to consider the spatial frequencies close to $f_x = f_y = 0$ in this case and use the approximation:

$$\exp(i z \sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}) \approx \exp[i k z (1 - 2\pi^2(f_x^2 + f_y^2)/k^2)], \quad (9.41)$$

which is the same as the Fresnel diffraction transfer function in Eq. (9.39).

9.5.1 Transport of Intensity Equation (TIE)

An interesting property satisfied by the diffraction field for z in Fresnel approximation zone as above is the relation between the z -intensity derivative of the field and the transverse gradient of the phase of the field [10]. Suppose we represent the diffraction field as

$$u(x, y, z) = \sqrt{I(x, y, z)} \exp[i\theta(x, y, z)]. \quad (9.42)$$

Here the intensity and phase of the field $u(x, y, z)$ are denoted by $I(x, y, z) = |u(x, y, z)|^2$ ignoring any additional constant factors and $\theta(x, y, z)$ respectively. We will first obtain an approximate wave equation satisfied by the para-axial Fresnel diffraction field $u(x, y, z)$. Using the transfer function for Fresnel diffraction in Eq. (9.39), we may represent the field $u(x, y, z)$ as

$$\begin{aligned} & u(x, y, z) \\ &= \exp(ikz) \iint df_x df_y U(f_x, f_y, 0) \exp[-i\pi\lambda z(f_x^2 + f_y^2) + i2\pi(f_x x + f_y y)]. \end{aligned} \quad (9.43)$$

Using this Fourier representation, it is easy to verify that $u(x, y, z)$ satisfies the following differential equation:

$$\left[i \frac{\partial}{\partial z} + \frac{1}{2k} \frac{\partial^2}{\partial x^2} + \frac{1}{2k} \frac{\partial^2}{\partial y^2} + k \right] u(x, y, z) = 0. \quad (9.44)$$

This equation is referred to as the para-axial or parabolic form of wave equation. If Eq. (9.44) is multiplied on the left by $u^*(x, y, z)$ and the conjugate equation is multiplied on the left by $u(x, y, z)$, then the two resultant equations on subtraction give

$$-\frac{2\pi}{\lambda} \frac{\partial I(x, y, z)}{\partial z} = \nabla_{\perp} \cdot [I(x, y, z) \nabla_{\perp} \theta(x, y, z)], \quad (9.45)$$

where ∇_{\perp} refers to the transverse vector operator $(\partial/\partial x, \partial/\partial y)$. This equation is known as the transport of intensity equation (TIE). From an experimental point of view this equation implies that in the Fresnel diffraction approximation, the z intensity derivative may be used to determine the phase $\theta(x, y, z)$ of a wavefront if the differential equation above can be solved. The intensity derivative can be obtained for example by recording two intensity images $I(x, y, z - \Delta z/2)$ and $I(x, y, z + \Delta z/2)$ with longitudinal translation of an area detector. The z -derivative may then be expressed as

$$\frac{\partial I}{\partial z} = \frac{I(x, y, z + \Delta z/2) - I(x, y, z - \Delta z/2)}{\Delta z} + O((\Delta z)^2). \quad (9.46)$$

In an experimental record of two intensities as above, the choice of Δz involves a trade-off. In particular smaller Δz provides a better approximation to the longitudinal intensity derivative. On the other hand, Δz needs to be sufficiently large so that the difference in intensities recorded by camera pixels on small translation of camera is above the noise floor of the camera readout. In order to obtain the phase $\theta(x, y, z)$ we first introduce an auxiliary function $\psi(x, y, z)$ such that

$$\nabla_{\perp} \psi = I \nabla_{\perp} \theta, \quad (9.47)$$

giving

$$-\frac{2\pi}{\lambda} \frac{\partial I}{\partial z} = \nabla_{\perp}^2 \psi. \quad (9.48)$$

This equation may be solved as

$$\psi = -\frac{2\pi}{\lambda} \nabla_{\perp}^{-2} \frac{\partial I}{\partial z}. \quad (9.49)$$

Further using the definition of the auxiliary function ψ above and taking transverse gradient gives

$$\nabla_{\perp} \psi = I \nabla_{\perp} \theta = -\frac{2\pi}{\lambda} \nabla_{\perp} \nabla_{\perp}^{-2} \frac{\partial I}{\partial z}. \quad (9.50)$$

The same idea used for obtaining ψ as in Eq. (9.49) can now be followed for getting the phase function $\theta(x, y, z)$. The final result can be expressed as

$$\theta(x, y, z) = -\frac{2\pi}{\lambda} \nabla_{\perp}^{-2} \nabla_{\perp} \cdot \left[\frac{1}{I} \nabla_{\perp} \nabla_{\perp}^{-2} \left(\frac{\partial I}{\partial z} \right) \right]. \quad (9.51)$$

The inverse Laplacian operator ∇_{\perp}^{-2} demands a brief discussion. In a manner similar to the Laplacian, the operator can be defined easily in the Fourier domain as

$$\nabla_{\perp}^{-2} g(x, y) = \mathcal{F}^{-1} \left[\frac{G(f_x, f_y)}{-4\pi^2(f_x^2 + f_y^2)} \right]. \quad (9.52)$$

Further in order to avoid division by zero, the operator may be practically implemented as

$$\nabla_{\perp}^{-2} g(x, y) \approx \mathcal{F}^{-1} \left\{ \frac{-4\pi^2(f_x^2 + f_y^2)G(f_x, f_y)}{[-4\pi^2(f_x^2 + f_y^2)]^2 + \epsilon^2} \right\}, \quad (9.53)$$

for small positive constant ϵ^2 . The operator can thus be readily implemented using the 2D FFT routines. In the special case when the intensity $I(x, y, z)$ of the field is constant I_0 in a particular plane $z = z_0$, the phase solution in Eq. (9.51) can be simplified considerably and may be written as:

$$\theta(x, y; z_0) = -\frac{2\pi}{\lambda I_0} \nabla_{\perp}^{-2} \frac{\partial I}{\partial z} \quad (9.54)$$

One interesting aspect of solution of the TIE equation is that the phase $\theta(x, y, z)$ is not represented in terms of the arctangent function but rather as a solution of a differential equation. The consequence of this is that the solution θ is not wrapped in the range $[-\pi, \pi]$ which is usually the artificial limitation on the recovered phase maps in experiments in interferometry or other phase estimation methods. TIE may therefore be used as a powerful tool for addressing the unwrapping problem [11]. As an illustration, we consider a continuous phase map with phase values in the range $(0, 2\pi)$ rad as shown in Fig. 9.5a. The corresponding wrapped phase map $\theta_{wrapped}(x, y)$ is shown in Fig. 9.5b. We may now construct an auxiliary field $u(x, y) = \exp[i\theta_{wrapped}(x, y)]$. On propagating this field by $\pm\Delta z$ we may compute the corresponding intensities and the longitudinal intensity derivative. On solving the TIE, one can recover the unwrapped phase as shown in Fig. 9.5c. It may be noted that in this computational approach for unwrapping, the choice of distance Δz is not limited by any detector noise.

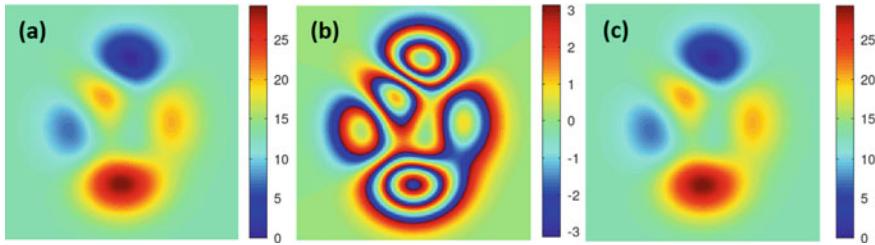


Fig. 9.5 Illustration of phase unwrapping using the transport of intensity equation (TIE). **a** Original phase map, **b** its wrapped version, **c** unwrapped phase map using TIE-based approach

9.5.2 Self-Imaging: Montgomery Conditions and Talbot Effect

Self-imaging is a curious phenomenon in which an object with certain spatial frequency content on illumination with a coherent beam produces intensity distribution at a distance that is similar to the object itself. We will first look at self-imaging from the full angular spectrum considerations. As we have studied in the last chapter, the diffraction problem may be described in the Fourier transform domain using the angular spectrum transfer function as

$$U(f_x, f_y; z) = \exp[iz\sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)}] U(f_x, f_y; 0). \quad (9.55)$$

It is clear that self-imaging will occur if the phase factor in the transfer function (assuming we are restricting ourselves to propagating components) is an integral multiple of 2π . We may denote the corresponding distances as z_m such that

$$z_m \sqrt{k^2 - 4\pi^2(f_x^2 + f_y^2)} = 2\pi m, \quad (9.56)$$

where m is an integer. The condition can be stated in another manner as

$$\left(\frac{m}{z_m}\right)^2 = \left(\frac{1}{\lambda}\right)^2 - (f_x^2 + f_y^2). \quad (9.57)$$

Objects whose spatial frequency content is restricted to the spatial frequencies (circles in Fourier transform space) as per the above condition will show self-imaging at distances z_m from the aperture. It is to be noted that the above restriction in spatial frequency space implies that such objects are inherently infinite in extent. The condition for self-imaging above is due to Montgomery [12]. The paraxial version of this general result is the Talbot effect, in which, self-imaging is demonstrated by periodic structures. Consider an amplitude grating described by a transmission function:

$$t(x, y) = \frac{1}{2}[1 + \cos(2\pi f_0 x)]. \quad (9.58)$$

Fourier transform of this transmission function is given by

$$T(f_x, f_y) = \frac{1}{2} \left[\delta(f_x, f_y) + \frac{1}{2}\delta(f_x - f_0, f_y) + \frac{1}{2}\delta(f_x + f_0, f_y) \right]. \quad (9.59)$$

On plane wave illumination of this grating, the Fresnel diffraction field at distance z from the grating may be expressed in Fourier domain as

$$\begin{aligned} U(f_x, f_y; z) &= T(f_x, f_y) \exp[ikz - i\pi\lambda z(f_x^2 + f_y^2)] \\ &= \frac{1}{2} \exp(ikz)\{\delta(f_x, f_y) + \frac{1}{2} \exp(-i\pi\lambda z f_0^2)[\delta(f_x - f_0, f_y) + \delta(f_x + f_0, f_y)]\}. \end{aligned} \quad (9.60)$$

We observe that for distances z_n for positive integer values of n such that

$$z_n = \frac{2n}{\lambda f_0^2} = \frac{2np^2}{\lambda}, \quad (9.61)$$

with $p = 1/f_0$ denoting the grating period, the exponential factor $e^{-i\pi\lambda z f_0^2}$ is equal to $\exp(-i2\pi n)$ and thus the Fourier transform $U(f_x, f_y; z)$ and the transfer function $T(f_x, f_y)$ are proportional to each other. Fresnel propagation of periodic structures therefore leads to a self-imaging effect at specific distances given by z_n . For a more general periodic object, the same conclusion can be made by noting that such periodic functions may be expressed as a Fourier series. Further it is easy to see that at distances $z_n/2$ the grating function reverses its contrast and the self-imaging with contrast reversal is observed. This effect is termed as the fractional Talbot effect.

9.5.3 Fractional Fourier Transform

The Fresnel diffraction can also be viewed as a fractional order Fourier transform [13]. It is well known that the Hermite-Gauss functions denoted by

$$\phi_n(x) = \frac{2^{1/4}}{\sqrt{2^n n!}} \exp(-\pi x^2) H_n(\sqrt{2\pi}x), \quad (9.62)$$

are eigenfunctions of the Fourier transform operator. Here $H_n(x)$ denote the Hermite polynomials given by

$$H_n(x) = (-1)^n \exp(x^2) \frac{d^n}{dx^n} \exp(-x^2). \quad (9.63)$$

The Hermite-Gauss functions on Fourier transforming reproduce themselves such that

$$\mathcal{F}\{\phi_n(x)\} = (-i)^n \phi_n(f_x). \quad (9.64)$$

Suppose we represent any finite energy function $g(x)$ as a linear combination of $\phi_n(x)$ as

$$g(x) = \sum_{n=0}^{\infty} a_n \phi_n(x), \quad (9.65)$$

where

$$a_n = \int_{-\infty}^{\infty} dx g(x) \phi_n(x). \quad (9.66)$$

The Fourier transform of $g(x)$ can be obtained using the Fourier transform property of the Hermite-Gauss functions as

$$G(f_x) = \sum_{n=0}^{\infty} a_n (-i)^n \phi_n(f_x) = \int_{-\infty}^{\infty} dx g(x) \sum_{n=0}^{\infty} (-i)^n \phi_n(x) \phi_n(f_x). \quad (9.67)$$

The fractional order Fourier transform of order α may now be introduced as

$$\mathcal{F}^{\alpha}\{g(x)\} = \sum_{n=0}^{\infty} a_n (-i)^{n\alpha} \phi_n(f_x). \quad (9.68)$$

We note that the transform is linear and additive in index α in the sense that two successive operations \mathcal{F}^{α} and \mathcal{F}^{β} are equivalent to $\mathcal{F}^{\alpha+\beta}$. Using Mehler's identity for Hermite functions, the transformation above may be expressed as

$$\mathcal{F}^{\alpha}\{g(x)\} = \int_{-\infty}^{\infty} dx g(x) K(x, f_x, \alpha), \quad (9.69)$$

where

$$K(x, f_x, \alpha) = \sqrt{1 - i \cot(\alpha\pi/2)} \times \\ \exp\{i\pi[(x^2 + f_x^2) \cot(\alpha\pi/2) - 2xf_x \csc(\alpha\pi/2)]\}. \quad (9.70)$$

The form of this kernel is similar to the Fresnel kernel and as $\alpha \rightarrow 1$ we observe that the Fresnel diffraction goes over to 2D Fourier transform that we will see in next section to be equivalent to the Fraunhofer diffraction. Fractional Fourier transform has found application in describing propagation of light through GRIN lenses, in image/signal coding, noise separation, etc.

9.6 Fraunhofer Diffraction

Fraunhofer diffraction model further approximates the Fresnel diffraction relation for even larger propagation distances z . We write the Fresnel diffraction relation in Eq. (9.37) in a slightly modified form:

$$u(x, y, z) = \frac{\exp(ikz) \exp[\frac{i\pi}{\lambda z}(x^2 + y^2)]}{i\lambda z} \times \iint_{S_A} dx' dy' u(x', y', 0) \exp[\frac{i\pi}{\lambda z}(x'^2 + y'^2)] \exp\left[-i\frac{2\pi}{\lambda z}(xx' + yy')\right], \quad (9.71)$$

and observe that the Fresnel diffraction is thus equivalent to a 2D Fourier transform of the input field $u(x', y', 0)$ multiplied by a quadratic phase factor. For distances z and the aperture sizes such that the quadratic phase factor is small:

$$\frac{\pi(x'^2 + y'^2)_{\max}}{\lambda} \ll z, \quad (9.72)$$

the Fresnel diffraction relation may be modified as

$$u(x, y, z) = \frac{\exp(ikz) \exp[\frac{i\pi}{\lambda z}(x^2 + y^2)]}{i\lambda z} \iint_{S_A} dx' dy' u(x', y', 0) \exp\left[-i\frac{2\pi}{\lambda z}(xx' + yy')\right]. \quad (9.73)$$

This relation is known as the Fraunhofer diffraction formula and we observe that the input and output fields in this case are related by a Fourier transform relationship apart from a quadratic phase factor. The coordinates (x, y) in the observation plane may be related to spatial frequency components as

$$f_x = \frac{x}{\lambda z}, f_y = \frac{y}{\lambda z}. \quad (9.74)$$

We observe that the Fourier transform relationship implies that the Fraunhofer diffraction cannot be associated with a convolution. Lateral shift of the input object therefore simply adds an overall phase factor to the diffraction pattern thus keeping its magnitude unchanged. A few examples of Fraunhofer diffraction patterns are shown in Fig. 9.6 where the diffraction pattern intensity is displayed on a log-scale.

It is interesting to observe some characteristic features that are typical of 2D Fourier transforms. For an object like the letter ‘A’, there is a streak perpendicular to each of the three main edge directions. The diffraction pattern for letter ‘O’ shows a circularly symmetric for the same reason. For a more complex object like the ‘cameraman’ image, it is difficult to assign any feature in the transform with the

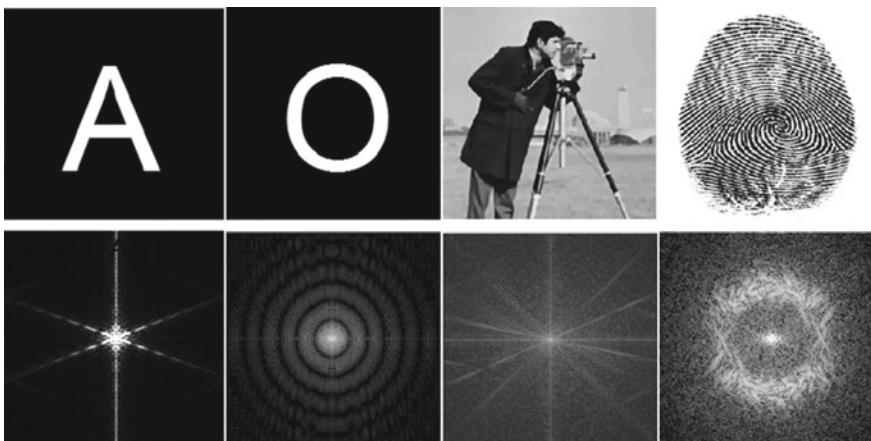


Fig. 9.6 Illustration of several objects and their corresponding Fraunhofer diffraction intensity patterns

grey-scale features in the picture, however, some streaks corresponding to the tripod of the camera are observed over a diffuse background. The energy concentration near zero frequency may also be noted. The fingerprint object is very interesting in that there is a local carrier frequency at every point in the picture which is oriented in different directions. The result is that the diffraction pattern has energy concentration near zero frequency as well as on an annular ring.

We conclude this chapter by noting that the Fresnel and Fraunhofer diffraction approximations described in this chapter are very useful for describing most of the typical laboratory experiments and imaging systems and we will use these models extensively in the coming chapters.

Problems

9.1 The spot of Arago: A plane monochromatic wave $\exp(ikz)$ is incident on a thin opaque circular disk of radius a much larger than wavelength λ , located in $z = 0$ plane and centered at $x = y = 0$. Calculate the on-axis diffraction field intensity as a function of z .

9.2 Show that a plane wave traveling in the direction of the k-vector (k_x, k_y) continues to be a plane wave at all propagation distances.

9.3 Pseudo-Eigenfunctions of Fresnel transform The set of angular prolate spheroidal functions of order zero $\{\phi_n(x)\}$ have the self- finite Fourier transform property:

$$\int_{-1}^1 dx \phi_n(x) \exp(-i2\pi xx') = \alpha_n \phi_n(x'),$$

where α_n 's are constants. Show that a prolate function multiplied by a quadratic phase function on Fresnel transform gives the same prolate function multiplied by another quadratic phase function. For simplicity you may assume Fresnel diffraction over an aperture of size $x : (-1, 1)$.

9.4 Starting with the paraxial wave equation and modifying the steps used to obtain the transport of intensity equation, find the transport of phase equation for the quantity $\partial\theta/\partial z$:

$$k^2 I^2 \frac{\partial\theta}{\partial z} = \frac{1}{2} I \nabla^2 I - \frac{1}{4} (\nabla I)^2 - I^2 (\nabla\theta)^2 + 2k^2 I^2.$$

9.5 Show that a more accurate expression for the Talbot (or self-imaging) distance corresponding to the grating is given by

$$z_T = \frac{\lambda}{1 - \sqrt{1 - \lambda^2/a^2}}.$$

9.6 An image is printed on a transparency using a printer with resolution M dots per mm and the transparency is used as an object in a F-Lens-F Fourier transform setup. The transparency is illuminated using a plane wave of wavelength λ traveling along the optical axis of the system. Estimate the size of an array sensor (in terms of number of pixels) that is required to observe the highest detail in the object transparency. Assume square pixels of dimension $p \mu\text{m}$.

References

1. M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, (Ed. 7) (Cambridge University Press, 1999)
2. K. Iizuka, *Elements of Photonics*, vol. 1 and 2 (Wiley, 2002)
3. E. Hecht, *Optics*, edn 5 (Pearson Education, 2017)
4. J.W. Goodman, *Introduction to Fourier Optics*, (Ed. 2) (McGraw-Hill, 1996)
5. D.J. Griffiths, *Introduction to Electrodynamics*, (Ed. 4), (Pearson Education, 2014)
6. L. Mandel, E. Wolf, *Coherence and Quantum Optics* (Cambridge University Press, 1995)
7. J.A. Ratcliffe, Some aspects of diffraction theory and their application to the ionosphere. *Rep. Prog. Phys.* **19**, 188 (1956)
8. W.R. Smythe, The double current sheet in diffraction. *Phys. Rev.* **72**, 1066 (1947)
9. D.G. Voelz, M.C. Roggemann, Digital simulation of scalar optical diffraction: revisiting chirp function sampling criteria and consequenses. *Appl. Opt.* **48**, 6132–6142 (2009)
10. M.R. Teague, Deterministic phase retrieval: a Green's function solution. *J. Opt. Soc. Am.* **73**, 1434–1441 (1983)
11. N. Pandey, A. Ghosh, K. Khare, Two-dimensional phase unwrapping using the transport of intensity equation. *Appl. Opt.* **55**, 2418–2425 (2016)
12. W.D. Montgomery, Self-imaging objects of infinite aperture. *J. Opt. Soc. Am.* **57**, 772–778 (1967)
13. H.M. Ozaktas, Z. Zalevsky, M. Alper Kutay, *The Fractional Fourier Transform: With Applications in Optics and Signal Processing* (Wiley 2001)

Chapter 10

Coherence of Light Fields



The light wavefields in our discussion so far have been treated as deterministic. In particular, we have solved the problems on wave propagation and diffraction starting with Maxwell equations which describe the relation between the electric and magnetic fields associated with the light waves. The generation of light waves in a realistic light source occurs from electronic transitions in a large number of atoms may be radiating independently. As a result, the description of light fields is necessarily statistical in nature and the quantities observable in the laboratory can only be explained by means of expectation values. As we shall see later in this chapter, due to the effects of source bandwidth and propagation of light waves, there exist correlations between fields at two space-time points (\mathbf{r}_1, t_1) and (\mathbf{r}_2, t_2) that are central to understanding the concepts like spatial and temporal coherence. Additionally, we will describe the higher order correlations and propagation of correlations that brings out interesting implications. In Chap. 7 we have developed several key concepts in theory of random processes that will be directly utilized in the analysis presented in this chapter.

10.1 Spatial and Temporal Coherence

We begin by defining important properties of optical fields, viz. the spatial and the temporal coherence, which will be understood in terms of the correlation function introduced in Eq. (7.46). We will introduce these concepts here for a single scalar component of the electric field of light waves, however, these concepts can be generalized to include vectorial fields in a straightforward manner [1, 2]. Given an arbitrary source of light we will define the complex degree of coherence to be the correlation function:

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle E^*(\mathbf{r}_1, t + \tau) E(\mathbf{r}_2, t) \rangle. \quad (10.1)$$

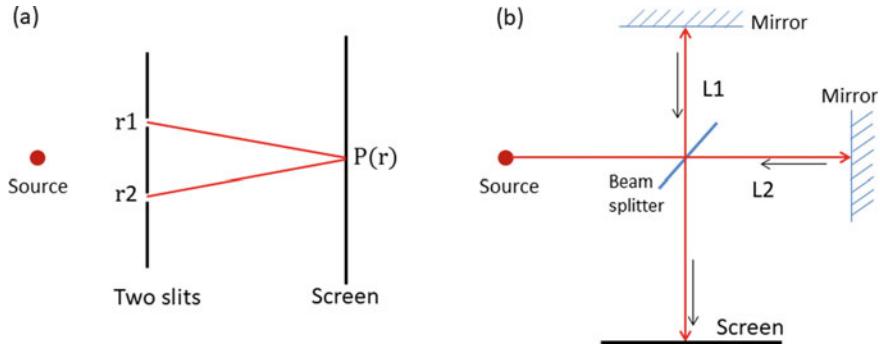


Fig. 10.1 Typical experiments **a** Young's double slit and **b** Michelson interferometer illustrating spatial and temporal coherence respectively

Here $E(\mathbf{r}, t)$ denotes the scalar electric field due to the source which we assume to be a stationary random process in time. We note that the observed intensity of light is proportional to the equi-space and equi-time correlation function $\Gamma(\mathbf{r}, \mathbf{r}, 0)$. The correlation function is typically normalized as follows:

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)}{\sqrt{\Gamma(\mathbf{r}_1, \mathbf{r}_1, 0) \Gamma(\mathbf{r}_2, \mathbf{r}_2, 0)}}, \quad (10.2)$$

so that, the degree of coherence γ satisfies the property

$$0 \leq |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| \leq 1. \quad (10.3)$$

The equi-time degree of coherence $\gamma(\mathbf{r}_1, \mathbf{r}_2, 0)$ is commonly associated with the spatial coherence of the source while the equi-space degree of coherence $\gamma(\mathbf{r}_1, \mathbf{r}_1, \tau)$ is associated with the temporal coherence of the source. Two typical optical experiments—Young's double slit experiment and the Michelson interferometer are shown in Fig. 10.1. In Young's double slit experiment, any near-axis interference fringes observed on the screen can be associated with the equi-time correlation between the field at the two slits. In the Michelson interferometer on the other hand, the two arms of the interferometer can be varied in length to study the field correlation at the same point in space but delayed in time. We will express these relations more formally in the following section.

10.1.1 Interference Law

In reference to Fig. 10.1a, we write the intensity observed at point P due to two point sources at \mathbf{r}_1 and \mathbf{r}_2 as

$$I(P) = \langle |k_1 E(\mathbf{r}_1, t - R_1/c) + k_2 E(\mathbf{r}_2, t - R_2/c)|^2 \rangle. \quad (10.4)$$

Here k_1, k_2 are geometrical factors and c is the speed of light in vacuum. The field at point P at time t is sum of the fields from the two point sources with the effect of time delay included. The distances R_1 and R_2 are defined as: $R_1 = |\mathbf{r}_1 - \mathbf{r}|$ and $R_2 = |\mathbf{r}_2 - \mathbf{r}|$. Assuming stationarity condition and defining $I_n = |k_n|^2 \langle |E(\mathbf{r}_n, t - R_n/c)|^2 \rangle$ for $n = 1, 2$ gives

$$I(P) = I_1 + I_2 + 2\operatorname{Re}\{k_1^* k_2 \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau = |\mathbf{r}_1 - \mathbf{r}_2|/c)\}. \quad (10.5)$$

When the observation screen is located sufficiently far away from the two pinholes, as per the diffraction formulation developed in the previous chapter, the constants k_1 and k_2 are purely imaginary and equal to

$$k_{1,2} \approx \frac{-i}{\bar{\lambda} R_{1,2}}. \quad (10.6)$$

Here $\bar{\lambda}$ is the mean wavelength of the spectrum of the incident light (assuming a narrowband source). In terms of the normalized coherence function $\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ the interference law can therefore be written as

$$\begin{aligned} I(P) &= I_1 + I_2 + 2\sqrt{I_1 I_2} \operatorname{Re}[\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)] \\ &= I_1 + I_2 + 2\sqrt{I_1 I_2} |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| \cos[\arg(\gamma)]. \end{aligned} \quad (10.7)$$

For a narrowband source, the fields $E(\mathbf{r}_{1,2}, t)$ have a form:

$$E(\mathbf{r}_{1,2}, t) = E_0(\mathbf{r}_{1,2}, t) \exp(i2\pi\nu_0 t), \quad (10.8)$$

where $E_0(\mathbf{r}_{1,2}, t)$ represents the slowly varying complex envelope and ν_0 is the central frequency of the source. The complex degree of coherence in such case may therefore be expressed as

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| \exp[i2\pi\nu_0\tau - i\alpha(\mathbf{r}_1, \mathbf{r}_2, \tau)]. \quad (10.9)$$

Here $\alpha(\mathbf{r}_1, \mathbf{r}_2, \tau)$ represents the slowly varying phase (due to the correlation of the envelope) and the term $2\pi\nu_0\tau$ with $\tau = (\mathbf{r}_1 - \mathbf{r}_2)/c$ is responsible for the observation of typical interference fringes. The fringe contrast or visibility near the observation point P is given by

$$V = \frac{I(P)_{\max} - I(P)_{\min}}{I(P)_{\max} + I(P)_{\min}} = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2} |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)|. \quad (10.10)$$

Interference fringes with good contrast can be observed over the region where the degree of coherence is significant. In Young's experiment setup, the quantity τ is equal to zero at points on the screen that are symmetric with respect to the two pinholes/slits. The fringe contrast is therefore maximum at the center of the pattern.

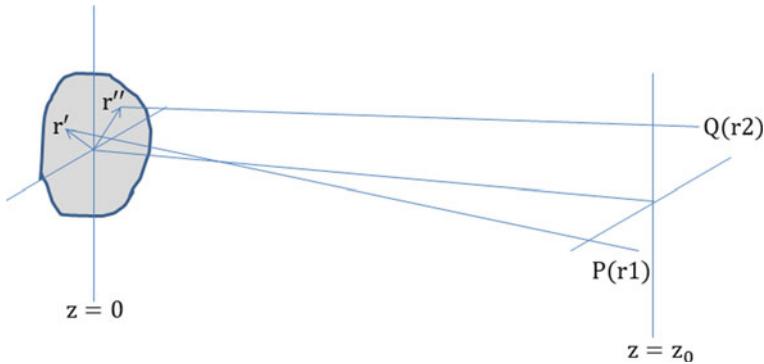


Fig. 10.2 Geometry for the van Cittert and Zernike theorem

For any source with finite bandwidth, the fringe contrast will slowly decrease as we move away from the center of the pattern. The spatial and temporal coherence are closely linked through the degree of coherence $\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ and can be conveniently separated in two experimental settings.

In Fig. 10.1a, b we observe two interferometric arrangements. In Fig. 10.1a, we observe that when interference fringes are observed near on-axis locations of the screen, the spatial coherence is of more importance. In the Michelson interferometer in Fig. 10.1b, the interference is observed between time-delayed version of the field at the same spatial location. As a result, the temporal coherence plays important role in this configuration.

10.2 van Cittert and Zernike Theorem

In this section, we will describe an important result concerning the propagation of the correlation function. We will consider a source in plane $z = 0$ and evaluate the equi-time correlation function at a plane $z = z_0$. If the plane $z = z_0$ is a number of wavelengths away from the source plane, we may simply express the field at point P due to a point source at \mathbf{r}' as a spherical wave. The equi-time correlation function of fields at points $P(\mathbf{r}_1)$ and $Q(\mathbf{r}_2)$ is given by [1]:

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, 0) = \iint d^2\mathbf{r}' d^2\mathbf{r}'' \Gamma(\mathbf{r}', \mathbf{r}'', 0) \frac{k^2 e^{-ik|\mathbf{r}_1 - \mathbf{r}'|} e^{ik|\mathbf{r}_2 - \mathbf{r}''|}}{(2\pi)^2 |\mathbf{r}_1 - \mathbf{r}'| |\mathbf{r}_2 - \mathbf{r}''|}. \quad (10.11)$$

When the observation plane is in the far field relative to the source, several approximations may be made in accordance with the Fraunhofer approximation. For example, $|\mathbf{r}_1| \gg |\mathbf{r}'|$ and $|\mathbf{r}_2| \gg |\mathbf{r}''|$, so that

$$|\mathbf{r}_1 - \mathbf{r}'| \approx |\mathbf{r}_1| - \frac{\mathbf{r}_1 \cdot \mathbf{r}'}{|\mathbf{r}_1|}, \quad \text{and} \quad |\mathbf{r}_2 - \mathbf{r}''| \approx |\mathbf{r}_2| - \frac{\mathbf{r}_2 \cdot \mathbf{r}''}{|\mathbf{r}_2|}. \quad (10.12)$$

Also the distances in the denominator of the integrand in Eq. (10.11) may be replaced by z_0 . We will consider the important case of spatially incoherent source for which

$$\Gamma(\mathbf{r}', \mathbf{r}'', 0) = I_0(\mathbf{r}') \delta(\mathbf{r}' - \mathbf{r}''). \quad (10.13)$$

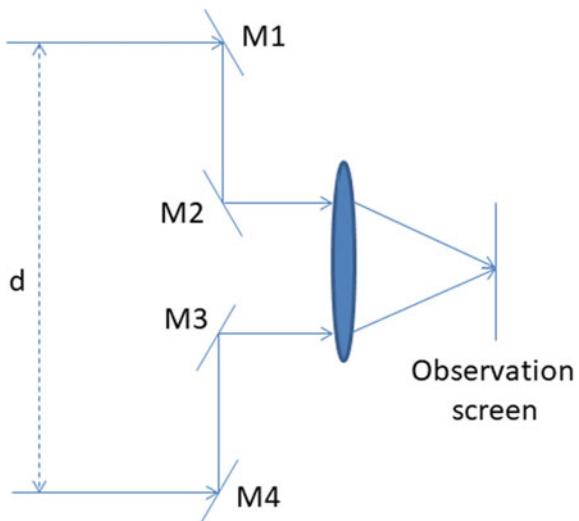
With these approximations the correlation function in the observation plane as in Eq. (10.11) may be expressed as

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, 0) \approx \frac{k^2 e^{-ik(|\mathbf{r}_1| - |\mathbf{r}_2|)}}{(2\pi z_0)^2} \int d^2 \mathbf{r}' I_0(\mathbf{r}') \exp[-ik\mathbf{r}' \cdot (\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2)]. \quad (10.14)$$

This is a statement of the far-zone form of the van Cittert and Zernike theorem. In the above equation, $\hat{\mathbf{r}}_1$ and $\hat{\mathbf{r}}_2$ denote unit vectors along \mathbf{r}_1 and \mathbf{r}_2 respectively. *We note an important result that the far-zone correlation function for a spatially incoherent source is proportional to the 2D Fourier transform of the source intensity pattern [1].* The correlation is to be understood in the following manner. Suppose two pinholes were placed at points $P(\mathbf{r}_1)$ and $Q(\mathbf{r}_2)$, then the interference fringes on a screen downstream will have visibility or contrast proportional to the magnitude $|\Gamma(\mathbf{r}_1, \mathbf{r}_2, 0)|$. Further the Fourier transform relation implies that the extent where $|\Gamma(\mathbf{r}_1, \mathbf{r}_2, 0)|$ is significant is inversely proportional to the angular size of the source as seen in plane $z = z_0$. This is the reason why light field from a far away star is spatially coherent over a large distance (several meters) at the earth's surface. This idea was effectively utilized by Michelson in designing the stellar interferometer instrument which is shown schematically in Fig. 10.3.

As illustrated in Fig. 10.3, the starlight from two collection mirrors $M1, M4$ separated by a baseline d is brought together and made to interfere on the observation screen. The visibility of fringes is observed as the baseline d is changed. Assuming a circular disk shape for the star, its angular diameter can be related to the baseline d for which the fringes show zero contrast (washed out fringe pattern). Michelson was able to measure the angular diameter of the star Betelgeuse to be 0.02 arc-seconds using a baseline d approximately equal to 6 m which is much larger than a typical telescope diameter. This principle is also used in radio-astronomical telescopes that involve large antenna arrays spread over several kilometers. The correlation of signals from two antennas provides a sample in the Fourier space corresponding to the star's intensity pattern $I(\mathbf{r}')$ as per Eq. (10.14). Clearly with a limited number of radio antennas, the number of Fourier space samples available is usually much less than the number of pixels in the desired image. Ideas like compressive sensing as discussed in Chap. 6 therefore find important application in this limited data Fourier inversion problem.

Fig. 10.3 Michelson stellar interferometer



The **imaging of a supermassive blackhole**, M-87 located at ~ 55 million light years away from the Earth is a recent illustration of the use of van Cittert and Zernike theorem [3]. To image M-87 from our planet, at radio wavelength 1.3 mm, having an angular size of $15\mu\text{arcsec}$ one needs a telescope of aperture size larger than the diameter of Earth. However, very long base line interferometry(VLBI) technique can achieve this resolution beyond the diffraction limit of a single telescope. In VLBI, the radio waves from the source are captured by an array of radio telescopes, each recording the received signals independently. The ground-based telescopes are spread across the globe, separated by large distances (long baselines). The signals from the telescopes are pair-wise correlated and time averaged. From van Cittert and Zernike theorem we know that the average correlation of measured signals from two points is equivalent to the Fourier transform of the intensity distribution of the source. These time averaged correlations are subjected to computational imaging techniques for the final reconstruction of the image of source. It is to be noted that the spatial frequency captured increases with the increase in baseline distance of a pair of telescope and hence the angular resolution will increase. Using 6–8 telescopes as the earth rotates, limited Fourier transform measurements are available. The whole VLBI system which is essentially a signal correlator and the associated computational imaging techniques collectively make the Event Horizon Telescope (EHT) which is a Earth-sized computational telescope.

10.3 Space-frequency Representation of the Coherence Function

From the definition of the coherence function in Eq. (10.1) it is clear that the coherence function satisfies two wave equations in the form:

$$\left(\nabla_i^2 - \frac{1}{c^2} \frac{\partial^2}{\partial \tau^2} \right) \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = 0 \quad \text{for } i = 1, 2. \quad (10.15)$$

Here the operator ∇_i for $i = 1, 2$ is with respect to the coordinates \mathbf{r}_1 and \mathbf{r}_2 . These relations are known as Wolf equations. The behavior of the coherence function can also be studied in the frequency domain using the cross-spectral density function defined through a Fourier transform relationship as

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \int_{-\infty}^{\infty} d\tau \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) \exp(-i2\pi\nu\tau). \quad (10.16)$$

The cross-spectral density is typically normalized as

$$\mu(\mathbf{r}_1, \mathbf{r}_2, \nu) = \frac{W(\mathbf{r}_1, \mathbf{r}_2, \nu)}{\sqrt{S(\mathbf{r}_1, \nu) S(\mathbf{r}_2, \nu)}}, \quad (10.17)$$

where $S(\mathbf{r}_i, \nu) = W(\mathbf{r}_i, \mathbf{r}_i, \nu)$ for $i = 1, 2$ represent the spectral density (or spectrum). The normalized quantity $\mu(\mathbf{r}_1, \mathbf{r}_2, \nu)$ is referred to as the spectral degree of coherence and has the property:

$$0 \leq |\mu(\mathbf{r}_1, \mathbf{r}_2, \nu)| \leq 1. \quad (10.18)$$

It was noted by [12] that the cross-spectral density can be thought of as a correlation function obtained from an ensemble of monochromatic realizations of the field. The cross-spectral density is Hermitian with respect to variables \mathbf{r}_1 and \mathbf{r}_2 and non-negative definite. Further assuming that the cross-spectral density is square integrable over domain of interest, we may represent it as a Mercer expansion:

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \sum_n W_n(\mathbf{r}_1, \mathbf{r}_2, \nu) = \sum_n \lambda_n(\nu) \phi_n^*(\mathbf{r}_1, \nu) \phi_n(\mathbf{r}_2, \nu). \quad (10.19)$$

This is known as the coherent mode representation of the cross-spectral density function. The parameters $\lambda_n(\nu)$ are the eigenvalues (positive due to non-negative definite property of $W(\mathbf{r}_1, \mathbf{r}_2, \nu)$) in the homogeneous integral equation:

$$\int d^2\mathbf{r}_1 W(\mathbf{r}_1, \mathbf{r}_2, \nu) \phi_n(\mathbf{r}_1, \nu) = \lambda_n(\nu) \phi_n(\mathbf{r}_2, \nu). \quad (10.20)$$

We may now introduce an ensemble of fields in frequency domain (in a manner analogous to the Karhunen-Loeve expansion):

$$u(\mathbf{r}, \nu) = \sum_n a_n(\nu) \phi_n(\mathbf{r}_1, \nu). \quad (10.21)$$

where a_n are random variables satisfying

$$\langle a_n^*(\nu) a_m(\nu) \rangle = \lambda_n(\nu) \delta_{m,n}. \quad (10.22)$$

The individual realizations $u(\mathbf{r}, \nu)$ satisfy the Helmholtz equation and further lead to the representation

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \langle u^*(\mathbf{r}_1, \nu) u(\mathbf{r}_2, \nu) \rangle_\nu, \quad (10.23)$$

where $\langle \dots \rangle_\nu$ represents an ensemble in the frequency domain. The individual terms in the summation in Eq. (10.19) can be seen to be fully spatially coherent by evaluating the corresponding normalized cross-spectral densities.

$$\mu_n(\mathbf{r}_1, \mathbf{r}_2, \nu) = \frac{W_n(\mathbf{r}_1, \mathbf{r}_2, \nu)}{\sqrt{W_n(\mathbf{r}_1, \mathbf{r}_1, \nu) W_n(\mathbf{r}_2, \mathbf{r}_2, \nu)}} = \frac{\phi^*(\mathbf{r}_1, \nu) \phi_n(\mathbf{r}_2, \nu)}{|\phi^*(\mathbf{r}_1, \nu)| |\phi_n(\mathbf{r}_2, \nu)|}. \quad (10.24)$$

As a result $|\mu_n(\mathbf{r}_1, \mathbf{r}_2, \nu)|$ is identically equal to 1. *The partially coherent monochromatic field $u(\mathbf{r}, \nu)$ as expressed in Eq. (10.21) can therefore be thought of as an incoherent sum of fully spatially coherent modes.* The space-frequency representation of the cross-spectral density has found extensive usage in the description of a number of interesting phenomena associated with partially coherent light fields.

Partially coherent fields as described above consist of an incoherent combination of fully spatially coherent modes $\phi_n(\mathbf{r}, \nu)$. The coherent modes typically get more oscillatory with their order n . Partially coherent field is often considered to be more robust to propagation in random media such as atmospheric turbulence [6]. This is because the individual coherent modes evolve in diverse manner as they propagate in the random medium. Since the contributions of individual modes add incoherently, after passing through the random medium, the overall beam profile has reduced speckle (or scintillations). Design of robust beams in this manner has a number of applications in free space communication and defence systems.

10.4 Intensity Interferometry: Hanbury Brown and Twiss Effect

We will now turn to a somewhat surprising interference effect—the intensity interference. So far in the discussions in this chapter, we have talked about interference in terms of the second-order field correlations. However, there is interesting infor-

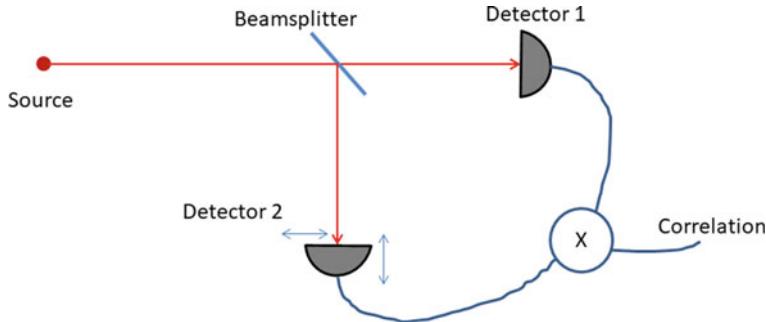


Fig. 10.4 Schematic diagram of Hanbury Brown and Twiss intensity interferometer

mation to be gained from fourth-order or intensity correlations as well. We will be particularly interested in thermal sources. Due to large number of atoms radiating independently in a thermal source, the probability distribution for the real and imaginary parts of the field of light from a thermal source may be assumed to have Gaussian distribution.

$$p(E_r, E_i) = \frac{1}{\pi E_0^2} \exp\left[-\frac{E_r^2 + E_i^2}{E_0^2}\right]. \quad (10.25)$$

Here E_r and E_i denote the real and imaginary parts of the light field respectively. Using the complex form of Gaussian moment theorem (7.38), we may calculate the intensity correlation at the two space-time points as follows:

$$\begin{aligned} & \langle E^*(\mathbf{r}_1, t_1) E(\mathbf{r}_1, t_1) E^*(\mathbf{r}_2, t_2) E(\mathbf{r}_2, t_2) \rangle \\ & \langle I(\mathbf{r}_1) \rangle \langle I(\mathbf{r}_2) \rangle + |\Gamma(\mathbf{r}_1, \mathbf{r}_2, t_1 - t_2)|^2. \end{aligned} \quad (10.26)$$

We observe that in the space-time region where $\Gamma(\mathbf{r}_1, \mathbf{r}_2, t_1 - t_2)$ is significant, the fourth-order correlation is in excess of $\langle I(\mathbf{r}_1) \rangle \langle I(\mathbf{r}_2) \rangle$. This effect was first described by R. Hanbury Brown and R. Quintin Twiss. Their experiments were first performed with radio antennas and later with optical sources. The optical version of the experiment is shown schematically in Fig. 10.4.

The radiation from a light source is split into two parts and the photo-currents from the two detectors are correlated. Detector 2 can be translated in transverse and longitudinal directions in order to study intensity correlations wherever $\Gamma(\mathbf{r}_1, \mathbf{r}_2, t_1 - t_2)$ is significant so that excess correlation in intensity can be observed. This effect of intensity correlation is surprising as traditionally interference is understood only as a second-order correlation effect. The excess intensity correlation effect is a manifestation of statistical property of thermal or Gaussian light fields. Such excess correlation is not observed for light from well-stabilized single mode laser thus providing a possibility of distinguishing light sources based on their intensity correlation properties. The Hanbury Brown and Twiss experiment was largely responsible for emergence

of what we now call ‘Quantum Optics’. We will discuss more about this effect later in the book in the context of what is known as the single pixel ghost imaging system.

10.5 Photon Counting Formula

In the previous section, we assumed that the photo-current in the detector is proportional to the average light intensity falling on the detector. When photon counting statistics for a given state of light field is of interest, the statistical distribution of number of photons can be determined correctly by considering two statistical processes, viz. the fluctuations in the light intensity and the statistical nature of the detection process itself. Suppose that a beam of light is incident on a photon counting detector with mean intensity $I(x, y, t)$ (in units of number of photons). Then the integrated intensity over the detector area is given by

$$W = \eta \int_0^T dt \int_A dx dy I(x, y, t). \quad (10.27)$$

Here η represents the detector quantum efficiency. When the light field is fluctuating with some statistical distribution $p(W)$ associated with the integrated intensity W , the distribution $p(n)$ of number of photon counts registered in series of trials each with integration time T is given by

$$p(n) = \int_0^\infty dW p(n|W) p(W). \quad (10.28)$$

The probability is determined as sum of all conditional probabilities so that the statistical fluctuation in the integrated intensity is accounted for. The conditional probability $p(n|W)$ is the standard Poisson distribution.

$$p(n|W) = \frac{\exp(-W) W^n}{n!}. \quad (10.29)$$

The photon counting distribution is therefore what is known as the Poisson transform of the distribution $p(W)$:

$$p(n) = \int_0^\infty dW \frac{\exp(-W) W^n}{n!} p(W). \quad (10.30)$$

This is the well-known photon counting formula due to Mandel. Some special cases of this formula are of interest. While discussing the statistics of light from a thermal source, we assumed a complex Gaussian distribution for the electric field. The corresponding intensity has an exponential distribution:

$$p(I) = \frac{1}{\langle I \rangle} \exp\left(-\frac{I}{\langle I \rangle}\right). \quad (10.31)$$

Here $\langle I \rangle$ denotes the mean intensity. If the measurement time T used in the definition of integrated intensity W is much smaller than the coherence time of the source, the intensity $I(t)$ is essentially a constant during the measurement interval. Further assuming a small-area detector, we can use the following approximation:

$$W \approx \eta A_{det} I(t) T, \quad (10.32)$$

where A_{det} is the detector area. Since W and I are now related by a proportionality constant, the probability density for W is also an exponential distribution.

$$p(W) = \frac{1}{\langle W \rangle} \exp\left(-\frac{W}{\langle W \rangle}\right). \quad (10.33)$$

The corresponding photon counting distribution obtained by using the Mandel formula is the geometric (or Bose-Einstein) distribution.

$$p(n) = \frac{\langle W \rangle^n}{(1 + \langle W \rangle)^{n+1}}. \quad (10.34)$$

The variance corresponding to this distribution is

$$(\Delta n)^2 = \langle W \rangle(1 + \langle W \rangle). \quad (10.35)$$

Another special case is that of well-stabilized single mode laser where the distribution $p(W)$ can be approximated as

$$p(W) = \delta(W - \langle W \rangle). \quad (10.36)$$

The resultant photon count distribution is a Poisson distribution.

$$p(n) = \frac{\exp(-\langle W \rangle)\langle W \rangle^n}{n!} \quad (10.37)$$

The variance for this distribution is given by

$$(\Delta n)^2 = \langle W \rangle. \quad (10.38)$$

The excess fluctuation in the case of thermal light over Poisson distribution is often referred to as the bunching of photons, since the photon counts are seen to be more clustered together in time (which is a consequence of the excess variance). The photon counting distribution therefore provides an interesting way to distinguish light from different sources even when other properties of the light beam such as intensity,

polarization, spectrum are same. For completeness we mention here that there are certain states of light that show sub-Poissonian nature of fluctuation. Description of such states requires a full quantum treatment.

10.6 Speckle Phenomenon

Speckle is a complex interference phenomenon commonly observed in the laboratory when a coherent laser beam scatters off a rough surface with surface features that induce almost uniformly random phase shifts in the range $[0, 2\pi]$ to the light beam at different locations. The observed scattered field at any point in space is therefore a superposition of a large number of randomly oriented phasors. Figure 10.5 shows a laser beam focused on a rough diffuser. The resultant diffraction pattern has blob-like bright regions which are known as speckles [4, 5]. The complex speckle field has interesting structure in that the zeros of real and imaginary parts are along continuous curves in the observation plane. The zeros of intensity pattern are therefore typically located at isolated points as shown in Fig. 10.6 [7]. If the wavelength of illuminating laser is tuned, the speckle pattern slowly evolves and shows de-correlation with initial speckle pattern [8]. So tuning of wavelength is in some sense equivalent to moving the diffuser. When the laser spot size on the rough surface has a diameter D , then on a plane which is at a distance z from the scattering surface, the average size of individual speckles is same as the diffraction-limited spot which is of the order of

$$d_{\perp} \approx \frac{\lambda z}{D}. \quad (10.39)$$

Individual speckles are 3D structures like jelly beans and the average longitudinal size of speckles is given by

$$d_{||} \approx \lambda \left(\frac{z}{D} \right)^2. \quad (10.40)$$

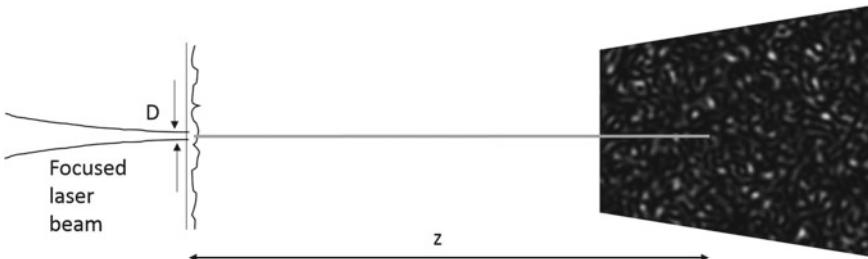


Fig. 10.5 Schematic diagram showing formation of speckle pattern

The lateral and longitudinal size is thus effectively of the order $\lambda/(NA)$ and $\lambda/(NA)^2$ with the numerical aperture subtended by the laser spot at the observation surface given by $NA = D/(2z)$. *The number of speckles in a single image frame is thus seen to be of the same order as the space-bandwidth product of the system.* The space-bandwidth product is associated with the number of resolvable elements each of which is of the size of individual speckle. Zeroes of a speckle pattern occur where both the real and imaginary parts of the speckle field are zero. The zeroes are typically isolated points, and their number again equals the space-bandwidth product. *The phase of speckle field near an intensity zero is also known to have a vortex-like structure* and this fact has been used recently in studies on small motion displacement detection on nanometric scale [7, 9, 10]. Speckle has a non-deterministic structure due to the randomness in the height distribution of the rough surface. In particular we will assume that for performing ensemble averages, the diffuser is being moved along the transverse direction so that the total field at a given observation point arises due to a different set of scattering centers in each realization of the speckle pattern. As the resultant field is a sum of large number of contributions from various scattering centers with phases distributed uniformly in $[0, 2\pi]$, the resultant field obeys complex Gaussian statistics (as per the central limit theorem). The probability distribution for the real and imaginary parts of scalar E -field denoted by E_r and E_i respectively is thus given by

$$p(E_{r,i}) = \frac{1}{\sqrt{\pi}E_0} \exp\left(-\frac{E_{r,i}^2}{E_0^2}\right). \quad (10.41)$$

The intensity of the speckle field $I = E_r^2 + E_i^2$ therefore follows an exponential statistics given by

$$p(I) = \frac{1}{I_0} \exp\left(-\frac{I}{I_0}\right), \quad (10.42)$$

where $I_0 = |E_0|^2$ is the mean intensity. The spatial statistics of a speckle is thus similar to the temporal statistics of thermal light. If the speckle pattern is generated by unpolarized laser, the resultant speckle intensity distribution may be obtained by convolution of two exponential distributions which gives the following probability distribution:

$$p(I) = \int_0^I dI' \frac{1}{I_0^2} \exp\left[-\frac{(I - I')}{I_0}\right] \exp\left(-\frac{I'}{I_0}\right) = \frac{I}{I_0} \exp\left(-\frac{I}{I_0}\right), \quad (10.43)$$

which is now a Gamma distribution. The limits of integration are selected as above since the exponential distribution is valid only in the region where the intensity is non-negative. Note that by adding two polarized speckles, the probability distribution has changed drastically near $I = 0$. Care must be taken if the above statistical distributions are to be applied to a single speckle pattern image as in Fig. 10.6a where individual speckles are spread over large number of pixels of an array detector. The

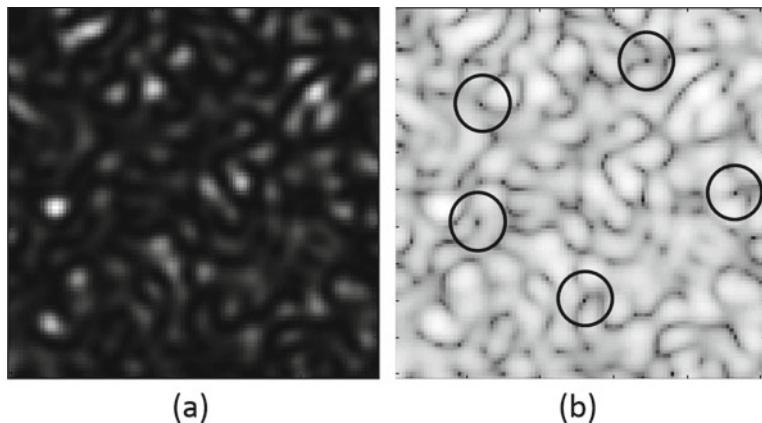


Fig. 10.6 **a** A typical speckle intensity pattern, **b** The same pattern shown on log-scale with circles showing some of the isolated intensity zeros

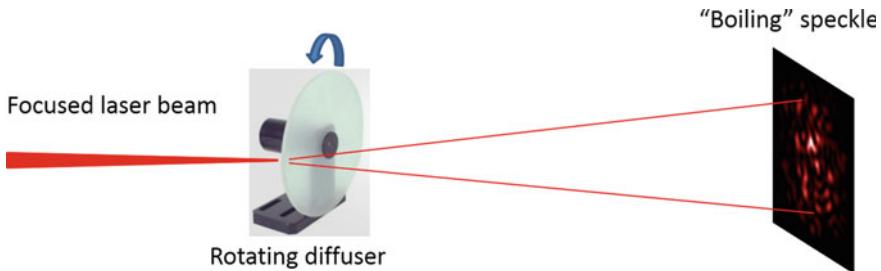


Fig. 10.7 Pseudo-thermal source using a rotating diffuser

distribution of intensity values of all pixels in such a speckle pattern record is certainly not exponential, as the pixel intensity values within a single speckle are correlated. The statistics will turn out to be exponential only if intensity values at a set of pixels that are mutually separated by a distance more than individual speckle size are selected for histogram purpose. The speckle size is also assumed to be extending over at least several pixels so as to obtain ‘point’ samples of the intensity distribution. An important laboratory tool in the study of speckle is the moving (or rotating) glass diffuser which on illumination by a laser spot produces a thermal-like source statistics although the native illumination is from a laser source [11]. As the glass diffuser translates, a new set of scattering centers keeps appearing in the laser spot. The speckle pattern at an observation point thus keeps fluctuating on temporal scale of the time required for a point on the diffuser to cross the laser spot. The coherence time of such a source is thus given by:

$$\tau_c = a/v, \quad (10.44)$$

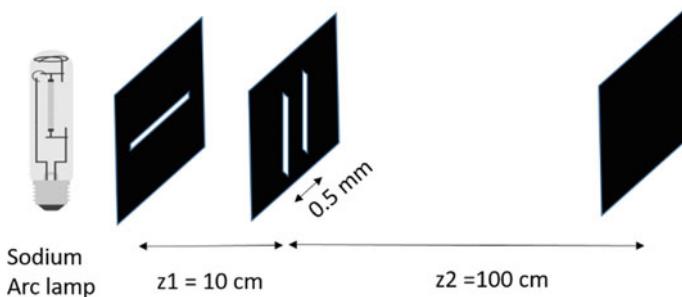


Fig. 10.8 Double slit experiment with orthogonal illuminating slit

where a is the laser spot size on the diffuser and v is the diffuser velocity at the laser spot position. The coherence time of the source is guided by diffuser motion and the transverse spatial coherence length near the observation point is same as the average size of the speckles.

Such source can give rise to excess intensity correlations as in Hanbury Brown Twiss experiment and therefore finds use in correlation imaging experiments. Further the rotating diffuser source is much brighter compared to a typical narrowband thermal source obtained by filtering a small spectral window from a thermal lamp. Such a source may be used for averaging out the undesirable effect of speckles in imaging systems by making the coherence time τ_c much smaller than the exposure time. The rotating diffuser is also a starting point in realizing a number of partially coherent sources, e.g. the Gaussian Schell-model source and some more exotic variants like the partially coherent vortex or polarization singular sources.

Problems

10.1 The setup shown has three slits of size $5 \text{ cm} \times 5 \mu\text{m}$ each as shown. State with appropriate reasoning if you think that stationary fringes will be observed on the screen placed at distance $z_2 = 100 \text{ cm}$.

10.2 A car located 1 km away has two head-lamps of size $5 \text{ cm} \times 5 \text{ cm}$ separated by 1 m. A double slit experiment is to be arranged using the light from the two lamps. Draw a sketch of the setup and with appropriate reasoning estimate the largest distance between the two pinholes for which you will observe fringes with contrast greater than 0.5. Does the setup need to use a narrowband filter? Should the filter be placed right in front of the car lamps or in your setup at 1 km from the car? If the car moves towards you, will you observe any change in the interference pattern?

10.3 Spectral interference: In analyzing Young's double slit setup, one may examine the correlation between $E^*(\mathbf{r}, t)$ and $E(\mathbf{r}, t + \tau)$ (see notation in Sect. 10.1.1). Show that the correlation function has the form:

$$\Gamma(\mathbf{r}, \mathbf{r}, \tau) = |a_1|^2 \Gamma(\mathbf{r}_1, \mathbf{r}_1, \tau) + |a_2|^2 \Gamma(\mathbf{r}_2, \mathbf{r}_2, \tau) + 2 \operatorname{Re}\{a_1^* a_2 \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau + |R_1 - R_2|/c)\}.$$

Here $R_{1,2}$ represent distances between the two pinholes and the observation point on the screen. Further using the Wiener-Khintchine theorem, determine the spectral density at point \mathbf{r} in terms of the spectral densities at \mathbf{r}_1 and \mathbf{r}_2 .

10.4 Two pinholes in a Young's experiment are illuminated by a plane collimated beam from (1) a stabilized single mode laser or (2) an LED with the same central wavelength as the laser. A spectrometer collects light from a small region at a distance z from the pinhole plane. Explain (qualitatively with sketches) the difference in observed spectra at different points on the screen for the two illuminations.

10.5 In Young's double slit arrangement, show that on the plane bisecting the line joining the two pinholes and perpendicular to that line, the quantity $|\mu(r_1, r_2, v)| = 1$ irrespective of the state of coherence of the field at the two pinholes.

References

1. L. Mandel, E. Wolf, *Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995)
2. J.W. Goodman, *Statistical Optics* (Wiley-Interscience, 2000)
3. The Event Horizon Telescope Collaboration et al., First M87 event horizon telescope results. IV. Imaging the central supermassive black hole. *Astrophys. J. Lett.* **875** (2019)
4. J.C. Dainty (Ed.), *Laser speckle and related phenomena* Top. Appl. Phys. 9, Springer (1975)
5. J.W. Goodman, *Speckle phenomena in Optics: Theory and Applications* (Roberts and Company, 2010)
6. G. Gbur, Partially coherent beam propagation in atmospheric turbulence. *JOSA A* **31**, 2038–2045 (2014)
7. N.B. Baranova, B.. Ya.. Zel'dovich, Dislocations of the wave-front surface and zeros of the amplitude. *Sov. Phys. JETP* **53**, 925–929 (1982)
8. N. George, A. Jain, Space and wavelength dependence of speckle intensity. *Appl. Phys. A* **4**, 201–212 (1974)
9. J.F. Nye, M.V. Berry, Dislocations in wavetrains. *Proc. R. Soc. Lond. A* **336**, 165–190 (1974)
10. W. Wang, T. Yokozeki, R. Ishijima, A. Wada, Y. Miyamoto, M. Takeda, S.G. Hanson, Optical vortex metrology for nanometric speckle displacement measurement. *Opt. Express* **14**, 120–127 (2006)
11. W. Martienssen, E. Spiller, Coherence and fluctuations in light beams. *Am. J. Phys.* **32**, 919–926 (1964)
12. E. Wolf, New theory of partial coherence in the space-frequency domain. Part I: spectra and cross spectra of steady-state sources. *J. Opt. Soc. Am.* **72**, 343–351 (1982)

Chapter 11

Polarization of Light



Our analysis of light fields so far has been mostly in terms of a single scalar component of the electric field of light waves. In this chapter, we present a brief discussion on polarization of light for completeness. The vector nature of the diffraction field was briefly discussed in relation to the exact Rayleigh-Sommerfeld diffraction theory. For the purpose of discussion, we will assume the light field to be monochromatic in nature. Polarization and allied topics such as propagation of light in anisotropic crystals are described well in several standard textbooks [1–3]. The material in this chapter is only meant to introduce the reader to some standard notation.

11.1 The Jones Matrix Formalism

For a linearly polarized plane electromagnetic wave traveling in z -direction, the \mathbf{E} and \mathbf{B} fields are perpendicular to each other in transverse direction. The \mathbf{E} field can thus have two independent components in the transverse plane that we may denote by

$$\mathbf{E}_1 = \hat{x} E_{10} \exp[i(kz - \omega t)], \quad (11.1)$$

and

$$\mathbf{E}_2 = \hat{y} E_{20} \exp[i(kz - \omega t)]. \quad (11.2)$$

These two polarizations can be considered as basis states. Further, assuming that E_{10} and E_{20} are complex valued allows us to describe a general state of the transverse electromagnetic field. The polarization may thus be denoted as a 2-element column vector:

$$\mathbf{E} = \frac{1}{\sqrt{|E_{10}|^2 + |E_{20}|^2}} \begin{pmatrix} E_{10} \\ E_{20} \end{pmatrix} \quad (11.3)$$

The special cases of linear polarization in x and y directions are denoted by column vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (11.4)$$

respectively. In the same notation right and left circularly polarized light is denoted as

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \text{and} \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}. \quad (11.5)$$

The elliptically polarized light is represented in a manner similar to the circularly polarized light with the magnitudes of the two components of the vector being unequal. The effect of a polarization sensitive device is typically represented by a 2×2 matrix known as the Jones matrix. Birefringent materials make some of the most important polarization sensitive devices. Suppose that a linearly polarized beam passes through sheet of a material where the x and y components of the polarization vector suffer different phase delays as a result of the polarization dependent refractive index in these two directions. The Jones matrix for such an element may be represented as

$$J = \begin{pmatrix} \exp(i\delta_x) & 0 \\ 0 & \exp(i\delta_y) \end{pmatrix} \quad (11.6)$$

If the thickness of the material is given by L then the two phase differences δ_x and δ_y are given as

$$\delta_x = \frac{2\pi}{\lambda} (n_x - 1)L, \quad (11.7)$$

and

$$\delta_y = \frac{2\pi}{\lambda} (n_y - 1)L. \quad (11.8)$$

Waveplates are commonly used polarization sensitive devices. When a linearly polarized light with polarization vector oriented at 45° to x -axis passes through the above material of thickness such that the two components exiting the material have a phase shift $|\delta_x - \delta_y|$ equal to $\pi/2$ or π , then such a device is called a quarter or half wave plate respectively. It is readily seen that the quarter wave plate (QWP) can convert a linearly polarized light (at 45°) into a left or right circularly polarized light depending on the sign of $(\delta_x - \delta_y)$. In the Jones matrix formalism a polarization rotator device is readily described by a rotation matrix. For example, the matrix for rotation by angle α (counter-clockwise) is given by

$$R(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (11.9)$$

A polarizer is a device which passes either x or y polarized light through it while blocking the other component. The x and y polarizers may thus be described by

$$P_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (11.10)$$

and

$$P_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (11.11)$$

respectively. When the incident linear polarization is tilted by angle θ to the x -axis, the effect of polarizer is given by

$$R(\theta) P_x R(-\theta) = \begin{pmatrix} \cos^2(\theta) & \cos(\theta) \sin(\theta) \\ \cos(\theta) \sin(\theta) & \sin^2(\theta) \end{pmatrix} \quad (11.12)$$

The sequence of operations above may be understood as follows. The incident polarization undergoes a basis (or coordinate axis) change followed by the action of polarizer in the new basis and finally transforming back to the original basis. Note that the polarizer acts as an x -polarizer in the transformed rotated basis.

11.2 The QHQ Geometric Phase Shifter

The geometric phase shifter is an interesting device which consists of a combination QHQ where Q stands for a quarter wave plate and H stands for a half wave plate. We assume a configuration where the fast axes of the quarter wave plates are aligned and that of the half wave plate is rotated by an angle θ with respect to the quarter wave plates. For an incident linear polarization that is 45° with respect to the fast axis of the quarter wave plate, the output field is given by the product:

$$J_{\pi/2} R(\theta) J_\pi R(-\theta) J_{\pi/2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \exp(2i\theta) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (11.13)$$

For relative rotation of the half wave plate by angle θ the incident linearly polarized wave thus gets phase shifted by an angle 2θ . This phase shift can be understood as the geometric phase shift. The QHQ device is an interesting method to introduce phase change in a beam without introducing additional path delay, e.g. in a phase shifting interferometer.

11.3 Degree of Polarization

The state of polarization of a beam of light can be described by a 2×2 “coherency” matrix :

$$C = \begin{pmatrix} \langle E_1^* E_1 \rangle & \langle E_1^* E_2 \rangle \\ \langle E_2^* E_1 \rangle & \langle E_2^* E_2 \rangle \end{pmatrix}, \quad (11.14)$$

where E_1 and E_2 denote the x and y components of the electric field. For example, for unit amplitude x -polarized light,

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (11.15)$$

and for right circularly polarized light, we have

$$C = \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}. \quad (11.16)$$

Since the matrix C is Hermitian it may be diagonalized by a unitary transformation as follows:

$$U J U^\dagger = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \quad (11.17)$$

Assuming $\alpha_1 > \alpha_2$, the degree of polarization of the light beam is defined as

$$\mathcal{P} = \frac{\alpha_1 - \alpha_2}{\alpha_1 + \alpha_2}. \quad (11.18)$$

The eigenvalues of C are determined by solving the equation:

$$\det[C - \alpha I] = 0. \quad (11.19)$$

The degree of polarization can therefore be further expressed as

$$\mathcal{P} = \sqrt{1 - 4 \frac{\det(C)}{[\text{tr}(C)]^2}}. \quad (11.20)$$

Here $\det(\dots)$ and $\text{tr}(\dots)$ stand for the determinant and the trace of the concerned matrix respectively. The important point to note is that when $\alpha_2 = 0$ we have fully polarized light whereas when $\alpha_1 = \alpha_2$ the light is fully unpolarized.

11.4 Non-uniformly Polarized Light

While the treatment of polarization in this chapter so far assumed uniform polarization across the transverse cross-section of the paraxial beams, in the current optics literature, inhomogeneous polarization states have also received much attention. A vector beam with non-uniform polarization may be expressed as

$$\mathbf{E}(x, y) = [\mathbf{e}_1 E_1(x, y) + \mathbf{e}_2 E_2(x, y)]. \quad (11.21)$$

Here, $\mathbf{e}_{1,2}$ denotes orthogonal polarization vectors (e.g. belonging to the $x - y$ basis or the circular basis). The orthogonal polarization components are assumed to be derived from the same source and as a result are coherent with respect to each other. A few popularly studied exotic non-uniform polarization states of light are depicted in Fig. 11.1. They are generated by combining distinct orbital angular momentum states of light in orthogonal polarizations [4] and are perfectly legitimate solutions of the wave equation. The background of the polarization pattern shows the intensity

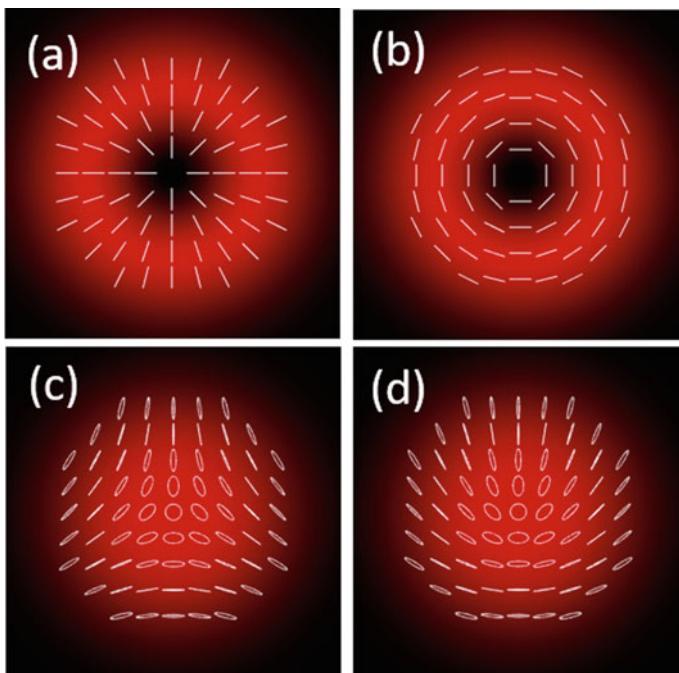


Fig. 11.1 Examples of non-uniform polarization states. The four states shown here are known as **a** radial polarization, **b** azimuthal polarization, **c** star polarization singularity and **d** lemon polarization singularity. Adapted with permission from [4] ©Taylor & Francis, UK

profile of the corresponding beam. This topic is not within the scope of this book, however, it is conceivable that such non-uniform polarization states may have an important role to play in future computational imaging systems.

Problems

11.1 Determine the state of polarization of a light beam whose electric field vector is described by

$$\mathbf{E}(r, \theta, z) = [\mathbf{e}_1 \exp(i\theta) + \mathbf{e}_2 \exp(-i\theta)] r \exp(-r^2/2a^2).$$

Here (r, θ) are polar coordinates given by $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$. The basis vectors $\mathbf{e}_{1,2}$ may be assumed to represent the right and left circularly polarized basis states.

11.2 The light field described in the previous problem is passed through a half wave plate. What will be the resultant polarization state of the beam?

11.3 A linear polarizer typically rejects one (say x) polarization component while passing the orthogonal (y) component. Design an element which will reject the left circularly polarized state while passing the right circularly polarized light.

11.4 Anti-glare screen: Show that the combination of linear polarizer and quarter wave plate can act as an anti-glare element for use with a display device.

11.5 The polarization state of monochromatic spatially coherent light at the two pinholes of Young's experiment is orthogonal with respect to each other. Show that the resultant field observed on the screen shows periodic variation of the polarization state [5].

References

1. M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, (Ed. 7) (Cambridge University Press, 1999)
2. L. Mandel, E. Wolf, *Coherence and Quantum Optics* (Cambridge University Press, 1995)
3. J.W. Goodman, *Statistical Optics* (Wiley-Interscience, 2000)
4. B. Kumar, P. Lochab, E. B. Kayal, D. P. Ghai, P. Senthilkumaran and K. Khare, Speckle in polarization structured light, *J. Mod. Opt.* **69**, 47–54 (2022)
5. D. Gossman, B. Perez-Garcia, R.I. Hernandez-Aranda, A. Forbes, Optical interference with digital holograms. *Am. J. Phys.* **84**, 508–516 (2016)

Chapter 12

Analysis of Optical Systems



This chapter applies several ideas developed so far in the book for wave optical analysis of optical systems [1–4]. The thin convex lens is shown to have capability of implementing the 2D Fourier transform operation. The afocal imaging system is also introduced which allows manipulation of images by Fourier space filtering. Some notations/terminologies are also introduced along the way. A few interesting cases of Fourier space image filtering are also illustrated with numerical examples.

12.1 Transmission Function for a Thin Lens

Lens elements are commonly used for the purpose of imaging. Starting with human eye, lens system is the most important part of commonly used imaging systems such as cameras and microscopes. We will consider thin lenses for the present discussion. In the ray optical model, a thin lens refers to a lens for which a ray parallel to optic axis entering one surface of the lens exits the other surface at approximately the same height from the optic axis. This is a useful approximation in determining the transmission function of a typical lens. We assume the radii of curvatures for the two lens surfaces to be R_1 and R_2 , respectively, as shown in Fig. 12.1. Denoting by n the refractive index of the lens material, the optical path difference encountered along a line at height h parallel to the optic axis and between the two surfaces bounding the lens is given by

$$OPD(x, y) = \frac{2\pi}{\lambda} [(L_0 - L(x, y)) + nL(x, y)]. \quad (12.1)$$

Here $(L_0 - L(x, y))$ denotes the air path length and the corresponding refractive index has been assumed to be equal to 1. Our goal is to express $L(x, y)$ in terms of

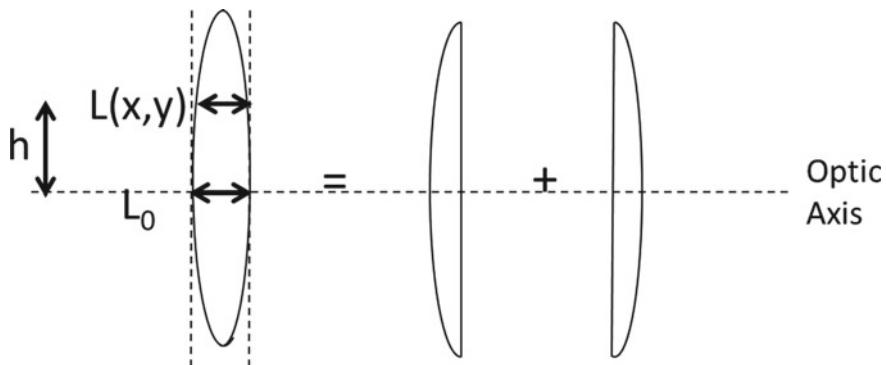


Fig. 12.1 Determination of the transmission function for a thin lens. L_0 and $L(x, y)$ denote the thickness of glass material on optic axis and at height h above the optic axis. The radii of curvatures of the two lens surfaces are R_1 and R_2 , respectively

the radii of curvature R_1 and R_2 of the two lens surfaces. As shown in Fig. 12.1, we divide the thin lens into two parts for convenience. The glass length $L(x, y)$ can now be expressed as

$$L(x, y) = L_0 - \left[R_1 - \sqrt{R_1^2 - (x^2 + y^2)} \right] - \left[-R_2 - \sqrt{R_2^2 - (x^2 + y^2)} \right]. \quad (12.2)$$

Note that here we have used the usual geometrical optics convention that the second surface with its center of curvature to the left has negative radius of curvature. Further, if we approximate the square roots in Eq. (12.1) by assuming that R_1 and $-R_2$ are much larger than (x, y) dimensions in the problem, the length $L(x, y)$ can be expressed as

$$L(x, y) = L_0 - \frac{(x^2 + y^2)}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right). \quad (12.3)$$

Substituting back in Eq. (12.1) gives

$$OPD(x, y) = \frac{2\pi}{\lambda} \left[nL_0 - (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \frac{(x^2 + y^2)}{2} \right]. \quad (12.4)$$

Finally, recognizing the expression $(n-1)(1/R_1 - 1/R_2)$ as the inverse focal length $1/F$ of the lens, we get the transmission function for a thin lens as follows:

$$t(x, y) = \exp[iOPD(x, y)] = \exp \left[-\frac{i\pi}{\lambda F} (x^2 + y^2) \right]. \quad (12.5)$$

In the above expression, we have omitted the constant phase factor involving the lens thickness L_0 along the optical axis as it is inconsequential.

12.2 Fourier Transforming Property of a Thin Lens

In this section, we will discuss the very important 2D Fourier transforming property of a lens which is key to several imaging and optical processing applications. We will consider a configuration involving a lens and free-space propagation as shown in Fig. 12.2. An object transparency with transmittance $t(x', y')$ is placed at a distance d in front of the lens. The output field is observed at the back focal plane of the lens. Assuming illumination of the object by a plane wave propagating along the optic axis, the calculation of the output $g(x, y)$ in terms of the input $t(x', y')$ involves the following steps:

1. Fresnel propagation from input plane to the lens.
2. Multiplication by lens transmission function.
3. Fresnel propagation from lens to output plane.

The Fresnel propagation formula developed in Chap. 9 and the lens transmission function obtained in the previous section may be used for the above steps and the result is stated directly as follows:

$$g(x, y) = \frac{e^{[i\pi/(F\lambda)(x^2+y^2)+ik(F+d)]}}{i\lambda F} \iint dx' dy' t(x', y') \exp[-i\frac{2\pi}{\lambda F}(xx' + yy')]. \quad (12.6)$$

We observe that the input-output fields above have a Fourier transform relation with an additional quadratic phase factor. *Further, if we choose $d = F$, the quadratic phase factor vanishes and the Fourier transform relationship is exact.* As per the definition in this book, the (x, y) coordinates in the output plane may be associated with the spatial frequencies in the input as

$$f_x = \frac{x}{\lambda F}, \quad f_y = \frac{y}{\lambda F}. \quad (12.7)$$

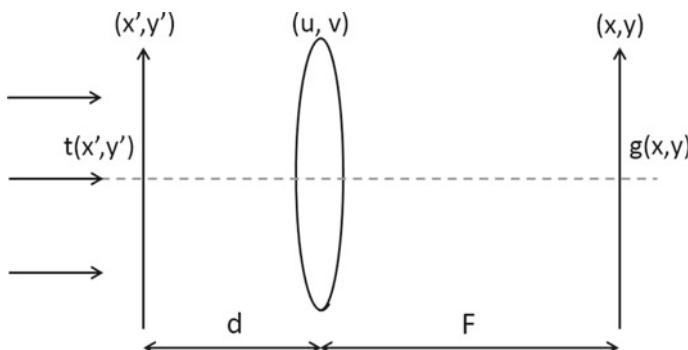


Fig. 12.2 Fourier transforming property of a lens

The 2D Fourier transforming property makes the lens a very important tool for 2D optical signal processing. The Fourier transforming property may be understood easily in terms of the angular spectrum picture for diffraction. The input $t(x', y')$ when illuminated by a plane wave of wavelength λ will result in an angular spectrum of plane waves. Each plane wave is focused in its back focal plane to a spot (ideally) representing a delta function. The location of this focal spot for a particular plane wave indexed by (f_x, f_y) is at the location $x = \lambda F f_x$ and $y = \lambda F f_y$ in the output plane. The output pattern is an interference of all the plane waves accepted by the lens aperture. Finally, we would like to mention that due to the dependence of the spatial frequency scaling on focal length F of the lens, the extent of the Fourier transform pattern may be controlled so that it is sampled well by a pixelated detector placed in the output plane. The Fraunhofer diffraction pattern which is the 2D Fourier transform of the object transmission function may thus be readily observed (with appropriate scaling) in the back focal plane of a lens.

12.3 Canonical Optical Processor

Having derived the Fourier transforming property of a convex lens, we will now consider an important imaging system—the canonical or 4F optical processor. This is a general imaging system model which may be used to describe a large number of imaging systems with minor modifications. The system of interest is as shown in Fig. 12.3. Continuing from previous section, we observe that the canonical optical system consists of two Fourier transform systems placed one after the other. When an object $g_{in}(x', y')$ in the input plane is illuminated with a plane wave of wavelength λ , the field in the intermediate plane is the scaled Fourier transform $G(f_x, f_y)$. For this reason, we will call this intermediate plane as the Fourier plane. The spatial frequencies are scaled in terms of coordinates (u, v) in this plane as $f_x = u/(\lambda F)$ and

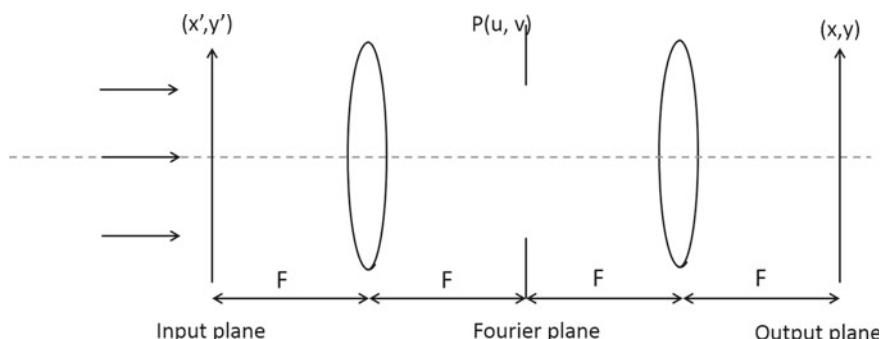


Fig. 12.3 Canonical Optical Processor system layout: the focal lengths of the two lenses are given by F , input plane coordinates are (x', y') , output plane coordinates are (x, y) and the intermediate or Fourier plane coordinates are (u, v)

$f_y = v/(\lambda F)$. A *pupil function* $P(u, v)$, which may in general have both amplitude and phase variations, controls the behavior of the spatial frequencies in the sense that it either blocks or passes a particular spatial frequency (f_x, f_y) or modifies the Fourier component $G(f_x, f_y)$ with both amplitude and phase weight. $P(u, v)$ can thus be understood as a Fourier plane filter. The output plane field is described by

$$\begin{aligned} g_{out}(x, y) &= \mathcal{F}\{G\left(\frac{u}{\lambda F}, \frac{v}{\lambda F}\right)P(u, v)\} \\ &= \iint dx' dy' g_{in}(x', y') p(x + x', y + y'). \end{aligned} \quad (12.8)$$

Here, $p(x, y)$ is the *coherent impulse response* of the system and has a Fourier transform relationship with the aperture $P(u, v)$ of the system as given by

$$p(x, y) = \frac{\exp(4ikF)}{i\lambda F} \iint du dv P(u, v) \exp\left[-i\frac{2\pi}{\lambda F}(ux + vy)\right]. \quad (12.9)$$

Note that the output $g_{out}(x, y)$ is a convolution of the input field $g_{in}(x, y)$ with a flipped impulse response $p(-x, -y)$. The flipping of coordinates here represents image inversion. The inversion may also be understood as due to two successive applications of 2D Fourier transform. In the further discussion, we will usually omit the effect of this inversion of coordinates. We remark that the 4F system model may be used in describing a large number of imaging and diffraction phenomena. For example, in describing diffraction phenomena using the angular spectrum approach, one may consider an effective aperture $P(u, v)$ to be the free-space transfer function that defines how the spatial frequencies in the input object are modified. We will use this system for study of several imaging phenomena as illustrated with some interesting filter choices next.

12.3.1 DC Block or Dark Field Imaging

DC blocking is an important technique that allows one to observe details in an image that would otherwise be hard to detect in presence of strong background. An example is the detection of planets near bright stars or the problem of imaging of solar atmosphere. A filter that blocks the dc component of light in Fourier plane of the 4F system and passes the high-frequency components may achieve this. We illustrate this technique with example of a high intensity 2D Gaussian object to mimic a bright star and a small weak intensity object present in the Gaussian tail (see Fig. 12.4a).

By blocking the frequencies near the point $(f_x, f_y) = (0, 0)$ in the Fourier plane, the bright background has been removed as seen in Fig. 12.4b. The overall dynamic range of the image intensity is therefore significantly modified and as a result the faint object is now visible. A similar approach is also used frequently in the dark-field mode in microscopes. In dark-field microscopy, the condenser lens used to illuminate the

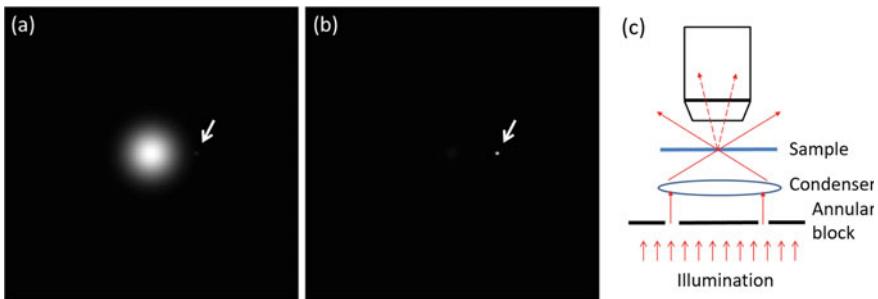


Fig. 12.4 Effect of dc blocking on a simulated bright star image with weak planet object nearby: **a** test object, **b** resultant image after dc blocking, **c** Annular illumination in a dark-field microscope. The dotted arrows represented large angle scattering by sample which contributes to image formation. The direct scattered light does not enter the objective lens

object/sample has a central circular block. The resulting oblique annular illumination means that the scattered light collected by the imaging objective lens only arises out of large angle scattering as shown in Fig. 12.4c. The resultant effect is similar to that of dc blocking. A disadvantage of this technique is that the direct scattered light is lost; however, the methodology clearly provides stunning visual imagery of the detailed high-frequency structure in the sample.

12.3.2 Zernike's Phase Contrast Microscopy

Phase contrast microscopy is an important Fourier filtering technique first described by F. Zernike in 1942. Commonly studied biological cell specimens are transparent in nature and so their contrast is very poor when observed under a bright-field illumination. The dc blocking technique may be used here but it has one disadvantage that it blocks most of the light. Zernike suggested the use of a quarter wave phase shifting element that significantly enhances the image contrast. The working of the Zernike method can be understood in an approximate manner as follows. Suppose that the incident light undergoes a phase change of $\psi(x, y)$ after passing through a transparent cell specimen. For a weak phase object, we may approximate

$$\exp[i\psi(x, y)] \approx 1 + i\psi(x, y). \quad (12.10)$$

The intensity $I(x, y)$ corresponding to this wavefront observed on an imaging sensor is given by

$$I(x, y) \approx |1 + i\psi(x, y)|^2 = 1 + [\psi(x, y)]^2. \quad (12.11)$$

For a weak phase object, the first term dominates over the square of the second term and reduces the overall image contrast. Blocking the dc component here amounts to

effectively eliminating the first term and the image intensity is proportional to the square of the phase imparted by the object. However, if a small quarter wave phase shifting element is placed near the zero frequency in the Fourier plane, the resultant image intensity is given by

$$I(x, y) \approx |i(1 + \psi(x, y))|^2 = 1 + 2\psi(x, y) + \psi(x, y)^2. \quad (12.12)$$

For a weak phase object, the image intensity is thus approximately linearly related to the phase change instead of square of the phase function and resultant images have significantly improved contrast. Commercial phase contrast microscopes typically use annular or ring illumination and a corresponding annular quarter wave phase shifter in the Fourier plane. This arrangement has advantage that it has good light throughput and further that from a user perspective aligning the ring image when inserting the Zernike phase plate is much easier. In the illustration, in Fig. 12.5, we simulate a weak phase object that is assumed to be imaged by a 4F system. The binary phase function (phase levels are 0 and $\pi/4$) is shown in Fig. 12.5a, and the intensity in the image plane without any phase filter in the Fourier plane is shown in Fig. 12.5b. The small contrast here arises due to finite aperture of the system. Figure 12.5c shows the effect of Zernike's phase contrast filter clearly illustrating the contrast enhancement. The phase contrast mechanism provides significantly enhanced images for Bio-science researchers when transparent cells are to be visualized using a microscope. We note that the phase range $\pi/4$ is not too small and as a result the enhanced contrast image also shows a halo. Even for biological samples like unstained blood cells, the weak phase approximation does not necessarily hold. The resultant detected image is then valuable only qualitatively as the features in the cell are visible better compared to usual bright-field microscopy when the phase contrast idea is employed. It may be noted that Eq. (12.12) does not explain the appearance of the eventual image in detail. In recent times, the phase contrast principle has also been used in cryo-electron microscopy where Zernike phase plates added to the Fourier plane have shown substantial improvement in the contrast of the recorded raw images that are



Fig. 12.5 Simulation of Zernike phase contrast: **a** simulated phase map of weak phase object where phase values of 0 and $\pi/4$ are used in the binary phase object for illustration, **b** low-contrast bright-field intensity image, **c** phase contrast image

typically low-contrast and highly noisy. While phase contrast is a powerful technique, we remark that it is a qualitative imaging method. This modality does not provide quantitative phase information about the sample object under consideration. We will discuss quantitative phase imaging in the context of digital holography later in the book.

12.3.3 *Highlighting of Edges with Vortex Filter*

A vortex filter is an interesting case for study and provides an equivalent of a derivative operation uniformly in all directions. We define the vortex filter as

$$P(\rho, \phi) = \exp(i\phi). \quad (12.13)$$

A physical vortex filter plate has unit amplitude everywhere and a spirally varying phase which changes uniformly from 0 to 2π . For 1D signal, a derivative-like operation is provided by the Hilbert transform (or equivalently a $-i\text{sgn}(v)$ filter) which converts a cosinusoidal signal to a sinusoidal one and vice versa with a sign change. The sine-cosine pair is commonly known by the name quadrature pair since the two signals are out of phase by $\pi/2$. This idea is often used in the $\exp(i\omega t)$ or complex representation in the study of wave phenomena. There is no straightforward generalization of the Hilbert transform concept in two or higher dimensions. The vortex phase as an isotropic Hilbert transform filter was first suggested by Larkin and co-workers in 2001 [5]. The spiral phase function behaves like a sign function along any radial line passing through the origin of the (f_x, f_y) plane. We illustrate the effect of the spiral phase filter in enhancing the edges in an image isotropically. This edge enhancing effect due to a vortex filter has been utilized successfully in microscopy. Figure 12.6a, b shows an amplitude object and its corresponding image using a 4F system when a vortex phase filter is placed in the Fourier transform plane of the system aligned with the optic axis. The unit amplitude of the filter implies that the energy in the input plane is not lost but preferentially re-distributed near the points of high magnitude of gradients in the object function.

12.3.4 *Apodization Filters*

The coherent impulse response of a 4F imaging system may have side-lobes analogous to the Gibbs oscillations discussed in Chap. 2, particularly when the Fourier plane aperture has hard edges at the boundaries. Such side-lobes can create ringing artifacts in the resultant image. For example, the impulse response corresponding to a rect-function aperture is a sinc function which has long tail and side-lobes. A suitable window function used along with the filter function significantly reduces the side-lobes—an effect known as apodization (or suppression of side-lobes). We

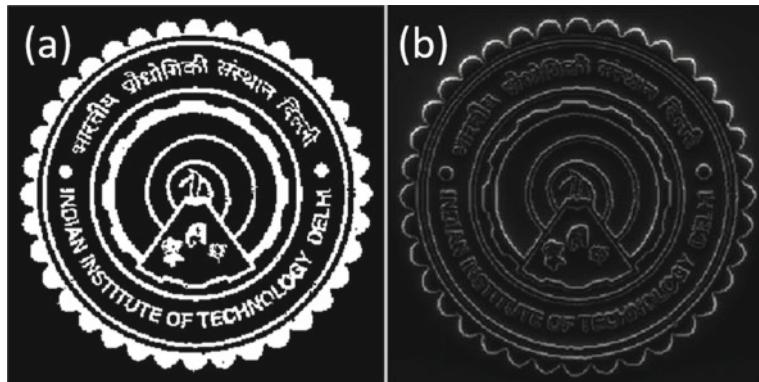


Fig. 12.6 Simulation of the effect of vortex phase filter on an image: **a** amplitude object, **b** effect of vortex phase filter



Fig. 12.7 Simulation for illustrating apodization: **a** test image, **b** test image filtered using a low-pass rect filter, **c** test image filtered using a Gaussian weighted low-pass rect filter.

illustrate this effect with an example of a low-pass rect filter for which a Gaussian apodization window is applied. The effect of a Gaussian window used with the rect filter will be to convolve the corresponding sinc-impulse response with the Fourier transform of the window function (another Gaussian). This averaging effectively reduces the side-lobes of the sinc function. In the illustration, we show a test image in Fig. 12.7a which is low-pass filtered using a simple rect filter and a Gaussian weighted rect filter as in Fig. 12.7b, c, respectively. The side-lobes of the sinc function are seen to create ringing or ghost-like artifacts near sharp edge features after filtering. These artifacts are seen to be mitigated in the Gaussian apodized low-pass filter. Neutral density apodization filter with radially decreasing transmission function is used in SLR cameras for giving a smooth defocused background region behind the main object being photographed.

12.4 Frequency Response of Optical Imaging Systems: Coherent and Incoherent Illumination

Having studied the effects of various Fourier plane filters, we now require a formal way to compare the performance of different imaging systems from the perspective of their spatial frequency response. Once again based on our discussion on the 4F system, we have already noted in Eq. (12.9) that the coherent impulse response $p(x, y)$ of a system is related to the aperture $P(u, v)$ by a Fourier relationship. We define the coherent transfer function $H(f_x, f_y)$ of the imaging system as the Fourier transform of $p(x, y)$ as follows:

$$\begin{aligned} H(f_x, f_y) &= \iint_{-\infty}^{\infty} dx dy p(x, y) \exp[-i2\pi(f_x x + f_y y)] \\ &= \iint_{-\infty}^{\infty} du dv P(u, v) \delta(f_x - \frac{u}{\lambda F}, f_y - \frac{v}{\lambda F}) \\ &= P(\lambda F f_x, \lambda F f_y). \end{aligned} \quad (12.14)$$

Here, we have omitted some inconsequential constant factors in the last step. While we have concentrated on the coherent imaging case so far, several commonly used imaging systems like cameras operate with illumination that is virtually spatially incoherent. In order to arrive at the input-output relation for such systems, we start by relating the two-point correlation function in the object and image planes. The correlation function is given by

$$\begin{aligned} \langle u^*(x_1, y_1) u(x_2, y_2) \rangle = \\ \iiint dx'_1 dy'_1 dx'_2 dy'_2 \langle u^*(x'_1, y'_1) u(x'_2, y'_2) \rangle p^*(x_1 - x'_1, y_1 - y'_1) p(x_2 - x'_2, y_2 - y'_2). \end{aligned} \quad (12.15)$$

For incoherent illumination, the two-point correlation in the input plane is given by a delta function:

$$\langle u^*(x'_1, y'_1) u(x'_2, y'_2) \rangle = I(x'_1, y'_1) \delta(x'_1 - x'_2, y'_1 - y'_2). \quad (12.16)$$

Further setting $x_1 = x_2$ and $y_1 = y_2$ in the output plane gives a relation between field intensities in the input and output planes:

$$I(x_1, y_1) = \iint dx'_1 dy'_1 I(x'_1, y'_1) |p(x_1 - x'_1, y_1 - y'_1)|^2 = I(x_1, y_1) * |p(x_1, y_1)|^2. \quad (12.17)$$

For spatially incoherent illumination, the function $|p(x, y)|^2$ thus takes the role of the system's impulse response. *The impulse response $|p(x, y)|^2$ is referred to as*

the incoherent impulse response or the point spread function (PSF) of the imaging system. Since $p(x, y)$ is related to the aperture of the system by a Fourier transform relation, the frequency response of the system for incoherent case is somewhat more involved. The frequency response for the incoherent illumination case is known by the name of optical transfer function (OTF) that is defined as

$$\begin{aligned} OTF(f_x, f_y) &= \frac{\mathcal{F}\{|p(x, y)|^2\}}{\mathcal{F}\{|p(x, y)|^2\}_{f_x=0, f_y=0}} \\ &= \frac{\iint dudv P(u + \frac{\lambda F f_x}{2}, v + \frac{\lambda F f_y}{2}) P^*(u - \frac{\lambda F f_x}{2}, v - \frac{\lambda F f_y}{2})}{\iint dudv |P(u, v)|^2}. \end{aligned} \quad (12.18)$$

Here, we have used the auto-correlation theorem in the last step, and the two terms under integral sign in the numerator have been made symmetric. In this form, the OTF is similar to the ambiguity function defined in Eq. (2.93). The absolute magnitude of the OTF is known as the magnitude transfer function (MTF). For illustration of the concepts above, we will consider a 1D 4F system with rectangular aperture $P(u) = \text{rect}(u/(2w_x))$. The coherent cut-off frequency of the system is obtained by replacing the variable u with $\lambda F f_x$ so that the highest spatial frequency passed by the system is given by $f_{cc} = (w_x)/(\lambda F)$. The coherent impulse response of the system is given by $p(x) = 2w_x \text{sinc}(2w_x x)$ and the OTF is given by convolution of two rect functions which is a triangle function:

$$OTF(f_x) = \Lambda\left(\frac{\lambda F f_x}{2w_x}\right). \quad (12.19)$$

The incoherent cut-off frequency is thus given by $f_{ic} = (2w_x)/(\lambda F)$ which is twice that of the coherent cut-off frequency. This result does not necessarily imply that incoherent systems are capable of resolving better. For example, suppose that in the above example, a coherently illuminated system can just resolve a signal $g(x) = \cos(2\pi f_0 x)$ with $f_0 = f_{cc} = (w_x/\lambda F)$. When incoherent illumination is used, only intensities are of concern so that the input to the system that can just be resolved is now $|g(x)|^2$ which has twice the extent in Fourier space as compared to $g(x)$.

The aberrations in an imaging system may be modeled as phase errors in the Fourier plane relative to an open aperture. For example, defocus may be modeled by inserting a quadratic phase mask in the Fourier plane. As per standard convention, the defocused wavefront in the Fourier plane aperture may be associated with a phase map given by

$$W(\rho) = \frac{2\pi}{\lambda} w_d \rho^2. \quad (12.20)$$

Here, w_d is the defocus coefficient and ρ represents the normalized radial coordinate in the Fourier plane aperture which varies from 0 to 1. The MTF curves for two different defocus values ($w_d = 0.25\lambda, 0.65\lambda$) are shown in Fig. 12.8a. We note that

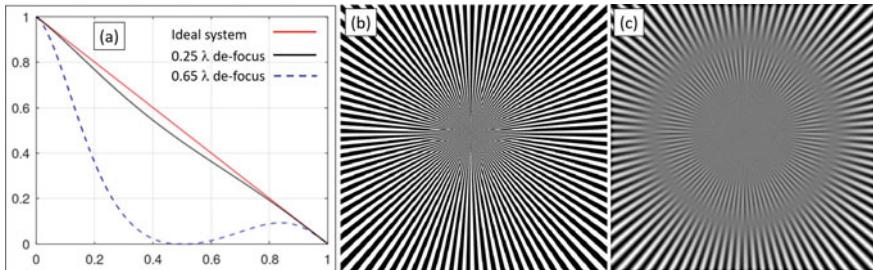


Fig. 12.8 Illustration of defocus aberration. **a** The MTF curves for ideal aberration-free system and for two defocus values $w_d = 0.25\lambda$, 0.65λ , **b** spokes object, **c** defocused image of the object for defocus of $w_d = 0.65\lambda$

all the three MTF curves in Fig. 12.8a reach zero at the same cut-off frequency (denoted as 1 on the x-axis). This is because from the definition of OTF in Eq. (12.18), we see that the cut-off frequency depends on the size of the Fourier plane aperture. The behavior of the three MTF curves is however quite different. For open aperture case, the MTF curve nearly fall-off in straight line. The MTF curve for small defocus $w_d = 0.25\lambda$ is very close to this open aperture MTF. For higher defocus of $w_d = 0.65\lambda$, we however see that the MTF reaches zero and then rises again. Correspondingly, the OTF shows zero-crossing which is accompanied by a π phase shift. In Fig. 12.8c, we illustrate the effect of defocus $w_d = 0.65\lambda$ on the spokes object in Fig. 12.8b. The spokes image is interesting in that the spatial frequency of the spokes features increases as we move radially to the center of the image. We observe that in Fig. 12.8c, the central part is almost blurred indicating that higher spatial frequencies beyond the cut-off are not passed by the imaging system. Relatively the edges of the images representing the low-spatial frequencies are reproduced nearly undisturbed. At an intermediate spatial frequency beyond the zero crossing of the OTF, we also observe contrast reversal due to the π phase shift. The effect of magnitude fall-off of OTF is thus to reduce contrast in high-frequency features. The phase of OTF on the other hand causes relative phase shift between different spatial frequencies resulting in aberration-induced image distortion. We will have more occasions to discuss the behavior of defocused system in the later chapter on PSF engineering. For now, we note that the OTF (or MTF) of two imaging systems can provide a helpful tool for quantitative comparison of the frequency response and their imaging performance.

We have provided a discussion of Fourier optics analysis of imaging systems by means of illustrations based on the 4F canonical optical system. The formalism for understanding the frequency response of imaging systems using coherent and incoherent illumination was also discussed in some detail. The formalism developed in this chapter will be useful in discussion of a number of computational optical systems in the third part of this book.

Problems

12.1 Consider an F-Lens-F optical Fourier transform system that you wish to set up in the laboratory. A phase-only transmissive liquid crystal SLM is to be placed at the input of the system and the Fourier transform output is recorded on a CCD camera. The SLM pixel size is $10 \mu\text{m}$ and the CCD camera pixel size is $5 \mu\text{m}$. Both the SLM and CCD have 1000×1000 pixels. The illumination is with a plane laser beam which fully illuminates the SLM active area and is blocked outside the SLM area. The phase function displayed on the SLM is such that the alternate columns of the SLM have phase values 0 and π .

- (a) What focal length F would you choose for the lens so that the diffraction peaks are well separated on the CCD camera?
- (b) Write a computer program to model this system in a computer and show the output image.

12.2 An ultra-thin annular (ring-like) slit of radius r_1 is placed in the input plane of the F-Lens-F Fourier transforming system. For illumination using a monochromatic (wavelength = λ) plane wave beam traveling along the optic axis of the system, find an expression for the resultant scalar field in the back focal plane of the lens. Find an expression for the field for propagation by a distance z_0 beyond the back-focal plane and show that the beam intensity profile has not changed due to propagation.

12.3 A 4F imaging system has a rectangular Fourier plane aperture of size $(2u_x \times 2u_y)$.

- (a) Find the coherent impulse response of the system.
- (b) Find the OTF of the system and sketch the MTF along the x -dimension.
- (c) What is the incoherent cut-off frequency for this system?
- (d) Sketch the approximate system MTF when the system has aberrations.

12.4 A resolution test target—a bar chart with $(1,0)$ periodic amplitude transmission function—serves as an input to a 2F-lens-2F imaging setup. The period of the input pattern is $1/f_o$ mm. Spatially incoherent illumination with wavelength λ is used in this experiment.

- (a) You have two lenses with aperture diameters $D_1 = (\lambda F f_0 / 2)$ and $D_2 = (3\lambda F f_0)$ and same focal length F available to you in the lab. Assuming a imaging system configuration with 1:1 magnification, which lens would you use for imaging the bar chart object and why?
- (b) A pixelated digital detector is to be used for recording the image in this experiment. What is the maximum permissible pixel size that can produce a meaningful image?

12.5 An imaging system has a centrally obscured square aperture of size $a \times a$. The central $a/3 \times a/3$ square part of the aperture is opaque. Sketch (approximately) the MTF curve for the system along the x -axis.

References

1. M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, (Ed. 7) (Cambridge University Press, 1999)
2. J.W. Goodman, *Introduction to Fourier Optics*, (Ed. 2) (McGraw-Hill, 1996)
3. K. Iizuka, *Elements of Photonics*, vol. 1 and 2 (Wiley, 2002)
4. J.W. Goodman, *Statistical Optics* (Wiley-Interscience, 2000)
5. K.G. Larkin, D.J. Bone, M.A. Oldfield, Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform. *J. Opt. Soc. Am. A* **18**(8), 1862–70 (2001)

Chapter 13

Imaging from Information Point of View



In the view of the general framework of computational imaging systems combining optical hardware and algorithms for image recovery as outlined in Chap. 1, we need a mathematical framework for comparison, analysis and possibly synthesis of novel imaging system concepts. As novel optical components (e.g. aspherics, freeform optics, binary/multi-level phase masks, etc.) get incorporated in imaging system design, the conventional metrics for imaging system quality (e.g. Strehl ratio) become difficult to use in practice. This is because image formation process does not end at the image detection stage. The image recovery algorithms have an important role to play as well. A central question one may study is the quantification of information that an image recovery algorithm may be able to retrieve after light waves carrying this information travel from object through the imaging system aperture and get detected by the imaging sensor. The problem of quantifying information in light waves was first addressed by Gabor in the 1950s [1] and subsequently by many researchers [2–10]. The treatment in this chapter is presented in terms of the canonical or 4F system as it can effectively mimic a large number of imaging systems. The focus of our analysis will be on basic elements of information that are transmitted by a canonical 4F imaging system.

13.1 Eigenmodes of a Canonical Imaging System

In this section, we will present a brief analysis of the 4F imaging system, and in particular, the information-carrying capacity of the system from the object to the image plane. In this context, we will discuss the eigenfunction analysis of the 4F imaging systems as presented in [11, 12]. This discussion is on the lines of the discussion on the prolate spheroidal functions in Sect. 3.4. For simplicity, a 1D notation will be used here. We have already seen in Eq. (12.9) that the impulse

response $p(x)$ of a 4F system is related to its aperture $P(u)$ by a Fourier transform relationship. Since the aperture of any practical imaging system is finite in extent, the impulse response $p(x)$ is a bandlimited function and therefore has a sampling expansion. Assuming the bandwidth of the impulse response to be $2B$, we have

$$p(x) = \sum_{m=-\infty}^{\infty} p\left(\frac{m}{2B}\right) \text{sinc}(2Bx - m). \quad (13.1)$$

Further assuming that the input object extends to a length $x : (-L, L)$, we will study the possibility of finding the eigenmodes $\psi_n(x)$ of the systems that obey the following eigenvalue equation:

$$\mu_n \psi_n(x) = \int_{-L}^L dx' p(x - x') \psi_n(x'). \quad (13.2)$$

Here, μ_n refers to the eigenvalue corresponding to the eigenfunction $\psi_n(x)$. The eigenvalue μ_n and the eigenfunction $\psi_n(x)$ may be complex-valued in general depending on the nature of $p(x)$. To maintain consistency with the analysis on prolate spheroidal functions in Sect. 3.4, we have used the impulse response $p(x - x')$ above instead of $p(x + x')$ as is expected from Eq. (12.8). This sign change only refers to inversion of the resultant image and is therefore inconsequential for the present discussion. We note that in the discussion of prolate functions in Sect. 3.4, a similar eigenvalue equation was studied for the case when $p(x) = \text{sinc}(2Bx)$. The analysis there solely depended on the fact that the sinc function has a sampling expansion. A similar analysis for the present problem gives rise to a discrete eigenvalue problem that is equivalent to the continuous eigenvalue problem in Eq. (13.2). The discrete eigenvalue problem may be stated as

$$A \mathbf{u}_n = \mu_n \mathbf{u}_n, \quad (13.3)$$

where the elements of the matrix A are defined as

$$A_{mk} = \int_{-L}^L dx' p\left(\frac{m}{2B} - x'\right) \text{sinc}(2Bx' - k), \quad m, k = -\infty, \dots, 0, \dots, \infty. \quad (13.4)$$

The vector \mathbf{u}_n is defined as

$$\mathbf{u}_n = \left[\dots \psi_n\left(\frac{m}{2B}\right) \dots \right]^T \quad (13.5)$$

and is seen to be made up of the Nyquist samples of the eigenfunction $\psi_n(x)$. We further note an interesting connection between the matrix A and the matrix $A^{(0)}$ in Eq. (3.22) that was defined in the analysis of the prolate spheroidal functions. The impulse response $p(m/(2B) - x')$ in the definition of matrix A above may be further expressed as a sampling expansion with respect to the variable x' to get

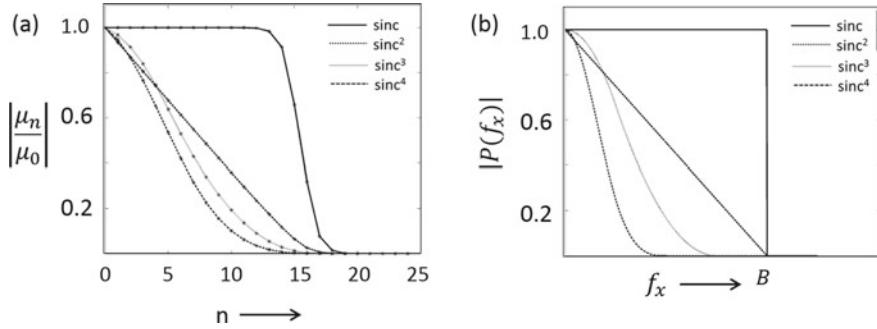


Fig. 13.1 Eigenfunction analysis of 4F system: **a** eigenvalues corresponding to the impulse responses $p(x) = \text{sinc}, \text{sinc}^2, \text{sinc}^3, \text{sinc}^4$ all having the same bandwidth, **b** Corresponding magnitudes of the aperture functions $|P(f_x)|$. The numerical values $L = 2, B = 2$ are used in these illustrations. Note the striking similarity between the eigenvalue fall-off trends and the magnitude $|P(f_x)|$ of the transfer function. Adapted with permission from [12] ©Institute of Physics, UK

$$\begin{aligned} A_{mk} &= \sum_{l=-\infty}^{\infty} p\left(\frac{m-l}{2B}\right) \int_{-L}^L dx' \text{sinc}(2Bx' - l) \text{sinc}(2Bx' - k) \\ &= \sum_{l=-\infty}^{\infty} p\left(\frac{m-l}{2B}\right) A_{lk}^{(0)}. \end{aligned} \quad (13.6)$$

The matrix A is therefore simply a product of the matrix $p_{ml} = p((m-l)/(2B))$ that is formed using the Nyquist samples of the impulse response and the matrix $A^{(0)}$ that was defined in Eq. (3.22) in the analysis of the prolate spheroidal functions. The solution of the discrete eigenvalue problem above gives the eigenvectors that may be sinc-interpolated to get the eigenfunctions $\psi_n(x)$. As an illustration of this sampling theorem-based procedure, we show in Fig. 13.1a the relative fall-off of eigenvalues for the four impulse responses $p(x) = \text{sinc}(2Bx), \text{sinc}^2(2Bx), \text{sinc}^3(2Bx/3), \text{sinc}^4(2Bx/2)$ all of which have the same bandwidth. The relative fall-off in the absolute magnitude of the Fourier transform $P(f_x) = \mathcal{F}\{p(x)\}$ is plotted in Fig. 13.1b for comparison. It is interesting to note that there is a striking similarity in the behavior of the eigenvalue plots and $P(f)$. The numerical values of $L = 2, B = 2$ and a matrix A of size 101×101 were used for this illustration [12]. With a reasoning similar to that for the prolate function case in Sect. 3.4, we note that the number of significant eigenvalues for an impulse response is approximately equal to the space-bandwidth product $(2L)(2B)$ of the system, although in detail the fall-off in eigenvalues depends on the specific impulse response. The set of eigenfunctions $\{\psi_n(x)\}$ may now be considered as the modes or the fundamental elements of information that may be transmitted by the 4F system from the input to the output plane. The function set is specific to the system impulse response. If the input to the system is represented as a linear combination of the eigenfunctions (at least in the sense of uniform convergence),

$$g_i(x) = \sum_n a_n \psi_n(x), \quad (13.7)$$

then the output of the system is easy to determine and is given by

$$g_o(x) = \sum_n \mu_n a_n \psi_n(x). \quad (13.8)$$

The modes with index $n >> 4BL$ are thus transmitted with very small eigenvalues and may get lost in any detection noise at the output end. Further, for a given noise level in detection at the output end, one may compare different systems in terms of their information-carrying capacity based on the nature of eigenvalue fall-off. We observe that the analysis presented in this section does not rely on any symmetry properties of the impulse response function $p(x)$ and as a result general cases involving system aberrations (represented by phase variations in Fourier plane) may also be handled using the same procedure. Such analysis typically leads to complex-valued eigenfunctions and eigenvalues. We may refer to such complex-valued eigenfunctions as the eigenwavefronts associated with the imaging system [11].

13.1.1 Eigenfunctions and Inverse Problems

The eigenmodes of a general 4F imaging system as may be determined using the procedure shown in the previous section are potentially useful for treating the inverse or image de-blurring problems. Suppose that the measured output to a 4F system is denoted by $\tilde{g}_{out}(x)$ which includes any noise at the detection end. We may represent the measured output as a linear combination of the eigenmodes $\psi_n(x)$ as

$$\tilde{g}_{out}(x) = \sum_n b_n \psi_n(x). \quad (13.9)$$

While the eigenmodes may not in general be orthogonal to each other, the representation above may be understood as a projection onto the space spanned by the eigenmodes. In other words, the coefficients b_n are selected such that the squared L2-norm error

$$\|\tilde{g}_{out}(x) - \sum_n b_n \psi_n(x)\|_2^2$$

is minimized. Having represented the output in terms of the eigenmodes, the estimated solution for the input is

$$\tilde{g}_{in}(x) = \sum_n \frac{b_n}{\mu_n} \psi_n(x). \quad (13.10)$$

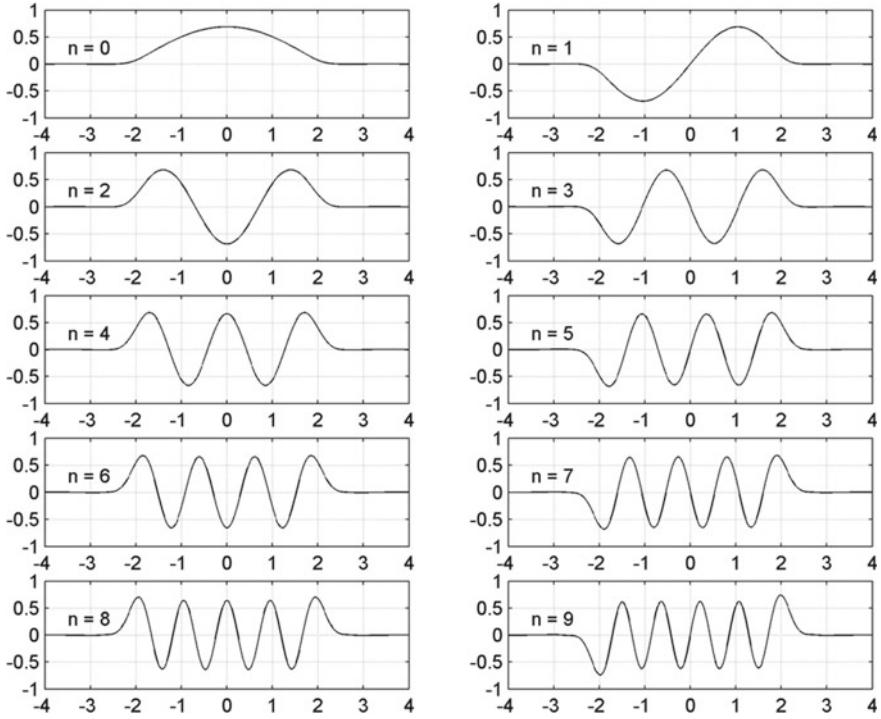


Fig. 13.2 First ten eigenfunctions corresponding to the sinc^4 impulse response. The numerical parameters used are $L = 2$, $B = 2$. Adapted with permission from [12] ©Institute of Physics, UK

We observe that the coefficients b_n contain the effect of noise in the measured data $\tilde{g}_{out}(x)$. As the index n of the terms in Eq. (13.10) increases beyond $(2B)(2L)$ we are in effect dividing the noisy coefficients b_n by small numbers. Division of the noisy coefficient by small numbers in general amplifies the noise in the inverse solution. A simple regularization method to avoid this is to truncate the series solution to a finite number of terms thus effectively reducing the information transmitted from object to image space. The noise at the image detection plane is thus seen to be the fundamental limiting factor in terms of information retrievable by means of an inverse solution as in Eq. (13.10). We illustrate these ideas with a specific example of the inverse problem with the eigenfunctions of the $\text{sinc}^4(Bx/2)$ impulse response. The first ten eigenfunctions for this impulse response for the parameters $L = 2$, $B = 2$ are shown in Fig. 13.2. A raised triangle object and the corresponding output blurred by the sinc^4 impulse response with 1% Gaussian random additive noise are shown in Fig. 13.3 a and b, respectively. The reconstruction of the object using 10, 14 and 18 eigenfunctions is shown in Fig. 13.4a, b, respectively. The solid curves in this figure show the recovery whereas the dotted curves show the test object. The illustration shows that adding terms with index $N > 4BL$ causes noise amplification due to division by small eigenvalues as discussed above. The noise in the measured blurred

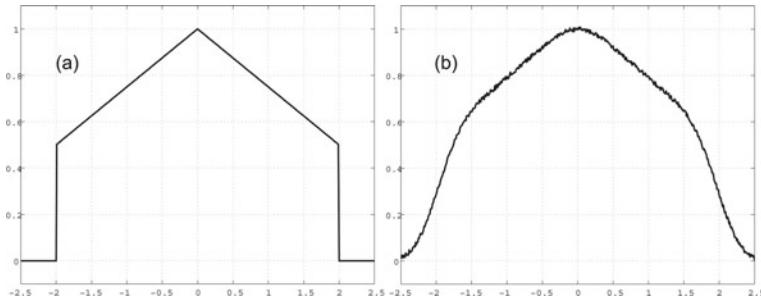


Fig. 13.3 **a** Raised triangle object, **b** Object blurred with sinc^4 impulse response with 1% additive random noise. Adapted with permission from [12] ©Institute of Physics, UK

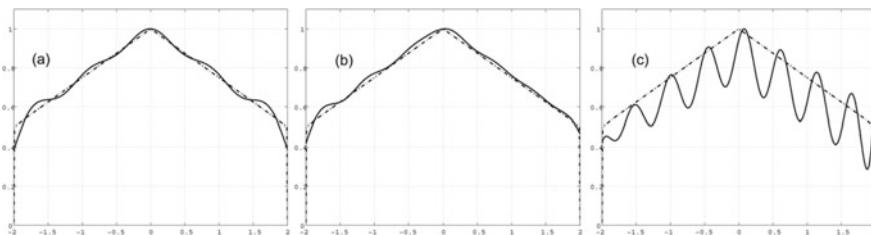


Fig. 13.4 Estimation of inverse solution with eigenfunctions for 1% Gaussian random noise. **a, b** show recoveries using 10, 14 and 18 eigenfunctions, respectively. Solid curves: recovery, Dotted curves: test object. Adapted with permission from [12] ©Institute of Physics, UK

data therefore inherently limits the number of terms that can be included in the series solution thereby limiting the achievable resolution. The addition of constraints on the solution (e.g. smoothness or sparsity as seen in Chap. 6) may allow the inclusion of some additional number of terms leading to further improvement in recovered object resolution. Practically, the best solution is however obtained only when higher resolution information is present in the raw data as we will see in later chapters on Fourier ptychography and structured illumination imaging.

We would like to point out that the sampling theorem-based approach for eigenfunction analysis of the 4F imaging system can be extended to address the singular value problem associated with this system as well. The singular functions offer a slight advantage over the eigenfunctions when used for the purpose of the inverse recovery problem as illustrated in Fig. 13.4. The eigenfunction analysis of a generic imaging system was presented in this chapter for a single scalar component of light field that is propagated from object to image space. Additional degrees of freedom including polarization and spectrum may be used to understand the full electromagnetic degrees of freedom available to an imaging system. We will not explore this topic in this book but point the reader to detailed work on this theme [13].

Problems

13.1 Singular value decomposition: Consider the singular value decomposition of a bandlimited kernel $p(x)$ as

$$p(x - x') = \sum_{m=0}^{\infty} \sigma_m \alpha_m(x) \beta_m(x'),$$

where $\sigma_m > 0$ are the singular values. The impulse response $p(x)$ is bandlimited to the frequency interval $(-B, B)$ as in Sect. 13.1. The functions $\alpha_m(x)$ and $\beta_m(x)$ satisfy the following relations:

$$\int_{-L}^L dx' p(x - x') \alpha_m(x') = \sigma_m \beta_m(x), \quad x : (-\infty, \infty)$$

and

$$\int_{-\infty}^{\infty} dx p^*(x - x') \beta_m(x') = \sigma_m \alpha_m(x'). \quad x' : (-L, L).$$

Show that the functions $\alpha_m(x)$ and $\beta_m(x)$ with $x : (-\infty, \infty)$ are bandlimited to the frequency interval $(-B, B)$. If \mathbf{h} stands for the matrix $p((n - m)/(2B))$, show that the eigenvectors $\mathbf{w}_{\alpha m}$ and $\mathbf{w}_{\beta m}$ corresponding to the discrete eigenvalue problems

$$\mathbf{h}^\dagger \mathbf{h} \mathbf{A}^{(o)} \mathbf{w}_{\alpha m} = (2B\sigma_m)^2 \mathbf{w}_{\alpha m}$$

and

$$\mathbf{h} \mathbf{A}^{(o)} \mathbf{h}^\dagger \mathbf{w}_{\beta m} = (2B\sigma_m)^2 \mathbf{w}_{\beta m}$$

yield Nyquist samples of the functions $\alpha_m(x)$ and $\beta_m(x)$ at intervals $1/(2B)$. Here $\mathbf{A}^{(o)}$ is the matrix defined in Eq. (3.22).

13.2 Show that the functions $\alpha_m(x)$ and $\beta_m(x)$ in the singular value decomposition in the previous problem satisfy the orthogonality relations:

$$\int_{-L}^L dx \alpha_m^*(x) \alpha_n(x) = \delta_{m,n}$$

and

$$\int_{-\infty}^{\infty} dx \beta_m^*(x) \beta_n(x) = \delta_{m,n}.$$

13.3 Show that the functions $\alpha_m(x)$ and $\beta_m(x)$ reduce to the prolate spheroidal basis functions (or Slepian functions) when $p(x) = \text{sinc}(2Bx)$.

13.4 Consider two canonical optical processor (or 4F) systems. One has an open aperture in the Fourier plane whereas the other has a quadratic phase in its Fourier

plane aperture (e.g. denoting 1λ defocus aberration). The two systems are otherwise identical. For a given detector at the output end with a specified detection SNR, which system will be able to recover more information (or degrees of freedom) at the output end? (You may include computational inverse solution as a part of the imaging system.)

13.5 Suggest a method to compute eigenmodes $\{\psi_n(x)\}$ associated with free-space diffraction over a fixed distance z_0 from a rectangular aperture by restricting attention to paraxial spatial frequencies.

References

1. D. Gabor, "Light and Information", *Progress in Optics*, vol. 1 (ed E. Wolf), 109-153 (North-Holland, Amsterdam) (1961)
2. C.K. Rushforth, R.W. Harris, Restoration, resolution and noise. *J. Opt. Soc. Am.* **58**, 539–545 (1968)
3. G. Toraldo di Francia, Degrees of freedom of an image. *J. Opt. Soc. Am.* **59**, 799–804 (1969)
4. M. Bendinelli, A. Consortini, L. Ronchi, B.R. Frieden, Degrees of freedom and eigenfunctions for the noisy image. *J. Opt. Soc. Am.* **64**, 1498–1502 (1974)
5. M. Bertero, E.R. Pike, Resolution in diffraction limited imaging, a singular value analysis I, the case of coherent illumination. *Opt. Acta* **29**, 727–746 (1982)
6. M. Bertero, P. Boccacci, E.R. Pike, Resolution in diffraction limited imaging, a singular value analysis II, the case of incoherent illumination. *Opt. Acta* **29**, 1599–1611 (1982)
7. M. bertero, P. Brianzi, P. Parker and E. R. Pike, "Resolution in diffraction limited imaging, a singular value analysis III, The effect of sampling and truncation of the data", *Opt. Acta*. 31, 181-201 (1984)
8. I.J. Cox, C.J.R. Sheppard, Information capacity and resolution in an optical system. *J. Opt. Soc. Am. A* **3**, 1152–1158 (1986)
9. A.W. Lohmann, R.G. Dorsch, D. Mendlovic, Z. Zalevsky, C. Ferreira, Space-bandwidth product of optical signals and systems. *J. Opt. Soc. Am. A* **13**, 470–473 (1996)
10. D.A.B. Miller, Spatial channels for communicating with waves between volumes. *Opt. Lett.* **23**, 1645–1647 (1998)
11. K. Khare, N. George, Sampling theory approach to eigenwavefronts of imaging systems. *J. Opt. Soc. Am. A* **22**, 434–438 (2005)
12. K. Khare, Sampling theorem, bandlimited integral kernels and inverse problems. *Inverse Probl.* **23**, 1395–1416 (2007)
13. R. Piestun, D.A.B. Miller, Electromagnetic degrees of freedom of an optical system. *J. Opt. Soc. Am. A* **17**, 892–902 (2000)

Part III

Computational Imaging Systems

Chapter 14

Digital Holography



Phase measurement is one of the most important problems in Optics. When light waves are scattered by an object, the information about the object is encoded in the scattered waves in the form of amplitude as well as phase variations (apart from possible changes in polarization, spectrum, etc.). The phase information in scattered light waves is often considered to be much more valuable compared to the amplitude information. All detectors operating at visible wavelengths respond to time-averaged energy incident on individual detection element like a pixel. Ordinary photography cameras do manipulate the phase of incoming light waves by means of a lens assembly but only record the intensity (or amplitude) information. Holography is an interferometric technique that records both amplitude as well as phase information. Holographic imaging was first demonstrated by Gabor [1] and a major advance in the form of an off-axis hologram was demonstrated by Leith and Upatnieks [2]. There are several excellent resources that discuss the history and recent developments of holography as provided in references to this chapter [3, 4]. The main idea behind the possibility of holographic imaging was already discussed in the context of Rayleigh-Sommerfeld-Smythe diffraction theory in Chap. 9. In particular we obtained the important result for free-space diffraction that if the tangential E-field is specified in a particular plane (say $z = 0$), then all the components of the E-field are completely determined in the right half space. Consider an object that on illumination by a coherent beam of light produces a certain tangential field $O(x, y, z = 0)$ in the plane $z = 0$ as shown in Fig. 14.1. Now suppose that the same field can be generated in the $z = 0$ plane by some other means so that an observer in the $z > 0$ region receives the same diffracted field as before. In that case even if the original object is not physically present to the left of the screen, the viewer of the scene will perceive its presence. It is clear that the field $O(x, y, z = 0)$ has both amplitude as well as phase in general which needs to be recorded and then re-created. One of the simplest means to generate the same field $O(x, y, z = 0)$ is to record it

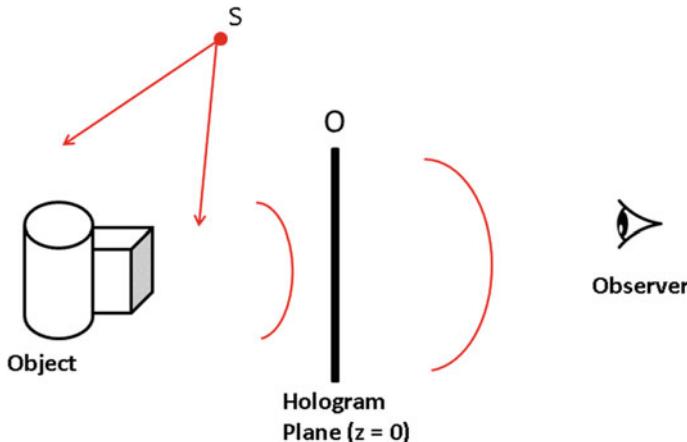


Fig. 14.1 Spatially coherent wavefront from source S illuminates an object to produce scattered field whose tangential scalar component in the hologram plane is given by $O(x, y, 0)$. If this field component can be recorded and generated in the hologram plane again, then the viewer on the right-hand side will get the same perception as if the object was present in its original location

interferometrically. As we have already studied in Chap. 10, an interference pattern with good fringe contrast (or visibility) may be recorded if the two interfering beams are mutually coherent. Typically a reference beam R is derived from the same source by means of a beamsplitter device and is interfered with the object beam O at the $z = 0$ plane which we will now refer to as the hologram plane. If a square-law area detector is placed in plane $z = 0$, the interference pattern H recorded on the detector is described as:

$$H = |R + O|^2 = |R|^2 + |O|^2 + R^*O + RO^*. \quad (14.1)$$

The phase information about the object wave O is recorded in an encoded form in the last two terms of the equation above. The amplitude as well as phase information about the object beam may be recorded this way using multiple configurations. For example, if the field $O(x, y)$ at the hologram plane represents the Fraunhofer or Fresnel diffraction pattern corresponding to the object being imaged, we may term the hologram as a Fourier transform hologram or a Fresnel hologram. If on the other hand the object field $O(x, y)$ corresponds to the image of the object of interest, as may be recorded by a $4F$ system, the hologram is termed as an image plane hologram.

In traditional film-based holography, the hologram pattern H used to be recorded on a high-resolution photographic film. The replay of the hologram is then achieved by physical re-illumination of the hologram with a counter-propagating or conjugate beam R^* . Assuming that the transmission function of the developed film is linearly proportional to the intensity pattern H , the re-illumination process of the hologram with the R^* beam creates a field profile R^*H in the hologram plane:

$$R^* H = R^* |R|^2 + R^* |O|^2 + R^{*2} O + |R|^2 O^*. \quad (14.2)$$

The cross terms in the above equation (third and fourth terms on right-hand side) are important as they are proportional to the complex object waves O and O^* . These two terms evolve via free-space diffraction and form real and virtual images of the original object. From early days of holography till mid-1990s, holographic imaging continued to be performed preferentially using high-resolution photographic films although hologram recording on an electronic array detector was demonstrated as early as in 1970s [5]. Our goal in this chapter is to concentrate on the topic of digital holography which has become a popular computational phase imaging modality over the last couple of decades. In digital holography, the hologram is recorded using an electronic array detector like CCD or CMOS sensors. Once the hologram frame H is recorded and is available in a numerical form, the hologram replay process cannot be carried out physically but can only be performed using numerical processing methods [6–11]. We will discuss some of the traditional methods for processing digital holograms (holograms recorded on array sensors) that are in-principle similar to the traditional film-based hologram replay. We will also present a brief discussion of some recent developments in using an optimization approach to complex object wave recovery from holograms. The new optimization approach will be seen to overcome the important limitations of the traditional numerical processing methods. Finally, we will point out an important distinction between 3D visual perception in display holography and computational 3D holographic image reconstruction. In particular *it will be shown that the traditional holographic replay does not produce a numerical field distribution identical to the original 3D object.* We will further present a methodology of what we call as the “true 3D” reconstruction. The overall discussion of these advanced topics brings out an important point that digital holography can go much beyond simple mimicking the holographic reconstruction as achieved by traditional film-based methods.

14.1 Some Practical Aspects of Digital Holography Systems

14.1.1 In-line and Off-axis Configurations

Digital holograms can be recorded in two main configurations, viz. in-line and off-axis geometry. The original Gabor hologram was recorded using the in-line configuration shown in Fig. 14.2a. In an in-line holography system, the objects are typically small in size (e.g. particulates, micro-organisms, etc.) that are illuminated by a plane or spherical wavefront which is large in extent compared to the objects. The weak scattering due to the small objects suggests that the resultant field on the detector is a sum of the illuminating beam as well as the scattered field from the objects. The corresponding interference pattern is commonly referred to as in-line hologram and the typical look of the interference pattern for this case is shown in Fig. 14.2b. From the

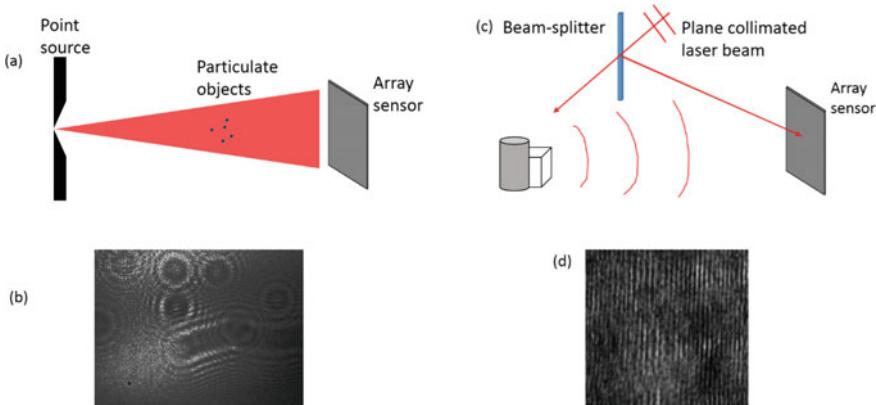


Fig. 14.2 In-line and off-axis holographic recording configurations are shown in **a, c**, respectively. The typical holograms recorded with these geometries are shown **b** and **d**

hologram in Fig. 14.2b, we observe interference fringe patterns due to several point objects and a thin thread-like object (seen as an elongated pattern). For a point object in the path of a spherical wave, the approximate form of the interference pattern may be expressed as

$$H \approx |A \exp[i \frac{\pi}{\lambda z_1} (x^2 + y^2)] + B \exp[i \frac{\pi}{\lambda z_2} ((x - x_0)^2 + (y - y_0)^2)]|^2. \quad (14.3)$$

Here z_1, z_2 refers to the distance between the point source and the point object from the array sensor and (x_0, y_0) are the transverse coordinates of the point object. The amplitudes A and B above have a typical relationship $|A| >> |B|$ due to weak scattering amplitude from the point object. Further the amplitudes A, B also fall off inversely with the distance between the point object or point source and the detector coordinate. Referring to Fig. 14.2a, we note that when multiple weak scattering centers are present in the reference beam path, the second term in Eq. (14.3) will be replaced by a summation of individual scattered fields. The resultant interference pattern contains multiple local circular fringe patterns which may or may not overlap. It may be noted that both the reference and object beams travel along common path in in-line configuration. The off-axis hologram configuration on the other hand involves a reference beam which is incident on the detector at an angle as shown in Fig. 14.2c relative to the nominal direction of the object beam. This configuration typically leads to straight line fringes as shown in Fig. 14.2d that are modulated due to the object beam. If the off-axis reference beam is denoted by a tilted plane wave $R = R_0 \exp(i 2\pi f_0 x)$ and the object beam at the detector is denoted by $O(x, y)$, the interference pattern recorded on the array sensor may be described as:

$$H = |R|^2 + |O|^2 + 2|R||O| \cos[2\pi f_0 x - \phi_O(x, y)]. \quad (14.4)$$

Here $\phi_O(x, y)$ represents the phase map associated with the complex-valued object wave at the detector plane. The fringes described as per the above equation are shown in Fig. 14.2d.

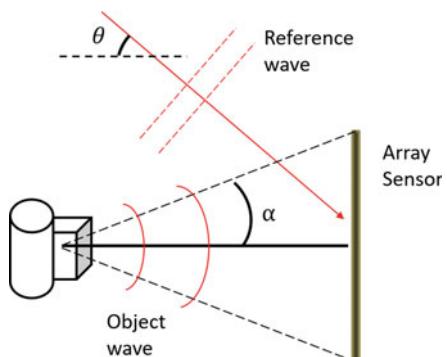
14.1.2 Sampling of Digital Holograms

When recording an interference pattern on an array sensor, it is important to remember that the oscillations in the interference fringe pattern need to be sampled appropriately at least as much as governed by the Nyquist sampling rate. If the angle between the k -vectors of two plane waves (derived from the same source) is given by θ as shown in Fig. 14.3, then the straight line interference fringes associated with this configuration have a period Λ given by

$$\Lambda = \frac{\lambda}{2 \sin \theta}. \quad (14.5)$$

Here λ is the illuminating wavelength. Clearly *as per the Nyquist criterion, we must have at least two samples of the interference pattern per fringe period along every pixel row which is normal to the fringe direction*. In practice it is suitable to have 4–5 pixels span one fringe period since any array sensor does not really measure point samples but rather integrates light irradiance over the pixel size. For sensor pixel size of $5 \mu\text{m}$ and $\lambda = 0.5 \mu\text{m}$, the nominal angle between two beams cannot be very large (at most 1–2 degrees) when recording a digital hologram. In this respect, digital holography differs from film-based holography as high-resolution films readily allow much larger angle (30–40 degrees) between two interfering beams. For circular fringes as in in-line holography similar considerations apply so that the largest visible circular fringe is sufficiently sampled. In the discussion hereafter we will assume that holograms have been recorded with the appropriate sampling considerations.

Fig. 14.3 Numerical aperture for the hologram recording geometry.
Numerical aperture limited by detector size is given by
 $NA = \sin \alpha$



14.1.3 Numerical Aperture of the Hologram Recording System

In both in-line and off-axis configurations shown in Fig. 14.2, it is important to understand that the numerical aperture ($NA = \sin \alpha$) subtended by the detector with respect to the nominal object position decides the extent of angular spectrum from the scattered object field received at the sensor (see Fig. 14.3). When we discuss the recovery of complex-valued object field in the hologram plane or consider the holographic reconstruction or replay process the effective NA plays an important role in deciding the recoverable information (in terms of lateral and longitudinal resolution) about the test object to be imaged. In particular, one may expect a lateral resolution of $\sim \lambda/(NA)$ and a longitudinal resolution of $\sim \lambda/(NA)^2$ from a configuration as shown in Fig. 14.3.

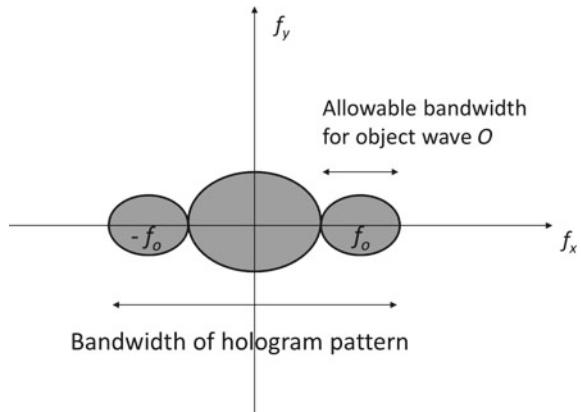
14.2 Complex Object Wave Recovery in the Hologram Plane

In this section, we will discuss three methods for recovery of complex object wave $O(x, y)$ from the recorded hologram(s) $H(x, y)$ as in Eq. (14.1). Following this, in the next section, we will examine the question of image reconstruction using the recovered object field $O(x, y)$.

14.2.1 Off-axis Digital Holography

Off-axis digital holography that was first proposed in the work of Leith and Upatnieks [2] uses a plane wave reference beam at a nominal angle θ relative to the object wave $O(x, y)$ (see Fig. 14.3). The reference beam may be described as $R = R_0 \exp(i2\pi f_0 x)$ where $f_0 = \sin \theta / \lambda$ and the resulting interference pattern can be described as in Eq. (14.4). The amplitude $|R|$ of the plane reference beam is essentially constant (or have very low frequency features in an experiment) across the hologram. The Fourier transform structure of the off-axis hologram is as shown in Fig. 14.4 and consists of three main regions where the energy in the hologram signal is concentrated. The first two term $|R|^2$ and $|O(x, y)|^2$ are located near the center of the (f_x, f_y) plane and are known as the dc terms. The cosine term in Eq. 14.4 can be split into two terms that are located at spatial frequency locations $(f_0, 0)$ and $(-f_0, 0)$. The information of interest about the complex object wave $O(x, y)$ is contained in one of the side lobes in the Fourier plane. The simplest approach to recover this information is to filter out the side-lobe by numerical processing. For this method to work, one must make sure that the three different lobes as in Fig. 14.4 do not overlap with each other substantially. Suppose the expected bandwidth of the

Fig. 14.4 Fourier transform representation of the off-axis hologram showing dc and cross terms



object wave is $2B$ along the x -direction, the width of the central lobe in the Fourier transform plane corresponding to the term $|O(x, y)|^2$ is $4B$ and the transform of $|R|^2$ is almost a delta function located at the zero frequency. In order to make sure that the dc and the cross terms remain separated, we must therefore have:

$$f_0 = \frac{\sin \theta}{\lambda} \geq 3B. \quad (14.6)$$

The reference beam angle is thus selected carefully to avoid overlap of terms in Fourier space and subsequent loss of information about the object wave. We observe that the allowable object wave bandwidth is much smaller than the total hologram bandwidth along the x -direction due to the non-overlap condition. In single-shot operation, one may process the digitally recorded hologram signal in following procedure which is commonly referred to as the Fourier transform method (FTM).

1. Compute the 2D Fourier transform of the off-axis hologram.
2. Locate the carrier-frequency peak (or centroid) and filter out appropriate region of the Fourier transform plane centered on it.
3. Shift the filtered region to zero frequency (or dc) in Fourier transform plane.
4. Compute inverse Fourier transform of the dc shifted field to obtain the complex object field $O(x, y)$ in the hologram plane.

This procedure was first described by Wolf [12] and later by Takeda [13]. In the step 2 above, locating the carrier-frequency peak in the spatial frequency domain is not always a straightforward matter. The finite resolution of the FFT function used to compute the Fourier transform of the off-axis hologram may make it difficult to locate the carrier-frequency peak accurately. If the hologram is for example represented on 2D array of size $N \times N$, the spatial frequency resolution on using the FFT function is $\Delta f_x = \Delta f_y = 1/N$. There is no guarantee that the peak due to the carrier frequency will fall at the center of a pixel in the Fourier transform domain. If the carrier-fringe

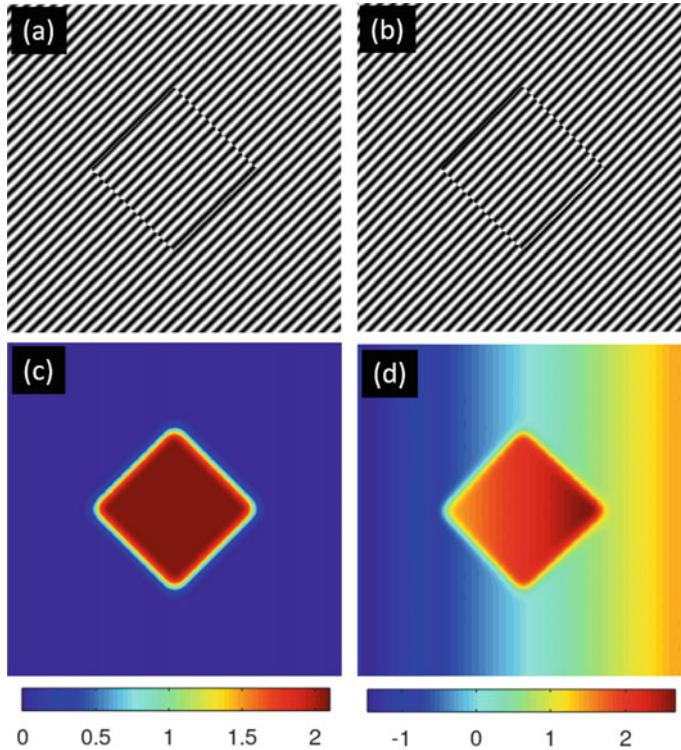


Fig. 14.5 Phase map obtained by following the nominal four step procedure for the Fourier transform method. **a, b:** Two holograms of the same step phase object where the fringe directions are slightly rotated with respect to each other, **c, d:** Phase maps show relative tilt phase due to sub-pixel shift in the carrier-frequency peak

frequency shifts this peak location by Δ_x , Δ_y , respectively ($|\Delta_x| < 0.5$ pixel and $|\Delta_y| < 0.5$ pixel), the resultant solution may show an overall tilt phase given by

$$\Delta\phi(x, y) = \frac{2\pi}{N}(\Delta_x x + \Delta_y y). \quad (14.7)$$

As an illustration in Fig. 14.5, we show two holograms for a square-shaped binary step phase object. The phase step in this illustration is equal to $2\pi/3$. The fringes in the two holograms look nearly same to the eye but have a slight relative rotation between them. If the four step procedure outline above is followed, we observe that the resultant phase map has a tilt phase background in one of the phase recoveries as shown in Fig. 14.5c, d. This tilt has appeared despite the fact that the object wave in both the holograms is identical. In order to address this fractional fringe detection issue, the pixels near the carrier-frequency peak need to be sampled at higher sampling density to estimate any sub-pixel peak shift. This is best achieved by evaluating the

local discrete Fourier transform (DFT) of the hologram near the carrier-frequency peak [14]. After evaluating the carrier-frequency peak (u_0, v_0) from the FFT of the hologram, we may thus proceed as follows: A local DFT centered on (u_0, v_0) may be computed as a product of three matrices:

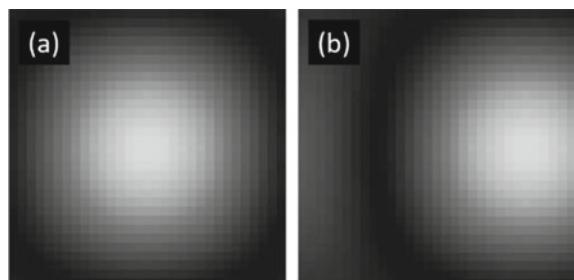
$$F(\mathbf{U}, \mathbf{V}) = \exp\left(-i \frac{2\pi}{N} \mathbf{U}\mathbf{X}^T\right) H(\mathbf{X}, \mathbf{Y}) \exp\left(-i \frac{2\pi}{N} \mathbf{Y}\mathbf{V}^T\right). \quad (14.8)$$

Here \mathbf{X}, \mathbf{Y} are spatial coordinate vectors given by $\mathbf{X} = \mathbf{Y} = [-N/2, -N/2 + 1, \dots, N/2 - 1]^T$ (assuming even N) in the hologram plane. The spatial frequency coordinates (\mathbf{U}, \mathbf{V}) are centered on (u_0, v_0) . If we want to sample a neighborhood of say 1.5 pixels surrounding (u_0, v_0) , this may be achieved by defining:

$$\begin{aligned} U &= u_0 + [-\frac{1.5}{2}, -\frac{1.5}{2} + \frac{1}{\alpha}, -\frac{1.5}{2} + \frac{2}{\alpha}, \dots, \frac{1.5}{2} - \frac{1}{\alpha}]^T \\ V &= v_0 + [-\frac{1.5}{2}, -\frac{1.5}{2} + \frac{1}{\alpha}, -\frac{1.5}{2} + \frac{2}{\alpha}, \dots, \frac{1.5}{2} - \frac{1}{\alpha}]^T. \end{aligned} \quad (14.9)$$

In the above relations, α is the up-sampling factor with which we wish to sub-sample the pixels near the carrier-frequency peak. For a hologram of size $N \times N$, this process matrix product has three matrices of size $(1.5\alpha \times N)$, $(N \times N)$ and $(N \times 1.5\alpha)$ which is marginal additional computation when $\alpha \ll N$. Typically size of the holograms is $N \sim 10^3$, whereas an up-sampling factor $\alpha \sim 10^1$ is sufficient. For the case illustrated in Fig. 14.5, the magnitude of the three matrix product in two cases for $\alpha = 20$ is shown in Fig. 14.6. The two matrices are of size 30×30 and represent the sub-sampled representation of 1.5 pixel near the pixel (u_0, v_0) in Fourier transform of the hologram. The sub-pixel shift corresponding to the tilt phase in Fig. 14.5d is clearly visible in Fig. 14.6b. This tilt can be corrected by finding the sub-pixel peak shifts Δ_x, Δ_y as in Eq. (14.7) from the matrix in Fig. 14.6b and subtracting the resultant $\Delta\phi(x, y)$ from the computed phase map in Fig. 14.5b. This sub-pixel peak shift correction is important for standardizing the usage of off-axis interferometry systems. The phase recovery result using the Fourier transform methodology as illustrated in Fig. 14.5 suggests that even after the sub-pixel shift correction, the methodology still has one drawback that the spatial resolution of the recovered phase

Fig. 14.6 a, b Show the magnitude of the three matrix product in Eq. (14.8) for the interference patterns in Fig. 14.5a, b, respectively



map is poor due to the low-pass filtering nature of this methodology which does not allow one to reconstruct the sharp features in the phase map. Note that this resolution limitation is purely due to the Fourier filtering methodology. It may be observed from Fig. 14.5 that the interference data itself contains the information about sharp edges in the form of fringe discontinuities. As we will discuss in the next sections, the lost resolution can be regained by a multi-step phase shifting methodology or by employing a more advanced optimization methodology for complex object wave reconstruction from a single interference data frame.

14.2.2 Phase Shifting Digital Holography

Phase shifting is a technique [15] may be used to regain the lost resolution in Fourier transform method. Multiple digital holograms are recorded on a digital array sensor with known phase shifts in the reference beam. For example if the R beam gets four pre-determined phase shifts $0, \pi/2, \pi$, and $3\pi/2$, the corresponding digital holograms as per Eq. (14.1) are given by

$$\begin{aligned} H_1 &= |R|^2 + |O|^2 + 2|R||O| \cos(\phi_O - \phi_R), \\ H_2 &= |R|^2 + |O|^2 - 2|R||O| \sin(\phi_O - \phi_R), \\ H_3 &= |R|^2 + |O|^2 - 2|R||O| \cos(\phi_O - \phi_R), \\ H_4 &= |R|^2 + |O|^2 + 2|R||O| \sin(\phi_O - \phi_R). \end{aligned} \quad (14.10)$$

In the above equation, the terms ϕ_O and ϕ_R denote the 2D phase functions corresponding to the object and reference beams, respectively, at the hologram plane. Known phase shifts can be added to the reference beam in several ways, e.g. introducing appropriate combination of wave retarder plates in the reference beam, introducing a QHQ (Q=quarter wave plate, H= half wave plate) geometric phase shifter in the reference beam, using a piezo-electric transducer to shift the location of an optical component like a mirror in the reference beam path, etc. By adding the known reference phase shifts to the reference beam, it is now possible to obtain both the sine and cosine quadratures of the unknown phase. If the reference phase ϕ_R is known, we can obtain the object beam phase as

$$\phi_O(x, y) = \phi_R(x, y) + \arctan\left(\frac{H_4 - H_2}{H_1 - H_3}\right). \quad (14.11)$$

Observe that unlike off-axis holography, the reference beam does not have to be an off-axis plane wave but can in general have any known phase profile. In Fig. 14.7 we diagrammatically show the complex object recovery for a step phase object (step size equal to $\pi/2$) using the phase shifting method. For illustration a spherical reference wave is used. The steps in the recovery of the square shaped step phase object are illustrated in Fig. 14.7 and we clearly see the phase object recovery with full

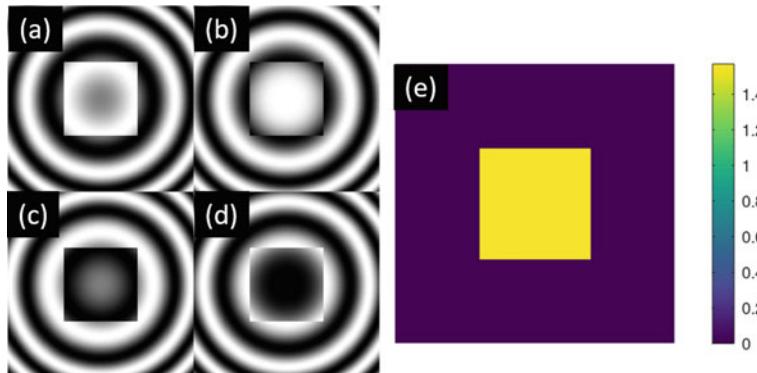


Fig. 14.7 Phase shifting digital holography, **a–d**: Four holograms for the square step phase object with reference beam phase shifts given by $0, \pi/2, \pi, 3\pi/2$, respectively. **e** Recovery of object phase using Eq. (14.11)

pixel resolution. At the cost of recording multiple frames the phase shifting digital holography regains the full detector resolution. It is required that during the recording of the multiple hologram frames, the object remains stationary. Also it is important to remember that while recording the four data frames as in Eq. (14.10) in the laboratory, each data frame contains noise (e.g. detector readout noise and photon noise), and as a result by processing the four frames as above there is a noise penalty to pay in terms of accuracy of the measured phase.

14.2.2.1 Accuracy Analysis of Classical Interferometers

The analysis for phase measurement accuracy is typically performed by analyzing the measurements at the two output ports of an interferometer as shown in Fig. 14.8. The unknown phase is assumed to be $(\pi/2 + \Delta\phi)$ for convenience. If N photons are incident on the interferometer as shown in Fig. 14.8, the detectors D_1 and D_2 (assumed to be ideal) at the two output ports of the interferometer will detect average number of photons $N_{1,2}$ given by

$$N_{1,2} = \frac{N}{2} [1 \pm \sin(\Delta\phi)]. \quad (14.12)$$

From the phase measurement perspective the signal of interest is

$$N_s = N_1 - N_2 = N \sin \Delta\phi \approx N \Delta\phi, \quad (14.13)$$

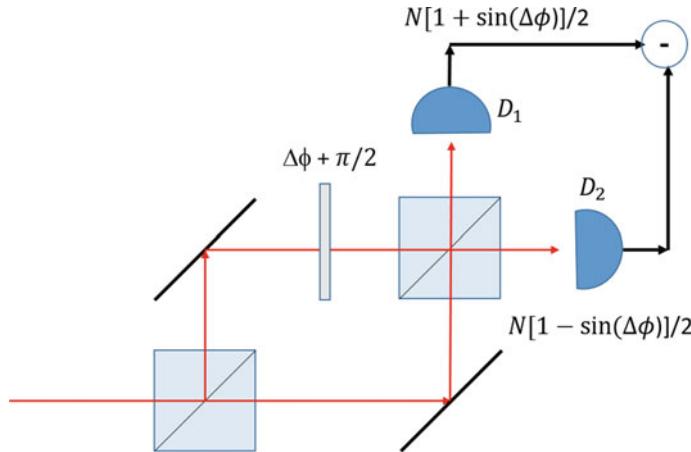


Fig. 14.8 Accuracy of phase measurement in an interferometer

for small angle $\Delta\phi$. For classical light sources and typical detection integration times, the photon counts at the two detectors $N_{1,2}$ follow Poisson statistics. The standard deviation in the signal N_s can then be determined as

$$\sigma_{N_s} = \sqrt{\sigma_{N_1}^2 + \sigma_{N_2}^2} = \sqrt{N}. \quad (14.14)$$

The signal-to-noise ratio for phase detection is therefore equal to $N_s/\sigma_{N_s} = \sqrt{N}\Delta\phi$. Assuming that the lowest detectability limit occurs when the SNR is equal to 1, we observe that the phase measurement accuracy is given by $1/\sqrt{N}$. The accuracy limit is commonly referred to as the shot-noise-limit (SNL) or the standard quantum limit. A number of efforts are underway to improve this phase measurement accuracy for highly sensitive interferometer devices like the advanced gravitational wave interferometer. The typically proposed approaches for this purpose involve adding squeezed vacuum state through the second input port of the interferometer. We however note that the whole accuracy analysis above is focused on single detected pixel of the interference signal at the two symmetrically placed point detectors. Measured phase signals are seldom single pixel in nature and as soon as the phase signal has a multi-pixel nature (in space or time), ideas such as image sparsity that we discussed in the first part of the book immediately become applicable. Accuracy improvement beyond SNL is then possible even with classical light interferometers. It is just that now the phase recovery must happen via an optimization algorithm that uses appropriate sparsity constraints. We will discuss this optimization approach to phase recovery from an interferogram in detail next.

14.2.3 Optimization Method for Complex Object Wave Recovery from Digital Holography

The Fourier transform method and the phase shifting method for complex object wave reconstruction are well-established, however, they have certain limitations due to the nature in which these methods approach the problem of information recovery from a digital hologram. For example, while Fourier transform method can perform single-shot holographic imaging, it by default achieves lower resolution than what the digital detector array used for recording the hologram is capable of. Due to sampling considerations, the allowable reference beam angle for off-axis hologram is already quite small and the non-overlap requirement for the dc and cross-term lobes in Fourier domain further reduces the attainable object resolution in this configuration. The phase shifting method regains the detector resolution, however, multiple hologram frames have to be recorded. Further, in terms of photon resources, the phase shifting method has to divide the total number detected photons into four camera records (for arbitrary phase values in $[0, 2\pi]$) and the SNR for each interference record is therefore poor. The optical system has to be stable during the multiple recordings and further the object needs to be stationary during the recording process. High-quality vibration isolation is not always possible in field environment where digital holographic systems may have to be deployed. Single-shot holographic imaging systems have a distinct practical advantage in this regard. The optimization approach discussed in the following is promising in that it can recover phase images with full pixel resolution even with a single-shot operation.

We start by writing the hologram equation Eq. (14.1) in a slightly modified form.

$$|O|^2 + 2|O||R|\cos(\phi_O - \phi_R) + (|R|^2 - H) = 0, \quad (14.15)$$

where $\phi_O(x, y)$ and $\phi_R(x, y)$ are the phase maps associated with the object and reference beams, respectively. For every pixel (x, y) , the above equation is a quadratic in $|O|$ and therefore we can write a solution:

$$|O| = -|R|\cos(\phi_O - \phi_R) \pm \sqrt{H - |R|^2 \sin^2(\phi_O - \phi_R)}. \quad (14.16)$$

At each pixel location (x, y) , one may choose the positive valued solution as we are solving for $|O|$. From the above solution, we note that $|O|$ is dependent on the phase ϕ_O even when the complex-valued reference beam is known via a calibration procedure. Any arbitrary unknown phase function ϕ_O along with the solution $|O|$ can therefore be used to construct a complex object wave which satisfies Eq. (14.1). The performance of any practical device which uses interference phenomenon depends on how well one can recover the complex object wave information. The ambiguity in the solution as discussed above is not desirable and a procedure for arriving at a meaningful solution must be discussed carefully as we will do next.

We have already discussed in Chapter 6 that an appropriate method for image reconstruction in presence of such ambiguities is to formulate the problem as a

constrained optimization problem [16–18]. In particular we consider the problem of minimization of a cost function of the form:

$$C(O, O^*) = ||H - (|R|^2 + |O|^2 + R^*O + RO^*)||^2 + \alpha \psi(O, O^*). \quad (14.17)$$

The first term above is the least square or L2-norm squared error that represents consistency of the solution O with the hologram data H . The reference beam is assumed to be known as is the case with the Fourier transform method or the phase shifting method. The second term $\psi(O, O^*)$ is the constraint term that models some desired property of the solution O depending on the Physics of the problem. For example, if we are recording a Fresnel hologram, we expect the solution for object wave $O(x, y)$ to have a smooth profile whereas in case of an image plane hologram, the solution $O(x, y)$ is expected to retain information about sharp edges in the object to be imaged. The optimization problem above is peculiar in that the cost function to be minimized is real and positive valued but the solution of interest $O(x, y)$ is complex-valued. For optimization purpose, it is necessary to evaluate the functional gradient of the cost function to find the descent direction in which the solution is to be progressed starting with some initial guess. We have already studied one such problem in Chap. 6 and recall that direction of steepest descent for δC is obtained when the variation δO is in the direction of $\nabla_{O^*} C$. The simplest gradient descent iteration for obtaining the solution $O(x, y)$ may now be written as

$$O^{(n+1)} = O^{(n)} - t[\nabla_{O^*} C]_{O=O^{(n)}}, \quad (14.18)$$

where $O^{(n)}$ denotes the guess solution for O at the n th iteration. The Wirtinger derivative $[\nabla_{O^*} C]$ may be evaluated as

$$\nabla_{O^*} C = -2[H - |R + O|^2] \cdot (R + O) - \alpha \nabla_{O^*} \psi. \quad (14.19)$$

The step size t may be selected using standard backtracking line search. Further the image recovery may also be performed by methodologies discussed in Section 6.5, so that, no free parameter α is required for object wave recovery. In the illustration in Fig. 14.9, we first show the effect of α on the phase solution $\arg(O)$. In this numerical example, both $|R|$ and $|O|$ are constants and equal to 1. The same off-axis hologram as in Fig. 14.5a has been used for the iterative reconstruction. Phase map reconstruction for three values of $\alpha = 0.1, 1, 10$ is shown. For penalty function we use the total variation (TV) penalty. From the reconstructions, the solution for regularization parameter $\alpha = 1$ appears to be the most appropriate for a step phase object. For $\alpha = 0.1$, the regularization is not enough and the fringe-like artifacts are present in the phase map. For $\alpha = 10$ there is over-regularization and the solution appears to loose the sharp edge feature of the step phase object. The illustration for $\alpha = 1$ however shows that sharp edge recovery with much higher resolution compared to the Fourier filtering method (see, e.g. Figure 14.5b) is possible via optimization approach. *Full pixel resolution in single-shot should thus be possible to achieve if the complex object wave recovery is formulated as an optimization*

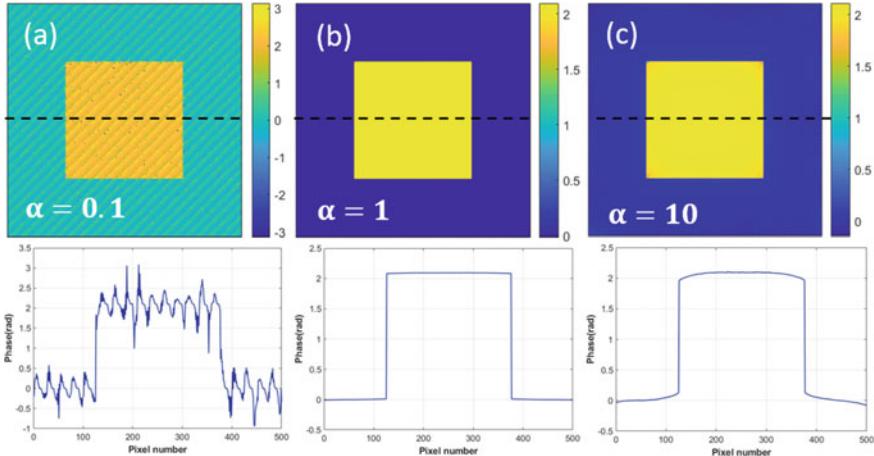


Fig. 14.9 Iterative reconstruction of phase map from single-shot off-axis interferogram. Reconstructions with three different values of regularization parameter $\alpha = 0.1, 1, 10$ are shown in **a**, **b**, **c**, respectively. The profiles of the phase recoveries along the dotted line are also plotted

problem. Similar to the discussion in Section 6.5, the selection of value of α is a tedious process in this example as well. The mean gradient descent (MGD) approach as discussed in the context of the de-convolution can also be applied in the present problem. In Fig. 14.10, we show phase reconstruction for an off-axis as well as on-axis interferogram using the MGD approach [19]. The sharp phase object feature can be reconstructed as seen in Fig. 14.10c, d from single-shot interferogram at full pixel resolution. Interestingly, the optimization procedure can recover the object wave for on-axis case as well, for which there is no solution otherwise except for using a multi-shot phase shifting method. We would however caution the readers that if the fringe density is too low as in Fig. 14.7, the penalties like TV which only consider the next nearest neighbor in their definition may not be the most appropriate for suppressing fringe artifacts and overall the algorithm may take longer to converge to appropriate solution.

14.2.4 Noise Advantage Offered by the Optimization Method

The accuracy of measured phase is very important for sensitive interferometric systems. While we have discussed this topic in Sect. 14.2.2.1, the traditional treatment presented there uses a single-pixel-based methodology for studying accuracy of classical interferometers. One of the ideas that has been explored in detail in the literature is to use the light states that offer sub-Poisson statistics of photon stream for improving phase accuracy beyond the shot noise limit. While single-pixel-based analysis of noise is followed in most texts, we wish to point out that phase signals of inter-

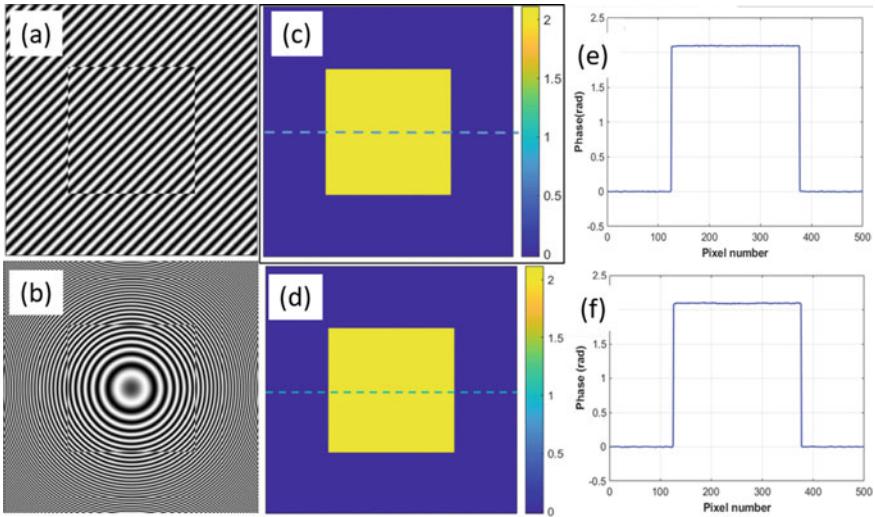


Fig. 14.10 Iterative reconstruction of phase map from single-shot off-axis as well as on-axis interferograms using MGD approach which does not require any regularization parameter α . Note that optimization method can in this case handle on-axis case as well. The profiles of the phase recoveries along the dotted line are plotted. Adapted with permission from [19] ©Optica

est are almost never single-pixel in practice. One may be interested in measuring a time-domain waveform at a single-pixel or in measuring a 2D phase front as is often the case in digital holography experiments. In such cases, the signal of interest will almost always have some structure. The in-built redundancy or sparsity of the phase signal to be reconstructed can then be utilized via an optimization algorithm to obtain RMS phase error that is better than the single-pixel-based shot noise limit. In the numerical illustrations shown in Fig. 14.10, the interferograms have been simulated with Poisson noise and the recovered RMS phase errors are observed to have error performance better than SNL. The noise advantage of utilizing the sparsity of phase map in the reconstruction algorithm can best be appreciated at low light levels where shot noise becomes prominent. In Fig. 14.11b we illustrate an experimental measurement of phase map due to a convex lens object [20]. The lens object is placed in one arm of a Mach-Zehnder interferometer arrangement. The QHQ phase shifter (see Chap. 11) is introduced in reference arm of the interferometer in order to introduce four phase shifts in the reference beam. The interference pattern is recorded on an EMCCD (Electron Multiplying CCD) array sensor which can operate in a photon counting mode. When four interferograms are recorded with high light level (> 5000 photons/pixel), the four interferograms processed using Eq. (14.11) provide a smooth ground truth phase map as shown in Fig. 14.11a. In Fig. 14.11c–e we show three interferograms in the same arrangement recorded by reducing the detected light level to 200, 50, 10 photons/pixel, respectively. In the second and third columns in Fig. 14.11c–e, we show phase reconstruction results obtained using four-step phase

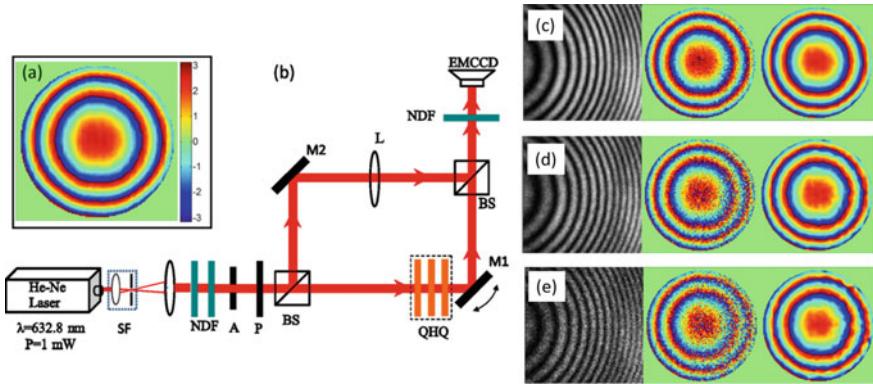


Fig. 14.11 Sub-shot-noise phase measurement using sparsity-based phase retrieval. **a, b** Show ground truth phase map for a lens object when measured using a Mach-Zehnder interferometer. **c–e** Typical interferograms at low light levels with 200, 50 and 10 photons/pixel. Phase reconstructions using four-step phase shifting method and single-shot optimization method are shown in two columns adjoining the interferograms. Photon budget for the two methods is kept equal. The pixel-to-pixel phase variation is not present in the optimization solution. Adapted with permission from [20] ©American Physical Society

shifting method and the optimization method, respectively. The total photon budget in the two methodologies is same, meaning that same photon flux is divided into four frames for the phase shifting method. Here the penalty function used was a sum of the (weighted) squared difference function of the form:

$$\psi(O, O^*) = \sum_{i=\text{all pixels}} \sum_{p,q \in \text{window}} w_{pq} |O_{i,p} - O_{i,q}|^2, \quad (14.20)$$

as this penalty is appropriate for a smooth object like a lens. In the above equation $O_{i,p}$ denotes p th pixel in a window centered on i -the pixel in the interereference record. The weight function $w_{p,q}$ is inversely proportional to the distance between pixels p, q in a window of 5×5 pixels. We observe that the phase shifting method operates individually on a pixel to pixel basis (see Eq. (14.11)) whereas in the optimization method the neighboring pixels interact with each other through the penalty function. As a result, when the light level is reduced, the phase shifting solution gets noisy whereas the optimization solution still remains smooth and similar to the ground truth solution. In this experiment [20], it was observed that at low light level of 10 photons/pixel in the interferogram data, the RMS phase accuracy of the recovered phase map for the optimization solution with respect to the ground truth phase map was more than 5-fold better than the SNL-based phase accuracy which is observed to hold for the phase shifting method. This result indicates the importance of signal sparsity in essentially beating the traditional limits like SNL for realistic applications where the phase map to be measured is often expected to have sparsity. Finally, we

note that the sparsity ideas are independent of the nature of light source used and therefore can further enhance performance of an interferometer that uses a non-classical light state with sub-Poisson photon statistics.

14.3 Digital Holographic Microscopy

The ability of measurement of quantitative phase information has found important applications in microscopic imaging of label-free or unstained cells [21]. Most light microscopy in bio-science laboratories or pathology clinics is currently performed using intensity-based imaging modalities such as bright-field or fluorescence. The contrast mechanism in these methods are dyes used for staining of cells or fluorescence agents that are used to tag particular parts of a cell. While these methodologies have been popular, they involve wet lab processing of cells that affects their natural state. Quantitative phase imaging offers an attractive alternative that uses the natural refractive index of cells as a contrast mechanism. When a plane laser beam is transmitted through a transparent cell, under weak scattering assumption, the phase of the transmitted beam is given by

$$\phi(x, y) = \frac{2\pi}{\lambda} \int dz n(x, y, z). \quad (14.21)$$

As depicted in Fig. 14.12, a plane wave passing through a cell sample undergoes a spatially dependent phase change as in this equation. Here $n(x, y, z)$ stands for the relative refractive index of the cell with respect to the surrounding medium, which is in general a function of coordinates (x, y, z) . The microscope objectives and the tube lens combine as depicted in Fig. 14.12 make a 4F-like (infinity corrected) imaging system in both the arms of the interferometer which is a transmission version of the Linnik type interferometer. A nominal image plane hologram for blood cell sample is shown for illustration. The off-axis hologram recorded in this balanced Mach-Zehnder interferometer configuration consists of straight line fringes which show bending at the location of cells indicating a phase modulation. With the prior discussion in this chapter, it is clear that the single-shot optimization method offers important advantage in terms of resolution as well as phase accuracy. We note that the optimization method fully operates in the image domain and as a result can also be used for region-of-interest (ROI) reconstruction from an image plane hologram. Phase reconstruction of an individual red blood cell is also shown in Fig. 14.12 as a surface plot where the color coding refers to the optical path difference experienced by the transmitted laser beam. The optimization method used for this illustration has used Huber penalty (see Problem 6.2). More details on the topic of ROI phase reconstruction from an image plane hologram can be found in [22]. This depth information is not available with commonly used bright-field microscopes. We note that the contrast provided by bright-field microscopes is due to absorption of light in stained cells. On the other hand, the quantitative phase information $\phi(x, y)$ as in Eq.

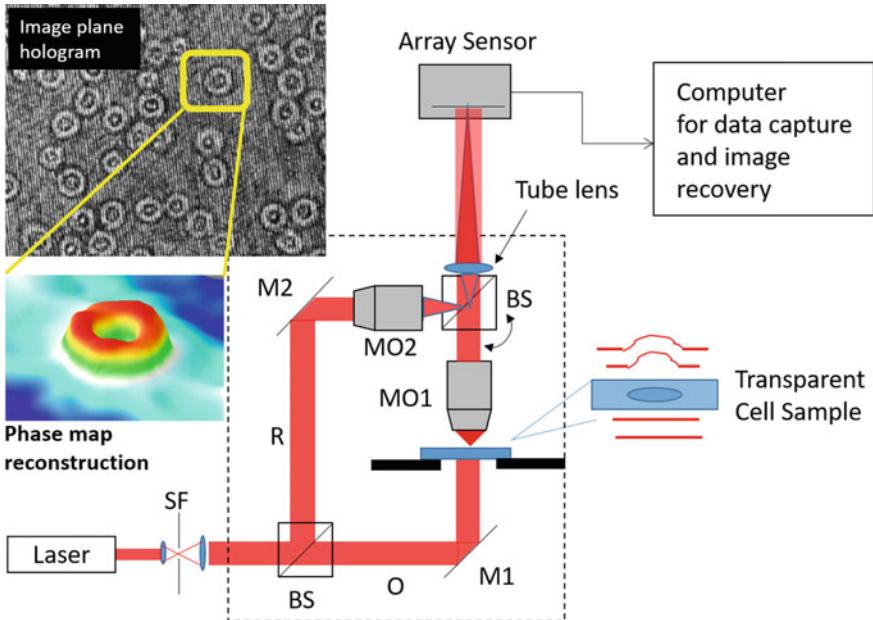


Fig. 14.12 A typical balanced digital holographic microscope (DHM) arrangement employing a Mach-Zehnder arrangement is shown, SF: spatial filter, M1, M2:mirrors, MO1, MO2: Microscope objectives, BS: Beamsplitter, R: Reference beam, O: Object beam. The curved arrow near the second beamsplitter shows how R beam may be tilted to introduce off-axis beam angle. Adapted with permission from [22] ©Taylor & Francis Group

(14.21) as provided by a DHM instrument depends on the locally varying refractive index of cells which depends on their chemical properties. DHM methodology is therefore useful in studying small refractive index changes that may happen in cells leading to phase changes. The accurately measured phase change can in fact serve as an important parameter for cell image-based pathological diagnosis [23]. While this methodology is not yet commonly used in clinics, the sensitive phase imaging methodology has great potential for future application development. From basic biosciences perspective, DHM offers an important modality for studying live cells in their natural state as no staining or fluorescent labeling of cells is required in this methodology.

Several variants of DHM system depicted in Fig. 14.12 have been explored in the literature. It is possible to use unbalanced interferometer systems with a single microscope objective in the object beam arm. Such a system design however requires compensation of aberrations due to the microscope objective from the recovered phase map [27, 28]. Compact 3D printed or mobile phone camera-based DHM system designs have also been studied by now and offer attractive possibility of low cost point-of-care diagnostic devices. It is important to remember that apart from phase reconstruction, a practical DHM system also requires additional algorithms

for auto-focusing, phase unwrapping, etc. to make a complete system. A useful phase unwrapping algorithm based on the transport of intensity equation has already been discussed in Sect. 9.5.1 which can be readily used with a DHM system. The auto-focusing in digital holography involves propagating the complex object field $O(x, y)$ to nearby planes where the definition of object features may be sharper thus correcting for any de-focus error in recording of the hologram. One of the powerful criteria for this purpose is to use the sparsity-of-gradient criterion. In particular, the recovered object field $O(x, y)$ may be propagated to various z distances as follows:

$$O_z(x, y) = O(x, y) * h_{AS}(x, y, z), \quad (14.22)$$

where $h_{AS}(x, y, z)$ is the angular spectrum impulse response (computation above is preferably performed using transfer function and FFTs). A gradient magnitude image $G = |\nabla_{\perp} O_z(x, y)|$ is computed and a sparsity measure like the Tamura coefficient (TC) defined as

$$TC = \sqrt{\frac{\sigma_G}{\langle G \rangle}}, \quad (14.23)$$

is then used to determine the best focus plane [26]. In the above equation, σ_G and $\langle G \rangle$ stand for the standard deviation and the mean of the gradient magnitude image G , respectively. In the best focus plane, the TC measure is usually observed to show a distinct peak. For unstained phase objects, best focus plane can also be located simply by observing the image plane hologram [29] as the sample is de-focused sequentially through focus. At the best focus plane the fringes in the hologram predominantly show phase modulation and minimal amplitude modulation. Such a technique can be used for deciding the focus plane while recording the hologram and provides a similar user experience like that with a bright-field microscope, particularly for bio-science researchers. We will discuss the topic of auto-focusing further in the context of in-line holography of particulates in the next section.

14.4 In-line Digital Holography of Particulates or Weak Scattering Objects

In-line digital holography can be a very useful tool when localization and tracking of particulates or other small objects like micro-organisms in a liquid are of interest. The interesting feature of an in-line holography system is its compact nature. Further, since the illumination and the scattered wave essentially follow the same path, holography can be performed with a source of somewhat lower temporal coherence. Given a recorded in-line hologram as shown in Fig. 14.2b, the main task for reconstruction problem is to propagate the hologram back to the original object volume. This may for example be performed using the angular spectrum propagation approach explained in Chap. 9. A back-propagation of the recorded hologram to two different distances

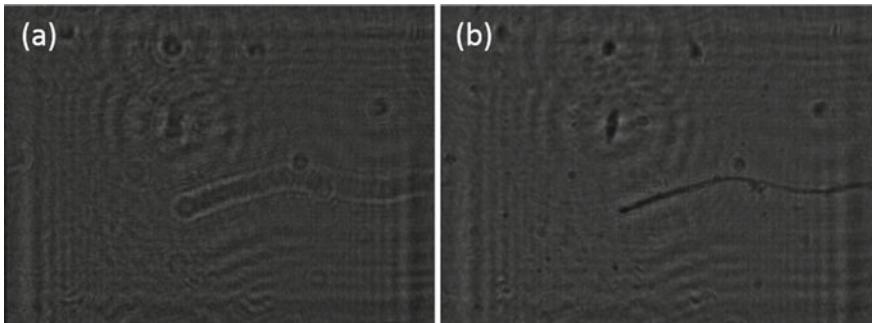


Fig. 14.13 a, b Computed intensity pattern obtained by numerical back-propagation of the in-line hologram shown in Fig. 14.2b to two different distances

is shown in Fig. 14.13. We clearly observe that in Fig. 14.13b, a number of features (a thread-like elongated objects and point-like objects) are sharp and focused. While propagating the inline hologram back, one needs to take into account the system magnification given by $M = (z_1 + z_2)/z_1$, where z_1 is the source to object distance and z_2 is the object to sensor distance. If an object is originally located at a nominal distance of z_2 from the sensor, then on back-propagation, the best focus distance for that object will be Mz_2 . Depending on the known bounding box where the sample is located (e.g. size of a cuvette), the back-propagation process can be limited to a range of distances. For opaque objects, sometimes a minimum intensity criterion is found to be useful for determining the best focus plane. The distance z_2 can be determined by applying focusing criterion like the Tamura coefficient discussed in previous section (see Eq. (14.23)) for automating the focus detection process. A number of auto-focusing criteria in in-line digital holography have been studied comparatively in the literature [30]. We wish to mention here that while the small weak scattering objects do appear to focus in the simple back-propagation-based reconstruction as shown in Fig. 14.13b, this reconstruction is actually corrupted by the twin image which is inevitable in the back-propagation process. This twin image can be eliminated if the reconstruction problem is formulated as an optimization problem employing object sparsity [31, 32]. We will discuss this topic more in the next Section where we discuss the possibility of a tomographic reconstruction in digital holography for sparse objects.

14.5 True 3D Image Reconstruction in Digital Holography

The early work in holography was mostly conducted using high-resolution photographic films for recording of holograms. Holographic image reconstruction was also mainly performed for the purpose of display holography or visual inspection by illuminating the recorded holographic film with the conjugate reference beam

as described in Eq. (14.2). In the modern era of digital holography the hologram is recorded on a digital array sensor and therefore the image reconstruction has to be performed numerically. The typical procedure of 3D image reconstruction in this case is to follow a two step procedure. The first step involves recovering the complex object wave $O(x, y)$ in the hologram recording plane. This field may then be back-propagated to original object volume as done in the previous section on in-line holography. It may be noted that unlike the back-propagation of in-line hologram, the back-propagation of the object wave $O(x, y)$ does not give rise to any twin image. The nature of the 3D reconstruction however needs more discussion.

We may examine the 3D reconstruction problem from the point of view of degrees of freedom in the 3D volume and the recovered object field $O(x, y)$ in the hologram plane. It is clear that recording of a hologram is a process of transferring information from a 3D volume onto a 2D detector. In that process we now have lesser number of pixel values in the detector plane compared to the number of 3D voxels in the object volume/ bounding box. As discussed in Sect. 14.2, the 2D complex-valued object field $O(x, y)$ is recovered from the recorded hologram(s) in the detector plane with full pixel resolution. Now the process of forming 3D image as in the traditional holographic replay is to simply back-propagate this field to the original object volume which is a linear operation. It is therefore surprising as to how one is able to generate more information in a 3D volume in the holographic replay process. *An important question to ask is whether this holographic replay (or back-propagation) of the numerically recovered object field $O(x, y)$ truly constitutes an inverse operation leading to a 3D reconstruction.* It may be emphasized here that the perception of focused object in physical hologram replay process is aided by focusing capability of observer's eyes that are absent when we are simply reconstructing the 3D object by numerical back-propagation of the object field from hologram plane to the original object volume.

We will consider a Fresnel zone hologram for discussion and introduce a simple model for total object field $O(x, y)$ that may result from a 3D object as shown in Fig. 14.14. Assuming a reflective object, we may divide the object function $U(x', y', z')$ into multiple slices along z -axis. The total object field may then be described by propagating field in the individual slice up to the hologram plane and then integrating all the propagated fields to get the object field $O(x, y)$. Here we have neglected secondary scattering from the object which is a reasonable approximation for the present purpose and has been used in holography traditionally [33]. The total object field may be expressed by an operator \hat{A} acting on the 3D object function $U(x', y', z')$:

$$\begin{aligned} O(x, y) &= \hat{A}U \\ &= \int dz' e^{ik(z'-z_d)} \iint dx' dy' U(x', y', z') h_F(x - x', y - y', z_d - z'). \end{aligned} \quad (14.24)$$

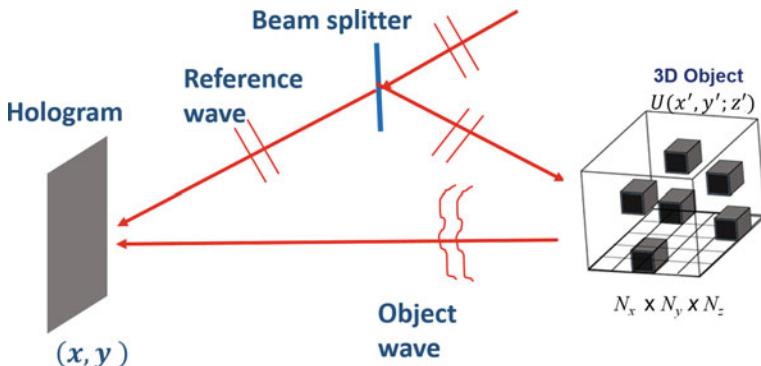


Fig. 14.14 Schematic explaining a simple model for hologram formation due to a 3D object described by $U(x', y', z')$. Adapted from [25] ©Institute of Physics, UK

Here, $h_F(x, y, z)$ stands for Fresnel propagator, z_d refers to the z -coordinate of the detector or the hologram recording plane. The additional phase factor $\exp[ik(z' - z_d)]$ arises since the illuminating wave has different phases for individual object planes. Further we would like to point out that the integral over z' is to be evaluated over the extent of the bounding box containing the 3D object. The interference of this total object field $O(x, y)$ and a reference field $R(x, y)$ is recorded. However the aim of the numerical processing methods in digital holography is to recover the object field $O(x, y)$ from the recorded hologram pattern. The accurately determined object field may therefore be treated as a starting point for our further analysis. Note that physical hologram replay involves using the recorded hologram pattern and therefore the replay field has dc and twin image terms. If the starting point is the numerically recovered object field $O(x, y)$, then the dc or twin terms do not arise in the computed back-propagated field. Next we may determine the adjoint operator \hat{A}^\dagger corresponding to the forward model in Eq. (14.24) by utilizing the scalar product identity:

$$\langle O_2, \hat{A}U_1 \rangle_{2D} = \langle \hat{A}^\dagger O_2, U_1 \rangle_{3D}. \quad (14.25)$$

In the above equation, $O_2(x, y)$ and $U_1(x, y, z)$ refer to an arbitrary 2D object field (at the detector) and a 3D object function, respectively. The suffixes $2D$ and $3D$ on the scalar products $\langle \dots, \dots \rangle$ refer to the fact that the scalar products have been evaluated on the 2D detector coordinates or the 3D volume coordinates. Writing out the above scalar product relation explicitly and rearranging the order of integration leads to the following interesting result:

$$[\hat{A}^\dagger O](x', y', z') = e^{-ik(z' - z_d)} \iint dx dy O(x, y) h_F^*(x - x', y - y', z_d - z'). \quad (14.26)$$

Apart from the phase factor in front, the above expression for the adjoint operator is precisely the back-propagation of the object field $O(x, y)$ to the original object volume as achieved by traditional holographic replay. We therefore conclude that *the traditional holographic replay corresponds to Hermitian transpose corresponding to the forward hologram formation model* in Eq. (14.24). The numerical back-propagation of $O(x, y)$ therefore does not truly reconstruct the original 3D object but provides an approximation to it. When physically replaying a hologram, the generation of $[\hat{A}^\dagger O]$ as an approximation to $U(x', y', z')$ is inevitable, however, in numerical 3D reconstruction, more choices may be available. The 3D reconstruction may for example be formulated as a problem of minimization of the cost function:

$$C(U, U^*) = \|O - \hat{A}U\|_2^2 + \alpha \psi(U, U^*). \quad (14.27)$$

As in Chapter 6, the constraint function could for example represent sparsity of the 3D object function $U(x', y', z')$. In the illustration next, we illustrate the 3D reconstruction via optimization as above for an object consisting of four letters A, B, C, D (see Fig. 14.15a) which is a sparse object in the 3D volume of interest. A hologram of this object is recorded and the corresponding object field magnitude on back-propagation of the object wave to the object volume is shown in Fig. 14.15b. As expected, as the object wave propagates from hologram plane back to the object volume, focusing effect is observed at the original location of the letters A, B, C, D . But since the four letters were in different planes initially, back-propagated field in any given plane now consists of a focused field (if an object was present in that plane) along with the de-focused background due to objects in the other planes. Human observers naturally pay more attention to the focused fields and ignore the de-focused background. But from a numerical standpoint we see that the reconstructed field is quite different from the original object field as in Fig. 14.15a. In Fig. 14.15c, we show the reconstruction of the 3D object field that results from solution of the optimization problem

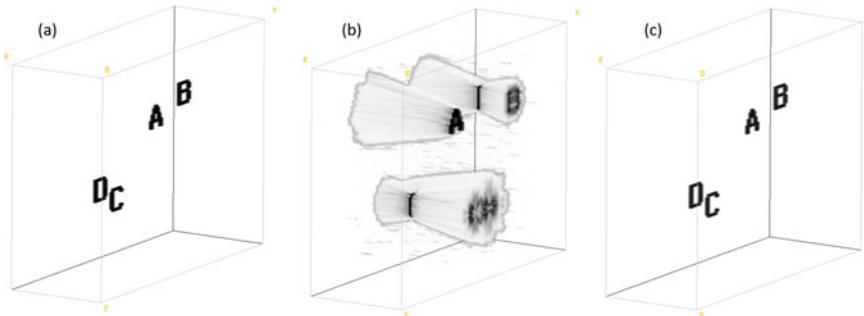
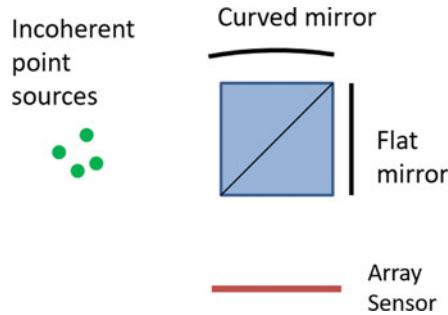


Fig. 14.15 **a** Ground truth 3D object, **b** Reconstructed field magnitude by simple back-propagation of the object field $O(x, y)$ from hologram plane back to the original object volume, **c** Reconstruction of the 3D image via solving a minimization problem for cost function in Eq. (14.27). Adapted from [25] ©Institute of Physics, UK

in Eq. (14.27). This “true 3D” reconstruction is much superior as the object field now predominantly exists only where the letters A , B , C , D were originally located. In this illustration a total variation penalty has been used for encouraging a sparse solution. More details on this illustration may be found in [25]. We emphasize that a CAD-model like 3D reconstruction as observed in Fig. 14.15c is not possible with any physical hologram replay process in which the reconstruction field cannot possibly vanish at slices where no object was originally present. Such reconstruction is thus possible only with digital holography. Digital holographic 3D image reconstruction does not necessarily have to mimic the physical hologram replay process as is a common practice. Instead optimization approach can inherently provide a superior 3D solution. Such true 3D reconstruction may not however be possible for non-sparse objects from degrees-of-freedom considerations. In such cases, additional diversity mechanisms may be required for measuring non-redundant data. The question of whether true 3D reconstruction of a general object is possible is however not just dependent on degrees-of-freedom considerations. The assumption about small secondary scattering that has been used here may not be valid as the 3D object gets more and more complex in its structure. Nevertheless the iterative approach presented here may be valid for 3D reconstruction of low complexity objects such as transparent cells or optical elements like lenses, etc. from their 2D holograms/interferograms. Finally while it is somewhat out of scope in the present discussion, the question of how many holographic measurements are required for an object of given sparsity is not answered fully in our opinion, although partial answers are available based on the literature on diffraction tomography or what is also sometimes referred to as holo-tomography [34].

In summary, this chapter considered the problem of holographic image recording by means of interference phenomena. A number of methods for recovering complex object field from single (or multiple) hologram(s) were discussed with attention to spatial resolution and phase measurement accuracy offered by them. Applications of digital holography to microscopy were considered. Finally some interesting questions regarding possibility of a true 3D/tomographic image reconstruction from hologram data was also considered briefly.

Fig. 14.16 System for holography with incoherent light



Problems

14.1 Holography with spatially incoherent illumination Consider multiple incoherent point sources as in a fluorescence microscopy sample (Fig. 14.16). The radiation from these sources is passed through a Michelson interferometer system which has a flat mirror in one arm and a curved mirror in another arm such that the path lengths along the optic axis are matched. Obtain an expression for the pattern formed on the array sensor at the output port of the interferometer. Simulate the pattern by assuming reasonable system geometry. Explain how this hologram pattern can be utilized to form a tomographic image of the original object. For an interesting twist on how multiple measurements of this nature can be treated as a single hologram with spatially coherent illumination, refer to [24].

14.2 Fringe projection profilometry: Tactile diagrams are commonly used for teaching visually impaired children. These diagrams are printed on a specialized polymer sheet such that the features protrude out of the sheet and can be understood by touch. In order to measure height profile of the tactile diagram, a sinusoidal fringe pattern is projected onto the tactile map surface and the surface is then imaged (see Fig. 14.17). Explain what computational and calibration steps will be required to recover the feature height information.

14.3 Spherical waves from an on-axis point source located at $z = 0$ and an in-line plane wave traveling in the $+z$ -direction (assume that the two sources are derived from the same laser beam and thus have required coherence properties) are used to form a hologram in plane $z = z_0$. The hologram was recorded using illumination wavelength λ_1 nm but for reconstruction purpose, a plane wave traveling in negative z direction and having a wavelength λ_2 nm is used. Obtain an expression for the hologram intensity pattern and simulate the pattern in computer. Find the location of the reconstructed real image.

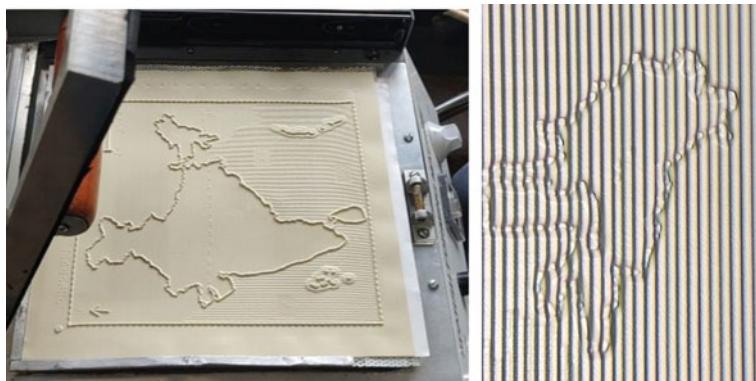


Fig. 14.17 Tactile diagram of a map and part of map with fringe projection

14.4 Fringe demodulation using Hilbert transform: Consider an off-axis hologram with straight line fringes oriented at an arbitrary angle in the camera record. The low frequency terms in the hologram may be estimated by low-pass filtering of the recorded fringe pattern and removed. Describe how the spiral phase transform (using the $\exp(i\phi)$ or spiral phase filter in the Fourier domain) is helpful for fringe demodulation process.

14.5 Traditionally an in-line hologram is processed by simple back-propagation as illustrated in Fig. 14.13. This method however has a problem that the reconstructed object consists of a background field associated with the intensity and twin terms in the hologram. Formulate an iterative optimization procedure in order to get rid of these extra terms so that the reconstruction of the object in the appropriate focus plane is quantitatively meaningful.

References

1. D. Gabor, A new microscopic principle. *Nature* **161**, 177–178 (1948)
2. E.N. Leith, J. Upatnieks, Reconstructed wavefronts and communication theory. *J. Opt. Soc. Am.* **52**, 1123–1128 (1962)
3. M.K. Kim, Principles and techniques of digital holographic microscopy. *SPIE Rev.* **1**, 018005-1–018005-50 (2010)
4. B. Javidi, A. Carnicer, A. Anand, G. Barbastathis, W. Chen, P. Ferraro, J.W. Goodman, R. Horisaki, K. Khare, M. Kujawinska, R.A. Leitgeb, P. Marquet, T. Nomura, A. Ozcan, Y. Park, G. Pedrini, P. Picart, J. Rosen, G. Saavedra, N.T. Shaked, A. Stern, E. Tajahuerce, L. Tian, G. Wetzstein, M. Yamaguchi, Roadmap on digital holography [invited]. *Opt. Express* **29**, 35078–35118 (2021)
5. J.W. Goodman, R.W. Lawrence, Digital image formation from electronically detected holograms. *Appl. Phys. Lett.* **11**, 77–79 (1967)
6. H. Chen, M. Shih, E. Arons, E.N. Leith, J. Lopez, D. Dilworth, P.C. Sun, Electronic holographic imaging through living human tissue. *Appl. Opt.* **33**, 3630–3632 (1994)
7. U. Schnars, W.P. Juptner, Direct reconstruction of holograms by a CCD target and numerical reconstruction. *Appl. Opt.* **33**, 179–181 (1994)
8. E. Cuche, F. Bevilacqua, C. Depeursinge, Digital holography for quantitative phase contrast imaging. *Opt. Lett.* **24**, 291–293 (1999)
9. E. Cuche, P. Marquet, C. Depeursinge, Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms. *Appl. Opt.* **38**, 6994–7001 (1999)
10. M. Liebling, T. Blu, M. Unser, Complex-wave retrieval from a single off-axis hologram. *J. Opt. Soc. Am. A* **21**, 367–377 (2004)
11. L. Yaroslavsky, *Digital Holography and Digital Image Processing* (Kluwer Academic, 2010)
12. E. Wolf, Determination of the amplitude and the phase of scattered fields by holography. *J. Opt. Soc. Am. A* **60**, 18–20 (1970)
13. M. Takeda, H. Ina, S. Kobayashi, Fourier transform method of fringe pattern analysis for computer based topography and interferometry. *J. Opt. Soc. Am.* **72**, 156–160 (1982)
14. M. Singh, K. Khare, Accurate efficient carrier estimation for single shot digital holographic imaging. *Opt. Lett.* **41**, 4871–4874 (2016)
15. I. Yamaguchi, T. Zhang, Phase shifting digital holography. *Opt. Lett.* **22**, 1268–1270 (1997)

16. K. Khare, P.T.S. Ali, J. Joseph, Single shot high resolution digital holography. *Opt. Express* **21**, 2581–2591 (2013)
17. P.T.S. Ali, K. Khare, J. Joesph, Quantitative phase imaging with single shot digital holography. *Opt. Commun.* **319**, 85–89 (2014)
18. M. Singh, K. Khare, Single-shot interferogram analysis for accurate reconstruction of step phase objects. *J. Opt. Soc. Am. A* **34**, 349–355 (2017)
19. S. Rajora, M. Butola, K. Khare, Mean gradient descent: an optimization approach for single-shot interferogram analysis. *J. Opt. Soc. Am. A* **36**, D7–D13 (2019)
20. M. Singh, K. Khare, A.K. Jha, S. Prabhakar, R.P. Singh, Accurate multipixel phase measurement with classical light interferometry. *Phys. Rev. A* **91**, 021802(R) (2015)
21. Y.K. Park, C. Depeursing, G. Propescu, Quantitative phase imaging in biomedicine. *Nat. Photon.* **12**, 578 (2018)
22. M. Singh, K. Khare, Single-shot full resolution region-of-interest (ROI) reconstruction in image plane digital holographic microscopy. *J. Mod. Opt.* **65**, 1127–1134 (2018)
23. J. Mangal, R. Monga, R. Mathur, A.K. Dinda, J. Joseph, S. Ahlawat, K. Khare, Unsupervised organization of cervical cells using bright-field and single-shot digital holographic microscopy. *J. Biophotonics* **12**, e201800409 (2019)
24. J. Rosen, G. Brooker, Digital spatially incoherent Fresnel holography. *Opt. Lett.* **32**, 912–914 (2007)
25. J. Birdi, S. Rajora, M. Butola, K. Khare, True 3D reconstruction in digital holography. *J. Phys. Photonics* **2**, 044004 (2020)
26. Y. Zhang, H. Wang, Y. Wu, M. Tamamitsu, A. Ozcan, Edge sparsity criterion for robust holographic autofocusing. *Opt. Lett.* **42**, 3824–3827 (2017)
27. E. Sánchez-Ortiga, A.I. Doblas, G. Saavedra, M. Martínez-Corral, J. García-Sucerquia, Off-axis digital holographic microscopy: practical design parameters for operating at diffraction limit. *Appl. Opt.* **53**(10), 2058–2066 (2014)
28. A.I. Doblas, E. Sánchez-Ortiga, M. Martínez-Corral, G. Saavedra, J. García-Sucerquia, Accurate single-shot quantitative phase imaging of biological specimens with tele-centric digital holographic microscopy. *J. Biomed. Opt.* **19**, 046022 (2014)
29. R. Malik, P. Sharma, S. Poulose, S. Ahlawat, K. Khare, A practical criterion for focusing of unstained cell samples using a digital holographic microscope. *J. Microsc.* **279**, 114–122 (2020)
30. T. Zimmerman, A. Park, S. Bianco, Evaluating autofocusing metrics for plankton holographic microscope image reconstruction. *Proc. SPIE* **11710**, 117100A (2021)
31. F. Soulez, L. Denis, C. Fournier, É. Thiébaut, C. Goepfert, Inverse-problem approach for particle digital holography: accurate location based on local optimization. *J. Opt. Soc. Am. A* **24**, 1164–1171 (2007)
32. L. Denis, D. Lorenz, E. Thiébaut, C. Fournier, D. Trede, Inline hologram reconstruction with sparsity constraints. *Opt. Lett.* **34**, 3475–3477 (2009)
33. D. Gabor, Theory of electron interference experiments. *Rev. Mod. Phys.* **28**, 260 (1956)
34. V. Balasubramani, A. Kuś, H.Y. Tu, C.J. Cheng, M. Baczevska, W. Krauze, M. Kujawińska, Holographic tomography: techniques and biomedical applications. *Appl. Opt.* **60**, B65–B80 (2021)

Chapter 15

Non-interferometric Phase Retrieval



Non-interferometric phase retrieval from intensity measurements is one of the most challenging computational imaging problems. In Chap. 14, we discussed interferometric approach to phase measurement. There are however practical situations (e.g. X-ray imaging, imaging through turbulent atmosphere) where setting up an interferometer is not always easy [1]. In Chap. 9, we already discussed one non-interferometric phase imaging technique—the transport of intensity equation (TIE)—which relates the longitudinal intensity derivative of a diffraction field (in Fresnel zone) to the transverse phase gradient. The solution of the problem was studied in that case with a direct Fourier transform-based method. In this chapter, we will describe non-interferometric iterative techniques that allow one to reconstruct an object from the absolute magnitude of its Fourier or Fresnel transforms when some additional constraints are imposed on the object. The Fourier phase retrieval problem may be stated as follows: *Determine an unknown object function $g(x, y)$ from its Fourier intensity $|G(f_x, f_y)|^2$ data.* Clearly any randomly selected phase function when associated with the magnitude $|G(f_x, f_y)|$ will correspond to a solution. However, the challenging task in this problem is to find a solution that satisfies physically meaningful properties that may be defined mathematically in the form of constraints. The constraints represent *a priori* information about the object. For example, the measured Fourier intensity data is a Fourier domain constraint. While the object/image domain constraints are object support defining the extent of object, positivity and reality(in case of amplitude-only object). The iterative phase retrieval procedures sequentially enforce the constraints in both the Fourier transform domain and the spatial domain until the solution is stabilized. The Fourier phase retrieval problem occurs naturally in study of Fraunhofer diffraction, where the measured diffraction intensity pattern $|G(f_x, f_y)|^2$ is not sufficient to reconstruct object due to lack of phase information. This particular case naturally arises in X-ray diffraction studies, where the recorded intensity pattern is proportional to the absolute squared magnitude of the Fourier transform of the electronic density in the crystal structure under

investigation. Another case where the phase retrieval problem occurs is that of astronomical imaging through turbulent atmosphere. For example, in the technique known as speckle interferometry, the short exposure images of a fixed star object $o(x, y)$ are recorded sequentially for multiple realizations of atmospheric turbulence. The resultant set of images may be described by

$$g_k(x, y) = o(x, y) * p_k(x, y), \quad (15.1)$$

where $p_k(x, y)$ denotes the combined PSF of the telescope system and the atmospheric turbulence effects. The name speckle interferometry arises from the fact that the individual short exposure images (short compared to scale of atmospheric turbulence fluctuation) have a speckle-like structure due to the atmosphere-induced random phase changes. If the images $g_k(x, y)$ are averaged, the averaged image will simply be equivalent to long-time averaging over atmospheric turbulence and the resultant image will have a large blur obscuring the star image. Instead, as first proposed by Labeyrie in [2], the averaging may be performed in the Fourier domain as

$$\langle |G_k(f_x, f_y)|^2 \rangle = |O(f_x, f_y)|^2 \langle |P_k(f_x, f_y)|^2 \rangle. \quad (15.2)$$

The quantity $\langle |P_k(f_x, f_y)|^2 \rangle$ may be measured separately by imaging a nearby unresolved (delta function like) guide-star object that sees similar atmospheric turbulence statistics. The result is that one obtains the Fourier transform intensity $|O(f_x, f_y)|^2$ corresponding to the object $o(x, y)$ of interest. The phase retrieval techniques may then be employed in order to compute the image of the star. A number of phase retrieval problems of this nature can be addressed by using iterative algorithms [3]. Some well-known algorithms such as the Error Reduction (ER) algorithm and the Fienup hybrid input-output algorithm will be discussed in detail in the following sections. Additionally, we will present how image sparsity ideas play an important role in phase retrieval algorithms.

15.1 Lensless Coherent X-ray Diffraction Imaging

Phase retrieval also finds an important application in X-ray coherent diffraction imaging (CDI). CDI is a lensless imaging technique in which a highly coherent X-ray beam scattered by an isolated specimen is measured on a detector at a far-field distance as shown in Fig. 15.1. From the measured far-field or Fourier intensity pattern, the object is then reconstructed back using phase retrieval methods. The history of CDI is even more older than phase retrieval in optics and imaging which dates back to 1952 with Sayre's work in crystallography [4]. It was in 1999 when the reconstruction from an experimental X-ray data for

the non-crystalline specimen was reported for the first time [5] which opened up the new field of coherent diffraction imaging [6]. After that the experimental hardware for CDI has improved at a rapid rate over the last two decades [7]. Historically, the phase retrieval methods have been developed in optics (or imaging) and crystallography quite independently as both the domains have many differences. However, since both optics and crystallography share some commonalities, the phase retrieval algorithms developed in optics like Fienup's HIO algorithm have been widely used for the crystallographic and CDI applications. An excellent review of phase retrieval in crystallography and optics can be found in [8].

The recording of coherent X-ray diffraction data is itself a challenging task that is being carried out by a number of dedicated instruments [7, 9]. Reconstructing stable solutions from such a data remains a challenging and fundamental task for the iterative phase retrieval algorithms. A coherent X-ray imaging record has missing or unreliable pixels due to detection noise and detector beam stop. The detector beam stop is usually applied to block the very high-intensity pixels near central frequency region so that the high-frequency signals can be accommodated within the dynamic range of the detector. The reconstruction of a unique and reliable solution from single incomplete diffraction data cannot be guaranteed by single run of the phase retrieval algorithms. Therefore, the algorithm is run over hundreds of times starting with random initial guesses to obtain multiple trial solutions. Finally, the solutions with least error are averaged together for the final solution recovery [10].

Coherent X-ray imaging is one of the powerful techniques for imaging of nano and micro-sized particles and has wide applications in structural biology, material science, chemistry, medicine, etc. Recent advancement in the new generation of X-ray sources called X-ray free electron lasers (XFELs) can now produce extremely short pulses ranging from tens of attosecond to less than hundreds of femto seconds [11]. With such short duration and intense pulses, the measurement of diffraction pattern of any biological specimen has become possible before the pulse actually damages the sample [12]. This has now enabled the entire new emerging science of single-particle imaging (SPI) or live imaging of cells, organelles, viruses, etc.

15.2 Error Reduction Algorithm

We will discuss some of the main ideas in phase retrieval starting with the Error Reduction (ER) algorithm. The ER algorithm originally used in electron microscopy requires two intensity measurements for phase retrieval [13] which is known as the Gerchberg-Saxton (GS) algorithm. However, the basic theme of the GS algorithm was practically found to be more important for problems involving single Fourier intensity measurement [14]. This version of GS algorithm is popularly known as

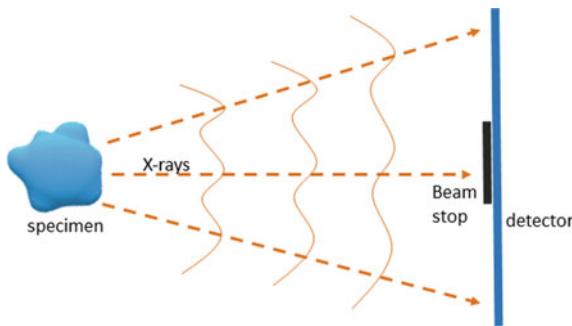


Fig. 15.1 Schematic diagram of lensless coherent X-ray diffraction imaging

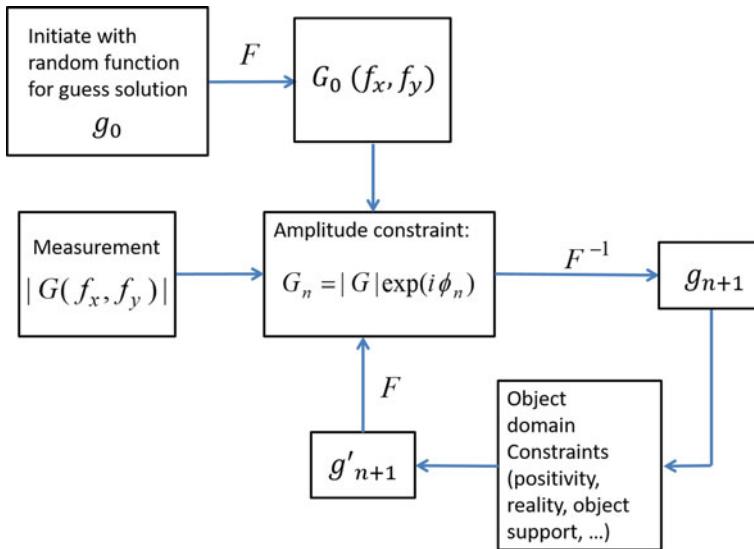


Fig. 15.2 Schematic diagram showing steps in the Error Reduction method for phase retrieval

an error reduction (ER) algorithm. The basic structure of the ER algorithm may be described by means of the flowchart in Fig. 15.2. Given the Fourier modulus data $|G(f_x, f_y)|$, an initial guess for the function $g(x, y)$ is made typically as a matrix of real or complex random numbers. Steps in the n -th iteration of the algorithm are as follows:

1. The function $g_n(x, y)$ is Fourier transformed to obtain a guess transform $G_n(f_x, f_y)$.
2. The amplitude of the transform $G_n(f_x, f_y)$ is replaced by the measurement $|G(f_x, f_y)|$ retaining its phase as is.
3. The modified Fourier transform is inverse transformed to get a guess g_{n+1} for the solution.

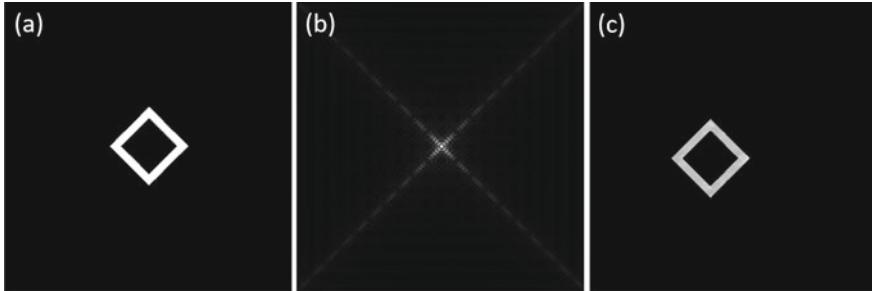


Fig. 15.3 Illustration of ER iteration for phase retrieval **a** Object, **b** its Fourier magnitude and **c** recovered solution after 500 GS iterations

4. Any prior information about the object is applied to the guess solution $g_{n+1}(x, y)$ to obtain a modified guess solution $g'_{n+1}(x, y)$. The constraints of interest may include positivity of the solution $g(x, y)$, information about object support, image sparsity, etc. The image pixels not satisfying the desired constraint are set to zero.
5. The modified solution is Fourier transformed to obtain the next guess $G_{n+1}(f_x, f_y)$ and one iteration is thus complete.

The iterative process can be stopped when the error norm $E = || |G| - |\mathcal{F}[g]| ||^2$ is below a threshold value. It was shown by Fienup in [15] that the iteration described above is equivalent to a steepest descent step for reduction of the error norm E . The illustration in Fig. 15.3 shows an example of phase retrieval of an amplitude binary object function using the ER iterative process. The solution shown here required 500 ER iterations. It may be noted that the diamond shape is translated in the recovered image. Translation of the object in this manner is an ambiguity which cannot be resolved as the original and translated objects have the same Fourier magnitude. The illustration for ER iteration in Fig. 15.3 used reality, positivity and support constraints as the object to be reconstructed is an amplitude-only object. In order to employ positivity constraint, one may retain only the positive numerical values in the solution within the support region. All other pixels (those outside the support and those with negative real parts within the support) are turned to zero in each iteration. While the ER iteration generally works well for simple amplitude binary objects as shown in Fig. 15.3a, the convergence to a reasonable solution is impractically slow when a phase object or an object with more detailed features is used. The iterative process in case of such objects is seen to stagnate to undesirable solutions and even large number of iterations do not make much change to the solution. The powerful hybrid input-output algorithm discussed next can significantly enhance the convergence to a desired solution as we will show next.

15.3 Hybrid Input-Output (HIO) Algorithm

An important advance over the ER method which is far superior in terms of handling the stagnation issues in phase retrieval is the hybrid input-output (HIO) algorithm proposed by Fienup [14, 15]. The algorithm maintains the basic structure of the ER algorithm but has an important negative feedback feature. For example, as shown in Fig. 15.4, consider the process that enforces the Fourier domain magnitude constraint. The current guess solution $g(x, y)$ is Fourier transformed and the amplitude of the resultant transform is replaced by the known modulus $|G|$. On inverse Fourier transforming, we get a new function $\hat{g}(x, y)$. The pixels in the new image may be divided into two sets: (i) set Γ : those satisfying the object domain constraints (e.g. positivity and support) and (ii) those which do not obey the object domain constraints. While the pixels of \hat{g} that obey the object domain constraints are retained as is, the pixels that do not obey the object domain constraints can be used to provide a negative feedback to the previous guess solution. While the operation illustrated in Fig. 15.4 is not linear in nature, this negative feedback idea as suggested in the HIO algorithm is known to improve the solution convergence significantly. The update step for the hybrid input-output algorithm is

$$\begin{aligned} g_{n+1}(x, y) &= \hat{g}_n(x, y), & (x, y) \in \Gamma \\ &= g_n(x, y) - \beta \hat{g}_n(x, y), & (x, y) \notin \Gamma. \end{aligned} \quad (15.3)$$

Here β is a positive parameter whose value is selected in range $(0.5, 1)$. In all the illustrations that follow, we have used $\beta = 0.7$ whenever HIO iteration is used. The effectiveness of the Fienup method for an object with higher complexity is illustrated in Fig. 15.5 where we use the cameraman as a real positive-valued object. The Fourier magnitude corresponding to this object and the reconstructions using 500 iterations of the ER and HIO algorithms are shown in Fig. 15.5a–c, respectively. For a computational window of 512×512 pixels, we place the cameraman object within the central 250×250 pixels. The support window size slightly less than half the computational window ensures that the Fourier domain intensity $|G(f_x, f_y)|^2$ is sampled as per the Nyquist rate. Note that the reconstructions are shown only over the support window. Once again the positivity and support constraints have been used for both ER and HIO iterations. We note that while the solution from ER iteration has not reached the desired solution, the HIO solution has nearly converged to the original cameraman object. There are a few artifacts in the solution that do not

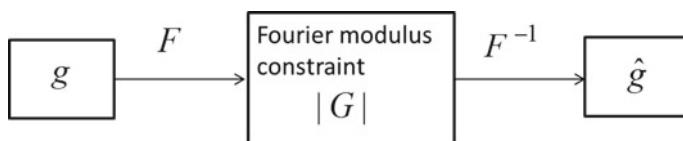


Fig. 15.4 Operation enforcing Fourier domain magnitude in phase retrieval iteration

go away easily; however, the solution is far better compared to what is recovered with the ER algorithm. The powerful negative feedback idea made phase retrieval an attractive approach from the point of view of system design and has since received a lot of attention in the optics, applied mathematics and electrical engineering community. While the HIO solution converges to the desired solution reasonably well in case of real positive-valued objects, the phase retrieval algorithms face difficulties when it is desired to recover complex-valued object functions. In particular, we note that a function $g(x, y)$ and its inverted complex conjugate (or twin) $g^*(-x, -y)$ have the same Fourier magnitude $|G(f_x, f_y)|$. Appearance of the twin in the HIO solution is illustrated in Fig. 15.6 where we use the letters ‘IITD’ as a binary phase object with phase values scaled in the range $[0, 2\pi/3]$. The phase object, its Fourier magnitude and the reconstructed phase are shown in Fig. 15.6a–c, respectively. The phase solution is corrupted with the twin (the letters ‘IITD’ are overlapped with their inverted and negative-valued replica). Clearly, if the object was not known *a priori*, it is difficult to make sense of the recovered solution when the twin is present. On a number of runs of HIO algorithm with different initial random guesses, a strong twin is observed to occur in significant number ($>90\%$) of runs. The twin once present in the solution does not go away easily by simply continuing with more HIO iterations. This effect is known as the twin stagnation. Getting rid of the twin image is usually considered to be a difficult problem. It was observed in [16] that when twin stagnation occurs, the solution pixels are roughly divided into two parts where each part takes on phase values corresponding to either of the twins. One of the solutions to the twin problem is to use a non centro-symmetric support (e.g. a triangular support) in the initial HIO iterations, which then encourages predominantly one of the twins to supersede over the other [17]. With the tools such as gradient-based penalty functions in hand in the context of image sparsity, we however make an interesting observation in [18]—*the solution with twin-stagnation artifact has lesser sparsity compared to any of the individual twins*. We may for example use the total variation (TV) of the solution $g(x, y)$ as a measure of fluctuations in it (this point will be formalized further in the next section). While the TV of the ground truth ‘IITD’ phase image is numerically equal to 1.9×10^3 , on a number of runs of 500 HIO iterations with random initial guess, the numerical value of TV of the stagnated solution as shown in Fig. 15.6c is observed to be $9(\pm 0.14) \times 10^3$. In [18] it was shown that the twin stagnation problem can be addressed effectively if the solution is nudged toward an image with lesser TV in every HIO iteration. Interestingly, it was observed that the algorithm can converge to either $g(x, y)$ or its twin $g^*(-x, -y)$ with equal chance.

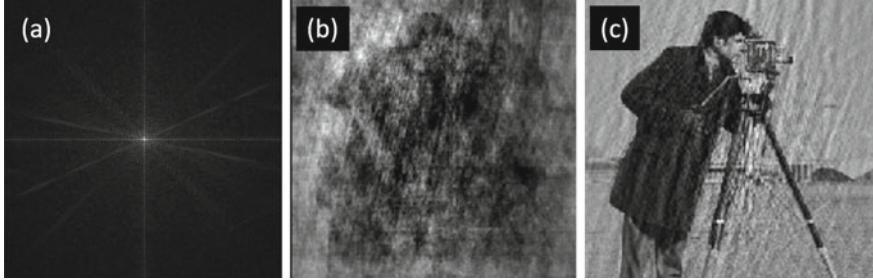


Fig. 15.5 Illustration of phase retrieval for a gray-scale real positive object. **a** Fourier magnitude shown as $|G(f_x, f_y)|^{0.25}$ for display, iterative recovery using 500 iterations of the **b** ER and **c** HIO approaches. The support part of reconstructions is shown

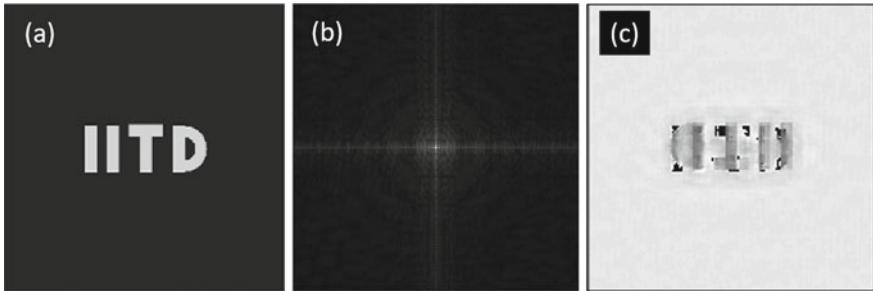


Fig. 15.6 Illustration of twin stagnation for **a** IITD' binary phase object. The two phase levels are 0 and $2\pi/3$. **b** Fourier transform of the phase object represented as $|G(f_x, f_y)|^{0.25}$ for display purposes, **c** Phase reconstruction with 500 HIO iterations showing twin stagnation. The support part of reconstructions is shown

15.3.1 The Relaxed-Reflect-Reflect (RRR) Algorithm for Phase Retrieval

We would like to mention here that there is an interesting generalization of the HIO algorithm—the relaxed-reflect-reflect (RRR) algorithm—which was proposed by Elser [19]. In RRR, a sequence of images is generated by starting with a random guess in the following manner:

$$\hat{g}_{n+1} = (1 - \gamma)\hat{g}_n + \gamma R_A R_B \hat{g}_n. \quad (15.4)$$

Here, γ is a constant in the range $(0, 1)$ (e.g. a standard choice is $\gamma = 0.5$). Further, the reflection operators are defined as

$$R_{A,B} = 2P_{A,B} - \hat{I}, \quad (15.5)$$

where $P_{A,B}$ represents a projection on the constraint(s) A or B . The constraint A may for example denote a combination of the object domain constraints like support or positivity and the constraint B may denote the Fourier magnitude constraint. It may be noted that in the RRR algorithm the current guess solution of interest g_n is given by

$$g_n = P_A[\hat{g}_n] \quad \text{or} \quad P_B[\hat{g}_n]. \quad (15.6)$$

The RRR algorithm can address a general class of multiple constraint feasibility problems. It has an interesting geometrical interpretation for which we refer the reader to the references [19, 21].

15.3.2 *The Relaxed Average Alternating Reflection (RAAR) Algorithm for Phase Retrieval*

The RAAR is a further generalization of the RRR theme [20] and the iterations are given by

$$\hat{g}_{n+1} = \frac{\beta}{2}(1 + R_A R_B)\hat{g}_n + (1 - \beta)P_B\hat{g}_n. \quad (15.7)$$

The parameter β may be selected to have a fixed value (e.g. $\beta = 0.9$) or it may be made to increase slowly to an upper bound of 1. The RAAR scheme is known to provide a more stable solution to the phase retrieval problem when noise and missing pixels in the data lead to a condition that there is no solution that satisfies both the space domain and Fourier domain constraints. RAAR is a popular choice among coherent X-ray imaging community.

15.4 Phase Retrieval Using Complexity Guidance

It may be noted that the object domain constraints (e.g. positivity, support) in the iterative algorithms for phase retrieval are applied somewhat independently in the sense that while modifying the guess solution in the object domain, no direct information from Fourier domain is utilized. The observation regarding the behavior of total variation of the HIO solution as in the previous section raises an important question. *Can the fluctuations in the solution be estimated in some manner beforehand from raw Fourier magnitude data?* If so, then this information can be used in the HIO iteration effectively to get rid of the stagnation artifacts. In this regard, we introduce a complexity parameter that measures expected degree of fluctuations in the phase retrieval solution. The complexity parameter may be defined as [22]

$$\zeta = \iint dx dy [|\nabla_x g(x, y)|^2 + |\nabla_y g(x, y)|^2]. \quad (15.8)$$

Note that the gradient-based definition means that its numerical value is higher when the function $g(x, y)$ has more structure or details in it. It is very easy to show using the energy theorem and the derivative property of Fourier transforms that the complexity ζ can also be expressed as

$$\zeta = 4\pi^2 \iint df_x df_y (f_x^2 + f_y^2) |G(f_x, f_y)|^2. \quad (15.9)$$

The above relation suggests that the complexity of the desired solution may be estimated directly from the measured Fourier intensity data even before initiating any phase retrieval algorithm. As we will show in the following discussion, the complexity information can be used advantageously to obtain a stagnation artifact-free solution. For the purpose of illustration, we now use the cameraman as a phase object with phase values scaled in the range $[0, 2\pi/3]$. Figure 15.7a shows the Fourier magnitude of the cameraman phase object. The phase reconstruction after 500 HIO iterations is shown in Fig. 15.7b. Note that in object domain only support constraint is used since the object is complex-valued. A faint inverted twin cameraman is clearly visible in this reconstruction where only the support part is shown. A superior artifact-free reconstruction with complexity guidance is shown in Fig. 15.7c. To understand the method of complexity guidance clearly, in Fig. 15.7d we show the complexity of the iterative solution as may be computed using Eq. (15.8). The blue straight line curve represents the ground truth complexity as obtained from the Fourier intensity data using Eq. (15.9). This is the complexity we desire for our solution. The magenta curve represents the solution's complexity during the HIO iterations. It is interesting to note that the solution's complexity is always higher compared to the ground truth complexity after a few initial iterations (we have initiated the solution with a constant zero phase in the support region). We may associate this higher complexity with the stagnation artifacts in the phase solution. In the complexity guidance methodology applied to the HIO algorithm, we apply recursive TV reduction within the support region of the solution after the HIO step (Eq. 15.3) by following the sub-iteration:

$$g \leftarrow g - \tau ||g||_2 [\nabla_{g^*} TV(g)] / ||\nabla_{g^*} TV(g)||_2. \quad (15.10)$$

The step size τ may be selected to be a constant small number (e.g. 0.005) so as to change the solution in small steps or selected by a line search. This recursive TV reduction is performed till the solution's complexity is brought close to the ground truth complexity as illustrated by the green curve in Fig. 15.7d within each HIO main iteration. The resultant phase solution in Fig. 15.7c is observed to be cleaner without any artifacts like (twin stagnation, stripes, etc.) those in the HIO solution. The number of recursive TV iterations required is not very large (typically less than 10) and as a result with some additional processing it is possible to arrive at a good-quality phase solution. The illustration in Fig. 15.7 was shown with nearly noiseless Fourier

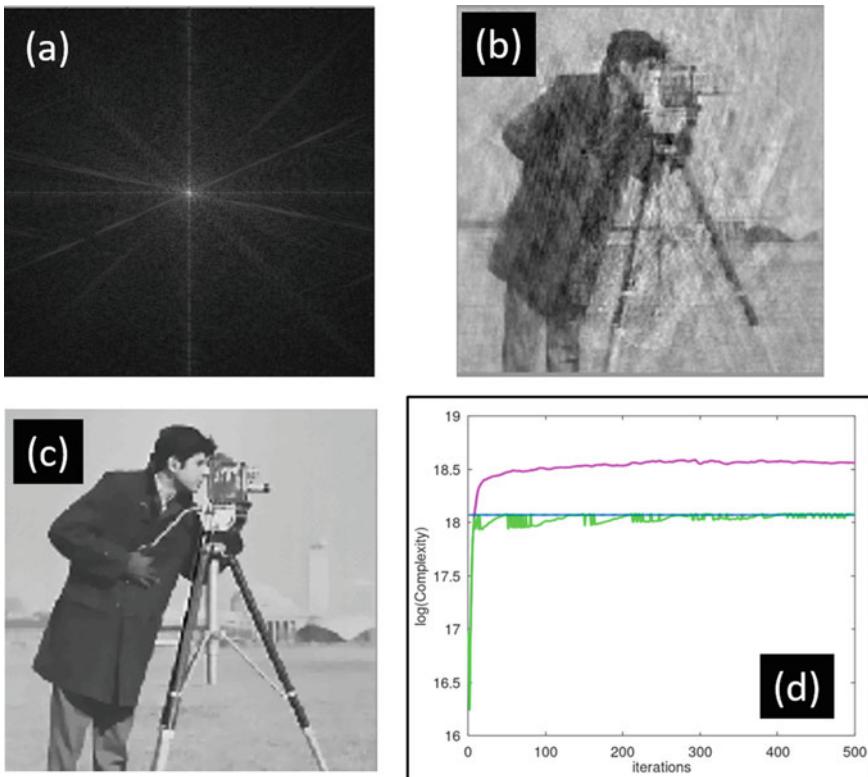


Fig. 15.7 Illustration of complexity-guided phase retrieval. **a** Fourier magnitude of cameraman phase object displayed as $|G(f_x, f_y)|^{0.25}$, **b** Phase recovery with 500 HIO iterations, **c** Phase recovery with 500 complexity-guided HIO iterations, **d** Behavior of solution’s complexity as a function of iterations. Blue constant curve: ground truth complexity, Magenta curve: complexity of HIO solution, Green curve: complexity of complexity-guided HIO solution

magnitude data. In real experiments, the Fourier magnitude data may be corrupted with noise. Further, due to limited dynamic range of the detector being used, the detector pixels corresponding to the central peak in the Fourier transform domain may get saturated. Note that the complexity as defined in Eq. (15.9) is minimally affected by the saturation of pixels near zero spatial frequency. The complexity guidance procedure in presence of substantial noise is somewhat more involved and requires careful control of solution complexity within and outside the support region. We will not discuss here this topic further but refer the readers to [23]. The interesting aspect about complexity guidance methodology is that the complexity information about the solution is available with phase-less incomplete data and can potentially benefit any version of the phase retrieval algorithm. Also, the methodology allows one to control the complexity of the solution precisely thus providing a guideline to the iterative phase retrieval algorithms.

15.5 Phase Retrieval with Multiple Intensity Measurements: Fourier Ptychography

While we have mainly discussed the phase retrieval problem in the chapter for Fourier magnitude data, the iterative algorithms may also be applied to situations where multiple intensity measurements are available. Traditionally such algorithms which utilize multiple intensity measurements fall under the category of phase diversity methods [24–29]. Here diversity refers to the fact that the same unknown optical wavefront to be measured was subjected to a diversity of optical systems before making multiple intensity measurements. For example, the diversity may be achieved via a series of defocus images, by introducing spiral phase diversity or having illumination diversity in recording of multiple intensity patterns. In this section, we will describe the Fourier ptychography system for achieving resolution beyond the diffraction limit via diversity of illumination angles. A Fourier ptychography system can be understood in terms of a 4F system model as shown in Fig. 15.8. The illumination of object placed in (x', y') plane is performed sequentially with tilted plane waves that are depicted with red and purple colored rays in Fig. 15.8a. When a plane wave $\exp(i2\pi f_0 y)$ illuminates an object $g(x, y)$, the Fourier transform of the resultant input object is given by its shifted replica $G(f_x, f_y - f_0)$ as shown by the purple circle in Fig. 15.8b. The size of the circles (red or purple) is decided by the size of the Fourier plane aperture of the 4F system. The tilted illumination therefore allows spatial frequencies outside the pass-band of the system (shown by the red circle) to contribute to image formation. Multiple angled illumination by plane waves in different directions allows us to increase the effective spatial frequency coverage in Fourier space. This amounts to obtaining higher spatial resolution beyond the native diffraction-limited resolution of the 4F system which is decided by the Fourier plane aperture size. One may therefore get a final coherent image with higher effective pixel density over the field of view of the system. The idea of Fourier ptychography was first proposed in the context of coherent X-ray imaging [30] and its optical implementation was realized by [31]. Note that each of the measurements corresponding to the angled illumination results

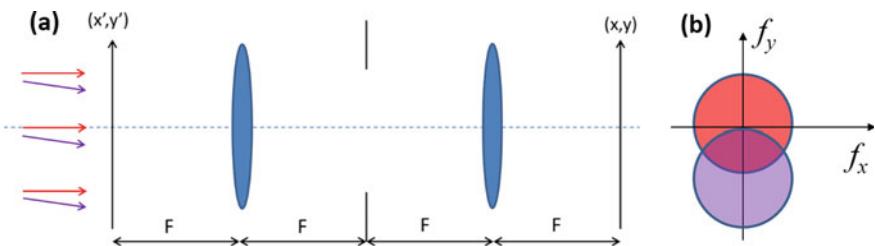


Fig. 15.8 **a** Fourier ptychography imaging modeled using a 4F optical processor. Illumination from two angles by a coherent plane wave is depicted with red and purple colored rays. **b** Corresponding regions of coverage in spatial frequency space are shown using red and purple circles. The size of the circles is governed by the Fourier plane aperture of the 4F system

in an intensity image at the output plane of the imaging system. The phase information is lost at this measurement stage. The reconstruction of complex-valued object function requires a phase retrieval algorithm that we will describe next.

Suppose use of a set of N angled illuminations results in intensity images denoted by $I_k(x, y)$ with $k = 1, 2, 3, \dots, N$. The aim is to recover a complex-valued object function $g(x, y)$ which satisfies constraint corresponding to all the recorded intensity images. The algorithm is similar in nature to the Gerchberg-Saxton algorithm in that we move in a cyclic fashion over all measurements and update the solution after satisfying each of the measurement amplitude constraints. The iterative algorithm is initiated with a random guess $g_0(x, y)$ at a higher resolution whose Fourier transform may be computed as $G_0(f_x, f_y)$. A possible initial guess for $g_0(x, y)$ may for example be a zero phase real object which is an interpolated version of the image amplitude $\sqrt{I_k(x, y)}$. In the $(m + 1)$ th iteration of the algorithm, the following steps are carried out:

1. The high pixel resolution image $g_m(x, y)$ is passed through k th Fourier band and downsampled to the size of the detector data $I_k(x, y)$. The amplitude of the resultant image is replaced by $\sqrt{I_k(x, y)}$ and keeping the phase unchanged.
2. The resultant image is inverse Fourier transformed with upsampling. The upsampling may for example be achieved by zero-padding the image to appropriate size prior to inverse transforming operation.
3. In the Fourier space, the region of Fourier transform $G_m(f_x, f_y)$ corresponding to the k th spatial frequency band is replaced by numerical values from the inverse transformed image in step 2 above. The numerical values of the Fourier transform $G_m(f_x, f_y)$ outside the spatial frequency band of interest are left unchanged.
4. These processing steps are performed sequentially for all the N spatial frequency bands.
5. The resultant Fourier transform $G_{m+1}(f_x, f_y)$ after cycling through all N spatial frequency bands may be inverse transformed to get the next guess $g_{m+1}(x, y)$ for the complex-valued solution.

Figure 15.9 shows nine raw intensity frames corresponding to the spatial frequency coverage shown by blue dotted circles. The central filled circle represents the native pass-band of the 4F system. The result of five iterations of the Fourier ptychography algorithm is illustrated in Fig. 15.10a, b where we show low-resolution image recorded by the native system pass-band (on-axis plane wave illumination), and the high-resolution reconstruction by cycling through all the nine bands. In Fig. 15.10c, we also show the Fourier magnitude corresponding to the recovered high-resolution solution after five iterations. The yellow circle in Fig. 15.10c shows the native pass-band of the 4F system. It is observed that the bandwidth of the resultant image is almost doubled due to the raw data from intensity images with angled illumination.

In practice, getting a Fourier ptychographic system working involves more efforts on carefully calibrating various illumination angles involved so that they may be incorporated in the algorithm without causing any mismatch. Further since current

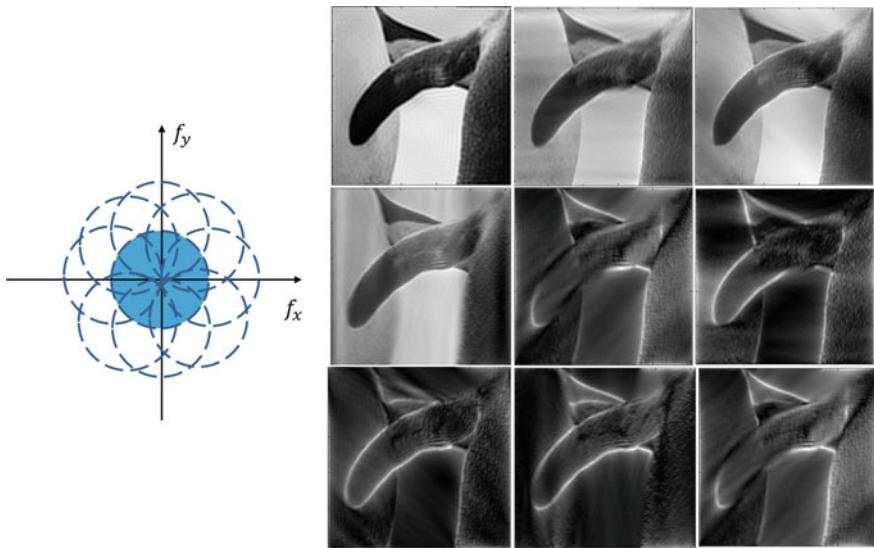


Fig. 15.9 Illustration of raw intensity images from nine Fourier bands in a Fourier ptychography system. The spatial frequency coverage is illustrated by circles in the spatial frequency plane

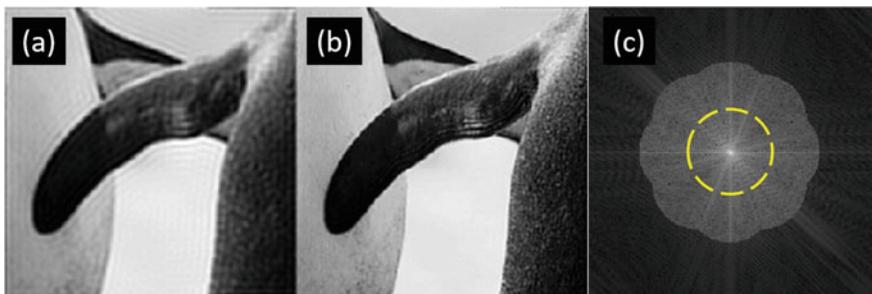


Fig. 15.10 Illustration of high-resolution ptychographic reconstruction for an amplitude object, **a** low-resolution image from a system with illumination by an on-axis plane wave, **b** high-resolution reconstruction after five iterations of Fourier ptychography algorithm, **c** Fourier transform magnitude of the resultant solution in **b** showing enhanced spatial frequency coverage. The yellow circle shows the native spatial frequency pass-band of the 4F system

ptychographic systems use LED (light emitting diode) arrays as source of angled illumination, efforts have to be made to make sure that the individual LEDs are similar in their illumination pattern. A calibration of this aspect is an additional requirement. Finally, sufficient distance must be maintained between a single LED and the sample so that the illumination reaching a sample within field of view of interest can be considered to be spatially coherent (or may be thought of as a plane wave making an angle with respect to the optic axis). The van Cittert-Zernike theorem (see Chap. 10) may be used for estimating the coherence area of the individual LEDs at the sam-

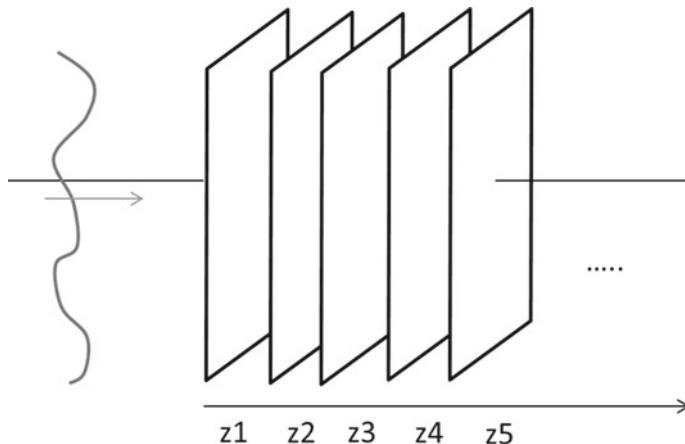


Fig. 15.11 Phase diversity imaging by multiple intensity measurements related by defocus. An unknown wavefront is traveling in $+z$ -direction and intensity measurements are made using an area sensor at different locations along z -axis

ple. The Fourier ptychography imaging system can be considered as an example of coherent structured illumination. Finally, we remark that the practical convergence of the ptychography algorithm within a few iterations as discussed here requires significant overlap between adjacent frequency bands. The image reconstruction in the incoherent version of structured illumination imaging is handled in a different manner and we will discuss this topic later in this book.

15.6 Phase Retrieval as an Optimization Problem

We have discussed a number of popular iterative approaches for phase retrieval in this chapter. While the Gerchberg-Saxton algorithm was shown to be equivalent to a steepest descent type algorithm by Fienup, it was not until last decade that phase retrieval problem gained a renewed attention as an optimization problem. We briefly mention two main approaches in this context. The first approach is known as the **Wirtinger Flow** algorithm [32], which works with data of the form:

$$y_r = |\mathcal{F}[s_r(x, y) g(x, y)]|^2 \quad r = 1, 2, 3, \dots, M. \quad (15.11)$$

Here $s_r(x, y)$ are typically random code patterns. The Wirtinger Flow aims to minimize the L_2 -norm squared error function:

$$C(g, g^*) = \sum_{r=1}^M [y_r - |\mathcal{F}[s_r(x, y) g(x, y)]|^2]^2. \quad (15.12)$$

The name Wirtinger Flow follows from the fact that the steepest descent direction for the cost function is evaluated using the complex or Wirtinger derivatives. Typical illustrations of Wirtinger Flow in the literature have used approximately 10 coded masks $s_r(x, y)$ for obtaining a stable recovery of the image $g(x, y)$. The **PhaseLift** [34] algorithm on the other hand converts the phase retrieval problem into a linear problem. In particular, we observe that the data y_r in Eq. (15.11) may alternatively be written as

$$y_r = |\langle F_r, g \rangle|^2 = \text{Tr}(F_r F_r^* g g^*). \quad (15.13)$$

Here, F_r represents the Fourier transform operation including multiplication by code function s_r . In the above equation, g represents a vectorized form of the solution $g(x, y)$. Denoting $G = gg^*$, the phase retrieval problem is now equivalent to finding a positive semi-definite matrix G of rank 1 such that $\text{Tr}(F_r F_r^* G) = y_r$. It was shown in [33] that this problem may be relaxed to the problem of minimizing $\text{Tr}(G)$ subject to the measurement constraints. Once the matrix G is recovered, the solution g is determined by its factorization up to a global phase factor. At present, a number of applications in optics and coherent diffraction imaging still depend on the more traditional algorithms like GS, HIO, RRR or their variants. The more advanced optimization methods discussed above may however become important in years to come if they prove to be more robust.

Problems

15.1 Computer-Generated Hologram (CGH): Computer-generated holograms refer to amplitude or phase patterns that may be fabricated on a glass element or displayed on spatial light modulators whose diffraction pattern yields a picture of interest. Write a Gerchberg-Saxton type algorithm for determining a constant amplitude phase mask whose diffraction pattern is an annular ring-shaped (or any other) intensity pattern. What constraints will you use in the Fourier domain and the object domain? Read and explain the concept of “hogel” from modern CGH design literature [35].

15.2 Detour phase hologram: Once a phase mask is determined as described in Problem 15.1, it is important to impart this phase map onto a spatially coherent beam. Detour phase hologram was an interesting method which allowed phase mask display using amplitude transmission element. What is the difficulty that one may

encounter when using an amplitude mask as a CGH? When the illumination is via a tilted wavefront, the phase of the incident field changes periodically in the CGH plane. Think of a scheme which on blocking appropriate parts in the CGH allows one to realize a phase mask pattern. Can the detour phase concept be generalized to incident wavefronts other than a tilted plane wave?

15.3 Scanned beam ptychography : While we have discussed the concept of Fourier ptychography in this chapter, another form of ptychography in the image domain also exists [36]. Consider any real positive-valued object $g(x, y)$ of your choice over which a small circular aperture $P(x, y)$ is scanned. At each of the positions of the aperture, a far field diffraction intensity pattern is recorded which may be given by $|\mathcal{F}\{P(x - x_j, y - y_j)g(x, y)\}|^2$. Here (x_j, y_j) denote the center locations of the circular aperture. Using the number of diffraction intensity patterns as data, write an ER-like algorithm to recover the whole object function $g(x, y)$. What is the effect of reducing the overlap between the neighboring aperture locations on the convergence of the algorithm?

15.4 Defocus diversity: Consider any binary text image of your choice over a 256×256 pixel grid of $5 \mu\text{m}$ pixels. The text part may be limited to the central 100×100 pixels of the computational window. Simulate angular spectrum propagation of this input field over a periodic interval of $\Delta z = 1 \text{ mm}$ up to a distance $z = 10 \text{ mm}$ (Fig. 15.11). At each plane, save the field intensity $I_k(x, y)$ to which you may add a 1% uniform random noise. Now treat the input plane object in the $z = 0$ plane as unknown and implement a Gerchberg-Saxton-type algorithm that cycles through all the intensity measurements and further uses a reality and positivity constraint in the $z = 0$ plane in order to recover the unknown object. How well does the recovery work if you use only the first two intensity images at $z = 1, 2 \text{ mm}$ as a data? How does the recovery change if you choose the two intensity images at $z = 5, 10 \text{ mm}$ as data? In your algorithm, does the sequence of accessing different planes in the iterative reconstruction make any difference to image recovery? (e.g. you may try propagation to different z -planes sequentially or in a randomized manner.)

References

1. J.R. Fienup, Phase retrieval algorithms: a personal tour. *Appl. Opt.* **52**(1), 45–56 (2013)
2. A. Labeyrie, Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images. *Astron. Astrophys.* **6**, 85–87 (1970)
3. Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, M. Segev, Phase retrieval with application to optical imaging: a contemporary overview. In: *IEEE Signal Process. Mag.* **32**(3), 87–109 (2015)
4. D. Sayre, Some implications of a theorem due to Shannon. *Acta Crystallogr. A* **5**, 843–843 (1952)
5. J.W. Miao, P. Charalambous, J. Kirz, D. Sayre, Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature* **400**, 342–344 (1999)

6. H.N. Chapman, K.A. Nugent, Coherent lensless X-ray imaging. *Nat. Photon.* **4**(12), 833–839 (2010)
7. J. Miao, T. Ishikawa, I.K. Robinson, M.M. Murnane, Beyond crystallography: diffractive imaging using coherent x-ray light sources. *Science* **348**(6234), 530–535 (2015)
8. R.P. Millane, Phase retrieval in crystallography and optics. *J. Opt. Soc. Am. A* **7**(3), 394–411 (1990)
9. C. Bostedt, S. Boutet, D.M. Fritz, Z. Huang, H.J. Lee, H.T. Lemke, A. Robert, W.F. Schlotter, J.J. Turner, G.J. Williams, Linac coherent light source: the first five years. *Rev. Mod. Phys.* **88**(1), 015007 (2016)
10. D. Shapiro, P. Thibault, T. Beetz, V. Elser, M. Howells, C. Jacobsen, J. Kirz, E. Lima, H. Miao, A.M. Neiman, D. Sayre, Biological imaging by soft x-ray diffraction microscopy. *Proc. Natl. Acad. Sci.* **102**(43), 15343–15346 (2005)
11. ...P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F.-J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, Ph. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H.-D. Nuhn, G. Pile, D. Ratner, J. Rzepieila, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, J. Galayda, First lasing and operation of an ångstrom-wavelength free-electron laser. *Nat. Photon.* **4**(9), 641–647 (2010)
12. H.N. Chapman et al., Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nat. Phys.* **2**(12), 839–843 (2006)
13. R.W. Gerchberg, W.O. Saxton, A practical algorithm for the determination of the phase from image and diffraction plane pictures. *Optik* **35**, 237–246 (1972)
14. J.R. Fienup, Reconstruction of an object from the modulus of its Fourier transform. *Opt. Lett.* **3**, 27–29 (1978)
15. J.R. Fienup, Phase retrieval algorithms: a comparison. *Appl. Opt.* **21**, 2758–2769 (1982)
16. M. Guizar-Sicairos, J.R. Fienup, Understanding the twin-image problem in phase retrieval. *J. Opt. Soc. Am. A* **29**(11), 2367–2375 (2012)
17. J.R. Fienup, C.C. Wackerman, Phase retrieval stagnation problems and solutions. *J. Opt. Soc. Am. A* **3**, 1897–1907 (1986)
18. C. Gaur, B. Mohan, K. Khare, Sparsity-assisted solution to the twin image problem in phase retrieval. *J. Opt. Soc. Am. A* **32**(11), 1922–1927 (2015)
19. V. Elser, I. Rankenburg, P. Thibault, Searching with iterated maps. *Proc. Natl. Acad. Sci.* **104**, 418–423 (2007)
20. D. Russel Luke, Relaxed averaged alternating reflections for diffraction imaging. *Invers. Prob.* **21**, 37–50 (2005)
21. R.P. Millane, V.L. Lo, Iterative projection algorithms in protein crystallography I. Theory. *Acta Crystallogr. A* **69**(5), 517–527 (2013)
22. M. Butola, S. Rajora, K. Khare, Phase retrieval with complexity guidance. *J. Opt. Soc. Am. A* **36**(2), 202–211 (2019)
23. M. Butola, S. Rajora, K. Khare, Complexity-guided Fourier phase retrieval from noisy data. *J. Opt. Soc. Am. A* **38**(1), 488–497 (2021)
24. A. Anand, G. Pedrini, W. Osten, P. Almoro, Wavefront sensing with random amplitude mask and phase retrieval. *Opt. Lett.* **32**, 1584–1586 (2007)
25. P. Almoro, G. Pedrini, W. Osten, Complete wavefront reconstruction using sequential intensity measurements of a volume speckle field. *Appl. Opt.* **45**, 8596–8605 (2006)
26. P. Gao, G. Pedrini, W. Osten, Phase retrieval with resolution enhancement by using structured illumination. *Opt. Lett.* **38**, 5204–5207 (2013)
27. P. Gao, G. Pedrini, C. Zuo, W. Osten, Phase retrieval using spatially modulated illumination. *Opt. Lett.* **39**, 3615–3618 (2014)
28. R.A. Gonsalves, Phase retrieval by differential intensity measurements. *J. Opt. Soc. Am. A* **4**, 166–170 (1987)
29. M.K. Sharma, C. Gaur, P. Senthilkumaran, K. Khare, Phase imaging using spiral phase diversity. *Appl. Opt.* **54**, 3979–3985 (2015)

30. H.M.L. Faulkner, J.M. Rodenburg, Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm. *Phys. Rev. Lett.* **93**(2), 023903 (2004)
31. A.M. Maiden, M.J. Humphry, F. Zhang, J.M. Rodenburg, Superresolution imaging via ptychography. *J. Opt. Soc. Am. A* **28**(4), 604–612 (2011)
32. E.J. Candes, X. Li, M. Soltanolkotabi, Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Trans. Inf. Theory* **61**(4), 1985–2007 (2015)
33. E.J. Candes, Y.C. Eldar, T. Strohmer, V. Voroninski, Phase retrieval via matrix completion. *SIAM Rev.* **57**(2), 225–251 (2015)
34. E.J. Candès, T. Strohmer, V. Voroninski, PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming. *Commun. Pure Appl. Math.* **66**, 1241–1274 (2013)
35. M. Yamaguchi, Light-field and holographic three-dimensional displays. *J. Opt. Soc. Am. A* **33**(12), 2348–2364 (2016)
36. A.M. Maiden, J.M. Rodenburg, An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy* **109**, 1256–62 (2009)

Chapter 16

Compact Multi-lens Imaging Systems



Digital imaging has become a widespread technology with miniature cameras getting embedded in devices such as mobile phones. While digital array (CCD or CMOS) sensors have virtually replaced film-based cameras for most imaging applications, the physics of image formation in these systems has largely remained the same. It is only in the last decade that novel system concepts are emerging that use unconventional optical design followed by image recovery algorithms [1]. In this chapter, we will discuss on some specific computational imaging systems that have received much attention in literature in recent times [2, 3]. The examples chosen here represent systems or ideas that truly bring out the advantages of integrating computational methods [4] with unconventional optical system design to achieve superior imaging performance [5–7]. The first system that we will discuss is the compact form factor multi-lens camera followed by the plenoptic or lightfield camera that allows retrospective computational focusing. Both the systems offer practical advantages that cannot be achieved through purely hardware or purely image processing approaches thus highlighting the importance of the computational imaging framework.

16.1 Compact Form Factor Computational Camera

Easy availability and manufacturability of miniature optical components have made it possible to realize compact optical imaging systems. Suppose we consider an ordinary digital camera and set out to miniaturize it, all components may be reduced proportionately except for the sensor pixels that have typical dimensions in the range $1\text{--}5 \mu\text{m}$. Suppose a microlens with the same F-number as an ordinary camera lens is used, reduction in lens aperture size implies that proportionately smaller number of pixels are available for recording the image with the same field of view. One interesting solution to this problem is to use multiple microlenses (typically a

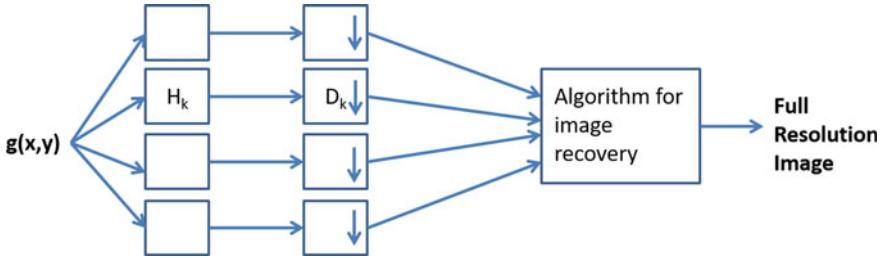


Fig. 16.1 Computational camera model from the viewpoint of generalized sampling. A given high-resolution scene is imaged using multiple imaging systems (denoted by H_k). Recording the output of these multiple imagers with limited number of pixels behind each microlens involves downsampling (denoted by D). All the low-resolution images are then combined by numerical processing to obtain a single high-resolution image

periodic microlens-array). The image recorded behind each of the microlenses is a low-resolution image but all the low-resolution images may be combined digitally to produce a single full-resolution image. The model for this computational camera relies on the idea of generalized sampling [8–10] which may be explained as in the schematic diagram shown in Fig. 16.1. A given picture $g(x, y)$ is seen by multiple microlens imagers with individual impulse responses given by $h_k(x, y)$. The recorded image behind an individual microlens may be given by a downsampled and translated version of $g_k(x, y) = g(x, y) * h_k(x, y)$. Even if all the individual impulse responses $h_k(x, y)$ are identical, a diversity in recorded images is possible due to relative fractional pixel shifts between the individual images that have been downsampled. This diversity may be exploited algorithmically to combine all the low-resolution images into a single high-resolution image. Formally the k th recorded sub-image may be represented as

$$g_k = \hat{D} \hat{H}_k \hat{S}_k g. \quad (16.1)$$

Here, the operators \hat{D} , \hat{H}_k , \hat{S}_k denote the operations of downsampling, blurring due to system impulse response and shift (or translation), respectively. Note that the downsampling operator is assumed to be the same for all elemental sub-images. The inversion problem of reconstructing the full high-resolution image may be expressed as constrained optimization problem of minimizing a cost function:

$$C(g) = \sum_k \|g_k - \hat{D} \hat{H}_k \hat{S}_k g\|^2 + \alpha \psi(g). \quad (16.2)$$

Here, the first term represents the data consistency and the second term represents any desirable constraint on the final image, e.g. minimal total variation or other appropriate penalty functions. When solving the optimization problem iteratively, a transpose (or adjoint) of the forward system operator for the k -th system

$$\hat{A}_k = \hat{D} \hat{H}_k \hat{S}_k \quad (16.3)$$

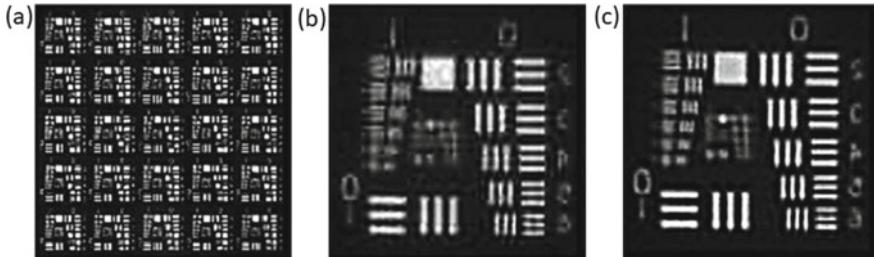


Fig. 16.2 Illustration of image recovery for a multi-lens thin camera system, **a** 25 elemental low-resolution images, **b** upsampled interpolated version of the central sub-image in the 5×5 low-resolution image array in **(a)**, **c** Reconstructed high-resolution image. The reconstruction was performed using the total variation penalty. [Multi-lens image data courtesy Dr. G. Unnikrishnan, Instruments Research Development Establishment, Dehradun, India]

is required. The transpose operator is readily represented as

$$\hat{A}^\dagger = \hat{S}_k^- \hat{H}_k^\dagger \hat{U}, \quad (16.4)$$

where \hat{S}_k^- , \hat{H}_k^\dagger , \hat{U} denote negative shift, convolution with inverted and conjugated impulse response $h_k^*(-x, -y)$ and upsampling operations, respectively. The down-sampling and upsampling operations may be performed by any standard interpolation routine. In Fig. 16.2, we show an example of high-resolution image recovery with a compact system consisting of 25 element microlens array. The elemental images, interpolated version of the central image in the 5×5 elemental image array and the recovered high-resolution image are shown. The total variation penalty function was used for the constraint term $\psi(g)$. We observe that the final reconstruction has higher resolution as well as lower noise background. The reconstruction is typically an iterative procedure as described in Chapter 6. If the system is modeled correctly, it may be possible in general to get a good final image quality even without high-quality microlens array. We note once again that a compact form factor system is possible here through a combination of system design and image recovery algorithms. Any amount of lens design sophistication or image processing done stand alone cannot easily achieve this goal in general. Finally, one may note that the above system is an incoherent imaging configuration where a digital array detector is used. In such systems, the resolution is typically decided by the pixel size (currently 2–5 μm range) rather than by the diffraction limit of the aperture. So the aim of multi-lens systems as described here is to achieve the same digital image resolution as that for a conventional camera.

16.2 Lightfield Cameras

In traditional camera designs, typically a multi-element lens is placed in front of a sensor and focusing is achieved by changing the distance between the sensor and the lens. Plenoptic imaging systems [11] record information in such a way that a single recorded frame allows retrospective focusing of images [12]. The image recorded at the sensor is not directly meaningful visually. An image reconstruction algorithm applied to the recorded data however enables focusing at various depths. The imaging systems of this nature are referred to as plenoptic or lightfield imaging systems. In this section, we will explain the working principle of lightfield systems with the help of some simple simulations. We will first discuss the concept of lightfield and how plenoptic systems record the lightfield using a microlens array. Image reconstruction algorithm for retrospective computational refocusing will be discussed next.

16.2.1 The Concept of Lightfield

In typical camera systems, every sensor pixel records a signal proportional to the total integrated light energy arriving at the pixel area. The information about the directional distribution of this energy is lost. Referring to Fig. 16.3, the lightfield function $L(u, v; x, y)$ represents the distribution of light energy reaching the image plane at the point (x, y) from an elemental area located at coordinates (u, v) in

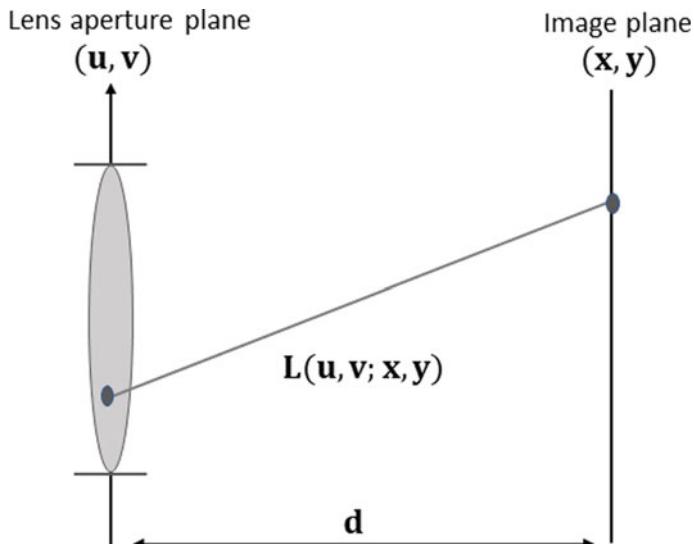


Fig. 16.3 Schematic of an imaging system. The lightfield function $L(u, v; x, y)$ represents the energy distribution from lens aperture plane to the image plane

the lens aperture plane. Conceptually, the lightfield function may be considered to have complete information about the directional distribution of light rays entering the imaging system. The plenoptic systems aim to record this complete information (the name “plenoptic” refers to complete information). *The lightfield is analogous to the Wigner distribution of a function as it contains information about the position as well as direction of a ray.* The image intensity recorded as a function of image plane coordinates (x, y) is obtained by integrating the lightfield function over the lens aperture coordinates:

$$I(x, y) = \iint dudv L(u, v; x, y). \quad (16.5)$$

We note that if the lightfield function is known between the lens aperture plane and an image plane, then it is possible to compute it between the lens aperture plane and any other defocused plane. As shown in Fig. 16.4, the ray of light representing optical energy flow between points $A(u, v)$ and $P(x, y)$ also cuts another plane at a point $P'(x', y')$. The coordinates (x', y') of the new plane are related to the coordinates (x, y) by a simple geometrical relation:

$$x' = \alpha x + \beta u, \quad (16.6)$$

$$y' = \alpha y + \beta v. \quad (16.7)$$

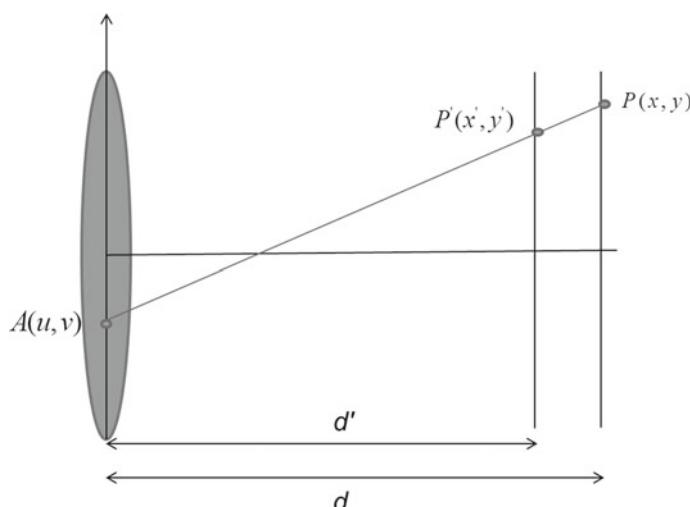


Fig. 16.4 Knowledge of the lightfield function $L(u, v; x, y)$ allows one to compute the transformed lightfield function $L(u, v; x', y')$ between lens aperture and another plane

Here, the parameters α and β are given by

$$\alpha = \frac{d'}{d}, \quad \beta = 1 - \frac{d'}{d}. \quad (16.8)$$

We may therefore obtain the lightfield function $L(u, v; x', y')$ as

$$L(u, v; x', y') = L(u, v; \alpha x + \beta u, \alpha y + \beta v). \quad (16.9)$$

The implication of this scale and shift transformation is that, when integrated with respect to (u, v) coordinates, the new lightfield function gives us a new image $I(x', y')$ formed in the new plane without any physical focusing requirement.

16.2.2 Recording the Lightfield Function with Microlens Array

In the last section, we have seen that *the knowledge of the lightfield function between two planes is sufficient for computational refocus*. The recording of the lightfield function is therefore the first aim of a lightfield imaging system. The lightfield function may be recorded by introducing a microlens array between the main lens aperture and the sensor array. The microlens array is positioned such that each microlens in the array images the main lens aperture on the sensor plane. The configuration is shown schematically in Fig. 16.5a. Figure 16.5b shows part of the pattern recorded

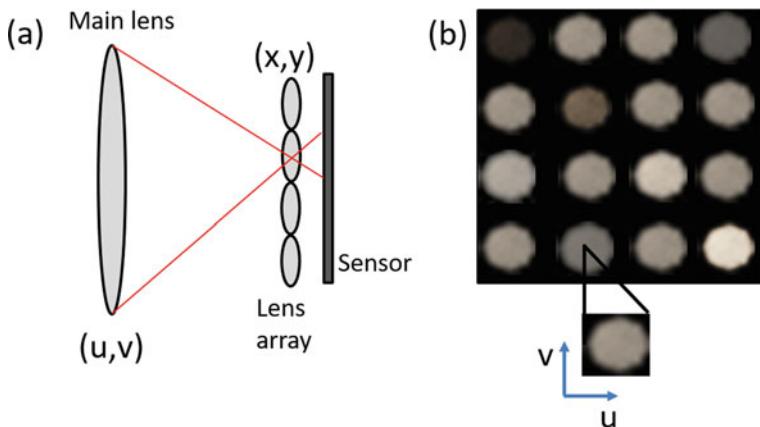


Fig. 16.5 **a** Schematic of a lightfield imaging system. The lenses in the microlens array are positioned to image the main lens aperture on the sensor. **b** Part of a typical sensor pattern. The pattern behind an individual microlens is a map of the lens aperture $((u, v)$ coordinates) as seen from the corresponding microlens position

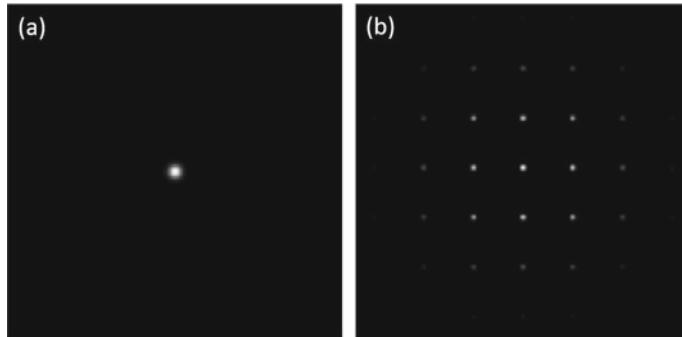


Fig. 16.6 Central portion of simulated data frame as recorded by a plenoptic camera for an on-axis point source. For simulation, we used a 31×31 microlens array with 32×32 pixels behind each microlens. **a** Nominal on-axis point source which is imaged by main camera lens on microlens array plane, **b** on-axis point source imaged by main lens in a plane defocused with respect to microlens plane

on the sensor array. Care must be taken to adjust the lens aperture so that the images behind the individual microlenses do not overlap. As shown in Fig. 16.5b, the pattern recorded behind an individual microlens is a map of the lens aperture as seen from the corresponding microlens position. Representing the microlens center coordinates as (x_j, y_j) , the pattern behind a single microlens may be denoted as $L(u, v; x_j, y_j)$. The collection of all individual images thus forms the complete lightfield function. In Fig. 16.6, we show a simulation for image data recorded by a plenoptic camera for an on-axis point source. For the simulation, we assumed 31×31 lens array. The lens size is such that there is an array of 32×32 pixels behind each microlens. The illustration is shown in order to provide a basic understanding of the plenoptic system and do not refer to a specific plenoptic camera. When the point source is located such that the source is imaged in the microlens plane, the raw detector data as shown in Fig. 16.6a is observed. We see that only the image area behind the on-axis microlens is illuminated. For a defocused on-axis source located farther away from the plenoptic camera, a focused image is formed somewhere between the main lens and the microlens array. As a result, the microlens array is illuminated by a cone of light rays covering a number of microlenses depending on the defocus distance. The central part of the eventual pattern appearing on the detector is shown in Fig. 16.6b. Note that a source located on the system's optical axis is also on the optical axis of the central microlens but is an off-axis source for other microlenses. The peaks in the detector pattern as in Fig. 16.6b are thus shifted relative to the microlens centers (except for the central microlens). The next task is to process this data to form an image of the point source. Note that the number of (x, y) samples available is the same as the number of microlenses and as a result nominal image resolution obtained in this example is equal to the number of microlenses (which is much less than the number of sensor pixels). In practical plenoptic cameras, the lens array consists of at least 256×256 lenses with lens diameter in the $100\text{--}150\,\mu\text{m}$ range. The plenoptic

camera thus requires a sensor with large active area. The lenslet (microlens) array also needs to be positioned carefully in the camera. For the data in Fig. 16.6a, image formation is a trivial task since only the central sub-image contains non-zero data and thus contributes to image formation. The lightfield function is a collection of all the sub-images. Since the microlens plane is the nominal image plane, a focused image is obtained as

$$I_1(x, y) = \iint dudv L(u, v; x, y). \quad (16.10)$$

The reconstructed image in this case is essentially obtained by integrating the total energy in the central sub-image and putting it in the central image pixel. All the other reconstructed pixel values are zero. For the data in Fig. 16.6b, we observe that the cone of light due to an on-axis point source is seen to reach several lenses and the corresponding sub-images on the detector have non-zero pixels. The pixels from the sub-images obtained from the raw detector data are first rearranged such that for a fixed (u, v) we get a 2D image $L(u, v; x, y)$ by appropriately sampling the pixel locations behind each of the microlenses. These images are shifted and scaled as per Eq. (16.9) to obtain the components $L(u, v; x', y')$ of the lightfield for fixed (u, v) at a desired z -plane. Integration over main lens aperture plane coordinates (u, v) is then performed for computing the image:

$$I_2(x', y') = \iint dudv L(u, v; x', y'). \quad (16.11)$$

The procedure above amounts to adding up all the individually transformed light-field components. Figure 16.7a–c shows reconstructions of the image of the on-axis point source (corresponding to data in Fig. 16.6b) at various depths. The scaling and shifting operations above were implemented using the standard bicubic interpolation scheme. For photography in ambient light, a general object/scene to be imaged may

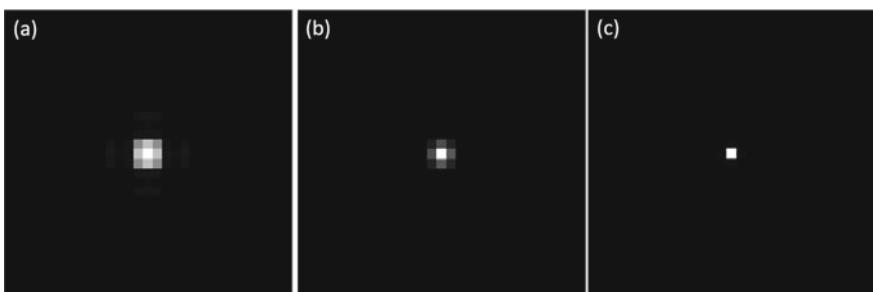


Fig. 16.7 Image reconstruction corresponding to raw data in Fig. 16.6b at various depths between the main lens and the microlens array plane. **a, b** Reconstruction for on-axis point source at defocused planes, **c** focused image of the point source when appropriate values of α and β as in Eq. (16.9) are used

be considered to be a collection of point sources at various depths. The raw data in that case may not be easy to interpret by visual inspection; however, the processing of the data is identical to the above case of on-axis point source.

The ability of post-capture focusing is an interesting aspect of the plenoptic camera which makes it particularly important for applications where focusing in real time is difficult (e.g. in case an object moving longitudinally with respect to the camera). An addition of lenslet array hardware in the camera optics along with the reconstruction algorithm therefore offers new possibilities for how photography cameras may work in the future [13, 14]. Plenoptic imaging systems have also been used for microscopy as well as in characterization of turbulence [15, 16].

Problems

16.1 Integral imaging or Lippmann photography is another lens array-based modality which aims to provide a 3D display/imaging capability. Integral imaging systems consist of a lens array, each lens of which captures multiple perspectives of a 3D object on a sensor as shown in Fig. 16.8. Image of the object is first recorded using a lens array. Alternatively if the object can be modeled computationally as a 3D point cloud, image of each point in the point cloud may be computed for every lens in the array. The recorded (or computed) images when displayed on a 2D display screen and viewed through the same lens array, an observer gets a perception of 3D object. Assuming an $M \times M$ lens array with lens diameter D , find the largest object at nominal distance of z_0 from lens array that will not cause overlap of sub-images. Is the observed image pseudo-scopic? Explain how this system is different from a holographic replay.

16.2 Consider the plenoptic camera shown in Fig. 16.5. What is the relation between the image side F-number of the main lens and the individual microlenses so that the images recorded by individual microlenses do not overlap?

16.3 Consider the compact multi-lens imaging system described in Sect. 16.1 operating in ambient (almost) spatially incoherent illumination. Explain if the overall

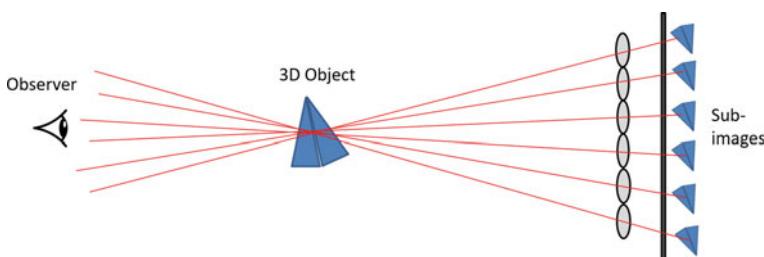
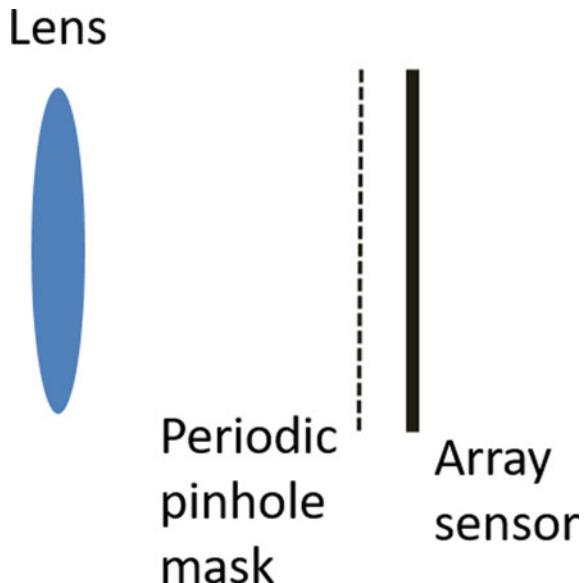


Fig. 16.8 Configuration of an integral imaging display

Fig. 16.9 Pinhole array-based plenoptic camera



computational image resolution is equal to the diffraction limit corresponding to the individual lenses or the whole lens aperture.

16.4 A camera system as shown in Fig. 16.9 consists of a main lens and a sensor. A periodic pinhole mask is placed in front of the sensor. Explain if this system can work like a plenoptic camera. Where will you position the pinhole array? What will be the effective resolution of the reconstructed images? See more details in [11].

16.5 Shack-Hartmann wavefront sensor (SHWS): The multi-lens array can be used with coherent wavefronts to measure the phase of the wavefronts in an SHWS. A plane wavefront incident normal to the lens array forms a periodic array of spots behind the lens array in their back-focal plane. A local tilt in the wavefront leads to displacement of the focal spot which may be used to estimate wavefront slope. For a given lens array arrangement, estimate the largest wavefront slope that such a configuration can measure.

References

1. S.A. Shroff, K. Berkner, Image formation analysis and high resolution image reconstruction for plenoptic imaging systems. *Appl. Opt.* **52**, D22–D31 (2013)
2. J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, Y. Ichikoa, Thin observation module by bound optics (TOMBO): concept and experimental verification. *Appl. Opt.* **40**, 1806–1813 (2001)

3. M. Shankar, R. Willett, N. Pitsianis, T. Schultz, R. Gibbons, R. Te Kolste, J. Carriere, C. Chen, D. Prather, D. Brady, Thin infrared imaging systems through multichannel sampling. *Appl. Opt.* **47**, B1–B10 (2008)
4. P. Milanfar, *Super Resolution Imaging*, 1st edn. (CRC Press, Boca Raton, 2011)
5. M. Elad, A. Feuer, Super-resolution reconstruction of image sequences. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 817–834 (1997)
6. K. Nitta, R. Shogenji, S. Miyatake, J. Tanida, Image reconstruction for thin observation module by iterative backprojection. *Appl. Opt.* **45**, 2839–2900 (2006)
7. K. Choi, T.J. Schultz, Signal processing approaches for image resolution restoration for TOMBO imagery. *Appl. Opt.* **47**, B104–B116 (2008)
8. A. Papoulis, Generalized sampling expansion. *IEEE Trans. Circuits Syst.* **24**, 652–654 (1977)
9. J.L. Brown, Multi-channel sampling of low pass signals. *IEEE Trans. Circuits Syst.* **28**, 101–106 (1981)
10. A.D. Portnoy, N. Pitsianis, X. Sun, D. Brady, Multichannel sampling schemes for optical imaging systems. *Appl. Opt.* **47**, B76–B85 (2008)
11. E.H. Adelson, J.Y.A. Wang, Single lens stereo with a plenoptic camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 99–106 (1992)
12. E.H. Adelson, J.Y.A. Wang, Single lens stereo with a plenoptic camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 99–106 (1992)
13. Commercial Lightfield system and literature at www.lytro.com
14. R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, P. Hanrahan, Light field photography with a hand-held plenoptic camera. *Comput. Sci. Stanf. Tech. Rep. CSTR* **2**, 1–11 (2011)
15. X. Xiao, B. Javidi, M. Martinez-Corral, A. Stern, Advances in three-dimensional integral imaging: sensing, display and applications. *Appl. Opt.* **52**, 546–560 (2013). (and references therein)
16. G. Wetzstein, I. Ihrke, D. Lanman, W. Heidrich, Computational plenoptic imaging. *Comput. Graph. Forum* **30**, 2397–2426 (2011)

Chapter 17

PSF Engineering



In Chap. 12, we studied the canonical $4F$ system model for imaging. It was observed that carefully designed amplitude and/or phase masks in the Fourier plane of the system lead to a number of visual enhancements. The ideas like dark-field, phase contrast, etc. have already been incorporated into microscopy systems and are being routinely used. The effect of the Fourier plane mask on the frequency response of the system was also discussed in terms of the optical transfer function (OTF). With the ready availability of computational power in the 1990s, a number of ideas emerged that suggested enhanced imaging performance by engineering PSF of the system [1–4]. A common interesting feature of these PSF engineering ideas is that they essentially perform controlled blurring of images at the recording stage in order to record a coded image which may not be immediately meaningful visually. On de-blurring the recorded images, the resultant system performance is however enhanced significantly with respect to depth of focus or achieving 3D tomographic imaging. It is well-known that for a typical imaging system, the depth of focus is given by $\lambda/(NA)^2$ where λ is the nominal wavelength of illumination and NA is the numerical aperture. The designs we will discuss here give a performance that is an order of magnitude higher than this typical depth of field value. The main aim of the designs to be discussed here is to make the system point spread function invariant with respect to defocus. This feature allows an image de-blurring process (like Wiener filter or iterative reconstruction) to bring all the object planes in focus simultaneously.

17.1 Extending Depth of Focus from a Focal Stack

Extending the effective depth of focus of a microscope system using a stack of images was first demonstrated by Hausler [5]. The idea behind Hausler's work is to realize that a 2D image recorded by an incoherent imaging system effectively contains contributions from a stack of slices of the sample. If an object is therefore moved through the system focus and all the individual images are integrated, then the effective PSF of the integrated image is nearly depth-invariant. This idea may be understood more

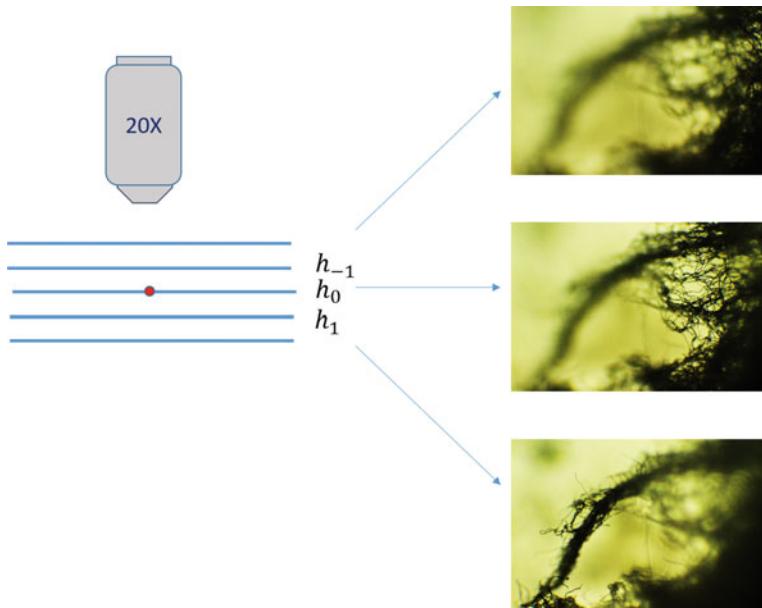


Fig. 17.1 Imaging of a z-stack of sample slices. The slice with red dot shows the perfect focus plane. Three different images of a 3D object are shown where different parts of the object are seen in focus in each of the images. [Image stack of fungus recorded with a 20X objective in collaboration with Darya Warner.]

clearly from Fig. 17.1. The object slices separated by distance Δz are shown next to a microscope objective lens and may be considered to be separated by the depth of focus of the imaging system. We will denote the system PSF for the perfect focus plane (shown with a red dot) as $h_0(x, y)$. The PSFs for two sides of the perfect focus plane are denoted as $\dots, h_{-2}(x, y), h_{-1}(x, y), h_1(x, y), h_2(x, y), \dots$, etc. Denoting the object function in different slices by $g_k(x, y)$ with $k = \dots, -2, -1, 0, 1, 2, \dots$, we observe that the total recorded image when the microscope is focused on the nominal central plane (shown with a red dot) may be described as

$$I_0(x, y) = \sum_k [h_k(x, y) * g_k(x, y)]. \quad (17.1)$$

If the objective lens is now translated by $m\Delta z$, we observe that the recorded image now corresponds to

$$I_{m\Delta z}(x, y) = \sum_k [h_{k+m}(x, y) * g_k(x, y)]. \quad (17.2)$$

Observe that the expression in the summation above has remained essentially the same as in Eq. (17.1) but just that the same slices are now contributing to the recorded

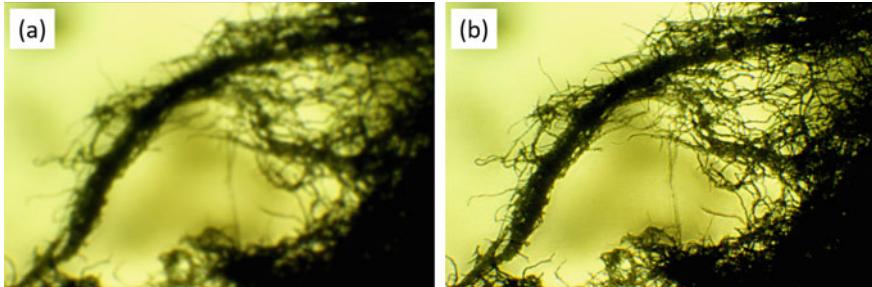


Fig. 17.2 Illustration of all-in-focus imaging: **a** Composite image obtained by aligning and adding the recorded image stack, **b** Deconvolution of the composite image by a single composite impulse response. [Image stack of fungus recorded with a 20X objective in collaboration with Darya Warner.]

image with respect to sequentially shifted defocused impulse responses. We observe that if all the recorded images in the stack are now added after appropriate alignment, the resultant image is expressible as

$$\sum_m I_{m \Delta z}(x, y) \approx \sum_m h_m(x, y) * \sum_k g_k(x, y). \quad (17.3)$$

In the first term above, the subscript k on the impulse response is omitted as all the impulse responses are now contributing to each slice. An important point to note here is that all slices $g_k(x, y)$ are now blurred by the same impulse response making the defocus blur approximately depth-invariant. A depth-invariant blur is amenable to single deconvolution for all-in-focus image recovery as we demonstrate in Fig. 17.2. Figure 17.2a shows the composite image obtained by aligning a stack of 25 images recorded with a microscope. The three individual images from the image stack shown in Fig. 17.1 each have some features in focus and others in defocus. Since the defocus blur is depth-dependent and the depth location of different features in the field of view cannot be known *a priori*; such images are difficult to de-blur by digital processing. The composite image in Fig. 17.2a however has features at all depths blurred by a nearly depth-invariant impulse response. The composite image may therefore be de-blurred using any of the direct or iterative algorithms discussed in Chaps. 5, 6. The sample in this case has a total depth which is over ten times larger than the depth of focus of the microscope. The de-blurring of the composite image is shown in Fig. 17.2b which is an all-in-focus image with features from all depths appearing sharp. Unlike Hausler's original technique, in the present illustration we have accounted for magnification changes in features at various depths by aligning the images in the stack with the affine transformation (translation, rotation, scaling) that is readily available in image registration tools which enhances the overall performance of the technique. In the next sections of this chapter, we will discuss wavefront coding techniques which achieve depth-invariant PSF by introducing phase masks in the pupil or aperture of the imaging system. While these techniques have received

a lot of due attention, Hausler's technique can be potentially incorporated into any microscope at low cost by a simpler method of collecting a stack of images and then deconvolving the composite image.

17.2 Extended Depth of Field with Cubic Phase Mask

The usage of cubic phase mask for extending the depth of field was first proposed by Cathey and Dowski in [6]. This is one of the early computational optical imaging systems that combined optical design with image recovery to obtain superior imaging performance. As we will discuss here, the cubic phase mask function may be arrived at by considering the behavior of the optical transfer function (OTF) of an imaging system for variable defocus. We will assume the system aperture function to be separable in x and y dimensions and limit our discussion to a 1D case for simplicity. Further, we will present the discussion in terms of a 4F or canonical optical processor. The OTF of the 4F system is given by

$$OTF(u) = \frac{\int dx P(x + u/2) P^*(x - u/2)}{\int dx |P(x)|^2}. \quad (17.4)$$

In terms of the symbols in Sect. 12.4 of Chap. 12, the variable $u = \lambda F f_x$ and P denotes the system aperture function. For an on-axis point source defocused with respect to the object or input plane of a canonical optical processor, the aperture function may be suitably represented as

$$P(x) = P_0(x) \exp(i\alpha x^2). \quad (17.5)$$

Here, $P_0(x)$ represents the system aperture (with any suitable phase mask) and the quadratic phase factor arises due to defocus of an object point from the nominal object plane. The parameter α controls the magnitude of defocus. The idea here is to design the aperture $P_0(x)$ such that the system OTF becomes almost independent of α . The OTF may now be expressed as

$$OTF(u) = A \int dx P_0(x + u/2) P_0^*(x - u/2) \exp(i\alpha x u), \quad (17.6)$$

where we have included the normalization in the constant A . Assuming a trial power law form for the aperture phase of the form:

$$P_0(x) = \exp(i\beta x^\gamma), \quad (17.7)$$

the OTF may be expressed as

$$OTF(u) = A \int dx \exp(i\theta(x)), \quad (17.8)$$

where

$$\theta(x) = [\beta(x + \frac{u}{2})^\gamma - \beta(x - \frac{u}{2})^\gamma] + \alpha xu. \quad (17.9)$$

The integrals with reasonably fast phase variation as above can be approximated by stationary phase method [7]. The idea behind the stationary phase method is that the integrand oscillates slowest near the stationary points of the phase function, and these regions on x -axis contribute to the overall integral the most. Denoting the stationary point of $\theta(x)$ by x_0 , the integral may be approximated as

$$A \int dx \exp(i\theta(x)) \approx \frac{A}{2} \sqrt{\frac{2\pi}{\theta''(x_0)}} \exp[i\theta(x_0)]. \quad (17.10)$$

If we want the amplitude term in the approximation above to be independent of the defocus parameter α , it is required that the second derivative of $\theta(x)$ at $x = x_0$ is independent of α (or $[\partial\theta''(x_0)/\partial\alpha] = 0$). Since the stationary point x_0 is determined using

$$\theta'(x_0) = \alpha u + \gamma \beta \left[\left(x_0 + \frac{u}{2}\right)^{\gamma-1} - \left(x_0 - \frac{u}{2}\right)^{\gamma-1} \right] = 0, \quad (17.11)$$

we require that x_0 be linear in α . This condition can be achieved if $\gamma = 3$ which will result in a cubic phase mask. For $\gamma = 3$, we observe that the stationary point is given by

$$x_0 = -\frac{\alpha}{6\beta}. \quad (17.12)$$

The system OTF is thus obtained as per Eq. (17.10):

$$OTF(u) \approx \frac{A}{2} \sqrt{\frac{\pi}{3\beta u}} \exp[-i\frac{\alpha^2 u}{12\beta} + i\frac{\beta u^3}{4}]. \quad (17.13)$$

We observe that the MTF of the system which is equal to the magnitude of the OTF is independent of the defocus parameter α . The phase part of OTF has two terms, of which the second one is independent of α . The first term in the phase function is linear in u and will thus cause translation or shift in the point spread function (PSF) of the system with changing defocus. The effect of this α -dependent term may however be made negligible by selecting a phase mask function with a large value of β . In the illustration in Fig. 17.3d-f, we show the PSF of a cubic phase mask system for three distinct values of the defocus parameter $\alpha = 0, 4, 8$ respectively with $\beta = 25$. For comparison, the system PSFs for a system without the cubic phase mask is shown for the same values of α in Fig. 17.3a, b, c. We see that the PSF for the cubic phase

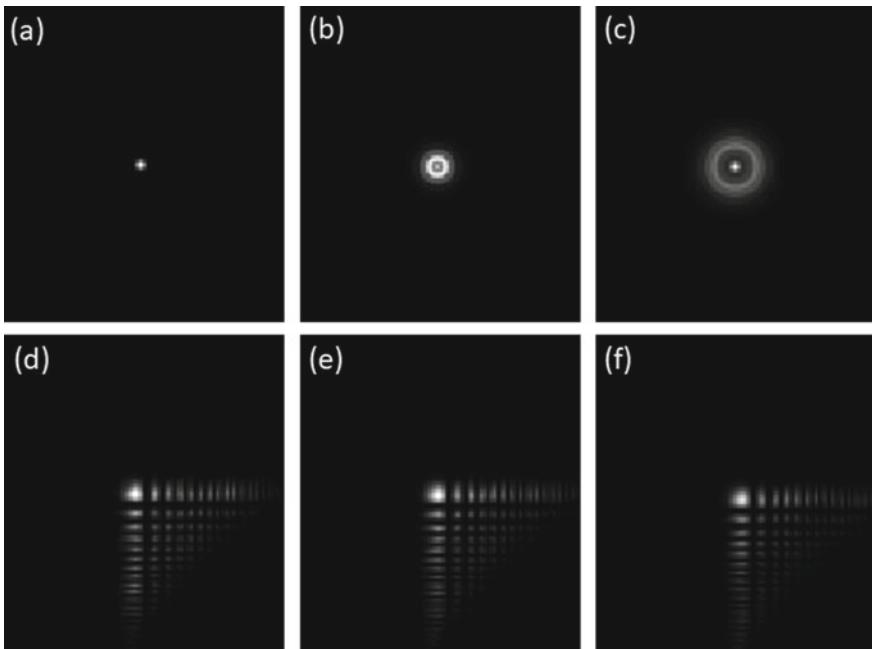


Fig. 17.3 Effect of cubic phase mask on system PSF. **a–c** PSF for systems with defocus parameter $\alpha = 0, 4, 8$ and no cubic mask. **d–f** System PSF for same values of α where a cubic phase mask $\beta(x^3 + y^3)$ has been introduced in the system aperture with $\beta = 25$. Here x and y are normalized coordinates in the aperture

mask system has remained almost the same over the range of α values. Objects at different depths relative to the nominal object plane are therefore blurred in a similar manner. Unlike the Hausler method discussed earlier, the cubic phase mask PSF is quite asymmetric. The advantage of nearly depth-independent blur function is that objects at different depths can be de-blurred by a single deconvolution operation. A combination of hardware design and digital processing thus makes it possible to have an imaging system with an extended depth of field. *The cubic phase function as discussed here for extending the depth of field is also known in the context of the so-called non-diffracting solutions of free-space diffraction problem that are known as the Airy beams.* Airy beams have found several recent applications in the manipulation and trapping of micro-particles in liquids. We will not discuss the topic of Airy beams further but refer the reader to references [8–10].

17.3 Extended Depth of Focus Using the Log-Asphere Lens

In this section, we will discuss an alternate model for obtaining depth-invariant PSF using a circularly symmetric lens design as first suggested in the work of Chi and George [11–13]. The basic idea behind this design is explained in Fig. 17.4. The lens element can be broken into concentric rings such that an elemental annular ring with radius r is designed to form an image of an object point at distance s from the lens. The resultant image of an object point as recorded on the array detector consists of a focused image due to one annular lens ring and defocused images due to other annular rings. As shown in Fig. 17.4, a lens element may be designed to have a depth of field ranging over distances s_1 to s_2 along the system optic axis. An annular ring of radius r can be considered to focus a point at a distance $s(r)$ from the lens such that

$$s(r) = s_1 + (s_2 - s_1) \frac{r^2}{R^2}, \quad (17.14)$$

where R is the radius of the lens element. The phase delay function for the lens may be derived using Fermat's principle. The optical path length for a point at $s(r)$ imaged at the on-axis point on the detector is given by

$$L(r) = \sqrt{s^2 + r^2} + \frac{\lambda}{2\pi} \phi(r) + \sqrt{z_0^2 + r^2}. \quad (17.15)$$

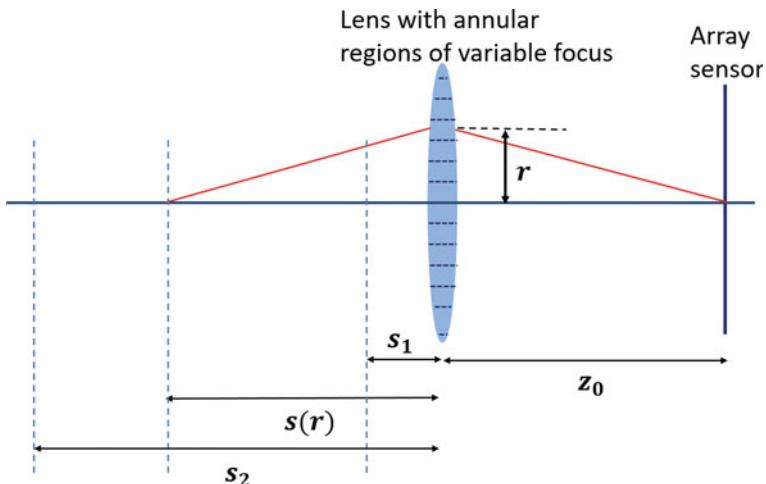


Fig. 17.4 Extended depth of field using variable focus lens: concept

Solving for the stationary point according to the condition $\partial L/\partial r = 0$ (Fermat's principle) provides a solution for the phase function of the desired lens given by

$$\begin{aligned}\phi(r) = & -\frac{2\pi}{\lambda} \left(\sqrt{r^2 + z_0^2} - z_0 \right) \\ & - \frac{\pi}{\lambda} \frac{1}{\gamma} \left[\log \left(2\gamma \left\{ \sqrt{r^2 + (s_1 + \gamma r^2)^2} + (s_1 + \gamma r^2) \right\} + 1 \right) - \log(4\gamma s_1 + 1) \right].\end{aligned}\quad (17.16)$$

Here, the parameter γ is defined as

$$\gamma = \frac{s_2 - s_1}{R^2}, \quad (17.17)$$

and z_0 is the distance between the lens and the detector. It may be noted that the first term in the expression for $\phi(r)$ may be identified with a phase function for a spherical wave converging toward the on-axis detector point from the lens aperture. The second logarithmic term arises out of Fermat's principle and is the reason for the name log-asphere lens. In Fig. 17.5a–c, we show the PSFs for the log-asphere lens over a depth range corresponding to the defocus values in Fig. 17.3a–c, respectively. The PSFs are again seen to be approximately depth-invariant. Further since the PSFs have a form of a symmetric blurring function, a relatively simple operation similar to an edge-enhancing filter can lead to the recovery of an all-in-focus image. More sophisticated iterative approaches to image recovery are also possible and give further enhancement in image quality [13]. The radially symmetric nature of the log-asphere design allows it to be expressed in the form of an even-order polynomial in radial coordinate r . Radially symmetric functions have a distinct advantage from a fabrication point of view over odd-order polynomial designs like the cubic phase.

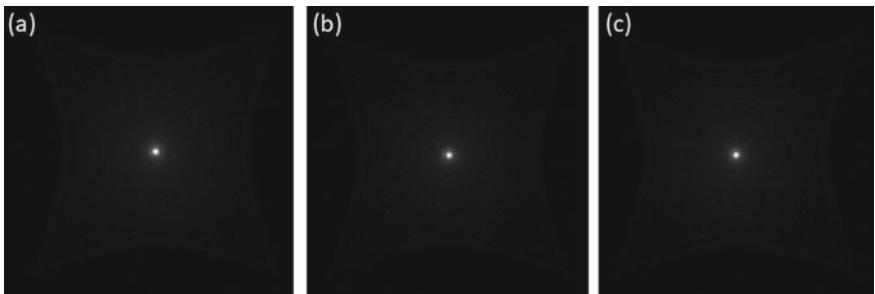


Fig. 17.5 a–c PSF for log-asphere lens for object points at variable depths corresponding to the defocus values in Fig. 17.3a–c

17.4 Rotating Point Spread Functions

Rotating point spread functions are another important development in the context of pupil function engineering [14–17]. The pupil functions in this case as first studied in the work of Schechner, Piestun and Shamir are linear combinations of the Laguerre-Gaussian (LG) functions [18]. The LG functions are the solutions of the para-axial wave equation in the cylindrical coordinate system. An LG mode is defined by means of two indices (l, p) such that the function $LG_{l,p}(r, \phi, z)$ is defined as

$$LG_{lp}(r, \phi, z) = \frac{C_{lp}}{w(z)} \left(\frac{r\sqrt{2}}{w(z)} \right)^{|l|} \exp \left[-\frac{r^2}{w^2(z)} - i \frac{\pi r^2}{\lambda R(z)} \right] \times L_p^{|l|} \left(\frac{2r^2}{w^2(z)} \right) \exp[-il\phi + i(2p + |l| + 1)\psi(z)]. \quad (17.18)$$

Here, C_{lp} is a normalization constant. Other parameters in the above equation are beam waist

$$w(z) = w_0 \sqrt{1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2}, \quad (17.19)$$

with w_0 being the initial ($z = 0$) spot size,

$$R(z) = z \left[1 + \left(\frac{\pi w_0^2}{\lambda z} \right)^2 \right] \quad (17.20)$$

is the radius of curvature of the beam and

$$\psi(z) = \arctan \left(\frac{\lambda z}{\pi w_0^2} \right) \quad (17.21)$$

is the Gouy phase. The functions $L_p^{|l|}$ are the generalized Laguerre polynomials. It was shown in the early work of Schechner, Piestun and Shamir that if wavefields are generated by coherent superposition of LG_{lp} modes such that mode labels (l_j, p_j) ($j = 1, 2, 3, \dots, N$) satisfy

$$\frac{p_{j+1} - p_j}{l_{j+1} - l_j} = \text{constant}, \quad (17.22)$$

the resultant intensity PSF patterns approximately retain their functional form over a distance such that the whole pattern rotates continuously with defocus. In Fig. 17.6, we show the behavior of intensity PSF obtained by linear superposition of the LG modes denoted by indices $(1, 1), (2, 2), (3, 3), (4, 4), (5, 5)$ with equal energy in each mode. The intensity PSF at three distinct planes is shown to illustrate the rotation of the pattern. If a phase mask approximating such a linear combination of LG

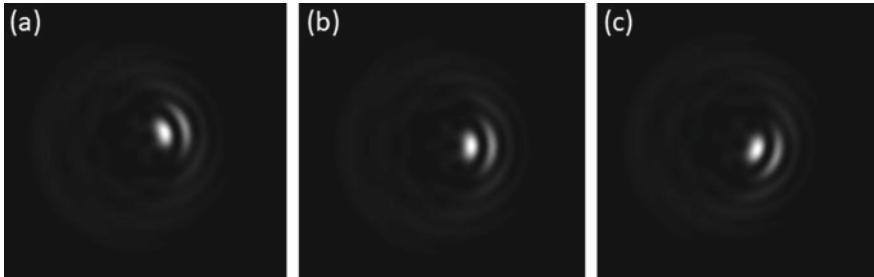


Fig. 17.6 a–c Illustration of rotating PSF with defocus. The PSF is generated by a mask which corresponds to a linear superposition of LG modes denoted by (l, p) indices $(1, 1), (2, 2), (3, 3), (4, 4)$ and $(5, 5)$. It is observed that the PSF retains its functional form but rotates continuously with defocus

modes is placed in the pupil plane of an imaging system, then distinct object planes at varying depths will be blurred by such rotating PSFs. Such pupil plane masks may be designed to produce PSF rotation of say 180 degrees over the desired depth. The PSF rotation in a sense provides depth encoding along with an extended depth of field feature. For example, when the recorded blurred image is de-blurred using a PSF with a particular orientation, the recovered image corresponds to a particular slice in the object along the depth dimension. Such optical sectioning capability is particularly important for fluorescence microscopy where individual sparsely distributed point sources are easily localized in depth by using such a rotating PSF. In recent years, localization microscopy techniques have gained prominence due to their super-resolution capability. The rotating PSF ideas further provide the possibility of 3D localization of fluorophores thereby enabling super-resolved 3D microscopy [19, 20].

Problems

17.1 Implement a computer program to find the incoherent PSF of an imaging system whose pupil function is a vortex phase function $\exp(i\phi)$.

17.2 Rotated vortex pair: Consider a pupil function which is made up of the multiplication of a rotated vortex pair with transmission functions $T_1(\rho, \phi)$ and $T_2(\rho, \phi)$ given by

$$T_{1,2}(\rho, \phi) = \exp[\mp i 2\pi \alpha R(\rho)\phi] \text{circ}(\rho/a_0).$$

Show that the composite transmission function

$$T(\rho, \phi) = T_1(\rho, \phi + \beta/2) T_2(\rho, \phi - \beta/2)$$

is independent of the polar angle ϕ . Further, if the radial function $R(\rho)$ is selected to be quadratic, show that the resultant rotated vortex pair phase mask leads to a varifocal system depending on the relative rotation angle β between the vortex pair. Explain the similarities with what is known as the Alvarez-Lohmann lens. For more details, the reader is referred to [21].

17.3 COSA phase mask The cropped oblique secondary astigmatism (COSA) phase mask [22] is defined by the phase function:

$$\psi(\rho, \phi) = \alpha (\rho^4 - \rho^2) \sin(2\phi). \quad (17.23)$$

Compute the PSF of an imaging system whose pupil function is modeled by the COSA phase. Observe the evolution of PSF with depth and strength α of the phase mask function.

17.4 The PSF of a system consists of multiple randomly placed impulses at (x_n, y_n) described as

$$h(x, y) = \text{jinc}^2(\dots) * \sum_{n=1}^N \delta(x - x_n, y - y_n).$$

The $\text{jinc}^2(\dots)$ function here represents the diffraction-limited spot size due to a circular system aperture. Determine a pupil phase mask that will lead to such a PSF. Show that the image recorded using such a PSF contains information about the image field beyond the physical sensor area [23].

17.5 Consider an imaging system whose pupil plane contains a random binary (0, 1) mask. Explain the effect of the mask on PSF and the computational depth of field.

References

1. J. Ojeda-Castañeda, P. Andres, A. Diaz, Annular apodizers for low sensitivity to defocus and to spherical aberration. *Opt. Lett.* **11**, 487–489 (1986)
2. J. Ojeda-Castañeda, R. Ramos, A. Noyola-Isgleas, High focal depth by apodization and digital restoration. *Appl. Opt.* **27**, 2583–2586 (1988)
3. J. Ojeda-Castañeda, L.R. Berriel-Valdos, Zone plate for arbitrarily high focal depth. *Appl. Opt.* **29**, 994–997 (1990)
4. W. Chi, K. Chu, N. George, Polarization coded aperture. *Opt. Express* **14**, 6634–6642 (2006)
5. G. Hausler, A method to increase the depth of focus by two step image processing. *Opt. Commun.* **6**, 38–42 (1972)
6. E.R. Dowski, W.T. Cathey, Extended depth of field through wavefront coding. *Appl. Opt.* **34**, 1859–1866 (1995)
7. L. Mandel, E. Wolf, *Coherence and Quantum Optics* (Cambridge University Press, 1995)
8. M.V. Berry, N.L. Balazs, Nonspreadng wave packets. *Am. J. Phys.* **47**, 264–267 (1979)
9. G.A. Siviloglou, J. Broky, A. Dogariu, D.N. Christodoulides, Observation of accelerating Airy beams. *Phys. Rev. Lett.* **99**, 213901 (2007)

10. I. Kaminer, R. Bekenstein, J. Nemirovsky, M. Segev, Nondiffracting accelerating wave packets of Maxwell's equations. *Phys. Rev. Lett.* **108**, 163901 (2012)
11. W. Chi, N. George, Electronic imaging using a logarithmic asphere. *Opt. Lett.* **26**, 875–877 (2001)
12. N. George, W. Chi, Extended depth of field using a logarithmic asphere. *J. Opt. A: Pure Appl. Opt.* **5**, S157–S163 (2003)
13. W. Chi, N. George, Computational imaging with logarithmic asphere: theory. *J. Opt. Soc. Am. A* **20**, 2260–2273 (2003)
14. R. Piestun, Y.Y. Schechner, J. Shamir, Propagation invariant wavefields with finite energy. *J. Opt. Soc. Am. A* **17**, 294–303 (2000)
15. A. Greengard, Y.Y. Schechner, R. Piestun, Depth from diffracted rotation. *Opt. Lett.* **31**, 181–183 (2006)
16. S.R.P. Pavani, R. Piestun, High efficiency rotating point spread functions. *Opt. Express* **16**, 3484–3489 (2008)
17. S. Prasad, Rotating point spread function via pupil function engineering. *Opt. Lett.* **38**, 585–587 (2013)
18. Y.Y. Schechner, R. Piestun, J. Shamir, Wave propagation with rotating intensity distributions. *Phys. Rev. E* **54**, R50–R53 (1996)
19. S.R.P. Pavani, M.A. Thompson, J.S. Biteen, S.J. Lord, N. Liu, R.J. Twieg, R. Piestun, W.E. Moerner, Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2995 (2009)
20. G.P.J. Laporte, D.B. Conkey, A. Vasdekis, R. Piestun, D. Psaltis, Double helix enhanced axial localization in STED-nanoscopy. *Opt. Express* **21**, 30984–30992 (2013)
21. J. Ojeda-Castañeda, Cristina M. Gomez-Sarabia, Tuning field depth at high resolution using pupil engineering. *Adv. Opt. Photonics* **7**, 814–880 (2015)
22. Y. Zhou, G. Carles, Precise 3D particle localization over large axial ranges using secondary astigmatism. *Opt. Lett.* **45**, 2466–2469 (2020)
23. R. Malik, R. Elangovan, K. Khare, Computational imaging with an extended field of view. *J. Opt.* **23**, 085703 (2021)

Chapter 18

Structured Illumination Imaging



As we have discussed so far in this book, computational imaging systems aim at breaking traditional limits by combining novel optics hardware and image recovery algorithms. Structured illumination is an important concept in this regard [1–6]. Microscopy systems using structured illumination are now commercially available for achieving image resolution performance beyond the diffraction limit. Structured illumination typically refers to high-frequency modulation of the illumination pattern in imaging applications. As we will see in this chapter, the knowledge of the high-frequency modulation pattern allows one to computationally recover images with a resolution higher than the diffraction-limited resolution offered by optics alone. The term “diffraction-limited resolution” itself needs to be looked at once again in view of the development of such computational imaging systems. In Chap. 15, we have already described the Fourier ptychographic imaging which is an example of coherent structured illumination in the form of tilted plane waves to achieve super-resolution imaging. In this chapter, our main focus will be on incoherent imaging systems that achieve super-resolution.

We will first explain the super-resolution concept with a simple picture of a microscope objective as in Fig. 18.1a where a sample is illuminated by a light beam traveling along the optic axis. For a usual microscope objective, the diffraction-limited resolution is decided by the numerical aperture which is related to the cone of the acceptance angle of the light rays. This is similar to the case of the 4F or canonical processor where the resolution is decided by the extent of the Fourier plane aperture. The light rays scattered from the object at larger angles compared to the light cone defined by the numerical aperture are not captured by the imaging system and hence do not contribute to image formation. The idea behind structured illumination is to modulate the illumination with a high spatial frequency pattern. The modulation of illumination patterns may be achieved by placing a structure like a grating near the object to be imaged. The role of this structure as shown in Fig. 18.1b is to scatter some of the high-frequency components (that were previously not captured by the optical system aperture) back into the aperture of the system. The computational problem now is to recover this high-frequency information in a visually suitable form

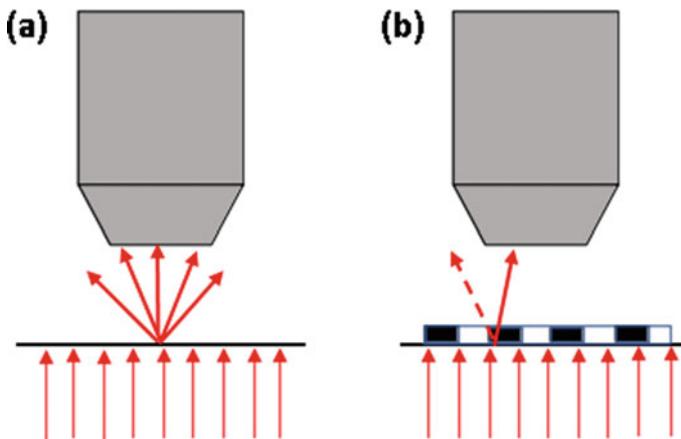


Fig. 18.1 Schematic diagram for understanding the basic principle behind structured illumination imaging. **a** When an object scatters light from illumination beam, some of it is not captured by the imaging system due to its limited aperture. **b** When a structure is present near the object, some of the light which would not have entered the aperture (dotted arrow) is now scattered by the high-frequency structure into the system aperture. The aim of the computational imaging algorithm is to recover this high-frequency information from the coded image data with the knowledge of the modulation pattern

using the knowledge of the modulation pattern. Typically multiple modulation patterns are used for capturing one high-resolution image beyond the diffraction-limited resolution capability of the imaging system.

18.1 Forward Model and Image Reconstruction

We will state the forward problem in structured illumination imaging as follows. Consider a modulation pattern $m_k(x, y)$ that effectively multiplies with the object $g(x, y)$ to be imaged. Here, k denotes the index of the modulation pattern used. If the imaging system point spread function (PSF) is denoted by $o(x, y)$, then the k th recorded image may be represented as a convolution:

$$q_k(x, y) = o(x, y) * [m_k(x, y)g(x, y)]. \quad (18.1)$$

Note that the recorded raw image has limited resolution due to the convolution with the PSF $o(x, y)$ since the corresponding OTF has a low-frequency cut-off in the Fourier space. For usual imaging systems that we have studied so far, the modulation pattern is typically $m_k(x, y) = 1$ across all pixels. The data $q_k(x, y)$ recorded with modulation however contains high-frequency information in a coded form which is to

be recovered computationally. While more general coding patterns are possible [7], in this chapter we will consider simple high-frequency sinusoidal modulation patterns to explain high-resolution image recovery in structured illumination imaging. In particular, we will start by considering a modulation pattern of the form:

$$m_1(x, y) = \frac{1}{2}[1 + \alpha \cos(2\pi f_0 x + \Phi_1)], \quad (18.2)$$

where α is a positive number in $(0, 1)$ representing the contrast of the illuminating fringes. It is instructive to look at the corresponding recorded image $q_1(x, y)$ in Fourier space:

$$\begin{aligned} Q_1(f_x, f_y) = \frac{1}{2} O(f_x, f_y) &[G(f_x, f_y) + \frac{\alpha}{2} G(f_x - f_0, f_y) e^{i\Phi_1} \\ &+ \frac{\alpha}{2} G(f_x + f_0, f_y) e^{-i\Phi_1}]. \end{aligned} \quad (18.3)$$

We see that the Fourier space extent of $Q_1(f_x, f_y)$ is limited by that of the system OTF as there is an overall multiplication by the function $O(f_x, f_y)$. At each location (f_x, f_y) in the Fourier domain, the data consists of a summation of three terms corresponding to $G(f_x, f_y)$ and its shifted versions with additional phases. By introducing three code patterns corresponding to three distinct values of Φ_k with $k = 1, 2, 3$, we can resolve this ambiguity. In particular, we can write a system of equations:

$$\begin{pmatrix} Q_1(f_x, f_y) \\ Q_2(f_x, f_y) \\ Q_3(f_x, f_y) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & e^{i\Phi_1} & e^{-i\Phi_1} \\ 1 & e^{i\Phi_2} & e^{-i\Phi_2} \\ 1 & e^{i\Phi_3} & e^{-i\Phi_3} \end{pmatrix} \begin{pmatrix} O(f_x, f_y)G(f_x, f_y) \\ \frac{\alpha}{2}O(f_x, f_y)G(f_x - f_0, f_y) \\ \frac{\alpha}{2}O(f_x, f_y)G(f_x + f_0, f_y) \end{pmatrix}. \quad (18.4)$$

At every sample point (f_x, f_y) in Fourier space, this system may be inverted to obtain the 2D Fourier transform $O(f_x, f_y)G(f_x, f_y)$ and its folded versions $O(f_x, f_y)G(f_x \pm f_0, f_y)$. We notice that the folded versions of the transform products $O(f_x, f_y)G(f_x \pm f_0, f_y)$ contain the high-frequency information that the system cannot record in the absence of the structured illumination. The modulation of the illumination thus offers new information that can now be utilized to get a high-resolution image using low-resolution optics. In practice, illumination modulation is applied in three directions to effectively extend the spatial bandwidth to about twice that of the extent of system OTF as shown in Fig. 18.2a, b. Three image records corresponding to three phase shifts are thus required along each of the three directions along which illumination with fringe modulation is applied. In general, after decoupling the three bands in each of the three illumination directions ($\theta = 0, 2\pi/3, 4\pi/3$) by solving a system of equations as in Eq. (18.4), we have information about nine overlapping bands in the Fourier space in the form of products of $O(f_x, f_y)$ with $G(f_x, f_y)$ and $G(f_x \pm f_0 \cos \theta, f_y \pm f_0 \sin \theta)$. Our aim is to recover $G(f_x, f_y)$ over the total available Fourier space, and this situation is what is explained with the

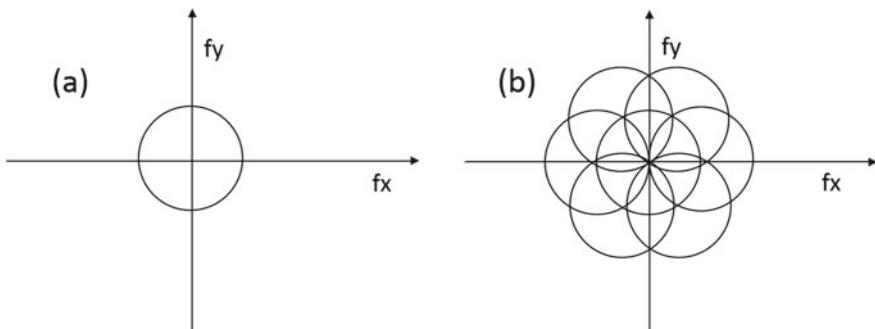


Fig. 18.2 **a, b** Diagrams illustrating Fourier space filling for system without and with structured illumination. The circle in **a** shows the extent of system OTF

generalized Wiener filter discussed in Sect. 5.5. Figure 18.3 illustrates the structured illumination imaging concept. In Fig. 18.3a–c, we show the cameraman object with three typical sinusoidal structured illuminations along directions $0, \pi/3$ and $2\pi/3$. The image of the cameraman object with resolution limited by the native pass-band of the imaging system is shown in Fig. 18.3d. When three phase-shifted illuminations are used in each of the three directions (as in Fig. 18.3a–c), the information in various Fourier bands can be decoupled as per Eq. (18.4). Knowledge of the shifted Fourier bands now can be utilized in the formula for the generalized Wiener filter and the corresponding high-resolution image reconstruction is shown in Fig. 18.3e. The Fourier band structure of interest is shown in Fig. 18.3f where the central dotted yellow circle represents the native system pass-band.

While a bare-bones simplistic discussion of the main ideas involved in structured illumination imaging is provided here, a practical SIM system needs to pay attention to a number of allied aspects such as accurate fringe orientation estimation, low photon counts, bleaching of the fluorescent samples, etc. Beyond the linear generalized Wiener filter-based recovery for SIM systems, a number of advanced approaches have been developed that are iterative optimization techniques. Such methods nominally pose the SIM reconstruction as a problem of minimizing a cost function of the form:

$$C(g) = \sum k \{q_k(x, y) - o(x, y) * [m_k(x, y) g(x, y)]\}^2 + \Psi(g). \quad (18.5)$$

The first term above is the data consistency term, while the second term is a penalty term as discussed in Chap. 6. Such an approach offers a certain advantage over linear recovery in terms of handling noise, reduced number of coded measurements, etc. [8–10].

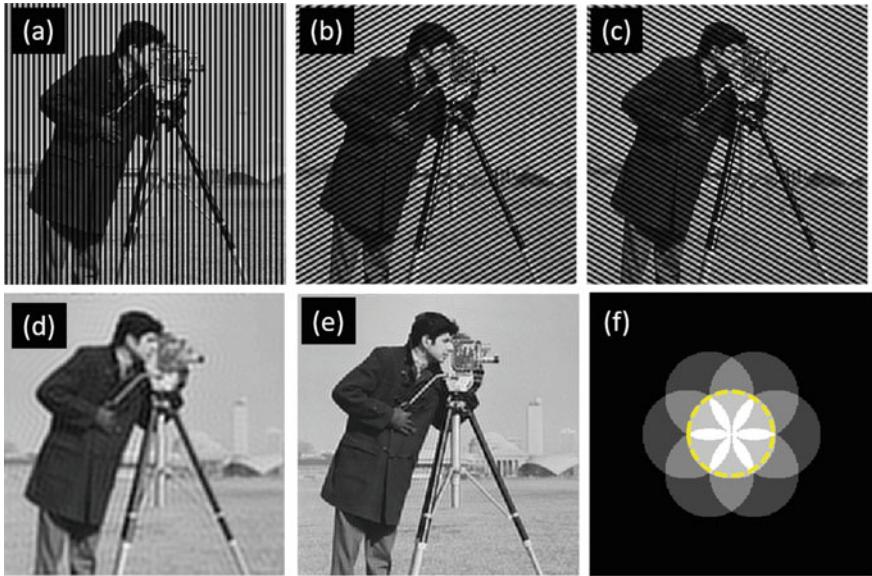


Fig. 18.3 **a–c** Object with structured sinusoidal illumination in three directions ($0, \pi/3, 2\pi/3$), **d** Low-resolution image obtained using native pass-band of imaging system, **e** High-resolution image recovery using generalized Wiener filter and **f** Effective band structure of the recovered image, dotted yellow circle shows native pass-band of the imaging system

18.2 Other Super-Resolution Microscopy Techniques

In this section, we will provide a brief discussion of two other important super-resolution microscopy modalities—stimulated emission depletion (STED) microscopy and stochastic reconstruction microscopy (STORM). Both these techniques have gained a lot of popularity in the bio-science community. The image reconstruction problem in these techniques is of a slightly different nature compared to structured illumination. These techniques depend on the behavior of fluorescent dye molecules to some extent.

18.2.1 Simulated Emission Depletion Microscopy (STED)

The main idea behind the STED technique is the possibility of selective deactivation of a fluorophore by a second illumination other than the one used for activating fluorescence [11]. In the STED technique, the usual fluorescence

process is halted by stimulated emission to a red-shifted level which can be filtered out. The interesting aspect of the methodology is that while the excitation is performed by a laser pulse with a Gaussian profile, the second depletion pulse sent in quick succession is shaped like an optical vortex (or donut-like) beam (see Fig. 18.4). The vortex beam has a null in the center whose size is smaller than the diffraction-limited spot size of the microscope. As a result, the fluorescence is depleted over a donut-like region leaving the central null intact. At a given location, the fluorescence is thus observed from a sub-diffraction area which determines the spatial resolution of a STED microscope. The STED systems operate in scanning mode and are therefore often coupled with a confocal scanning microscope. STED microscopes enable resolution down to the range of 50–100 nm.

18.2.2 Stochastic Optical Reconstruction Microscopy (STORM)

The resolution of a fluorescence microscope is usually limited by diffraction or alternatively by the numerical aperture of the objective lens. The STORM approach uses blinking fluorescent dyes. Further, the excitation illumination is controlled such that only a small subset of fluorescent centers are activated in one exposure frame. The recorded single image frame thus typically consists of multiple sparsely spaced PSF patterns (image of a number of sparsely located individual point sources). The centroids of individual PSF spots in the image can be computed with accuracy better than the diffraction limit (see Fig. 18.5), and this idea is utilized to form a composite high-resolution image by combining a large number of individual image frames after localizing their PSF centroids [12–14]. The accuracy in the determination of the centroid is dependent on the number of photon counts in a PSF spot and any pixelation effects. STORM systems usually employ highly sensitive electron multiplier CCD detectors. With accurate centroid detection algorithms, the resolution of a STORM system can reach in the 20–30 nm range.

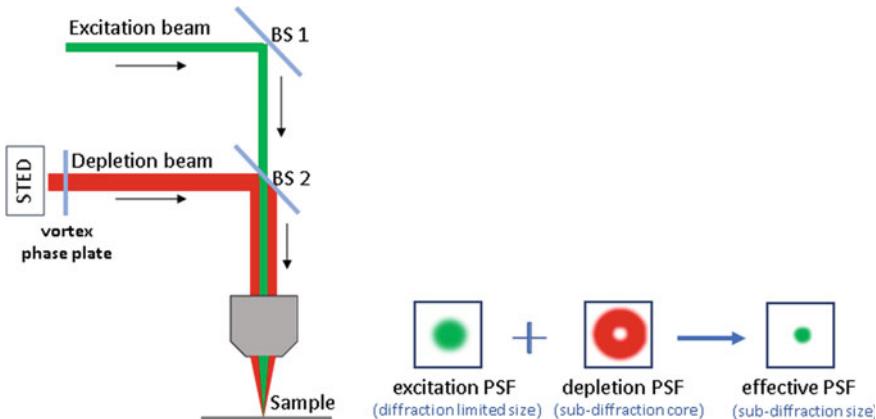


Fig. 18.4 Typical setup of STED microscope

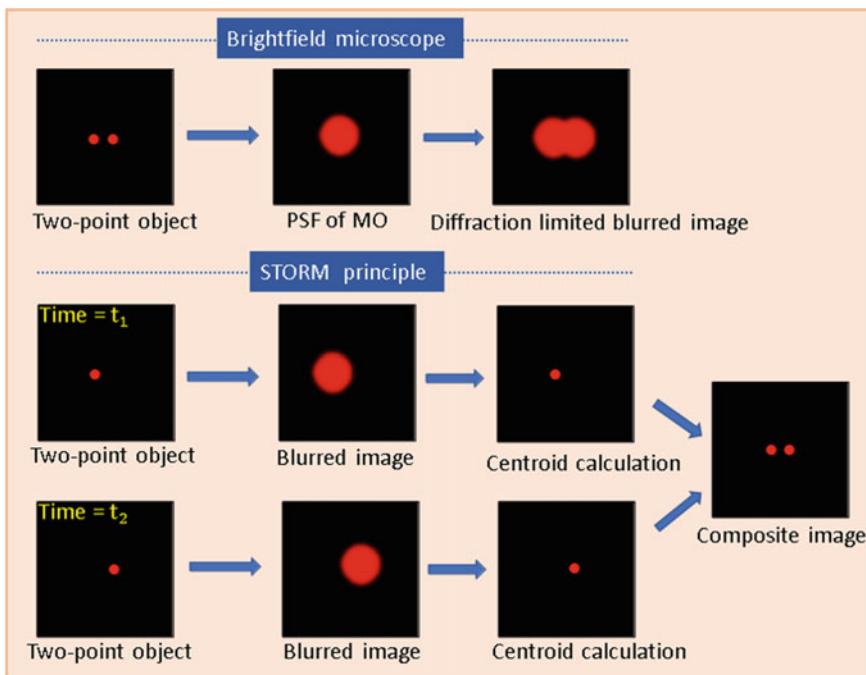


Fig. 18.5 Illustration showing the principle of STORM

Problems

18.1 Super-resolution with the Gerchberg-Papoulis algorithm Implement an algorithm similar to the Error Reduction method in phase retrieval to determine an object $g(x, y)$ of known support $S(x, y)$ from the knowledge of its Fourier transform $G(f_x, f_y)$ over a finite region $W(f_x, f_y)$. Denoting the Fourier transform of the n th guess solution $g^{(n)}(x, y)$ (which has been limited to support $S(x, y)$) by $G'_n(f_x, f_y)$, the $(n + 1)$ th iteration may proceed as

$$G^{(n+1)}(f_x, f_y) = G(f_x, f_y)W(f_x, f_y) + G'_n(f_x, f_y)[1 - W(f_x, f_y)].$$

By using an additional constraint such as positivity on $g(x, y)$, show that the above iteration starts filling the Fourier space beyond the known window region $W(f_x, f_y)$. For simulation purposes, you may use a simple two-point object [15, 16].

18.2 Show that a structured illumination microscopy system which uses the same objective lens for illumination and imaging can enhance the image resolution by a maximum factor of 2 beyond the native diffraction limit of the system.

18.3 Pixel reassignment[17]: A scanning confocal imaging system is shown in Fig. 18.6. Explain what happens to the effective imaging PSF when the pinhole is displaced laterally by distance Δx from its ideal confocal position. Note that the system PSF in this case is the product of illumination PSF and the imaging PSF.

18.4 Digital super-resolution: Find the resolution of a camera with F-number equal to 2 and compare it with a typical pixel size of $5 \mu\text{m}$. If the array sensor is given a number of sub-pixel displacements (periodic or random) and multiple image frames

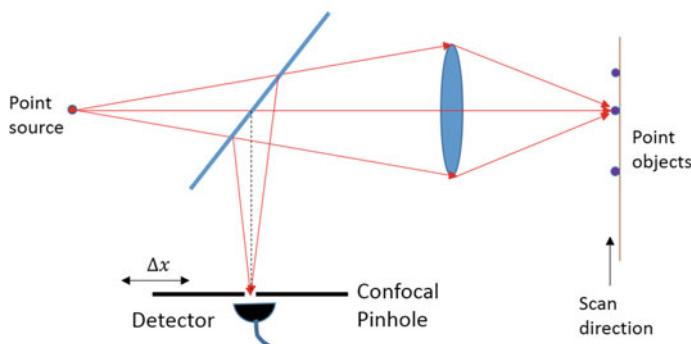


Fig. 18.6 Scanning confocal imaging system

of a scene are captured, show that it is possible to generate an image with a higher pixel number compared to that of the camera sensor. Will such a system have a higher resolution compared to the diffraction limit of the imaging lens?

18.5 Explain how the rotating PSF concept discussed in Chap. 17 may be used for achieving super-resolution in lateral as well as depth directions.

References

1. M.G.L. Gustafsson, Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Microscopy* **198**, 82–87 (2000)
2. P. Kner, B.B. Chun, E.R. Griffis, L. Winoto, M.G. Gustafsson, Super-resolution video microscopy of live cells by structured illumination. *Nat. Methods* **6**, 339–342 (2009)
3. M. Saxena, G. Eluru, S.S. Gorthi, Structured illumination microscopy. *Adv. Opt. Photon.* **7**, 241–275 (2015)
4. P. Gao, Y. Ma, K. Wen, M. Liu, J. Zheng, K. Chu, Z.J. Smith, L. Liu, Recent advances in structured illumination microscopy. *J. Phys.* **3**, 024009 (2021)
5. K. Samanta, J. Joseph, An overview of structured illumination microscopy: recent advances and perspectives. *J. Opt.* **23**, 123002 (2021)
6. R. Heintzmann, T. Huser, Super-resolution structured illumination microscopy. *Chem. Rev.* **117**, 13890–13908 (2017)
7. E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain, A. Sentenac, Structured illumination microscopy using unknown speckle patterns. *Nat. Photon.* **6**, 312–315 (2012)
8. S.A. Shroff, J.R. Fienup, D.R. Williams, Phase-shift estimation in sinusoidally illuminated images for lateral superresolution. *J. Opt. Soc. Am. A* **26**, 413–424 (2009)
9. Y. Wu, H. Shroff, Faster, sharper, and deeper: structured illumination microscopy for biological imaging. *Nat. Methods* **15**, 1011–1019 (2018)
10. Y. Ma, K. Wen, M. Liu, J. Zheng, K. Chu, Z.J. Smith, L. Liu, P. Gao, Recent advances in structured illumination microscopy. *JPhys: Photon.* **3**, 024009 (2021)
11. S.W. Hell, J. Wichmann, Breaking the diffraction resolution limit by stimulated emission: stimulated emission depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994)
12. E. Betzig, G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, H.F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006)
13. M. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796 (2006)
14. B. Huang, S. Jones, B. Brandenburg, X. Zhuang, Whole cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution. *Nat. Methods* **5**, 1047–1052 (2008)
15. R.W. Gerchberg, Super-resolution through error energy reduction. *Opt. Acta* **21**, 709–720 (1974)
16. A. Papoulis, A new algorithm in spectral analysis and band-limited signal extrapolation. *IEEE Trans. Circuits Syst.* **25**, 735–742 (1975)
17. C.J.R. Sheppard, M. Castello, G. Tortarolo, T. Deguchi, S.V. Koho, G. Vicidomini, A. Diaspro, Pixel reassignment in image scanning microscopy: a re-evaluation. *J. Opt. Soc. Am. A* **37**, 154–162 (2020)

Chapter 19

Image Reconstruction from Projections



In this chapter, we will briefly describe computational imaging concepts that use projections of the object of interest as the input data. Projection-based imaging naturally occurs in diagnostic healthcare systems, e.g. X-ray-based imaging systems, positron emission tomography (PET), magnetic resonance imaging (MRI), single-photon emission computed tomography, and to some extent in ultrasound imaging to an approximation [1–3]. These minimally invasive computational imaging systems have made tremendous difference to healthcare practices by providing suitable tools for imaging of internal body parts of human patients for diagnostic purposes and for carrying out image guided surgeries. In addition to healthcare systems, the projection-based ideas also find application in radio-astronomical imaging using large antenna arrays [4] where correlations between signals recorded at distant antennas provide Fourier space information about the source star structure as per the van Cittert-Zernike theorem. The projection-based concepts are inherently present in 3D structure reconstruction problems in cryo-electron microscopy which has become increasingly important for structural biology community [5]. Optical diffraction tomography and holo-tomography of biological samples also utilize projection-based ideas coupled with holographic recording of Biological samples such as cells. We have already developed some background on projection-based imaging in Chap. 2 where we discussed the Fourier slice theorem that connects a projection with the Fourier transform of the object to be imaged. The common line theorem which is a consequence of the Fourier slice theorem was also described. In this chapter, we will revisit the topic in more detail. The discussion in the following sections will be limited to 2D image reconstruction from projection data. Currently used medical X-ray computed tomography systems use a cone-beam X-ray configuration. The image reconstruction algorithms in this case are more involved, however, their basic principles can be traced back to the ideas discussed in the following sections. Historically the projection-based image reconstruction ideas originated from a field of mathematics called integral geometry. The basic mathematical contributions by Johann Radon (1917) and Fritz John (1938) which showed the possibility of recovering 2D or 3D functions from their line integrals are key to the modern developments and applications of the projection-based imaging systems.

19.1 X-ray Projection Data

When X-rays pass through an object such as a human patient, they get attenuated due to the absorption of the X-rays at each location within the object. Further for energies over 40 keV or so and in typical clinical settings, the diffraction effects if any are negligible and the X-ray propagation can be treated with a ray or geometrical optics model. Attenuation through a material can be modeled as follows. Referring to Fig. 19.1 we observe that if the incident X-ray beam intensity before entering the object is denoted by I_0 , the intensity I after passing through the object can be approximately represented by the well-known Beer-Lambert law:

$$I = I_0 \exp \left[- \int ds \mu(\mathbf{r}) \right]. \quad (19.1)$$

Here ds is an elemental length along the ray and $\mu(\mathbf{r})$ is the position (and hence material) dependent X-ray absorption coefficient. For example, in case of a human patient, the absorption coefficient is higher for bones as compared to soft tissue.

Following the notation in Sect. 2.11 the quantity

$$-\log \left(\frac{I}{I_0} \right) = \int ds \mu(\mathbf{r}) \quad (19.2)$$

may be associated with a line integral through the image $\mu(x, y)$ of the absorption coefficient at some angle θ represented as

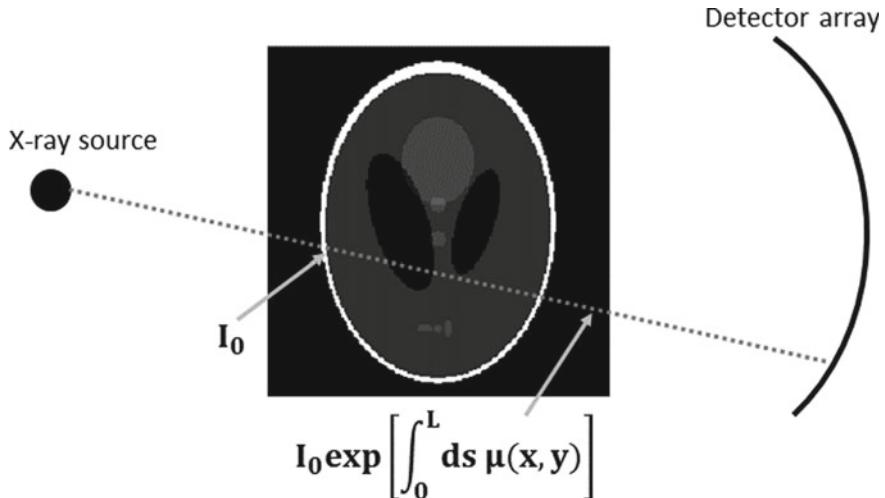


Fig. 19.1 Schematic representation showing mechanism of recording X-ray projection data through an object

$$p_\theta(t) = \iint dx dy \mu(x, y) \delta(x \cos \theta + y \sin \theta - t). \quad (19.3)$$

Readers may refer to Fig. 2.5 for the notation in the above equation. The projection data thus has two dimensions: the rotation angle θ and the location t along the detector plane. The projection data for a point object located at (x_0, y_0) is given by

$$p_\theta(t) = \delta(x_0 \cos \theta + y_0 \sin \theta - t). \quad (19.4)$$

As a function of t and θ and for constant (x_0, y_0) the projection data for a point object is a sinusoidal curve in the $t - \theta$ space. The projection data $p_\theta(t)$ is thus often referred to as a sinogram. As an illustration, we show the sinograms for a two-point object and a phantom object. Both the sinograms are plotted for parallel set of projections with 500 samples on the t axis and 1000 angular samples around the circle. While the sinogram for a two-point object is fairly easy to decipher, for a more complex object the sinogram does not provide much information about the object on simple visual inspection. Image reconstruction from projections therefore involves a linear reconstruction algorithm that needs to be designed carefully.

19.2 Image Reconstruction from Projection Data

In Sect. 2.11 we have already studied the Fourier slice theorem which states that the Fourier transform of $p_\theta(t)$ with respect to t provides a slice of the transform $M(f_x, f_y) = \mathcal{F}\{\mu(x, y)\}$ through the center of the Fourier space in the projection direction. In other words,

$$\mathcal{F}_t\{p_\theta(t)\} = M(v \cos \theta, v \sin \theta). \quad (19.5)$$

As shown in Fig. 19.3 the projection along one direction denoted by θ thus fills the Fourier space data along one line through the center of Fourier space. Projections along multiple θ directions therefore fill the whole 2D Fourier space densely. We further rewrite the Fourier expansion of $\mu(x, y)$ in order to relate the resulting expression to the Fourier slice theorem.

$$\begin{aligned} \mu(x, y) &= \iint_{-\infty}^{\infty} df_x df_y M(f_x, f_y) \exp[i2\pi(f_x x + f_y y)] \\ &= \int_0^{2\pi} d\theta \int_0^{\infty} v dv M(v \cos \theta, v \sin \theta) \exp[i2\pi v(x \cos \theta + y \sin \theta)] \\ &= \int_0^{\pi} d\theta \int_{-\infty}^{\infty} |v| dv M(v \cos \theta, v \sin \theta) \exp[i2\pi v(x \cos \theta + y \sin \theta)]. \end{aligned} \quad (19.6)$$

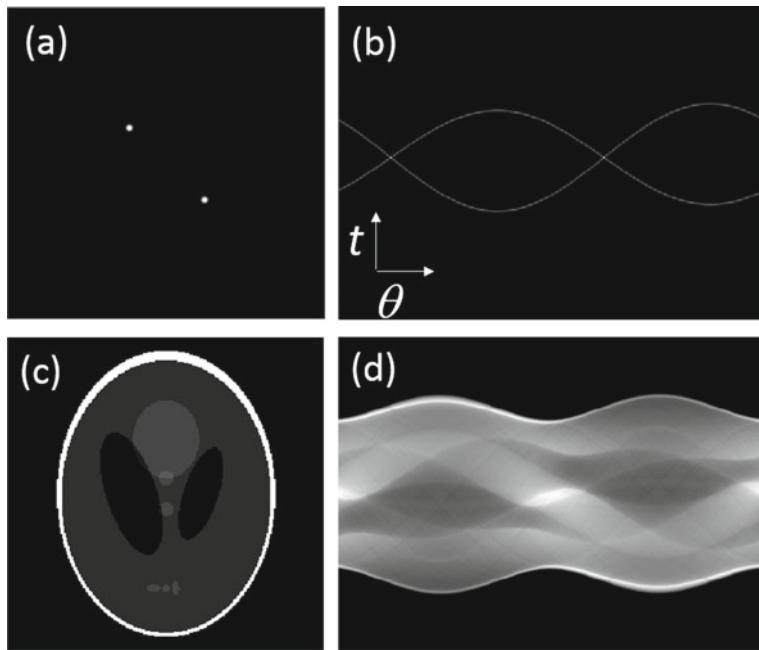


Fig. 19.2 Illustration of sinogram data for **a, c** Two point object, and phantom object; **b, d** Sinograms corresponding to (a), (c)

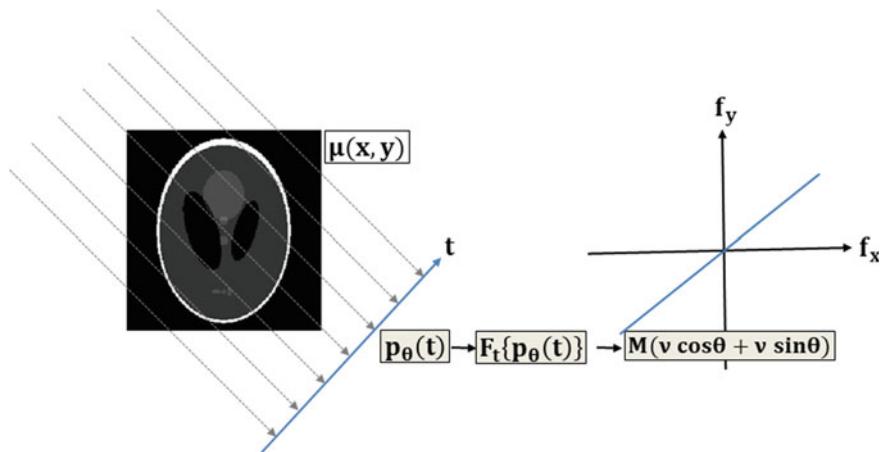


Fig. 19.3 Schematic representation showing the use of Fourier slice theorem for image reconstruction from projection data

Here we have used the polar coordinates $v = \sqrt{f_x^2 + f_y^2}$ and $\theta = \arctan(f_y/f_x)$. Further in the last line of equation above we have varied θ from 0 to π and correspondingly changed the limits on integration with respect to v so as to cover the whole (f_x, f_y) plane. Using the Fourier slice theorem and the above equation, we immediately see the steps required for reconstructing the image $\mu(x, y)$ as follows:

1. Take 1D Fourier transform of the projection data $p_\theta(t)$ along the $t -$ direction. This transform is equal to a slice through the 2D Fourier transform as represented by $M(v \cos \theta, v \sin \theta)$ for angle θ denoting the projection direction.
2. Multiply the 1D Fourier transform by the filter $|v|$ which is known as the ramp filter.
3. Take inverse 1D Fourier transform.
4. Integrate over all angles θ to calculate the contribution of each projection to a given image point $\mu(x, y)$.

The last step of integrating over all angles is known as backprojection and this operation is mathematically equivalent to the adjoint or transpose of the forward projection operation. The procedure outlined above is known as the filtered backprojection method and it was first suggested by Ramachandran and Lakshminarayanan [6]. In their honor the ramp filter is sometimes referred to as “Ram-Lak” filter. The filtered-backprojection method and its variants are used regularly in almost all current medical computed tomography (CT) scanners for 2D (slice) and 3D imaging of human patients [7–9]. In Fig. 19.4 we show the result of unfiltered and filtered backprojection of the data in Fig. 19.2d. We observe that simple (unfiltered) backprojection is not an inverse operation and typically gives rise to a blurred image and the full high-resolution reconstruction is provided when the ramp filter is employed [10]. The ramp filter may also be interpreted in an alternative way by writing:

$$|v| = -i \operatorname{sgn}(v) (i v). \quad (19.7)$$

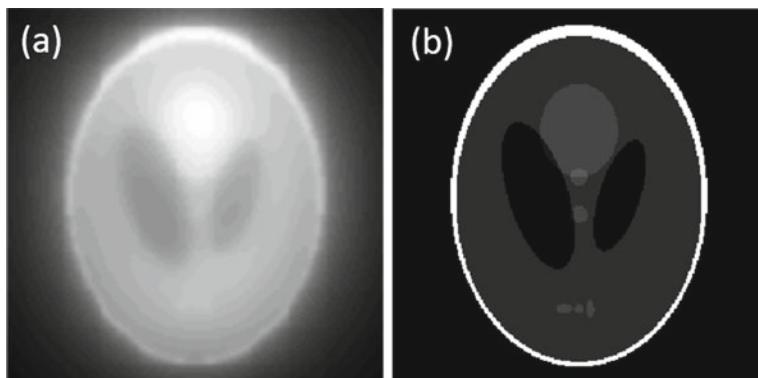


Fig. 19.4 Image reconstruction from sinogram data using **a** simple backprojection and **b** filtered backprojection. The illustration uses 1000 views around a circle

The filter function $-i \operatorname{sgn}(v)$ may be associated with Hilbert transformation and the term (iv) may be associated with a derivative operation. The ramp filtering can then also be thought of as a derivative of each projection (along t – axis) followed by a directional Hilbert transform in the image domain. We will not discuss this topic in further detail but mention that the derivative operation being a local operation, this method has some advantages in relation to memory access pattern in a practical implementation of projection-based reconstruction.

In the discussion of projection-based reconstruction we have ignored any noise issues so far. In X-ray-based projection imaging systems, it is necessary to keep the X-ray dose to the patient as minimum as possible. As the X-ray intensity is reduced, the detected X-ray signal prior to log-processing (as in Eq. 19.2) is corrupted by noise leading to noisy projections that have to be used in the reconstruction. In recent times, there is extensive ongoing research on exploring efficient iterative algorithms that are more suitable at handling noise as compared to the direct filtered- backprojection-based methods. The problem of image reconstruction in iterative methods reduces to a constrained optimization problem which may be stated as follows. Minimize the cost function $C(\mu)$ given by

$$C(\mu) = \|W(p) \cdot (p - \hat{A}\mu)\|^2 + \Psi(\mu). \quad (19.8)$$

Here p refers to the projection data, \hat{A} refers to the forward projection operation and $\Psi(\mu)$ refers to the set of constraints on the image $\mu(x, y)$. $W(p)$ is a data dependent weight function which gives appropriate weight to the error $(p - \hat{A}\mu)$ depending on the noise estimate in each measured value p so as to give less importance to measurements that are more noisy. The first term is thus a weighted data fitting term and instead of error, this term can have other forms like log-likelihood that can model the noise statistics much better. The constraints can be, e.g. positivity of the image, minimum total variation or other edge preserving penalties. The projection (\hat{A}) and backprojection (\hat{A}^T) operations are computationally expensive for large 3D images and specialized hardware is required for their fast and accurate implementation on commercial medical scanners. Iterative image reconstruction is becoming more widespread in medical CT scanners due to its image quality advantages at low X-ray dose [11–14]. More recently deep learning-based methods are also being employed in the CT reconstruction problem. Imaging of dynamic objects (like beating heart), incorporation of multiple X-ray energy projection data into recording and reconstruction protocol are some of the current research areas in X-ray projection-based imaging.

19.3 Holographic Tomography

In Sect. 14.3, we described the digital holographic microscopy system which provides a quantitative phase image of biological samples like cells. The phase map, $\phi(x, y) = k_0 \int dz n(x, y, z)$, is a projection of the refractive index $n(x, y, z)$ of the sample being

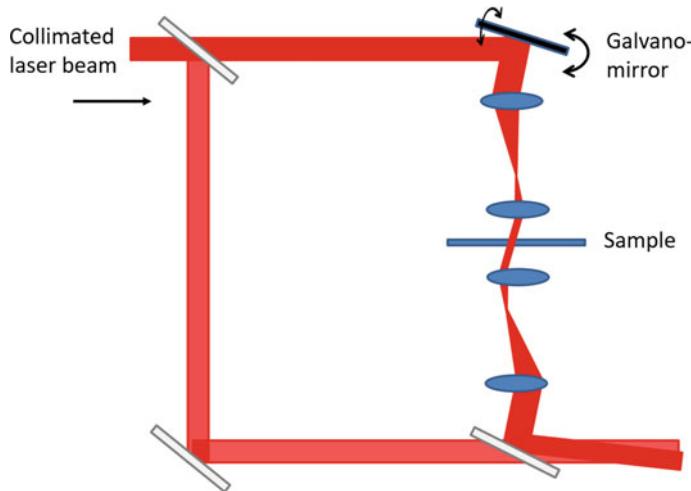


Fig. 19.5 Digital holographic tomography system for 3D index distribution imaging. The object beam and the reference beam are shown in different shades for clarity. The galvanometer mirror arrangement allows one to record holograms of the same sample with illumination from various angles

imaged, where k_0 is a constant. The setup in Fig. 19.5 may be used to perform such an imaging experiment. The galvanometer mirror arrangement allows one to record holograms of the same sample with illumination from various angles. De-modulation of each of the holograms provides a phase map which is equal to the projection of the object's refractive index distribution from different viewpoints. The generated phase maps can thus be utilized to perform a tomographic reconstruction of the refractive index profile of the sample. If the sample is held stationary, it may be noted that projections at angles closer to horizontal will not be available. This missing angle information will create slightly inaccurate representation of the 3D index distribution. Some of the missing angle information can be obtained if the sample can be tilted [15]. This will however make the system hardware more complex. A limitation of the picture presented for the holographic tomography here is that it essentially assumes a weak scattering object model (Born approximation) where the phase map recovered from an individual hologram recording can be considered as a projection of refractive index. This approximation is not valid if we are dealing with a highly scattering sample. The discussion in Sect. 14.5 regarding this point applies here too although we will not expand on this point any further.

Another important computational imaging problem which essentially uses holographic tomography ideas is the single particle 3D image reconstruction [5, 16] in cryo-electron microscopy (cryoEM). CryoEM design is like a 4F optical system. The macro-molecules (or virus like particles) of interest that are imaged using cryoEM systems are essentially weak phase objects as seen by electron beams. Therefore, a flat image is observed in the focus plane. In order to get some image contrast, the sample is de-focused as this transfers some phase information into amplitude of the

beam. The recorded pattern for each single particle is however like an in-line hologram. The recorded micrographs may be shown to have information related to single particle projection. Instead of orienting a single particle (which is too small), cryoEM systems record projection images of multiple similar kind of particles that are randomly oriented. These multiple projections whose angular orientation is not known *a priori* need to be used for reconstruction of 3D model of the virus like particles. The ideas like the common line theorem (see Fig. 2.6) are useful for this purpose. The major problem in cryoEM, unlike optical holographic tomography, is that the recorded micrographs are generally very noisy due to limits on highest allowable electron dosage. The image processing chain in cryoEM therefore involves elaborate procedures for particle classification and alignment followed by averaging over large number of similar projections to reduce noise. Recently quantitative phase recovery for single particles using the noisy cryoEM micrographs has also been demonstrated [17]. There has been a surge in research activity in the cryoEM reconstruction problem in the last decade and several challenges need to be addressed to get down to atomic scale resolution for the 3D reconstructions.

In conclusion, in this chapter, we have discussed some of the basic principles involved in projection-based imaging systems. The main idea of projection slice theorem was discussed in some detail. The typical X-ray computed tomography systems utilize a fan or cone beam geometry. In 2D problems, the fan beam geometry is equivalent to a parallel projection geometry (after re-assignment of rays). However in 3D X-ray cone beam geometry, a single rotation of source in a circle around the patient leads to missing data problem and helical source motion is required for artifact free reconstruction. We will not discuss these more complex projection geometries, however, it may be noted that the ideas discussed in this chapter do carry over to more advanced projection configurations as well.

Problems

19.1 Explain how you may be able to determine the outer boundary of the object being imaged using the sinogram data.

19.2 Determine the spatial domain version of the ramp filter which has been truncated to highest frequency ν_0 . Further show that the simple backprojection operation corresponds to image domain convolution by filter $1/r$ with $r = \sqrt{x^2 + y^2}$.

19.3 Typical CT scanners employ a fan beam geometry rather than the parallel beam X-ray geometry described in this chapter. Show that the 2D fan beam geometry (where a point source emits X-rays over an angular range) is equivalent to the parallel beam geometry.

19.4 Computed Tomography data with reduced X-ray dose may be recorded with reduced number of view angles or by keeping the same number of view angles but

reducing the X-ray beam intensity in each view. Provide your thoughts on which methodology will be better from the perspective of eventual image quality.

19.5 You are given two projections of a 3D object from different projection directions. Is this information sufficient to determine the relative orientation between the two projections?

References

1. A.C. Kak, M. Slaney, *Principles of Computerized Tomographic Imaging* (Society for Industrial and Applied Mathematics, 2001)
2. P. Suetens, *Fundamentals of Medical Imaging* (Cambridge University Press, 2009)
3. F. Natterer, *The Mathematics of Computerized Tomography* (Wiley, 1986)
4. R. Bracewell, *The Fourier Transform and Its Applications* (McGraw Hill, 1986)
5. J. Frank, *Three-Dimensional Electron Microscopy of Macromolecular Assemblies* (Oxford University Press, 2006)
6. G.N. Ramachandran, A.V. Lakshminarayanan, Three-dimensional reconstructions from radiographs and electron micrographs: application of convolution instead of Fourier transform. Proc. Natl. Acad. Sci. **68**, 2236–2240 (1971)
7. L.A. Feldkamp, L.C. Davis, J.W. Kress, Practical cone-beam algorithm. J. Opt. Soc. Am. A **1**, 612–619 (1984)
8. A. Katsevich, Theoretically exact filtered-backprojection type inversion algorithm for spiral CT. SIAM J. Appl. Math. **32**, 681–697 (2002)
9. Y. Zou, X. Pan, Exact image reconstruction on PI-lines from minimum data in helical cone-beam CT. Phys. Med. Biol. **49**, 941–959 (2004)
10. F. Noo, R. Clackdoyle, J. Pack, A two step Hilbert transform method for 2D image reconstruction. Phys. Med. Biol. **49**, 3903–3923 (2004)
11. B. DeMan, J. Nuyts, P. Dupont, G. Marchal, P. Suetens, An iterative maximum likelihood polychromatic algorithm for CT. IEEE Trans. Med. Imaging **20**, 999–1008 (2001)
12. I. Elbakri, J. Fessler, Statistical image reconstruction for polyenergetic X-ray computed tomography. IEEE Trans. Med. Imaging **21**, 89–99 (2002)
13. E. Sidky, X. Pan, Image reconstruction in helical cone-beam computed tomography by constrained total variation minimization. Phys. Med. Biol. **53**, 4777–4807 (2008)
14. J. Nuyts, B. DeMan, J. Fessler, W. Zbijewski, F.J. Beekman, Modelling the physics in the iterative reconstruction for transmission computed tomography. Phys. Med. Biol. **58**, R63–R96 (2013)
15. V. Balasubramani, A. Ku, H.Y. Tu, C.J. Cheng, M. Baczevska, W. Krauze, M. Kujawińska, Holographic tomography: techniques and biomedical applications. Appl. Opt. **60**, B65–B80 (2021)
16. E.V. Orlova, H.R. Saibil, Structural analysis of macromolecular assemblies by electron microscopy. Chem. Rev. **111**, 7710–7748 (2011)
17. A. Pant, M. Banerjee, K. Khare, Quantitative phase imaging of single particles from a cryoEM micrograph. Opt. Commun. **506**, 127588 (2022)

Chapter 20

Correlation Imaging



In most of the imaging systems discussed in this book so far, the image of the object to be recovered existed in the recorded data in the form of a system transfer function depending on whether the modality was coherent or incoherent. In this chapter, our focus will be on correlation imaging modalities where information is present in spatial correlation of seemingly random field intensity patterns. The first topic we will consider is the single-pixel ghost imaging. This will be followed by a discussion on imaging through random media using the so-called speckle memory effect.

20.1 Single-Pixel Ghost Imaging

The specialty of the technique referred to as “ghost imaging” is that it uses a bucket detector or a single-pixel detector to collect light that has interacted with the object of interest to form its image [1–5]. As a result, the resolution of one image record is just one pixel. The aim of this section is to provide a simple signal-processing-based explanation of image formation in this modality so as to make this topic more accessible. The point of view presented here is also amenable to newer and more efficient ways of utilizing this computational imaging modality. A ghost imaging system may be schematically described as shown in Fig. 20.1. As proposed initially [6], a source of light producing correlated photon pairs (e.g. parametrically downconverted photons in a non-linear optical process) is employed. It may be noted that the signal and idler photons generated in this process are time- as well as angle-correlated. The signal photon is transmitted through (or reflected from) the object of interest. This photon is detected by a bucket detector (practically a condenser lens which collects all the light and focuses it on a photo-detector) with complete disregard for its spatial location. The second idler photon which has never interacted with the object of interest is simultaneously detected on a position-sensitive detector. Our goal here is to recover the transmission function $T(x, y)$ of the unknown object from this arrangement which produces two signals—a photo-current $i_2(t)$ and area detector

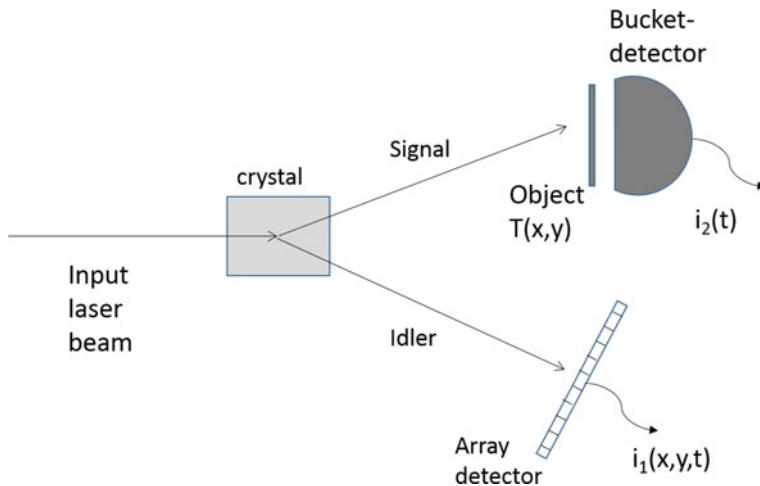


Fig. 20.1 Ghost imaging setup schematic using correlated photon pair from parametric downconversion process

frames denoted by $i_1(x, y; t)$. The estimated image of the object is then formed by the correlation of the two signals $i_2(t)$ and $i_1(x, y; t)$:

$$\tilde{T}(x, y) = \langle i_2(t) i_1(x, y; t + \tau) \rangle. \quad (20.1)$$

The time delay τ is adjusted appropriately according to the optical path difference in the two arms of the ghost imaging configuration. The correlation above is performed either by an electronic circuit or by numerically multiplying the time-dependent image $i_1(x, y; t)$ recorded on the array sensor with the photo-current $i_2(t)$ from the bucket detector and averaging over time. The image formation is possible due to the time and position correlation of the downconverted photon pair. If a photon count is detected by the bucket detector in a time interval Δt at time t , and a photon count is also registered by a pixel (x_0, y_0) in the corresponding frame on the area detector, the correlation of the two signals will lead to localizing the photon that was transmitted by the unknown object. It is easy to see that the strength of correlation for a given pixel location (x_0, y_0) on the area sensor will be proportional to the transmission function $T(x_0, y_0)$.

In the early days of ghost imaging research, it was generally considered that a quantum source of light producing correlated photons is required for implementing a ghost imaging system. However, it was soon clear that classical thermal (or pseudo-thermal) light can also give rise to required correlations [7–9] that have been known since the Hanbury Brown and Twiss experiment described in Chap. 10. For example, consider a system schematic as shown in Fig. 20.2 where a pseudo-thermal source is produced by a laser spot focused on a rotating ground glass diffuser. A beam-splitter in the arrangement in Fig. 20.2 ensures that the array detector in the reference arm

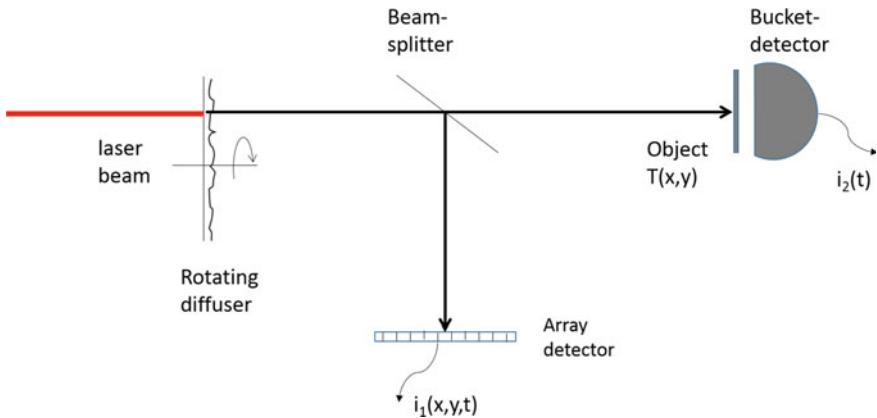


Fig. 20.2 Ghost imaging setup schematic using dynamic speckle from rotating ground glass diffuser that acts as a pseudo-thermal source

records an image of the speckle intensity pattern $i_1(x, y, t)$ and at the same time a bucket detector records a signal represented by the scalar product:

$$i_2(t) = \iint dx dy T(x, y) i_1(x, y, t), \quad (20.2)$$

between the speckle intensity pattern and the transmission function $T(x, y)$ of the object of interest. The correlation between the signals from the array and bucket detectors as above gives rise to

$$\begin{aligned} & \langle i_1(x, y, t) i_2(t + \tau) \rangle \\ &= \iint dx' dy' T(x', y') \langle i_1(x', y', t) i_1(x, y, t + \tau) \rangle \\ &= \iint dx' dy' T(x', y') \langle i_1(x', y') \rangle^2 [1 + |\gamma(x - x', y - y', \tau)|^2]. \end{aligned} \quad (20.3)$$

The intensity-intensity correlation above for a pseudo-thermal source has been factorized using the Gaussian moment theorem (see Eq. 7.38). The first term in the above equation integrates out to a constant background while the second term is a convolution of the absolute square of the coherence function $|\gamma(x - x', y - y', \tau)|^2$ with the image $T(x, y)$ to be estimated. The form of the coherence function can be controlled by selecting the appropriate size for the laser spot on the diffuser and the diffuser velocity, so that $|\gamma|^2$ can have a very narrow spread. Further, if the form of $|\gamma|^2$ is known, the image $T(x, y)$ can be recovered by any of the suitable image deconvolution algorithms as discussed in Chaps. 5 and 6 applied to the correlation function. The correlation above thus helps in forming the image. In the following illustration, we use a 100×100 pixel “IITD” object and perform the speckle-based

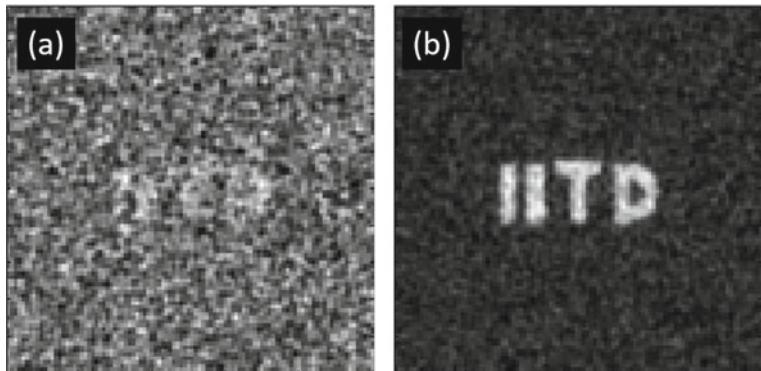


Fig. 20.3 Ghost images obtained by **a** 10^4 and **b** 10^6 speckle intensity frames for ensemble averaging. The IITD object has 100×100 pixels

averaging described above numerically. The reconstructed images after averaging over $N = 10^4$ and 10^6 realizations of speckle patterns are shown in Fig. 20.3a, b, respectively. We remark that the resolution of the resultant image is guided by the speckle size at the object transparency $T(x, y)$. The ghost imaging system may be simplified further if instead of the rotating glass diffuser, a spatial light modulator screen is used to generate the speckle patterns. In such cases, the speckle intensity patterns at the object can be numerically computed and the reference arm with an array detector is then not necessary.

20.1.1 A Signal-Processing Viewpoint of Ghost Imaging

The speckle-based averaging for image formation as suggested above is not very intuitive especially when a large number of photons are being detected in a single speckle intensity frame almost simultaneously in a realistic ghost imaging experiment. Further, a large number of speckle realizations (as compared to the number of pixels in the image), required for image formation, makes the overall image formation process somewhat inefficient from an information content point of view. A question one may further ask is whether the pseudo-thermal source is an absolute necessity for ghost imaging to work. We will answer these issues by providing a signal-processing approach to image formation in ghost imaging.

We will denote the sequence of speckle patterns used for correlation purpose as $\{\phi_n(x, y)\}$ with $n = 1, 2, \dots, N$. It is clear that the correlation of the bucket detector signal with the signal at the array detector amounts to the following process for the formation of the estimated image $\tilde{T}(x, y)$:

$$\tilde{T}(x, y) = \sum_{n=1}^N c_n \phi_n(x, y), \quad (20.4)$$

where the coefficients c_n 's are given by the scalar product between $\phi_n(x, y)$ and $T(x, y)$:

$$c_n = \iint dx dy T(x, y) \phi_n(x, y). \quad (20.5)$$

Clearly the speckle pattern realizations $\phi_n(x, y)$ are simply being treated as random keys and the image is being formed by their linear combination. Since the functions $\phi_n(x, y)$ representing the speckle intensity patterns are not orthogonal to each other, the process of the simple linear combination above is inefficient. In fact, since $\phi_n(x, y)$ are likely to be linearly independent functions, one should require no more speckle pattern realizations than the number of pixels in the image to be reconstructed. Further, if the expected sparsity in the object $T(x, y)$ is accounted for appropriately in the image reconstruction process, the number of speckle realizations much less than the number of pixels in the image should also allow for a good estimation of $T(x, y)$. In the illustration below, we model the ghost image formation problem as an optimization problem where we minimize the cost function of the form:

$$C(\tilde{T}) = \|g - S \tilde{T}\|_2^2 + \alpha TV(\tilde{T}). \quad (20.6)$$

Here, g is a vector formed by the sequence of bucket detector signals, matrix S consists of random keys (speckle patterns) arranged as row vectors and \tilde{T} denotes the required estimated image matrix arranged as a single column vector. Figure 20.4

Fig. 20.4 Ghost image obtained 5000 speckle intensity frames for ensemble averaging and optimization-based reconstruction. The IITD object has 100×100 pixels



shows the ghost image recovery with 5000 realizations of speckle patterns using the optimization-based solution which is much better compared to the image formed by simplistic averaging as in Fig. 20.3. This methodology for ghost imaging was proposed first with the name “single-pixel camera” [10]. The single-pixel camera used binary (0, 1) type illumination patterns that were obtained using a micro-mirror array device. The results on superior image reconstruction with reduced measurements as above once again highlight the power of the sparsity-constrained optimization procedure, this time for the ghost imaging modality.

20.2 Imaging Through a Scattering Medium Using Speckle Memory Effect

In this section, we will describe computational imaging systems that are capable of imaging objects located behind a scattering medium. Such media (e.g. turbid liquid, ground glass, scattering tissue, etc.) can transmit light but the high degree of scattering in them means that an object placed behind a slab of such a medium cannot be visualized directly. Referring to Fig. 20.5, a point monochromatic source placed behind such a medium will produce a random speckle pattern on an array sensor placed on the other side of the scattering medium. Since the light radiated by the point source is essentially completely scrambled by the scattering medium, it appears that the problem of locating the point source may be an impossible task. Such a task is however made possible by what is known as the speckle memory effect. Suppose now that the point source is translated laterally by a distance $\Delta x < (\lambda z_0)/(\pi L)$ (λ = wavelength, z_0 = distance between the point source and the scattering medium and L = thickness of scattering medium), then it is observed that the speckle pattern just moves laterally by a distance $\Delta x' = \Delta x(z_1/z_0)$ (z_1 = distance between scattering medium and array sensor), but the pattern itself does not change substantially [11],

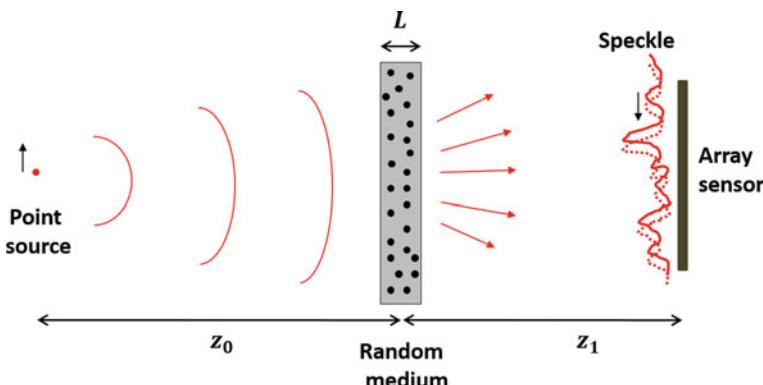


Fig. 20.5 Geometry of speckle memory effect

[12]. For an incoherent object which is small compared to the memory effect length Δx , the scattering medium thus essentially acts like a space-invariant system whose impulse response is given by the speckle pattern $s(x, y)$. For such a *small object*, we may therefore express the detected sensor pattern as

$$I(x, y) = I_{obj}(x, y) * s(x, y). \quad (20.7)$$

Here the symbol $*$ denotes convolution. The auto-correlation of the detected sensor pattern may now be written as

$$I(x, y) \star I(x, y) = [I_{obj}(x, y) \star I_{obj}(x, y)] * [s(x, y) \star s(x, y)]. \quad (20.8)$$

If the auto-correlation of the random speckle pattern $s(x, y)$ is small in spatial extent, its auto-correlation can be treated approximately as a delta distribution. This gives us an important result that

$$I(x, y) \star I(x, y) \approx [I_{obj}(x, y) \star I_{obj}(x, y)]. \quad (20.9)$$

The auto-correlation of the detected speckle pattern is thus equal to (a slightly blurred) auto-correlation of the object $I_{obj}(x, y)$ of interest that we wish to image. Further, the determination of $I_{obj}(x, y)$ from its auto-correlation is nothing but the phase retrieval problem which may be handled by algorithms discussed in Chap. 15. The relation to the phase retrieval problem here follows from the fact that the Fourier transform of the auto-correlation is equal to the squared Fourier magnitude (auto-correlation theorem). In Fig. 20.6, we show a letter object (“D”), its auto-correlation, the resultant speckle pattern in an arrangement like that in Fig. 20.5 and the auto-correlation of the speckle pattern. In line with the discussion above, the illumination is assumed to be monochromatic but spatially incoherent. The similarity of the two auto-correlations in Fig. 20.6c, d may be noted. It is clear that the speckle size may be controlled by the effective lateral dimension of the scattering medium, and a good approximation of the object’s auto-correlation is thus measurable from a single-shot pattern seen behind the diffuser. The quality of the auto-correlation may be improved if multiple patterns are recorded by moving the diffuser and the resultant auto-correlations of the random patterns are averaged. This unusual methodology where information is being recovered from seemingly random recorded intensity patterns truly demonstrates the power of computational imaging and how far we have come from a single-lens imaging model.

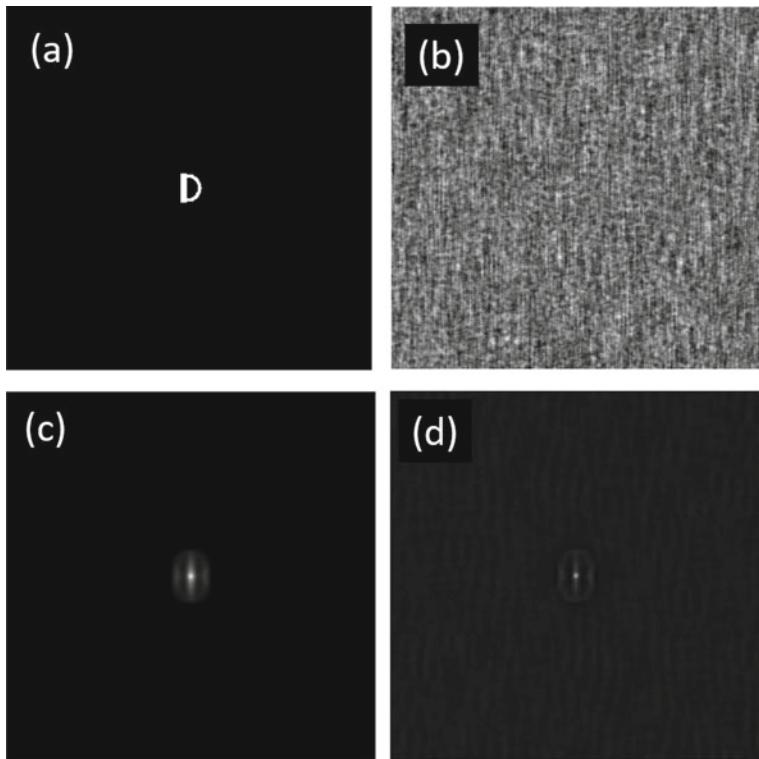


Fig. 20.6 Illustration of speckle memory effect. **a** Letter D object illuminated with incoherent beam, **b** resultant pattern seen on array sensor (see arrangement in Fig. 20.5) and **c**, **d** Auto-correlation of (a) and (b), respectively

Problems

20.1 Consider a ghost imaging system employing a low-resolution 4×4 pixels array detector instead of a single bucket detector. Formulate the reconstruction problem and explain the possible advantages of such a system.

20.2 Two cameras image a dynamic scene. One of the cameras has high spatial resolution but a slow frame rate. The other camera has low spatial resolution but a high frame rate. How would you utilize the image data stream from the two cameras to your advantage?

20.3 Why does the ensemble averaging approach in Eq. (20.3) require a very large number of speckle frames before a reasonable image is formed with the ghost imaging system?

20.4 As a more general case of the speckle memory effect, assume that the diffusing screen has been illuminated separately by two independent wavefronts $U_1(x, y)$ and

$U_2(x, y)$ at different times leading to speckle patterns $I_1(x, y)$ and $I_2(x, y)$ on the observation screen in Fig. 20.5. Show that the correlation of speckle intensity patterns contains information about $U_1^*U_2$ which is essentially the interference corresponding to the two probe fields. Explain the significance of this result [11].

20.5 An early demonstration of correlation imaging was shown by Albert Rose in [13]. An object was point scanned by directing light from a source with Galvano mirrors. The return light was synchronously detected using a photo-multiplier tube. A picture of a human face could be reconstructed at very low photon counts in this process. Given such data, how would you model the reconstruction process based on what you have learnt? (See [14] if you are interested after you have thought about the problem.)

References

1. B.I. Erkmen, J.H. Shapiro, Ghost imaging: from quantum to classical to computational. *Adv. Opt. Photonics* **2**, 405–450 (2011)
2. J.H. Shapiro, R.W. Boyd, The physics of ghost imaging. *Quantum Inf. Process.* **11**, 949–993 (2012)
3. B.I. Erkmen, Computational ghost imaging for remote sensing. *J. Opt. Soc. Am. A* **29**, 782–789 (2012)
4. O. Katz, Y. Bromberg, Y. Silberberg, Compressive ghost imaging. *Appl. Phys. Lett.* **95**, 131110 (2009)
5. J. Cheng, Ghost imaging through turbulent atmosphere. *Opt. Express* **17**, 7916–7921 (2009)
6. T.B. Pittman, Y.H. Shih, D.V. Strekalov, A.V. Sergienko, Optical imaging by means of two-photon quantum entanglement. *Phys. Rev. A* **52**, R3429-3432 (1995)
7. R.S. Bennink, S.J. Bentley, R.W. Boyd, Two-photon coincidence imaging with a classical source. *Phys. Rev. Lett.* **89**, 113601 (2002)
8. A. Gatti, E. Brambilla, L.A. Lugiato, Correlated imaging, quantum and classical. *Phys. Rev. A* **70**, 013802 (2004)
9. A. Gatti, E. Brambilla, M. Bache, L.A. Lugiato, Ghost imaging with thermal light: comparing entanglement and classical correlation. *Phys. Rev. Lett.* **93**, 093602 (2004)
10. M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, T. Sun, K.F. Kelly, R.G. Baraniuk, Single pixel imaging via compressive sampling. *IEEE Signal Proc. Mag.* **25**, 83–91 (2008)
11. I. Freund, Looking through walls and around corners. *Phys. A: Stat. Mech. Appl.* **168**(1), 49–65 (1990)
12. O. Katz, P. Heidmann, M. Fink, S. Gigan, Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nat. Photonics* **8**, 784–790 (2014)
13. A. Rose, The sensitivity performance of the human eye on an absolute scale. *J. Opt. Soc. Am.* **38**, 196–208 (1948)
14. A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. NC Wong, J.H. Shapiro, V.K. Goyal, First-photon imaging. *Science* **343**, 58–61 (2014)

Index

A

Afocal imaging system, 165
Analytic signal, 109
Angular spectrum method, 125
Aperture stop, 118
Apodization, 172
Arago spot, 140
Armijo rule, 73
ASD-POCS, 87
Auto-correlation function, 100

B

Backtracking line search, 73
Bandlimited extrapolation, 44
Bandpass sampling theorem, 38
Baye's rule, 96
Bessel function, 26
Black hole imaging, 148
Bochner's theorem, 100
Bucket detector, 281

C

Canonical optical processor, 168
Cauchy principal value, 19, 21
Central limit theorem, 101
Central moments, 98
Characteristic function, 99
Chief ray, 118
Chirp phase function, 23
Circ function, 26
Coherence function, 149
Coherency matrix, 162
Coherent impulse response, 169

Coherent mode representation, 149
Coherent transfer function, 174
Comb function, 21
Common line theorem, 28, 278
Compact form factor camera, 237
Complex degree of coherence, 143
Complexity guided phase retrieval, 225
Complex signal representation, 109
Compressive imaging, 71, 81
Computational imaging, 3
Computed tomography, 275
Computer generated hologram, 232
Conditional probability, 96
Convolution, 61
Correlation coefficient, 99
Correlation function, 144
Correlation imaging, 281
COSA phase mask, 259
Covariance matrix, 98
Cross spectral density, 149
Cryo-electron microscopy, 277
Cubic phase mask, 252

D

DC blocking filter, 169
De-blurring problem, 64
Defocus diversity, 233
Degree of polarization, 162
Delta function, 15, 17
Diffraction, 121
Digital super-resolution, 268
Digital holographic microscopy, 206
Dirichlet sufficiency conditions, 9, 14
Discrete Fourier transform, 51, 53

E

- Eigenfunction analysis of imaging systems, 179
 Entrance pupil, 118
 Ergodic process, 105
 Error reduction algorithm, 218
 Evanescent waves, 129
 Event Horizon Telescope, 148
 Exit pupil, 118
 Expectation value, 97

F

- Fast Fourier Transform (FFT), 51
 Fermat's principle, 115
 Fienup's HIO algorithm, 222
 Filtered backprojection, 275
 FINCH, 214
 Fourier phase retrieval, 217
 Fourier Ptychography, 228
 Fourier series, 9
 Fourier slice theorem, 27
 Fourier transform, 14
 Fourier transform hologram, 190
 Fourier Transform Method (FTM), 195
 Fractional Fourier transform, 137
 Fraunhofer diffraction, 131, 139
 Fresnel diffraction, 131
 Fresnel hologram, 190
 Fringe projection profilometry, 214
 FT property of lens, 167

G

- Gaussian distribution, 101
 Gaussian function, 22
 Gaussian moment theorem, 103
 Gaussian or Normal distribution, 101
 Geman prior, 93
 Generalized functions, 15
 Generalized sampling, 238
 Generalized Wiener filter, 66, 264
 Geometric phase shifter, 161
 Gerchberg-Papoulis algorithm, 268
 Gerchberg-Saxton algorithm, 219
 Ghost imaging, 281
 Gibbs phenomenon, 12
 Gradient descent iteration, 202

H

- Half wave plate, 160
 Hanbury Brown Twiss interferometer, 150
 Helmholtz equation, 123

Hilbert transform, 109

Holographic tomography, 213, 276
 Huber penalty, 92

I

- Ill-posedness, 62
 Image deconvolution, 71, 78
 Image denoising, 71
 Image plane hologram, 190
 Image reconstruction, 3
 Imaging systems, 1
 Impulse response, 60
 Inline digital holography, 208
 In-line hologram, 191, 192
 Integral Imaging, 245
 Intensity interferometry, 150
 Interference law, 144
 Inverse filter, 63
 Inverse problems, 62

J

- Jinc function, 26
 Joint probability, 96
 Jones matrix, 159
 Jordan's lemma, 20

K

- Kosambi-Karhunen-Loeve expansion, 107

L

- Laguerre-Gaussian function, 257
 Least-square solution, 82
 Lightfield camera, 240
 Linear systems, 59
 Linnik interferometer, 206
 Localization microscopy, 258, 266
 Log asphere, 255

M

- Mach-Zehnder interferometer, 204
 Magnitude transfer function, 175
 Mandel's theorem, 110
 Marginal ray, 118
 Maximum likelihood, 92
 Maxwell equations, 122
 Mean Gradient Descent (MGD), 89, 203
 Michelson interferometer, 144
 Michelson stellar interferometer, 147
 Moments, 97
 Montgomery conditions, 136

- N**
Nyquist rate, 35
- O**
Off-axis hologram, 191, 192
Off-axis digital holography, 194
Optical transfer function, 175
Optimization, 78
Optimization method for digital holography, 201
- P**
PhaseLift, 232
Phase unwrapping, 136
Phase shifting digital holography, 198
Photon counting formula, 152
Plenoptic imaging, 240
Point spread function, 175
Poisson summation formula, 34
Polarization, 159
Principal Component Analysis (PCA) basis, 108
Probability, 95
Prolate spheroidal functions, 39, 108, 179
Propagating waves, 128
Properties of Fourier transform, 24
Pseudo-eigenfunctions of Fresnel transform , 140
Pseudo thermal source, 156
Pupil function, 169
- Q**
Quarter wave plate, 160
- R**
Radio-astronomical telescope, 147
Ram-Lak filter, 275
Ramp filter, 275
Random variable, 97
Random process, 103
Rayleigh resolution criterion, 69
Rayleigh-Sommerfeld solution, 127
Rayleigh-Sommerfeld-Smythe relations, 128
Ray transfer matrix, 116
Rect function, 14
Relaxed Average Alternating Reflection algorithm, 225
Relaxed-Reflect-Reflect algorithm, 224
Rotating diffuser, 156
- S**
Sampling theorem, 33
Scanned beam ptychography, 233
Scanning confocal imaging, 268
Self-imaging, 136
Shack-Hartmann wavefront sensor, 246
Shannon sampling theorem, 34
Shot-noise-limit, 200
Sign function, 19, 20
Singular value decomposition (SVD), 185
Sinogram, 273
Slepian functions, 39, 108
Space-bandwidth product, 38, 181
Space-frequency representation, 149
Space-invariance, 60
Spatial and temporal coherence, 143
Speckle, 154
Speckle interferometry, 218
Speckle memory effect, 286
Spectral density, 106
Spectral interference, 158
Standard deviation, 98
Stationary phase method, 253
Stationary process, 104
Statistically independent events, 96
STED Microscopy, 265
Step function, 18
STORM, 266
Structured illumination, 261
- T**
Talbot effect, 136
Thin lens transmission function, 165
Tomographic image reconstruction, 213
Total square variation, 91
Total variation, 71, 72, 223
Transfer function, 130
Transport of intensity equation, 133
Twin stagnation problem, 223
- U**
Uniform convergence, 12
- V**
van Cittert-Zernike theorem, 146
Variance, 98
Varifocal system, 259

Very long base line interferometry (VLBI),
148
Vortex phase filter, 172

W

Weyl representation, 127
Wiener-Khintchine theorem, 106
Wiener filter, 64
Wigner distribution, 28

Wirtinger derivative, 78, 202
Wirtinger Flow, 231

Y

Young's double slit experiment, 144

Z

Zernike's phase contrast microscope, 170